



HAL
open science

Mesure et prévision de la fiabilité des circuits intégrés numériques de dernière génération sous l'effet du vieillissement

Justin Sobas

► **To cite this version:**

Justin Sobas. Mesure et prévision de la fiabilité des circuits intégrés numériques de dernière génération sous l'effet du vieillissement. Electronique. Université de Bordeaux, 2023. Français. NNT : 2023BORD0421 . tel-04400829

HAL Id: tel-04400829

<https://theses.hal.science/tel-04400829v1>

Submitted on 17 Jan 2024

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 PRÉSENTÉE
POUR OBTENIR LE GRADE DE
DOCTEUR
DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES PHYSIQUES ET DE L'INGÉNIEUR
SPÉCIALITÉ ÉLECTRONIQUE

Par **Justin SOBAS**

Mesure et prévision de la fiabilité des circuits intégrés numériques de
dernière génération sous l'effet du vieillissement

Sous la direction de : **François MARC**

Soutenue le 13 décembre 2023

Membres du jury :

Lirida Naviner De Barros,	Professeure, LTCI (Paris)	Rapporteur
Vincent Beroulle,	Professeur, LCIS (Valence)	Rapporteur
Christophe Jego,	Professeur, IMS (Bordeaux)	Président du jury
Jean-Claude Clément,	Ingénieur, Thales (Palaiseau)	Examineur
Sylvain Jouanolle,	Ingénieur, DGA-MI (Rennes)	Encadrant Industriel
François Marc,	Maître de conférence HDR, IMS (Bordeaux)	Directeur de thèse

Table des matières

Table des matières	iii
Remerciements	vii
Liste des figures	ix
Liste des tableaux	xv
Acronymes	xvii
Introduction générale	1
1 Du transistor au FPGA, de la fiabilité du transistor à celle du FPGA	5
1.1 Introduction	5
1.2 Le transistor	5
1.2.1 Bilan historique et économique du transistor	5
1.2.2 Principe de fonctionnement du MOSFET	7
1.2.3 Évolution du FET	10
1.2.4 L'architecture FinFET	11
1.3 Circuits élémentaires CMOS dans les FPGAs	12
1.3.1 L'inverseur	13
1.3.2 La porte de transfert	14
1.3.3 Les interconnexions	15
1.4 Field-Programmable Gate Array (FPGA)	17
1.4.1 Évolution des Programmable Logic Devices (PLDs)	17
1.4.2 Structure globale	19
1.4.3 La LookUp Table (LUT)	20
1.4.4 Le routage	22
1.4.5 La famille de FPGA Zynq UltraScale+	24
1.5 Mécanismes de dégradation et fiabilité des transistors	25
1.5.1 Introduction	25
1.5.2 Bias Temperature Instability (BTI)	26
1.5.3 Le HCI (Hot Carriers Injection)	31
1.5.4 Time Dependant Dielectric Breakdown (TDDB)	35
1.5.5 La fiabilité dans l'industrie	36
1.5.6 De la mesure à la fiabilité	38
1.5.7 Lois de vieillissement	39
1.5.8 Facteurs d'accélération (AF)	41
1.6 Méthodes et mesures de dégradation des FPGAs	42
1.6.1 Introduction	42
1.6.2 Méthode Transition Probability (TP)	43
1.6.3 Mesure par Ring Oscillator (RO)	45
1.6.4 Mesures de vieillissement des FPGAs	47
1.6.5 Modélisation de la fiabilité	50

1.6.6	Des techniques complémentaires pour les bancs de test	51
1.6.7	Bilan des mesures de dégradation sur FPGA	53
1.6.8	Conclusion	54
1.7	Synthèse	55
2	Développement et caractérisation du banc de test	57
2.1	Introduction	57
2.2	Description générale du banc de test	58
2.3	Banc d'oscillateur en anneau	60
2.3.1	L'oscillateur en anneau	60
2.3.2	Implémentation des ROs dans le FPGA	61
2.4	Circuit de mesure	62
2.4.1	Mesure de fréquence et du rapport cyclique	63
2.4.2	Temps de propagation d'un front montant et d'un front descendant	64
2.5	Utilisation du GPS comme référence de temps	68
2.6	Régulation de la température et tension interne du FPGA	72
2.6.1	SYStem MONItoring (SYSMON)	73
2.6.2	Régulation en température	74
2.6.3	Régulation en tension	80
2.6.4	Validation du système de régulation	81
2.6.5	Conditions de température et de tension possibles	82
2.7	Validation du banc de vieillissement	83
2.7.1	Présentation du test	84
2.7.2	Caractérisation du RO et compensation post-mesure	84
2.7.3	Résultats pour différentes configurations internes	86
2.8	Conclusion	87
3	Analyse et extraction des dégradations dans les FPGAs	89
3.1	Introduction	89
3.2	Banc de test et plan d'expérience	90
3.3	Mesures initiales des RO	91
3.3.1	Introduction	91
3.3.2	Période des ROs	92
3.3.3	Rapport cyclique des ROs	97
3.3.4	Caractérisation et modélisation en température et tension des ROs	98
3.3.5	Conclusion	100
3.4	Analyse préliminaire des dégradations des ROs	100
3.4.1	Introduction	100
3.4.2	Résultats des 5103 ROs	101
3.4.3	Effet de la tension et de la température	102
3.4.4	Effet de la fréquence de stress	104
3.4.5	Effet du rapport cyclique	106
3.4.6	Effet de l'architecture du RO	107
3.4.7	Conclusion	112
3.5	Extraction des dégradations par ressources	112
3.5.1	Introduction	112
3.5.2	Extraction à t_0 du temps de propagation des ressources de LUT et de routage	113
3.5.3	Extraction des dégradations de ressources de LUT et de routage	119
3.5.4	Extraction du temps de propagation en montée et descente	122
3.5.5	Extraction des dégradations des PMOS et NMOS pour un stress statique	124
3.5.6	Conclusion	129
3.6	Conclusion	129

4 Modélisation et prévision de la fiabilité des FPGAs	133
4.1 Introduction	133
4.2 Modélisation des dégradations pour un stress statique	134
4.2.1 Introduction	134
4.2.2 Dégradations du temps de propagation dans un RO	134
4.2.3 Dégradation dans les ressources logiques	137
4.2.4 Ressources de routage	143
4.2.5 Conclusion	146
4.3 Modélisation des dégradations pour un stress dynamique : rapport cyclique	146
4.3.1 Effet du rapport cyclique de stress sur les dégradations	147
4.3.2 Modèle dynamique à partir des modèles statiques	147
4.3.3 Modélisation du mécanisme de récupération	149
4.3.4 Conclusion	153
4.4 Modélisation des dégradations pour un stress dynamique : fréquence	153
4.5 Comparaison des dégradations dans un FPGA Artix 28nm HKMG et Zynq UltraScale+ 16nm FinFET	156
4.6 Modélisation des défaillances	158
4.6.1 Introduction	158
4.6.2 Seuil de défaillance	158
4.6.3 Extrapolation des dérivées	159
4.6.4 Modélisation statistique de la durée de vie	160
4.6.5 Conclusion	163
4.7 Conclusion	164
Conclusion et perspectives	167
Conclusion générales	167
Perspectives	168
Publications	171
Articles	171
Colloques	171
Bibliographie	173
Résumé/Abstract	185

Remerciements

"La pierre n'a point d'espoir d'être autre chose qu'une pierre. Mais, de collaborer, elle s'assemble et devient temple." Antoine de Saint-Exupéry

Cette thèse n'est pas le fruit de mon travail personnel mais s'installe dans un décor collaboratif dans lequel j'y ai apporté mon petit grain de connaissance, de travail et d'imagination. Je ne peux pas commencer cette thèse sans remercier toutes ces personnes sans qui ces lignes ne seraient pas là. Je souhaite commencer en remerciant la communauté scientifique au sens large puisque mes travaux sont construits sur une base de connaissances acquises au fil des siècles. Plus précisément, je tiens à remercier Mohammad Naouss et Julien Coutet dont les thèses ont servi de socle pendant les trois années de thèse.

Je veux pleinement remercier François Marc, mon directeur de thèse, qui s'est montré largement disponible afin de m'assister à différents plans. Premièrement sur la transmission de la connaissance en électronique puisqu'en commençant la thèse je ne connaissais même pas le fonctionnement d'un transistor. Deuxièmement, sur les aspects plus techniques allant de la conception VHDL au développement du banc de test ou encore, du traitement de données à la modélisation. Finalement, sur le travail de relecture et de correction, déjà de cette thèse, où j'ai compté une seule page sans trace de son stylo rouge, mais aussi des deux articles publiés. Je termine cette thèse satisfait de ce que j'ai appris et accompli, c'est en grande partie grâce à toi François, alors merci.

Je tiens à remercier Mme. Lirida Naviner et M. Vincent Berouille d'avoir accepté d'être les rapporteurs de ma thèse, j'espère que la lecture de ce mémoire vous sera plaisante.

Je souhaite remercier la Direction Générale de l'Armement ainsi que Bordeaux INP qui ont financé cette thèse. Je remercie spécialement Christian Moreau et Sylvain Jouanolle de la DGA qui ont régulièrement suivi mes travaux.

Durant les trois années de thèse, mon lieu de travail était au laboratoire IMS (Intégration du Matériaux au Système) sur le campus universitaire de Bordeaux. Ainsi, je remercie la direction du laboratoire ainsi que tous les services techniques qui ont œuvré pour me donner un cadre professionnel (bureau, réseau local, salle à manger ...) et un cadre technique (salles et instruments de mesure, outillage...). Dans le laboratoire IMS j'ai fait parti du groupe conception dont je remercie tous les membres pour le climat sympathique et les sorties de groupe annuelles (bateau sur la Garonne, restaurants...). Dans le groupe conception j'ai fait parti de l'équipe M4C (Model for Circuit) dont je tiens à remercier chacun des membres : Yifan, Lucas, Houssein, Chhandak, Bruno, Cristell, François, Risab, Louise, Nil et évidemment Marina la responsable. Je n'oublierai pas nos réunions du jeudi matin suivies d'un copieux *coffee break*. Finalement, je remercie mes collègues de bureau : Karthy, Barnabé, Djeber, Adrien, Soumya, avec qui j'ai partagé aussi bien sur le plan professionnel que personnel. Je repars de cette thèse avec des amitiés qui sont nouées pour de longues années.

Mes parcours musical, académique et professionnel sont les piliers de mon développement en terme de : méthode de travail, connaissance et structuration. Je remercie tous les enseignants et collègues qui ont contribué à mon développement. Je tiens à remercier particulièrement Olivier Brouard, mon professeur de cor, pour m'avoir transmis l'amour à la musique et le goût au beau. Je remercie Christophe Farges, le directeur de mon master, qui a rendu possible mon rebondissement du BTS vers l'université, il était mon encadrant académique d'alternance et m'a accompagné jusqu'en thèse. Je remercie aussi Laurent Boyer, ingénieur chez Dassault Aviation, qui était mon encadrant professionnel d'alternance et qui m'a transmis la rigueur et l'organisation dans le travail.

Une thèse c'est un ascenseur émotionnel avec des moments de grande satisfaction et des moments d'extrême détresse. Heureusement, j'ai un entourage qui m'a toujours soutenu. Pour cela j'adresse toute ma gratitude à mes parents pour les valeurs et l'éducation que j'ai reçue. Je remercie particulièrement mon père qui a soigneusement relu tout le manuscrit en portant une attention particulière sur l'orthographe. Je remercie aussi ma belle famille chez qui j'ai parfois télé-travaillé, notamment cet été pour rédiger dans le sud de l'Espagne au bord de la piscine avec une bière fraîche en main...je le conseille lorsque la motivation pour la rédaction est au plus bas. Finalement et surtout, j'adresse toute ma gratitude à ma femme Begoña qui est une source de réconfort, de joie, de motivation et d'espoir. Je ne compte pas le nombre de fois qu'elle m'a relevé pour aller de l'avant dans cette thèse. J'ai eu le bonheur d'accueillir notre fille Rita dans cette dernière année de thèse. Du haut de ses un an c'est elle qui me fait relativiser les moments difficiles du travail.

À ma famille,

Liste des figures

1.1	Loi de Moore 1965 et 1975 : Pr�evision empirique de l'�volution du nombre de composants dans les circuits int�gr�s	6
1.2	Loi de Moore revisit�e : More than Moore	7
1.3	Impact du semi-conducteur dans l'industrie et l'�conomie	7
1.4	Repr�sentation en coupe d'un transistor NMOS en mode d�pletion et inversion	8
1.5	Caract�ristique de transfert et caract�ristique de sortie d'un transistor NMOS	9
1.6	�volution de la densit� de transistors entre 1997 et 2027 dans les circuits int�gr�s num�riques	11
1.7	Sch�ma d'un MOSFET et d'un FinFET	12
1.8	Comparaison des caract�ristiques (DIBL et SS) et des performances (vitesse et puissance) entre le MOSFET et le FinFET	12
1.9	Fonctionnement d'un inverseur CMOS : sch�ma et caract�ristiques	13
1.10	Fonctionnement d'une porte de transfert CMOS : sch�ma et caract�ristiques	14
1.11	Vue en coupe d'un circuit avec trois niveaux d'interconnexions.	15
1.12	Repr�sentation des effets capacitifs dans les lignes d'interconnexion	16
1.13	�volution de R, C et RC dans les lignes d'interconnexion pour diff�rents noeuds technologiques	16
1.14	Frise chronologique de l'�volution des Programmable Logic Devices (PLDs)	17
1.15	Architecture d'un <i>Programmable Logic Array</i> (PLA)	18
1.16	Architecture d'un FPGA et d'une slice	19
1.17	Processus de configuration d'un FPGA illustr�	20
1.18	Fonctionnement d'une LookUp Table (LUT)	20
1.19	Architecture interne d'une LUT	21
1.20	Architecture d'un buffer tristate impl�ment� dans une LUT	22
1.21	Architecture du routage dans un FPGA	23
1.22	Architecture d'un <i>Programmable Interconnect Point</i> (PIP)	24
1.23	Pr�sentation du FPGA Zynq UltraScale+	25
1.24	Les d�gradations du BTI suivent une loi puissance	26
1.25	Effet de la tension et de la temp�rature sur le BTI	27
1.26	M�canisme de r�cup�ration du BTI lors d'un stress dynamique	28
1.27	Mod�lisation semi-empirique du recovery pour le BTI	28
1.28	Repr�sentation sch�matique du mod�le de <i>Reaction-Diffusion</i>	30
1.29	Mesure du BTI dans les FinFETs	30
1.30	Repr�sentation du BTI dans un FinFET	31
1.31	Principales caract�ristiques du <i>Hot Carrier Injection</i> (HCI)	32
1.32	Repr�sentation sch�matique du HCI dans un MOSFET	34
1.33	Effet du HCI en fonction du nombre d'ailettes dans le FinFET	34
1.34	Effet du HCI en fonction de la largeur d'ailette et de la fr�quence de stress dans un FinFET	35
1.35	Principales caract�ristiques du <i>Time Dependant Dielectric Breakdown</i> (TDDB)	36
1.36	Analyse des Modes de D�faillance, de leurs Effets et de leur Criticit� (AMDEC)	37
1.37	Approche statistique de la fiabilit�	39

1.38	Courbe en baignoire	40
1.39	Schéma de la méthode Transaction Probability (TP)	43
1.40	Chronogramme de la méthode Transaction Probability (TP)	44
1.41	Caractéristique de la méthode TP	44
1.42	Schéma du circuit pour mesurer la fréquence et le rapport cyclique d'un Ring Oscillator (RO)	46
1.43	Banc de test pour la mesure de fréquence de l'oscillation du signal de sortie d'un RO avec la méthode électromagnétique	46
1.44	Principe du circuit de mesure de la fréquence d'oscillation d'un RO avec la méthode <i>Undersampling</i>	47
1.45	Mesure de dégradation avec la méthode TP et RO	48
1.46	Mesure de dégradation avec la méthode <i>undersampling</i>	49
1.47	Mesure de dégradation avec la méthode <i>électromagnétique</i>	49
1.48	Mesure de dégradation avec la méthode RO et des compteurs	50
1.49	Schéma d'un RO avec un multiplexeur en entrée pour choisir le mode de fonctionnement du circuit : stress ou mesure	51
1.50	Plan en étage du FPGA configuré pour mesurer les dégradations des ressources de routage	52
1.51	Schéma d'un court circuit dans un FPGA pour augmenter la température	53
2.1	Vue d'ensemble de l'architecture du banc de test pour mesurer le vieillissement d'un FPGA	59
2.2	Schéma d'un Ring Oscillator (RO)	60
2.3	Schéma d'implémentation d'un RO dans un FPGA	61
2.4	Plan d'étage (floorplan) du FPGA après implémentation des 567 ROs	62
2.5	Schéma du circuit pour la mesure de fréquence, du rapport cyclique et son complémentaire des ROs	63
2.6	Chronogramme des signaux d'entrées et de sorties du circuit de mesure RO	64
2.7	Décomposition de la période d'un RO en la somme de temps de propagation de fronts descendants et fronts montants dans les étages	65
2.8	Identification des chemins de conduction dans une LUT en fonction du niveau d'entrée et de la fonction de la LUT	66
2.9	Banc de test pour la mesure de stabilité de l'oscillateur à quartz et du récepteur GPS	69
2.10	Mesures ΔM_{ij} avec deux oscillateurs à quartz et deux récepteurs GPS différents	69
2.11	Histogramme de la différence entre les mesures faites avec un deux récepteurs GPS différents et un même quartz et les mesures faites avec deux oscillateurs à quartz différents et un même récepteur GPS	70
2.12	Effet de la température, de la tension et de la consommation sur la fréquence d'oscillation d'un RO dans un FPGA	72
2.13	Diagramme du fonctionnement global du SYSMON	73
2.14	Explication du fonctionnement d'un module Peltier	74
2.15	Montage du module Peltier sur le FPGA et principe de fonctionnement du système thermique	75
2.16	Comparaison des modèles de connaissance et de comportement du module Peltier avec les mesures et évolution de la température interne du FPGA lors d'une Séquence Binaire Pseudo Aléatoire (SBPA)	77
2.17	Schéma électrique du filtre RLC avec la résistance R qui représente le module Peltier	79
2.18	Abaque pour dimensionner une capacité et une inductance dans un filtre RLC	79
2.19	Comparaison entre le modèle et la mesure pour une réponse indicielle	80
2.20	Plan d'étage du FPGA avec l'implémentation d'un banc de 1000 ROs, des 3 ROs NEAR et, des 3 ROs FAR	81
2.21	Évolution de la température et tension interne du FPGA ainsi que de la fréquence des ROs lors de l'activation et désactivation du banc des 1000 ROs	82

2.22	Caractérisation du système thermique en fonction du PWM et de la tension interne du FPGA	83
2.23	Caractérisation de la fréquence et du rapport cyclique d'un RO en fonction de la tension et pour différentes températures	84
2.24	Validation de la compensation post-mesure	85
2.25	Résultats pour la mesure de dégradation après un 1000 heures de vieillissement	87
3.1	Plan d'organisation pour la présentation des mesures expérimentales	90
3.2	Photo du banc de test pour le vieillissement et la mesure de la dégradation de neuf FPGAs	91
3.3	Conditions de stress en température et en tension des neuf FPGAs	91
3.4	Périodes initiales de tous les ROs dans les neuf FPGAs	92
3.5	Comparaison de la période des ROs mesurée avec la période attendue par Vivado	94
3.6	Comparaison de la variation inter-FPGA des performances mesurée et attendue par Vivado	95
3.7	Comparaison de la variation intra-FPGA des performances mesurée et attendue par Vivado	96
3.8	Rapports cycliques initiales de tous les ROs mesurés dans les neuf FPGAs	97
3.9	Coefficients des modèles de caractérisation de la fréquence des ROs en fonction de la tension et de la température interne du FPGA	98
3.10	Comparaison de la compensation post-mesure avec un modèle commun à tous les ROs et un modèle propre à chaque RO	99
3.11	Dérive fréquentielle relative de tous les ROs dans les neuf FPGAs chacun avec une condition de stress en température et tension différente	102
3.12	Effet de la tension de stress sur la variation de la dérive fréquentielle relative des ROs	103
3.13	Effet de la température de stress sur la variation de la dérive fréquentielle relative des ROs	103
3.14	Effet de la fréquence de stress sur la variation de la dérive fréquentielle relative des ROs	105
3.15	Dégradation relative finale par rapport à la dégradation finale pour un stress statique DC0 en fonction de la fréquence de stress	106
3.16	Effet du rapport cyclique de stress sur la variation de la dérive fréquentielle relative des ROs	107
3.17	Influence du type de ressource dans le RO (routage ou logique) sur la dérive fréquentielle relative des ROs	108
3.18	Identification du chemin de propagation d'un niveau haut et bas dans une LUT et des portes logiques vieilles avec différents nombres d'entrées utilisées	109
3.19	Effet du nombre d'entrées de la LUT utilisées sur la dérive fréquentielle relative finale des ROs	110
3.20	Influence de l'entrée de la LUT utilisée sur la dérive fréquentielle relative finale des ROs en fonction des conditions de stress	111
3.21	Influence de la fonction de la LUT sur la dérive fréquentielle relative finale des ROs en fonction des conditions de stress	111
3.22	Site logique (SLICEL), où les BEL logiques sont magenta, les BEL de routage sont vertes, les site pins sont rouges et les site wires sont jaunes [1]	113
3.23	Différentes ressources de routage dans un FPGA Xilinx	114
3.24	Histogrammes des ressources utilisées pour implémenter les 567 ROs dans un FPGA	115
3.25	Validation des régressions pour prédire la période initiale d'un RO	116
3.26	Temps de propagations initiales des ressources de routage et de LUT	117
3.27	Proportion du temps de propagation initial généré par les ressources de LUT par rapport au temps de propagation total dans un RO pour différentes architectures de RO	119
3.28	Validation de la première méthode de régression pour la prédiction des dégradations	120

3.29	Validation de la deuxième méthode de régression pour la prédiction des dégradations	121
3.30	Cohérence de l'extraction des dégradations dans les ressources de routage et de LUT	122
3.31	Rapport des dégradations finales dans les ressources de LUT	122
3.32	Évolution des dégradations du temps de propagation dans un RO, dans les ressources de LUT et de routage en fonction du temps de propagation initial	123
3.33	Validation de la régression pour l'extraction du temps de propagation initial à l'échelle du transistor	126
3.34	Validation de la régression pour l'extraction de la dégradation du temps de propagation à l'échelle du transistor	127
3.35	Dégradation du temps de propagation d'un PMOS et du NMOS	128
4.1	Évolution du temps de propagation d'un front dans un RO avec un stress statique	135
4.2	Identification du chemin de conduction dans une LUT à 4 entrées configurée en buffer où seule l'entrée I1 est utilisée	136
4.3	Effet sur le temps de propagation d'un front montant et descendant en sortie d'une LUT en fonction du stress en entrée d'une LUT	136
4.4	Évolution du temps de propagation d'un front dans les ressources de LUT avec un stress statique	137
4.5	Pré-facteur et exposant identifiés pour un modèle puissance	138
4.6	Comparaison du modèle 2 et de la mesure du temps de propagation d'un front descendant dans les ressources de LUTs avec un stress statiques	139
4.7	Paramètres b , γ , Ea et A du modèle 2	139
4.8	Paramètre d'amplitude A en fonction du temps de propagation initial et validation de la régression pour prédire l'amplitude du coefficient A à partir des entrées de LUT utilisées dans le RO	141
4.9	Comparaison du modèle 3 et de la mesure du temps de propagation d'un front descendant dans les ressources de LUTs avec un stress statiques	142
4.10	Validation du modèle 3 pour prédire les dégradations finales dans les ressources de LUT pour un stress statique	142
4.11	Comparaison du modèle et de la mesure du temps de propagation d'un front descendant dans les ressources de routage avec un stress statiques	143
4.12	Paramètre d'amplitude A en fonction du temps de propagation initial dans les ressources de routage et validation de la régression pour prédire l'amplitude du coefficient A à partir du temps de propagation initial dans les ressources de routage	144
4.13	Comparaison du modèle final et de la mesure du temps de propagation d'un front descendant dans les ressources de routage avec un stress statiques	145
4.14	Validation du modèle pour prédire les dégradations finales dans les ressources de routage pour un stress statique	146
4.15	Dégradations relatives finales dans les ressources de LUT et de routage en fonction du rapport cyclique de stress	147
4.16	Comparaison modèle dynamique sans prendre en compte la récupération du BTI et mesure de dégradation pour différents rapports cycliques de stress	148
4.17	Caractéristique du modèle semi-empirique de la récupération du BTI	149
4.18	Comparaison modèle dynamique en tenant compte de la récupération du BTI et mesure de dégradation pour différents rapports cycliques de stress	150
4.19	Paramètres du modèle de récupération du BTI	151
4.20	Validation du modèle pour prévoir les dégradations finales dans les ressources de LUT et de routage pour un stress statique et dynamique	152
4.21	Validation du modèle pour prévoir les dégradations finales dans les ROs pour un stress statique et dynamique	153
4.22	Effet de la fréquence de stress sur la dégradation du temps de propagation dans un RO	154
4.23	Dégradations relatives finales dans les ressources de LUT en fonction de la fréquence de stress	154

4.24 Évolution de la dégradation dans les ressources de LUT pour différentes fréquence de stress	155
4.25 Dégradations relatives finales dans les ressources de routage en fonction de la fréquence de stress	155
4.26 Comparaison des dégradation dans un FPGA 16nm FinFET et un FPGA 28nm MOSFET	157
4.27 Comparaison du seuil de défaillance extrait de Vivado avant vieillissement et après vieillissement des ROs	159
4.28 Modélisation des instants de défaillance avec une loi log-normales	160
4.29 Modélisation des instants de défaillance avec une loi log-normales et identification des facteurs d'accélération	161
4.30 Taux de défaillance des ROs avec comme seuil de défaillance Vivado et 10%	163
4.31 Évolution du taux de défaillance en fonction du temps de fonctionnement et du seuil de défaillance	163

Liste des tableaux

1.1	Équations simplifiés du courant de drain dans un transistor NMOS et PMOS pour différents modes de fonctionnement	9
1.2	Coefficients des facteurs d'accélération en tension (loi exponentielle) et en température (loi d'Arrhenius) obtenus pour la modélisation de la fiabilité des FPGAs dans les conditions de stress stimulant le BTI (hautes températures et hautes tensions)	51
1.3	Présentation des conditions et des résultats des test de vieillissement de FPGA	53
2.1	Conditions de stress interne	61
3.1	Résultats des variations de performances obtenus avec la mesure et extrait depuis Vivado	97
3.2	Coefficients du modèle de puissance pour quantifier l'impact de la tension et de la température de stress sur le vieillissement	104
3.3	Temps de propagation en fonction de l'entrée de la LUT extrait par la régression linéaire (3.12)	118
3.4	Temps de propagation dans différentes ressources à t_0 pour un RO L1 B	125
4.1	Coefficients du modèles (4.4) pour prédire l'effet du vieillissement pour un stress statique sur les LUT	141
4.2	Coefficients du modèles (4.6) pour prédire l'effet du vieillissement pour un stress statique sur les LUT	145
4.3	Coefficients pour les modèles de β_0 , β_1 , B_0 et B_1	152
4.4	Paramètres du modèles (4.16) extraits pour un FPGA Zynq UltraScale+ 16nm FinFET et un FPGA Artix 28nm HKMG	157
4.5	Paramètres extraits à partir du modèle (4.22) dans le cadre de notre étude sur le Zynq UltraScale+ 16nm FinFET et ceux obtenus dans [2] pour l'Artix 28nm HKMG	162

Acronymes

A

AC : Alternative Current;
ADC : Analog-to-Digital Converter;
AF : Acceleration Factor;
AF_T : AF Temperature;
AF_V : AF Voltage;
AHI : Anode Hole Injection;
AHR : Anode Hydrogen Release;
AMDEC : Analyse des Modes de Défaillance, de leurs Effets et de leur Criticité;
APU : Application Processor Unit;
ASIC : FPGA Programmable Logic;
AXI : Advanced eXtensible Interface;

B

BEL : Basic Element of Logic;
BEOL : Back End Of Line;
BRAM : Block RAM;
BTI : Bias Temperature Instability;

C

CDF : Cumulative Distribution Function;
CI : Circuit Intégré;
CLB : Configurable Logic Block;
CMOS : Complementary MOS;
CPLD : Complex Programmable Logic Devices;
CPT : ComPteur;
CUT : Circuit Under Test;

D

DC : Direct Current;
DG : Double Grille;
DGA : Direction Générale d'Armement;
DRP : Dynamic Reconfiguration Port;
DSP : Digital Signal Processor;
DUCY : DUty CYcle;

E

EEPROM : Electrically Erasable PROM;
EM : ÉlectroMagnétique;

F

FEOL : Front End Of Line;
FET : Field Effect Transistor;
FF : Flip-Flop;
FIT : Failure In Time;
FMEA : Failure Modes, Effects and Criticality Analysis;
FPGA : Field-Programmable Gate Array;

G

GAA : Gate All Around;
GND : Ground;
GPS : Global Positioning System;

H

HCI : Hot Carrier Injection;
HDL : Hardware Description Language;
HKMG : High- κ Metal Gate;

I

IA : Intelligence Artificielle;
IDE : Integrated Development Environment;
IO : Input/Output;
IP : Intellectual Property;

L

L1 B : Lut 1 Buffer;
L1 I : Lut 1 Inverser;
L2 XB : Lut 2 Xor Buffer;
L2 XI : Lut 2 Xor Inverser;
L3 XB : Lut 3 Xor Buffer;
L3 XB I1 : Lut 3 Xor Buffer Input 1;
L3 XB I5 : Lut 3 Xor Buffer Input 5;
L5 XB : Lut 5 Xor Buffer;
LED : Light Emitting Diode;
LUT : Look-Up Table;

M

MOS : Metal-Oxide-Semiconductor;
MOSFET : MOS Field Effect Transistor;
MTTF : Mean Time To Fail;

N

NBTI : Negative BTI;
NMOS : N-type doping MOS;

P

PAL : Programmable Array Logic;
PBTI : Positive BTI;
PCB : Printed Circuit Board;
PDF : Probability Density Function;
PI : Proportionnelle Intégrale;
PIP : Programmable Interconnect Point;
PL : FPGA Programmable Logic;
PLA : Programmable Logic Array;
PLD : Programmable Logic Device;
PLL : Phase Locked Loop;
PMIC : Power Management Integrated Circuit;
PMOS : P-type doping MOS;
PROM : Programmable Read Only Memory;
PS : FPGA Processing System;
PWM : Pulse Width Modulation;

R

RAM : Random Access Memory;
RAZ : Remise À Zéro;
RD : Reaction Diffusion;
RF : Recovery Fraction;
RO : Ring Oscillator;
RO LP : RO Long Path;
RO SP : RO Short Path;
ROUT : ROUTage;
RTL : Register Transfer Level;

S

SBPA : Séquence Binaire Pseudo Aléatoire;

SHE : Self Heating Effect;

SliceL : Slice Logic;

SliceM : Slice Memory;

SPI : Serial Peripheral Interface;

SPLD : Simple Programmable Logic Device;

SP : Semi-conducteur dopé P;

SN : Semi-conducteur dopé N;

SR : Sample Register;

SYSMON : SYStem MONitoring;

T

TAC : Transistion Activity Counter;

TCG : Test Clock Generator;

TCL : Tool Command Language;

TDDDB : Time Dependant Dielectric Break-down;

TG : Transmission Gate;

TP : Transition Probability;

TPA : Transition Probability Analyser;

TSMC : Taiwan Semiconductor Manufacturing Company;

TTF : Time To Fail;

TTOM : Transient Trap Occupancy Model;

TVG : Test Vector Generator;

U

UART : Universal Asynchronous Receiver/-Transmitter;

US : UnderSampling;

UTE : Union Technique de l'Électricité;

V

VAF : Voltage Acceleration Factor;

VHDL : VHSIC Hardware Description Language;

VHSIC : Very High Speed Integrated Circuit;

Introduction générale

L'évolution de notre société est marquée par quatre révolutions industrielles. La première révolution commence au XVIII^e siècle avec la découverte de la machine à vapeur conduisant à la mécanisation de la production. Au XIX^e siècle, la découverte de l'électricité donne le jour à la deuxième révolution industrielle qui se manifeste par la production sur ligne d'assemblage. C'est dans les années 1950 que la troisième révolution industrielle a débuté avec l'invention du transistor, découlant sur l'automatisation partielle de l'industrie grâce aux circuits intégrés numériques. Aujourd'hui nous sommes en plein dans la quatrième révolution industrielle : l'industrie 4.0, construite autour de la gestion des données et qui débouche, par exemple, sur l'intelligence artificielle rendant un système autonome. Les trois dernières révolutions sont chronologiquement dépendantes : il est impossible de faire un système intelligent sans circuit électronique comme il est impossible de faire un circuit électronique sans électricité.

L'électronique est plus que jamais présente dans la vie quotidienne. En 2020, 1000 milliards de circuits numériques ont été fabriqués, soit 130 circuits numériques pour chaque être humain sur terre [3]. Avec la diminution de la taille du transistor, le nombre de composants dans un seul circuit numérique explose, à l'image du System On Chip (SoC) M1 ultra chez Apple développé en 2022 [4]. La finesse de gravure du M1 Ultra est de 5 nanomètres, la puce contient 114 milliards de transistors et peut réaliser 22000 milliards d'opérations par seconde. Entre la communication, internet ... se sont 29000 gigaoctets par seconde de données qui sont émises ou stockées. Ainsi, il est prévu que la capacité totale des systèmes de stockage de données, ou *cloud*, soit de 100 zettaoctets (10^{21} octets) [5], correspondant au nombre de litres d'eau dans tous les océans. Presque 5 milliards de personnes utilisent quotidiennement internet et l'utilisation moyenne est de 7 heures par jours.

L'électronique génère un bilan économique colossal. La valeur du marché mondial des semi-conducteurs en 2021 était d'environ 550 milliards de dollars. Le marché mondial des puces devrait dépasser 1000 milliards de dollars d'ici à 2030 soit un tiers du PIB de la France [6]. En 2020, pendant la crise du Covid, le marché mondial des composants électroniques a connu une pénurie causée par une demande dépassant les capacités de production. Parce que l'électronique est présente dans la majorité des secteurs industriels, cette pénurie a eu des répercussions importantes, par exemple un recul d'un tiers de la production automobile dans certains pays d'Europe. Face à cette pénurie et dans un contexte géopolitique en tension, les grandes économies ont réalisé des investissements importants visant à augmenter leur production de circuits numériques. Par exemple, le "Chip Act" est un accord au niveau européen qui consiste en un investissement de 43 milliards d'euros dans l'industrie électronique pour que la production de circuits numériques en Europe soit de 20% de la production mondiale en 2030 (10% en 2022).

Le transistor est le composant à la base des circuits numériques. Depuis sa création en 1947 jusqu'à aujourd'hui, le transistor n'a cessé d'évoluer afin de répondre aux besoins sociétaux, industriels et économiques que nous venons de décrire. La principale évolution du transistor réside dans la diminution de sa taille. En dessous de 100 nanomètres (années 2000) les performances du transistor sont devenues insuffisantes, par rapport à la demande du marché, ce qui a nécessité l'utilisation de nouveaux matériaux pour générer des contraintes au niveau du canal de conduction, augmenter la perméabilité et améliorer la conduction au niveau de la grille (High- κ Metal

Gate HKMG). En dessous de 22 nanomètres (années 2010), même avec l'utilisation de nouveaux matériaux, la diminution de la taille du transistor tout en préservant de bonnes performances a touché ses limites. Dans ce contexte, le transistor FinFET est développé pour continuer la course à la miniaturisation tout en conservant de bonnes performances électrostatiques. Les circuits numériques de dernière génération sont conçus avec des transistors FinFET à l'exemple du SoC M1 ultra chez Apple. Ce qui caractérise le FinFET, c'est son canal vertical qui sort du plan, en forme d'ailette (en anglais Fin), contrairement au MOSFET qui a un canal horizontal.

La fabrication d'un circuit numérique est sujet à un compromis entre le coût, la performance et la fiabilité. La fiabilité d'un circuit numérique sous l'effet du vieillissement est impactée par la dégradation au niveau des transistors [7, chapitre 15] mais aussi par des défauts sur le packaging [7, chapitre 16]. Dans cette thèse nous nous intéressons uniquement à la dégradation des transistors. La fiabilité d'un transistor sous l'effet du vieillissement est affectée par trois principaux mécanismes de défaillances : le Bias Temperature Instability (BTI), le Hot Carrier Injection (HCI) et le Time Dependant Dielectric Breakdown (TDDB). Le mode de défaillance commun aux trois mécanismes est la dégradation progressive de la tension de seuil du transistor ayant pour conséquence l'augmentation du temps de propagation d'un signal dans un circuit numérique. Alors que les trois mécanismes de dégradation sont largement observés et modélisés dans les transistors planaires, la question suivante se pose : qu'en est-il de la fiabilité sous l'effet du vieillissement des transistors FinFET? L'évolution géométrique du FinFET va-t-elle avoir un effet négatif stimulant davantage les mécanismes (BTI, HCI et TDDB) et provoquant des dégradations du temps de propagation plus importante dans un circuit numérique?

Pour connaître l'effet des mécanismes de défaillance (BTI, HCI et TDDB) sur la fiabilité d'un transistor, il est nécessaire de réaliser des tests de vieillissement accéléré et de mesurer les dégradations. Les dégradations peuvent être mesurées au niveau du transistor mais cela nécessite de fabriquer des cartes de tests spécifiques, de posséder des pointes de mesure parfois de quelques micromètres et d'avoir une instrumentation spécifique pour placer les pointes et mesurer la tension de seuil. Monopoliser un tel banc de test avec l'accès au transistor est coûteux, c'est pourquoi le choix de réduire la durée du test de vieillissement devient nécessaire. Cependant, pour observer des dégradations pour une courte durée de test (quelques dizaines d'heures), il devient nécessaire que les conditions de vieillissement appliquées soient très éloignées des conditions opérationnelles (parfois deux fois la tension nominale). Le risque d'appliquer de telles conditions d'accélération du vieillissement est d'observer des mécanismes de défaillances qui ne sont pas représentatifs des mécanismes de dégradations dans les conditions opérationnelles.

Une alternative de la mesure des dégradations de la tension de seuil au niveau du transistor consiste à mesurer les dégradations du temps de propagation d'un signal au niveau du circuit numérique puisque la dégradation de la tension de seuil provoque la dégradation du temps de propagation d'un signal qui le traverse. De par sa configurabilité et sa modularité le FPGA est un candidat parfait pour implémenter une structure de test visant à mesurer la dégradation du temps de propagation. De plus, pour quelques centaines d'euros, des FPGAs conçus à partir de FinFET sont disponibles sur le marché. Le développement d'un banc de test avec des FPGAs est bien moins coûteux que la mesure directe sur le transistor. Ainsi des tests de plus longues durées (des milliers d'heures) peuvent être réalisés dans des conditions de vieillissement proches des conditions opérationnelles. Cependant, en mesurant la dégradation du temps de propagation d'un signal dans un FPGA nous observons en réalité un effet secondaire des mécanismes de dégradations sur des milliers de transistors compliquant la modélisation physique.

Dans ces travaux nous considérons que les dégradations dans les FPGA sont représentatives des dégradations dans les autres circuits numériques (processeur, Digital Signal Processor, etc...). En effet, en considérant que les deux circuits sont conçus avec la même technologie de transistor

(FinFET), la principale différence entre un FPGA et un autre circuit numérique c'est l'architecture. Dans cette thèse nous considérons qu'en choisissant un grand nombre de conditions de vieillissement et notamment différentes architectures de circuits dans un même FPGA, les résultats que nous obtiendrons permettront de couvrir les différences architecturales, à l'échelle du transistor, entre les circuits numériques.

Pourquoi est-il important de connaître la fiabilité d'un circuit numérique? Sur le marché public la performance et le coût d'un circuit numérique importe plus que sa fiabilité. Le téléphone portable en constitue un exemple flagrant. En moyenne, un français change de téléphone portable tous les 4 ans et, bien souvent, non pas parce qu'il est dégradé ou défaillant, mais pour en acheter un nouveau plus performant. Dans les secteurs sensibles tels que la défense, l'aéronautique ou encore le spatial, la fiabilité est une donnée d'entrée essentielle pour l'analyse de sûreté de fonctionnement ou de gestion de la qualité. Par exemple, dans l'aéronautique, afin de garantir la sécurité des passagers et du personnel, le constructeur d'avion ne peut pas intégrer un système avec un taux de défaillance supérieur à 10^{-9} par heure de vol si la conséquence de la défaillance de ce système mène au décès d'une personne. Parce que les études de fiabilité sont coûteuses en temps et en ressources financières, il existe des référentiels de fiabilité accessibles par les entreprises pour prédire la fiabilité d'un système. En 2004, la Direction Générale de l'Armement (DGA) a encouragé un consortium d'entreprises françaises à développer un référentiel de fiabilité : FIDES qui est alimenté par le retour d'expérience et par des études de vieillissements accélérés. C'est dans ce contexte que la DGA ainsi que Bordeaux INP financent ma thèse.

La sujet de cette thèse se fragmente en deux objectifs principaux. Le premier est de développer un banc de test capable de vieillir et de mesurer en continu les dégradations dans des FPGAs de dernière génération, avec une technologie 16 nanomètres FinFET, pendant un an soit plus de 10000 heures. Le deuxième objectif est d'analyser les mesures et de produire un modèle semi-empirique des dégradations dans un FPGA en nous basant sur les mécanismes de défaillance au niveau du transistor et ce afin de prédire la fiabilité sous l'effet du vieillissement. Ainsi, nous diviserons cette thèse en quatre chapitres.

Dans le chapitre 1, en nous appuyant sur la littérature, nous poserons les notions utiles pour comprendre les travaux qui seront présentés et qui s'articulent autour de deux thématiques : le FPGA et la fiabilité. Parce que le transistor est le composant élémentaire à la base d'un FPGA, nous présenterons son principe de fonctionnement et nous introduirons le FinFET. Nous décomposerons l'architecture d'un FPGA au plus bas niveau, en détaillant par exemple l'architecture interne d'une Look-Up Table, ce qui nous permettra de comprendre le fonctionnement d'un FPGA. Nous ferons un point de la littérature portant sur les mécanismes de défaillance (HCI, BTI et TDDB) au niveau du transistor et nous focaliserons l'état de l'art sur l'effet des mécanismes de défaillance sur le FinFET. Finalement, nous présenterons les méthodes de mesures de dégradations dans un FPGA et nous résumerons une dizaine de tests de vieillissement utilisant ces méthodes afin de nous guider dans le développement de notre banc de test.

Dans le chapitre 2 nous présenterons le banc de test que nous avons développé pour vieillir et mesurer les dégradations des FPGAs pendant plus de 10000 heures et dans des conditions (température et tension) proches des conditions d'utilisation opérationnelles. Pour mesurer les dégradations dans un FPGA, le banc de test se base sur l'implémentation de 567 oscillateurs en anneaux avec différentes configurations et l'implémentation de compteurs pour mesurer la fréquence et le rapport cyclique d'oscillation des oscillateurs en anneaux dans le FPGA. Nous souhaitons vieillir les FPGAs dans des conditions proches des conditions nominales afin que les mécanismes de dégradations soient représentatifs de ceux présents dans les conditions d'utilisation normales. Cependant, dans ces conditions, les dégradations sont faibles et requièrent un moyen de mesure précis pour être observées. De plus, la fréquence d'oscillation d'un oscillateur en anneau est sensible

au vieillissement mais encore plus à la température et la tension. Ainsi, la précision et la stabilité de mesure seront les points focaux dans le développement du banc de test. Afin d'améliorer la stabilité et la précision de mesure nous utiliserons un récepteur GPS comme référence de temps dans le circuit de mesure ainsi qu'un système de régulation de la température et de la tension interne du FPGA. Finalement, nous présenterons les résultats d'un test de vieillissement sur un seul FPGA pendant 1000 heures pour confirmer les performances du banc de test à mesurer des dégradations dans un FPGA.

Dans le chapitre 3 nous présenterons le banc de test final constitué de neuf FPGAs et ce avec la même configuration, mais avec chacun une condition de température et de tension différente. Avant le vieillissement nous avons mesuré tous les oscillateurs en anneaux (fréquence et rapport cyclique) pour connaître l'état initial des oscillateurs en anneaux et des FPGAs. Nous comparerons les mesures initiales avec les données de temps de propagation extraites du logiciel de conception Vivado. Après 8000 heures de vieillissement et de mesures de dégradation de 5103 oscillateurs en anneaux répartis dans neuf FPGAs, nous présenterons les résultats de dégradation et nous montrerons l'effet de la tension, de la température, du rapport cyclique de stress, de la fréquence de stress et de l'architecture de l'oscillateur en anneau sur les dégradations de la fréquence d'oscillation de l'oscillateur en anneau. Dans un FPGA, un oscillateur en anneau contient des ressources logiques et des ressources de routage. La dégradation de la fréquence d'un oscillateur en anneau dépend donc de la dégradation de ces deux types de ressources. Dans la fin de ce chapitre nous présenterons une méthode consistant à séparer de la mesure du temps de propagation d'un oscillateur en anneau, les temps de propagation générés par les ressources logiques et par les ressources de routage. Dans cette continuité nous montrerons comment extraire la dégradation du temps de propagation dans les ressources logiques et les ressources de routage.

Dans le chapitre 4 nous développerons un modèle semi-empirique des dégradations dans tous les oscillateurs en anneaux. Pour cela nous modéliserons séparément les dégradations dans les ressources logiques et les ressources de routage, ce qui nous permettra d'avoir un modèle commun pour toutes les architectures d'oscillateurs en anneaux. Dans un premier temps, nous modéliserons les dégradations dans le cas d'un stress statique pour différentes températures et tensions. Ensuite, nous modéliserons les dégradations pour un stress dynamique en séparant l'effet du rapport cyclique et de la fréquence du signal de stress. Ainsi, le modèle que nous développerons pourra prévoir les dégradations du temps de propagation d'un signal dans un oscillateur en anneau en fonction de la tension, la température, le rapport cyclique du signal de stress, la fréquence du signal de stress et l'architecture de l'oscillateur en anneau. Afin de montrer l'évolution des dégradations avec un FPGA utilisant une technologie de transistor planaire, nous comparerons nos dégradations avec celles mesurées et modélisées dans un test de vieillissement similaire sur un FPGA Artix 28 nanomètres HKMG. Finalement, en fixant un seuil de défaillance et en extrapolant nos mesures jusqu'à ce seuil, nous modéliserons la répartition des défaillances afin d'en extraire le taux de défaillance.

Chapitre 1

Du transistor au FPGA, de la fiabilité du transistor à celle du FPGA

1.1 Introduction

Les travaux de cette thèse s'articulent autour de deux thématiques : le FPGA et la fiabilité. Le composant élémentaire qui est à la base de la conception d'un FPGA est le transistor. Aussi, la fiabilité du FPGA sous l'effet du vieillissement dépend de mécanismes de défaillance à l'échelle du transistor. Ainsi, pour comprendre les travaux de "mesure et prédiction de la fiabilité des FPGAs" que nous présenterons dans les prochains chapitres, nous allons dans ce chapitre introduire le transistor (section 1.2) ce qui nous permettra de présenter le fonctionnement d'un FPGA (section 1.4), puis, nous présenterons l'état de l'art des mécanismes de défaillance présent dans le transistor (section 1.5) afin d'introduire l'état de l'art des études de vieillissement visant à mesurer la fiabilité des FPGAs (section 1.6).

1.2 Le transistor

Le transistor est le composant actif à la base de tous les circuits numériques intégrés, et donc du FPGA. Afin de comprendre le fonctionnement d'un FPGA au plus bas niveau, il est nécessaire de connaître le fonctionnement du transistor. Dans cette section, nous contextualisons l'émergence du transistor dans l'industrie de l'électronique avant d'expliquer le principe de fonctionnement du MOSFET. Parce que la taille du transistor ne cesse de diminuer, nous présentons l'évolution du transistor au fil des années avec les principales avancées technologiques pour introduire le FinFET. Le FinFET étant le transistor présent dans le FPGA dont nous mesurons la fiabilité, nous montrons son architecture et ses caractéristiques.

1.2.1 Bilan historique et économique du transistor

Créé en 1947 par un groupe de chercheurs des Laboratoires Bell, le transistor est une des inventions révolutionnaires du XX siècle qui impacte notre quotidien. Le transistor est le composant élémentaire à la base du Circuit Intégré (CI), du processeur, de la cellule mémoire, etc. L'objectif principal de l'industrie de la micro-électronique est de réduire la taille du composant afin d'augmenter la densité de transistors sur une même surface et ainsi de gagner en performance et en coûts de développement.

En 1965, Gordon Moore observa deux points essentiels concernant l'évolution de l'intégration des transistors et de leur coût de fabrication [8]. Premièrement, le coût revient d'un composant en fonction de sa densité dans un circuit intégré possède un minimum. Deuxièmement, pour ce coût minimum le nombre de transistors dans un CI est multipliée par deux chaque année. Ces observations se basent sur l'analyse du marché du circuit intégré entre 1962 et 1965. La figure 1.1a, extraite de [8], est le résultat de cette observation et présente le prix par composant en fonction du

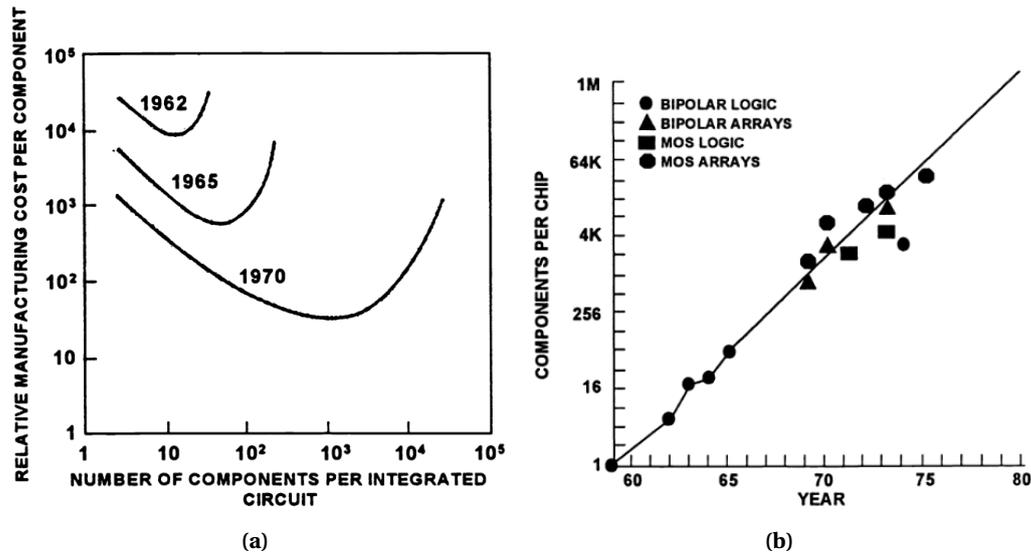


FIGURE 1.1 – (a) Coût par composant en fonction du nombre de composant par circuit intégré entre 1962 et 1970 [8] (b) Nombre de composants par circuit intégré entre 1959 et 1975 à l'origine de la loi de Moore [9]

nombre de composant dans le circuit intégré. Le prix par composant minimum est obtenu pour 10 composants par circuit en 1962 contre 80 composants par circuit en 1965, soit bien un facteur 2^3 en trois ans d'intervalle.

En 1975, Gordon Moore montre dans [9] que le nombre de composant par circuit intégré, soit sa complexité, continue d'augmenter par deux chaque année. La figure 1.1b montre que le nombre de composants est passé de 1 à 65000 par circuit entre 1959 et 1975 soit un facteur 2^{16} en accord avec la projection faite en 1965. La taille des circuits ont augmenté d'un facteur de 20 sur cette même période. Cela montre que le nombre de composant par circuit a augmenté grâce à l'augmentation de la densité des composants, et donc par l'amélioration des outils de lithographie.

En 1995, Gordon Moore montre dans [10] que la nouvelle tendance du nombre de composants par circuit entre 1975 et 1995 est d'un facteur deux tous les deux ans (tous les ans avant 1975). La cause de la diminution de la pente de l'intégration est l'augmentation importante du coût des outils de lithographie en fonction de la diminution de la taille du composant. En 1968 le coût était de 12000 dollars pour une technologie de $10\mu m$ contre 3.5 millions de dollars en 1995 pour une technologie de $0.4\mu m$. Selon [11], l'industrie du circuit intégré a suivi cette dernière prédiction de Gordon Moore [10] jusqu'en 2010.

La diversification des systèmes dans une même puce complète la loi de Moore avec une nouvelle tendance appelée "More-than-Moore" [12]. La figure 1.2 montre que le développement de CI est porté aujourd'hui par deux axes : "More than Moore" (en bleu) et "More Moore" (en rouge). L'axe "More than Moore" correspond à la diversification des systèmes dans une même puce et le "More Moore" correspond à la miniaturisation des composants. L'évolution de la technologie dans cette dernière décennie, au même titre que l'émergence des objets connectés (Internet of Things IoT) ou encore de l'Intelligence Artificielle (IA), a poussé l'industrie au développement de CI comprenant des fonctionnalités numériques et analogiques interfacées dans un même boîtier (System in Package SiP) ou dans une même puce (System on Chip SoC).

L'industrie du semi-conducteur est en croissance permanente. La figure 1.3a montre que le semi-conducteur est présent dans différents secteurs d'activités tels que la communication, l'informatique, le transport et d'autres encore [13]. En passant de 300 milliards d'unités à 1100 milliard d'unités, la production de semi-conducteurs est multipliée par quatre entre 2001 et 2021 [14]. Le bilan économique du marché mondial du semi-conducteur suit la même tendance, comme en témoigne la figure 1.3b. Un facteur quatre sépare le bilan de 2001 avec 147 milliards de dollars de

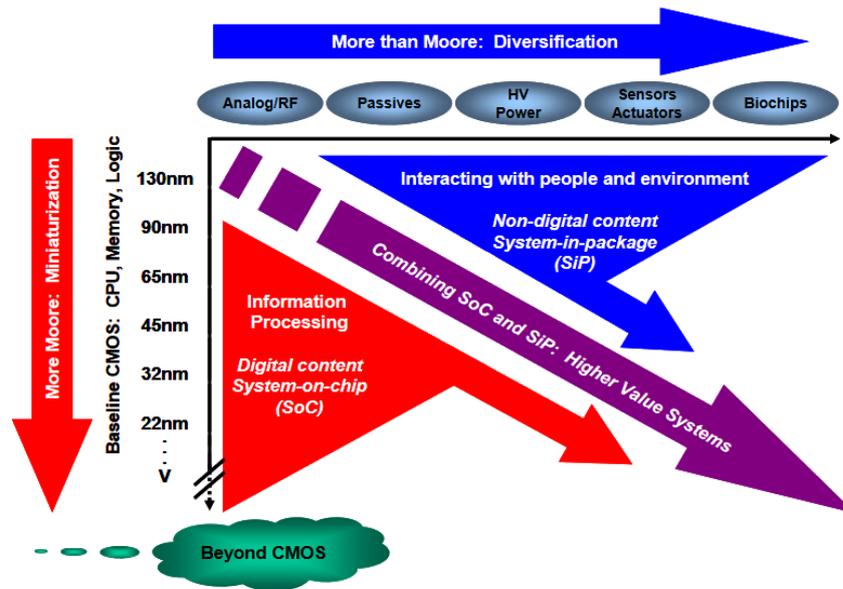


FIGURE 1.2 – Évolution des circuits intégrés en considérant la miniaturisation du composant (more Moore) et la diversification de système sur dans un même circuit (more than Moore)[12]

2021 avec 543 milliards de dollars. On constate que le marché économique du semi-conducteur est en majorité contrôlé par l'Asie (Asie/Pacifique+Chine), suivis par les États-Unis et enfin le Japon et l'Europe. La prévision de l'évolution du marché du semi-conducteur est en croissance pour les années à venir avec l'intention de produire un transistor sub-nanométrique pour 2034 [14]. La course à la miniaturisation à un coût considérable pour les fondeurs de semi-conducteur, ce qui entraîne la diminution du nombre de fondeurs dans le monde. En 2017-2018 on ne comptait que 3 entreprises capables de produire une technologie de 10nm : Intel, TSMC et Samsung.

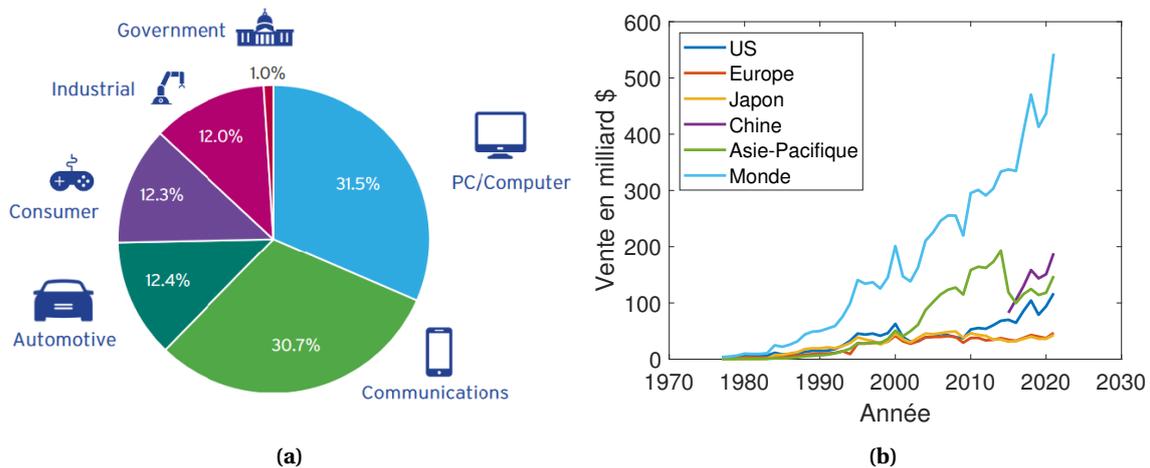


FIGURE 1.3 – (a) Répartition des composants issus de l'industrie du semi-conducteur par secteur d'activité [13] (b) Vente de composants semi-conducteur par année pour différentes régions du monde (tracé à partir des données récupérées dans "World Semiconductor Trade Statistics WSTS " [15])

1.2.2 Principe de fonctionnement du MOSFET

Le MOSFET (Metal-Oxide-Semiconductor Field-Effect Transistor) a été un des transistors à effet de champs les plus présents dans les circuits numériques depuis la création des circuits intégrés en 1980 jusqu'en 2010. À ce jour c'est le FinFET qui est à base des circuits numériques de

dernière génération. Parce que nous étudierons la fiabilité du FPGA conçu avec du FinFET, qui est une évolution du MOSFET, nous introduisons dans cette section le fonctionnement du MOSFET.

Afin de mieux comprendre les explications qui vont suivre, un schéma en coupe d'un transistor est donné sur la figure 1.4. Le MOSFET (Metal-Oxide-Semiconductor Field-Effect Transistor) est un transistor à effet de champs qui est divisé en 4 parties : La source, la grille, le drain et le substrat. Un FET classique est fabriqué à partir du silicium qui est un matériau du groupe *IV* dans la table périodique, c'est à dire avec quatre électrons sur la bande de valence, et qui possède la caractéristique électrique de semi-conducteur. Pour contrôler le transistor, des contacts métalliques sont déposés sur la source, la grille, le drain et le substrat. Une couche d'oxyde de silicium isole électriquement le substrat du contact métallique de grille.

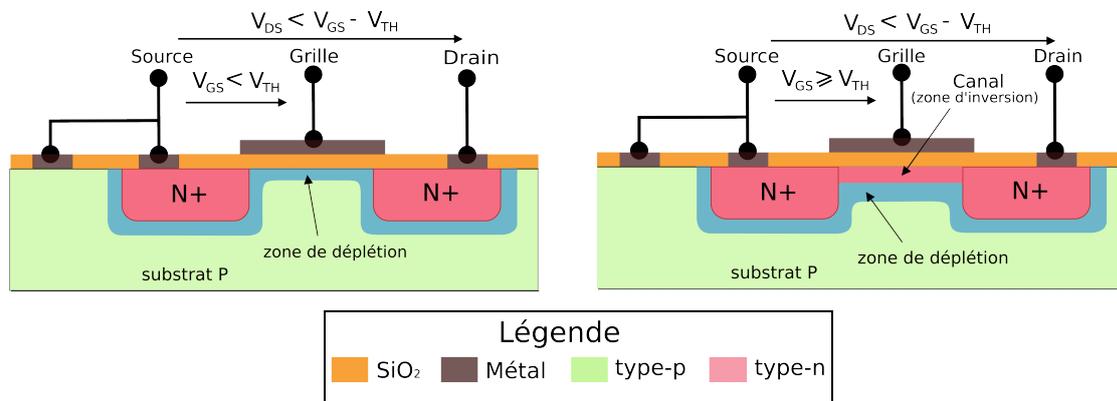


FIGURE 1.4 – Représentation en coupe d'un transistor NMOS avec le canal en mode déplétion (gauche) et inversion (droite)

Parce que le silicium à température ambiante a une grande résistivité, des impuretés sont introduites afin d'augmenter la concentration de porteurs libres et ainsi de diminuer la résistivité électrique. Il existe deux types d'impuretés : l'arsenic et le phosphore du groupe *V* (dopage de type-n) et le bore du groupe *III* (dopage de type-p). Le dopage de type-n est utilisé pour augmenter la densité d'électrons et le dopage de type-p est utilisé pour augmenter la densité de trous. Un transistor est de type NMOS lorsque la source et le drain ont un dopage de type-n et le substrat un dopage de type-p. Dans un transistor NMOS la source et le drain sont des zones denses en électrons contrairement au substrat qui est dense en trous. Inversement, dans un transistor PMOS la source et le drain sont de type-p et le substrat de type-n.

Dans le cas d'un NMOS, une différence de potentiel positive est appliquée entre la grille et le substrat générant un champ électrique. Le peu d'électrons présents dans le substrat se déplacent en sens opposé du champ électrique et s'accumulent au niveau de l'oxyde. De plus, le champ électrique aplanit les niveaux d'énergie entre la source et le substrat ce qui rend possible le déplacement d'électrons de la source vers l'oxyde de grille. Plus le champ électrique est important et plus la densité d'électrons au niveau de l'oxyde devient importante formant ainsi un canal de conduction de type-n entre la source et le drain qui sont aussi denses en électrons. Dans un premier temps, lorsque les premiers électrons commencent à s'accumuler on parle de déplétion puis lorsque la densité d'électrons devient plus importante que celle des trous on parle d'inversion. Dans le cas d'un PMOS il faut une différence de potentiel négative entre la grille et le substrat pour qu'il y ait une accumulation de trous au niveau de l'oxyde, formant ainsi un canal de conduction de type-p. Une fois le canal de conduction établi dans le MOSFET il suffit d'appliquer un champ électrique entre la source et le drain pour avoir un transport d'électrons pour le NMOS et de trous pour le PMOS.

NMOS		
Bloqué	$V_{GS} \leq V_T$	$I_{DS} = 0$
Linéaire	$V_{GS} > V_T, V_{DS} \leq V_{GS} - V_T$	$I_{DS} = \beta \left((V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right)$
Saturé	$V_{GS} > V_T, V_{DS} > V_{GS} - V_T$	$I_{DS} = \frac{\beta}{2} (V_{GS} - V_T)^2$
PMOS		
Bloqué	$V_{SG} \leq V_T $	$I_{SD} = 0$
Linéaire	$V_{SG} > V_T , V_{SD} \leq V_{SG} - V_T $	$I_{SD} = \beta \left((V_{SG} - V_T) V_{SD} - \frac{V_{SD}^2}{2} \right)$
Saturé	$V_{SG} > V_T , V_{SD} > V_{SG} - V_T $	$I_{SD} = \frac{\beta}{2} (V_{SG} - V_T)^2$

TABLE 1.1 – Équations simplifiées du courant de drain dans un transistor NMOS et PMOS pour différents modes de fonctionnement

En fonction de la tension grille-source (V_{GS}) et de la tension drain-source (V_{DS}), le transistor fonctionne suivant trois régimes. Nous présentons le fonctionnement d'un NMOS qui est similaire à celui d'un PMOS avec comme différence le sens de la polarisation des tensions V_{GS} et V_{DS} .

Lorsque V_{GS} est inférieur à la tension de seuil du transistor (V_T) le canal commence à se dépléter (voir figure 1.4). Cependant, le canal n'est pas établi et par conséquent aucun courant ne peut passer de la source au drain. En réalité, même dans ce régime il existe un faible courant non désiré appelé I_{OFF} . Dans cette condition, la figure 1.5a montre que le transistor est en régime bloqué.

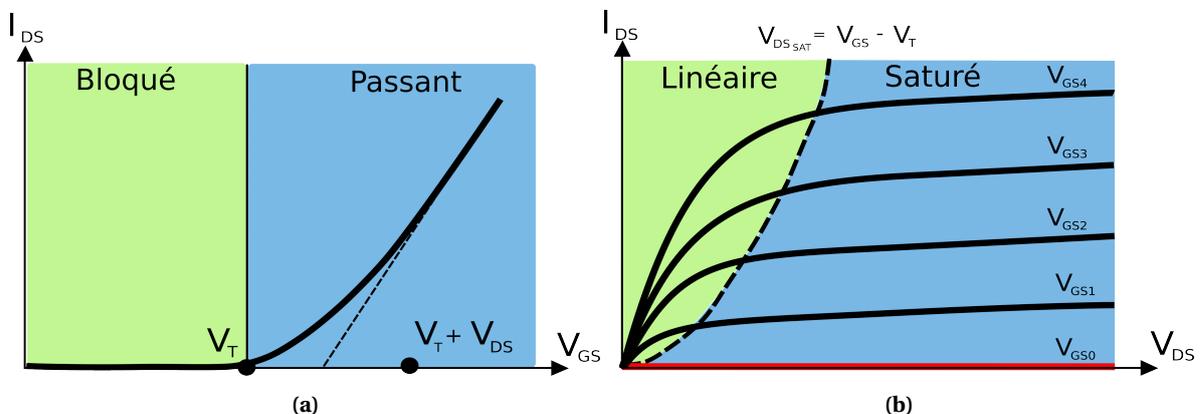


FIGURE 1.5 – (a) Caractéristique de transfert I_{DS} en fonction de V_{GS} d'un transistor NMOS (b) Caractéristique de sortie I_{DS} en fonction de V_{DS} pour un transistor NMOS avec $V_{GS0} < V_T < V_{GS1} < V_{GS2} < V_{GS3} < V_{GS4}$

Lorsque V_{GS} est supérieur à V_T le canal s'inverse et rend ainsi le transistor passant (voir figure 1.5a). La caractéristique de sortie présente sur la figure 1.5b montre que pour un V_{GS} fixe supérieur à V_T , le régime de fonctionnement du transistor, linéaire ou saturé, dépend alors de la tension V_{DS} . Lorsque $V_{DS} < V_{GS} - V_T$ alors le transistor est en régime linéaire puisque le courant de sortie évolue linéairement suivant V_{DS} . Ce régime correspond à la zone verte de la figure 1.5b. Lorsque $V_{DS} > V_{GS} - V_T$, le transistor est en régime saturé puisque la tension V_{DS} n'a presque plus d'effet sur le courant de sortie qui reste stable. L'intensité de courant de sortie dépend de V_{GS} et évolue presque quadratiquement. Ce régime correspond à la zone bleu de la figure 1.5a.

Finalement, nous retenons que le transistor fonctionne suivant trois régimes : bloqué, passant linéaire et passant saturé. Le tableau 1.1 récapitule les conditions pour être dans les différents modes de fonctionnement et donne le courant de drain associé pour un transistor NMOS et PMOS.

1.2.3 Évolution du FET

Dans cette section nous présentons l'évolution du FET depuis sa création jusqu'à aujourd'hui afin de comprendre dans quel contexte le transistor FinFET a été introduit dans les circuits numériques. Les années 1960 sont marquées par l'émergence du MOSFET dans l'industrie. Une revue depuis la création du MOSFET à son implémentation dans des VLSI (Very-Large-Scale Integration) est faite dans [16]. De sa création à aujourd'hui le MOSFET n'a cessé d'évoluer afin d'être plus performant, d'avoir des dimensions plus petites et un coût plus faible. La figure 1.6 révèle trois périodes qui ont marqué l'évolution du MOSFET et que nous détaillons.

Le dimensionnement proportionnel [1970-2000]

Dans [17] Robert Dennard explique comment diminuer les dimensions du transistor avec un facteur proportionnel k . La diminution du canal $\frac{L}{k}$ accentue l'effet canal court qui diminue la tension de seuil du transistor est en conséquence augmente le courant I_{OFF} . Une explication plus complète de l'effet canal court est donnée dans [18, chap VI]. Afin de limiter l'effet canal court il est possible de diminuer la largeur du canal qui dépend de la largeur du drain et de la source $\frac{X_d}{k}$ ce qui implique l'augmentation de la concentration de dopant kN_A . Pour avoir un meilleur contrôle en tension du canal il faut diminuer l'épaisseur de l'oxyde de grille $\frac{T_{ox}}{k}$ ce qui rend le transistor plus sensible au *breakdown*. Par conséquent, il est nécessaire de réduire la tension de grille $\frac{V_G}{k}$. Comme nous venons de voir, la simplicité du dimensionnement des transistors réside dans la relation linéaire entre les différentes grandeurs. Cependant, l'épaisseur d'oxyde de grille est devenu le paramètre qui mettra fin au dimensionnement avec la méthode Dennard. En effet, à partir du noeud 90nm l'oxyde de grille a une épaisseur de 1.2nm et les charges peuvent passer au travers de l'oxyde par effet tunnel augmentant considérablement le courant de fuite. Une description détaillée de l'effet tunnel est donnée dans [18, chap IV.6].

Le dimensionnement par équivalence [2000-2020]

L'application de contraintes mécaniques a permis d'améliorer la mobilité des charges [19]. Pour le PMOS, il s'agit de SiGe qui est déposé dans le drain et la source pour appliquer une contrainte de compression au niveau du canal afin d'améliorer la mobilité des trous. Pour le NMOS, il s'agit de SiN qui est déposé dans le drain et la source pour appliquer une contrainte de tension sur le canal et ainsi améliorer la mobilité des électrons. L'application de contraintes a permis d'améliorer les performances du transistor, cependant, le courant de fuite au niveau de l'oxyde reste toujours problématique.

À partir du noeud 45nm [20] l'oxyde de silicium utilisé comme diélectrique de la grille est remplacé par un matériau dit *high- κ* , c'est à dire avec une constante diélectrique élevée. L'utilisation d'un tel matériau permet d'augmenter l'épaisseur du diélectrique, de diminuer ainsi les courants de fuites par effet tunnel, tout en préservant un bon contrôle du champ électrique de grille qui dépend de la capacité de l'oxyde de grille ($C_{ox} \propto \frac{\epsilon_k}{t_{ox}}$). La diminution de la taille de la grille a pour conséquence l'augmentation de la résistance et donc la diminution de performance. C'est pourquoi, un métal avec une résistivité plus faible que le polysilicium est utilisé pour la grille. Avec l'utilisation des contraintes, du *high- κ* et de la grille en métal, il a été possible d'atteindre le noeud 32nm en maintenant de bonnes performances.

À partir de 22nm [21] le MOSFET planaire est remplacé par le transistor 3D FinFET. Le FinFET a des qualités de contrôle électrostatique de la grille qui ont permis de diminuer la taille du transistor tout en préservant des bonnes performances. Une explication plus détaillée du FinFET sera donnée dans la section 1.2.4.

Vers un dimensionnement sub-nanonométrique [2020-futur]

Cette période décrite dans [22] permet d'avoir une vision de l'évolution du transistor pour les années à venir. Deux grands axes orientent cette évolution. Le premier est le développement

de nouveaux transistors qui sont développés dans la continuité du CMOS, à l'exemple du "Gate All Around (GAA)" attendu en production pour 2025 [23] et du "Tunneling" FET [24], ou des technologies qui sont différentes du CMOS classique, "Beyond CMOS" [25], et qui permettront de le remplacer. Le deuxième axe est l'évolution de l'architecture du circuit. Jusqu'à présent, comme on peut voir sur la figure 1.11, les transistors dans un circuit sont répartis sur un niveau logique unique. Il est prévu pour 2031 la production de circuit avec un empilement de niveaux logiques [23].

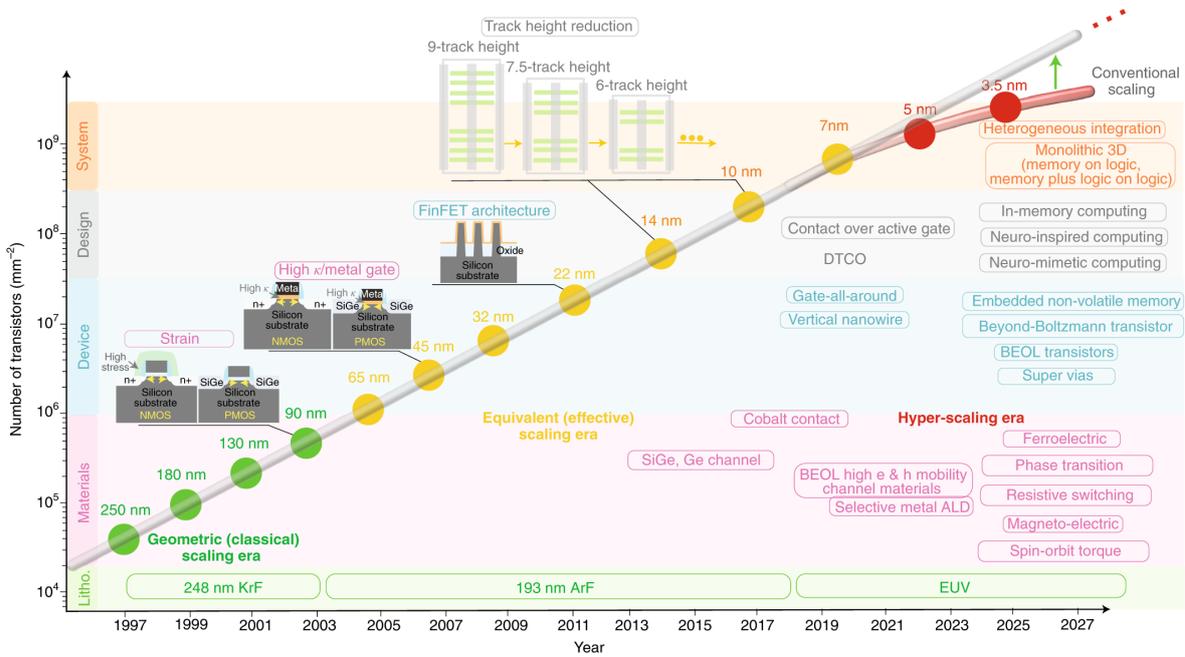


FIGURE 1.6 – Évolution de la densité de transistors entre 1997 et 2027 montrant l'évolution technologique du transistor divisée en trois périodes : dimensionnement proportionnel (points vert), dimensionnement par équivalence (points jaune) et le dimensionnement sub-nanométrique (rouge). Cette évolution de dimensionnement est possible grâce à l'introduction de nouveaux outils de lithographie (zone verte), de nouveaux matériaux (zone rose), de nouveaux dispositifs et architectures de conception (zone bleu) et à de nouveaux systèmes (zone orange) [22]

1.2.4 L'architecture FinFET

L'architecture FinFET qui est le transistor utilisé dans les circuits numériques de dernière génération. Le principe de fonctionnement du FinFET est identique à celui du MOSFET.

La réduction de la largeur du canal d'un MOSFET est limitée principalement par deux paramètres : le *Drain Induced Barrier Lowering* (DIBL) et le *Subthreshold Slope* (SS) [26]. Le DIBL, exprimé en mV/V, correspond au rapport de la variation de la tension de seuil sur la variation de la tension du drain. Le SS correspond à la pente du logarithme du courant de drain en fonction de la tension de la grille lorsque la tension est inférieure à la tension de seuil. Lorsqu'un MOSFET a un canal court, alors le DIBL ainsi que le SS sont importants ce qui a pour conséquence la diminution de la tension de seuil et l'augmentation du courant I_{OFF} .

Les FET à double grille (DG), dans lesquels une deuxième grille est ajoutée à l'opposé de la grille traditionnelle, sont reconnus depuis longtemps pour leur potentiel à mieux contrôler les effets de canal court (DIBL et SS) [27]. Le FinFET est un transistor, de la famille DG-FET, avec un canal vertical qui sort du plan, en forme d'ailette (en anglais *Fin*), contrairement au MOSFET qui a un canal horizontal (voir figure 1.7a et 1.7b). La densité de courant que peut conduire le transistor ne dépend plus uniquement de la largeur du canal (W) mais aussi de sa hauteur (H). La grille d'un

FinFET recouvre le canal qui est posé sur le substrat améliorant ainsi le contrôle électrostatique du transistor. Comme on peut voir sur la figure 1.8a, le DIBL et le SS d'un FinFET sont inférieurs à ceux d'un MOSFET [26] et par conséquent la consommation ainsi que la vitesse de commutation du transistor sont améliorées (voir figure 1.8b) [28].

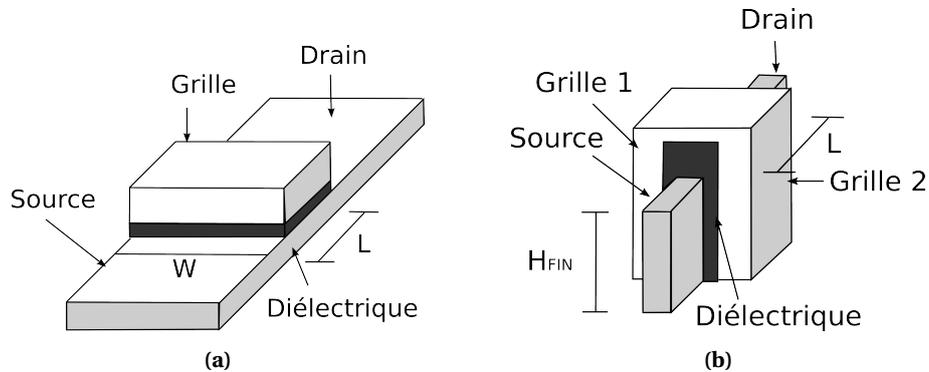


FIGURE 1.7 – (a) Schéma d'un MOSFET (b) Schéma d'un FinFET

De plus, le processus de fabrication du FinFET s'inscrit dans la continuité de celui du MOSFET, ce qui est un avantage considérable au regard du coût et du temps de développement. Par exemple, il est possible d'utiliser le même *wafer* et le même processus de fabrication pour fabriquer les tranchées d'isolations dans un MOSFET que pour fabriquer les ailettes du FinFET [26]. Pour avoir un courant de sortie acceptable, il est nécessaire de connecter les FinFETs en réseaux, c'est pourquoi on parle de courant par unité de surface. La densité d'intégration est délimitée par l'espacement (le pas) entre deux FinFETs dont la réduction a permis de passer progressivement du noeud 22nm au noeud 5nm [29, 30].

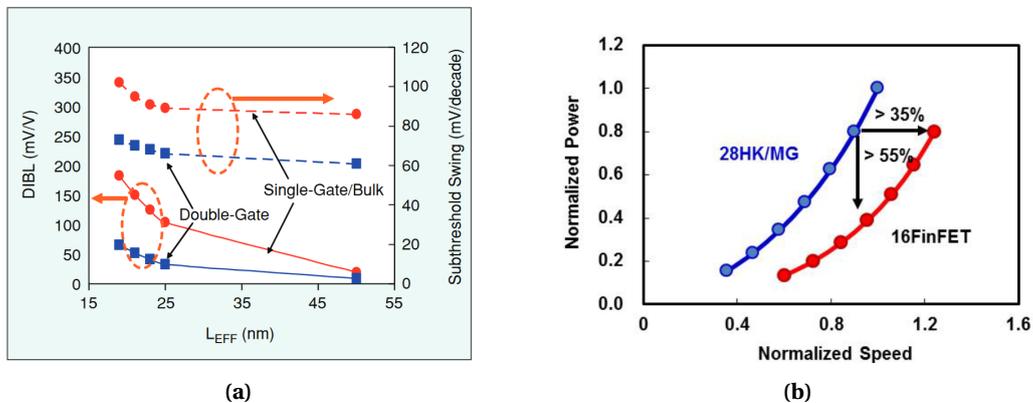


FIGURE 1.8 – (a) Comparaison entre le MOSFET et FinFET du DIBL et du SS en fonction de la longueur L [26] (b) Comparaison des performances entre le MOSFET et le FinFET [28]

1.3 Circuits élémentaires CMOS dans les FPGAs

Jusqu'à présent nous avons introduit le transistor comme un composant isolé. Les circuits numériques utilisent les architectures *Complementary Metal Oxide Semi-conductor* (CMOS) associant des transistors PMOS et NMOS. Dans un FPGA par exemple, le CMOS est utilisé pour faire des portes élémentaires telles que l'inverseur ou la porte de transfert. Un circuit numérique possède parfois de millions de portes élémentaires qui sont connectées entre elles et vers les entrées et sorties de la puce grâce à une architecture d'interconnexion. Pour cela dans cette section, nous introduisons le fonctionnement de l'inverseur et de la porte de transfert qui sont fortement présents dans un FPGA ainsi que l'architecture d'interconnexion.

1.3.1 L'inverseur

L'inverseur CMOS est un élément de base des circuits numériques [31]. Comme nous pouvons voir sur la figure 1.9a la grille du NMOS et du PMOS sont connectées ensemble et forment l'entrée de l'inverseur. Les drains du NMOS et du PMOS sont connectés ensemble et forment la sortie. La source du PMOS est connectée à la tension d'alimentation V_{dd} et la source du NMOS est connectée à la masse du circuit.

Lorsque l'entrée de l'inverseur n'a pas changé depuis un certain alors on considère que celui-ci est en régime statique. Dans ce régime, si $V_{IN} = V_{dd}$ alors $V_{OUT} = 0V$ et si $V_{IN} = 0V$ alors $V_{OUT} = V_{dd}$.

Lorsque l'entrée de l'inverseur change, alors celui-ci est en régime transitoire le temps que la sortie se stabilise. La figure 1.9b montre la caractéristique de sortie correspondant au courant passant dans les transistors en fonction du niveau de la tension de sortie. Si $V_{IN} = V_{dd}$ alors le NMOS est passant et le PMOS bloqué. Par conséquent, le courant présent dans le condensateur en sortie se décharge progressivement par le NMOS jusqu'à ce que $V_{OUT} = 0V$. Si $V_{IN} = 0V$ alors le PMOS est passant et le NMOS bloqué. Par conséquent, le courant passant par le PMOS charge progressivement le condensateur en sortie jusqu'à ce que $V_{OUT} = V_{dd}$. Donc, en régime transitoire, un front montant en entrée génère un front descendant en sortie tandis qu'un front descendant en entrée génère un front montant en sortie.

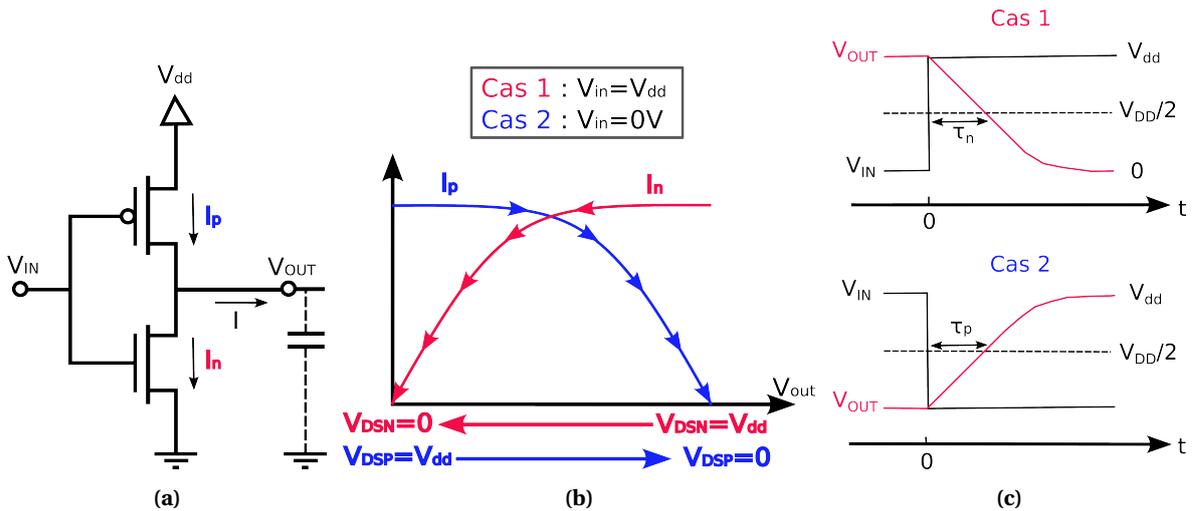


FIGURE 1.9 – (a) Schéma électrique d'un inverseur avec un condensateur en sortie pour modéliser la connection à l'étage logique suivant (b) Courant passant dans les transistors pendant les transitions (I_p pour le PMOS et I_n pour le NMOS) en fonction du niveau de tension en sortie V_{OUT} et avec l'écoulement du temps représenté par les flèches sur les courbes (c) Temps de propagation d'un inverseur dans le cas d'un front descendant en sortie (haut) et d'un front montant en sortie (bas)

Le temps de propagation d'un inverseur est une caractéristique importante dans les circuits numériques. Lorsqu'il y a un changement d'état au niveau de l'entrée, le temps de propagation est défini comme la différence de temps entre le moment où $V_{OUT} = \frac{V_{dd}}{2}$ et $V_{IN} = \frac{V_{dd}}{2}$. Une représentation de l'évolution de la tension d'entrée et de sortie en fonction du temps est donnée en figure 1.9c. Alors que la figure 1.9b montre que l'évolution du niveau de V_{OUT} dépend du temps de charge et de décharge du condensateur, c'est à dire à la capacité aux transistors PMOS et NMOS de conduire du courant. Il est possible d'exprimer approximativement le temps de propagation par :

$$\tau = \frac{CV_{dd}}{2I_{ON}} \quad (1.1)$$

où I_{ON} est le courant (I_n ou I_p) passant dans le NMOS ou le PMOS en fonction du niveau logique transmit. Comme le PMOS conduit le courant par transport de trous et le NMOS par trans-

port d'électrons et que la mobilité des trous est plus faible que celle des électrons, l'inverseur possède des temps de charge (τ_P) et de décharge (τ_N) différents [32]. Il est possible d'équilibrer les temps de propagation des deux modes en modifiant la largeur des transistors (W). Dans les circuits numériques de dernière génération, les transistors PMOS et NMOS sont équilibrés en terme de temps de propagation.

1.3.2 La porte de transfert

La porte de transfert (*Transmission Gate* TG) est un circuit dont la fonction est de transmettre ou non un signal. Avec un ensemble de portes de transferts il est possible de réaliser des fonctions essentielles dans les circuits numériques tel que le multiplexeur dans les LUT (Look-Up Table) des FPGA (voir section 1.4.3). Dans la porte de transfert CMOS, représenté sur la figure 1.10a, un PMOS et un NMOS sont montés en parallèle, les grilles sont commandées par une tension V_C pour le NMOS et son complémentaire $\overline{V_C}$ pour le PMOS. Avant la porte de transfert, les concepteurs utilisaient des *passtransistors*. Le *passtransistor* est constitué uniquement d'un type de transistor, souvent du NMOS. Cependant, l'utilisation d'un seul type de transistor crée un déséquilibre du temps de propagation d'un front montant et descendant et génère des chutes de tension en sortie ce qui limite la succession d'étage de *passtransistor*. L'utilisation d'un NMOS et d'un PMOS dans les TG rend possible l'acheminement d'un niveau logique haut et bas avec des chutes de tension en sortie bien plus faible que les chutes de tension présentes dans les *pass transistors* et donc, des temps de propagation des fronts montants et descendants équilibrés.

Lorsque $V_C = 0$ les transistors NMOS et PMOS sont bloqués. Dans ce cas, le circuit se comporte comme un interrupteur ouvert. Lorsque $V_C = V_{dd}$ les transistors NMOS et PMOS sont passants, alors un signal peut être transmis. Lorsque les transistors sont passants, deux cas de figure sont possibles : transmettre un niveau logique haut ou transmettre un niveau logique bas. Deux schémas de l'évolution du mode de fonctionnement des transistors NMOS et PMOS pour la transmission d'un niveau haut et bas sont donnés en figure 1.10b et 1.10c.

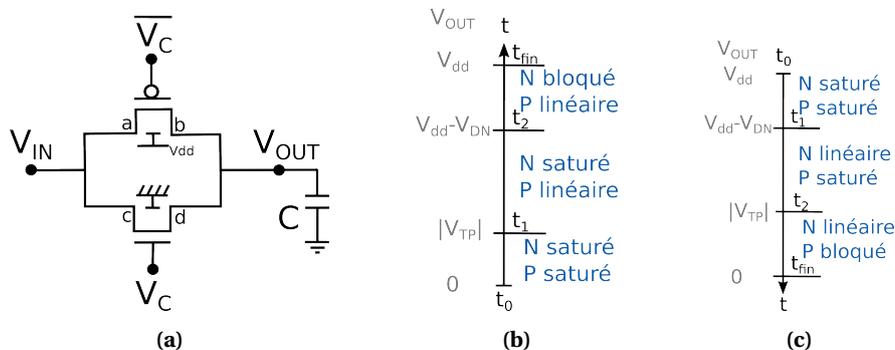


FIGURE 1.10 – (a) Schéma d'une porte de transfert (b)&(c) Évolution du mode de fonctionnement du NMOS et du PMOS pour la transmission d'un niveau haut & bas

Transmission d'un niveau logique haut

À t_0 , nous considérons que $V_{IN} = V_{dd}$, $V_C = V_{dd}$ et $V_{OUT} = 0$. Dans ce cas, sur le schéma 1.10a, la source et le drain sont respectivement a et b pour le PMOS et d et c pour le NMOS. Entre t_0 et t_1 la tension grille-source des deux transistors est supérieure à leurs tensions de seuil respective, les deux transistors sont donc passants et conduisent le courant vers le condensateur en mode saturé car la tension drain-source remplit les conditions énoncées dans le tableau 1.1. Le condensateur se charge et la tension V_{OUT} augmente jusqu'à atteindre une tension égale à $|V_{TP}|$. Alors, entre t_1 et t_2 , le transistor PMOS conduit le courant en mode linéaire car $V_{DSp} < V_{SGp} - V_{Tp}$, les conditions du NMOS n'ont pas changées. Les deux transistors continuent de conduire le courant, augmentant ainsi la tension V_{OUT} qui, en conséquence, diminue la tension V_{GSn} jusqu'à ce que $V_{OUT} = V_{dd} -$

V_{TN} . Alors le transistor NMOS se bloque mais le courant continu d'être acheminé par le PMOS entre t_2 et t_{fin} jusqu'à ce que $V_{OUT} = V_{dd}$.

Transmission d'un niveau logique bas

Sur le schéma 1.10a, la source et le drain sont respectivement b et a pour le PMOS et c et d pour le NMOS. Tout en suivant la même logique de la précédente explication, dans un premier temps les deux transistors conduisent le courant en mode saturé pour décharger le condensateur, la tension V_{OUT} diminue alors progressivement. La diminution progressive de V_{OUT} va d'abord rendre le mode de conduction du NMOS linéaire puis bloquer le PMOS.

Résistance de la porte de transfert

La porte de transfert se comporte comme un montage de deux résistances en parallèle qui varient suivant le mode de conduction des transistors et suivant le niveau de tension entre le drain et la source qui varie aussi en fonction du niveau de tension en sortie de la porte. Deux points importants sont à noter ici. Le premier point c'est que dans les deux modes de transmissions le PMOS et NMOS contribuent à conduire le courant. Le deuxième point c'est que nous pouvons considérer la porte de transfert comme un montage RC en série avec comme constante de temps $\tau_{TG} = r_{TG}C_{out}$ pour transmettre un signal de l'entrée vers la sortie.

1.3.3 Les interconnexions

Dans les circuits intégrés numériques, les interconnexions connectent les transistors pour conduire l'alimentation, la masse, les signaux logiques et les impulsions d'horloge. En général on retrouve trois types d'interconnexion avec des longueurs différentes : local (10nm-100µm), intermédiaire (500µm-900µm) et global (1000µm et plus). Les interconnexions locales sont utilisées pour connecter différents transistors entre eux. Les interconnexions intermédiaires le sont pour connecter deux blocs logiques. Les interconnexions globales transportent les signaux d'alimentation, de masse et d'horloge vers tout le reste du circuit. Avec l'augmentation de la densité de transistors, l'interconnexion est organisée par niveau, à l'exemple de la figure 1.11 où l'on distingue trois niveaux de métal (M1, M2 et M3). La bille de soudure fait le lien entre le circuit et le monde extérieur (capteurs, alimentation, circuits etc) interagissant avec le circuit. Le niveau de métal M1 est physiquement en contact avec la logique du circuit (transistors). On compte à ce jour des circuits intégrés numériques avec jusqu'à 17 niveaux d'interconnexion [33].

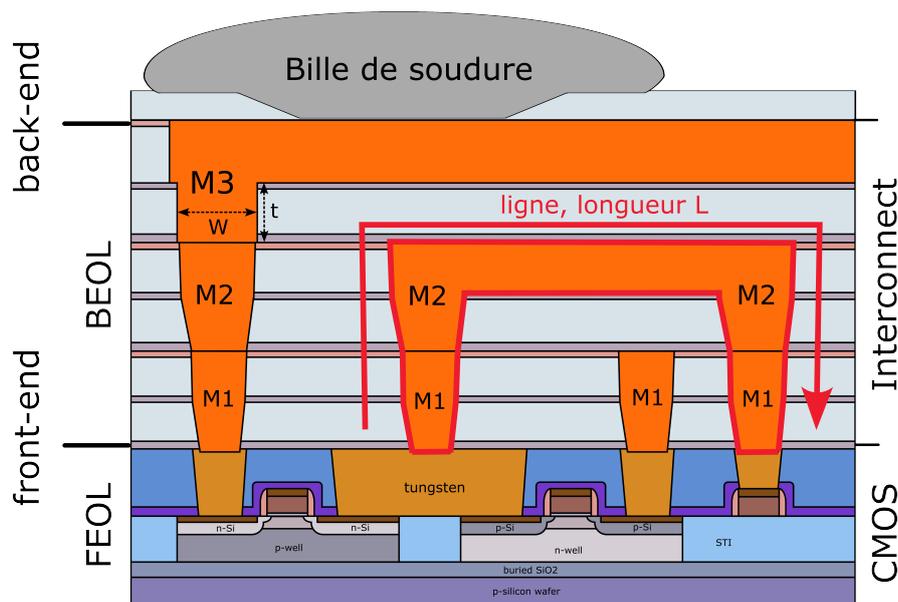


FIGURE 1.11 – Vue en coupe d'un circuit avec trois niveaux d'interconnexions.

Temps de propagation

Les lignes d'interconnexions se comportent approximativement comme un circuit RC série avec une constante du temps de propagation $\tau = RC$. La résistance d'une ligne d'interconnexion dépend de la résistivité électrique et des dimensions de la ligne. Comme illustré sur la figure 1.12, il existe trois principaux effets capacitifs d'une ligne d'interconnexion. Afin d'augmenter la densité de transistors dans les circuits intégrés numériques, la dimension des lignes d'interconnexions n'ont cessé de diminuer ces dernières années. La réduction de la taille (largeur et épaisseur) d'une ligne d'interconnexion augmente sa résistance et donc augmente le temps de propagation [34–36]. Pour réduire la résistance d'une ligne, le cuivre ($\rho = 1.7 \times 10^{-8} \Omega.m$) est utilisé à la place de l'aluminium ($\rho = 2.7 \times 10^{-8} \Omega.m$) [34]. Ensuite, l'utilisation de diélectrique dit "low- κ " a permis de réduire la capacitance de la ligne d'interconnexion [37]. Les matériaux utilisés dans l'interconnexion sont constamment améliorés pour préserver des performances identiques tout en diminuant la taille des lignes et en augmentant leurs longueurs proportionnellement à l'augmentation de la taille des circuits numériques.

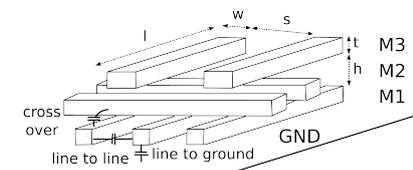


FIGURE 1.12 – Représentation des effets capacitifs dans les lignes d'interconnexion

La courbe de la figure 1.13 [34] permet de comprendre l'évolution des performances de l'interconnexion au cours des deux dernières décennies. On peut voir l'évolution de la résistance (R), de la capacitance (C) et de RC par unité de longueur en fonction du noeud technologique. L'unité de longueur correspond à la distance minimale entre deux lignes pour chaque noeud technologique. Les paramètres R, C et RC sont calculés à partir des données provenant d'autres articles pour différents noeuds technologiques. Malgré l'utilisation du cuivre la résistance par unité de longueur augmente avec la diminution du noeud. L'utilisation de matériaux "low- κ " a accentué la diminution de la capacitance à partir du noeud 90nm. On peut voir sur la figure 1.13 que la diminution de la capacitance par unité de longueur a permis de compenser l'augmentation de la résistance par unité de longueur jusqu'au noeud 32nm. Après le noeud 32nm le temps de propagation RC par unité de longueur a commencé à augmenter.

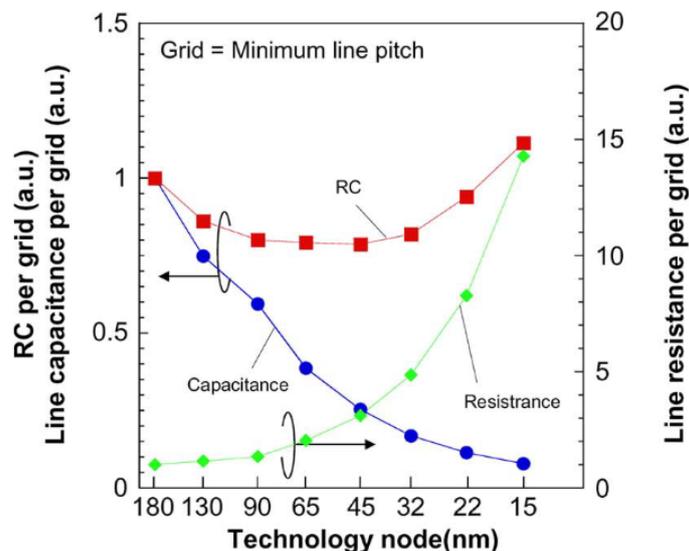


FIGURE 1.13 – Évolution de R, C et RC dans les lignes d'interconnexion pour différents noeuds [34]

1.4 Field-Programmable Gate Array (FPGA)

Le FPGA est le circuit numérique de dernière génération que nous avons utilisé pour développer un banc de test visant à vieillir et mesurer ses dégradations pour prédire la fiabilité des circuits numériques de dernière génération. Pour comprendre les fonctionnalités du banc de test que nous introduisons dans le chapitre 2, l'analyse des dégradations dans le chapitre 3 et la modélisation des dégradations dans le chapitre 4, il est essentiel de connaître certains principes de fonctionnement des FPGA. Ainsi, dans cette partie, après avoir contextualisé l'émergence des FPGA, nous décrivons le fonctionnement de celui-ci en détaillant l'architecture des différents éléments qui le constituent. Nous décrivons le FPGA en nous basant sur le FPGA Zynq UltraScale+ 16nm qui est celui que nous avons utilisé tout au long de cette thèse.

1.4.1 Évolution des Programmable Logic Devices (PLDs)

Dans la section 1.2 nous avons introduit à ce jour il existe un nombre important de circuits numériques que l'on peut classer en quatre catégories :

- Circuit logique standard : à l'exemple des circuits de la famille 74H, il peut avoir plusieurs entrées et sorties mais il réalise une même fonction logique (NOT, AND, OR etc) ;
- Circuit à fonctionnement programmable : tel que le μ processeur ou le μ contrôleur, il exécute des instructions machines ordonnées par un programme informatique ;
- Circuit personnalisé : tel que l'ASIC (*Application-Specific Integrated Circuit*), il contient un grand nombre de fonctionnalités complexes et le circuit est conçu pour répondre à un besoin spécifique ;
- Circuit à logique programmable : tel que les PLDs, un même circuit peut avoir plusieurs configurations architecturales possibles pour répondre à un besoin spécifique ;

Parce que nous mesurons la fiabilité des FPGAs, nous allons dans cette partie contextualiser l'émergence du FPGA en présentant l'évolution des PLDs dont le FPGA est un aboutissement. Pour présenter l'évolution des PLDs, nous nous sommes appuyé sur la lecture de [38, chapitre 3] et de [39, chapitre 3]. La figure 1.14 présente une frise chronologique de l'évolution des PLDs. Le *Simple Programmable Logic Devices* (SPLD), le *Complex Programmable Logic Devices* (CPLD) et le FPGA sont des sous-catégories du PLD.

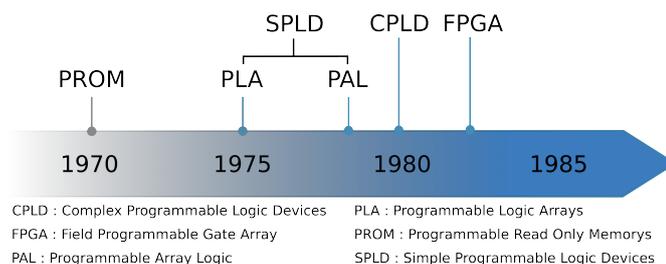


FIGURE 1.14 – Frise chronologique de l'évolution des Programmable Logic Devices (PLDs)

Développé en 1970, le *Programmable Read Only Memory* (PROM) est la technologie à l'origine des circuits logiques programmable. Le PROM est programmable une seule fois et enregistre une information qui est conservée même lorsque le circuit qui la contient n'est plus alimenté. Il existe des PROM reconfigurable tel que le *Erasable PROM* (EPROM) ou le *Electrically Erasable PROM* (EEPROM).

En 1975 apparaît le *Programmable Logic Array* (PLA). La figure 1.15a présente un schéma simplifié d'un PLA configuré avec trois entrées (a, b et c) et trois sorties (w, x et y). Une porte inverseur est placée au niveau de la sortie pour générer le signal complémentaire de chaque entrée. En fonction de la configuration de la matrice AND et OR, les entrées sont acheminées ou non vers

des portes AND dont les sorties sont acheminées ou non vers des portes OR. Pour être configurable, le PLA utilise des PROM. Sur la figure 1.15a seule la logique combinatoire est représentée, en rajoutant des bascules au niveau des sorties le PLA peut aussi faire de la logique séquentielle. Le PLA permet d'implémenter des fonctions complexes, cependant, le signal traverse un nombre important de connexions configurables ce qui génère un temps de propagation important.

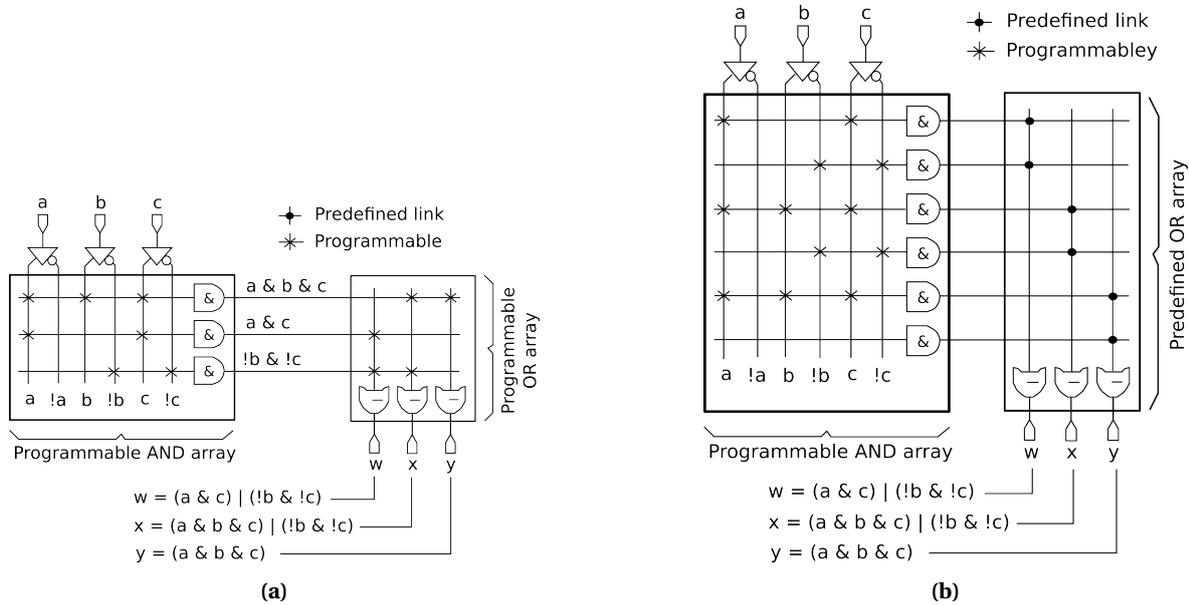


FIGURE 1.15 – Architecture simplifiée d'un (a) PLA et d'un (b) PAL configuré à trois entrées et sorties [39]

Afin d'améliorer le temps de propagation d'un signal dans un PLA, à la fin des années 70', le *Programmable Array Logic* (PAL) a été conçu. Contrairement au PLA où les matrices AND et OR sont configurables, dans le PAL, seule la matrice AND est configurable et la matrice OR est pré-configurée (voir figure 1.15b). Parce que le signal traverse moins de connexions programmables que dans un PLA le temps de propagation est amélioré. Cependant, le nombre de fonctions générées par le PAL sont limitées par le nombre de portes OR implémentées.

L'émergence des trois catégories de SPLD (Simple PLD) soulève un dilemme conceptuel inévitable : soit le circuit est plus rapide mais consomme plus d'énergie à cause du nombre de portes logiques (PROM et PAL), soit le circuit est moins rapide mais consomme moins d'énergie (PLA). C'est dans ce contexte que le CPLD (Complex PLD) est créé dans les années 1980. Le CPLD est un circuit qui combine des EPROM, qui sont des blocs logiques reconfigurables, avec du CMOS. Cette combinaison des deux technologies permet de diminuer considérablement la consommation du circuit. Pour qu'un SPLD soit fonctionnel, il est nécessaire que la matrice d'interconnexion connecte toutes les entrées et sorties aux blocs logiques. L'innovation dans les CPLD est la conception d'une matrice d'interconnexion configurable qui complexifie l'architecture du circuit mais qui permet au circuit de moins consommer d'énergie et d'améliorer le temps de propagation d'un signal.

Dans les années 80' il y a eu un écart important entre les SPLD/CPLD qui sont reconfigurables mais qui assurent des fonctions peu complexes et les ASIC qui implémentent des fonctions complexes mais qui sont non reconfigurables. C'est dans ce contexte que le FPGA est développé. Grâce à l'utilisation de LUTs (Look-Up Tables), de bascules et de multiplexeurs, par rapport aux SPLD et CPLD, le FPGA assure des fonctions plus complexes. La miniaturisation des composants de la création du FPGA jusqu'à aujourd'hui a rendu les FPGAs plus denses. Alors que le premier FPGA commercialisé par Xilinx (XC2064) en 1985 a 64 blocs logiques [40], en 2020, le FPGA XCZU3EG que nous utilisons a 141120 blocs logiques [41].

1.4.2 Structure globale

Après avoir introduit l'évolution des PLDs dont le FPGA est l'aboutissement, dans cette partie nous présentons l'architecture global d'un FPGA. La figure 1.16a présente un schéma de principe de l'architecture d'un FPGA. Des blocs d'entrées/sorties IO (rouge) connectent le FPGA vers l'extérieur. Les blocs logiques (*Configurable Logic Blocks* CLB), en bleu, contiennent la fonctionnalité du FPGA. Le réseau d'interconnexion distribue les signaux depuis les IO et entre les blocs logiques. Pour ce faire il est composé de lignes d'interconnexions (courtes ou longues), en noir, qui vont dans les quatre directions, ainsi que de matrices d'interconnexions (vert) qui sont configurées pour raccorder une ligne à une autre afin de créer un chemin. Des matrices d'interconnexions IMUX (orange) servent d'interfaces entre le réseau d'interconnexion et les CLB. Toutes les composantes du FPGA possèdent des mémoires de configuration contrôlées par un flux de données (bitsream).

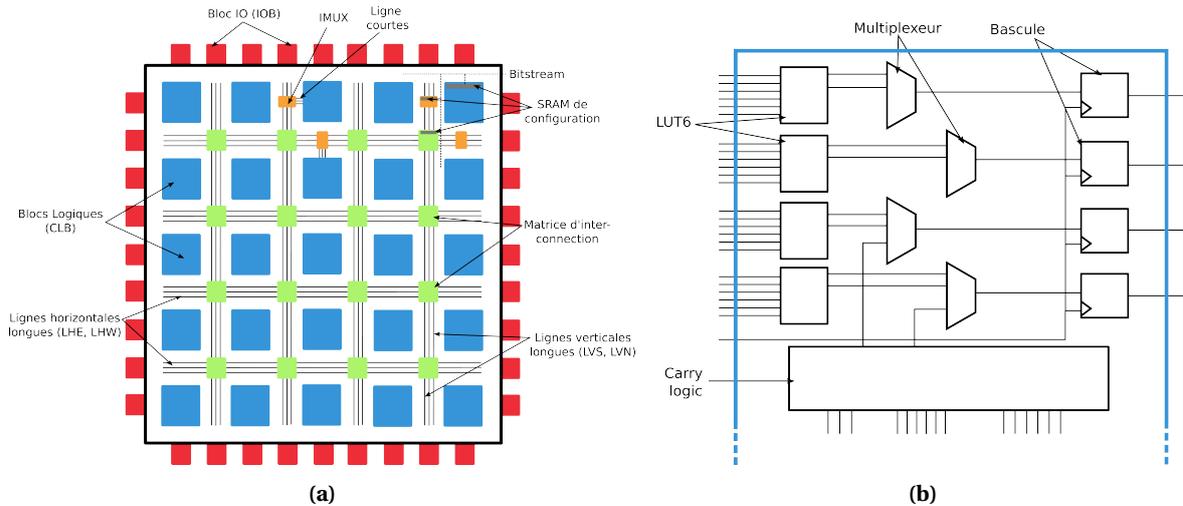


FIGURE 1.16 – (a) Vue d'ensemble d'un FPGA (b) Architecture interne de la moitié d'une sliceL

La CLB est la principale ressource pour la mise en oeuvre d'un circuit combinatoire et séquentiel à usage général. Dans les FPGA Zynq UltraScale+, chaque CLB contient ce qui est appelé une slice [42]. Il existe deux types de slices : Les sliceM (Mémoire) sont utilisées pour stocker l'information et les sliceL (Logique) pour les applications logiques. En plus d'une slice, la CLB contient une matrice d'interconnexion pour assurer son routage (voir partie 1.4.4). La figure 1.16b présente une vue d'ensemble de l'architecture interne d'une sliceL d'un FPGA Zynq UltraScale+. Une sliceL contient huit LookUp Tables (LUTs) à six entrées, seize bascules (Flip-Flops FF), huit multiplexeurs ainsi qu'un carry logic. Les LUTs sont des générateurs de fonctions logiques combinatoires. La fonction principale d'une FF est de passer d'une logique combinatoire à une logique séquentielle en synchronisant l'entrée sur l'horloge du module par exemple. Le rôle des multiplexeurs consiste à assembler plusieurs sorties de LUT afin de réaliser des fonctions plus complexes qu'une unique LUT à six entrées. Enfin, le *carry logic* est fourni pour effectuer des additions et des soustractions arithmétiques rapides.

L'intérêt du FPGA c'est sa configurabilité et sa polyvalence. La figure 1.17 décrit le processus de configuration d'un FPGA qui se réalise en plusieurs étapes. Dans l'étape de spécification du besoin, l'utilisateur définit les fonctionnalités ainsi que les contraintes auxquelles le FPGA doit répondre. Le langage de description du matériel (*Hardware Description Language* HDL) est utilisé pour traduire les fonctionnalités et les contraintes précédemment définies en une description comportementale ou structurelle. Le *Very High Speed HDL* (VHDL) et le *Verilog* sont les deux langages HDL les plus couramment utilisés. Un synthétiseur intégré à l'IDE (Vivado chez Xilinx) transforme la description HDL en un schéma *Register Transfer Level* (RTL) qui est une méthode de description des architectures logiques. À ce stade de la conception, il est possible de réaliser des

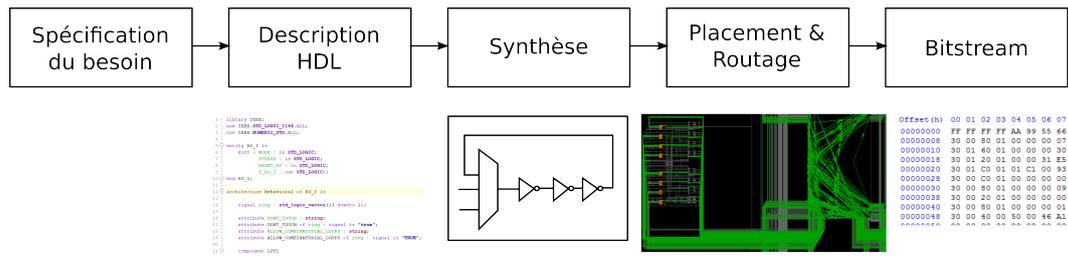


FIGURE 1.17 – Processus de configuration d'un FPGA illustré

simulations pour valider le circuit électronique synthétisé. Un algorithme de placement et de routage transforme le circuit électronique synthétisé en un circuit électronique utilisant les ressources (logique et routage) du FPGA utilisé. Par un exemple, le circuit électronique de la figure 1.17 est un oscillateur en anneau avec un multiplexeur et trois inverseurs. L'algorithme de placement et de routage utilisera les ressources logiques disponibles, soit une LUT configurée en multiplexeur et trois LUT configurées en inverseur. La connexion entre chaque porte logique sera faite par le réseau d'interconnexion en configurant les matrices d'interconnexions afin de connecter la sortie de la première LUT vers l'entrée la suivante etc. À ce niveau de la conception, il est possible de réaliser une simulation de l'architecture en tenant compte des ressources utilisées et notamment des temps de propagation. Enfin, un dernier algorithme traduit les informations de placement et de routage en un fichier bitstream. Le bitstream est un fichier binaire dont chaque bit est stocké dans une cellule mémoire de configuration dans le FPGA.

1.4.3 La LookUp Table (LUT)

La LUT est un circuit électronique configurable qui est largement utilisé dans les CLB d'un FPGA pour assurer la logique combinatoire. Une LUT à n entrées peut réaliser n'importe laquelle des 2^{2^n} fonctions combinatoires possibles grâce à ses 2^n bits de configuration stockant la table de vérité. La modularité d'une LUT est son plus grand avantage et c'est pourquoi il est tant utilisé dans le FPGA. Dans le banc de test que nous avons développé et qui sera présenté dans le prochain chapitre, nous utilisons les LUTs pour implémenter des oscillateurs en anneaux afin de mesurer les dégradations dans un FPGA. Pour que le développement du banc de test soit juste et afin d'interpréter les dégradations des oscillateurs en anneaux, il est nécessaire de connaître le fonctionnement et l'architecture d'une LUT. Ainsi, dans cette partie nous présentons dans un premier temps le fonctionnement général d'une LUT et, dans un deuxième temps, en nous basant sur différents brevets développés par Xilinx, nous détaillons l'architecture interne d'une LUT dans un FPGA Zynq UltraScale+ que nous utilisons.

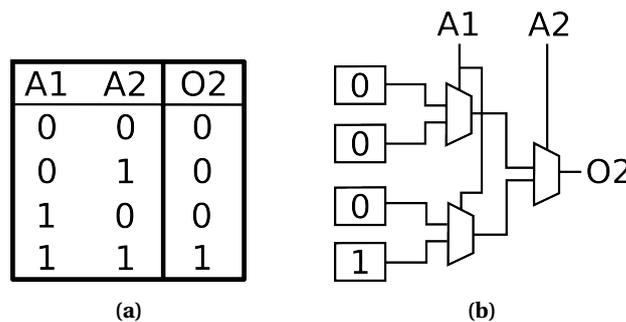


FIGURE 1.18 – (a) Table de vérité d'une porte AND à deux entrées (b) Architecture d'une LUT à deux entrées configuré en porte ET

Fonctionnement général d'une LUT

La table de vérité d'une porte ET est représentée sur la figure 1.18a. La figure 1.18b représente le schéma d'une LUT2 (LUT à deux entrées) configurée pour assurer la fonction d'une porte ET. La LUT2 est constituée de multiplexeurs commandés par les entrées A1 et A2. On constate que l'état de sortie des multiplexeurs dépend des bits de configurations qui correspondent à la sortie O2 de la table de vérité 1.18a. En effet, considérons que les multiplexeurs transmettent l'entrée 1 (en haut) lorsque le signal de contrôle vaut 0, sinon ils transmettent l'entrée 2 (en bas). Dans ce cas, la sortie O2 vaut 1 seulement si A1 = 1 et A2 = 1 sinon la sortie O2 vaut 0; nous avons bien une LUT2 qui assure la fonction d'une porte ET.

Description de l'architecture interne d'une LUT dans un FPGA Xilinx

Les FPGAs Zynq UltraScale utilisent des LUT à six entrées (de A1 à A6) et deux sorties indépendantes (O5 et O6). La LUT peut implémenter :

- Une fonction booléenne à six entrées;
- Deux fonctions booléennes différentes à cinq entrées, à conditions que les deux fonctions partagent les mêmes entrées;
- Deux fonctions booléennes différentes à trois et deux entrées indépendantes;

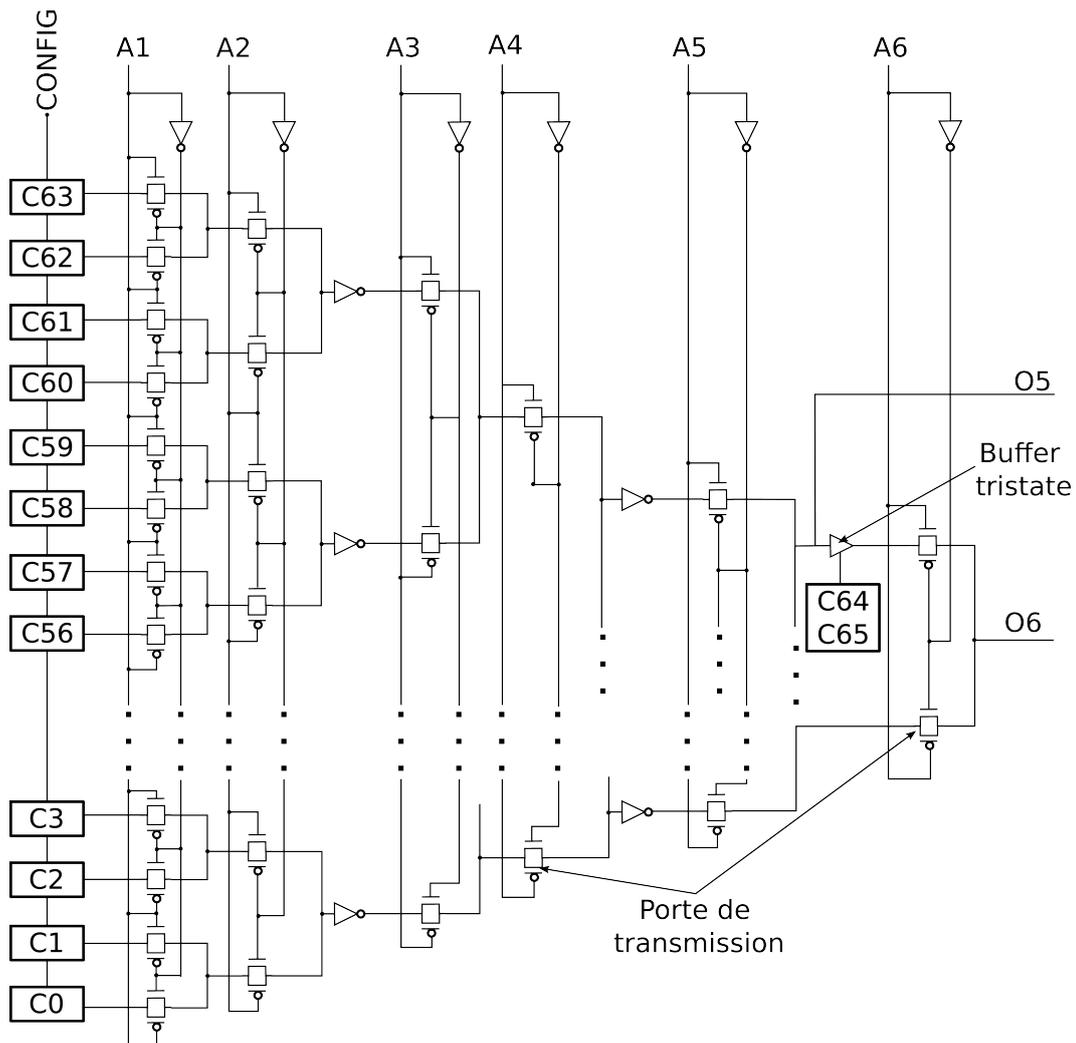


FIGURE 1.19 – Reconstruction de l'architecture interne d'une LUT basée sur l'interprétation des brevets de Xilinx [43–45]

La figure 1.19 représente l'architecture interne d'une LUT. Cette représentation se base sur les informations données dans le document technique du FPGA Zynq UltraScale+ [42] et de trois brevets de chez Xilinx [43–45].

Le premier brevet [43] présente une architecture de LUT à base de portes de transfert. Les architectures de LUT qui ont précédé ce brevet utilisaient des *pass-transistor* n'utilisant que des NMOS. Bien que le temps de propagation d'un signal dans un *pass-transistor* soit plus court que dans une porte de transmission, une chute de tension de l'ordre de la tension de seuil est présente en sortie. Avec la diminution de la taille des transistors, la tension de fonctionnement a diminué et, par conséquent la chute de tension en sortie d'un *pass-transistor* n'est plus devenu négligeable par rapport à la tension de fonctionnement, surtout pour des montages en cascade comme on retrouve dans les LUTs.

Le deuxième brevet [45] introduit l'utilisation d'un buffer tristate (voir figure 1.20a) afin d'utiliser une LUT en plusieurs LUTs de plus petites tailles. Deux bits, C64 et C65, configurent le circuit et déterminent l'état de sortie. Comme on peut voir sur la table de vérité 1.20b, la sortie peut prendre la valeur de l'inverse de l'entrée, ou alors, la sortie prend la valeur de GND ou V_{DD} , dans ce cas la sortie est déconnectée de l'entrée. Dans la LUT, le buffer tristate est placé entre le cinquième et le sixième étage (voir figure 1.19). Lorsque la LUT est utilisée pour implémenter une fonction booléenne à six entrées, alors le buffer tristate transmet le signal du cinquième étage vers le sixième. Lorsque la LUT est utilisée pour implémenter deux fonctions booléennes alors le buffer tristate déconnecte le cinquième étage du sixième et l'état en sortie du buffer tristate dépend alors de la fonction implémentée.

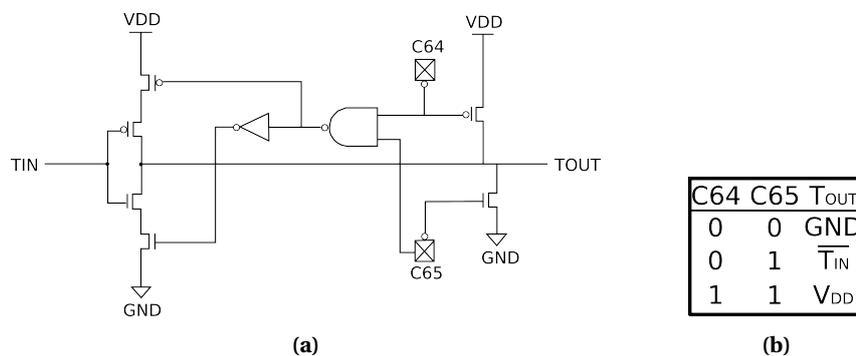


FIGURE 1.20 – (a) Exemple d'une architecture de buffer tristate utilisé pour transformer une LUT en plusieurs LUTs plus petites [45] (b) Table de vérité du buffer tristate

Enfin le troisième brevet [44] propose une architecture interne d'une LUT à six entrées. Dans cette architecture on constate que les portes de transfert fonctionnent par paire pour assurer la fonction d'un multiplexeur à deux entrées pilotées par les six signaux d'entrées de la LUT (A1-A6) et leurs complémentaires. La configuration de la LUT se fait par les 64 bits de configuration (C0-C63). Des inverseurs sont disposés entre les étages pour amplifier le niveau de tension.

1.4.4 Le routage

Dans un FPGA, la communication entre les CLB se fait à travers des circuits de routage qui ont un effet important sur les temps de propagation. C'est pourquoi nous allons décrire l'architecture interne du routage dans un FPGA.

Dans un FPGA, le routage doit répondre aux trois critères suivants : flexibilité du routage vers les autres blocs pour avoir une bonne modularité, un temps de propagation minimale et un espace occupé minimal. De par sa complexité, le routage d'un FPGA a suscité de nombreux brevets visant à en optimiser l'architecture. La figure 1.21 est une représentation légèrement simplifiée d'une

CLB avec les ressources de routage (en vert) et une slice (en bleu) d'après le brevet [46]. Dans la slice on retrouve une partie logique contenant l'architecture présentée sur la figure 1.16b, un multiplexeur d'entrée (IMUX) et de sortie (OMUX) qui interfacent la cellule logique et le réseau d'interconnexion. En sortie du bloc logique, les signaux sont : soit redirigés dans la même slice, soit redirigés vers la matrice d'interconnexion de la slice. Sur la figure 1.21 on distingue trois types de fils de connexions :

- Les chemins courts (N, S, W, E) communiquent les signaux vers les CLB voisines ;
- Les hex-lines (6VN, 6VS, 6HW, 6HE) communiquent les signaux vers les CLB qui sont à une distance de 3 ou 6 CLB avec la possibilité d'allonger le routage en les connectant à une nouvelle hex-lines ;
- Les chemins longs (LVN, LVS, LHW, LHE) communiquent les signaux vers les CLB qui sont à une distance de 12 CLB ;

Le routage des CLB est identique dans tout le FPGA, simplifiant l'architecture du FPGA et donc la gestion de l'implémentation. Sur la figure 1.21, les triangles représentent des points de connexions programmables entre deux fils (Programmable Interconnection Points PIPs). Les carrés repré-

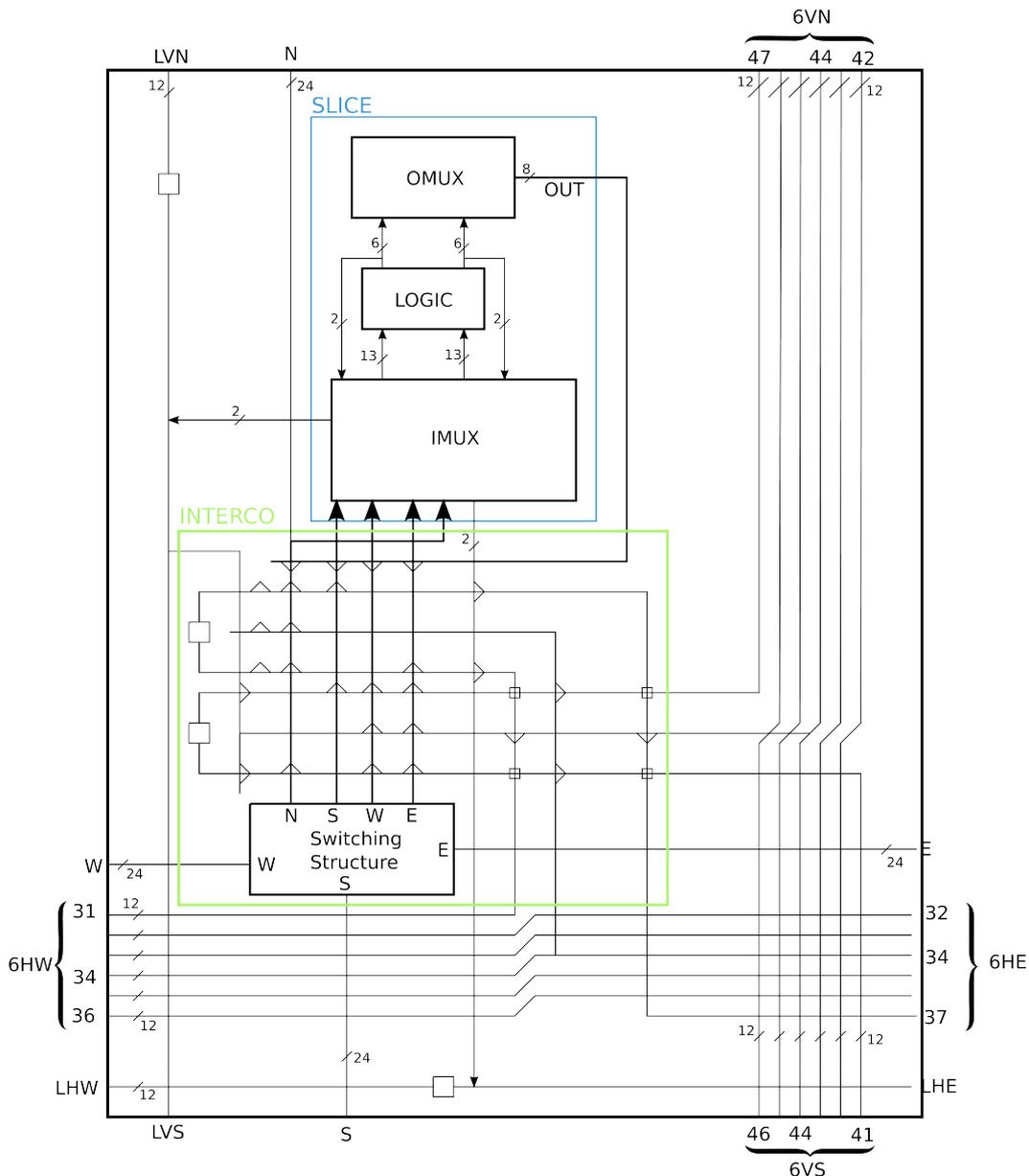


FIGURE 1.21 – Architecture du routage dans une CLB [46]

sentent des connexions de bus programmable qui sont eux même des matrices de PIPs. Le brevet [47] propose une architecture de CLB avec un accès direct vers la structure d'interconnexion afin de supprimer le bloc OMUX. Le brevet[48] propose une architecture d'interconnexion configurable pour fonctionner à faible puissance ou à haute vitesse en fonction de l'application.

La structure de base d'un PIP est la porte de transmission. Il existe trois architectures de PIPs [49, 50] : le cross-point (figure 1.22a), le break-point (figure 1.22b) et le multiplexeur (figure 1.22c). Le PIP cross-point connecte deux lignes de directions différentes alors que le PIP break-point connecte deux lignes de même direction. Il existe deux catégories de PIP multiplexeur : le décodé et le non décodé. Le multiplexeur décodé utilise 2^k cross-point qui sont configurés avec k bits pour sélectionner le signal d'entrée transmis vers la sortie. L'état de sortie dépendra donc des k bits de configuration et des entrées. Le multiplexeur non décodé utilise un bit de configuration par cross-point, dans ce cas seulement un bit de configuration sera sélectionné pour transmettre l'entrée choisit vers la sortie. L'architecture coumpound (figure 1.22d) utilise quatre cross-point et deux break-point afin de créer un noeud avec quatre files de connexions interconnectables dans toutes les directions. Selon [49] les architectures d'interconnexion des FPGAs les plus récents utilisent le PIP multiplexeur non-décodé.

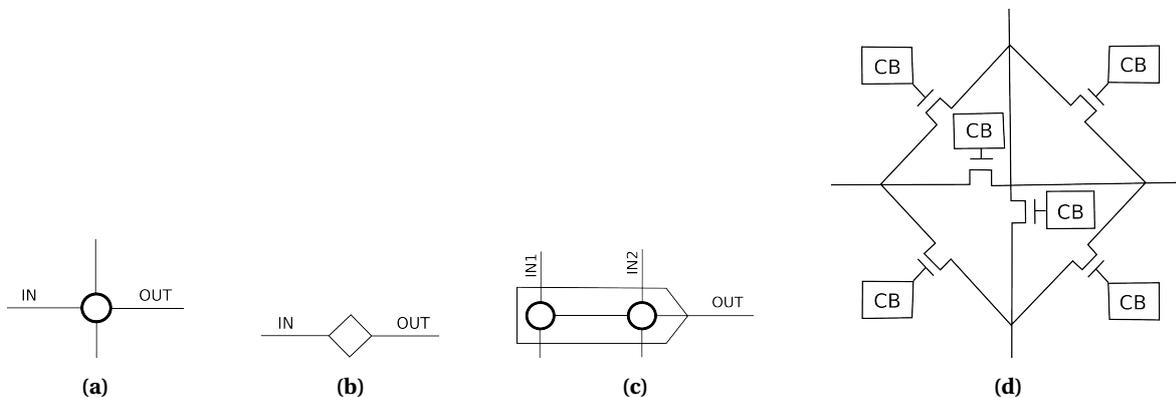


FIGURE 1.22 – Architecture d'un PIP (a) cross-point (b) break-point (c) multiplexeur (d) coumpound [49, 50]

1.4.5 La famille de FPGA Zynq UltraScale+

Depuis la première commercialisation en 1985 jusqu'à aujourd'hui, il est évident que l'architecture des FPGAs n'a cessé d'évoluer. Les deux axes d'améliorations des FPGAs suivent la tendance introduite dans la figure 1.2 : l'évolution d'un circuit passe par la miniaturisation des composants (More Moore) et par la diversification des fonctionnalités (More than Moore). Dans cette partie nous allons introduire ces nouvelles fonctionnalités que l'on retrouve notamment dans le FPGA Zynq UltraScale+ que nous utilisons.

Récemment encore, il existait deux approches distinctes pour concevoir des systèmes numériques dans l'industrie : une approche séquentielle (software) qui peut être intégrée dans des processeurs ou bien des Digital Signal Processors (DSP) et une approche parallèle (hardware) implémentée dans les FPGAs. Les fabricants de circuits intégrés se sont tournés vers le développement de puces capables de réaliser des tâches séquentielles mais aussi des tâches parallèles. Le FPGAs est alors devenu un System on Chip (SoC) très puissant d'où son nom FPSoC (Field Programmable SoC). La combinaison dans une seule puce de processeurs intégrés avec un FPGA a étendu les applications du FPGA dans tous les domaines. Une revue des améliorations récentes des FPSoC et des applications industriels est effectuée dans [51].

Le FPGA Zynq UltraScale+ est considéré comme un FPSoC. Un schéma blocs de son architecture interne est donné sur la figure 1.23 [52]. Le FPGA est divisé en deux parties : la partie processeur (Processing System PS) et la partie logique programmable (Programmable Logic PL).

La PL correspond à l'architecture du FPGA décrite dans la section 1.4.2 avec quelques fonctionnalités supplémentaires telles que des DSP pour exécuter des algorithmes, des blocs RAM pour stocker de l'information, un system monitoring pour contrôler les paramètres tels que la tension interne, et des connecteurs I/O à haut débit pour communiquer les données vers l'extérieur. La PS possède différents éléments tels que quatre APU (Application Processor Unit), deux processeurs temps réels et bien d'autres encore. Il est possible de communiquer depuis la PL vers la PS et inversement en utilisant le protocole AXI (Advanced eXtensible Interface). Xilinx a développé l'IDE Vitis pour programmer la PS.

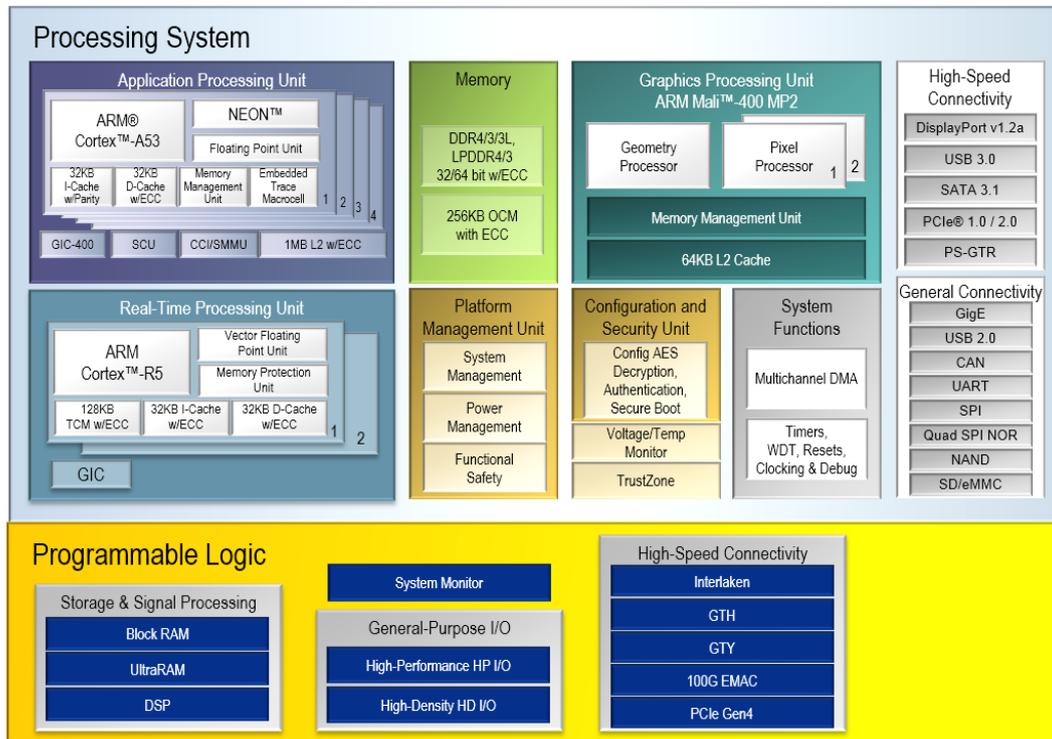


FIGURE 1.23 – Diagramme de blocs d'un FPGA Zynq UltraScale+ [52] où les CLB sont comprises dans la Programmable Logic

1.5 Mécanismes de dégradation et fiabilité des transistors

1.5.1 Introduction

Une définition officielle de la fiabilité est proposée par l'Union Technique de l'Électricité (UTE) : "La fiabilité est l'aptitude d'un dispositif à accomplir une fonction requise dans des conditions données pour une période de temps donnée". Dans un FPGA, qui est le dispositif à l'étude, la cessation d'une fonction requise peut être par exemple une diminution trop importante de la fréquence de fonctionnement d'un circuit combinatoire. Cette défaillance fonctionnelle au niveau circuit est en réalité causée par des mécanismes de défaillances au niveau du transistor tel que le BTI (Bias Temperature Instability), le HCI (Hot Carrier Injection) ou le TDDB (Time Dependant Dielectric Breakdown). À partir d'un nombre d'observations de défaillance suffisamment élevé, il est possible de réaliser une analyse statistique des défaillances pour prédire la fiabilité du FPGA. Cependant, les mécanismes de défaillances comme le BTI ou le HCI ne génèrent pas de défaillances cataclysmiques tel que la diminution trop importante et soudaine de la fréquence du FPGA. En effet, le BTI et le HCI ont pour effet une défaillance paramétrique qui dans ce cas est la diminution progressive de la fréquence du FPGA. La défaillance paramétrique de la fréquence FPGA peut être longue (parfois des dizaines d'années) avant de pouvoir observer une défaillance cataclysmique. Il est donc capital de connaître la physique de ces mécanismes de dégradation afin

de développer une modélisation la plus juste.

Dans cette section nous présentons les mécanismes de défaillance présent dans les transistors sous l'effet du vieillissement. Parce que nous étudions la fiabilité d'un FPGA Zynq UltraScale+ avec une technologie 16nm FinFET, nous nous penchons sur l'évolution de ces mécanismes dans les FinFET. Finalement, nous présentons les outils de statistique qui sont nécessaire pour modéliser la fiabilité d'un circuit numérique.

1.5.2 Bias Temperature Instability (BTI)

Le NBTI est connu pour être un problème de fiabilité depuis les années 1970' [53]. Avec plus de 5600 publications sur le sujet entre 2006-2016 [54], le NBTI a suscité un grand intérêt dans la communauté du semi-conducteur. Le BTI est un mécanisme de défaillance présent lorsque la grille d'un transistor est soumis à un champ électrique important et pour une température forte. Le mode de défaillance du BTI est l'augmentation de la tension de seuil du transistor. Dans une cellule logique, cela a pour conséquence l'augmentation du temps de propagation d'un signal. Dans le cas du PMOS on parle de NBTI (Negative BTI) et dans le cas du NMOS on parle de PBTI (Positive BTI). Dans le prochain paragraphe, nous décrivons le NBTI sur le PMOS. Le mécanisme du PBTI sur le NMOS est par analogie identique.

Principales caractéristiques du NBTI

Une revue complète du NBTI [55] présente les caractéristiques du NBTI acceptées par la communauté, les méthodes de mesures ainsi que les différentes théories des modèles physiques du NBTI. Nous allons ici regrouper les principales caractéristiques.

Loi de puissance

Le NBTI dépend du temps de vieillissement et suit une loi puissance avec un exposant au temps de 1/4. La figure 1.24 présente les résultats en échelle log-log obtenus dans [56] où la tension de seuil est mesurée au niveau du transistor pendant quelques minutes avec différentes méthodes de mesure et pour des conditions qui stimulent le NBTI ($T_{stress} = 125^{\circ}\text{C}$ et $V_{stress} = -3\text{V} \approx 2V_{nom}$). On constate que la dérive de la tension de seuil suit une loi puissance avec un exposant de 0.27. Un exposant au temps proche de 1/4 est aussi observé dans les tests de vieillissement de courtes durée (<10 heures) [56–58]. Pour des tests de vieillissement de longue durée (>1000 heures), un exposant

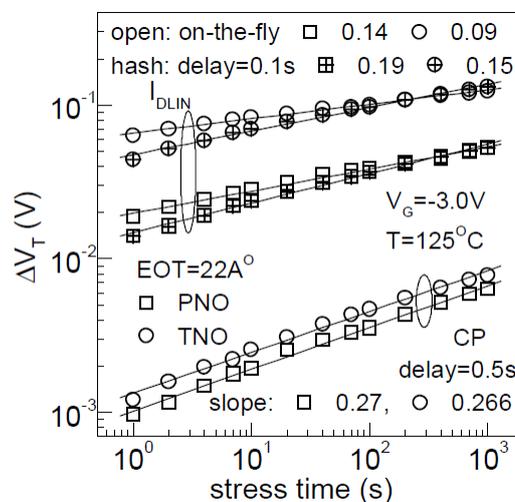


FIGURE 1.24 – Évolution de ΔV_{TH} en fonction du temps mesuré avec plusieurs méthodes de mesure et modélisé par une loi puissance avec un exposant au temps $n \approx 0.25$ [56]

de 1/6 a été observé dans [59]. Nous donnerons une explication possible de cette différence de comportement dans le modèle de Reaction-Diffusion.

Effet de la température et de la tension

L'amplitude du NBTI dépend de la tension de grille appliquée. La figure 1.25a montre l'amplitude des dégradations en fonction de la tension de stress appliquée sur deux types de transistor : le SiON (Si-OxyNitride) et le HKMG (High-k Metal Gate). Les points en noir correspondent à la mesure des dégradations de la tension de seuil. L'effet de la tension de stress est modélisé avec une loi puissance dont la forme simplifiée est $AF_V = (V_{stress} - V_{op})^\Gamma$. Le coefficient Γ est compris entre $4V^{-1}$ et $6V^{-1}$ suivant la méthode de mesure et le type de transistor considéré. Dans [60] on constate que le facteur d'accélération en tension peut être modélisé avec une loi exponentielle : $AF_V = e^{\gamma(V_{stress}-V_{op})}$. Dans la littérature on retrouve deux notations du coefficient du facteur d'accélération en tension : Γ et γ , nous utiliserons la notation γ .

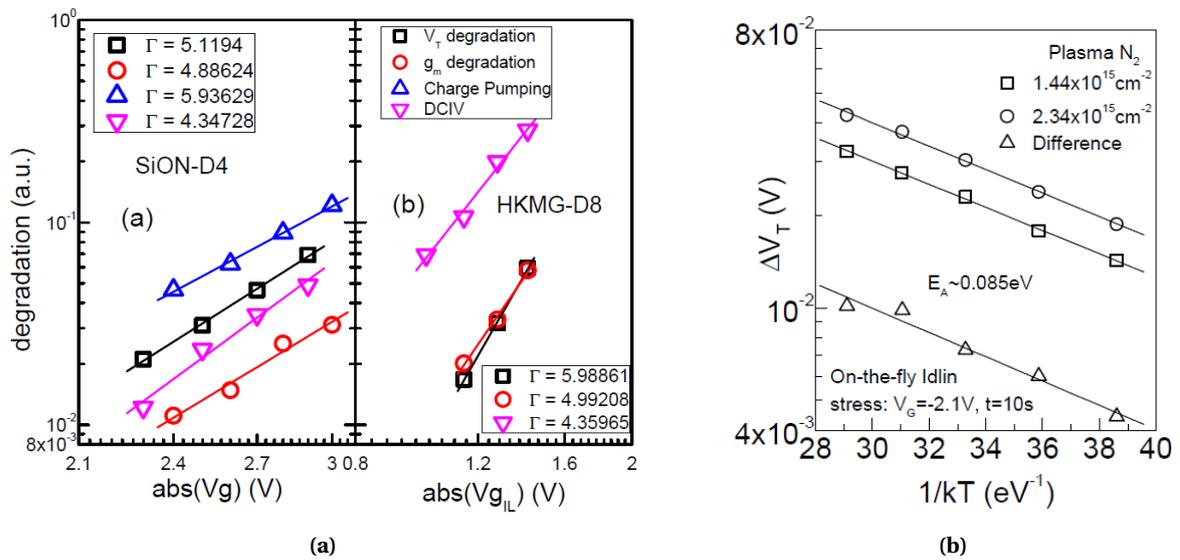


FIGURE 1.25 – (a) Amplitude des dégradations en fonction de la tension de stress pour deux technologies de transistor : le SiON et le HKMG. Chaque couleur correspond à une méthode de mesure. Extraction du paramètre (Γ) du facteur d'accélération en tension modélisé avec une loi puissance [61] (b) Extraction de l'énergie d'activation pour modéliser le facteur d'accélération en température avec une loi d'Arrhenius ($E_a \approx 0.085 eV$) sur un transistor avec l'oxide de grille nitruré au plasma et pour deux densités de N_2 [62]

L'amplitude du NBTI dépend de la température de jonction. La figure 1.25b présente la dégradation de la tension de seuil en fonction de l'inverse du produit de la constante de Boltzmann (k_B) et de la température de stress. Le facteur d'accélération en température est modélisé avec une loi d'Arrhenius : $AF_T = e^{\frac{E_a}{k_B T}}$. L'énergie d'activation extraite est $E_a \approx 0.085 eV$. Dans les études que nous venons de présenter, les mesures de dégradation sont réalisées au niveau du transistor. Dans cette thèse, nous mesurons les dégradations au niveau du circuit, c'est pourquoi nous utiliserons un modèle simplifié de l'impact du BTI sur le temps de propagation [2] avec une loi d'Arrhenius pour l'accélération en température et une loi exponentielle pour l'accélération en tension tel que :

$$\Delta\tau = A e^{\gamma V_{gs}} e^{\frac{E_a}{k_B T}} t^n \quad (1.2)$$

où A est l'amplitude de la dégradation, γ le coefficient du facteur d'accélération en tension et E_a l'énergie d'activation du facteur d'accélération en température.

Récupération du BTI

L'amplitude des dégradations dépend du temps lorsque le transistor est stressé et lorsqu'il n'est pas stressé [63]. La figure 1.26 montre l'évolution de la dégradation de la tension de seuil pour un

stress statique ($V_g = -2.5V$) et pour un stress dynamique avec une phase de stress ($V_g = -2.5V$) et de non stress ($V_g = +1.5V$). On constate que la dégradation pour un stress statique est continue. Pour un stress dynamique, en phase de stress la tension de seuil est dégradée alors qu'en phase de non stress la tension de seuil retourne vers sa tension de seuil initiale (récupère). Ce phénomène est appelé mécanisme de récupération du BTI. De manière générale pour mesurer le NBTI le transistor est stressé-mesuré-stressé en boucle. Cependant, lorsque le transistor est mesuré le phénomène de récupération est actif. C'est pourquoi les méthodes de mesures ont évolué visant à raccourcir le temps de mesure afin d'étudier séparément les dégradations mais aussi les récupérations liées au NBTI. Une revue des méthodes de tests est donnée dans [64].

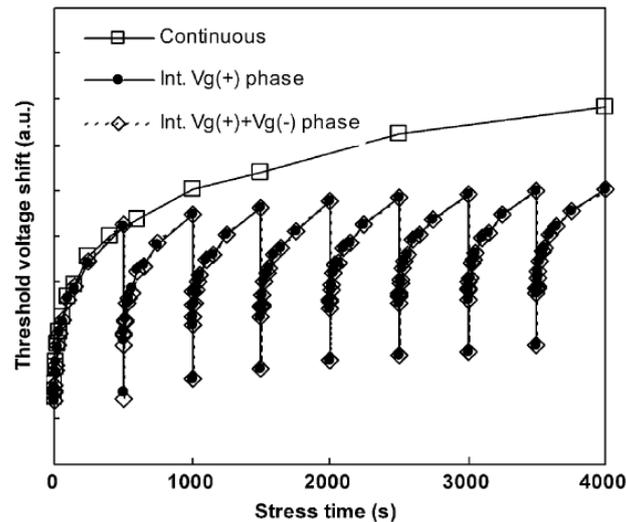


FIGURE 1.26 – Mesure de dégradation de la tension de seuil d'un transistor pour un stress statique (carré) et pour un stress dynamique (losange) mettant en évidence la récupération du BTI [58]

Effet de la fréquence de stress

L'amplitude des dégradations dépend de la fréquence du signal qui stress le transistor. Une étude réalisée dans [65, chapitre 21.3.3] présente l'effet de la fréquence de stress sur les dégradations dues au BTI. Dans cette étude les mesures sont réalisées sur des transistors isolés et les résultats sont présentés sur la figure 1.27. La figure présente la dégradation relative en fonction de la fréquence de stress. La dégradation relative correspond au rapport des dégradations mesurées lorsque le transistor subit un stress dynamique sur les dégradations mesurées lorsque le transistor subit un

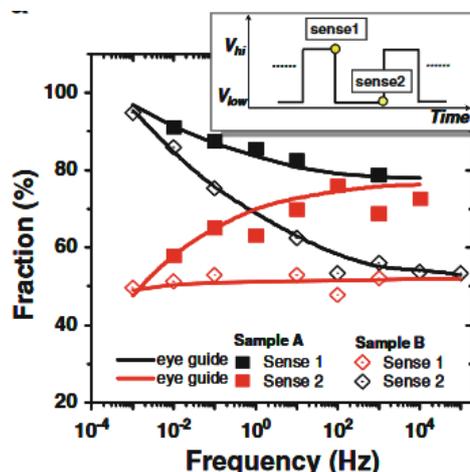


FIGURE 1.27 – Dégradations finales relatives en fonction de la fréquences de stress [65]

stress statique. Les points noirs correspondent aux mesures prises à la fin du temps de stress tandis que les points rouges correspondent aux mesures prises à la fin du temps de récupération. On constate une dépendance de l'amplitude des dégradations pour une fréquence de stress inférieure à 100 Hz. Au dessus de 100 Hz, l'amplitude des dégradations est indépendante à la fréquence de stress et les dégradations sont plus faibles. Dans [65, chapitre 21.3.3] et [66] les mécanismes de dégradation et de récupération du BTI sont modélisés comme étant un circuit de deux résistances, R_e et R_c , en parallèles et d'un condensateur C en série. Alors les dynamiques de temps sont : $\tau_e = R_e C$ et $\tau_c = R_c C$. Dans le BTI, τ_e correspond à la constante de temps de la dégradation et τ_c à la constante de temps de la récupération. La probabilité d'occupation d'un piège est indépendante de la fréquence de stress lorsque $f_{stress} \gg \max(1/\tau_e, 1/\tau_c)$.

Modèle de Reaction-Diffusion (RD)

Afin de modéliser au plus bas niveau de la physique les dégradations mesurés et présentés dans la précédente partie, des modèles physiques ont été développés. Ces modèles ont pour but d'interpréter les observations et de prédire les dégradations à partir des conditions d'utilisation et la technologie. À ce jour, les deux principaux modèles sont celui de Reaction-Diffusion (RD) [54] et le defect-centric [67][68]. Afin de comprendre le phénomène physique du NBTI, nous allons dans cette thèse introduire le modèle RD. Les explications qui vont suivre se basent sur l'étude de [69–71].

Afin d'améliorer les performances électriques d'un transistor, celui-ci est passivé avec des atomes d'hydrogène pour compléter les liaisons pendantes de silicium ou encore d'oxygène dans l'oxyde de grille (voir figure 1.28a). Lorsqu'un champ électrique est présent au niveau de la grille du PMOS, celui-ci peut provoquer la rupture d'une liaison $Si-H$ au niveau de l'interface du canal et de l'oxyde de grille (voir figure 1.28b). Cette rupture libère un atome d'hydrogène et génère un défaut qui est la liaison pendante $Si\cdot$ tel que : $Si-H \rightarrow Si\cdot + H$. L'atome d'hydrogène libéré se diffuse dans l'oxyde de grille jusqu'à voler un atome d'hydrogène à une liaison $O-H$ ou $Si-H$ (voir figure 1.28b). Ce phénomène provoque la dimérisation de l'hydrogène et génère un deuxième défaut tel que : $Si-O-H + H \rightarrow Si-O\cdot + H_2$. La diffusion de l'hydrogène suit une loi puissance dépendante au temps, à la température et au champ électrique. Le modèle de RD prédit une indépendance à la fréquence de stress [72]. Deux interprétations sont possibles avec le modèle RD [73, 74]. Soit le NBTI est dû à la diffusion d'hydrogène et, dans ce cas :

$$N_{IT}(t) = \sqrt{\frac{k_F N_0}{2k_R}} (D_H t)^{1/4} \quad (1.3)$$

où $N_{IT}(t)$ est le nombre de pièges d'interface à un instant t , N_0 est le nombre initial de liaisons $Si-H$ non cassées, D_H est la constante de diffusion de l'hydrogène, k_F est la constante de vitesse de dissociation et k_R est la constante de vitesse de récupération. Soit le NBTI est dû à la diffusion de dihydrogène, et dans ce cas :

$$N_{IT}(t) = \left(\frac{k_F N_0}{2k_R}\right)^{2/3} (D_{H_2} t)^{1/6} \quad (1.4)$$

Nous avons vu dans la partie 1.5.2 que les mesures de NBTI ont un exposant au temps de 1/4 lorsque le temps de test est court et un exposant de 1/6 lorsque le temps de test est plus long. Ce qui veut dire que sur du court terme les dégradations sont dues à la diffusion d'hydrogène et sur du long terme elles sont dues à de la diffusion de dihydrogène.

Lorsque le champ électrique de grille devient suffisamment faible, la mécanique inverse se produit (voir figure 1.28c). Dans l'oxyde de grille : $Si-O\cdot + H_2 \rightarrow Si-O-H + H$ et à l'interface entre l'oxyde de grille et l'interface $Si\cdot + H \rightarrow Si-H$. On appelle ce phénomène : récupération (recovery). Le modèle de RD ne prend en compte que les défauts d'interface. Un modèle complémentaire,

Transient Trap Occupancy Model (TTOM) peut être associé pour prendre en compte les pièges déjà existants dans l'oxyde et dans le substrat. Le TTOM décrit les dégradations du BTI pour un stress de courte durée ou pour une tension et une température de stress importantes.

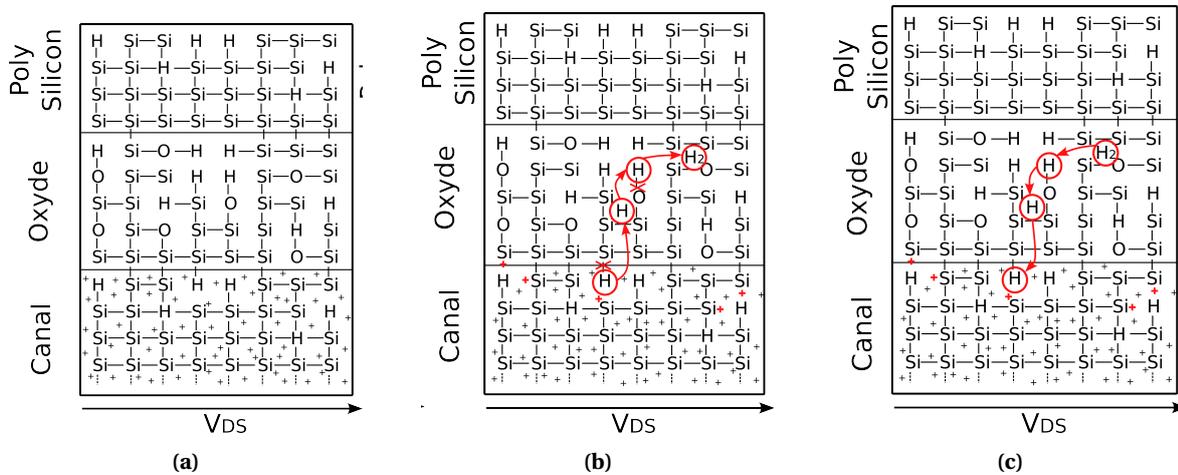


FIGURE 1.28 – Représentation du NBTI dans un MOSFET (a) pas de NBTI (b) dégradation (c) Récupération

Le BTI dans les FinFETs

Alors que le PBTI était négligeable dans les transistors SiON, il est devenu tout aussi important que le NBTI dans les transistors HKMG [75, 76]. Mais qu'en est-il du BTI dans les FinFETs?

Des études de fiabilités réalisées par Intel pour la technologie FinFET 22nm [21, 77] (voir figure 1.29a) ou par TSMC pour la technologie FinFET 16nm [28] montrent que le NBTI est identique sur les PMOS alors que le PBTI est devenu négligeable par rapport à la génération de NMOS précédente. Dans [78–81], une augmentation du NBTI et une diminution du PBTI sont observées. Le BTI est un mécanisme de défaillance dépendant de la surface de l'interface entre le substrat et l'oxyde [81]. Dans un MOSFET, cette surface correspond au produit de la longueur par la largeur du canal. Dans un FinFET, la hauteur de l'aileron devient un paramètre important, comme on peut voir sur la figure 1.30.

L'orientation du cristal de l'aileron du côté latéral d'un FinFET joue un rôle important pour le BTI. Il a été observé qu'une orientation (100) possède moins de NBTI qu'une orientation (110),

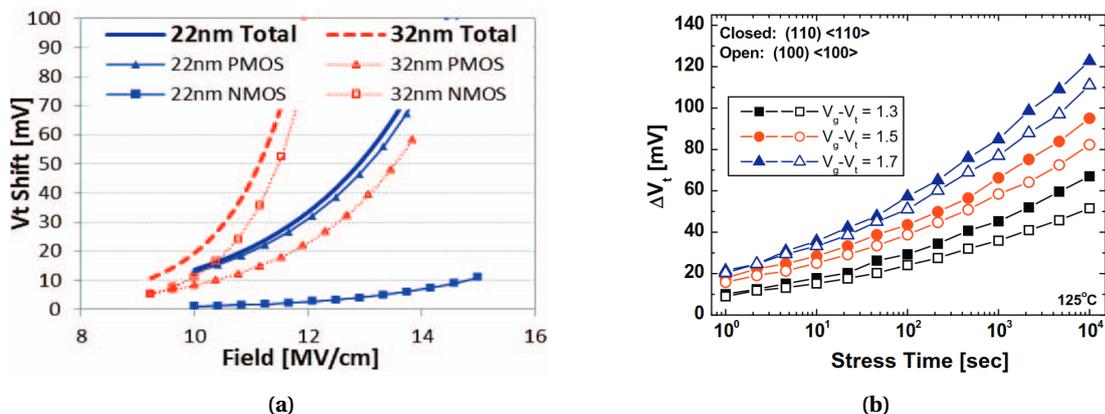


FIGURE 1.29 – (a) Comparaison du NBTI et PBTI dans un FinFET 22nm et un MOSFET 32nm [77] (b) Comparaison du NBTI dans un FinFET pour deux orientations (100) et (110) [78]

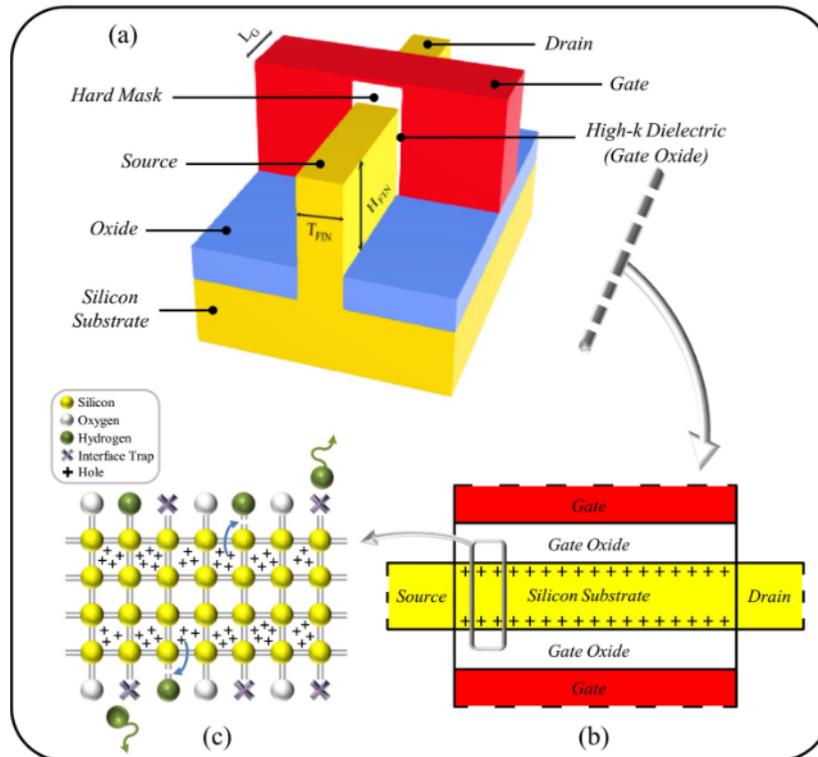


FIGURE 1.30 – Schéma d'un FinFET avec une vue tridimensionnelle, une vue en coupe et la représentation du BTI dans un FinFET [81]

voir figure 1.29b, car le nombre de liaisons pendantes est plus important dans une orientation (110) que dans une orientation (100) [78, 80]. L'étude présentée dans [78] montre que la largeur de l'aileron n'a pas d'impact sur le BTI et que le phénomène de récupération est identique entre une technologie planaire et FinFET. L'étude réalisée dans [81] montre que le modèle RD prédit une dégradation plus importante dans les FinFETs que dans les MOSFETs. Cependant, au niveau circuit les simulations se basant sur des modèles de dégradations montrent une dégradation du temps de propagation plus importante pour les MOSFETs. Toujours en se basant sur des modèles de dégradation, une simulation au niveau du circuit montre une dégradation similaire entre un circuit avec une technologie 20nm planaire et un circuit avec une technologie 14nm FinFET [79]. Ce dernier résultat est cohérent avec ce qu'on observe sur la figure 1.29a et dans [28] : par rapport à une technologie planaire, dans le FinFET le NBTI a augmenté et le PBTI a diminué. Donc, dans un circuit fait de PMOS et de NMOS il est cohérent d'observer des dégradations équivalentes en un circuit avec une technologie 20nm planaire et un circuit avec une technologie 14nm FinFET.

1.5.3 Le HCI (Hot Carriers Injection)

Le HCI est un mécanisme de défaillance connu depuis les années 1980 et mis en évidence par la mesure dans [82, 83]. Son mode de défaillance, identique au BTI, consiste en l'augmentation progressive de la tension de seuil et, par conséquent, l'augmentation du temps de propagation d'un signal dans une cellule logique. Les conditions favorables pour stimuler le HCI sont une tension de grille importante et une température de jonction basse. Le HCI est présent seulement dans les transistors qui sont en régime saturés. Nous verrons par la suite que ces conditions ont évolué avec la diminution du canal, notamment. La complexité de l'interprétation physique du HCI a suscité un grand nombre de travaux de modélisation [84]. Dans cette partie, comme pour le BTI, nous présentons les caractéristiques principales du HCI observés lors de mesure et, nous introduisons le modèle de l'électron chanceux qui est un des modèles physique du HCI. Finalement, nous étudions l'évolution du HCI dans les FinFETs.

Les caractéristiques principales du HCI

Le HCI est un mécanisme de défaillance dépendant du temps de vieillissement et une bonne approximation de la modélisation des défaillances se fait avec une loi de puissance. Dans [85, table 1], un résumé de 16 études portant sur la mesure du HCI dans les transistors révèle un exposant au temps qui est compris entre 0.3 et 0.7.

L'amplitude des dégradation dépend de la tension de grille et de la tension du drain-source. La figure 1.31a présente les dégradation d'un transistor avec un long canal en fonction de la tension de grille et pour une tension drain-source de 15V. On constate que l'amplitude des dégradations est plus importante lorsque $V_G \approx V_D/2$ [82]. La figure 1.31b présente les dégradations d'un transistor avec un canal court (40nm). Sur le graphique de droite, on constate que l'amplitude des dégradations de la tension de seuil est plus importante lorsque $V_G = 2.4V$ (en rouge) que lorsque $V_G = 1V$ (en noir) en sachant que $V_D = 2.4V$. Donc cette étude [86] montre que pour un transistor avec un canal court, les dégradations dues au HCI sont plus importantes lorsque $V_G = V_D$. La figure 1.31c présente la dégradation de la tension de seuil mesuré sur plusieurs transistors avec des stress de température et de la tension du drain qui sont différents et une tension de grille constante (3V). Pour une même température de 100°C (points noirs) on constate que l'amplitude des dégra-

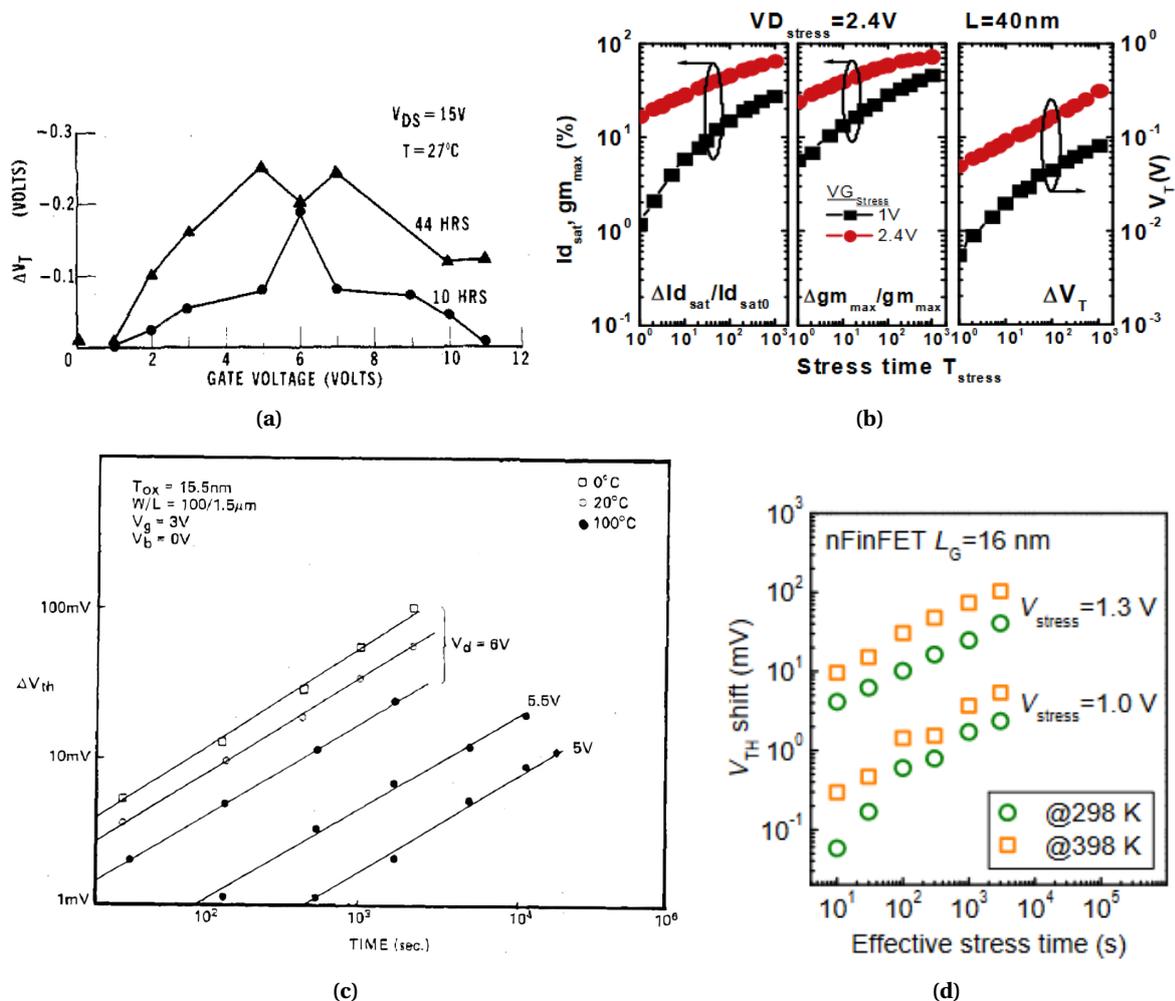


FIGURE 1.31 – (a) Dégradation de la tension de seuil ΔV_{TH} en fonction de la tension de grille appliquée pour un transistor à canal long sachant que $V_{DS} = 15V$ [82] (b) Dégradation, dans un transistor à canal court, du courant de drain (gauche), de la transconductance (milieu) et de la tension de seuil (droite) en fonction du temps pour deux tension de grilles 1V (noir) et 2.4V (rouge) sachant que $V_{DS} = 2.4V$ [86] (c) Dégradation de la tension de seuil ΔV_{TH} en fonction du temps pour trois températures $0^\circ C$, $20^\circ C$ et $100^\circ C$, pour trois tension de drain et dans un transistor à canal long [87] (d) Dégradation de la tension de seuil ΔV_{TH} en fonction du temps pour deux températures $25^\circ C$, $125^\circ C$ et dans un transistor à canal court [88]

dations est plus importante pour une tension de drain de 6V que pour une tension de drain de 5.5V et encore plus que pour une tension de drain de 5V.

L'amplitude des dégradation dépend de la température de jonction du transistor. Pendant longtemps la communauté scientifique a considéré que le HCI était majoritairement présent à basse température [87]. Sur la figure 1.31c, on constate que, pour une tension de grille fixe (3V) et une tension de drain fixe (6V), l'amplitude des dégradations est plus importante à 0°C qu'à 20°C et plus encore à 100°C. Cependant, des études plus récentes montrent que, dans les transistors à canal court, le HCI est aussi stimulé par les températures hautes [88, 89]. C'est aussi ce qu'on observe sur la figure 1.31d où les dégradations sont plus importantes à 125°C qu'à 25°C. Une explication physique de ce changement de comportement du HCI en fonction de la température est donné dans [89, chapitre II.C]. À haute température, c'est l'exposant au temps qui permet de dissocier les dégradations dues au BTI et au HCI [90, part II]. Si la dégradation de l'interface par des porteurs chauds est dominant alors l'exposant au temps observé est entre 0.5 et 1, tandis que, pour du BTI c'est à dire lorsque le piégeage des porteurs du substrat dans les défauts de l'oxyde est le principal mécanisme de dégradation l'exposant au temps est entre 0.1 et 0.2 [90, part II].

Finalement, un dernier paramètre qui joue un rôle important dans le HCI réside dans le type de stress du signal interne : statique (DC) ou dynamique (AC). Pour les transistors à canal long il a été constaté qu'un stress AC provoque plus de dégradation qu'un stress DC [91], tandis que l'inverse est observé pour les transistors à canal court [88]. Lorsque le transistor est soumis à un stress AC, alors la dégradation dépend linéairement du nombre de commutation [92, 93] tel que :

$$\Delta A \propto (f \times t)^n \quad (1.5)$$

où ΔA est la dégradation, f la fréquence de stress et n l'exposant au temps.

Le modèle de l'électron chanceux

En 1985, Hu et al. ont développé le modèle de l'électron chanceux [94] qui est le premier modèle physique pour le HCI. Lorsque le courant passe dans le canal, certaines charges (électrons dans le NMOS et trous dans le PMOS) acquièrent une énergie suffisante et sont capables de rompre une liaison Si – H, générant ainsi un piège de surface. La densité de porteurs chauds est plus importante au niveau du drain [95] car c'est à ce niveau que le champs électrique est plus important et que la charge a une probabilité plus importante d'acquérir beaucoup d'énergie. Une illustration du HCI selon le modèle de l'électron chanceux est présente sur la figure 1.32. La densité de piège d'interface s'exprime :

$$N_{IT} = C \left(t \frac{I_d}{W} e^{\frac{-\varphi_i}{q\lambda E_m}} \right)^n \quad (1.6)$$

où $\frac{I_d}{W}$ correspond à la densité de courant dans le canal, φ_i est l'énergie minimum qu'une charge doit avoir pour créer un piège, λ est le libre parcours moyen d'un porteur "chaud" et l'ensemble $\frac{I_d}{W} e^{\frac{-\varphi_i}{q\lambda E_m}}$ correspond finalement à la densité de charge qui a une énergie supérieure à φ_i . La densité de piège exprimé dans l'équation (1.6) suit une loi puissance avec un exposant au temps compris entre 0.5 – 1.

Le HCI dans les FinFET

Des études montrent dans les transistors avec un canal court l'augmentation de la température augmente les dégradations sous l'effet du HCI [88, 89]. Alors que le FinFET possède des performances électrostatiques meilleures que les transistors planaires, la géométrie du FinFET veut qu'il ait une moins bonne conductivité thermique qui provoque de l'auto-échauffement et donc une augmentation de la température de jonction [85]. La figure 1.33a montre l'évolution de la résistance thermique en fonction de différentes technologies de transistors. On constate que la

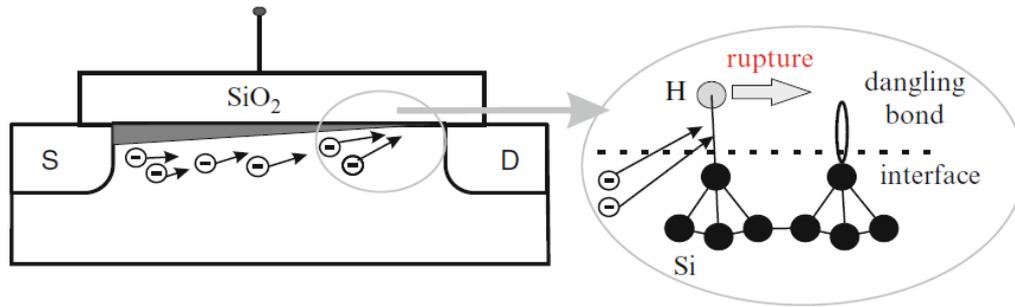


FIGURE 1.32 – Représentation schématique du HCI. Les porteurs chauds qui entrent en collision avec l’interface diélectrique/Si dans un MOSFET peuvent déposer leur énergie, produisant ainsi des dommages. Ce dommage est associé à la dissociation des liaisons Si – H. Si une telle liaison est rompue, une liaison pendante subsiste. Cette liaison pendante, électriquement active, peut piéger des électrons/trous et aggraver la mobilité des porteurs [90][Part.I-Chap.IV]

résistance thermique augmente entre les technologies planaire et FinFET. Cette même résistance thermique augmente considérablement lorsque plusieurs ailettes (multi-fin) sont assemblées ensemble.

La figure 1.33b présente les dégradations de la tension de seuil de transistors FinFET en fonction du nombre d’ailettes qui le constitue. On constate que les dégradations sont plus importantes lorsque le transistor est constitué de 50 ailettes que lorsqu’il est constitué de 10 ailettes ou plus encore de 6 ailettes. Ce résultat est cohérent avec la précédente observation de l’augmentation de la résistance thermique du transistor pour un multi-fin. Puisque l’augmentation de la résistance thermique provoque l’augmentation de la température de jonction et génère des dégradations plus importantes dans les transistors à canal court dues au HCI.

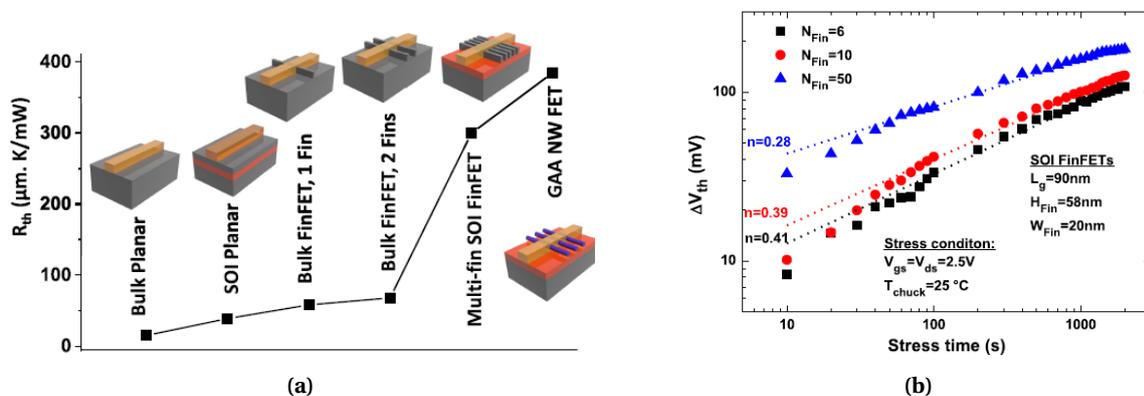


FIGURE 1.33 – (a) Évolution de la résistance thermique en fonction de la technologie de transistor [85] (b) Dégradation de la tension de seuil V_T en fonction du temps pour différents nombres d’ailettes utilisées dans le transistor [96]

Pour confirmer que la dynamique des dégradations observées sur la figure 1.33b sont bien dues à l’effet du nombre d’ailettes et donc de l’augmentation de la température, dans [96] les dégradations sont mesurées sur un FinFET avec 6 ailettes et pour plusieurs températures : 25°C, 55°C et 85°C. En comparant la figure 1.33b et [96, figure 3] on constate une dynamique des dégradations presque identique entre les dégradations avec 10 ailettes et 55°C ainsi qu’entre 50 ailettes et 85°C. Donc le nombre d’ailettes dans un FinFET augmente la température interne qui accélère le vieillissement.

Finalement, sur la figure 1.33b on constate que l’exposant au temps des dégradations diminue lorsque le nombre d’ailettes augmente ($n = 0.28$ pour 50 ailettes et $n = 0.41$ pour 6 ailettes). Cette figure montre que la dynamique des dégradations est dépendante à la température. D’après

[96], cela indique qu'une température élevée peut favoriser l'injection de porteurs dans la couche d'oxyde et les piéger dans les défauts de l'oxyde.

L'amplitude de dégradation dépend de la largeur de l'ailette du FinFET [97, 98]. Comme le montre la figure 1.34a, plus l'ailette est large et moins nous observons des dégradations. D'après [97], la cause est l'augmentation du SHE avec la diminution de la largeur de l'ailette, donc une température plus importante, qui stimule davantage le HCI.

La figure 1.34b montre l'évolution des dégradations dans des FinFET pour plusieurs signaux de stress statique et dynamiques. On constate alors que l'amplitude des dégradations sont plus importantes pour un stress statique. On observe aussi que l'amplitude des dégradations décroît lorsque la fréquence de stress augmente [99].

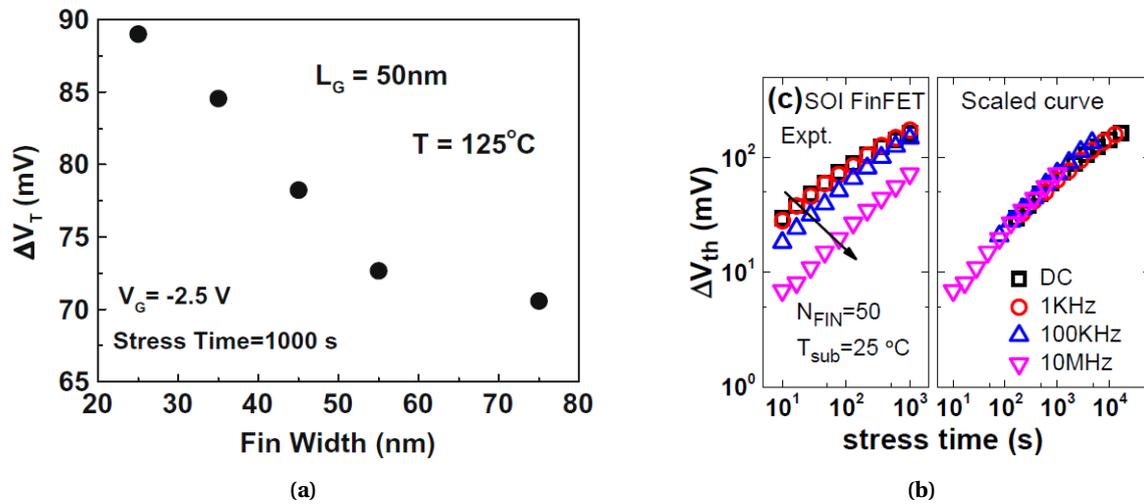


FIGURE 1.34 – (a) Dégradation de la tension de seuil V_T en fonction de la largeur de l'ailette d'un FinFET [98] (b) Dégradation de la tension de seuil V_T en fonction du temps pour différentes fréquence de stress [99]

1.5.4 Time Dependant Dielectric Breakdown (TDDB)

Dans les transistors, le TDDB est un mécanisme de dégradation redouté puisqu'il peut générer une défaillance catalectique. Lorsqu'il y a un champ électrique appliqué sur la grille d'un transistor, l'oxyde se dégrade avec la création de pièges (voir figure 1.35a). Le TDDB, aussi appelé claquage d'oxyde, se divise en trois étapes (voir figure 1.35b). La première étape correspond au temps pour qu'un piège prenne place dans la grille. Dans la seconde étape se produit une accumulation de pièges dans l'oxyde qui provoquant ainsi un léger courant de fuite. Les pièges peuvent être la conséquence du HCI ou du BTI. La troisième étape commence lorsque la densité de pièges dans l'oxyde est telle qu'un chemin conducteur se forme. Alors un courant de fuite conséquent passe dans le chemin conducteur formé de pièges, ce qui va d'abord dégrader le transistor (soft breakdown), puis le détruire (hard breakdown). Il existe trois modèles qui expliquent la génération des pièges dans le cadre du TDDB :

- Le modèle anode hole injection (AHI) [100],
- Le modèle thermo-chimique [101],
- Le modèle anode hydrogen release (AHR) [102].

Selon le modèle AHI, le temps du claquage de l'oxyde dépend de $1/E$, tel que :

$$t_{BD}(t) = \tau_0 e^{\frac{G}{E_{ox}}} \quad (1.7)$$

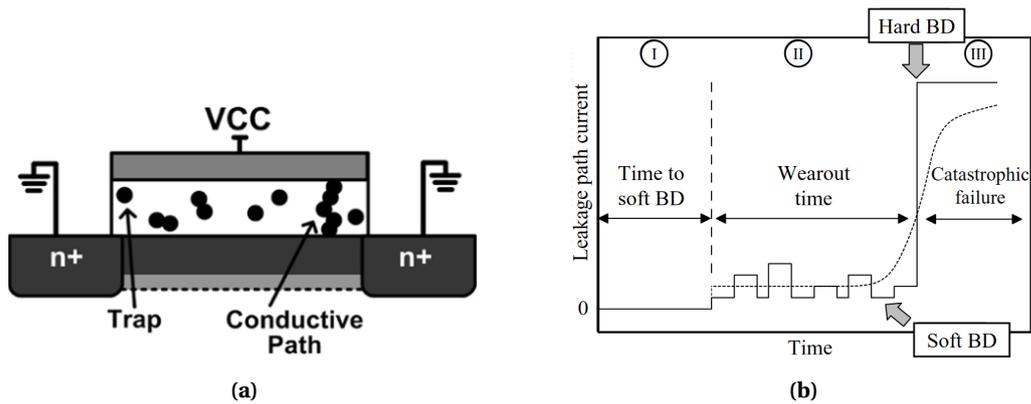


FIGURE 1.35 – (a) Schéma de principe du TDDB [103] (b) Les différentes phases de dégradation du TDDB [104]

où E_{OX} est le champ électrique qui traverse l'oxyde, τ_0 et G sont des constantes. Selon le modèle thermo-chimique le temps du claquage de l'oxyde dépend de E tel que :

$$t_{BD}(t) = t_0 e^{-\gamma E_{OX}} \quad (1.8)$$

où γ est une constante. L'étude de la fiabilité d'un tel mécanisme se fait par une étude statistique du nombre de défaillances au cours du temps (voir section 1.5.5). La fonction de répartition des défaillances peut suivre la loi de Weibull ou encore log-normale.

Des études montrent que le TDDB est aussi présent dans les FinFETs [105–108]. Dans [105] le modèle suivant est utilisé pour prédire le temps de défaillance (T_{BD}) du au TDDB :

$$T_{BD} = \tau_0 V^{-n} e^{\frac{E_A}{kT}} e^{\frac{1}{\beta} \ln(-\ln(1-F))} \quad (1.9)$$

où τ_0 est une constante, n le facteur d'accélération en tension (VAF), E_A l'énergie d'activation et β le coefficient d'ajustement de la loi de Weibull. Dans [105–107] l'énergie d'activation E_A est compris entre $0.6eV$ et $1.4eV$. Dans [105] et [106] le coefficient β de la loi de Weibull est respectivement égal à 1.6 et 1.5. Dans [106] le VAF est de ≈ 50 et dans [108] il est de ≈ 45 . Comme nous l'avons déjà évoqué pour le HCI, l'auto-échauffement est plus important dans le FinFET et, par conséquent, le TDDB est aussi plus important [107]. Dans [108] une comparaison entre un FinFET 14nm et 10nm montre que le TDDB a plus d'impact sur la technologie 10nm, par ailleurs, le coefficient β reste identique ce qui garantit que le mécanisme de défaillance observé est le même.

1.5.5 La fiabilité dans l'industrie

Avant de présenter une utilisation de la fiabilité dans l'industrie, nous complétons la définition de la fiabilité "la fiabilité est l'aptitude d'un dispositif à accomplir une fonction requise dans des conditions données pour une période de temps donnée" par celle du mot défaillance : "Une défaillance est la cessation de l'aptitude à assurer une fonction requise". Parce que les systèmes sont de plus en plus complexes et qu'ils sont composés de plusieurs couches de sous-systèmes qui sont elles mêmes composées de plusieurs circuits numériques, la fiabilité n'est pas étudiée au niveau du composant (transistor) mais au niveau système. On distingue alors plusieurs modes de défaillance génériques :

- Plus de fonction : La fonction cesse de se réaliser ;
- Pas de fonction : La fonction ne se réalise pas lorsqu'elle est sollicitée ;
- Fonction dégradée : La fonction ne se réalise pas dans un état optimal ;
- Fonction intempestive : La fonction se réalise lorsqu'elle n'est pas sollicitée ;

Pour le mode de défaillance "fonction dégradée" il est possible de fixer un seuil paramétrique, souvent lié à la performance du système pour accomplir la fonction, à partir duquel on considère le mode de défaillance comme "plus de fonction".

Dans la définition de la fiabilité, les "conditions données" correspondent à l'environnement physique dans lequel le dispositif évolue. Par exemple, pour qualifier les systèmes d'un satellite, il est primordial de réaliser des études de fiabilité dans un environnement de radiations fortes. "Une période de temps donnée", peut s'exprimer en :

- Calendaire : Nombre d'années depuis la mise en service d'un dispositif (en générale de 30 ans pour un avion) ;
- Heure de fonctionnement : Nombre d'heures comptabilisées seulement lorsque le dispositif est activé (environ 20000 heures pour une ampoule à LED) ;
- Cycle : Nombre de fois que le dispositif réalise un cycle complet (nombre de décollages/atterrissages d'un avion) ;

Mais alors à quoi sert l'étude de la fiabilité? Dans les secteurs sensibles tels que la défense, l'aéronautique ou encore le spatial, la fiabilité est une donnée d'entrée pour l'analyse de sûreté de fonctionnement ou de gestion de la qualité. Dans les années 1950 la méthode d'Analyse des Modes de Défaillance, de leurs Effets et de leur Criticité (AMDEC), en anglais Failure Modes, Effects and Criticality Analysis (FMEAC), est développée. L'objectif de cette méthode est d'identifier la criticité (matériel et humain) qu'une défaillance peut avoir. La criticité est calculée par le produit de trois indices que sont l'indice de fréquence, l'indice de gravité et l'indice de détection. L'indice de fréquence est défini à partir de la fiabilité, l'indice de gravité à partir des dommages provoqués et l'indice de détection à partir de l'usage. La figure 1.36 est un exemple de tableau avec les critères correspondant à un niveau de l'indice. Si la criticité dépasse un certain seuil fixé par un organisme (l'autorité, l'entreprise ou le client), l'entreprise prend des mesures pour réduire la criticité en instaurant par exemple de la maintenance qui diminuera l'indice de détection ou en améliorant la fiabilité du composant ce qui diminuera l'indice de fréquence.

Note F	Fréquence ou probabilité d'apparition	Note G	Gravité	Note D	Probabilité de non-détection
10	Permanent	10	Mort d'homme ou catastrophe environnementale	10	Aucune probabilité de détection
5	Fréquent	5	Conséquences financières et/ou matérielles	5	Un système de détection est en place mais n'est pas infallible
1	Invraisemblable	1	Pas grave	1	Le système de détection est infallible

FIGURE 1.36 – Exemple d'une grille de notation qui peut être utilisé dans l'AMDEC pour calculer la criticité

À l'image de l'AMDEC, on constate que la fiabilité est une donnée d'entrée importante dans l'industrie. Parce que les études de fiabilités sont coûteuses en temps et en ressources financières, il existe des référentiels de fiabilité accessibles par les entreprises pour prédire la fiabilité d'un système. Ces référentiels sont alimentés par le retour d'expérience et par des études de vieillissements accélérés (ce que nous faisons dans cette thèse). Deux référentiels américains sont utilisés pour la prédiction de la fiabilité de composants électronique : MIL-HDBK-217F et UTE C 80-810. En 2004 la Direction Générale de l'Armement (DGA) a encouragé un consortium d'entreprises françaises à développer un référentiel de fiabilité plus précis : FIDES. Ce référentiel est mis à jour régulièrement afin de couvrir les améliorations technologiques. Une définition précise et complète du profil de vie, comme les conditions d'exploitation et les contraintes environnementales des systèmes, l'a rendu unique.

1.5.6 De la mesure à la fiabilité

Les circuits numériques sont des systèmes non réparables et on distingue deux types de défaillance : la défaillance catalectique et la défaillance paramétrique. Pour faire une étude de fiabilité à partir des mesures de défaillances catalectiques, en principe, il suffit de compter en fonction du temps le nombre de circuits numériques défaillants observés. Cependant, l'observation de défaillance catalectique dans un circuit numérique est souvent longue, parfois des années. Pour ne pas faire un test de vieillissement de plusieurs années sans avoir la certitude d'observer une défaillance catalectique, une solution est de mesurer les défaillances paramétriques dans un circuit numérique, ce que nous faisons sur des FPGAs. Le BTI et le HCI sont des mécanismes de défaillance qui génèrent des défaillances paramétriques au niveau du transistor qui se manifestent dans un circuit numérique par la diminution de la fréquence de fonctionnement par exemple. Nous avons montré que le BTI et le HCI génèrent des dégradations continues qui peuvent être modélisé par une loi de puissance rendant l'extrapolation des dégradations à des temps de fonctionnement plus importants. Ainsi, à partir de la mesure de dégradation et en fixant un seuil de fonctionnement minimal, il est possible de noter le temps ou la dégradation dépasse le seuil et devient une défaillance. Si toutefois dans le temps de test la mesure de dégradation ne dépasse pas le seuil, alors il est possible d'extrapoler cette dégradation.

Nous noterons l'instant où une défaillance est observée le TTF (Time To Fail). En vieillissant et en mesurant les défaillances d'un grand nombre de dispositifs, il est alors possible d'utiliser les outils de la statistique descriptive pour prédire la fiabilité du dispositif en question. Pour présenter ces outils, nous considérons un échantillon de 5000 dispositifs dont nous simulons la distribution des TTF suivant une loi normale de moyenne 5000 heures et d'écart type 500 heures. La figure 1.37 présente quelques graphiques de statistique (bleu) tracés à partir des résultats de simulation et les courbes de probabilités (orange) tracés pour une distribution normale.

La figure 1.37a représente la fréquence cumulée et indique à chaque instant la probabilité qu'un dispositif soit en panne. Ce graphique correspond à la fonction de répartition (Cumulative Distribution Function CDF) des défaillances qu'on note aussi $F(t)$ (F pour failure). La durée de vie d'un dispositif peut être représentée par une variable aléatoire T , dont la fonction de défaillance est définie par :

$$F(t) = P[T \leq t] \quad (1.10)$$

où $P[T \leq t]$ est la probabilité qu'un dispositif soit en panne à une date T antérieur à t . Un diagramme complémentaire peut être facilement obtenu à partir de celui de la figure 1.37a, en évaluant à partir de $F(t)$ la fonction $R(t) = 1 - F(t)$. La fonction $R(t)$ est appelée fonction de survie ou fonction de fiabilité et permet d'évaluer quel est, à un instant t , la probabilité qu'un dispositif soit encore fonctionnel.

L'histogramme de la figure 1.37b représente le quotient du nombre de dispositifs en pannes par le produit du nombre total de dispositifs et d'un intervalle de temps (largeur d'un bâton). Cet histogramme représente la densité de probabilité (Probability Density Function PDF) de défaillance qui s'exprime en continu :

$$f(t) = \frac{dF(t)}{dt} \quad (1.11)$$

La densité de défaillance représente la probabilité qu'un dispositif tombe en panne dans l'intervalle de temps $]t, t + dt]$ divisée par la durée de l'intervalle.

Finalement, le taux de défaillance (voir figure 1.37c) est la proportion de systèmes en panne à un instant $[t + dt]$ sachant qu'ils étaient encore fonctionnels à l'instant t , divisé par dt . En continu le taux de défaillance s'écrit :

$$\lambda(t) = \frac{f(t)}{R(t)} \quad (1.12)$$

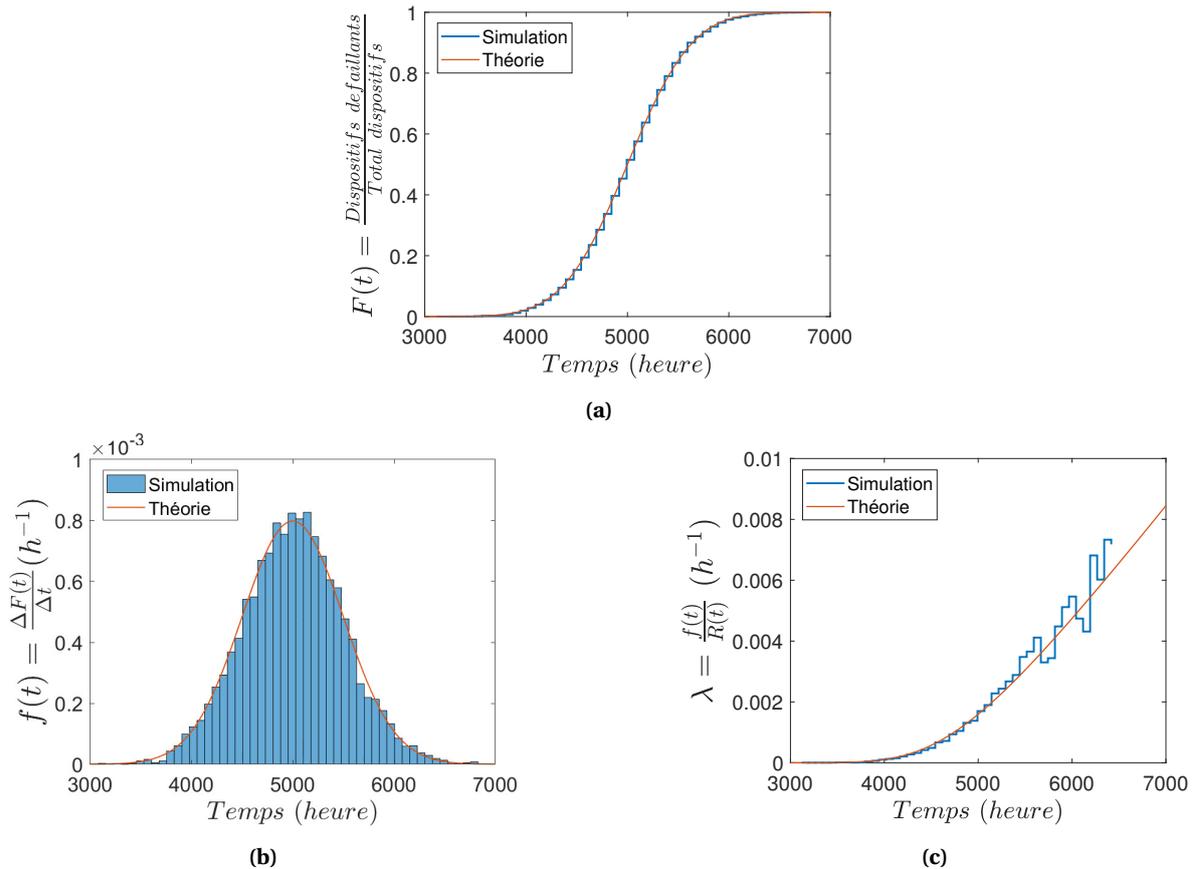


FIGURE 1.37 – (a) Fonction de défaillance $F(t)$ pour un échantillon de 5000 de dispositifs (b) Densité de probabilité $f(t)$ de la fonction de défaillance $F(t)$ (c) Taux de défaillance $\lambda(t)$

où $\lambda(t)$ est une probabilité de panne par heure. Sur la figure 1.37c on constate un effet de bord pour les derniers instants de mesure. En effet, le taux de défaillance est le quotient de $f(t)$ par $R(t)$ qui tendent vers 0 lorsque le temps augmente ce qui va directement impacter le calcul du taux de défaillance dans les derniers instants de mesure.

En électronique, parce que le taux de défaillance des transistors est relativement faible pour une durées de fonctionnement raisonnable (inférieur à 30 ans), il est possible d'exprimer son taux de défaillance en FIT (Failure In Time) avec $1\text{FIT} = 10^{-9}h^{-1}$.

Le moment d'ordre 1 d'une loi de défaillance aussi utilisé pour caractériser la fiabilité d'un système est le MTTF (Mean Time To Fail). Ce paramètre correspond au temps moyen pour qu'un dispositif soit défaillant, c'est-à-dire son espérance de vie telle que :

$$\text{MTTF} = E[t] = \int_0^{+\infty} t \cdot f(t) \cdot dt \quad (1.13)$$

1.5.7 Loïs de vieillissement

La courbe en baignoire

La figure 1.38 représente la fonction du taux de défaillance $\lambda(t)$ qui décrit le vieillissement de la plupart des composants en électronique. Cette courbe est aussi appelée courbe en baignoire pour la forme de la fonction.

Cette courbe se décompose en trois régions :

- Jeunesse : Dans cette région, les défauts apparaissent en début de vie du composant et sont le plus souvent liés à un défaut de fabrication. Il est possible d'éliminer les défauts de jeu-

nesse en ajoutant une étape post-fabrication de déverminage. Les défauts de jeunesse révèle un manque de qualité dans le processus de fabrication. L'étape de déverminage est un pansement, cependant, il serait plus approprié d'améliorer le processus de fabrication pour réduire les défauts de jeunesse.

- Maturité : Contrairement au taux de défaillance décroissant de la région de jeunesse, la région suivante se caractérise par un taux de défaillance à peu près constant. Les défaillances qui se produisent au début de cette région (partie gauche de la courbe) sont souvent liés à des phénomènes extérieurs tels que les décharges électrostatiques (ESD).
- Vieillesse : Dans cette dernière région le taux de défaillance croissant est dû à une accumulation de mécanismes de défaillance paramétriques pendant la vie utile. Par exemple, le BTI ou le HCI dégradent progressivement le fonctionnement d'un transistor. Lorsque le temps de fonctionnement d'un circuit numérique augmente, la probabilité pour que la dégradation d'un transistor dépasse le seuil de défaillance augmente, ce qui explique que le taux de défaillance est croissant.

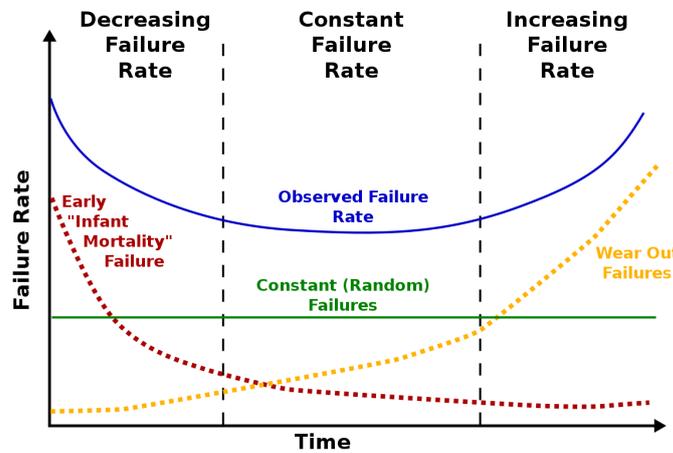


FIGURE 1.38 – Évolution du taux de défaillance pendant la durée de vie d'un dispositif représenté par la fameuse courbe en baignoire

La figure 1.38 montre que le taux de défaillance global correspond à la somme des trois taux de défaillance associés à la région de jeunesse, de maturité et de vieillesse. Par exemple, le taux de défaillance de la vieillesse (jaune) est croissant comme celui présenté sur la figure 1.37c. Cela montre que les défaillances présentés sur la figure 1.37c sont des défauts de vieillesse. Il existe plusieurs lois statistiques qui décrivent la distribution des défaillances dans le temps d'un composant électronique. Nous allons décrire les principales distributions qui sont : exponentielle, Weibull et log-normale.

Distribution exponentielle

Parce que son taux de défaillance est constant, la distribution exponentielle est la plus utilisée pour décrire la période de maturité d'un dispositif. On constate sur la courbe en baignoire 1.38 que pendant la maturité du dispositif (en vert) le taux de défaillance est constant. De plus, la loi de distribution exponentielle est la plus simple d'un point de vue mathématique. La fonction de densité ($f(t)$), la fonction de répartition ($F(t)$) et le taux de défaillance ($\lambda(t)$) prennent la forme de :

$$f(t) = \lambda_0 e^{-\lambda_0 t} \quad F(t) = 1 - e^{-\lambda_0 t} \quad \lambda(t) = \frac{f(t)}{1 - F(t)} = \lambda_0 \quad (1.14)$$

Quand le taux de défaillance est constant, alors le MTTF correspond simplement à son inverse tel que $MTTF = \frac{1}{\lambda_0}$.

Distribution log-normale

Lorsque le logarithme d'une variable aléatoire est distribué selon une loi normale, alors, on parle de distribution log-normale. La fonction de densité, la fonction de répartition et la fonction du taux de défaillance s'écrivent :

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} e^{-\frac{(\ln(t)-\ln(\mu))^2}{2\sigma^2}} \quad F(t) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left[\frac{\ln(t) - \mu}{\sigma\sqrt{2}} \right] \quad \lambda(t) = \frac{\frac{1}{t\sigma\sqrt{2\pi}} e^{-\frac{(\ln(t)-\ln(\mu))^2}{2\sigma^2}}}{\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left[\frac{\ln(t) - \mu}{\sigma\sqrt{2}} \right]} \quad (1.15)$$

Parce que c'est le logarithme du temps qui est distribué selon une loi normale, les coefficients μ et σ^2 ne correspondent pas exactement à la moyenne et à la variance comme dans une loi normale. En effet, la moyenne qu'on notera μ' pour la différencier est $\mu' = e^{\mu + \sigma^2/2}$ et la variance $\sigma'^2 = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$. Dans une loi log-normale, le paramètre μ modélise l'amplitude du logarithme de la variable aléatoire et le coefficient σ^2 est la variance du logarithme de la variable aléatoire. Dans l'équation (1.15), la fonction erf est la fonction d'erreur :

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (1.16)$$

Distribution de Weibull

Comme la distribution log-normale, la loi de Weibull est une distribution à deux paramètres. Le premier paramètres α ajuste l'échelle du temps et le deuxième paramètre β la forme de la courbe. En modifiant les paramètres α et β il est possible de représenter les trois régions de la courbe en baignoire. La densité de probabilité, la fonction de répartition et la fonction du taux de défaillance s'écrivent :

$$f(t) = \frac{\beta t^{\beta-1} e^{-(t/\alpha)^\beta}}{\alpha^\beta} \quad F(t) = 1 - e^{-(t/\alpha)^\beta} \quad \lambda(t) = \frac{\beta t^{\beta-1}}{\alpha^\beta} \quad (1.17)$$

Lorsque $\beta = 1$ alors on est dans le cas particulier où la distribution de Weibull correspond à une distribution exponentielle avec un taux de défaillance constant ($\lambda_0 = \frac{1}{\alpha}$). Bien que les distributions log-normale et de Weibull semblent à premier abord répondre au même besoin, celles-ci sont complémentaires.

Les distributions log-normale tendent à s'appliquer lorsqu'une dégradation est graduelle et se produit au fil du temps en raison d'effets de diffusion (par exemple le BTI présenté dans la section 1.5.2), de corrosion ou de réaction chimique. D'autre part, la distribution de Weibull semble être applicable dans le cas où le premier des nombreux défauts se propage jusqu'à la rupture. Les ruptures diélectriques (par exemple le TDDDB présenté dans la section 1.5.4), les défaillances de condensateurs et les fractures de céramiques sont généralement décrites par des distributions de Weibull.

1.5.8 Facteurs d'accélération (AF)

Les composants électroniques de dernière génération sont devenus très fiables. Pour faire une étude de fiabilité dans les conditions opérationnelles d'utilisation, il faudrait faire des tests de vieillissement, parfois pendant des dizaines d'années. Le test de vieillissement accéléré est un moyen de mesurer la fiabilité d'un composant pour un temps de test bien plus court. Les moyens d'accélérer le vieillissement dépendent des mécanismes de défaillance et sont souvent des contraintes d'environnements : température fixe ou cyclique, une sur-tension, un environnement très humide et bien d'autres. Le facteur d'accélération est un moyen d'obtenir la fiabilité d'un composant dans les conditions d'utilisations opérationnelles à partir des mesures de fiabilité obtenues en réalisant un test de vieillissement accéléré. Le facteur d'accélération peut s'exprimer

comme étant le rapport du MTTF dans les conditions opérationnelles d'utilisation et du MTTF issu du test de vieillissement accéléré, tel que :

$$AF = \frac{MTTF_{\text{opérationnel}}}{MTTF_{\text{stress}}} \quad (1.18)$$

L'étude présentée dans [109] introduit un logiciel pour calculer la durée de vie d'un dispositif dans les conditions d'utilisation opérationnelle en extrayant les facteurs d'accélération de multiples stress à partir des données de mesure d'un test de vieillissement accéléré. Ainsi, plusieurs modèles de facteurs d'accélération pour différents stress sont proposés dans [109]. Dans cette partie, nous présentons uniquement les facteurs d'accélération en température et en tension qui seront utilisés par la suite.

Le facteur d'accélération en température reprend la loi d'Arrhenius initialement inventée pour décrire la vitesse d'une réaction chimique en fonction de la température. Cette loi s'exprime par :

$$AF_T = e^{-\frac{E_a}{k_B} \left(\frac{1}{T_{\text{opérationnel}}} - \frac{1}{T_{\text{stress}}} \right)} \quad (1.19)$$

où E_a l'énergie d'activation, k_B la constante de Boltzmann, $T_{\text{opérationnel}}$ la température d'utilisation opérationnelle et T_{stress} la température de contrainte.

Le facteur d'accélération en tension suit une loi exponentielle :

$$AF_V = e^{\gamma(V_{\text{stress}} - V_{\text{opérationnel}})} \quad (1.20)$$

où γ est le paramètre d'accélération électrique en V^{-1} , V_{stress} et $V_{\text{opérationnel}}$ sont respectivement la tension de contrainte et la tension d'utilisation opérationnelle (0.85V).

Lorsque le composant est soumis à plusieurs contraintes en parallèle, comme détaillé dans FIDES [110], le facteur d'accélération prend la forme de :

$$AF = e^{\left(\sum_{i=1}^S \alpha_i \cdot (g(X_{0i}) - g(X_i)) \right)} = AF_V \times AF_T \times \dots \quad (1.21)$$

où S est le nombre de contraintes appliquées, α est un vecteur de paramètres qui dépend de la loi de défaillance, g correspond au modèle de facteur d'accélération considéré et X correspond à un changement de variable. Donc le facteur d'accélération global correspond au produit des différents facteurs d'accélération.

1.6 Méthodes et mesures de dégradation des FPGAs

1.6.1 Introduction

Les FPGAs sont des circuits numériques dont l'architecture interne est configurable. Il est constitué de cellules logiques tel que la Look-Up Table (LUT) ou le multiplexeur qui peuvent implémenter des circuits logiques combinatoires. Un FPGA contient aussi des bascules qui synchronisent les signaux internes à une horloge globale afin d'implémenter de la logique séquentielle. En décrivant l'architecture interne d'une LUT dans la figure 1.19, nous avons montré que le composant élémentaire dans un FPGA est le transistor. Les mécanismes de défaillance tel que le BTI (Bias Temperature Instability) ou le HCI (Hot Carrier Injection) ont pour conséquence la dégradation de la tension de seuil du transistor ce qui à pour effet, au niveau du circuit, d'augmenter le temps de propagation d'un signal dans un circuit numérique tel que le FPGA.

Des méthodes ont été développées visant à mesurer les dégradations dans un FPGA pour en prédire sa fiabilité sous l'effet du vieillissement. Ces méthodes s'appuient sur la mesure du temps de propagation d'un signal dans un circuit combinatoire. Puisque le temps de propagation dans

un circuit combinatoire ne dépend que du temps de propagation dans les ressources du circuit, la mesure du temps de propagation permet d'observer l'effet de l'augmentation de la tension de seuil dû au BTI et au HCI. Bien que les mesures soient réalisées au niveau du FPGA, elles permettent d'observer l'effet des mécanismes de défaillance des transistors de dernière génération, tel que le 16nm FinFET dans notre cas, sur le fonctionnement du FPGA.

Parce que nous étudions les dégradations dans un FPGA sous l'effet du vieillissement seulement, dans cette section, nous présentons les deux principales méthodes développées pour mesurer l'évolution du temps de propagation dans un FPGA sous l'effet du vieillissement. Puis nous confrontons les résultats obtenus dans différents test de vieillissement de FPGA afin d'en ressortir les éléments importants qui nous ont guidés dans le développement de notre banc de test présenté dans le chapitre 2.

1.6.2 Méthode Transition Probability (TP)

Cette méthode, développée dans [111], consiste à mesurer la probabilité de transition d'un signal en sortie du *Circuit Under Test* (CUT). La figure 1.39 introduit l'architecture pour la mesure du temps de propagation d'un signal dans la CUT avec la méthode TP. La probabilité de transition d'un signal est définie ici comme la probabilité pour que la sortie de le CUT change d'état lorsque le stimuli d'entrée est appliqué au circuit. Les signaux d'entrées de la CUT proviennent du *Test Vector Generator* (TVG) et sont synchronisés avec l'horloge du *Test Clock Generator* (TCG). Le signal en sortie de la CUT est échantillonné par la même horloge que le stimuli d'entrée. Si le temps de propagation du signal dans la CUT est inférieur à la période d'horloge de la TCG, alors l'état du signal échantillonné par le *Sample Register* (SR) correspondra à l'état présent du stimuli d'entrée. Si le temps de propagation du signal dans la CUT est supérieur à la période d'horloge de la TCG, alors l'état du signal échantillonné par le SR correspondra à l'état précédent du stimuli d'entrée. Le TCG est un ensemble de *Phase-Locked Loop* (PLL) qui génère une horloge avec un balayage des fréquences, en commençant par les basses fréquences et en les augmentant peu à peu. La résolution de la mesure du temps de propagation (Δt) dépend du pas de la fréquence de balayage (Δf) :

$$\Delta t \approx \frac{\Delta f}{f^2} \quad (1.22)$$

Avec $\Delta f = 0.25\text{MHz}$ et $f = 500\text{MHz}$ la résolution de la mesure est de $\Delta t = 1\text{ps}$ [111]. Le *Transition Activity Counter* (TAC) compte le nombre de transitions du signal et définit ainsi la probabilité de transition qui correspond à :

$$D(y) = \frac{N_{trans}}{K} \quad (1.23)$$

où N_{trans} est le nombre de transitions du signal et K est le nombre de périodes d'horloge de la TCG.

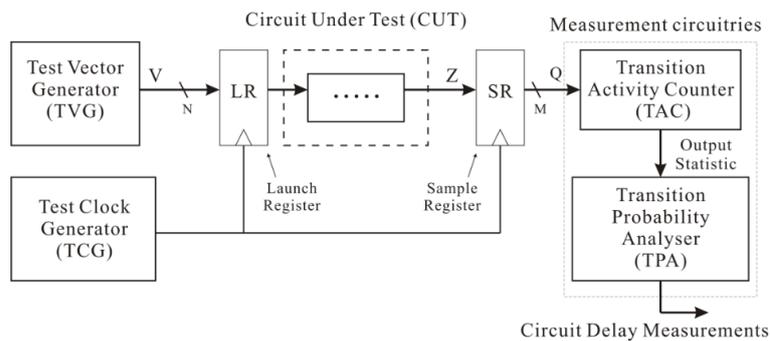


FIGURE 1.39 – Architecture générale pour la mesure du temps de propagation d'un chemin combinatoire avec la méthode TP [111]

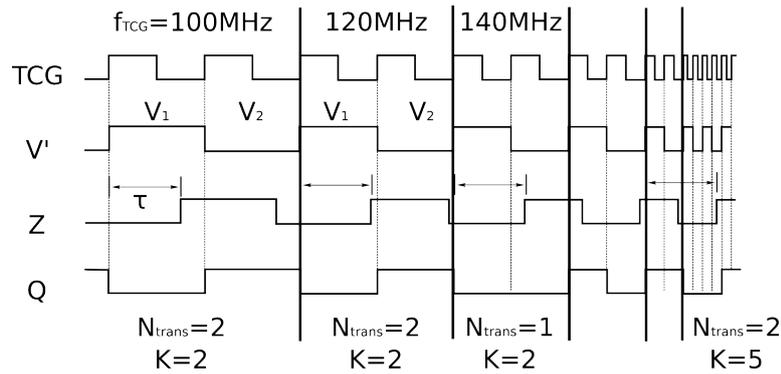


FIGURE 1.40 – Chronogramme des signaux internes de l'architecture TP présenté sur la figure 1.39 en considérant que le TVG génère un signal sur 1 bit dont la fréquence est deux fois plus faible que f_{TCG} et on considère que le CUT est une chaîne de buffers. Le signal V' correspond au signal de TVG en sortie du "launch register"

Afin de mieux comprendre le fonctionnement de la méthode TP, nous introduisons un exemple basique en étudiant le chronogramme présent sur la figure 1.40. Dans cet exemple la CUT est constituée d'une chaîne de buffers. Considérons que le TCG fournit un signal d'horloge qui augmente de 20MHz chaque deux périodes d'horloge. Le TVG génère un signal de stress d'un bit qui alterne entre un niveau logique haut puis bas et dont la fréquence correspond à $\frac{f_{TCG}}{2}$. Sur le chronogramme nous notons V' le signal généré par TVG et synchronisé avec l'horloge de TCG. Le signal Z en sortie de la CUT correspond au signal V' avec un retard τ provenant du temps de propagation du signal V' dans la CUT. Finalement Q correspond au signal Z mais synchronisé sur le signal d'horloge TCG. Lorsque $f_{TCG} = 100\text{MHz}$ ou $f_{TCG} = 120\text{MHz}$ on constate que le retard τ est inférieur à une période d'horloge TCG, par conséquent $N_{trans} = K$. Lorsque $f_{TCG} > 120\text{MHz}$ on constate que le retard est supérieur à la période d'horloge TCG, par conséquent on constate que $N_{trans} < K$. Plus la fréquence d'horloge augmente et plus N_{trans} diminue. Dans cet exemple on en déduit que $\tau \approx \frac{1}{140} = 7\text{ns}$.

Un résultat caractéristique de la probabilité de transition $D(y)$ en fonction de la fréquence d'horloge est présenté sur la figure 1.41. Lorsque la fréquence d'horloge est basse $N_{trans} = K$ soit $D(y) = 1$ puisque le temps de propagation du signal dans la CUT est inférieur à la période d'horloge. Puis, à partir d'une certaine fréquence f_{max} , $D(y)$ décroît ce qui veut dire que le temps de propagation du signal dans la CUT devient supérieur à la période d'horloge. C'est à partir de f_{max} que le temps de propagation du signal dans la CUT est défini.

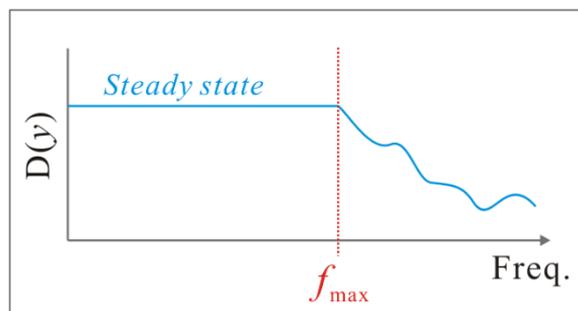


FIGURE 1.41 – Probabilité de transition $D(y)$ en fonction de la fréquence d'horloge TCG [111]

Un des avantages de la méthode TP est sa capacité à mesurer le temps de propagation de tous les chemins combinatoires implémentés dans un FPGA sans contrainte d'architecture de la CUT. De plus, grâce au TVG, il est possible de mesurer le temps de propagation de la CUT pour différentes entrées ce qui permet de connaître l'entrée pour laquelle le temps de propagation dans la CUT est le plus critique. Pour les inconvénients, on relève la complexité de la méthode qui néces-

site un circuit pour générer une horloge à balayage de fréquences. Le temps de propagation est mesuré grâce à une horloge possédant une certaine gigue. L'étude [111] montre l'effet de la gigue d'horloge sur les résultats de la mesure du temps de propagation et l'importance d'implémenter un circuit de traitement des mesures (TPA) qui complexifie, là encore, l'architecture de la méthode TP.

1.6.3 Mesure par Ring Oscillator (RO)

L'oscillateur en anneau (RO) est un circuit composé de n portes logiques où le signal de sortie est rebouclée sur l'étage d'entrée et dont le signal doit être inversé à chaque passage dans le circuit générant ainsi un signal dont la période :

$$T_{RO} = 2 \times \sum_{i=1}^n \tau_i \quad (1.24)$$

où τ_i correspond au temps de propagation dans la porte logique i .

Dans un circuit, un RO peut être utilisé pour :

- Caractériser les performances d'un circuit : En implémentant des RO dans un FPGA, l'auteur [112] mesure les variations stochastiques et systématiques ;
- Fabriquer des *Physical unclonable function* (PUF) pour garantir la sécurité d'un système : Dans [113] Douadi et al. mesurent l'effet de la température sur le fonctionnement d'un RO-PUF, implémenté dans un FPGA, révélant sa vulnérabilité lors d'une *Power-Off Attack* (POA) ;
- Mesurer le vieillissement d'un FPGA...

L'utilisation des ROs pour mesurer le vieillissement d'un FPGA a été découvert et présenté dans différents travaux [114–117]. Cette méthode consiste à mesurer la fréquence d'oscillation du signal de sortie d'un oscillateur en anneau (RO) afin de calculer le temps de propagation d'un signal qui le traverse. Dans un FPGA les portes logiques sont généralement implémentées par des LUTs.

Pour mesurer la fréquence du RO plusieurs méthodes ont été développées : compteur (CPT) [117], électromagnétique (EM) [115] et le UnderSampling (US) [118].

Mesure avec un compteur

Pour la méthode compteur [117], le circuit de mesure, présenté sur la figure 1.42, est implémenté en utilisant les ressources internes du FPGA. Le circuit mesure non seulement la fréquence d'oscillation du RO mais aussi son rapport cyclique. Le *measured signal* correspond au signal de sortie du RO. Un compteur utilise le signal de sortie du RO pour compter le nombre de périodes du RO lorsque le signal *enable* est actif. En connaissant le temps actif du signal *enable* (t_{enable}) nous pouvons calculer la fréquence d'oscillation du RO :

$$f_{RO} = \frac{number_{cycle}}{t_{enable}} \quad (1.25)$$

avec $number_{cycle}$ la donnée en sortie du compteur de fréquence.

Les deux compteurs du bas sont utilisés pour mesurer le rapport cyclique du signal en sortie du RO. Une horloge de référence avec une fréquence supérieur à la fréquence d'oscillation du RO est utilisée pour générer un signal *carry* avec le premier compteur. Le deuxième compteur compte le nombre de périodes d'horloge lorsque le niveau logique du signal en sortie du RO est à '1' et jusqu'à ce que le signal *carry* généré par le précédent compteur soit à '1'. Ainsi, en connaissant la taille du compteur générant le signal *carry* (k) nous pouvons calculer le rapport cyclique d'oscillation du RO :

$$\alpha_{RO} = \frac{value_{duty\ cycle}}{2^k - 1} \quad (1.26)$$

avec $value_{duty\ cycle}$ la donnée en sortie du compteur du rapport cyclique.

L'avantage d'utiliser cette méthode de mesure réside dans l'utilisation uniquement de ressources disponibles dans un FPGA ce qui réduit la complexité et le coût du banc de test. Cependant, le circuit de mesure consomme des ressources du FPGA et réduit le nombre de ressources disponibles pour implémenter d'autres fonctionnalités ou plus de RO par exemple.

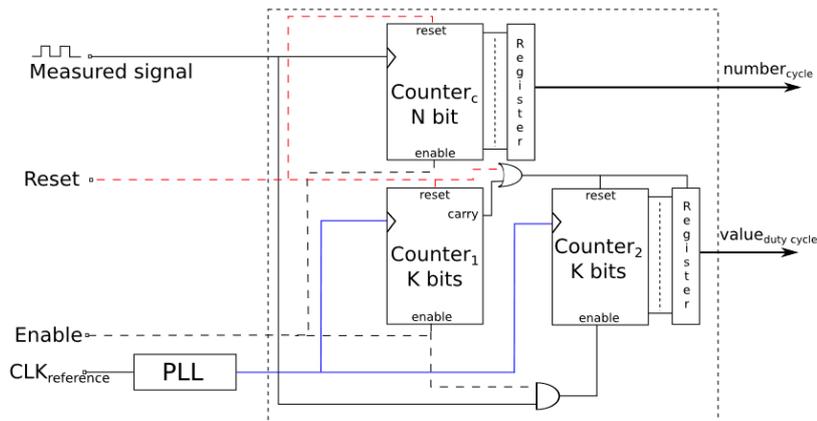


FIGURE 1.42 – Schéma du circuit de mesure de la fréquence et du rapport cyclique de l'oscillation du signal de sortie d'un RO avec la méthode du compteur [117]

Mesure électromagnétique

La méthode électromagnétique [115] se base sur le principe que le flux magnétique généré par un circuit intégré dépend de sa consommation instantanée de courant. La figure 1.43 présente le banc de test utilisant la méthode électromagnétique. Les oscillateurs en anneaux sont implémentés sur toute la surface du FPGA. Une sonde qui mesure le rayonnement électromagnétique balaye les zones du FPGA où les ROs sont implémentés. Les mesures de la sonde sont d'abord transmises à un oscilloscope qui amplifie, traite et échantillonne le signal avant d'envoyer les données vers un ordinateur pour une nouvelle phase de traitement qui aboutit au résultat de la fréquence des RO. Sur la figure 1.43 on constate que le résultat en sortie de ce banc de test est une cartographie de la fréquence des ROs mesurés dans les différentes zones du FPGA.

Contrairement à la méthode du compteur, la méthode de mesure électromagnétique est un circuit de mesure externe au FPGA ce qui libère des ressources dans le FPGA. Cependant, cette méthode requiert un équipement coûteux et la fréquence du RO est obtenue après plusieurs phases de traitement. Aussi, cette méthode ne permet pas de mesurer le rapport cyclique du signal d'oscillation d'un RO.

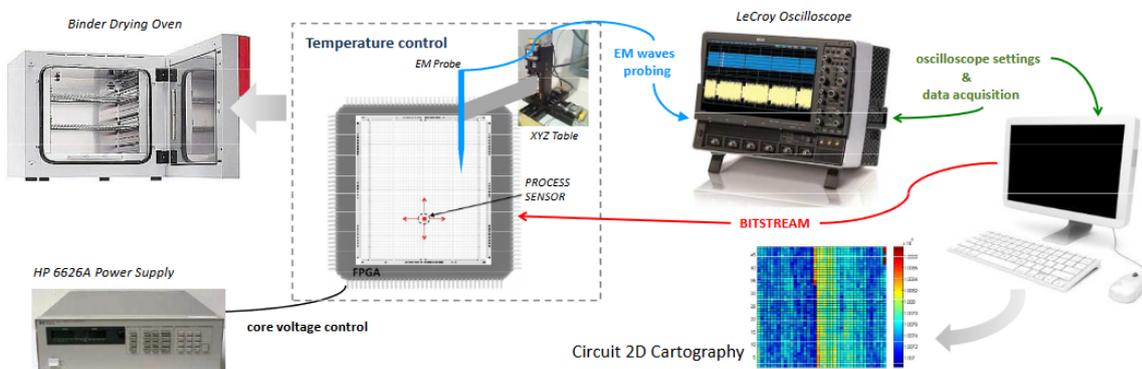


FIGURE 1.43 – Banc de test pour la mesure de fréquence de l'oscillation du signal de sortie d'un RO avec la méthode électromagnétique [115]

Mesure undersampling

La méthode de *undersampling* se base sur le principe d'échantillonnage d'un signal qui est fondée sur le théorème de Nyquist-Shannon [118]. La figure 1.44 présente le principe de fonctionnement de la mesure de fréquence et de rapport cyclique du signal en sortie d'un RO. Le signal en sortie d'un RO est échantillonné par la méthode de *undersampling* qui consiste à échantillonner un signal avec une fréquence inférieure à la fréquence de Nyquist [118]. Le signal du RO échantillonné est stocké dans des mémoire Block RAM (Random Access Memory) qui est une ressource dans le FPGA. Puis à partir des données d'échantillonnage du signal du RO le rapport cyclique peut facilement être calculé :

$$\alpha_{RO} = \frac{\sum_{j=1}^n x_j}{n} \quad (1.27)$$

avec n le nombre d'échantillons stockés dans la BRAM et x_j est le j^{eme} échantillon du signal du RO dont la valeur est soit '1' ou '0'. La fréquence du signal en sortie du RO est obtenue à partir du produit de la fréquence d'échantillonnage et du rapport du nombre de transitions du signal sur le nombre d'échantillons. Plus de détails concernant cette méthode sont donnés dans [116, 119]. Comme pour la méthode du compteur, l'avantage de l'*undersampling* réside dans l'implémentation du circuit numérique dans le FPGA. De plus, le traitement de l'échantillonnage est fait dans un processeur et peut être réalisé ultérieurement puisque les échantillons sont stockés dans des BRAM. En revanche, l'utilisation des BRAM pour stocker l'échantillonnage du signal du RO occupe un nombre important de ressources dans le FPGA.

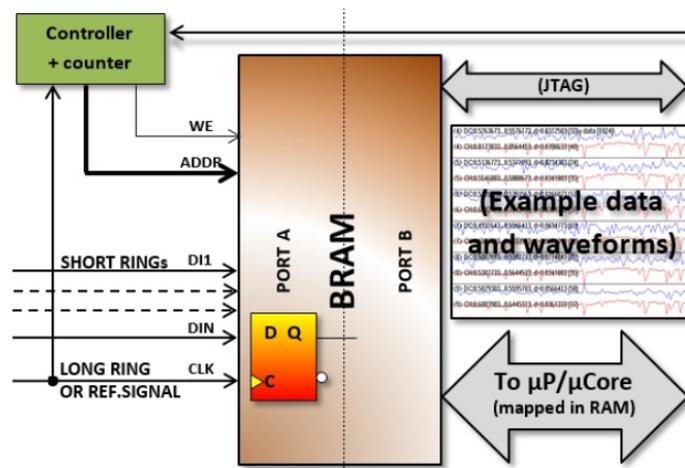


FIGURE 1.44 – Principe du circuit de mesure de la fréquence d'oscillation d'un RO avec la méthode *Undersampling* [118]

Face aux trois méthodes que nous venons de présenter (compteur, électromagnétique et *undersampling*) pour mesurer la fréquence et le rapport cyclique d'oscillation d'un RO, nous avons fait le choix de la méthode du compteur. En effet, il semble que cette méthode est la plus simple à mettre en place et qu'elle ne requiert pas d'instrumentation ou de connaissance supplémentaire. Nous présenterons notre architecture dans le chapitre 2.

1.6.4 Mesures de vieillissement des FPGAs

Mesurer le temps de propagation d'un signal dans un FPGA est un bon indicateur quant à l'état de vieillissement de celui-ci. Nous venons de présenter plusieurs méthodes visant à mesurer le temps de propagation d'un signal dans un circuit combinatoire du FPGA. Dans cette partie nous présentons des résultats de tests de vieillissement sur FPGA.

Mesure avec la méthode TP

Dans [114] l'auteur mesure la dégradation du temps de propagation en utilisant la méthode de *Transition Probability* (TP) et la méthode du compteur pour mesurer la fréquence d'un RO afin de comparer les résultats des deux méthodes. Pour la méthode TP, le circuit sous test est le même que le RO excepté que la sortie n'est pas rebouclée sur l'entrée. Ainsi la figure 1.45 montre la mesure de dégradation de la fréquence d'oscillation d'un RO mesuré avec la méthode TP et avec la méthode du compteur. On constate que les deux méthodes donnent des résultats équivalents, montrant que les deux méthodes ont une qualité de mesure équivalente. De plus, cela confirme la capacité des deux méthodes à mesurer avec justesse la fréquence d'un RO. Les données qui sont tracées sur la figure 1.45 sont en réalité des moyennes de plusieurs mesures. Une moyenne est appliquée sur les mesures pour atténuer l'instabilité de la fréquence d'oscillation d'un RO. Finalement, on constate que l'amplitude des dégradations maximum sont de 13% en 1500 heures de vieillissement ce qui, nous verrons, est important comparé aux résultats d'autres tests de vieillissement.

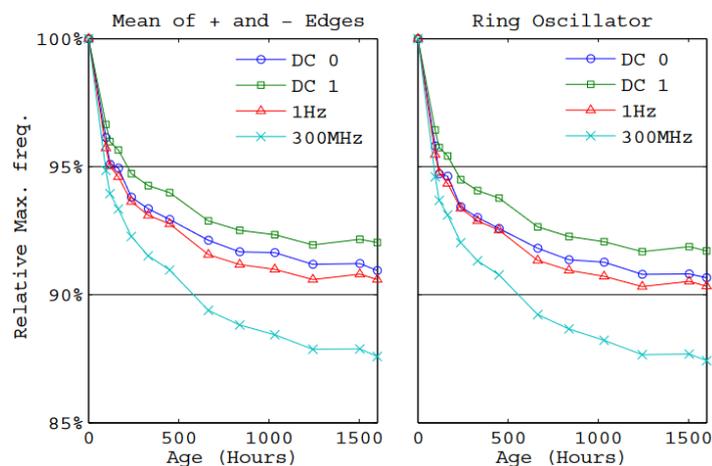


FIGURE 1.45 – Mesure de dégradation de la fréquence relative d'oscillation d'un RO mesuré avec la méthode de *Transition Probability* (gauche) et la méthode du compteur (droite). $T_{stress} = 75^{\circ}C$ et $V_{stress} = 1.8V_{nom}$ [114]

Mesure avec la méthode d'*undersampling*

Dans [119] l'auteur présente la capacité de la méthode d'*undersampling* à mesurer les dégradations de fréquence de RO dans un FPGA. Sur la figure 1.46 nous avons rapporté un des résultats obtenus. La figure 1.46 montre que la fréquence du RO subissant un stress statique de niveau '0' (rouge) se dégrade plus vite que celle du RO qui est toujours en auto-oscillation (vert). L'amplitude des dégradations dans le cas du RO avec un stress statique sont de 0.2% en seulement 7 jours (168 heures) de mesures et pour des conditions de stress qui ne sont pas précisées. Par ailleurs, nous relevons sur la figure 1.46 l'effet important du bruit des mesures qui empêche l'auteur de réaliser une modélisation plus précise.

Mesure avec la méthode électromagnétique

Dans [120] l'auteur présente les dégradations de la fréquence relative de RO en utilisant la méthode électromagnétique dont on reporte un des résultats sur la figure 1.47. Parce que les mesures avec cette méthode ne sont pas automatisées, les mesures sont réalisées à quatre instants pendant 15 jours de vieillissement. Il n'est donc pas possible d'observer en détail l'évolution des dégradations et donc encore moins de modéliser la dégradation de la fréquence des ROs. Cependant, la figure montre que lorsque les ROs sont stressés (du jour 0 au jour 7) alors la fréquence d'oscillation des ROs diminue, cependant, lorsque les ROs ne sont plus stressés (du jour 7 au jour 14) la

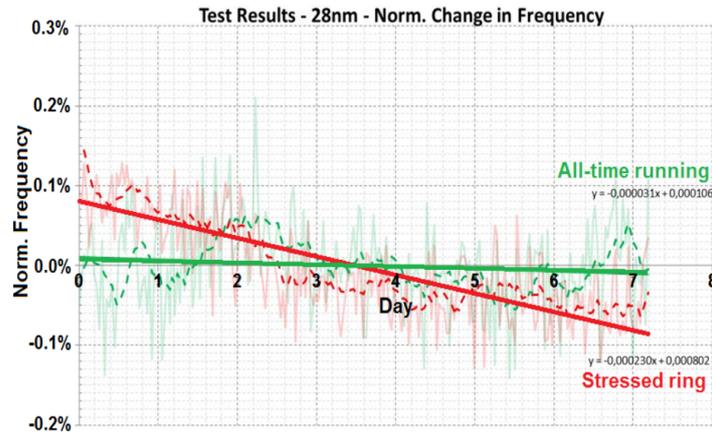


FIGURE 1.46 – Mesure de dégradation de la fréquence relative de ROs implémenté dans un FPGA Virtex 28nm, mesuré avec la méthode *undersampling* et dans des conditions de température et de tension non connues [119]. La courbe verte représente les dégradations pour un RO dont le seul stress est sa propre fréquence d’oscillation. La courbe rouge représente les dégradations d’un RO avec un stress statique de niveau '0'

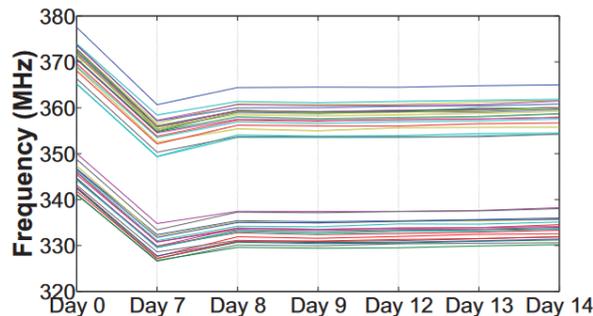


FIGURE 1.47 – Dégradation de la fréquence d’oscillation de plusieurs ROs (différentes couleurs). Les mesures sont réalisées au jour 0, jour 7, jour 8 et jour 14. Du jour 0 au jour 7 : 80°C et 1.5V_{nom}. Du jour 7 au jour 14 : 25°C et V_{nom}. [120]

fréquence des ROs revient vers leurs fréquences initiales. Ceci montre un effet possible du mécanisme de récupération du BTI au niveau du transistor qui se manifeste au niveau du circuit en l’augmentation de la fréquence du RO.

Mesure avec la méthode du compteur

L’étude réalisée dans [121] montre les dégradations de fréquence d’oscillation de RO pour, à priori, le plus long test de vieillissement (12000 heures). De plus, le banc de test inclut 45 FPGAs qui sont stressé dans différentes conditions de températures (de -30°C à 115°C), de tensions (de V_{nom} à 1.5V_{nom}) et pour des fréquences et rapports cycliques de stress différents. La figure 1.48 rapporte les dégradations de fréquence relative pour différentes conditions de stress. On constate clairement l’augmentation de l’amplitude des dégradation lorsque la tension de stress augmente. Pour une tension de stress élevée il est possible de distinguer l’effet du rapport cyclique de stress sur les dégradations, cependant, lorsque la tension de stress diminue il est plus difficile de dissocier les dégradations en fonction des conditions de stress. Sur cette figure nous voyons clairement l’importance de la précision des mesures pour observer des dégradations de la fréquence du RO pour différentes configurations lorsque la tension et la température de stress tendent vers les conditions d’utilisation opérationnelle.

Une modélisation du facteur d’accélération en tension est obtenue avec une loi puissance et une loi exponentielle, montrant ainsi la justesse et la cohérence des deux modèles. En utilisant une loi d’Arrhenius pour le facteur d’accélération en température, l’effet du BTI sur la dégradation

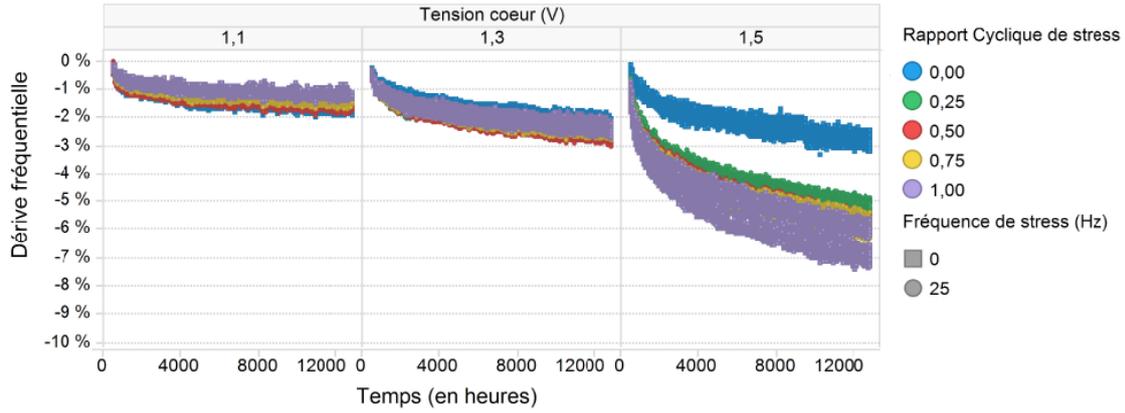


FIGURE 1.48 – Mesure de dégradation de la fréquence relative de RO pour une tension de stress de $1.1V_{nom}$, $1.3V_{nom}$ et $1.5V_{nom}$, une température à 90°C et pour différents rapport cyclique [121]

de la fréquence d'oscillation du RO (νf) est modélisé avec une loi puissance :

$$\nu f = A \exp\left(\gamma \frac{V}{V_{nom}}\right) \exp\left(\frac{-Ea}{k_B T}\right) \cdot t^n \quad (1.28)$$

Nous utiliserons ce même modèle pour modéliser les dégradations que nous observerons. Pour modéliser l'effet du HCI sur les dégradations de la fréquence du RO, l'auteur propose de soustraire aux dégradations avec un stress de haute fréquence les dégradations avec un stress de basse fréquence afin d'éliminer les dégradations dues au BTI à basse température :

$$\nu f_{\text{HCI}}(\alpha_{stress}, f_{stress}) = \nu f(\alpha_{stress}, f_{stress}) - \nu f(\alpha_{stress}, 25\text{Hz}) \quad (1.29)$$

L'auteur considère que l'amplitude des dégradations dues au HCI dépend du nombre de commutations du transistor et par conséquent de la fréquence de stress. Ainsi les dégradations sont tracées en fonction du $temps \times f_{stress}$ et montrent un bon alignement pour une fréquence de stress entre 50MHz et 800MHz.

1.6.5 Modélisation de la fiabilité

L'étude réalisée dans [122] présente la dégradation de la fréquence d'oscillation d'un RO en fonction de la racine carré du temps. Cette représentation, peu commune, est ensuite utilisée pour modéliser la dérive de la fréquence du RO (α_{slope}) :

$$\alpha_{slope} = \frac{\Delta f}{f_0 \times \sqrt{t}} \quad (1.30)$$

À partir de l'extraction du modèle de dérive, l'auteur extrapole les mesures jusqu'à obtenir une dégradation de la fréquence relative de 10% en notant le temps de défaillance TTF (Time To Fail) pour chacun des ROs : $TTF = \left(\frac{10\%}{\alpha_{slope}}\right)^2$. Ensuite, il semble que l'auteur a identifié le taux de défaillance sans avoir modélisé la répartition des TTF en exprimant le taux de défaillance en FIT (Failure In Time) comme étant $FIT = \frac{10^9}{TTF}$. Le BTI et le HCI sont modélisés séparément, cependant, l'identification des deux mécanismes est simplement faite à partir des conditions de stress (température et tension) et sans analyse du comportement des dégradations qui permet de vérifier avec plus de certitude le mécanisme de dégradation observé. Finalement, l'auteur extrait les facteurs d'accélération en température et en tension avec respectivement une loi d'Arrhenius et une loi exponentielle (voir la table 1.2). La même équipe de recherche a réalisé une étude assez proche sur la modélisation de la fiabilité d'un FPGA Zynq UltraScale+ [123]. Cette fois-ci la modélisation de la répartition des défaillances est faite avec une loi de Weibull. Les facteurs d'accélération sont

TABLE 1.2 – Coefficients des facteurs d’accélération en tension (loi exponentielle) et en température (loi d’Arrhenius) obtenus pour la modélisation de la fiabilité des FPGAs dans les conditions de stress stimulant le BTI (hautes températures et hautes tensions)

	$Ea(eV)$	$\gamma(V^{-1})$
[122]	0.52	17.2
[123]	0.49	11.7
[121]	0.81	11.9

modélisés avec les mêmes modèles que la précédente étude et sont notés dans la table 1.2.

Dans [121] l’auteur modélise aussi la fiabilité du FPGA à partir de mesure de dégradation de la fréquence d’oscillation des ROs. En fixant un seuil de défaillance à 10% l’auteur modélise la répartition des temps de défaillances (TTF) avec une loi log-normale en identifiant les facteurs d’accélération dont les coefficients sont reportés dans la table 1.2. On constate que les coefficients des facteurs d’accélération de température et de tension extraits dans les trois études 1.2 sont du même ordre de grandeur ce qui nous conforte quant à la justesse de la modélisation de la fiabilité du FPGA à partir des mesures de dégradation de la fréquence de RO.

1.6.6 Des techniques complémentaires pour les bancs de test

Utilisation d’un multiplexeur en entrée du RO

Dans [117] l’auteur introduit l’utilisation d’un multiplexeur à l’entrée de l’oscillateur en anneau comme nous voyons sur la figure 1.49. Ainsi, l’utilisateur peut choisir le mode de fonctionnement du circuit : stress ou mesure. En mode mesure, le signal en sortie du circuit est rebouclé à l’entrée ainsi le circuit est un RO et sa fréquence peut être mesurée. En mode stress, le signal d’entrée est connecté à un signal de stress dont l’utilisateur peut choisir les caractéristiques du signal (statique, dynamique, fréquence et rapport cyclique). Cela permet d’étudier l’effet du type de stress sur la dégradation du RO.

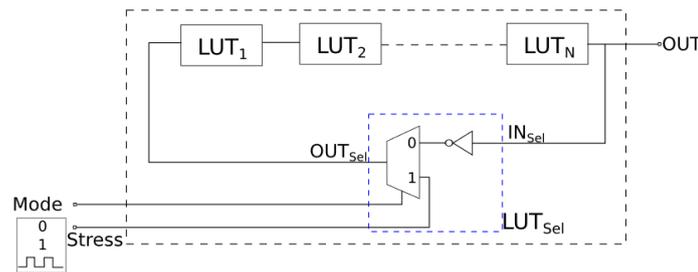


FIGURE 1.49 – Schéma d’un RO avec un multiplexeur en entrée pour choisir le mode de fonctionnement du circuit : stress ou mesure [117]

À partir de la mesure de fréquence (f_{CUT}) et du rapport cyclique (α) d’oscillation du RO, l’auteur [117] montre comment calculer le temps de propagation moyen d’un front montant et descendant dans une LUT du RO constitué de N LUTs :

$$t_{fall} = \frac{1 - \alpha}{f_{CUT} \times N} \quad (1.31)$$

$$t_{rise} = \frac{\alpha}{f_{CUT} \times N} \quad (1.32)$$

Ainsi, l'auteur [117] montre que le temps de propagation d'un front descendant est plus dégradé lorsque le RO subit un stress statique de niveau '0' que lorsque le RO subit un stress statique de niveau '1'. L'effet inverse est observé pour le temps de propagation d'un front montant. Finalement, à partir des mesures, le même auteur [117] étudie et modélise le mécanisme de dégradation BTI [124] et HCI [125].

Un oscillateur en anneau pour mesurer la dégradation des ressources de routage

Dans [114, 126] l'auteur présente une nouvelle architecture de RO consistant à mesurer les dégradations dans les ressources de routage du FPGA. La figure 1.50 montre que cette nouvelle architecture de RO est constituée de LUTs disposées à l'opposées dans le FPGA et qui par conséquent requiert un grand nombre de ressources de routage. Le RO peut s'étendre verticalement (connexion entre une LUT violette et rouge) ou horizontalement (connexion entre une LUT jaune et verte). Ainsi il est possible de mesurer les dégradations des ressources de routage permettant d'assurer des connexions verticales ou horizontales dans le FPGA. Cette architecture est aussi appelée Long Path (LP) à cause de la longueur du chemin qui sépare deux LUTs dans un même RO. Les résultats montrent que les RO LP (dégradation maximum $\approx 10\%$) se dégradent moins que les RO classiques (dégradation maximum $\approx 15\%$) constitués principalement de LUTs. En revanche les RO LP utilisant des ressources de routage permettant d'assurer des connexions verticales ou horizontales dans le FPGA se dégradent de la même façon.

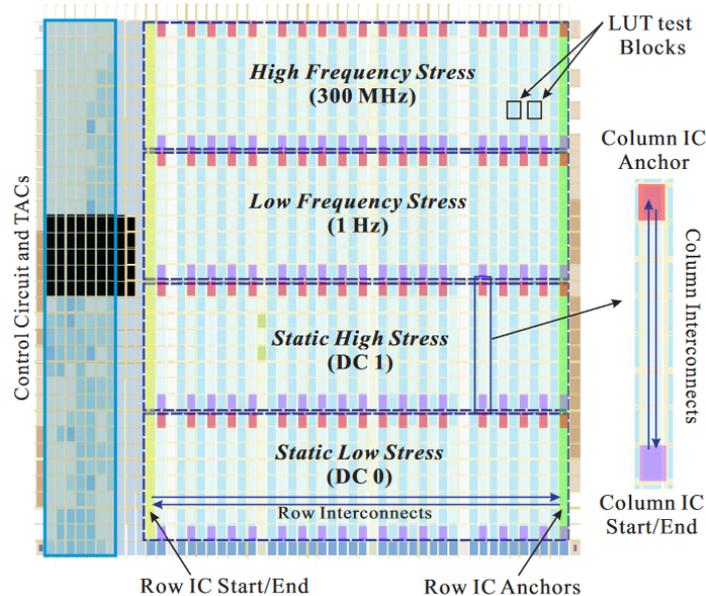


FIGURE 1.50 – Plan en étage du FPGA configuré pour mesurer les dégradations des ressources de routage [114]

Des courts circuits pour augmenter la température interne du FPGA

Dans [127] l'auteur montre une technique pour augmenter la température interne en implémentant plus de 20000 courts-circuits dans le FPGA. Dans un FPGA, un court-circuit correspond à la connexion de la sortie d'une cellule logique avec un niveau logique '1' vers la sortie d'une autre cellule logique avec un niveau logique '0'. Par exemple la figure 1.51 présente le schéma d'un court circuit fait à partir d'une LUT et d'une bascule. Pour implémenter un tel circuit dans le FPGA, l'auteur utilise RapidWright [1] un logiciel de conception avec une description structurée bas niveau donnant plus de souplesse que Vivado. Avec cette technique, l'auteur montre que la température interne du FPGA peut atteindre 177°C sans utiliser de moyens matériel supplémentaires. Avec

une telle température, une dégradation de plus de 5% est mesurée pour moins de 1000 heures de vieillissement.

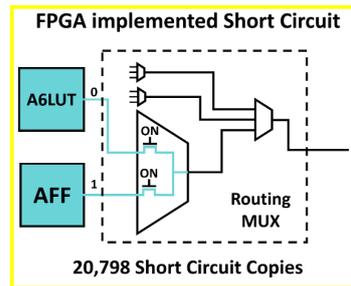


FIGURE 1.51 – Schéma d'un court circuit dans un FPGA en connectant la sortie d'une LUT avec un niveau logique '0' et la sortie d'une bascule avec un niveau logique '1' [127]

Dans notre banc de test nous utilisons seulement les deux premières techniques (multiplexeur et RO avec chemin long). En revanche, la troisième technique n'est pas utilisée puisqu'elle est incompatible avec notre système de régulation en température.

1.6.7 Bilan des mesures de dégradation sur FPGA

Dans la table 1.3, nous avons fait un récapitulatif de la plupart des tests de vieillissement de FPGA publiés et utilisant les méthodes présentées précédemment. Nous avons montré dans les mesures de dégradations (figures 1.45, 1.46, 1.47, 1.48) que la précision de mesure est un paramètre important du banc de test afin de dissocier l'effet des conditions de stress lorsque la tempé-

TABLE 1.3 – Présentation des conditions et des résultats des test de vieillissement de FPGA

Ref.	Méthode	FPGA	T (h)	Volt. $\frac{V_{test}}{V_{nom}}$	Temp.	Max dérive (%)	Précision (%)
[117, 124, 125]	CPT	Cyclone III	1300	1	25°C	2.4	0.5
			360	1.5	25°C	5.5	
[119]	US	Virtex7	144	n.c	n.c	1.5	0.35
		Kintex7	168	n.c	n.c	0.2	
[114, 126]	TP & CPT	Cyclone III	1800	1.8	75°C	15	n.c
				1	25°C	0.6	
[120]	EM	Spartan6	168	1.5	80°C	5.17	n.c
[122]	CPT	Spartan6	200	2	150°C	2.89	n.c
		Zynq7000			-35°C	1	
[127]	CPT	Artix7	864	1	177.5°C	5.13	0.016 (Moyenne*)
[123]	CPT	ZynqUS+	100	1.3	107°C	0.33	n.c
[128]	CPT	Artix7	332	1	35°C	0.3	Moyenne*
					125°C	1.8	
[2, 121]	CPT	Artix7	12000	1.3	-30°C	0.6	0.072
				1.5	115°C	10	
[129]	CPT	ZynqUS+	1000	1.25	100°C	0.5	0.009

*Moyenne : Le résultat est une moyenne de plusieurs mesures

rature et la tension sont proches des conditions opérationnelles. C'est pour cela que dans la table 1.3 en plus de rapporter les dégradations maximales nous notons la précision de mesure du banc de test.

Des dérives de 1% à 15% sont obtenus avec des conditions de stress en tension de 1.8 à 2 fois la tension nominale [114, 120, 122] ou encore avec un stress en température de 177°C [127]. Il est certain qu'avec des stress en tension et en température élevés, loin des conditions opérationnelles d'utilisations, il est possible d'observer des dégradations pour un temps de test relativement courts. Cependant, ce sur-stress peut avoir comme conséquence l'apparition de mécanismes de défaillance qui ne sont pas représentatif de ceux présents dans les conditions d'utilisation opérationnelles. Afin d'observer des dérives de temps de propagation avec un stress proche des conditions nominales et dans un temps de test nominal, il est nécessaire que le banc de test soit dans un environnement stable pour réaliser des mesures précises. En effet, le temps de propagation dépend fortement de la température et de la tension. Dans la plupart des tests, la température et la tension sont fixées mais non régulées, sauf dans [120, 123] où le FPGA est placé dans un four contrôlé en température. Néanmoins, le contrôle de la température est fait à un niveau global et ne régule pas les variations locales dans le FPGA. Afin d'alléger le bruit, le résultat correspond à une moyenne de plusieurs mesures [114, 120, 127, 128], cependant, une telle manipulation produit une perte d'information. Sans avoir fait de moyenne, une précision de mesure de 0.35% est obtenue dans [119] et une précision de 0.5% est obtenue dans [117]. Une telle précision peut être insuffisante pour observer des dérives dans un temps de test court et pour des conditions de stress proches des conditions nominales.

Le banc de test que nous avons développé et qui sera présenté dans le chapitre 2 possède une précision de mesure de 0.009% et, à priori dans l'état de l'art, se classe comme étant le banc de test le plus précis.

1.6.8 Conclusion

Parce que le BTI et le HCI ont pour effet d'augmenter la tension de seuil du transistor ce qui à pour conséquence d'augmenter le temps de propagation d'un signal dans un FPGA, dans cette section nous avons introduit les deux principales méthodes utilisées pour mesurer le temps de propagation d'un signal dans un circuit combinatoire dans un FPGA : la Transition Probability (TP) et le Ring Oscillator (RO). Nous avons introduit trois circuits de mesure différents pour la méthode du RO : *undersampling*, électromagnétique et le compteur.

Dans un deuxième temps, nous avons présenté des résultats de test de vieillissement sur FPGA utilisant les différentes méthodes. En introduisant une étude mesurant les dégradations avec la méthode TP et du RO avec des compteurs, on constate que ces deux méthodes donnent des résultats équivalents. Nous avons vu que les résultats obtenus avec la méthode de l'*undersampling* et électromagnétique sont trop bruités ou manquent de données compliquant ainsi une étude et une modélisation approfondie des dégradations. Bien souvent, les tests de vieillissement utilisent des stress de température et de tension largement supérieurs aux conditions opérationnelles ($T_{stress} = 177^{\circ}\text{C}$ ou $V_{stress} = 2V_{nom}$) afin d'observer des dégradations de l'ordre de quelques pourcent pour un temps de test relativement court (< 1000 heures). Le risque d'appliquer un tel stress est d'observer des mécanismes de défaillance qui ne sont pas représentatif de ceux présent dans les conditions d'utilisation opérationnelles.

Les différentes études, que nous avons présenté, visant à mesurer la dégradation des FPGA mettent en évidence l'importance de développer un banc de test capable de mesurer les dégradations avec précision afin d'observer des dégradations dans des conditions de stress proches des conditions opérationnelles et pour un temps de test acceptable. Dans le banc de test que nous avons développé, nous utiliserons la méthode du RO avec des compteurs puisque que c'est la méthode la plus simple à implémenter et qui offre la meilleur précision de mesure. Finalement, nous

avons montré que notre banc de test, par rapport à l'état de l'art, se classe comme étant le banc de test le plus précis. Nous détaillerons le banc de test que nous avons développé dans le chapitre 2.

1.7 Synthèse

Dans ce chapitre nous avons balayé les notions nécessaires pour comprendre les travaux de mesure et de prévision de la fiabilité des FPGAs que nous présenterons dans les prochains chapitres.

Le transistor est un composant à la base de la conception de tous les circuits numériques. Afin d'améliorer la performance des circuits numériques, la taille du transistor n'a cessé de diminuer en repoussant les limites technologiques avec l'introduction de nouveaux matériaux et l'émergence du FinFET. Dans les circuits numériques, les transistors sont assemblés par paire, de NMOS et PMOS, formant un ensemble complémentaire CMOS. C'est à partir du CMOS que les portes logiques, telles que l'inverseur ou la porte de transferts, sont conçues.

Le FPGA est un circuit numérique qui est l'aboutissement de l'évolution des Programmable Logic Devices (PLDs). La particularité du FPGA, par rapport à d'autres circuits numériques, c'est sa reconfigurabilité. Le FPGA est une imbrication de cellules logiques interconnectées entre elles et avec les entrées et sorties du FPGA. Les cellules logiques ainsi que la matrice d'interconnexion sont configurables (et reconfigurables) par l'utilisateur en utilisant un langage de description matériel tel que le VHDL ou le verilog. Les cellules logiques sont constituées de Look-Up Table (LUT). La LUT est un circuit combinatoire configurable pouvant implémenter différentes fonctions logiques. Par exemple, pour mesurer le vieillissement du FPGA nous implémenterons des oscillateurs en anneaux où chaque étage correspondra à une LUT configurée en buffer, inverseur etc. Des brevets de chez Xilinx décrivent l'architecture interne d'une LUT qui est constituée principalement de portes de transfert et d'inverseurs. L'architecture d'une LUT ressemble à celle d'un multiplexeur où les entrées de la LUT correspondent à l'entrée de commande du multiplexeur et les bits de configuration de la LUT correspondent aux entrées du multiplexeur. Des brevets de chez Xilinx présentent l'architecture de la matrice d'interconnexion qui joue un rôle essentiel pour la performance d'un FPGA. En effet, la matrice d'interconnexion (routage) doit pouvoir connecter toutes les cellules logiques entre elles sans générer une augmentation trop importante du temps de propagation d'un signal. Parce qu'un FPGA contient des dizaines de milliers de cellules logiques, l'architecture du routage au niveau du transistor est bien plus complexe que l'architecture d'une LUT. Il est important de retenir que l'architecture interne d'une LUT se base sur des brevets de chez Xilinx qui sont anciens (2004-2008) et ne sont pas forcément identiques à l'architecture interne de la LUT qui se trouve dans le FPGA que nous utiliserons.

L'état de l'art de la fiabilité micro-électronique révèle trois principaux mécanismes de défaillance affectant le transistor sous l'effet du vieillissement : le Bias Temperature Instability (BTI), le Hot Carrier Injection (HCI) et le Time Dependant Dielectric Breakdown (TDDB). Les trois mécanismes de défaillance ont un mode de défaillance commun qui est la dégradation progressive de la tension de seuil du transistor, ils génèrent une défaillance paramétrique. Le TDDB peut aussi générer une défaillance cataclysmique lorsqu'une accumulation de piège dans l'oxide devient trop importante est créée un chemin conducteur : c'est le claquage d'oxide.

Le BTI est décrit par le modèle de réaction-diffusion comme étant de la diffusion d'hydrogène générant des pièges d'interfaces qui captent les charges dans le canal. Selon le modèle réaction-diffusion la diffusion de l'hydrogène suit une loi de puissance avec un exposant au temps de 0.25. Les tests de vieillissement publiés montrent par la mesure que la dégradation de la tension de seuil suit une loi de puissance avec un exposant 0.25. Pour que le transistor soit stressé il faut qu'un champ électrique soit appliqué sur la grille. Lorsque le transistor n'est pas stressé, la tension de seuil revient vers sa tension de départ, c'est ce qu'on appelle le mécanisme de récupération.

Les conditions favorables pour stimuler le BTI sont les températures hautes, les tensions hautes et un stress statique pour ne pas avoir de mécanisme de récupération. Le mécanisme du BTI est différent suivant la nature du transistor, ainsi, le BTI est divisé en deux sous-mécanismes le NBTI dans le PMOS et le PBTI dans le NMOS. En comparaison avec la technologie planaire, dans le FinFET l'effet du NBTI a légèrement augmenté alors que le PBTI a considérablement diminué. Ainsi l'effet du BTI est moins important dans un CMOS FinFET qu'un CMOS planaire.

Le HCI est décrit par le modèle de l'électron chanceux comme étant la cause d'électrons qui acquièrent une certaine énergie les rendant suffisamment puissant pour rompre une liaison de silicium et d'hydrogène est générant ainsi un piège de surface. Selon le modèle de l'électron chanceux la densité de piège générée suit une loi de puissance avec un exposant au temps entre 0.5 – 1. Les mesures montrent aussi que la dégradation de la tension de seuil suit une loi de puissance le même exposant. Le transistor subit des dégradations dues au HCI lorsqu'un courant le traverse. Les conditions favorables pour stimuler le HCI sont les températures basses, les tensions hautes et un stress dynamique de haute fréquence. Concernant les transistors avec un canal court et donc aussi le FinFET, le HCI semble (selon notre interprétation de l'état de l'art) devenir plus important pour des températures hautes.

Le BTI, le HCI et le TDDB ont pour mode de défaillance la dégradation progressive de la tension de seuil d'un transistor, ce qui a pour effet d'augmenter le temps de propagation d'un signal dans un circuit numérique. Grâce à sa reconfigurabilité et sa modularité, le FPGA est un candidat idéal pour mesurer les dégradations du temps de propagation d'un signal. Deux principales méthodes sont développées pour mesurer le temps de propagation d'un signal dans un chemin combinatoire du FPGA : Transition Probability (TP) et l'oscillateur en anneau (RO). Plusieurs moyens de mesures sont utilisés pour la technique du RO : *undersampling*, électromagnétique et le compteur. Les tests de vieillissement visant à mesurer les dégradations dans des FPGAs révèlent l'importance de la précision de mesure pour observer des dégradations dans des conditions de stress proches des conditions opérationnelles. Ces tests de vieillissement montrent que la méthode de mesure du RO avec des compteurs est la plus précise. De plus, cette méthode est la plus simple d'un point de vue matériel et conception. Ainsi, dans notre banc de test nous implémenterons des oscillateurs en anneaux (ROs) dans le FPGA et nous implémenterons des compteurs pour mesurer la fréquence.

Dans le prochain chapitre nous présentons le banc de test que nous avons développé pour mesurer la dégradation de la fréquence d'oscillateur en anneau. Lors du développement du banc de test, nous avons mis un point d'honneur sur la précision de mesure afin d'observer des dégradations pour des conditions de stress proches des conditions opérationnelles.

Chapitre 2

Développement et caractérisation du banc de test

2.1 Introduction

À ce jour, on compte trois principaux mécanismes qui dégradent les transistors sous l'effet du vieillissement : le BTI, le HCI et le TDDDB. Ces mécanismes ont un mode de dégradation commun qui réside dans l'augmentation de la tension de seuil du transistor ayant pour conséquence l'augmentation du temps de propagation d'un signal dans les circuits intégrés numériques. À partir d'un certain seuil de dégradation de la tension de seuil, le transistor cesse de fonctionner, une panne catalectique est alors observée. Cependant, l'observation de panne catalectique au niveau du transistor est presque impossible pour un temps de test et des conditions de stress raisonnables. D'ailleurs, dans les tests de vieillissement allant jusqu'à 12000 heures (voir table 1.3), aucune panne catalectique n'est observée. C'est pourquoi, la prédiction de la fiabilité doit être basée sur la mesure de la dérive paramétrique telle que la tension de seuil du transistor.

Dans la plupart des études de fiabilité des transistors, la dégradation de la tension de seuil est mesurée directement aux bornes du transistor [56, 58, 61, 62, 82, 87, 88]. Cependant, un tel accès nécessite :

- Le développement d'une carte de test dédiée à la mesure de dégradations des transistors ;
- Des pointes parfois nanométriques en fonction des dimensions du transistor mesuré ;
- Une instrumentation spécifique pour placer les pointes et mesurer la tension de seuil ;

Monopoliser un tel banc de test avec l'accès au transistor est coûteux, c'est pourquoi, le choix de réduire la durée du test de vieillissement devient nécessaire. Afin d'observer des dégradations pour un court temps de test (quelques dizaines d'heures), les conditions de vieillissement appliquées sont très éloignées des conditions opérationnelles (parfois deux fois la tension nominale). Le risque d'appliquer de telles conditions d'accélération du vieillissement est d'observer des mécanismes de défaillances qui ne sont pas représentatifs des mécanismes de dégradations dans les conditions opérationnelles.

Un objectif de cette thèse est de développer un banc de test capable de vieillir et mesurer les dégradations, pendant plus de 10000 heures, de circuits numériques avec des transistors 16nm FinFET, dans des conditions de température et de tension proches des conditions opérationnelles. Parce que les dégradations dans des conditions de température et de tension proches des conditions opérationnelles sont faibles, nous souhaitons développer un banc de test avec une grande précision de mesure. Afin de réaliser une analyse de la fiabilité complète, le banc de test doit vieillir les circuits numériques sous différentes conditions de température et de tension mais aussi pour différents signaux de stress statique et dynamique.

Dans la section 1.6 nous avons introduit les méthodes pour mesurer la dégradation de nouvelles technologies de transistors en utilisant des FPGAs. À partir des mesures de dégradation faites au niveau circuit, c'est à dire au niveau du FPGA, des modèles de HCI [125] et de BTI [124] peuvent être développés. Les FPGAs sont des produits disponibles sur le marché public pour quelques centaines d'euros et utilisent les technologies de transistors les plus récentes. Nous avons fait le choix du FPGA Zynq UltraScale+ (ZU3EG) de chez Xilinx car il est fabriqué à partir de FinFET 16nm. De plus, le FPGA Zynq UltraScale+ contient un système de mesure de la température et de la tension interne du FPGA qui nous sera utile pour réguler la température et la tension aux conditions de vieillissement souhaitées. Pour mesurer la dégradation dans un FPGA, nous implémenterons des Ring Oscillators (ROs) car c'est un circuit dont la fréquence d'oscillation du signal de sortie est sensible au vieillissement. En effet, le vieillissement a pour effet d'augmenter la tension de seuil du transistor dont la conséquence est l'augmentation du temps de propagation du signal dans les étages du RO soit la diminution de la fréquence d'oscillation du signal en sortie du RO. Parmi les différentes méthodes de mesures présentées dans la section 1.6, nous avons fait le choix du compteur qui mesure la fréquence du RO car elle ne requiert aucune instrumentation de mesure externe, ce qui simplifie le banc de test. Finalement, nous utilisons l'Avnet96-V2 comme carte de développement qui contient le FPGA Zynq UltraScale+ ainsi, nous n'avons pas à fabriquer de notre propre carte de développement. La conception du banc de test s'inscrit dans la continuité des thèses de Mohammad Naouss [130] et de Julien Coutet [2] qui ont tout deux conçus un banc de test pour mesurer la fiabilité de FPGA. La thèse de M. Naouss présente une preuve de concept de l'utilisation du FPGA pour mesurer des dégradations tandis que J. Coutet réalise déjà des test sur plusieurs milliers d'heures et dans différentes conditions de vieillissement. Au regard des deux précédentes thèses, l'avancée technique dans la conception du banc de test doit être l'amélioration de la précision de mesure des dégradations qui se manifeste par une meilleur métrologie pour garantir un environnement (température, tension, référence de temps) stable.

Dans ce chapitre nous présenterons le fonctionnement général du banc de test que nous avons développé dans la section 2.2. Dans la section 2.3 nous introduirons le fonctionnement général d'un RO et nous détaillerons les architectures implémentées. Nous présenterons le circuit de mesure dans la section 2.4 et nous justifierons l'utilisation du GPS comme référence de temps dans le circuit de mesure dans la section 2.5. Dans la section 2.6 nous présenterons les systèmes de régulation de température et de tension développés et nous montreront leurs efficacité. Enfin, dans la section 2.7 nous analyserons la qualité du banc de test à mesurer les dégradations à partir d'un test de vieillissement de 1000 heures.

2.2 Description générale du banc de test

Avant de détailler les différentes parties du banc de test, nous allons dans cette section présenter une description générale du banc de test pour mesurer les dégradations d'un FPGA. Le schéma du banc de test est présenté sur la figure 2.1. Sur cette figure, la partie grise représente la carte de développement Ultra96-V2 de chez Avnet [131]. Sur cette carte de développement, il y a un FPGA Zynq UltraScale+ [132] de chez Xilinx (bleu) conçu avec une technologie 16nm FinFET de chez TSMC. Le FPGA Zynq UltraScale+, comme expliqué dans la section 1.4.5, est divisé en deux parties : la Programmable Logic (PL) et le Processor System (PS). Parce que l'architecture de la PL correspond à celle d'un FPGA classique présenté sur la figure 1.16, nous l'avons utilisé pour implémenter les RO ainsi que l'architecture de mesure. Nous avons utilisé la PS pour héberger les régulateurs de température et de tension car le processeur est plus adapté pour les tâches de calcul. Aussi le FPGA Zynq UltraScale+ contient un SYSMON (SYStem MONitoring) qui mesure la tension et la température interne du FPGA et qui sera utilisé comme donnée d'entrée dans les régulateurs de tension et de température.

La carte de développement Ultra96-V2 contient un Power Management Integrated Circuit (PMIC)

qui gère l'alimentation de tous les composants présent sur la carte et donc aussi du FPGA. Nous utiliserons le PMIC comme actionneur pour réguler la tension du FPGA. Concernant la régulation en température, nous utilisons le module Peltier (vert), en contact direct avec le FPGA, comme actionneur. Afin de changer la polarité du Peltier, c'est à dire pour contrôler si le Peltier réchauffe ou refroidit le FPGA, ses deux bornes d'alimentation sont connectées à deux ponts en H (motor driver) différents alimentés par un générateur 7V fixe et contrôlés par deux signaux PWM. Un filtre RLC est utilisé pour lisser les signaux en sortie du pont en H. Afin de protéger le FPGA, qui fonctionne avec des tensions d'alimentation de 3.3V et 0.85V, du circuit d'alimentation du Peltier, nous avons utilisé des optocoupleurs entre le signal PWM en sortie du FPGA et le motor driver. L'optocoupleur a la capacité de transmettre un signal sans qu'il n'y ait aucune connection électrique. D'un côté, un signal commande une LED et de l'autre côté un photo-transistor est commandé par la lumière de la LED. Ainsi nous avons la certitude que l'alimentation du Peltier ne pourra pas endommager le FPGA. Pour que la connection entre les différents composants externes à la carte de développement (en vert) soit fiable, nous avons conçu un PCB sur lequel nous avons soudé ces composants.

Un récepteur GPS est utilisé pour fournir la référence de temps au circuit de mesure. Finalement, les données qui sont mesurées dans le FPGA sont envoyées en série vers un ordinateur qui enregistre les données sur son disque local et les envoie vers un serveur pour protéger les données d'une panne de l'ordinateur.

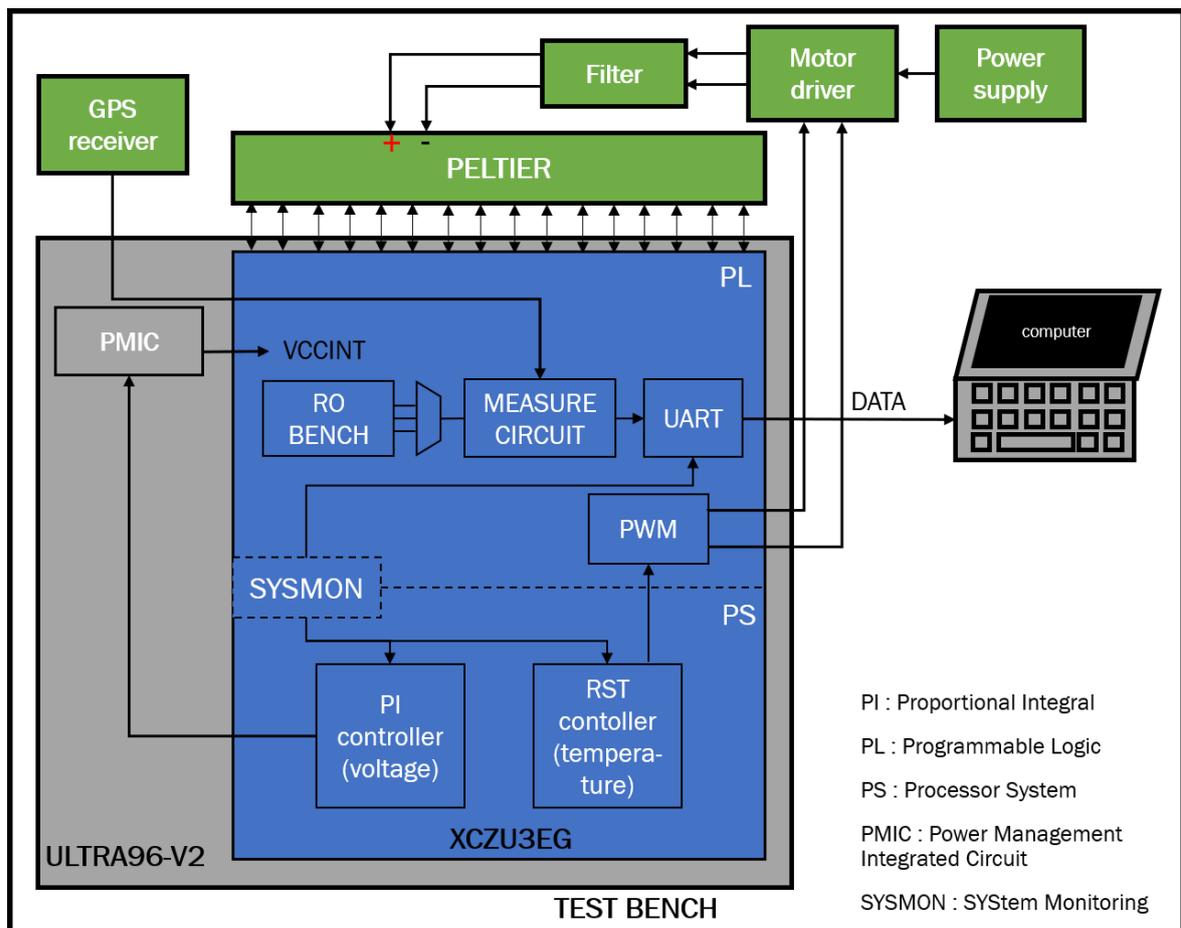


FIGURE 2.1 – Vue d'ensemble de l'architecture du banc de test pour mesurer le vieillissement d'un FPGA

2.3 Banc d'oscillateur en anneau

2.3.1 L'oscillateur en anneau

Afin de mesurer le vieillissement des FPGA, nous implémentons des oscillateurs en anneaux (Ring Oscillators ROs) dans le FPGA dont nous mesurerons la dégradation de la fréquence d'oscillation du signal de sortie. L'oscillateur en anneau est un circuit composé de portes logiques montées en série où le signal de sortie est rebouclé vers l'entrée du circuit. Tous types de portes logiques peuvent être utilisés, la seule contrainte pour que le signal en sortie oscille est la présence d'un nombre impair de portes logiques qui inversent le signal. Sur la figure 2.2a un RO composé de trois inverseurs est représenté. La période du signal de sortie du RO dépend du temps de propagation du signal dans les trois inverseurs (τ_1 , τ_2 et τ_3) selon la relation de l'équation (1.24). Dans sa thèse, Mohammad Naouss [130] propose des architectures de ROs constituées de portes logiques différentes, par exemple deux buffers et un inverseur. L'avantage de changer les portes logiques dans un RO est d'étudier la fiabilité de plusieurs blocs fonctionnels à partir d'une même méthodologie de mesure.

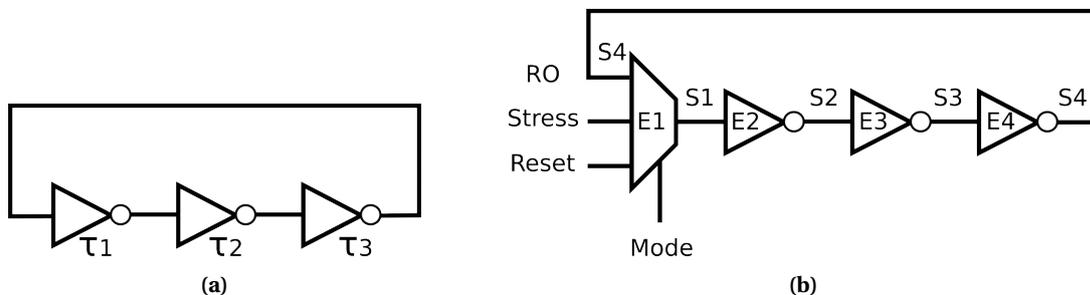


FIGURE 2.2 – Schéma logique d'un oscillateur en anneau (a) avec trois inverseurs (b) avec un multiplexeur pour contrôler le signal d'entrée du RO

Afin de contrôler le signal d'entrée de l'oscillateur en anneau, nous avons ajouté un multiplexeur à l'entrée de celui-ci (voir figure 2.2b). De cette façon, notre CUT possède trois modes de fonctionnement :

- Mode RO : Puisque la sortie est rebouclée sur l'entrée, le CUT est un RO. C'est dans ce mode que nous mesurons la période d'oscillation du RO.
- Mode stress : L'entrée du CUT est un signal de stress avec une fréquence et un rapport cyclique (DUty CYcle DUCY) prédéfini. Avec différents signaux de stress nous pouvons étudier l'impact de la fréquence et du rapport cyclique sur le vieillissement. Dans la suite du document nous appellerons ce stress le stress interne pour bien le dissocier du stress externe (température et tension) commun à tous les ROs d'une même puce. Nous avons 21 signaux de stress internes qui sont détaillés dans le tableau 2.1.
- Mode reset : L'entrée du CUT est un signal statique. Sans le mode reset, la transition entre le mode stress et le mode RO est d'une période d'horloge soit de 10ns. Si la période d'oscillation du RO est supérieure à la période de transition alors le signal de stress peut rester piégé dans le circuit pendant le mode RO pour la mesure. Dans ce cas, nous ne mesurons plus la fréquence du signal du RO mais un signal dont la fréquence est un multiple impair du RO. Afin de remédier à ce phénomène, nous utilisons le signal reset, qui est un signal statique, afin d'évacuer le signal de stress avant de passer au mode RO.

Pour limiter l'influence du mode RO et du mode reset sur le vieillissement du RO, il est important de réduire au maximum le temps de mesure et du reset. C'est pourquoi, le CUT est en mode RO pour être mesuré 0.3% du temps, en mode stress 99.6% du temps et en mode reset 0.1% du temps. Pour chaque stress interne nous avons implémenté 27 ROs de différentes architectures. Voici une liste des 21 stress internes appliqués dans notre banc de test :

n° Stress	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
fréquence	DC0	DC1	100Hz		50Hz	11Hz	1MHz		500KHz	111KHz	100MHz				600MHz			1500MHz			
rapport cyclique	0	1	0.1	0.25	0.5	0.33	0.1	0.1	0.25	0.5	0.33	0.1	0.1	0.25	0.5	0.75	0.9	0.5	0.5	0.75	0.5

TABLE 2.1 – Conditions de stress interne

2.3.2 Implémentation des ROs dans le FPGA

Dans un FPGA, quelque soit la fonction des portes logiques, celles-ci sont implémentées par une LUT configurée pour répondre à la fonction requise. De plus, ces LUTs sont reliées à travers des matrices d'interconnexions. La figure 2.3 correspond à l'implémentation dans un FPGA du même RO que de la figure 2.2b, c'est à dire avec un multiplexeur et trois inverseurs. Le signal entre chaque étage du RO passe par une matrice d'interconnexion constituée de Programmable Interconnection Point (PIP). Au travers de cet exemple on constate que, dans un FPGA, le temps de propagation du RO dépend des LUTs mais aussi du circuit d'interconnexion. Afin d'étudier le vieillissement des différentes ressources dans un FPGA, LUT et routage, nous avons conçu deux catégories de RO : les Short Path (SP) et les Long Path (LP). Les SP sont des ROs avec 13 étages de LUTs, multiplexeur inclus, et dont le placement et le routage dans le FPGA sont contraints pour que le RO soit implémenté dans une ou deux slices afin de limiter les ressources de routage. Les LP sont des RO avec deux étages de LUT, multiplexeur inclus, qui sont implémentés dans le FPGA, à l'opposé l'une de l'autre, afin d'avoir un maximum de ressources de routage. Donc, les RO SP sont utilisés pour mesurer le vieillissement des LUT et les RO LP le vieillissement du routage.

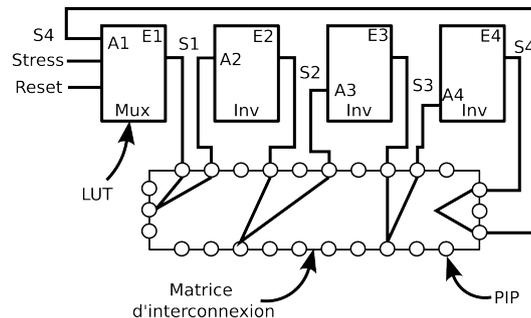


FIGURE 2.3 – Exemple d'implémentation d'un oscillateur en anneau dans un FPGA, avec une LUT configurée en multiplexeur et trois LUT configurées en inverseurs et avec le circuit d'interconnexion entre chaque étage du RO.

Afin d'observer l'impact de la configuration d'une LUT sur le vieillissement, nous avons conçu huit types d'architectures de RO SP. Dans les huit architectures, le multiplexeur en entrée inverse le signal lorsqu'il est en mode RO. Voici une liste descriptive des types d'architectures des SP :

- L1 I : Les LUTs sont configurées en inverseur et une seule entrée de la LUT est utilisée.
- L1 B : Les LUTs sont configurées en buffer et une seule entrée de la LUT est utilisée.
- L2 XI : Les LUTs sont configurées en xor2 et deux entrées de la LUT sont utilisées. La deuxième entrée est imposée à un niveau haut pour que la porte xor se comporte comme un inverseur.
- L2 XB : Les LUTs sont configurées en xor2 et deux entrées de la LUT sont utilisées. La deuxième entrée est imposée à un niveau bas pour que la porte xor se comporte comme un buffer.
- L3 XB : Les LUTs sont configurées en xor3 et trois entrées de la LUT sont utilisées, avec deux niveaux d'entrées imposées pour que la porte se comporte comme un buffer.
- L3 XB I1 : Même architecture que le L3 XB mais on impose que le signal du RO entre par l'input A1 de la LUT.

- L3 XB I5 : Même architecture que le L3 XB mais on impose que le signal du RO entre par l'input A5 de la LUT.
- L5 XB : Les LUTs sont configurées en xor et cinq entrées de la LUT sont utilisées, dont quatre entrées imposées pour que la porte se comporte comme un buffer.

Ces différentes architectures nous permettront d'étudier l'impact sur le vieillissement de la configuration de la LUT (fonction), du nombre d'entrée de LUT utilisées et de l'entrée de LUT utilisée. Aussi, nous pouvons distinguer le RO par son mode de fonctionnement :

- RO de type buffer : Chaque étage (sauf le multiplexeur) se comporte comme un buffer (L1 B, L2 XB, L3 XB, L3 XB I1, L3 XB I5, L5 XB et LP);
- RO de type inverseur : Chaque étage se comporte comme un inverseur (L1 I et L2 XI);

Dans chaque FPGA, nous avons implémenté 567 ROs qui correspondent à : 21 signaux de stress internes \times 9 architectures \times 3 ROs pour chaque architecture. La figure 2.4 présente le plan en étage du FPGA après implémentation des ROs ou chaque point violet est un RO. On constate que les ROs sont rangés en ligne par stress interne et en colonne par architecture. Les ROs ont été répartis sur toute la surface du FPGA afin de limiter une surchauffe locale non contrôlée. Lors du placement et du routage des ROs dans le FPGA, nous avons contraint le logiciel de conception à implémenter chaque RO dans une localisation bien précise en utilisant la contrainte LOC. Cependant, nous avons laissé le logiciel définir les ressources de routages et les entrées de LUT utilisées pour chaque RO, sauf pour l'architecture L3 XB I1 et L3 XB I5 que l'on contraint à utiliser les entrées de LUT A1 et A5 respectivement. Par conséquent, il est important de noter que l'architecture physique après placement et routage entre deux ROs de mêmes architectures logiques (RO I, RO B, etc) peut être différente.

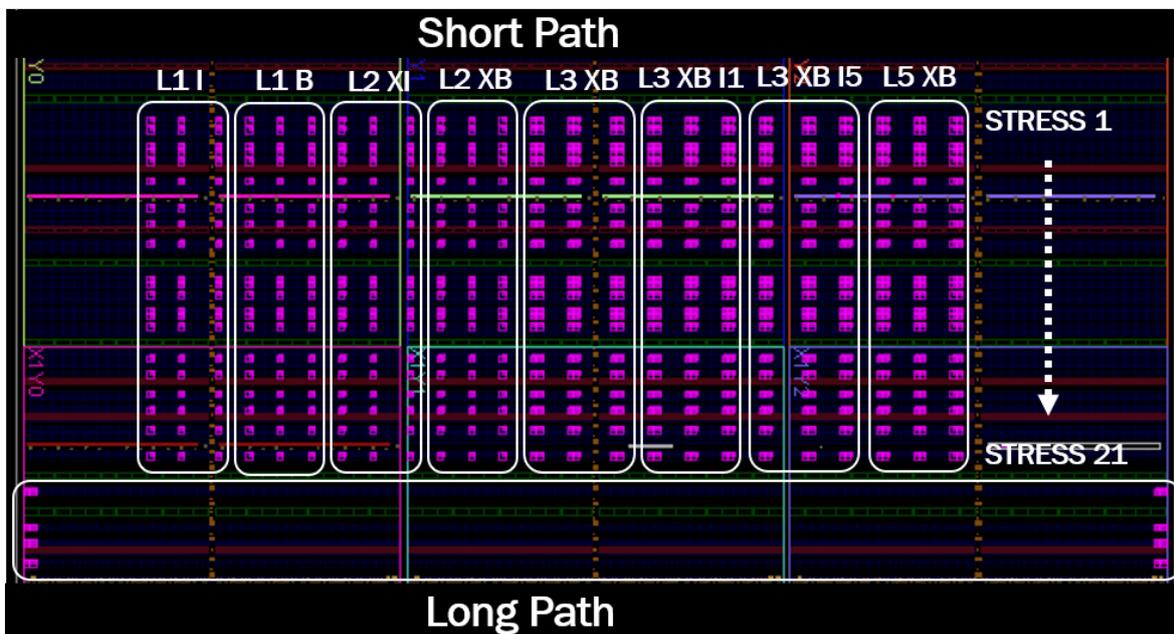


FIGURE 2.4 – Plan d'étage (floorplan) du FPGA après implémentation des 567 ROs

2.4 Circuit de mesure

Dans le FPGA, les circuits mesurés sont les oscillateurs en anneaux et nous allons dans cette section détailler l'architecture du circuit de mesure développé. Le circuit doit mesurer l'évolution du temps de propagation d'un signal dans le RO, ce qui, nous permettra de connaître son état de dégradation.

2.4.1 Mesure de fréquence et du rapport cyclique

Afin de simplifier le banc de test, le circuit de mesure est développé en interne dans le FPGA et il mesure la fréquence du signal de sortie du RO ainsi que son rapport cyclique. Avec la mesure de fréquence nous pouvons facilement connaître la période du RO, et, avec la mesure du rapport cyclique, nous verrons dans la sous-section suivante que nous pourrions extraire le temps de propagation d'un front descendant et d'un front montant dans le RO.

L'architecture du circuit qui mesure la fréquence et le rapport cyclique des ROs se base sur les travaux de Mohammad Naouss [117]. La figure 2.5 présente l'architecture du circuit que nous avons développé dans cette thèse. La figure 2.6 présente le fonctionnement du circuit de mesure avec un chronogramme de ses signaux d'entrée et de sortie. Pour mesurer la fréquence du RO nous utilisons un compteur avec comme signal d'horloge, le signal du RO. Ce compteur compte le nombre de périodes de RO pendant un temps prédéfini. Ce "temps prédéfini" est contrôlé par un signal de référence de temps qui, autorise à compter le nombre de période de RO lorsqu'il est actif, sinon, le compteur est remis à zéro. Le signal de référence de temps est actif pendant une seconde, ce qui veut dire que la dernière valeur comptée correspond au nombre de périodes du RO pendant une seconde, c'est à dire à la fréquence du RO en hertz.

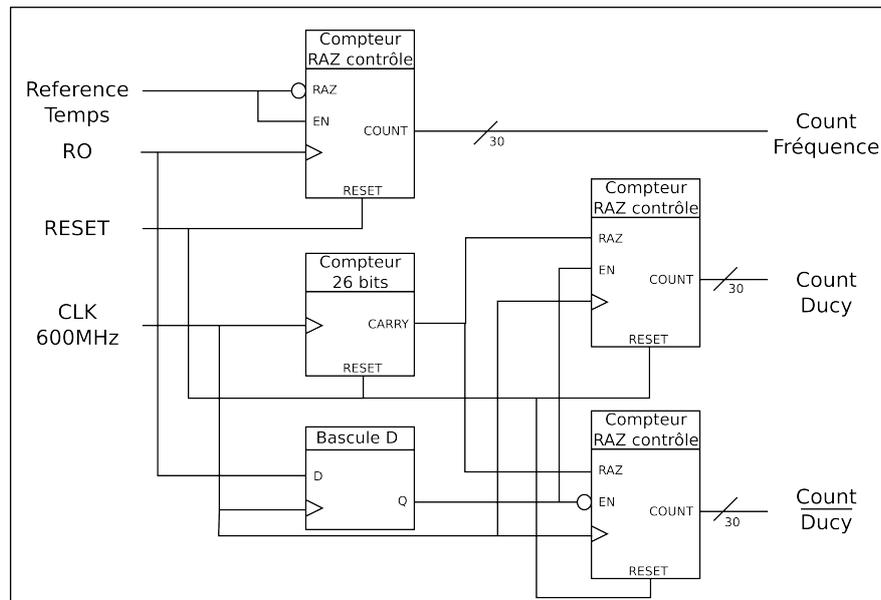


FIGURE 2.5 – Schéma du circuit pour la mesure de fréquence, du rapport cyclique et son complémentaire des ROs

Pour le rapport cyclique, nous utilisons le même type de compteur avec des signaux d'entrées différents. L'horloge du compteur vient d'une PLL (Phase-Locked Loop) qui génère une fréquence de 600MHz à partir d'un signal externe issu d'un oscillateur à quartz. L'autorisation de compter est contrôlée par le signal du RO après synchronisation avec l'horloge de la PLL. Finalement, la retenue d'un compteur 26 bits est utilisée pour la remise à zéro. Le compteur du rapport cyclique compte le nombre de périodes de la PLL lorsque le signal du RO est actif et ce pendant 2^{26} périodes de la PLL. Le rapport cyclique (α) correspond à :

$$\alpha = \frac{\text{Count Ducky}}{2^{26}} \quad (2.1)$$

Sur le chronogramme, le rapport cyclique est de $\alpha = \frac{33555934}{2^{26}} = 50.002\%$. Afin de confirmer l'architecture de mesure du rapport cyclique nous mesurons aussi le complémentaire du rapport cyclique. Pour mesurer le complémentaire du rapport cyclique nous avons implémenté un deuxième

compteur identique dont l'entrée Enable est le signal du RO complémenté. Aussi bien pour la mesure de fréquence que du rapport cyclique, un registre récupère la dernière valeur compté par le compteur lorsque le signal de remise à zéro (RAZ) est à '0' sinon, l'état du registre ne change pas si le signal RAZ est à '1'.

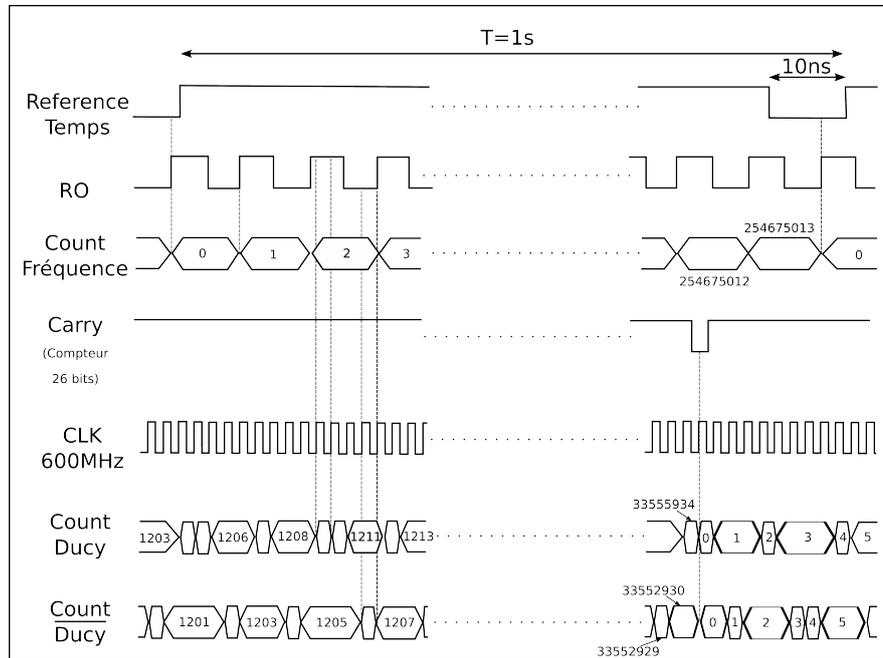


FIGURE 2.6 – Chronogramme des signaux d'entrées et de sorties du circuit de mesure RO avec une fréquence de 254675013Hz et un rapport cyclique de 50.002%.

Un circuit de mesure ne peut mesurer qu'un RO à la fois. En implémentant autant de circuit de mesure que de ROs, soit 567 circuits de mesure, nous pouvons mesurer tous les ROs en parallèle. Cependant, implémenter une telle quantité de circuit de mesure occupe une grande quantité de ressources dans le FPGA et, là aussi, peut générer des surchauffes locales non contrôlées. C'est la raison pour laquelle, comme la figure 2.1 le montre, nous avons implémenté des multiplexeurs entre le banc de ROs et le circuit de mesure. Ainsi, chaque circuit peut mesurer plusieurs ROs en séries. Nous avons implémenté 21 multiplexeurs pour 21 circuits qui mesurent 27 ROs en série.

2.4.2 Temps de propagation d'un front montant et d'un front descendant

Comme nous l'avons vu dans la section 1.5.2, le Bias Temperature Instability est un mécanisme de défaillance qui agit différemment suivant que le transistor est un PMOS (NBTI) ou un NMOS (PBTI). Dans les portes logiques tels l'inverseur ou la porte de transfert, les PMOS et NMOS sont complémentaires. Par exemple, dans un inverseur ou une porte de transfert, un front descendant en sortie est conduit par le NMOS tandis qu'un front montant est conduit par le PMOS. Afin de dissocier l'effet du BTI sur le temps de propagation d'un front descendant et d'un front montant, nous allons, à partir des mesures de fréquence et de rapport cyclique des ROs, extraire le temps de propagation d'un front montant et d'un front descendant dans une LUT. Le temps de propagation d'un front montant et d'un front descendant dans une LUT correspond au temps de changement de l'état de sortie de la LUT lorsqu'un changement est appliqué à l'entrée de la LUT.

Les différentes architectures de RO que nous avons implémenté dans le FPGA sont composées de LUT qui se comportent comme un buffer ou comme un inverseur pendant le mode RO (voir la liste 2.3.2). Donc, pendant la mesure il existe deux catégories de RO : le RO buffer composé de douze buffers et un inverseur, le RO inverseur composé de treize inverseurs. Nous avons représenté ces deux catégories de RO (avec moins d'étages) ainsi que le chronogramme détaillé sur la

figure 2.7. Dans un premier temps, nous considérons le temps de propagation d'un front montant et descendant en sortie d'une LUT configurée en inverseur et configurée en buffer.

RO inverseur

Parce qu'un RO inverseur est constitué seulement de LUTs configurées en inverseur, il y a deux types de fronts différents qui se propagent : front montant et front descendant en sortie d'une LUT inverseur. La figure 2.7 montre que le temps de niveau bas du signal du RO correspond à la somme des temps de propagation de deux fronts montant et d'un front descendant en sortie de la LUT inverseur. Inversement, le temps de niveau haut du signal du RO correspond à la somme des temps de propagation de deux fronts descendant et d'un front montant en sortie de la LUT inverseur. Plus généralement, nous pouvons établir la relation suivante :

$$T_H = \alpha T_{RO} = \left(\frac{n_I + 1}{2}\right) \tau_{f_i} + \left(\frac{n_I - 1}{2}\right) \tau_{r_i} \quad (2.2)$$

$$T_L = (1 - \alpha) T_{RO} = \left(\frac{n_I - 1}{2}\right) \tau_{f_i} + \left(\frac{n_I + 1}{2}\right) \tau_{r_i} \quad (2.3)$$

où n_I est le nombre de LUT inverseur dans le RO, α est le rapport cyclique du RO, T_L et T_H sont les temps de niveau bas et haut du signal du RO, τ_{f_i} et τ_{r_i} sont les temps de propagation d'un front descendant et d'un front montant en sortie d'une LUT inverseur. En résolvant le système linéaire défini par les équations (2.3) et (2.2) on obtient pour l'inverseur :

$$\tau_{f_i} = \left(\frac{1}{2n_I} + \frac{1}{2} - \alpha\right) T_{RO} \quad (2.4)$$

$$\tau_{r_i} = \left(\frac{1}{2n_I} + \alpha - \frac{1}{2}\right) T_{RO} \quad (2.5)$$

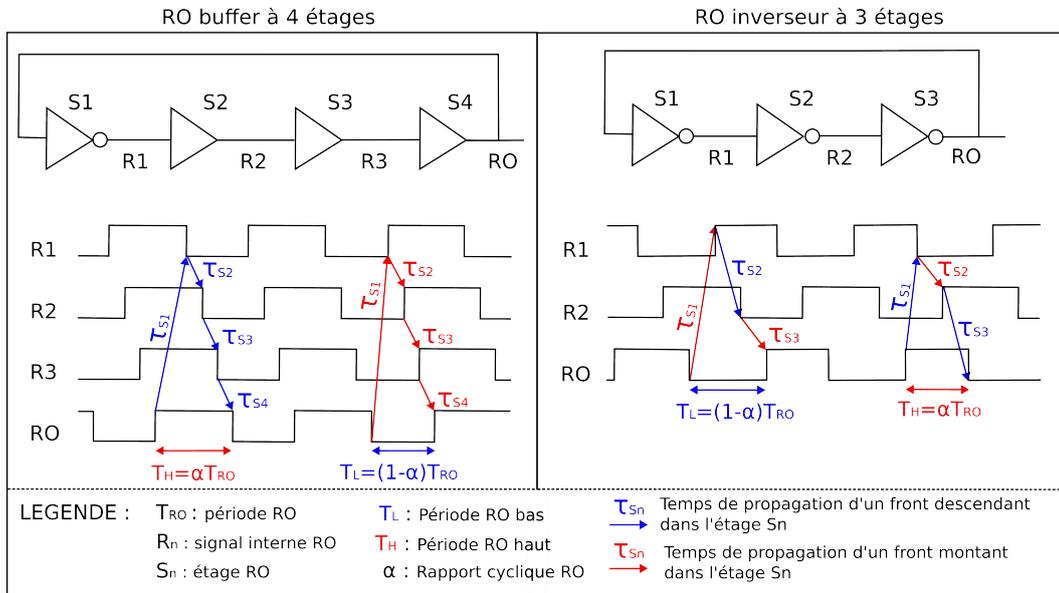


FIGURE 2.7 – Décomposition de la période d'un RO en la somme de temps de propagation de fronts descendants et fronts montants dans les étages

RO buffer

Parce que le RO buffer est composé d'un inverseur et de trois buffer, il y a quatre types de front différents qui se propagent : front montant et descendant en sortie d'une LUT buffer, front

montant et descendant en sortie d'une LUT inverseur. La figure 2.7 montre que le temps de niveau haut du signal RO correspond à la somme du temps de propagation d'un front descendant en sortie de la LUT inverseur et des trois LUT buffers. Inversement, le temps de niveau bas de la période du RO correspond à la somme du temps de propagation d'un front montant en sortie de la LUT inverseur et des trois LUT buffers. Plus généralement, nous pouvons établir la relation suivante :

$$T_H = \alpha T_{RO} = n_B \times \tau_{f_B} + n_I \times \tau_{f_I} \quad (2.6)$$

$$T_L = (1 - \alpha) T_{RO} = n_B \times \tau_{r_B} + n_I \times \tau_{r_I} \quad (2.7)$$

où n_B est le nombre de LUT buffer dans le RO, τ_{f_B} et τ_{r_B} sont les temps de propagation d'un front descendant et d'un front montant en sortie d'une LUT buffer. Les autres paramètres sont les mêmes que ceux définis dans les équations (2.2) et (2.3).

Concernant l'extraction du temps de propagation d'un front descendant et montant dans un buffer, nous constatons qu'il y a quatre inconnus (τ_{f_B} , τ_{r_B} , τ_{f_I} et τ_{r_I}) pour deux équations. Pour ne pas mélanger les mesures d'un RO buffer avec les mesures d'un RO inverseur, nous n'utiliserons pas le résultat obtenu dans les équations (2.4) et (2.5). Nous allons donc étudier comment réduire le nombre d'inconnues afin d'extraire les temps de propagation d'un front descendant et montant dans un buffer.

Sur la figure 2.8 nous avons représenté la propagation d'un signal à l'intérieur d'une LUT d'après l'architecture présentée sur la figure 1.19. La figure 2.8 présente quatre cas de configurations : niveau bas en sortie dans une LUT buffer (figure 2.8a), niveau haut en sortie dans une LUT buffer (figure 2.8b), niveau bas en sortie dans une LUT inverseur (figure 2.8c) et niveau haut en sortie dans une LUT inverseur (figure 2.8d). Les LUT sont des LUT6 avec une seule entrée utilisée (A1), identique dans les quatre cas de configurations. Les lignes où se trouve un niveau bas sont représentées en bleu, celle où se trouve un niveau haut sont représentées en rouge. Le chemin de conduction du bit de configuration à la sortie de la LUT est représenté par une ligne en gras. Finalement, en passant de la figure 2.8a à la figure 2.8b on observe l'effet de la propagation d'un front montant en sortie de la LUT, et en passant de la figure 2.8b à la figure 2.8a on observe l'effet d'un front descendant.

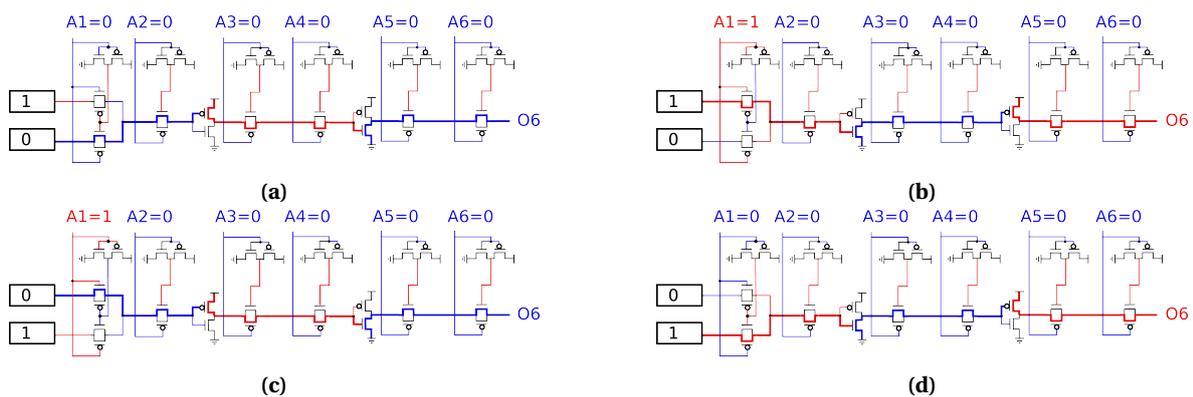


FIGURE 2.8 – Schéma simplifié de l'architecture supposée d'une LUT6 (voir figure 1.19) avec comme seule entrée utilisée l'entrée A1. Identification du chemin de conduction en gras et représentation d'un niveau bas en bleu et d'un niveau haut en rouge dans le cas (a) d'un niveau bas en sortie d'une LUT buffer (b) d'un niveau haut en sortie d'une LUT buffer (c) d'un niveau bas en sortie d'une LUT inverseur (d) d'un niveau haut en sortie d'une LUT inverseur.

On constate que le temps de propagation d'un signal dans une LUT dépend principalement du chemin entre le bit de configuration et la sortie. Pour la transmission d'un front descendant en sortie, on constate que les types de transistors (NMOS et PMOS) transmettant le signal dans

une LUT buffer (figure 2.8a) ou une LUT inverseur (figure 2.8c) sont presque identiques. La seule différence se trouve au niveau de l'inverseur d'entrée qui contrôle les portes de transferts du premier étage. On peut faire le même constat pour un front montant entre une LUT buffer et une LUT inverseur. En négligeant l'inverseur à l'entrée de la LUT, on peut considérer que la configuration de la LUT pour la transmission d'un front descendant en sortie ou d'un front montant en sortie est identique dans le cas d'une LUT buffer ou inverseur. C'est à dire que $\tau_{f_B} = \tau_{f_I} = \tau_f$ et $\tau_{r_B} = \tau_{r_I} = \tau_r$, donc :

$$T_H = \alpha T_{RO} = (n_I + n_B) \tau_f = n \tau_f \quad (2.8)$$

$$T_L = (1 - \alpha) T_{RO} = (n_I + n_B) \tau_r = n \tau_r \tau_{r_I} \quad (2.9)$$

À présent il est possible de résoudre le système d'équations 2.8 et 2.9 et d'exprimer le temps de propagation d'un front descendant et montant tel que :

$$\tau_{f_B} = \frac{\alpha T_{RO}}{n} \quad (2.10)$$

$$\tau_{r_B} = \frac{(1 - \alpha) T_{RO}}{n} \quad (2.11)$$

Comparaison calcul de τ_{fall} et τ_{rise} pour un RO inverseurs et un RO buffers

Nous venons de constater que suivant la composition d'un RO, l'extraction du temps de propagation d'un front montant et descendant en sortie d'une LUT n'a pas la même expression et dépend la mesure du rapport cyclique d'un RO qui peut être bruitée. Afin d'observer des dégradations dans des conditions de stress proches des conditions opérationnelles, nous devons avoir la meilleure précision de mesure possible. Ainsi, nous allons montrer l'effet du bruit de mesure du rapport cyclique du RO sur le calcul du temps de propagation d'un front descendant obtenu pour un RO de type buffer et de type inverseur. Le cas que nous étudions s'étend évidemment au calcul du temps de propagation d'un front montant.

Considérons que la mesure du rapport cyclique de l'oscillateur en anneau possède une erreur ϵ_α qui génère une erreur $\epsilon_{\tau_{f_B}}$ et $\epsilon_{\tau_{r_B}}$ pour le calcul du temps de propagation d'un front descendant fait respectivement avec un RO buffers et un RO inverseurs. En introduisant ces erreurs dans l'équation (2.10) pour un RO buffers :

$$\tau_{f_B} + \epsilon_{\tau_{f_B}} = \frac{(\alpha + \epsilon_\alpha) T_{RO}}{n} \quad (2.12)$$

par identification :

$$\epsilon_{\tau_{f_B}} = \frac{\epsilon_\alpha T_{RO}}{n} \quad (2.13)$$

À présent, en introduisant ces erreurs dans l'équation (2.4) pour un RO inverseurs :

$$\tau_{f_I} + \epsilon_{\tau_{f_I}} = \left(\frac{1}{2n_I} + \frac{1}{2} - (\alpha + \epsilon_\alpha) \right) T_{RO} \quad (2.14)$$

soit :

$$\epsilon_{\tau_{f_I}} = \epsilon_\alpha T_{RO} \quad (2.15)$$

En comparant les équations (2.13) et (2.15) on constate qu'une erreur ϵ_α sur la mesure du rapport cyclique du RO génère une erreur n fois plus importante sur le calcul de τ_{f_I} pour un RO inverseurs comparé au calcul de τ_{f_B} pour un RO buffers. C'est pourquoi, dans la suite de la thèse, les temps de propagation d'un front montant et descendant en sortie d'une LUT ne seront étudié que pour les RO de type buffer. D'ailleurs, c'est pour cela que nous avons implémenté plus d'architecture de RO de type buffer (sept architectures) que de type inverseur (deux architectures).

2.5 Utilisation du GPS comme référence de temps

La section 2.4.1 montre que la fréquence d'oscillation du RO est mesurée en comptant son nombre de périodes pendant une seconde fournie par une référence de temps. Pour avoir une bonne précision de mesure, il est nécessaire que la référence de temps utilisée par le circuit de mesure soit stable. Dans les différentes études de fiabilité utilisant la méthode du RO avec un compteur pour mesurer la fréquence [2, 114, 117, 123], la référence de temps provient d'un oscillateur à quartz souvent déjà présent sur la carte de développement. Cependant, l'oscillateur à quartz est sensible : à court terme à la température et à long terme au vieillissement. L'effet de la température sur l'oscillateur à quartz a pour conséquence de bruyier la mesure de la fréquence du RO et l'effet du vieillissement de l'oscillateur à quartz a pour conséquence de dévier la mesure de la fréquence du RO. Lors de la mesure de vieillissement du RO, en utilisant l'oscillateur à quartz, le résultat obtenu peut être égal à la somme du vieillissement des deux technologies (oscillateur à quartz et oscillateur en anneau). Comme notre étude ne porte que sur le FPGA, il est important d'isoler le vieillissement de celui-ci. Finalement, la stabilité à court terme de l'oscillateur à quartz peut être caractérisé par la gigue cycle-cycle qui correspond à la différence de la période entre deux cycles d'horloge. Le fabricant de l'horloge intégrée sur notre carte de développement a mesuré une gigue cycle-cycle de 5 ps pour une fréquence de 25 MHz soit une stabilité relative de 10^{-7} [133, table 16]. Une référence de temps avec une gigue cycle-cycle plus faible permettra d'atténuer le bruit de mesure de la fréquence d'oscillation du RO.

L'horloge atomique est un système qui fournit une référence de temps qui se base sur la fréquence du rayonnement électromagnétique émis par un électron lors du passage d'un niveau d'énergie à un autre et dont la stabilité relative est de 10^{-16} . Intégrer une horloge atomique dans notre circuit de mesure est trop coûteux et complexe, cependant, on retrouve sur le marché des récepteurs GPS à bas coût qui fournissent une référence de temps provenant du système GPS qui utilise des horloges atomiques. Le récepteur GPS que nous utilisons [134] fournit un signal 1PPS (Pulse Per Second), c'est à dire avec un période de 1 seconde, et dont le rapport cyclique est de 10%. La fabricant du récepteur GPS garantit une référence de temps du signal 1PPS avec une erreur maximale de 10^{-8} seconde. En plus d'avoir une meilleur stabilité que le quartz, l'utilisation d'un tel signal comme référence de temps est une solution pour s'affranchir de l'impact des variations environnementales et du vieillissement. Dans cette partie, nous présenterons la comparaison des résultats de mesure de stabilité de la référence de temps fournie par le récepteur GPS et celle fournie par l'oscillateur à quartz.

Architecture du test et mesures

Afin de mesurer la stabilité du récepteur GPS et de l'oscillateur à quartz nous avons développé dans un FPGA un banc de test que nous présentons sur la figure 2.9. Nous utilisons deux récepteurs GPS qui génèrent chacun un signal (GPS1 et GPS2) avec une fréquence de 1 Hz . Nous utilisons le signal de deux oscillateurs à quartz dont la fréquence est de 100 MHz . Le signal Q1 provient de l'oscillateur à quartz présent sur la même carte de développement que le FPGA dans lequel l'architecture de test (figure 2.9) est implantée, tandis que le signal Q2 provient d'un oscillateur à quartz provenant d'une deuxième carte de développement.

Dans le FPGA, après avoir synchronisé tous les signaux d'entrée sur l'horloge locale (Q1), nous utilisons des compteurs pour générer quatre références de temps de 30 secondes issues des quatre signaux (G1, G2, Q1 et Q2). Finalement, six compteurs, avec une remise à zéro contrôlée sont utilisés pour mesurer le nombre de périodes d'horloge issues d'un oscillateur à quartz comptés pendant les 30 secondes données par les trois autres de référence de temps. La notation de la donnée de sortie M_{ij} correspond au nombre de période de l'horloge i comptées pendant les 30 secondes fournies par la référence de temps j .

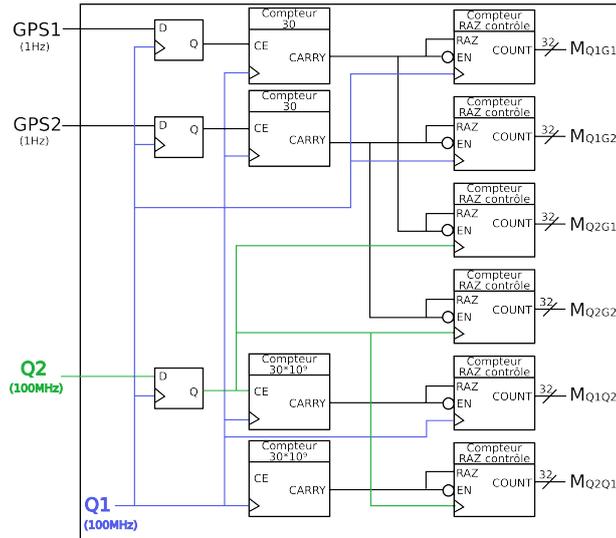


FIGURE 2.9 – Banc de test pour la mesure de stabilité de l’oscillateur à quartz et du récepteur GPS

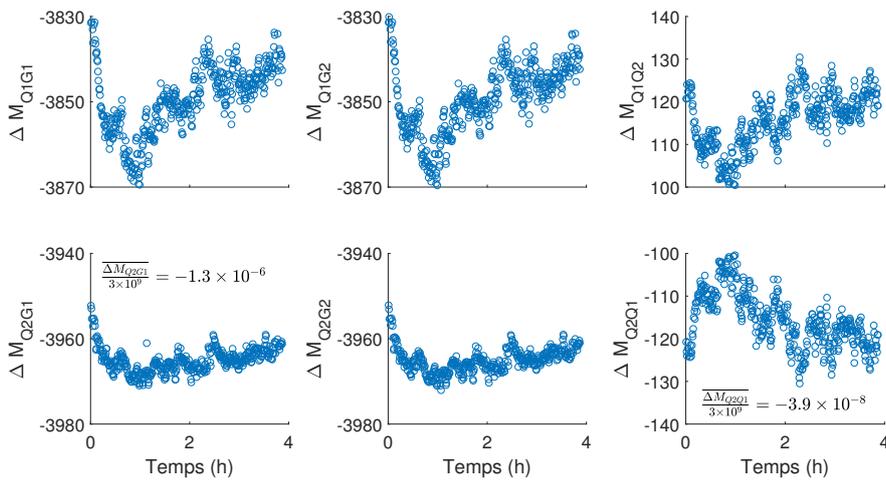


FIGURE 2.10 – Mesures ΔM_{ij} avec deux oscillateurs à quartz et deux récepteurs GPS différents. Nous avons calculé la différence relative moyenne pour ΔM_{Q1G2} et ΔM_{Q2Q1} .

La figure 2.10 représente la différence entre la mesure et la valeur du nombre de périodes théorique :

$$\Delta M_{ij} = M_{ij} - \frac{T_{RAZ}}{T_{CLK}} \quad (2.16)$$

où T_{RAZ} est la période théorique du signal de référence de temps utilisé pour la remise à zéro ($T_{RAZ} = 30s$), T_{CLK} est la période théorique de l’horloge de fréquence 100MHz ($T_{CLK} = 10^{-8}s$). On constate sur la figure 2.10 une différence négative pour les mesures dont la référence de temps est générée à partir du récepteur GPS (M_{Q1G1} , M_{Q1G2} , M_{Q2G1} et M_{Q2G2}), ce qui veut dire que la référence de temps fournie par les oscillateurs à quartz est plus grande que celle générée par le récepteur GPS. De la même manière, les mesures de M_{Q1Q2} et M_{Q2Q1} montrent que la référence de temps généré par l’oscillateur à quartz Q2 est plus grande que celle de l’oscillateur à quartz Q1. On remarque également des variations sur la mesure qui sont probablement causées par des perturbations de température. On observe une ressemblance plus importante entre les mesures effectuées avec un même oscillateur à quartz et deux récepteurs GPS comparé aux mesures effectuées avec deux oscillateurs à quartz et un même GPS. De plus, le bruit est de même amplitude pour les mesures de M_{QG} que pour les mesures de M_{QQ} . Ces deux dernières remarques nous donnent une

première idée quant au fait que le signal du GPS soit plus stable que celui de l'oscillateur à quartz. Nous allons confirmer cette observation avec une analyse statistique des mesures.

Étude statistique

Dans cette partie, nous allons utiliser les données de M_{Q1G1} , M_{Q1G2} , M_{Q2G1} , et M_{Q2G2} pour comparer la stabilité du récepteur GPS à celle de l'oscillateur à quartz. Les deux graphiques de gauche sur la figure 2.11 montre les histogrammes de l'écart des mesures faites avec un même oscillateur à quartz et deux récepteur GPS différents ($M_{Q1G1} - M_{Q1G2}$ et $M_{Q2G1} - M_{Q2G2}$). Les deux graphiques de droite sur la figure 2.11 montrent les histogrammes de l'écart des mesures faites avec un même récepteur GPS et deux oscillateurs à quartz différents ($M_{Q1G1} - M_{Q2G1}$ et $M_{Q1G2} - M_{Q2G2}$). On constate que la variance des différences avec un même oscillateur à quartz et deux récepteur GPS différents est plus de 1000 fois plus faible que la variance des différences avec un même récepteur GPS et deux oscillateurs à quartz différents. Ce résultat est un deuxième indicateur confirmant la meilleure stabilité du signal fournit par le récepteur GPS.

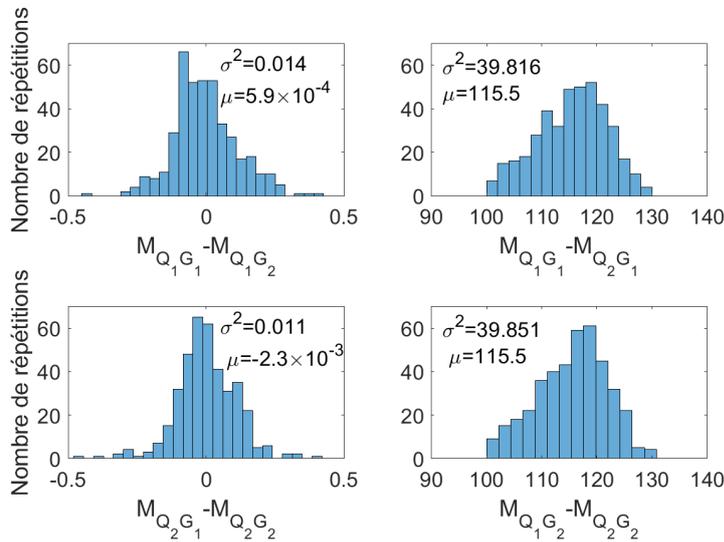


FIGURE 2.11 – Histogramme de la différence entre les mesures faites avec un deux récepteurs GPS différents et un même quartz (gauche) et les mesures faites avec deux oscillateurs à quartz différents et un même récepteur GPS (droite)

Pour que la suite soit plus lisible, nous remplacerons l'indice i par Q_p et l'indice j par G_n . Nous considérons que les périodes du GPS et de l'oscillateur à quartz à un instant donné correspondent à la somme d'une période moyenne et d'une erreur :

$$T_{Q_n}(t) = \overline{T_{Q_n}} (1 + \epsilon_{Q_n}(t)) \quad (2.17)$$

on peut alors écrire :

$$M_{Q_pG_n}(t) = \frac{\overline{T_{G_n}} (1 + \epsilon_{G_n}(t))}{\overline{T_{Q_p}} (1 + \epsilon_{Q_p}(t))} \quad (2.18)$$

où $\overline{T_{Q_p}}$ est la périodes moyennes de l'oscillateur à quartz p et $\overline{T_{G_n}}$ est la période de mesure obtenue avec le GPS (30s), $\epsilon_{Q_p}(t)$ et $\epsilon_{G_n}(t)$ sont deux variables aléatoires d'espérances nulles représentant l'erreur relative de l'oscillateur à quartz p et du GPS n . En faisant un développement limité du premier ordre de l'équation 2.18 et en négligeant les termes du second ordre, on peut écrire :

$$M_{Q_pG_n}(t) = \overline{M_{Q_pG_n}} (1 - \epsilon_{Q_p}(t) + \epsilon_{G_n}(t)) \quad (2.19)$$

où $\overline{M_{Q_pG_n}} = \frac{\overline{T_{G_n}}}{\overline{T_{Q_p}}}$ et correspond à la moyenne des mesures.

En considérant indépendantes les sources d'erreurs aléatoires de deux oscillateurs à quartz, de deux récepteurs GPS et d'un oscillateur à quartz avec un récepteur GPS :

$$\sigma^2 [M_{Q_1G_1} - M_{Q_2G_1}] = \sigma^2 \left[\overline{M_{Q_1G_1}} (1 - \epsilon_{Q_1} + \epsilon_{G_1}) - \overline{M_{Q_2G_1}} (1 - \epsilon_{Q_2} + \epsilon_{G_1}) \right] \quad (2.20)$$

$$\sigma^2 [M_{Q_1G_1} - M_{Q_2G_1}] = \sigma^2 \left[\overline{M_{Q_1G_1}} - \overline{M_{Q_2G_1}} - \overline{M_{Q_1G_1}} \epsilon_{Q_1} + \overline{M_{Q_2G_1}} \epsilon_{Q_2} + \left(\overline{M_{Q_1G_1}} - \overline{M_{Q_2G_1}} \right) \epsilon_{G_1} \right] \quad (2.21)$$

$$\sigma^2 [M_{Q_1G_1} - M_{Q_2G_1}] = \overline{M_{Q_1G_1}}^2 \overline{\epsilon_{Q_1}^2} + \overline{M_{Q_2G_1}}^2 \overline{\epsilon_{Q_2}^2} + \left(\overline{M_{Q_1G_1}} - \overline{M_{Q_2G_1}} \right)^2 \overline{\epsilon_{G_1}^2} \quad (2.22)$$

où $\overline{\epsilon_{Q_p}^2}$ et $\overline{\epsilon_{G_n}^2}$ sont les variances de l'erreur des quartz et des récepteurs et représentent leurs stabilités respectives. L'équation (2.22) correspond à la variance de la différence entre deux mesures faites avec deux quartz différents (Q_1 et Q_2) et un même récepteur GPS (G_1). Nous pouvons faire le même développement pour la variance de la différence entre deux mesures faites avec un quartz (Q_1) et deux récepteurs GPS différents (G_1 et G_2), ainsi nous obtenons le système suivant :

$$\begin{cases} \sigma^2 [M_{Q_1G_1} - M_{Q_2G_1}] = \overline{M_{Q_1G_1}}^2 \overline{\epsilon_{Q_1}^2} + \overline{M_{Q_2G_1}}^2 \overline{\epsilon_{Q_2}^2} + \left(\overline{M_{Q_1G_1}} - \overline{M_{Q_2G_1}} \right)^2 \overline{\epsilon_{G_1}^2} \\ \sigma^2 [M_{Q_1G_1} - M_{Q_1G_2}] = \overline{M_{Q_1G_1}}^2 \overline{\epsilon_{G_1}^2} + \overline{M_{Q_1G_2}}^2 \overline{\epsilon_{G_2}^2} + \left(\overline{M_{Q_1G_2}} - \overline{M_{Q_1G_1}} \right)^2 \overline{\epsilon_{Q_1}^2} \end{cases} \quad (2.23)$$

On suppose l'erreur des deux récepteurs GPS identique ainsi que l'erreur des deux oscillateurs à quartz identique. On peut donc écrire que : $\overline{\epsilon_{G_1}^2} = \overline{\epsilon_{G_2}^2} = \overline{\epsilon_G^2}$ et $\overline{\epsilon_{Q_1}^2} = \overline{\epsilon_{Q_2}^2} = \overline{\epsilon_Q^2}$. Ainsi nous pouvons simplifier le système (2.23) et l'écrire sous forme matricielle en exprimant la variance des erreurs en fonction des variances des différences de mesures simultanées :

$$\begin{bmatrix} \overline{\epsilon_Q^2} \\ \overline{\epsilon_G^2} \end{bmatrix} = \begin{bmatrix} \left(\overline{M_{Q_1G_1}}^2 + \overline{M_{Q_2G_1}}^2 \right) & \left(\overline{M_{Q_1G_1}} - \overline{M_{Q_2G_1}} \right)^2 \\ \left(\overline{M_{Q_1G_2}} - \overline{M_{Q_1G_1}} \right)^2 & \left(\overline{M_{Q_1G_1}}^2 + \overline{M_{Q_1G_2}}^2 \right) \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2 [M_{Q_1G_1} - M_{Q_2G_1}] \\ \sigma^2 [M_{Q_1G_1} - M_{Q_1G_2}] \end{bmatrix} \quad (2.24)$$

À partir des données présentées sur la figure 2.10 et sur la figure 2.11, nous calculons la variance de l'erreur de l'oscillateur à quartz et du récepteur GPS.

$$\begin{bmatrix} \overline{\epsilon_Q^2} \\ \overline{\epsilon_G^2} \end{bmatrix} = \begin{bmatrix} 3.1 \times 10^7 & 1.3 \times 10^4 \\ 3.5 \times 10^{-7} & 3.0 \times 10^7 \end{bmatrix}^{-1} \begin{bmatrix} 39.816 \\ 0.014 \end{bmatrix} = \begin{bmatrix} 1.3 \times 10^{-6} \\ 4.7 \times 10^{-10} \end{bmatrix} \quad (2.25)$$

on obtient alors :

$$\begin{bmatrix} \sigma(\epsilon_Q) \\ \sigma(\epsilon_G) \end{bmatrix} = \begin{bmatrix} 1.1 \times 10^{-3} \\ 2.2 \times 10^{-5} \end{bmatrix} \quad (2.26)$$

Le résultat obtenu montre que l'écart type de l'erreur relative du récepteur GPS est inférieure à celle de l'oscillateur à quartz, prouvant ainsi que la stabilité du récepteur GPS est meilleure. Cependant, nous obtenons un résultat supérieur à celui attendu et annoncé au début de cette même section ($\overline{\epsilon_Q} \approx 10^{-7}$ et $\overline{\epsilon_G} \approx 10^{-8}$). Cette différence est liée à l'approximation du développement limité du premier ordre (équation (2.19)) et à l'utilisation de la moyenne et de la variance des mesures rajoutant encore de l'incertitude. En renouvelant la même opération sur le système, mais avec $M_{Q_1G_2} - M_{Q_2G_2}$ et $M_{Q_2G_1} - M_{Q_2G_2}$, on retrouve des résultats significativement identiques $\sigma(\epsilon_Q) = 1.1 \times 10^{-3}$ et $\sigma(\epsilon_G) = 9.7 \times 10^{-5}$.

En conclusion de cette partie, nous venons de mettre en évidence que le récepteur GPS possède une meilleure stabilité que l'oscillateur à quartz. Cette meilleure stabilité contribuera à l'amélioration de la précision de mesure de fréquence des RO. En plus d'avoir une meilleure stabilité à court terme, le récepteur GPS offre une stabilité sur du long terme. En effet, la référence de temps générée par le système GPS est maintenue par des experts tandis que l'oscillateur à quartz est un

composant sensible, non seulement à la température, mais aussi au vieillissement. L'utilisation du récepteur GPS nous permettra d'obtenir des mesures de fréquence de RO non biaisées par le vieillissement de la référence de temps.

2.6 Régulation de la température et tension interne du FPGA

L'oscillateur en anneau est le circuit que nous utilisons pour mesurer les dégradations dans un FPGA, le vieillissement ayant pour effet de diminuer la fréquence d'oscillation du RO. Cependant, le RO est bien plus sensible à la température et à la tension d'alimentation. Le principal influenceur de la température et de la tension interne du FPGA est la variation de la consommation électrique dans le FPGA qui provoque des variations de la température de l'ordre de quelques degrés ainsi que des variations de la tension interne du FPGA de quelques millivolts.

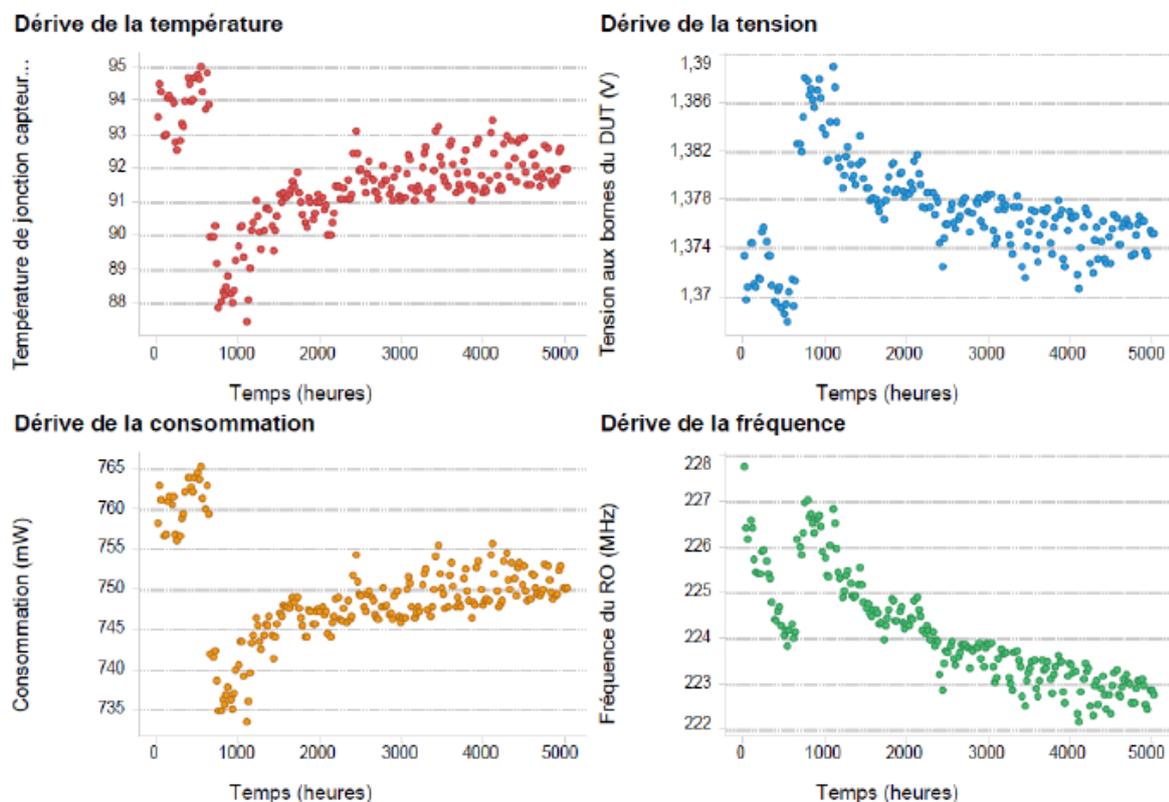


FIGURE 2.12 – Évolution de la température, de la tension et de la consommation interne d'un FPGA Artix ($V_{nom} = 1V$) et impact sur la fréquence d'oscillation d'un RO [2, fig.90].

Afin d'illustrer et de quantifier l'effet de la température et de la tension sur la fréquence d'oscillation d'un RO, la figure 2.12 présente l'évolution pendant 5000 heures de la température, de la tension et de la consommation interne du FPGA ainsi que la dérive de la fréquence d'oscillation d'un RO. On constate à 1000 heures une diminution soudaine de la consommation du FPGA de l'ordre de 25mW ce qui a pour effet de diminuer la température de 6°C et d'augmenter la tension de 16mV (soit 1.6% de V_{nom}). Cette variation de température et de tension a pour effet d'augmenter soudainement la fréquence relative du RO de 1.3% alors que la dérive de fréquence relative causée par 4000 heures de vieillissement (entre 1000 heures et 5000 heures) est seulement de 1.8%. De plus, entre 1000 heures et 5000 heures de vieillissement, on observe que la température et la tension interne ne sont pas stables mais dérivent dans le temps. Par conséquent, la dérive de la fréquence d'oscillation du RO observée entre 1000 heures et 5000 heures est le résultat de la somme de l'effet du vieillissement avec l'effet de la dérive de température et de tension.

La figure 2.12 confirme l'intérêt de réguler la température et la tension interne du FPGA afin que la dérive de fréquence du RO que nous mesurons soit seulement due au vieillissement. Dans cette section nous présentons les systèmes de régulation mis en place pour contrôler la température et la tension interne du FPGA. Dans les systèmes de régulation nous utilisons les mesures de température et de tension interne du FPGA faites par le SYSMON. Finalement, pour mettre en évidence les bonnes performances de nos systèmes de régulation, nous présentons les résultats d'une expérience dans laquelle nous faisons varier la température et la tension interne du FPGA.

2.6.1 SYStem MONItoring (SYSMON)

Comme le montre la figure 2.1, le SYSMON est un système intégré dans le FPGA. Le SYSMON est composé d'un convertisseur analogique-numérique (ADC) que l'on retrouve sur le diagramme de fonctionnement général de la figure 2.13. Un multiplexeur en amont de l'ADC sélectionne le signal analogique à convertir qui peut être soit un signal externe au FPGA soit un signal interne provenant des capteurs de température et tension intégrés dans le FPGA. Lorsque le signal analogique est converti, la donnée numérique est stockée dans un registre. Le Dynamic Reconfiguration Port (DRP) est un module qui fait tampon entre les registres du SYSMON et l'utilisateur. Des protocoles de communications sont utilisés pour recevoir ou envoyer les données depuis la Programmable Logic (PL) vers le SYSMON.

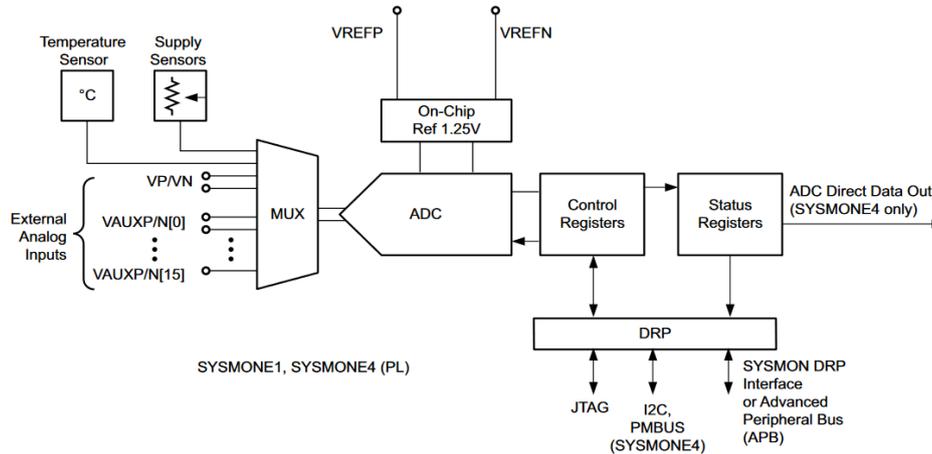


FIGURE 2.13 – Diagramme du fonctionnement global du SYSMON [135]

Dans notre banc de test, nous avons utilisé le SYSMON pour connaître la température et la tension interne du FPGA afin de les réguler avec précision. Pour implémenter le SYSMON dans notre architecture FPGA du banc de test, nous avons utilisé l'Intellectual Property (IP) System Management Wizard [136]. En sortie de l'IP System Management Wizard nous récupérons la mesure de température et de tension sous forme de vecteur de 16 bits. Aussi, nous configurons l'IP pour que l'ADC du SYSMON convertisse et stocke dans ses registres les mesures de température et de tension en boucle.

La fonction de transfert pour obtenir la température en degrés Celsius à partir de la donnée binaire de 16 bits stockée dans le registre du SYSMON est :

$$T(^{\circ}\text{C}) = \frac{do_out \times 507.59}{2^{16}} - 279.42 \quad (2.27)$$

et la fonction de transfert pour obtenir la tension en Volt :

$$V(\text{V}) = \frac{do_out}{2^{16}} \times 3 \quad (2.28)$$

Le quantum, c'est à dire le pas de conversion, pour la température est de $\frac{507.59}{2^{16}} = 7.7 \times 10^{-3} \text{ }^\circ\text{C}$ et pour la tension de $\frac{3}{2^{16}} = 4.6 \times 10^{-5} \text{ V}$. La fréquence d'échantillonnage de l'ADC est de 100MHz, nous obtenons un grand nombre de mesures de température et de tension par seconde. Afin d'éliminer du bruit de la mesure de température et de la tension, nous faisons calculer par la PL une moyenne glissée sur 2^{13} mesures. Les données en sorties des modules de moyennes glissées sont envoyées vers les systèmes de régulation et vers l'ordinateur pour connaître la température et la tension interne du FPGA simultanément à la mesure de fréquence des RO.

2.6.2 Régulation en température

Le principe général de la régulation en température est présenté sur la figure 2.1. La mesure de température obtenue par le SYSMON est moyennée et envoyée dans un régulateur hébergé dans la PS et qui commande, par le biais d'un générateur PWM implémenté en VHDL dans la PL, l'alimentation d'un module Peltier. Dans cette partie, nous allons détailler les étapes de conception du régulateur en température.

Le module Peltier

La figure 2.14a décrit le fonctionnement d'un module Peltier. Un module Peltier est constitué de semi-conducteurs, dopés n et p, reliés par une partie métallique, généralement du cuivre. Les jonctions sont protégées par une couche de céramique qui est un isolant électrique mais un conducteur thermique. Les jonctions sont assemblées électriquement en série et thermiquement

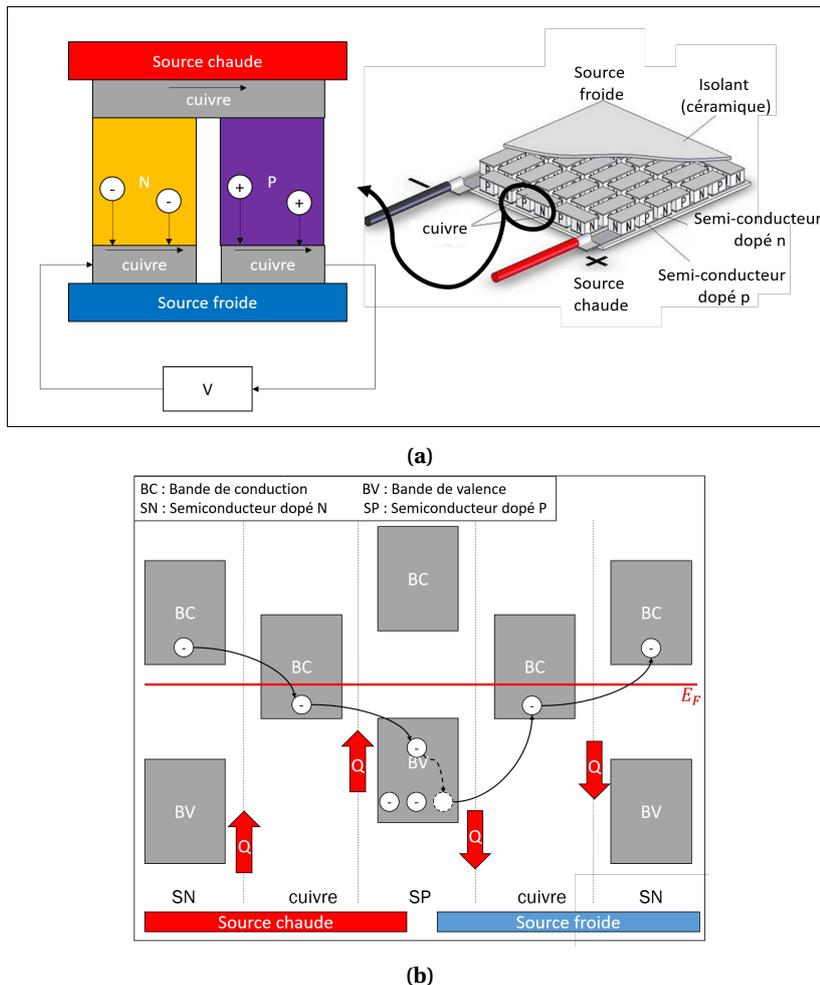


FIGURE 2.14 – (a) Schéma d'un module Peltier (à droite) et zoom sur une paire de semi-conducteur (à gauche) [137] (b) Explication du flux de chaleur dans le Peltier avec l'approche de la théorie des bandes

en parallèle. Le module Peltier possède deux modes de fonctionnement : machine thermique et générateur électrique. Dans le mode machine thermique, un courant injecté dans le module Peltier produit un flux de chaleur entre les deux faces. En mode générateur électrique, une différence de température appliquée sur les deux faces du module génère une tension électrique. Nous utiliserons le module Peltier en mode machine thermique pour contrôler la température du FPGA.

En mode machine thermique, la conduction électrique se fait dans trois matériaux différents : Cuivre, semi-conducteur dopé n (SN) et semi-conducteur dopé p (SP). Lorsque des matériaux présentant des structures de bandes différentes sont assemblés entre eux, l'équilibre énergétique de la structure est assuré par l'égalisation des différents niveaux de Fermi. La figure 2.14b présente le trajet suivi par un électron dans le Peltier à partir de la théorie des bandes. L'électron perd de l'énergie électrique quand il passe de SN vers le cuivre. De la même manière, lorsque l'électron se recombine avec un trou dans la bande de valence du SP, celui-ci perd de l'énergie électrique. L'énergie électrique perdue par l'électron est transformée en énergie thermique. En continuant le voyage de l'électron, on constate qu'il gagne de l'énergie électrique entre SP-Cuivre et Cuivre-SN. Cette énergie électrique que l'électron gagne provient de l'énergie thermique absorbée dans son environnement. C'est ainsi que le flux thermique est généré dans le Peltier.

Identification du modèle

Avant de développer un circuit de régulation, il faut connaître le modèle du système asservi. Pour cela nous avons développé deux modèles : un modèle de connaissance à partir des équations thermoélectriques et, un modèle comportemental identifié à partir de mesures réalisées sur le système.

Modèle de connaissance

La figure 2.15 montre un schéma de principe de notre système thermique. Le FPGA est en contact avec le module Peltier qui est lui-même en contact avec un radiateur thermique, le tout étant ventilé. Sur la figure 2.15, nous présentons un Peltier qui refroidit le FPGA. Cependant, en changeant la polarité d'alimentation du Peltier, il est aussi possible de réchauffer le FPGA.

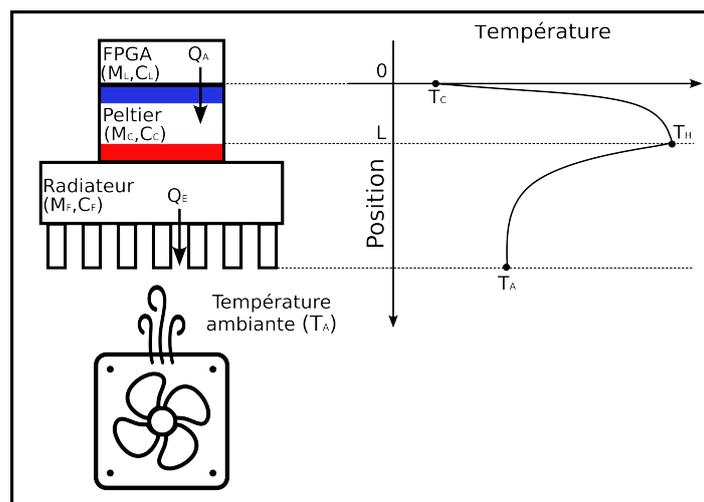


FIGURE 2.15 – Montage du module Peltier sur le FPGA et principe de fonctionnement du système thermique

Le développement du modèle de connaissance est construit à partir des travaux réalisés dans [138]. On retrouve quatre principaux phénomènes thermiques dans un module Peltier :

- Effet Peltier;

- Conduction thermique;
- Effet Joule;
- Effet Thomson;

L'effet Peltier produit un flux de chaleur lors de la circulation d'un courant entre deux matériaux conducteurs différents. Les trois autres effets sont non désirés mais induits par la nature du module. Par conduction thermique, un flux de chaleur du chaud vers le froid est présent. Le courant circulant dans le module provoque par effet Joule une source de chaleur interne. L'effet Thomson est mis en évidence lorsqu'il y a présence d'un gradient de température et un courant électrique, par conséquent il se produit une génération ou une absorption de chaleur. L'effet Thomson est un phénomène qui est négligeable par rapport aux autres phénomènes présents dans un module Peltier.

À partir de la figure 2.15, nous écrivons l'équilibre thermique en trois points du module Peltier : côté froid, côté chaud, intérieur module. En considérant la distribution de température comme homogène et le flux de chaleur Q_a provenant du FPGA comme étant uniforme, l'équilibre thermique côté froid s'exprime :

$$(M_L C_L + M_C C_C) \frac{dT_C}{dt} = Q_a - Q_k - I\alpha T_C \quad (2.29)$$

où :

- La conduction thermique du côté froid selon la loi de Fourier :

$$Q_k = -kS \left. \frac{\partial T(x, t)}{\partial x} \right|_{x=0} \quad (2.30)$$

- $T(x, t)$: La distribution de température dans le module Peltier;
- k : Conductivité thermique du Peltier;
- S : Surface transversale;
- L'effet Peltier : $I\alpha T_C$
 - I : Le courant circulant dans le Peltier;
 - α : Coefficient Peltier;
 - T_C : La température côté froid du Peltier; Notons que $T_C = T(0, t)$.

De la même manière, nous écrivons l'équilibre thermique côté chaud :

$$(M_F C_F + M_H C_H) \frac{dT_H}{dt} = Q_0 - hS_F (T_H - T_C) - I\alpha T_H \quad (2.31)$$

où :

- La conduction thermique du côté chaud :

$$Q_0 = -kS \left. \frac{\partial T(x, t)}{\partial x} \right|_{x=L} \quad (2.32)$$

- Convection forcée du ventilateur : $hS_F (T_H - T_C)$
 - h : Coefficient de convection;
 - S_F : Surface du radiateur thermique;
 - T_H : La température côté chaud du Peltier; Notons que $T_H = T(L, t)$.

Finalement, nous écrivons l'équilibre thermique dans le module thermoélectrique :

$$C_Y \frac{\partial T(x, t)}{\partial t} = k \frac{\partial^2 T(x, t)}{\partial x^2} - \frac{\tau}{S} I \frac{\partial T(x, t)}{\partial x} + \frac{\rho}{S^2} I^2 \quad (2.33)$$

où :

- La conduction thermique dans le Peltier : $k \frac{\partial^2 T(x,t)}{\partial x^2}$;
- L'effet Thomson : $\frac{1}{S} I \frac{\partial T(x,t)}{\partial x}$, avec τ le coefficient de Thomson ;
- L'effet Joule : $\frac{\rho}{S^2} I^2$, avec ρ la résistivité électrique du module ;

Dans [138], l'auteur développe un modèle théorique qui exprime la fonction de transfert de la température côté froid du Peltier (T_C) en fonction du courant (I). Pour aboutir au modèle théorique l'auteur réalise plusieurs étapes de simplifications. La première étape consiste à linéariser les équations 2.29, 2.31 et 2.33, autour d'un point d'équilibre. Ensuite, les équations sont exprimées après un développement du coefficient de Peltier en série de Taylor et en éliminant les ordres supérieurs et les termes d'équilibres. En appliquant la transformée de Laplace et en considérant le flux de chaleur Q_a ainsi que la température ambiante T_a constante une première fonction de transfert est obtenue. Cette fonction de transfert et ensuite simplifiée en négligeant l'effet Thomson dans le module Peltier ainsi, la fonction de transfert s'exprime :

$$G_I(s) = \frac{\tilde{T}_C(s)}{\tilde{I}(s)} = -K \frac{\frac{s}{z} + 1}{\left[\frac{s}{p_1} + 1\right] \left[\frac{s}{p_2} + 1\right]} \quad (2.34)$$

avec,

$$z = \frac{Sk\alpha(\bar{T}_C - \bar{T}_H) + \alpha\bar{T}_C L h S_F + \frac{2\rho\bar{I}}{S} L^2 h S_F - 2\rho\bar{I}kL}{\left[\frac{1}{2}S\alpha L^2 C\gamma + L\alpha(M_F C_F + M_H C_H)\right] \bar{T}_C + \frac{\rho L^2}{S} (M_F C_F + M_C C_C) \bar{I}} \quad (2.35)$$

$$K = \frac{Sk\alpha(\bar{T}_C - \bar{T}_H) + \alpha\bar{T}_C L h S_F + \frac{2\rho\bar{I}}{S} - 2\rho\bar{I}kL}{ap_1 p_2} \quad (2.36)$$

$$a = S(M_L C_L + M_C C_C) \frac{C\gamma L^2}{2} + (M_L C_L + M_C C_C)(M_F C_F + M_H C_H)L + S(M_F C_F + M_H C_H) \frac{C\gamma L^2}{2} \quad (2.37)$$

$$b = Ak(M_L C_L + M_C C_C) + (M_L C_L + M_C C_C)hS_F L + Sk\left(hS_F \frac{C\gamma L^2}{2k} + (M_F C_F + M_H C_H)\right) + S^2 k C\gamma L \quad (2.38)$$

$$c = SkhS_F \quad (2.39)$$

$$\Delta_D(s) = b^2 - 4ac \quad p_1 = \frac{-b + \sqrt{\Delta_D(s)}}{2a} \quad p_2 = \frac{-b - \sqrt{\Delta_D(s)}}{2a} \quad (2.40)$$

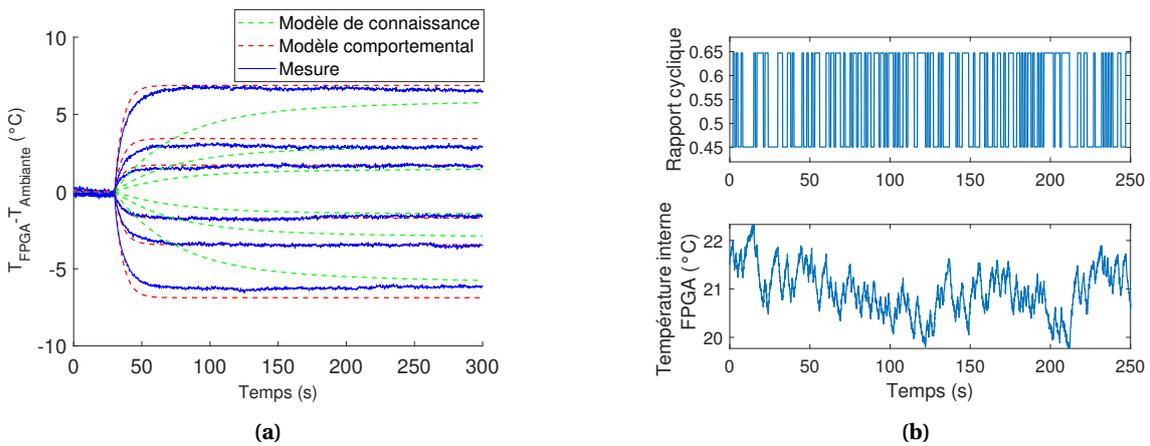


FIGURE 2.16 – (a) Comparaison des modèles de connaissance et de comportement avec les mesures pour des échelons du rapport cyclique : ± 0.1 , ± 0.2 , ± 0.4 (b) Évolution de la température interne du FPGA (en bas) pour l'application d'une Séquence Binaire Pseudo Aléatoire (en haut)

Pour l'application numérique, les paramètres du module Peltier sont extraits à partir des courbes de sa datasheet [139] et en utilisant la méthode par les moyennes [140]. Les résultats de cette extraction sont : $\alpha = 0.02848 \text{ V.K}^{-1}$, $\rho = 0,00058 \text{ } \Omega.m^2$ et $k = 1,91 \text{ W.m}^{-1}.\text{K}^{-1}$. Les résultats extraits

sont proches de ceux utilisés dans [138]. Les autres paramètres¹ sont mesurés ou extraits depuis leurs datasheets respectifs. La figure 2.16a compare le modèle de connaissance avec les mesures. Nous voyons que l'amplitude du modèle de connaissance tend vers la valeur mesurée, cependant, la dynamique du modèle de connaissance est plus lente que la mesure. Pour améliorer ce modèle comportemental il serait intéressant de modéliser la distribution de la température dans le radiateur thermique plutôt que de considérer la température homogène.

Modèle comportemental

Afin d'avoir un modèle plus cohérent avec le comportement de notre système thermique, nous avons réalisé une identification du modèle par le comportement expérimental. La Séquence Binaire Pseudo Aléatoire (SBPA ou PRBS en anglais) est un signal d'excitation reconnu pour exciter un système dans tous les spectres de fréquences susceptibles de contenir des constantes de temps. À partir des mesures obtenues lors d'une SBPA, il est possible de procéder à une identification de la fonction de transfert. La figure 2.16b présente la séquence de l'entrée stimulant le module Peltier ainsi que l'évolution de la température au même instant. À partir des mesures avec la SBPA, nous extrayons un modèle du premier ordre :

$$G(z^{-1}) = \frac{-0.0012z^{-1}}{1 - 0.9829z^{-1}} \quad (2.41)$$

Le signal étant échantillonné, le modèle identifié s'écrit en temps discret avec l'utilisation de la transformée en Z. La période d'échantillonnage est de 0.1s. Nous voyons, sur la figure 2.16a, une meilleure cohérence entre les mesures et le modèle comportemental, ce qui justifie l'utilisation du modèle comportemental pour la suite. Dans la suite de cette section, nous verrons que le modèle comportemental (2.41) qui est du premier ordre permet d'avoir des performances de régulation largement suffisante, donc, l'utilisation d'une fonction de transfert avec un ordre supérieur n'est pas nécessaire.

Loi de commande

À partir du modèle comportemental, une loi de commande RST est mise en place. Un régulateur RST est un organe de contrôle à trois degrés de liberté. Ce type de régulateur est couramment utilisé pour des systèmes à temps discret [141]. Nous avons utilisé le logiciel MIRCOD [142] pour calculer les paramètres du régulateur. Sur la figure 2.1, le régulateur RST est hébergé dans la PS. La sortie du régulateur RST est une valeur de rapport cyclique envoyée vers un module PWM (Pulse Width Modulation) et qui génère un signal PWM. Le signal PWM contrôle l'alimentation du module Peltier par le biais d'un pont en H et d'un filtre RLC.

Filtre RLC

L'étude réalisée dans [143] montre que l'utilisation du PWM génère des cycles thermiques qui fatiguent les jonctions de soudure, ce qui entraîne une augmentation de la résistance électrique du module et donc une dégradation des performances du Peltier prématurée. Comme nous devons réaliser des mesures sur du long terme, il est important que notre banc de test soit lui-même fiable. C'est pourquoi nous rajoutons un filtre RLC qui lisse la tension d'alimentation du Peltier. Afin de contrôler la polarité du Peltier, réchauffer ou refroidir, les deux bornes du module Peltier sont commandées par chacun un signal PWM. Deux filtres RLC sont nécessaire pour lisser respectivement les deux signaux PWM. Le schéma électrique équivalent est représenté sur la figure 2.17. La fonction de transfert de ce filtre du deuxième ordre prend la forme de :

$$H(j\omega) = \frac{U_s}{U_e} = \frac{K}{\left(j\frac{\omega}{\omega_c}\right)^2 + 2mj\frac{\omega}{\omega_c} + 1} \quad (2.42)$$

1. $S = 0,0003 \text{ m}^2$, $S_F = 0,00459 \text{ m}^2$, $\gamma = 200 \text{ kg.m}^{-3}$, $L = 0,003 \text{ m}$, $M_C = M_H = 0,014 \text{ kg}$, $M_L = 0,05 \text{ kg}$, $M_F = 0,4 \text{ kg}$, $C_C = C_H = 500 \text{ J.kg}^{-1}.\text{K}^{-1}$, $C_L = 250 \text{ J.kg}^{-1}.\text{K}^{-1}$, $C_F = 850 \text{ J.kg}^{-1}.\text{K}^{-1}$, $\alpha = 0,02848 \text{ V.K}^{-1}$, $\rho = 0,00058 \text{ }\Omega.\text{m}^2$, $k = 1,91 \text{ W.m}^{-1}.\text{K}^{-1}$

avec :

$$\omega_c = \sqrt{\frac{R + R_L}{RLC}} \quad m = \frac{(RR_L C + L)\omega_c}{2(R + R_L)} \quad K = \frac{R}{R + R_L} \quad (2.43)$$

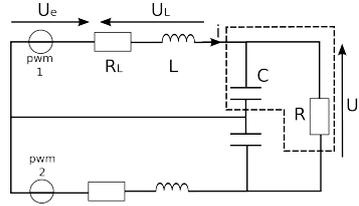


FIGURE 2.17 – Schéma électrique du filtre RLC avec la résistance R qui représente le module Peltier

L'objectif est de dimensionner le condensateur et la bobine pour que le filtre RLC réponde à nos critères. Le RLC que nous concevons doit atténuer l'amplitude du signal PWM de 40dB, il doit avoir une bonne stabilité ainsi qu'un temps de réponse inférieur à 0.01s. Ces différentes contraintes se traduisent pour l'atténuation :

$$H_{dB}(j\omega_{PWM}) = 20 \log_{10} |H(j\omega_{PWM})| = 20 \log_{10}(K) - 20 \log_{10} \left(\sqrt{\frac{\omega_{PWM}^4}{\omega_c^4} + 4m^2 \left(\frac{\omega_{PWM}^2}{\omega_c^2} \right) + 1} \right) \geq 40dB \quad (2.44)$$

pour la stabilité :

$$m = \frac{RR_L C + L}{2(R + R_L)} \omega_c \approx \frac{\sqrt{L}}{2R\sqrt{C}} \leq 0.7 \quad (2.45)$$

et pour le temps de réponse :

$$t_{95\%} \approx \frac{3}{m\omega_c} \approx \frac{6RLC}{(RR_L C + L)} \leq 0.01 \quad (2.46)$$

Sur la figure 2.18 nous avons tracé la capacité du condensateur en fonction de l'inductance de la bobine en fixant le critère d'atténuation (-20dB à -60dB), le critère de stabilité (0.3 à 0.9) et le critère de rapidité (0.1 ms à 1 ms). Ce graphique représente un abaque qui aide à choisir les valeurs du condensateur et de la bobine en fonction des trois critères précédemment définis. Nous avons choisi une inductance de $L = 4.6\mu\text{H}$ et une conductance de $C = 351\mu\text{F}$ ce qui donne $H_{dB}(j\omega_{PWM}) = -40\text{dB}$, $m = 0.7$ et $t_{95\%} = 0.17\text{ms}$.

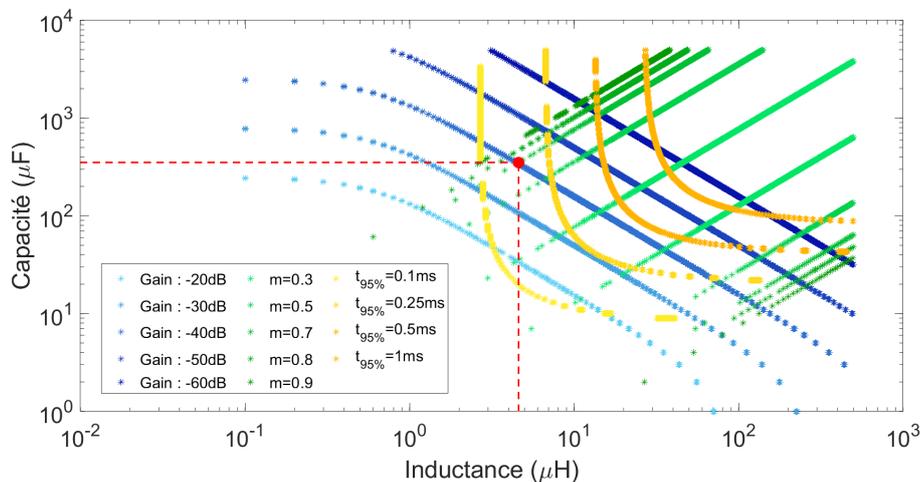


FIGURE 2.18 – Abaque pour dimensionner une capacité et une inductance dans un filtre RLC du second ordre en fonction des critères d'atténuation ($H_{dB}(j\omega_{PWM})$), de stabilité (m) et du temps de réponse ($t_{95\%}$)

Avec ce système de régulation, nous contrôlons la température interne du FPGA de 15°C à 115°C avec une précision de 0.05°C.

2.6.3 Régulation en tension

La carte de développement Ultra96, sur laquelle le FPGA est implémenté, contient deux Power Management Integrated Circuit (PMIC) [144]. Les PMIC contrôlent l'alimentation des composants présents sur la carte et, notamment l'alimentation du FPGA. Les PMIC possèdent un circuit de régulation Proportionnelle Intégrale (PI) afin d'avoir une alimentation stable en sortie du PMIC. Cependant, lorsque la consommation du FPGA augmente, on observe sur la figure 2.12 une chute de la tension interne du FPGA pour une même tension externe. La stabilité des ROs dépend de la tension interne, il est donc nécessaire de réguler la tension interne du FPGA pour que celle-ci reste stable. Une brève description du système de régulation de la tension est donnée sur la figure 2.1 : la mesure de la tension interne du FPGA est envoyée vers un régulateur PI, hébergé dans la PS, dont, la sortie est une consigne de tension envoyée en I2C pour modifier le registre du PMIC qui commande la tension interne du FPGA.

Sur la carte Ultra96, un multiplexeur interface la PS et les deux PMIC [145]. Pour établir une ligne de communication entre la PS et les PMIC, il faut d'abord configurer le registre du multiplexeur. Lorsqu'un chemin de communication est établi entre la PS et le PMIC, les paramètres de configuration de celui-ci sont stockés dans des registres qui sont accessibles en lecture/écriture avec le protocole I2C. Les adresses des registres d'un PMIC sont divisées en 24 pages de 256 adresses, soit 6144 adresses de registres possibles. Une liste des registres avec la description de leurs fonctionnalités est disponible dans [146]. Par exemple, pour modifier la tension du FPGA il faut écrire la nouvelle valeur de tension dans le registre 21h.

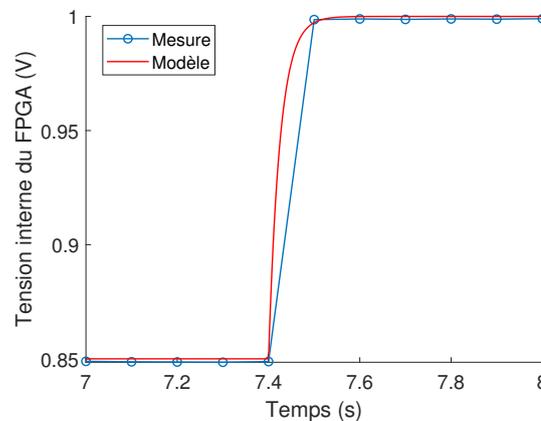


FIGURE 2.19 – Comparaison entre le modèle et la mesure pour une réponse indicielle de 0.15V

La figure 2.19 présente l'évolution de la tension interne du FPGA lorsqu'un échelon de 150mV est commandé par le PMIC. L'échantillonnage de la mesure de la tension interne est de 0.1s. On constate que la dynamique de notre système est largement plus rapide ($< 0.1s$) que l'échantillonnage du circuit de régulation. En conséquence nous n'aurons pas besoin d'apporter de l'avance de phase ce qui nous oriente vers le choix d'un régulateur PI. Le modèle identifié avec un premier ordre est :

$$G_V = \frac{1}{\frac{1}{40}s + 1} \quad (2.47)$$

Pour définir les paramètres du régulateur PI, nous utilisons la méthode de placement aux fréquences et nous obtenons :

$$C_{pi}(s) = C_0 \times \frac{1 + \frac{s}{\omega_i}}{\frac{s}{\omega_i}} \quad (2.48)$$

où $C_0 = 0.4032$ et $\omega_i = 283 \text{ rad.s}^{-1}$.

Avec ce système de régulation, nous contrôlons la tension interne du FPGA de $V_{nom} = 850 \text{ mV}$ à $1.3V_{nom}$. Au dessus de $1.3V_{nom}$, des alarmes de sécurité bloquent la carte. En effet, le fabricant du FPGA préconise une tension d'utilisation comprise entre 825 mV et 875 mV . La précision de la régulation est limitée par le pas de commande du PMIC qui est de 3.9 mV .

2.6.4 Validation du système de régulation

À présent, nous testerons les bonnes performances de la régulation en température et en tension. Pour ce faire, nous avons implémenté dans un FPGA un banc de 1000 oscillateurs en anneaux (RO). Ce banc de ROs est cycliquement actif puis inactif avec une période de 10 minutes. Cette configuration du FPGA est représentative des conditions de test de fiabilité. En plus du banc de ROs, 6 ROs sont toujours actifs pour mesurer les variations de fréquence tout au long de l'expérience. Ces 6 ROs sont divisés en deux groupes qui correspondent à l'implémentation géographique dans le FPGA (voir figure 2.20) :

- FAR : Loin du banc des 1000 ROs,
- NEAR : Proche du banc des 1000 ROs.

Sur la figure 2.21, nous avons tracé la température et la tension interne du FPGA ainsi que la variation relative de la fréquence des ROs FAR et NEAR tel que :

$$\Delta f_R(t) = \frac{f(t) - f_{V_{nom}, 30^\circ\text{C}}}{f_{V_{nom}, 30^\circ\text{C}}} \quad (2.49)$$

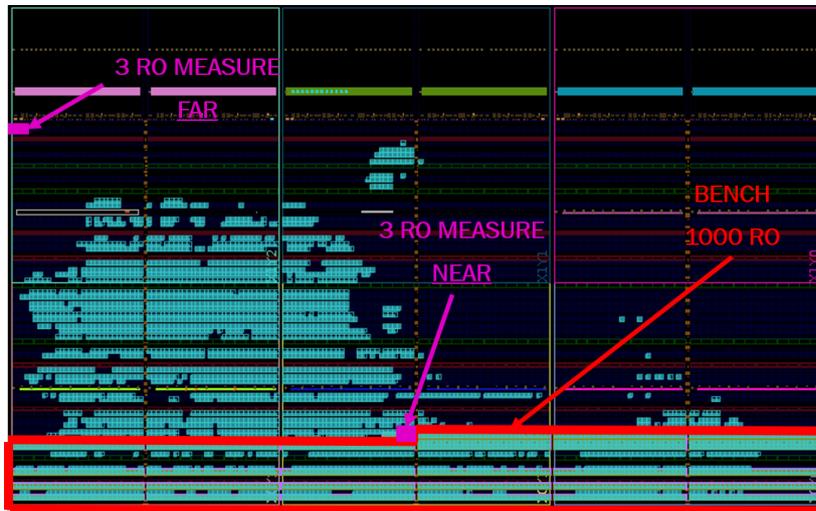


FIGURE 2.20 – Plan d'étage du FPGA avec l'implémentation d'un banc de 1000 ROs, des 3 ROs NEAR et, des 3 ROs FAR

L'expérience est divisée en trois phases de chacune 30 minutes :

- Phase 1 : Régulations inactives en température et en tension,
- Phase 2 : Régulation active en température et inactive en tension,
- Phase 3 : Régulations actives en température et en tension.

Phase 1 : Lorsque le banc est actif sans régulation, on constate une augmentation de température de 6°C et une baisse de la tension interne du FPGA de 10 mV . La conséquence directe de la variation de la température et de la tension est une diminution de la fréquence d'oscillation des ROs de 1.7% . Par ailleurs, le temps de stabilisation de la fréquence, environ 2 minutes, n'est pas

négligeable et suit la même dynamique que le temps de stabilisation de la température.

Phase 2 : La température est régulée à 30°C. Cette température est inférieure aux températures précédentes, entraînant une augmentation globale de la fréquence des RO. Avec la régulation en température, l'amplitude de la variation en fréquence est diminuée de -0.4% quand le banc passe de l'état actif à inactif. Cette variation, assez faible, manifeste que la température n'est pas le paramètre le plus influant sur la fréquence des ROs. Enfin, on constate que la fréquence se stabilise presque instantanément : inférieure à 2 secondes avec régulation contre 2 minutes sans régulation.

Phase 3 : La tension est régulée à 850mV. La régulation en tension nous apporte une stabilité nette sur la fréquence des ROs. Lors de la régulation en tension, on observe des sauts de tension de l'ordre de 4mV. Ils sont la conséquence du pas de commande du PMIC. On constate que, même après régulation, les ROs FAR sont plus stables que les ROs NEAR. Cette information confirme l'importance d'avoir une bonne répartition des ROs dans le FPGA. C'est bien ce que nous avons et que la figure 2.4 montre.

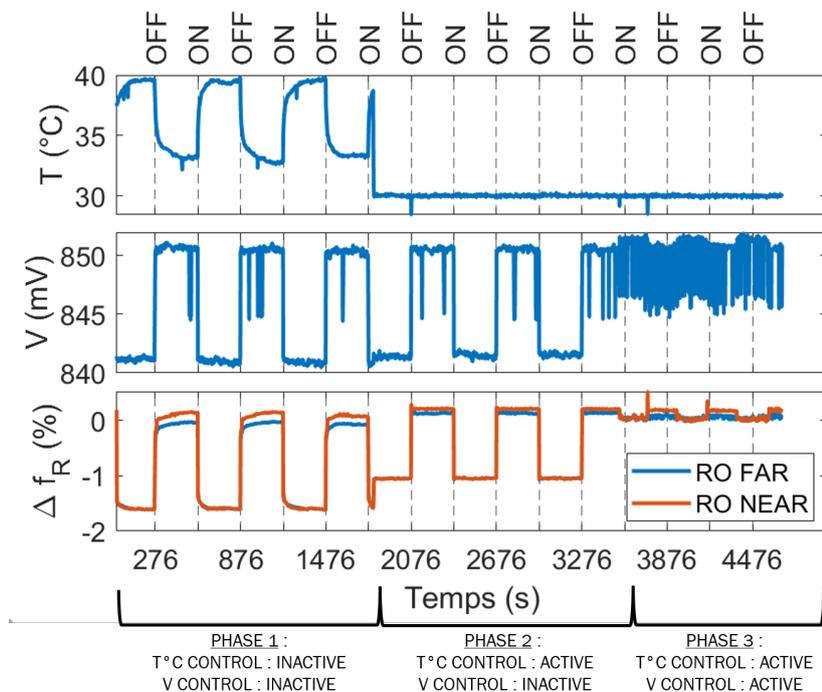


FIGURE 2.21 – Évolution de la température et tension interne du FPGA ainsi que de la fréquence des ROs lors de l'activation et désactivation du banc des 1000 ROs

2.6.5 Conditions de température et de tension possibles

Dans le banc de test que nous développons, les régulations de température et de tension serviront aussi à fixer la condition de stress de température et de tension du FPGA. Dans cette partie, nous présentons une caractérisation des systèmes de régulation afin de connaître les conditions de stress de température et de tension asservissable.

Nous avons implémenter les systèmes de régulation de température et de tension dans un FPGA. L'alimentation du FPGA est régulée pour des tensions fixes ($V = V_{nom}$, $V = 1.1V_{nom}$, $V = 1.2V_{nom}$ et $V = 1.3V_{nom}$). Pour chaque seuil de tension, le banc de test balaye avec un pas de 5% tous les rapports cycliques du PWM qui alimente le Peltier. Pour que la température dans le FPGA

soit stabilisée, le temps entre chaque pas de 5% est d'une minute. La température enregistrée correspond à la moyenne de 100 mesures prises lors des 10 dernières secondes avant d'incrémenter le PWM. Toute cette opération est faite pour le mode réchauffage et le mode refroidissement du Peltier. En mode réchauffage, la borne négative du Peltier est connectée au PWM tandis que la borne positive est connectée à la masse, à l'inverse du mode refroidissement.

La figure 2.22 présente les résultats du test précédemment décrit. Chaque point correspond à la moyenne de 100 mesures de température pour un PWM spécifique. Dans le mode réchauffage, la température augmente lorsque le rapport cyclique du PWM augmente, dans le mode refroidissement, la température diminue lorsque le rapport cyclique du PWM augmente. Dans le mode refroidissement, la température minimale est de 15°C. Pour le mode réchauffage, la température maximale dépend de la tension interne du FPGA. En effet, sur la carte de développement Ultra96-V2 des alarmes de coupures sont utilisées par le fabricant pour protéger le FPGA d'une surtension et d'une température trop élevée. Par exemple, on constate pour $V = 1.3V_{nom}$ que la température maximale est de $T = 90^{\circ}\text{C}$, au delà de cette température la carte de développement s'éteint. Donc, lors du choix des conditions de stress des FPGAs, il sera important de prendre en compte les limites du circuit de régulation mais aussi des limites imposées par la carte de développement.

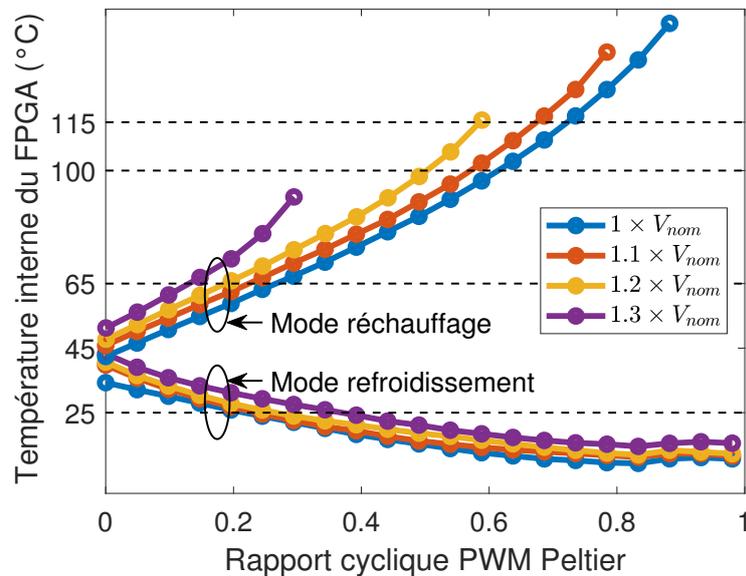


FIGURE 2.22 – Température interne du FPGA en fonction du rapport cyclique du PWM gérant l'alimentation du Peltier et de plusieurs tensions d'alimentation que nous réglons. La borne positive du Peltier est connectée à la masse pour le mode réchauffer tandis que pour le mode refroidir c'est la borne négative du Peltier qui est connectée à la masse.

En conclusion, dans cette partie nous avons présenté le système de régulation en température et en tension que nous avons développé. Nous pouvons ainsi réguler la température et la tension interne du FPGA avec une précision de 0.05° et de 3.9mV . Grâce aux deux systèmes de régulation, nous améliorons la précision de la mesure de fréquence des ROs qui est aussi sensible à la température et la tension.

2.7 Validation du banc de vieillissement

Afin de valider le fonctionnement du banc de test et sa capacité à mesurer des dérives d'oscillateur en anneau, nous avons fait un premier test de vieillissement pendant 1000 heures. Nous présentons dans cette partie les résultats obtenus.

2.7.1 Présentation du test

L'architecture du banc de test correspond à la figure 2.1 détaillée dans les précédentes sections. Pour ce test, nous avons utilisé un seul FPGA. Nous avons choisi une condition de stress : $T = 100^\circ\text{C}$ et $V = 1.25V_{nom}$ pour évaluer l'importance des dérives de fréquence pendant ce test relativement bref. Les mesures de fréquence et de rapport cyclique des 567 ROs dans le FPGA sont envoyées et enregistrées dans un ordinateur toutes les dix minutes. Nous envoyons aussi les mesures de température et de tension qui serviront pour la compensation post-mesure.

2.7.2 Caractérisation du RO et compensation post-mesure

Nous avons vu dans la section 2.6 que la résolution du système de régulation en température (0.05°C) et en tension (3.9mV) est supérieure au pas de mesure de la température ($7.7 \times 10^{-3}^\circ\text{C}$) et de la tension ($4.6 \times 10^{-2} \text{mV}$). Sachant que le RO est sensible aux variations de température, nous avons effectué une caractérisation de la fréquence et du rapport cyclique en fonction de la température et de la tension. Au moment où les ROs sont mesurés et enregistrés, nous enregistrons aussi la température et la tension interne du FPGA. À partir de ces mesures nous construisons un modèle qui nous permettra de compenser l'effet des erreurs résiduelles de régulation de la tension et de la température. Ainsi, nous pouvons réaliser une compensation post-mesure à partir de la caractérisation réalisée en amont du test de 1000 heures et des mesures.

Caractérisation du RO

Avant le test de 1000 heures, nous avons réalisé la caractérisation d'un seul RO en fonction de la tension : de 650mV à 1150mV et ce pour quatre températures : 25°C , 50°C , 75°C et 100°C . Nous retrouvons les mesures de cette caractérisation sur la figure 2.23. Nous avons modélisé l'évolution de la fréquence avec un polynôme du troisième degré pour la tension et du premier degré pour la température :

$$M_f(\Delta V(t), \Delta T(t)) = \frac{f(\Delta V(t), \Delta T(t))}{f_{V_0, T_0}} = 1 + p_{10}\Delta V(t) + p_{01}\Delta T(t) + p_{20}\Delta V(t)^2 + p_{11}\Delta V(t)\Delta T(t) + p_{30}\Delta V(t)^3 + p_{21}\Delta V(t)^2\Delta T(t) \quad (2.50)$$

avec $\Delta V(t) = V(t) - V_0$ et $\Delta T(t) = T(t) - T_0$. Graphiquement on constate que la tension a un impact plus important sur la fréquence du RO que la température. Aussi, nous voyons que pour les tensions inférieures à la tension nominale la fréquence augmente lorsque la température augmente tandis que pour les tensions supérieures, la fréquence diminue lorsque la température augmente.

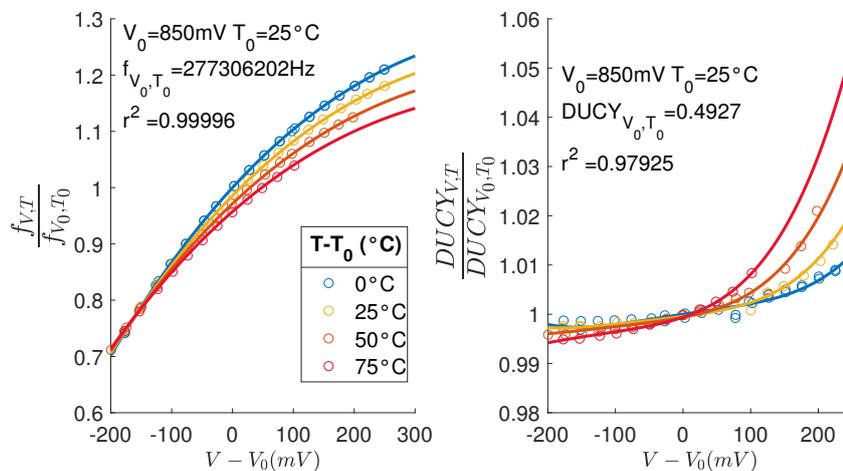


FIGURE 2.23 – Caractérisation de la fréquence (gauche) et du rapport cyclique (droite) d'un RO en fonction de la tension et pour différentes températures

Le manque de mesures à haute température et haute tension est lié aux alarmes de mise en sécurité de la carte.

Concernant la modélisation du rapport cyclique, nous avons utilisé un polynôme de degré 4 pour la tension et de degré 2 pour la température. Nous appellerons ce modèle $M_{DUCY}(\Delta V(t), \Delta T(t))$. Sur la figure 2.23, l'échelle des ordonnées pour le rapport cyclique est différente de celle pour la fréquence. On constate que le rapport cyclique est relativement stable pour des tensions inférieures à V_{nom} . Au delà de cette tension, le rapport cyclique croît rapidement en fonction de la tension, ce qui veut dire qu'il existe un équilibre entre les PMOS et NMOS qui n'est plus conservé à partir d'une certaine tension. Par ailleurs, on constate que le rapport cyclique mesuré aux conditions nominales montre le bon équilibre entre PMOS et NMOS ($DUCY_{V_0, T_0} = 0.4927$).

Compensation post-mesure

Afin de compenser les variations de température et de tension autour de la consigne, nous avons appliqué une compensation post-mesure telle que, sur la fréquence :

$$f_{V_c, T_c}(t) = f_{V_m, T_m}(t) \frac{M_f(\Delta V_c(t), \Delta T_c(t))}{M_f(\Delta V_m(t), \Delta T_m(t))} \quad (2.51)$$

où V_m et T_m désignent la tension et la température mesurée, V_c et T_c désignent la tension et la température de consigne, $f_{V_m, T_m}(t)$ est la fréquence mesurée à la température et tension mesurées tandis que $f_{V_c, T_c}(t)$ est la fréquence mesurée et ramenée à la température et à la tension de consigne. Nous avons aussi appliqué une compensation sur les mesures de rapport cyclique en utilisant son modèle de caractérisation.

Sur la figure 2.24a nous observons l'évolution de la fréquence relative d'un RO pendant 1000 heures de vieillissement ainsi que l'évolution de la température et de la tension interne du FPGA. Après 1000 heures de test nous observons une dérive de la fréquence relative de 0.3%. La fréquence

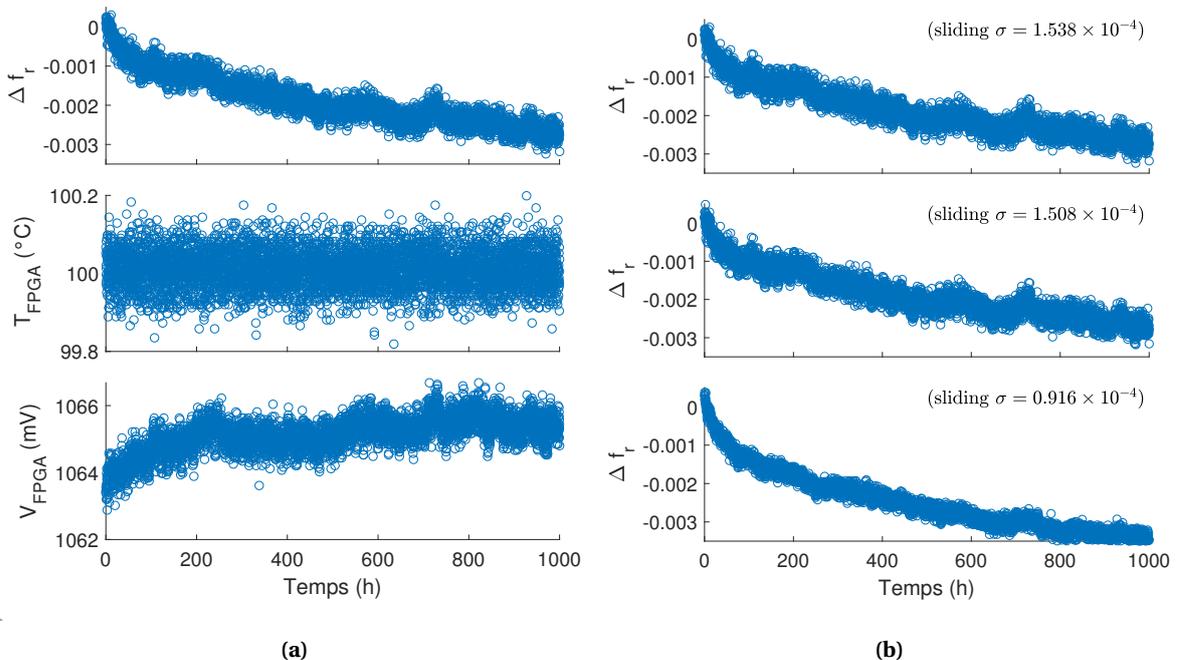


FIGURE 2.24 – (a) Évolution de la fréquence relative (haut), de la température (milieu) et de la tension (bas) en fonction du temps sans compensation post-mesure (b) Évolution de la fréquence relative en fonction du temps sans compensation (haut), après une compensation de la température uniquement (milieu) après une compensation en température et tension

relative s'exprime :

$$\Delta f_r = \frac{f(t) - f(t_0)}{f(t_0)} \quad (2.52)$$

On constate que la température est stable tandis que nous observons une légère dérive de la tension ($2mV$). La dérive de la tension est inférieure au pas de commande du PMIC ($3.9mV$) ce qui explique qu'elle ne soit pas corrigée.

Nous avons tracé sur le haut de la figure 2.24b la même dérive que celle de la figure 2.24a. Après application de la compensation post-mesure des variations de température uniquement on constate que la dérive reste inchangée et le bruit de mesure n'est que très légèrement amélioré. Le système de régulation en température est donc suffisamment précis et les légères variations de la température interne ($< 0.1^\circ C$) ont un impact négligeable sur la fréquence du RO. Lorsque nous appliquons la compensation post-mesure des variations de tension, on visualise une nette amélioration du bruit de mesure de la fréquence. Par ailleurs, l'amplitude de la dérive fréquentielle est plus importante qu'avant compensation. En effet, sur la figure 2.24a nous voyons une dérive de la tension dont l'impact est l'augmentation de la fréquence. La compensation post-mesure élimine cet impact de la dérive de la tension sur la fréquence. Finalement, après compensation post-mesure nous observons une dérive fréquentielle causée par le seul vieillissement. Notons qu'après compensation post-mesure nous obtenons une très bonne précision de mesure l'ordre de 9×10^{-5} .

2.7.3 Résultats pour différentes configurations internes

Nous avons tracé sur la figure 2.25 plusieurs dérives fréquentielles de ROs pour différentes architectures de ROs et différents stress internes. Il est important de noter que les mesures tracées sont brutes, aucun traitement sur les données n'étant effectué, hormis la compensation post-mesure. De manière générale, après 1000 heures de vieillissement dans des conditions de stress proches des conditions d'utilisation nous observons très clairement des dérives fréquentielles des ROs. De plus, nous pouvons déjà distinguer l'impact du stress interne ou de l'architecture du RO sur le vieillissement.

Sur la figure 2.25a on voit que le vieillissement du RO L1 B, qui est un RO de la catégorie short path, est plus sensible au stress statique (DC0 et DC1) et de basse fréquence (100Hz) qu'aux stress de hautes fréquences. Sur la figure 2.25b on compare ici tous les ROs de la catégorie short path, on constate un faible impact de l'architecture du RO sur le vieillissement. Sur la figure 2.25c nous voyons que les ROs LP (Long Path) ont des dégradations maximales avec un stress DC1 et des dégradations minimales avec un stress DC0. Si nous comparons la figure 2.25a et la figure 2.25c, nous observons que l'amplitude des dégradations sont plus importantes pour les ROs short path que les RO long path.

Finalement, nous avons modélisé les données de la figure 2.25a avec une loi puissance puis nous avons tracé le résultat en échelle log-log sur la figure 2.25d. Les conditions de stress en température et en tension ainsi que l'exposant au temps de la loi puissance (0.28 à 0.35) nous laisse à penser que le mécanisme de défaillance observé est du Bias Temperature Instability (BTI). L'exposant au temps est légèrement supérieur à celui défini dans le modèle de Reaction-Diffusion (0.25) [74]. Notre modélisation du mécanisme de dégradation est effectuée à partir des mesures au niveau du circuit, comportant de nombreux PMOS et NMOS, contrairement aux modélisations classiques qui utilisent des mesures réalisées sur un transistor isolé. Les résultats présentés sur la figure 2.25 sont donc des dégradations globales de plusieurs transistors PMOS (NBTI) et NMOS (PBTI) dégradés.

À ce niveau de l'étude nous souhaitons simplement montrer les bonnes performances du banc de test. Avec ce test de vieillissement de 1000 heures, nous avons mis en évidence que le banc

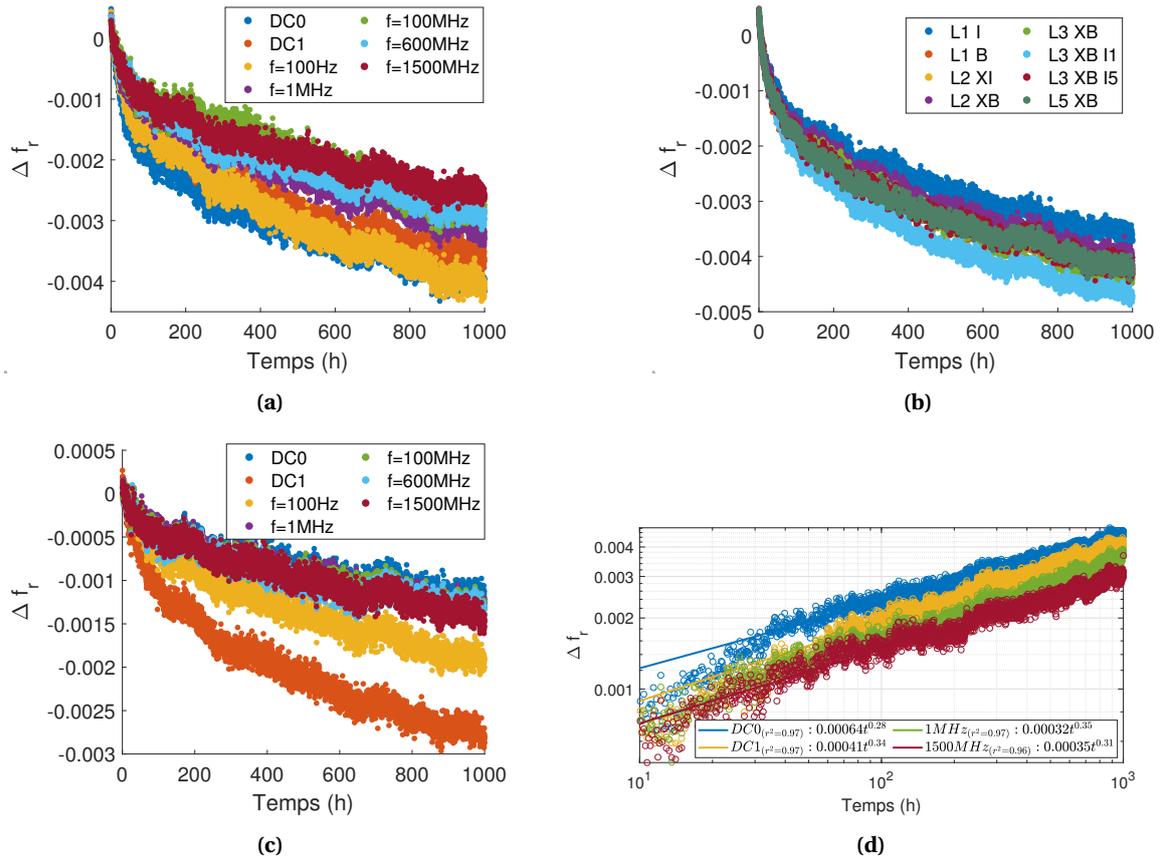


FIGURE 2.25 – Dérives fréquentielles de ROs obtenus avec $T = 100^\circ\text{C}$ et $V = 1.25V_{nom}$ et avec des paramètres d'architecture ou de stress internes différents (a) Catégorie RO : short path; Architecture RO : L1 B; Fréquence stress : variable; Rapport cyclique de stress : 0.5 (b) Catégorie RO : short path; Architecture RO : variable; Fréquence stress : DC1; Rapport cyclique de stress : DC1 (c) RO : LP; Fréquence stress : variable; Rapport cyclique de stress : 0.5 (d) Catégorie RO : long path; Architecture RO : L1 B; Fréquence stress : variable; Rapport cyclique de stress : 0.5 avec une échelle en log-log

de test est capable de mesurer des dérives fréquentielles de RO avec une précision relative de 9×10^{-5} . Grâce à cette précision de mesure nous observons des dérives de ROs en 1000 heures de vieillissement avec un stress en température et tension proche des conditions d'utilisation. Ce banc de test répond complètement à nos attentes en terme de performances.

2.8 Conclusion

Dans ce chapitre, nous avons détaillé l'architecture du banc de test qui servira pour le vieillissement et la mesure de dégradation des FPGA. Dans le FPGA, nous implémentons des oscillateurs en anneaux qui sont les circuits sous test. Nous avons présenté les différentes architectures d'oscillateurs en anneaux implémentées dans le FPGA. L'implémentation de plusieurs architectures de ROs nous permettra d'étudier le vieillissement de différentes ressources dans le FPGA tel que les ressources de routages (RO Long Path) ou les ressources logiques (RO Short Path). Sur notre banc, un signal de stress interne est envoyé dans le RO lorsqu'il n'est pas mesuré, c'est à dire 99.6% du temps. Nous appliquons 21 stress internes différents (statiques et dynamiques) afin de voir leurs effets sur le vieillissement. Au total, nous avons implémenté 567 ROs par FPGA.

Nous avons présenté l'architecture du circuit qui mesure la fréquence et le rapport des ROs. Nous avons montré comment extraire le temps de propagation d'un front montant et descendant dans une LUT à partir de la mesure de fréquence de rapport cyclique du RO. Enfin, nous avons

montré par la mesure que la référence de temps fournie par un récepteur GPS est plus précise et plus stable que celle fournie par l'oscillateur à quartz présent sur la carte de développement. Cette étude justifie l'utilisation du récepteur GPS comme référence de temps pour la mesure de fréquence des ROs dans un but d'améliorer la stabilité et la précision de mesure du banc de test.

Parce que le RO est sensible non seulement au vieillissement mais aussi à la température et la tension d'alimentation, nous avons développé un système de régulation en température et en tension. Nous régulons ainsi la température interne du FPGA avec une précision de 0.05°C et la tension interne avec une précision de 3.9mV . De plus, en parallèle de la mesure du RO, nous envoyons la mesure de température et de tension afin de réaliser des compensations post-mesures. Nous n'avons pas développé de solution pour détecter le vieillissement des composants utilisés pour l'asservissement. Nous faisons l'hypothèse que les circuits internes de mesure de tension et de température (SYSMON) sont conçus pour être peu sensibles aux caractéristiques des transistors, donc à leurs vieillissements. Néanmoins, une attention particulière sera apportée sur le comportement de la dérive du temps de propagation pour confirmer que nous observons uniquement l'effet des mécanismes de dégradations des transistors.

Finalement, nous avons réalisé un premier test de vieillissement de 1000 heures avec $T_{stress} = 100^{\circ}\text{C}$ et $V_{stress} = 1.25V_{nom}$. Grâce à la précision de mesure de notre banc de test, nous avons observé pour un temps de test relativement court et dans des conditions de stress proches des conditions de fonctionnement opérationnel, des dégradations inférieures à 0.5% avec une précision de mesure relative de 9×10^{-5} . À priori, selon l'étude de la majorité des bancs de vieillissement à base de FPGA (voir table 1.3), le banc de test que nous avons développé est le plus précis.

Chapitre 3

Analyse et extraction des dégradations dans les FPGAs

3.1 Introduction

Les mécanismes de dégradations tel que le BTI, le HCI ou le TDDB ont pour conséquence d'augmenter le temps de propagation de la transition d'un signal dans un circuit numérique. Afin de mesurer la fiabilité des FPGAs de dernière génération, nous avons développé un banc de test qui vieillit et mesure les dégradations. Pour mesurer les dégradations dans un FPGA, nous avons implémenté des oscillateurs en anneaux (ROs) dont nous mesurons l'évolution de la fréquence du signal de sortie qui dépend directement du temps de propagation du signal dans le RO. La cause élémentaire du vieillissement est le temps. Cependant, afin d'étudier l'effet d'autres paramètres sur le vieillissement des ROs, ceux-ci sont soumis à d'autres contraintes tel que la température, la tension d'alimentation, les signaux de stress internes (statique ou dynamique) circulant dans le RO et l'architecture du RO.

Après avoir fait vieillir et mesuré la dégradation de neuf FPGAs soumis à différentes températures et tensions de stress pendant presque 10000 heures, nous allons, dans ce chapitre, présenter les résultats. La figure 3.1 présente une vue d'ensemble du contenu de ce chapitre.

Dans la section 3.2, nous présenterons le banc de test ainsi que les conditions de températures et de tensions choisies pour le vieillissement et la mesure de dégradation des neuf FPGAs. L'architecture développée pour chaque FPGA correspond à celle présentée dans le chapitre précédent (voir figure 2.1).

Ensuite, dans la section 3.3, avant de commencer le vieillissement des FPGAs, nous avons mesuré tous les ROs afin de comparer les performances initiales des FPGAs avec les données de simulation extraites de Vivado. Aussi, nous avons caractérisé la fréquence de tous les ROs en fonction de la température et de la tension interne du FPGA afin d'étudier l'efficacité d'utiliser un modèle propre à chaque RO pour la compensation post-mesure des erreurs résiduelles de la régulation de température et de tension.

Dans la section 3.4 nous présenterons les résultats de la dégradation des ROs et nous ferons une analyse préliminaire du vieillissement des ROs sous l'effet de la température, la tension, le stress interne dépendant de la fréquence et du rapport cyclique, ainsi que de l'architecture du RO.

La section 3.4 révélera un vieillissement différent suivant l'architecture du FPGA et plus précisément entre les ressources de routage et les ressources logiques dans un RO (LUT). C'est pourquoi, dans la section 3.5 nous présenterons une méthode pour dissocier le temps de propagation dans les ressources de routage de celui dans les ressources de LUT, ce qui nous permettra dans le chapitre 4 de modéliser séparément le vieillissement des ressources de routage de celui des LUTs. Finalement, en nous basant sur l'extraction de la dégradation dans les ressources de LUT, nous ferons l'extraction des dégradations dans les PMOS et NMOS et ce afin d'identifier le NBTI et le PBTI.

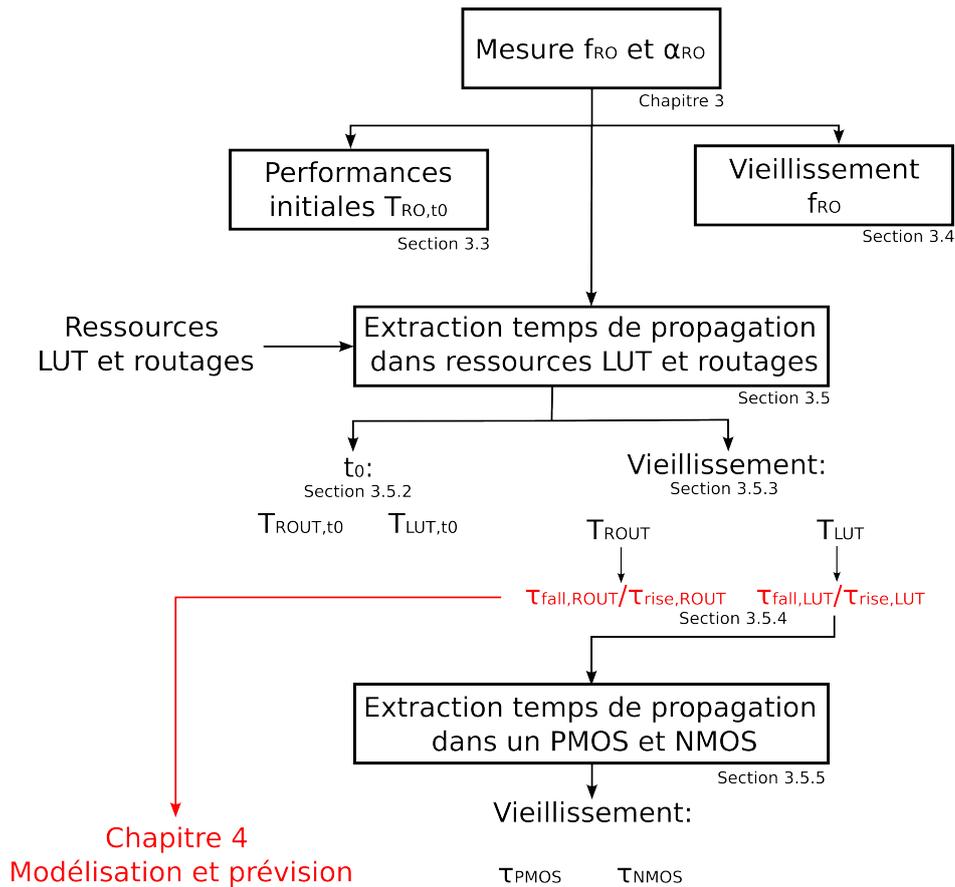


FIGURE 3.1 – Plan d'organisation pour la présentation des mesures expérimentales

3.2 Banc de test et plan d'expérience

Afin de répondre à la problématique de cette thèse qui consiste à mesurer et prévoir la fiabilité des FPGAs, nous avons développé un banc de test qui vieillit et mesure la dégradation de neuf FPGAs, chacun contenant 567 ROs. Le banc de test doit vieillir les FPGAs pendant plus de 10000 heures. À ce jour, nous présenterons les résultats pour 8000 heures de vieillissement. La figure 3.2 montre une photo du banc de test développé. Seules les conditions de température et de tension sont différentes entre les FPGAs. La carte de développement Ultra96 sur laquelle le FPGA est présent est en noir. Nous avons conçu une carte, en vert sur la photo, afin d'y mettre tous les composants externes à la carte de développement : un filtre RLC, un pont en H et des optocoupleurs pour interfacer entre la puissance et la logique. Une carte Arduino (en bleu) gère la communication des données depuis chaque FPGA vers l'ordinateur. Comme on peut le voir sur la figure 2.1, chaque FPGA communique les données avec un protocole UART. Un seul port UART est présent sur la carte Arduino, cependant, neuf sorties (une par FPGA) sont dédiées à autoriser le FPGA qui doit communiquer avec l'Arduino et mettre en veille la communication des autres FPGAs. Donc, l'Arduino reçoit les données de chaque FPGA successivement et cela une fois par heure. Les données reçues par la carte Arduino sont ensuite transférées en UART vers un ordinateur qui stocke les données dans un fichier puis enregistre ce fichier sur un serveur externe.

Le choix des conditions de stress en température et en tension dépend des mécanismes de défaillance que nous souhaitons observer. De manière générale, les conditions idéales sont :

- pour le BTI ou le TDDB : Température et tension hautes ;
- pour le HCI : Température basse et tension haute ;

Sur la figure 3.3 nous voyons les conditions de stress que nous avons choisies pour le vieillissement des neuf FPGAs. L'identification des mécanismes de défaillance en fonction des conditions

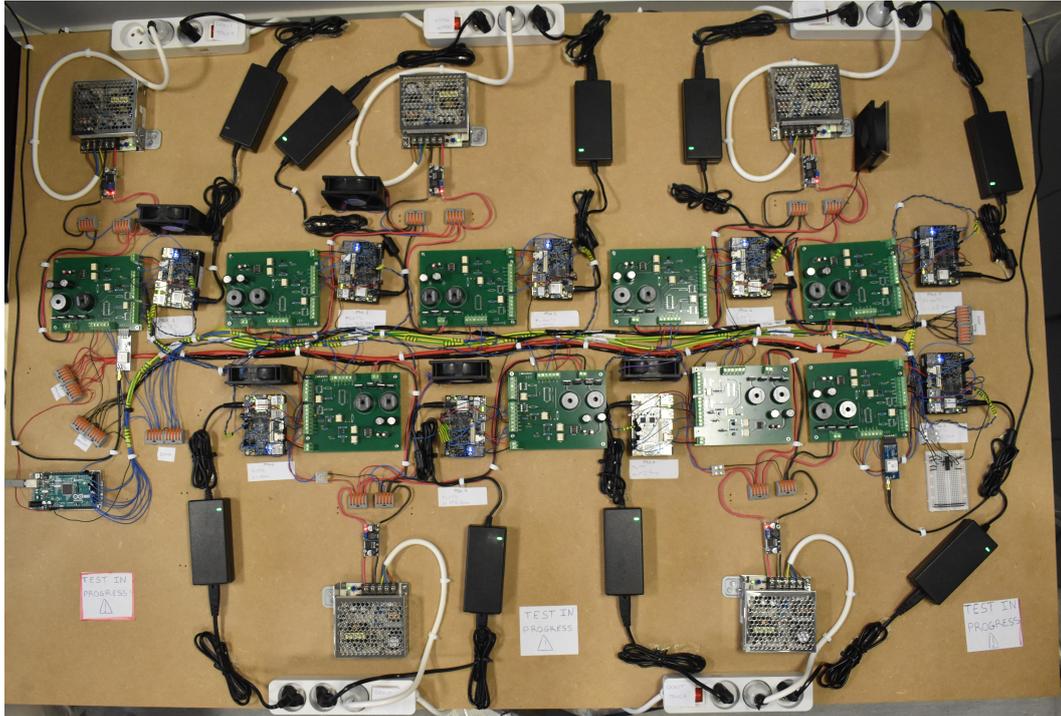


FIGURE 3.2 – Photo du banc de test pour le vieillissement et la mesure de la dégradation de neuf FPGAs

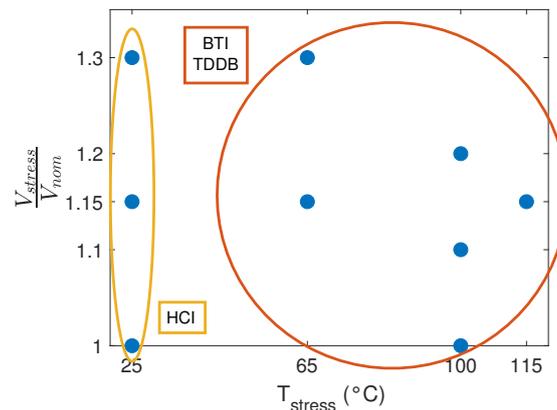


FIGURE 3.3 – Conditions de stress en température et en tension des neuf FPGAs

est préliminaire, nous ferons une identification plus précise lors de l'analyse des mesures. Les conditions de stress sont choisies en prenant soin de respecter les limites fonctionnelles des systèmes de régulation et des alarmes de protection dans la carte de développement que nous avons présenté dans la section 2.6.5. Enfin, il est important de noter que nous choisissons un stress en température et en tension proche des conditions d'utilisations afin d'avoir une prédiction de la fiabilité qui reflète au mieux les dégradations réelles.

3.3 Mesures initiales des RO

3.3.1 Introduction

Avant de vieillir les FPGAs, nous avons pris soins de mesurer tous les ROs dans les conditions de fonctionnement opérationnelles ($T = 25^\circ\text{C}$ et $V = V_{\text{nom}}$). Pour connaître l'état initial de nos FPGAs, dans cette section la période initiale des ROs obtenue à partir de la mesure de fréquence est comparée avec les données de simulation extraites dans le logiciel de conception Vivado. Afin de

garantir le respect des contraintes de temps, Vivado prédit le temps de propagation d'un signal dans un chemin combinatoire ou séquentiel. Dans un chemin combinatoire, comme c'est le cas des oscillateurs en anneaux, le temps de propagation dépend uniquement du retard accumulé dans les ressources logiques (LUT) et dans les ressources de routage.

Avant de vieillir les FPGAs, nous avons caractérisé la fréquence d'oscillation de tous les ROs en fonction de la température et de la tension. Dans cette section, nous montrons l'importance d'utiliser un modèle propre à chaque RO pour la compensation post-mesure des erreurs résiduelles de la régulation de température et de tension afin d'avoir une meilleure précision des résultats de dégradation.

3.3.2 Période des ROs

Mesure

Avant le vieillissement des FPGAs, nous avons mesuré la fréquence de tous les ROs dans les neuf FPGAs une fois par heure pendant dix heures. Pendant les mesures, les neuf FPGAs sont réglées à $T = 25^{\circ}\text{C}$ et $V = V_{nom}$ et ont une architecture logique identique qui est celle décrite dans le chapitre 2. Sur la figure 3.4 nous avons tracé la période moyenne de tous les ROs dans les neuf FPGAs calculée à partir de la fréquence des ROs mesurée pendant dix heures. Chaque graphique de la figure 3.4 correspond à une architecture logique de RO décrite dans la section 2.3. Dans chaque graphique, les lignes de couleurs correspondent à une architecture physique d'un RO qui est identique sur tous les FPGAs. La figure 3.4 montre que la période de tous les ROs se situe entre 2ns et 7ns soit une fréquence comprise entre 500MHz et 150MHz .

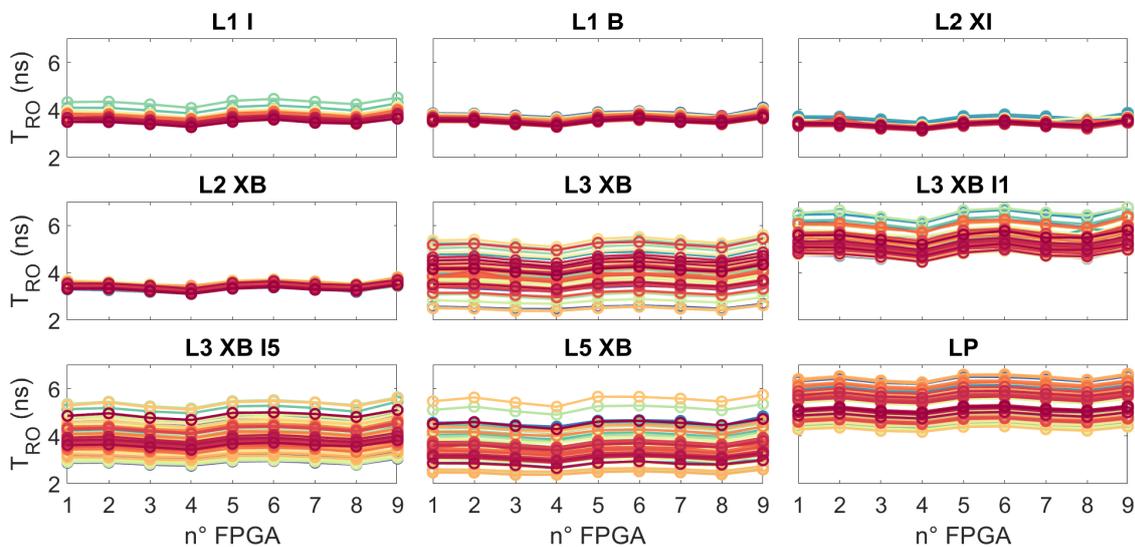


FIGURE 3.4 – Périodes des ROs en fonction du numéro FPGA ($T = 25^{\circ}\text{C}$ et $V = V_{nom}$) pour les neuf architectures de RO. Chaque ligne de couleur correspond à une architecture physique de RO.

Sur la 3.4 on constate peu de variation de la période d'un RO de même architecture physique entre les FPGA. De plus, les variations d'un FPGA à l'autre similaires entre toutes les architectures de RO semblent indiquer que les variations entre FPGAs sont principalement dues aux variations de la technologie. Graphiquement et globalement on relève une dispersion inférieure à 10%. Cette faible dispersion entre les FPGAs met en évidence une faible dispersion technologique lors de la fabrication des FPGAs.

Concernant les ROs L1 I, L1 B, L2 XI et L2 XB (les quatre premiers graphiques) on observe aussi peu de variations entre les ROs de même architecture et dans un même FPGA. Cependant, pour

les ROs L3 XB, L3 XB I1, L3 XB I5 et L5 XB (les cinq derniers graphiques), on observe des variations de plus du double entre les périodes minimales et maximales. Pourtant, la description logique (VHDL) est identique pour tous les ROs d'une même architecture. Cependant, l'implémentation dans le FPGA est gérée par Vivado qui peut utiliser des ressources physiques différentes. C'est l'utilisation de ces ressources physiques différentes qui explique la dispersion de la période des ROs de même architecture logique dans un même FPGA. Nous avons contraint l'implémentation des ROs L1 I, L1 B, L2 XI et L2 XB dans une slice tandis que les ROs L3 XB, L3 XB I1, L3 XB I5 et L5 XB sont dans deux slices. Lors de l'implémentation des ROs sur deux slices, le logiciel a un choix de ressources disponibles plus important que pour l'implémentation dans une slice. Donc, deux slices plutôt qu'une, favorisent une architecture interne différente des ROs. C'est pourquoi nous observons plus de variation de la période des ROs implémentés sur deux slices.

La figure 3.4 montre aussi que la période des ROs LP (Long Path) est du même ordre de grandeur de certains RO SP (Short Path) implémentés sur deux slices. Pourtant dans le RO LP, il y a seulement deux étages de LUT séparés par un long chemin de routage, tandis qu'il y en a 13 dans le RO SP. Cela montre que les ressources de routage contribuent grandement au temps de propagation. Aussi, en moyenne, la période des ROs d'architectures L3 XB I1 est plus importante que celle des ROs d'architecture L3 XB I5. Si nous considérons l'architecture de la LUT décrite dans [44], lorsque l'entrée I1 de la LUT est utilisée le chemin parcouru par le signal est plus long que lorsque c'est l'entrée I5. Cela explique la différence de période entre les RO L3 XB I1 et les ROs L3 XB I5 qui, respectivement, n'utilisent que des entrées I1 et I5 de la LUT.

Comparaison avec Vivado

Dans la littérature, les études montrent comment extraire [147–150], analyser les variations [112, 151, 152] et contrôler [153–155] le temps de propagation d'un chemin dans un FPGA en utilisant les données de Vivado ou les mesures expérimentales. Une légère comparaison entre les données de Vivado et les mesures expérimentales du temps de propagation est donnée dans [152]. Cependant, l'état de l'art ne montre aucune étude complète comparant les données de Vivado avec des mesures expérimentales du temps de propagation. Dans cette partie, nous utilisons les périodes des ROs mesurées et présentées sur la figure 3.4. Aussi, nous avons écrit un script TCL pour extraire les temps de propagation prédits par Vivado pour chaque RO afin de comparer les périodes de ROs mesurées avec les périodes de ROs prédites par le fabricant.

Afin de couvrir les variations de processus de fabrication (process corner), Vivado retourne quatre bornes de temps de propagation : Fast Min, Fast Min, Slow Max et Slow Min. Fast et Slow correspondent au FPGA le plus rapide et le plus lent en sortie de fabrication tandis que Max et Min correspondent aux ressources dans le FPGA avec le maximum et le minimum de temps de propagation. Donc, Fast et Slow couvrent les variations inter-FPGA et Max et Min couvrent les variations intra-FPGA.

Sur la figure 3.5a nous avons rapporté la période donnée par Vivado et celle mesurée. Chaque point correspond à un RO et les quatre couleurs identifient la borne fixée par Vivado pour être comparée avec la mesure. Cette figure montre que les périodes mesurées se situent entre les deux bornes limites attendues par Vivado Fast Min (meilleure performance) et Slow Max (pire performance). De plus, on constate une cohérence et une linéarité entre la mesure et les données Slow Min et Fast Max de Vivado. Afin d'observer les différences entre les architectures de RO, nous représentons les mesures avec la référence Slow Min dans la figure 3.5b en identifiant chaque architecture logique. On observe que la référence Slow Min sous-estime la mesure des ROs LP et RO L3 I5. Ces deux architectures ont en commun que le routage est le principal contributeur au temps de propagation. Les autres architectures logiques de ROs sont globalement identiques à la borne Slow Min.

Parce que les références de temps de propagation données par Vivado dépendent de deux pa-

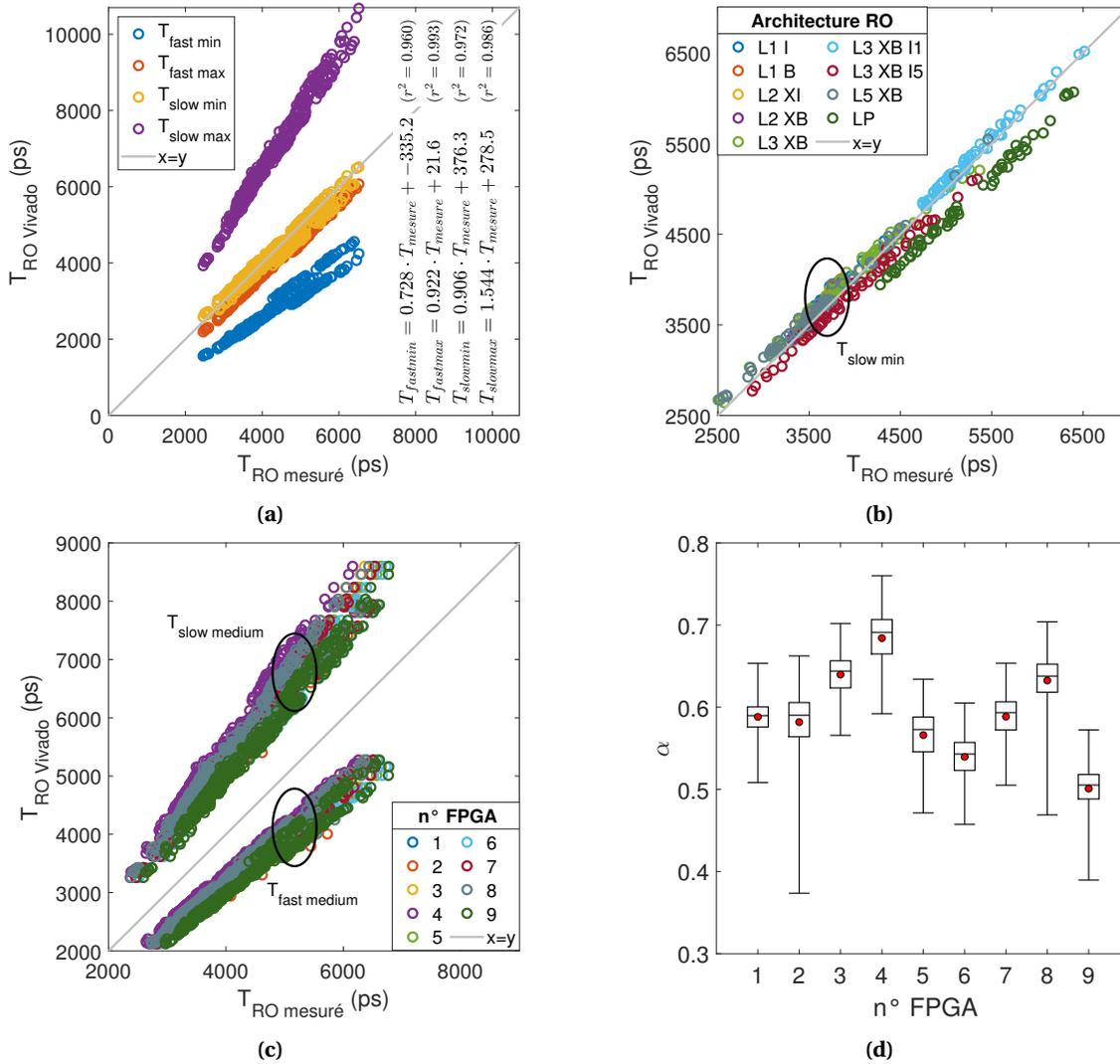


FIGURE 3.5 – Comparaison de la période du RO mesurée et extraite de Vivado pour (a) tous les ROs d’un FPGA, (b) Slow Min avec une couleur différente selon l’architecture du RO (c) Fast Medium et Slow Medium avec une couleur différente selon le FPGA (d) dispersion de la coordonnée barycentrique α

ramètres (Fast/Slow and Max/Min) il est ambiguë de situer la performance d’un FPGA par rapport aux prédictions de Vivado. Par exemple, sur la figure 3.5a nous voyons que la mesure est cohérente avec la référence Slow Min (FPGA le plus lent en considérant les composants internes les plus rapides) et Fast Max (FPGA le plus rapide en considérant les composants internes les plus lents). C’est pourquoi nous avons créé deux nouvelles références en considérant les performances intra FPGA comme moyenne. Grâce à ces nouveaux paramètres, nous connaissons l’état de performance de notre FPGA par rapport aux attentes du fabricant Xilinx. Les références sont :

$$T_{X,medium} = \frac{T_{X,max} + T_{X,min}}{2} \quad (3.1)$$

où X peut être Fast ou Slow. Nous avons tracé sur la figure 3.5c $T_{Fast,medium}$ et $T_{Slow,medium}$ en fonction de la mesure et avec une couleur par FPGA. Graphiquement, nous voyons que les mesures du FPGA9 sont les plus éloignées de la référence $T_{Fast,medium}$ et les plus proches de la référence $T_{Slow,medium}$ contrairement au FPGA4. Nous pouvons déjà en conclure que le FPGA9 possède la pire performance et le FPGA4 les meilleures. À partir des données tracées sur la figure 3.5c nous écrivons les mesures sous forme de coordonnées barycentriques par rapport aux références $T_{Fast,medium}$ et $T_{Slow,medium}$ tel que :

$$T_{Mesure} = \alpha \cdot T_{Fast,medium} + (1 - \alpha) \cdot T_{Slow,medium} \quad (3.2)$$

À partir de l'équation (3.2) nous extrayons la coordonnée barycentrique α pour tous les ROs de chaque FPGA :

$$\alpha = \frac{T_{\text{Mesure}} - T_{\text{slow,medium}}}{T_{\text{fast,medium}} - T_{\text{slow,medium}}} \quad (3.3)$$

Si $\alpha = 1$ (resp. $\alpha = 0$) alors les performances du FPGA mesurées correspondent aux meilleures (resp. pires) performances attendues par Vivado. Nous avons tracé la distribution de la coordonnée barycentrique α pour chaque FPGA dans un graphique écart inter-quartile sur la figure 3.5d. Les performances de nos FPGA se situent entre 50% et 70% des meilleures performances attendues par Vivado. La dispersion de la coordonnée barycentrique α est due à la dispersion des données de Vivado, au bruit de mesure et aussi à la dispersion intra-FPGA des performances.

Variations inter/intra FPGA

Dans cette partie, nous calculons la variation de la période d'oscillation d'un RO entre les neuf FPGA (inter FPGA) et dans un même FPGA (intra FPGA). Nous comparons les résultats obtenus à partir des mesures et des données extraites de Vivado.

Variations inter FPGA

Dans les neuf FPGAs, l'implémentation physique de chacun des 567 ROs est identique. Pour les 567 ROs on regroupe la période mesurée dans les neuf FPGAs, nous avons donc 567 groupes avec dans chaque groupe neuf périodes, d'un RO de même architecture physique, mesurées dans neuf FPGAs différents. Pour chaque groupe nous calculons la dispersion inter FPGA :

$$v_{\text{inter}} = \frac{T_{\text{RO,slow,inter}} - T_{\text{RO,fast,inter}}}{T_{\text{RO,average,inter}}} \quad (3.4)$$

où v_{inter} est la variation inter FPGA pour un RO répété dans les neuf FPGA, $T_{\text{RO,slow}}$ est la période maximale dans un groupe de neuf ROs de même architecture physique mesurés dans neuf FPGAs différents, $T_{\text{RO,fast}}$ est la période minimale ce groupe et $T_{\text{RO,average}}$ est la période moyenne du groupe.

Afin d'extraire la variation inter FPGA attendue par Vivado, nous avons calculé v_{inter} pour la période des 567 ROs en utilisant les données Fast Max et Slow Max. Sur la figure 3.6 nous traçons la variation inter FPGA obtenu pour chacun des 567 ROs sous forme d'histogramme en séparant les RO SP des RO LP. On constate alors que la variation inter FPGA calculé à partir des mesures sont presque dix fois moins importantes de celle attendue par Vivado. De plus, la figure 3.6 montre que la variation inter FPGA est plus importante pour les RO SP que pour les RO LP. Ce dernier résultat

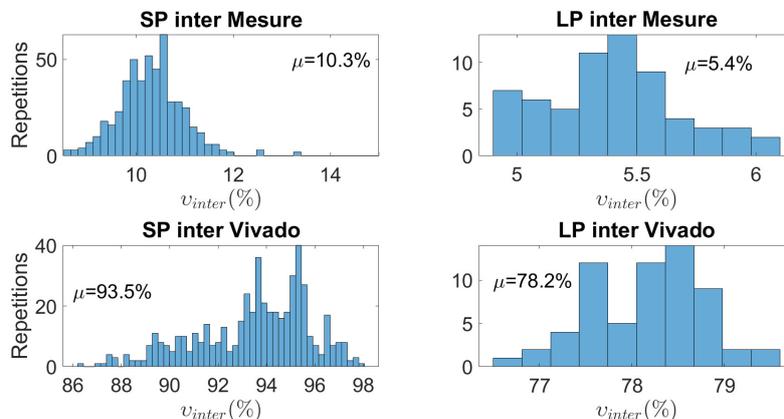


FIGURE 3.6 – Variation des performances inter-FPGA mesurée (en haut) et extraite depuis Vivado (en bas) pour des ROs SP (à gauche) et des ROs LP (à droite)

est obtenu aussi bien à partir de la mesure que des données Vivado ce qui conforte la cohérence de notre analyse.

Variations intra FPGA

Parce que Vivado gère l’implémentation des ROs dans le FPGA, deux ROs de même architecture logique (VHDL) peuvent avoir des ressources différentes. Afin d’obtenir les variations de performance intra FPGA, nous avons sélectionné des groupes de ROs utilisant exactement les mêmes ressources, c’est à dire les mêmes types de ressources de routage et les LUTs de même configurations utilisant les mêmes entrées. Dans chaque FPGA, nous obtenons neuf ROs dont l’architecture physique se répète au moins deux fois. Pour les neuf ROs nous regroupons les mesures de période des ROs de même architecture physique dans un même FPGA, nous obtenons ainsi neuf groupes par FPGA soit 81 groupes pour les neuf FPGA. Pour chaque groupe nous calculons la dispersion intra FPGA :

$$v_{intra} = \frac{T_{RO,max} - T_{RO,min}}{T_{RO,average}} \tag{3.5}$$

où v_{intra} est la variation intra FPGA dans un groupe de ROs d’architecture physique identique dans un même FPGA, $T_{RO,max}$ est la période maximale de ce groupe, $T_{RO,min}$ est la période minimale du groupe et $T_{RO,average}$ est la période moyenne du groupe.

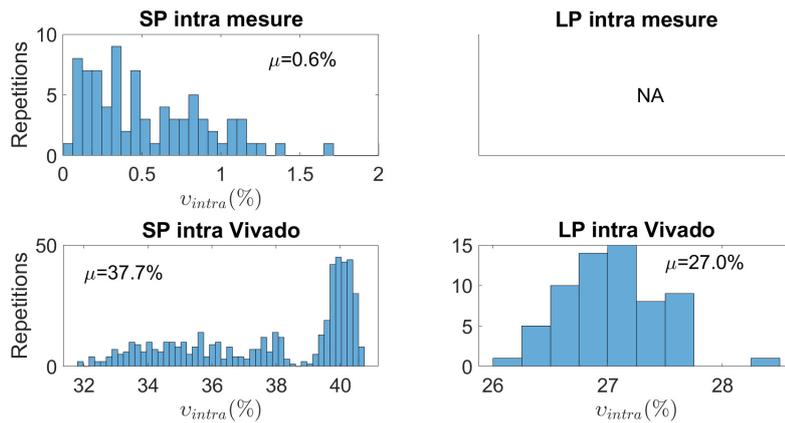


FIGURE 3.7 – Variation des performances intra FPGA (haut) mesuré pour des ROs SP et (bas) extrait depuis Vivado (gauche) pour des ROs SP et (droite) de RO LP

Afin d’extraire la variation intra FPGA attendue par Vivado, nous avons calculé v_{intra} pour la période des 567 ROs en utilisant les données Fast Max et Fast Min. Sur la figure 3.7 nous traçons la variation intra FPGA obtenu pour chacun des 567 ROs sous forme d’histogramme en séparant les RO SP des RO LP. Dans les FPGAs, nous n’avons pas de RO LP dont l’architecture physique se répète. C’est pourquoi, sur figure 3.7 nous n’avons pas de données pour les ROs LP intra mesure. On constate que la variation intra FPGA calculé à partir des mesures sont là aussi comme les dispersions intra FPGA biens moins importantes de celle attendue par Vivado.

Comparaison des résultats avec [152]

La table 3.1 visualise la moyenne des variations intra et inter FPGA obtenues avec la mesure et Vivado. Les variations calculées à partir des données Vivado représentent les limites maximales. On constate un écart important entre les variations attendues par Vivado et celles obtenues par la mesure. L’étude de variation réalisée sur une technologie similaire [152] met en évidence une variation inter FPGA entre 6% et 9% en fonction de l’architecture du circuit mesuré, ce qui est cohérent avec notre résultat qui est entre 5% et 11%. Concernant les variations intra FPGA, nous obtenons une variation 10 fois moins importante. En effet, dans [152] environ 20000 ROs sont répartis dans toutes les régions du FPGA, contrairement à notre étude de variation intra FPGA où

seulement quelques FPGAs utilisant les mêmes ressources sont utilisés.

TABLE 3.1 – Résultats des variations de performances obtenus avec la mesure et extrait depuis Vivado

Variation	Catégories RO	Mesure	Vivado
Intra FPGA	Short Path	0.6%	37.7%
	Long Path	NA	27.0%
Inter FPGA	Short Path	10.3%	93.5%
	Long Path	5.4%	78.2%

En conclusion, dans cette partie nous avons montré comment calculer les variations inter et intra FPGA à partir de la période d'oscillation de ROs. Aussi, en comparant les résultats obtenus à partir des mesures de ceux obtenus à partir des données Vivado, on constate une surestimation de Vivado d'un facteur dix pour la variation inter FPGA et encore plus importante pour la variation intra FPGA.

3.3.3 Rapport cyclique des ROs

Le banc de test que nous avons conçu mesure non seulement la dégradation de la fréquence d'oscillation des ROs mais aussi son rapport cyclique. Comme nous avons détaillé dans la section 2.4.1, à partir du rapport cyclique et de la fréquence du signal de sortie du RO nous pouvons calculer le temps de propagation d'un front montant et descendant dans un RO. Dans cette partie, nous présentons les mesures initiales du rapport cyclique de tous les ROs. Sur la figure 3.8 nous avons tracé le rapport cyclique initial, chaque point représente le rapport cyclique mesuré sur un RO. Chaque graphique correspond à une architecture logique de RO. Parce que la précision du résultat du temps de propagation d'un front montant ou descendant est plus précis à partir des RO de type buffer (voir section 2.4.1), sur la figure 3.8 nous avons représenté seulement les architectures RO de type buffer. Chaque ligne de couleur correspond aux mesures faites dans un FPGA, l'architecture implémentée est la même dans tous les FPGAs, donc chaque RO a la même architecture physique dans chacun des neuf FPGAs.

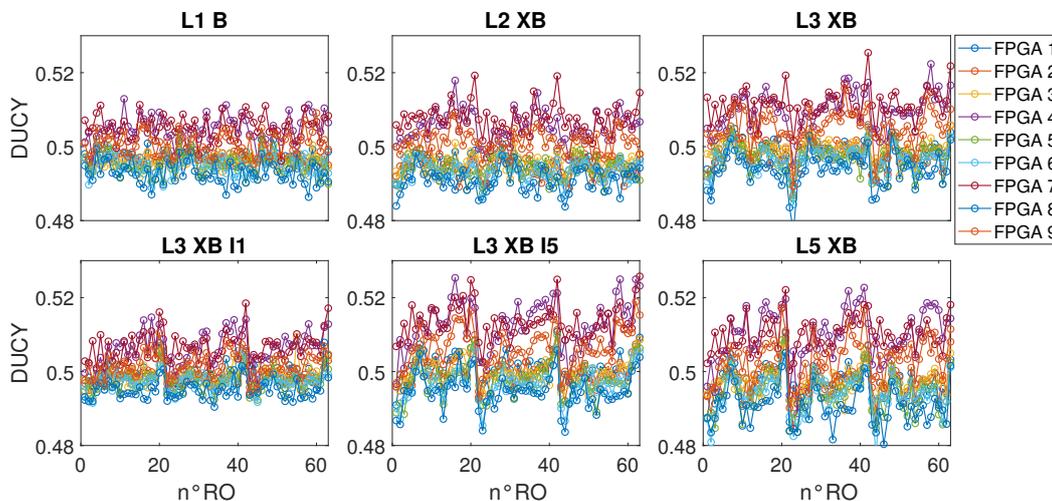


FIGURE 3.8 – Rapport cyclique des ROs mesurés dans les neuf FPGAs ($T = 25^\circ\text{C}$ et $V = V_{nom}$) et pour les architectures de RO de type buffer

On constate que le rapport cyclique est centré autour de 50%. La moyenne du rapport cyclique de toutes les architectures de RO et de tous les FPGAs confondus est de $\mu(\text{DUCY}) = 49.98\%$. Le temps de propagation d'un front montant et d'un front descendant dans les LUTs est donc initialement identique. Par ailleurs, on observe de la dispersion entre les FPGAs. Par exemple, le rapport cyclique du FPGA1 est le plus bas tandis que le rapport cyclique du FPGA4 est le plus haut. Ces différences sont certainement liées à une dispersion lors de l'équilibrage des NMOS et PMOS pendant la fabrication du FPGA.

3.3.4 Caractérisation et modélisation en température et tension des ROs

Dans la section 2.7, lors de la campagne de vieillissement préliminaire effectuée sur un seul FPGA, nous avons modélisé la fréquence et le rapport cyclique d'un RO en fonction de la température et de la tension interne du FPGA. À partir du modèle d'un seul RO, nous avons présenté la réalisation de compensations post-mesures. Avant de commencer le vieillissement des neuf FPGAs, nous avons caractérisé tous les ROs dans tous les FPGAs. Comme pour le test préliminaire, nous avons modélisé la fréquence des ROs avec un polynôme à deux variables (température et tension) du premier ordre pour la température et du troisième ordre pour la tension. Pour le rapport cyclique du RO, nous avons utilisé un polynôme du deuxième ordre pour la température et du quatrième ordre pour la tension. Nous avons introduit la forme du modèle dans l'équation (2.50).

L'étude réalisée dans [2] utilise un modèle de fréquence et de rapport cyclique commun à tous les ROs pour la compensation post-mesure afin de simplifier le traitement des mesures. L'objectif de cette thèse est d'obtenir des mesures de dégradation avec la plus grande précision. Aussi, on doit s'interroger sur la nécessité d'utiliser un modèle par RO pour avoir une compensation de la fréquence et du rapport cyclique avec la meilleure précision. Sur la figure 3.9, nous avons tracé les quatre coefficients les plus représentatifs du modèle de fréquence (voir équation (2.50)) extraits pour chaque RO. Cette figure montre la variation des coefficients en fonction de l'architecture du RO.

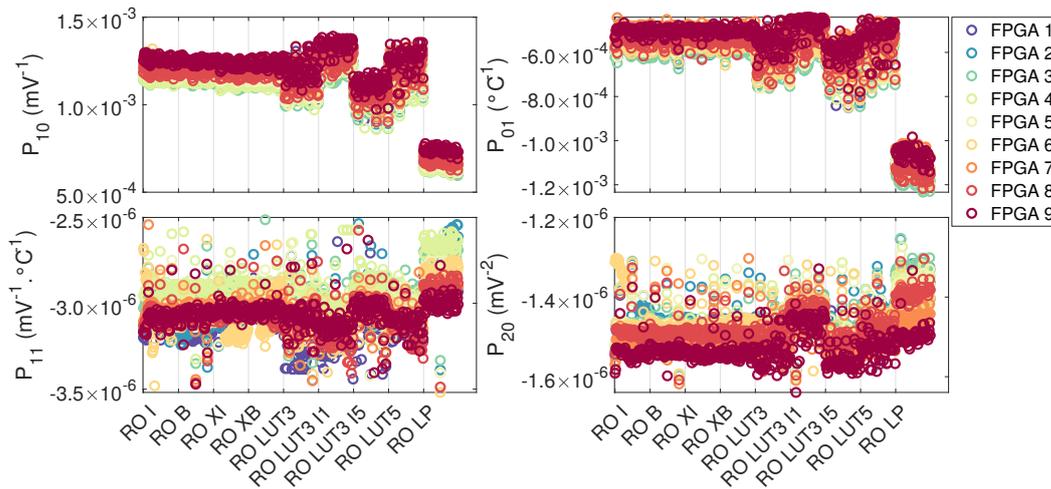


FIGURE 3.9 – Coefficients des modèles de caractérisation de la fréquence des ROs en fonction de la tension et de la température interne du FPGA en fonction de l'architecture du RO et pour différents FPGAs

Pour les ROs LP, les coefficients P10 et P01, qui sont respectivement les facteurs du premier ordre pour la tension et la température, sont bien inférieurs aux autres architectures. Les ROs LP sont majoritairement composés de ressources de routages qui sont visiblement moins sensibles aux variations de tension et de température. La figure 3.9 montre aussi une dépendance des coefficients au FPGA dans lequel le RO évolue. Donc, la figure 3.9 manifeste très clairement que les co-

efficients du modèle de la fréquence du RO dépendent de son architecture ainsi que du FPGA dans lequel il est implémenté. Ce même constat vaut pour les coefficients de la modélisation du rapport cyclique. Finalement, sur la figure 3.9 on remarque une dispersion notable des coefficients pour des ROs de même architecture physique mesurés dans différents FPGA. D'ailleurs, graphiquement on constate que cette dispersion est plus importante que la dispersion entre des ROs d'architecture logique identique mais d'architecture physique différente et mesurés dans un même FPGA.

Bien que la figure 3.9 dévoile déjà l'intérêt d'utiliser un modèle par RO, nous allons en plus comparer la prédiction de la différence de fréquence Δf_{calcul} en fonction de la mesure de différence de fréquence Δf_{mesure} définies par :

$$\Delta f_{mesure} = f_{mesure}(V, T) - f_{mesure}(V_{nom}, T_{nom}) \quad (3.6)$$

$$\Delta f_{calcul} = f_{modele}(V, T) - f_{modele}(V_{nom}, T_{nom}) \quad (3.7)$$

avec $T_{nom} = 25^\circ\text{C}$. Pour f_{mesure} on récupère la fréquence mesurée dans chacune des conditions de température et de tension. Pour f_{calcul} on utilise un modèle qui nous permet, à partir de la tension et de la température, de calculer la fréquence. Pour chaque RO, nous utilisons les mesures de la caractérisation de la fréquence avec une tension qui varie entre 650mV et 1100mV et une température entre 25°C et 100°C (voir figure 2.23).

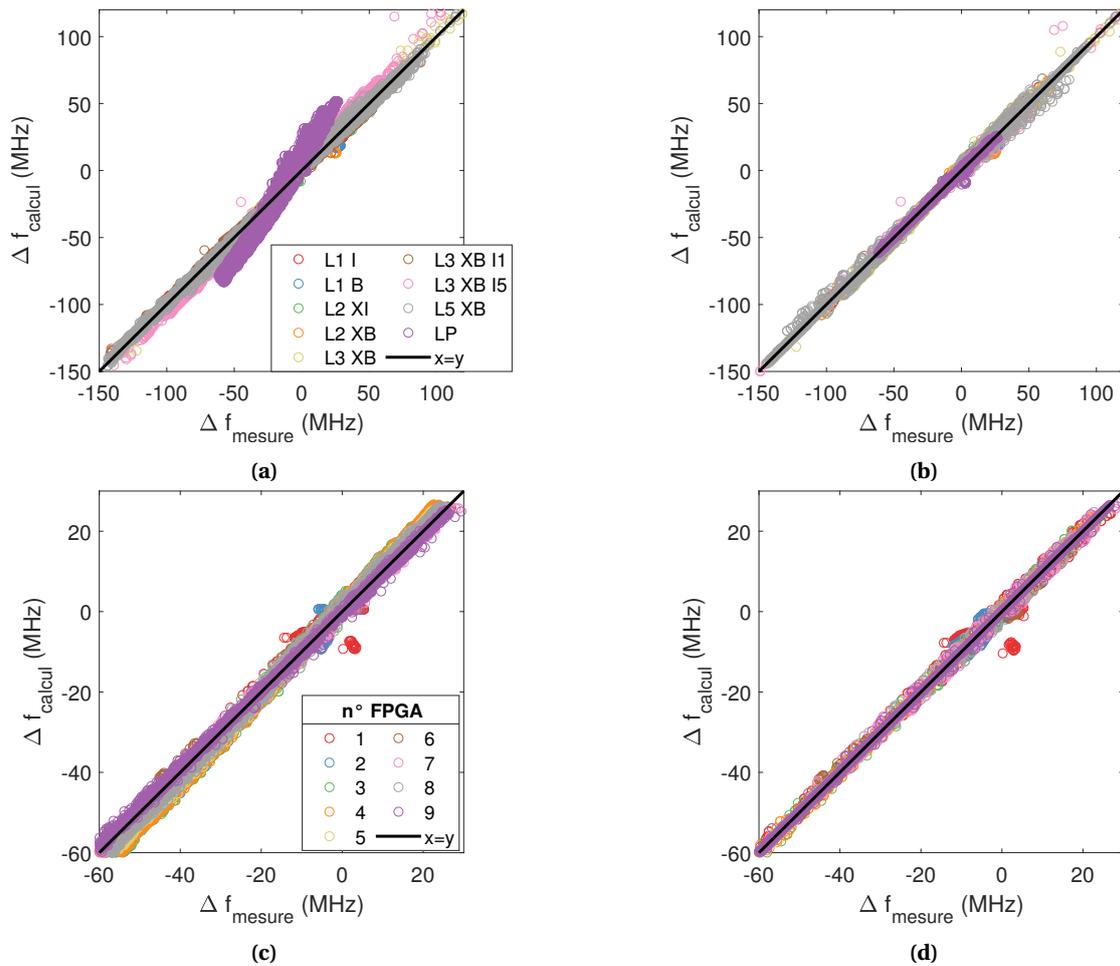


FIGURE 3.10 – Comparaison entre la mesure et la prédiction de fréquence faite sur (haut) un FPGA et toutes les architectures de RO (a) avec un modèle commun (b) avec un modèle par RO (bas) sur tous les FPGA et une architecture de RO (c) avec un modèle commun (d) avec un modèle par RO. La température varie entre 25°C et 100°C et la tension varie entre $0.75V_{nom}$ et $1.3V_{nom}$.

La figure 3.10a compare la fréquence mesurée et prédite en utilisant un modèle commun à tous les RO dans un seul FPGA. Le modèle commun est celui d'un RO LI I dans le FPGA1. On observe des différences significatives entre la mesure et la prédiction de la fréquence pour l'architecture RO LP. En effet, nous avons observé, sur la figure 3.9, un écart conséquent des coefficients du modèle de fréquence entre les RO SP (Short Path) et les RO LP. En utilisant un modèle commun à tous les ROs d'architectures différentes, nous obtenons un coefficient de corrélation de 0.9871. La figure 3.10b compare la fréquence mesurée et prédite en utilisant un modèle par RO et montre, avec un coefficient de corrélation de 0.9913, une meilleure cohérence entre la prédiction et la mesure qu'en utilisant un modèle commun.

La figure 3.10c compare la fréquence mesurée et prédite en utilisant un modèle commun pour des ROs de même architecture (LP) implémenté dans différents FPGAs. Le modèle commun est celui d'un RO LP dans le FPGA1. On observe de la dispersion entre les FPGA qui correspond à la dispersion entre les coefficients observés sur la figure 3.9. En utilisant un modèle commun à tous les ROs implémentés dans différents FPGA, nous obtenons un coefficient de corrélation de 0.9929. La figure 3.10d compare la fréquence mesurée et prédite en utilisant modèle par RO et nous observons moins de dispersion ainsi qu'un meilleur coefficient de corrélation 0.9935.

Comme nous venons de voir, pour appliquer une compensation post-mesure avec précision il est important d'avoir un modèle propre à chaque RO. En effet, l'impact de la tension et de la température sur la fréquence et le rapport cyclique des ROs dépend de son architecture et du FPGA. Donc, nous utiliserons un modèle propre à chaque RO pour la compensation post-mesure.

3.3.5 Conclusion

Dans cette section, nous avons présenté les périodes d'oscillations initiales de tous les RO dans les neuf FPGAs. Nous avons observé des dispersions importantes entre des ROs d'architecture logique (VHDL) identique. Ces dispersions montrent que le choix des ressources (entrée LUT, routage) pour l'implémentation d'un RO a un impact plus important que la fonctionnalité (configuration LUT) du RO. Lors de l'étude de fiabilité, il sera important de considérer l'architecture interne du RO et pas seulement son architecture logique. Nous avons profité des mesures de périodes des ROs aux conditions nominales pour les comparer avec les données Vivado. Nous observons une cohérence entre les mesures et les références Fast Max et Slow Min, ce qui nous conforte sur la capacité du banc de test à mesurer la période des ROs. Finalement, la mesure des variations intra/inter FPGA montre que Vivado surestime la dispersion des performances. La mesure initiale du rapport cyclique des ROs manifeste que le temps de propagation d'un front montant et descendant dans les LUTs est équilibré ($\pm 2\%$) et cela quelle que soit sa configuration. Enfin, nous avons vu l'importance d'utiliser un modèle de fréquence et de rapport cyclique des ROs propre à chaque RO afin d'obtenir une compensation post-mesure précise et donc des mesures de dégradations de meilleure qualité.

3.4 Analyse préliminaire des dégradations des ROs

3.4.1 Introduction

Le banc de test que nous avons développé et que l'on voit sur la photo 3.2 vieillit et mesure la dégradation des FPGAs. Les mécanismes de dégradation (BTI, HCI et TDDB) présents dans les circuits numériques ont pour conséquence d'augmenter le temps de propagation d'un signal dans le circuit. Pour mesurer la dégradation d'un FPGA, nous avons implémenté des ROs dont nous mesurons la fréquence et le rapport cyclique d'oscillation qui dépendent du temps de propagation du signal dans les ressources de LUT et de routages qui constituent un RO dans un FPGA. Pour étudier l'effet de la température et de la tension sur la fiabilité des FPGAs, notre banc de test vieillit neuf FPGAs chacun soumis à un stress de température et de tension différent. Mis à part

cette différence, l'architecture implémentée dans les neuf FPGAs est identique. Pour étudier l'effet du signal circulant dans un FPGA sur sa fiabilité, chacun des ROs est en mode stress plus de 99% du temps et reçoit en entrée un signal de stress interne statique ou dynamique avec une certaine fréquence et un certain rapport cyclique. Pour étudier la fiabilité pour différentes configurations du FPGA, nous avons implémenté différentes architectures de ROs.

Après avoir fait vieillir 5103 ROs implémentés dans neuf FPGAs pendant 8000 heures, dans cette section nous présentons les résultats. Premièrement nous présentons les résultats des 5103 ROs afin d'avoir une vue d'ensemble de la dégradation de tous les ROs dans les neuf FPGAs. Ensuite, nous présenterons les dégradations des RO en mettant en avant l'effet de la tension et de la température sur le vieillissement. Aussi, nous analyserons l'effet de la fréquence et du rapport cyclique du signal de stress interne. Finalement, nous analyserons les dégradations pour différentes architectures de ROs.

3.4.2 Résultats des 5103 ROs

Afin d'obtenir des courbes avec une bonne lisibilité et des modèles représentatifs des dégradations, sur les résultats que nous présenterons dans la suite de ce manuscrit, nous appliquons le traitement sur les données comme suit :

- Filtre : On élimine les données aberrantes qui correspondent à moins de 0.1% de toutes les données;
- Compensation : Voir section 2.7;
- Moyenne : Chaque donnée correspond à une moyenne de dix mesures, soit une moyenne sur 10 heures.

Avant d'étudier au cas par cas l'influence des différents stress sur le vieillissement, la figure 3.11 expose la dérive de la fréquence relative (Δf_R) en fonction du temps de tous les ROs, soit 567 ROs, dans chacun des FPGA, où :

$$\Delta f_R = \frac{f(t) - f(t_0)}{f(t_0)} \quad (3.8)$$

avec $f(t_0)$ la fréquence moyenne sur les 30 premières heures de vieillissement. Dans un premier temps nous utilisons Δf_R pour représenter les dégradations, indépendamment des ressources utilisées dans le RO. En effet, les dégradations absolues dépendent du nombre de ressources vieilles dans le RO et, donc de sa fréquence initiale. La variation de la fréquence relative du RO représente l'augmentation moyenne de la tension de seuil des transistors vieillis dans le RO.

La figure 3.11 montre des dégradations qui sont comprises entre 0% et 1%. L'impact de la température et de la tension sur les dégradations est clair. Grâce à la précision de mesure du banc de test nous observons des dégradations, inférieures à 0.2%, dans des conditions de stress identiques aux conditions d'utilisation : V_{nom} et 25°C. On constate des variations de dégradation importantes pour une même condition de stress en température et tension. Par exemple, pour $T_{stress} = 115^\circ\text{C}$ et $V_{stress} = 1.15V_{nom}$, il y a un facteur quatre entre le RO le moins dégradé et le plus dégradé. Ces différences de dégradations manifestent l'impact important de la configuration interne du FPGA (fréquence et rapport cyclique du stress, architecture RO) sur le vieillissement.

Pendant le test de vieillissement, autour de 2000 heures, les cartes de développement des FPGAs 100°C & 1.2 V_{nom} et 115°C & 1.15 V_{nom} ont cessé de fonctionner et autour de 7500 heures de fonctionnement la carte du FPGA 100°C & 1.1 V_{nom} a cessé de fonctionner. Nous ne connaissons pas la cause de la panne, cependant, ce sont les FPGAs avec les conditions de stress en température et en tension les plus importantes. Avant la panne, nous n'avons pas observé de comportement particulier sur les ROs. La cause provient certainement d'un composant présent sur la carte de développement qui est externe au FPGA.

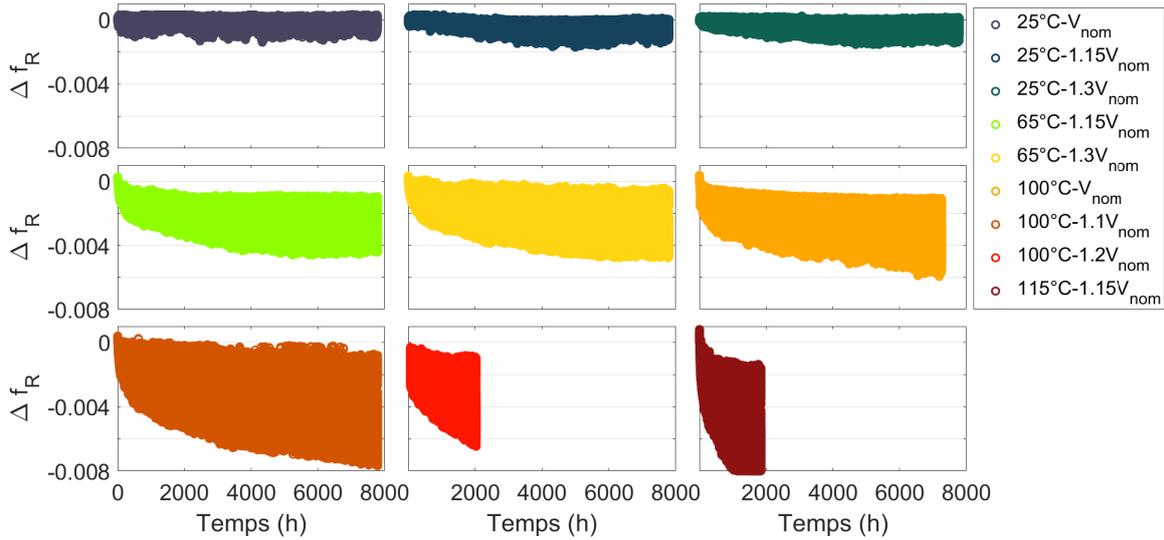


FIGURE 3.11 – Dérive fréquentielle relative de tous les ROs dans les neuf FPGAs chacun avec une condition de stress en température et tension différente

Nous avons cherché la présence de TDDB dans le comportement des ROs. Nous considérons qu'en fonction du mode de défaillance du TDDB au niveau d'un transistor (voir section 1.5.4), il peut y avoir plusieurs modes de défaillances au niveau du RO :

- Fréquence du RO nulle : cas du hard breakdown;
- Saut de fréquence spontané : cas du hard breakdown;
- Changement de comportement de la dérive fréquentielle du RO : cas du soft breakdown;

Nous n'avons pas observé ces modes de défaillance dans les ROs. Donc, nous n'avons pas observé de TDDB dans notre test de vieillissement.

3.4.3 Effet de la tension et de la température

Afin d'étudier l'effet de la tension de stress sur les dégradations, nous traçons sur la figure 3.12 les dérives fréquentielles de ROs avec une même température (100°C) de stress et trois tensions de stress différentes (V_{nom} , $1.1V_{nom}$ et $1.2V_{nom}$). Les dérives fréquentielles des ROs Short Path (SP) pour toutes les conditions de stress internes sont tracées, soit 504 ROs. Nous avons tracé les données en échelle linéaire et logarithmique. Sur l'échelle logarithmique, nous traçons la moyenne de tous les ROs avec des bornes qui représentent l'écart type.

À partir de la dérive moyenne de tous les ROs, nous avons modélisé le comportement des dégradations avec une loi puissance, tel que :

$$\Delta f_R(t) = at^b \quad (3.9)$$

où a est le facteur au temps et représente l'amplitude des dégradation tandis que b est l'exposant au temps et représente la dynamique de dégradation. Sur la figure 3.12 nous voyons que l'exposant au temps se situe entre 0.24 et 0.29, ce qui est proche de l'exposant au temps qui caractérise le modèle de Reaction-Diffusion pour le BTI (0.25, voir section 1.5.2). Par ailleurs, nous analysons les dégradations pour un stress en haute température ($T_{stress} = 100^\circ\text{C}$) ce qui favorise le BTI et nous suggère que nous observons du BTI.

Afin d'observer l'effet de la température sur le vieillissement, nous avons tracé sur la figure 3.13 les dérives fréquentielles de ROs stressés par une même tension ($1.15V_{nom}$) et trois températures

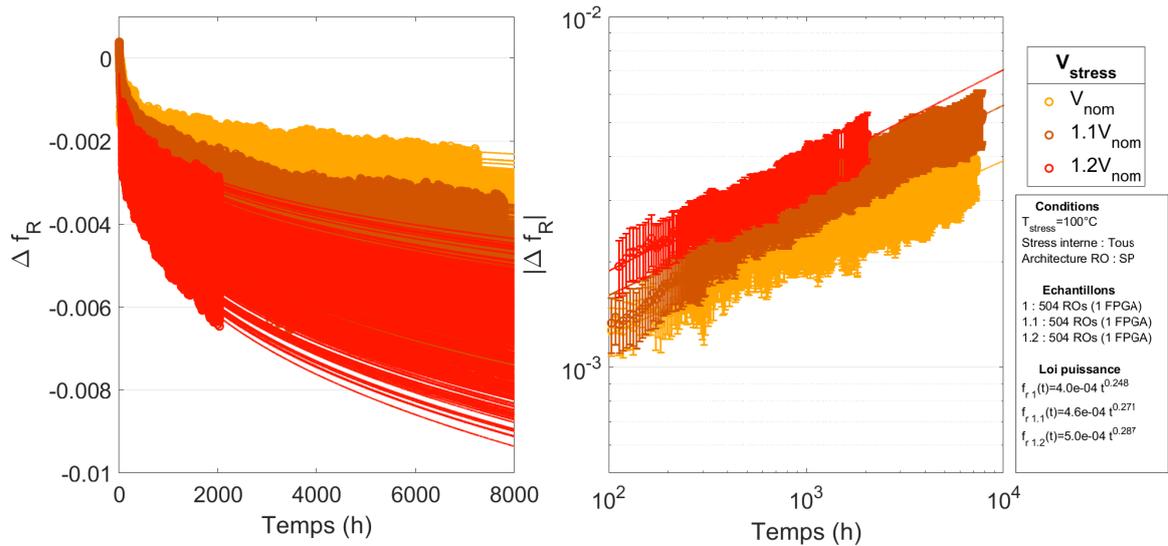


FIGURE 3.12 – Effet de la tension de stress sur la variation de la dérive fréquentielle relative des ROs SP (gauche) en échelle linéaire (droite) en échelle log-log. Sur le graphique de droite nous avons tracé la moyenne des mesures des ROs. Sur les deux graphique les points sont les mesures et les lignes sont le modèle de puissance identifiés.

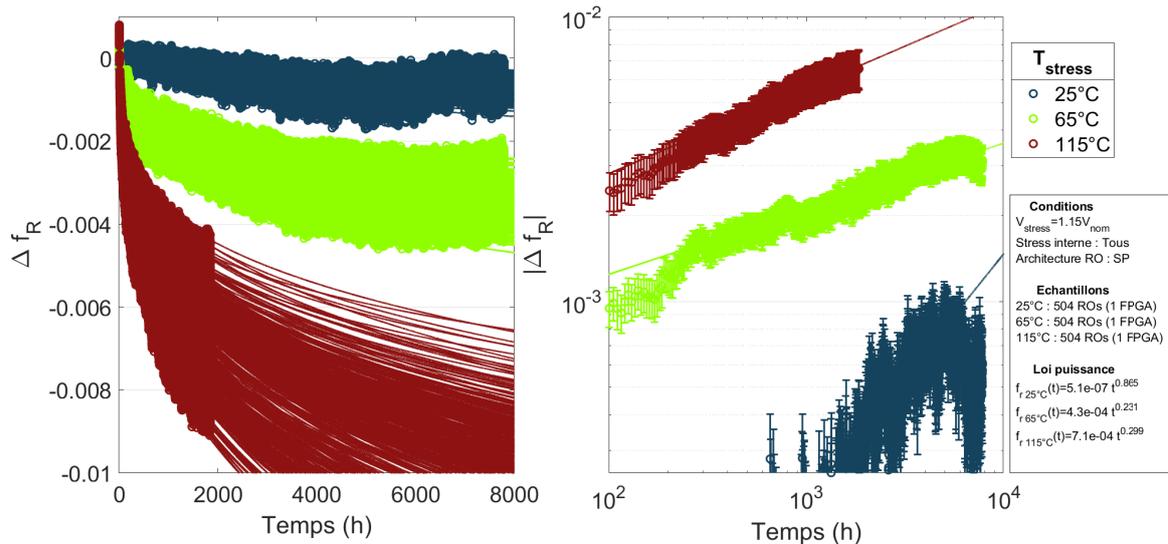


FIGURE 3.13 – Effet de la température de stress sur la variation de la dérive fréquentielle relative des ROs (droite) en échelle linéaire (gauche) en échelle log-log. Sur le graphique de droite nous avons tracé la moyenne des mesures des ROs. Sur les deux graphique les points sont les mesures et les lignes sont le modèle de puissance identifiés pour chaque condition.

(25°C, 65°C et 115°C). Pour les mesures à 25°C on constate après 6000 heures que la dérive de fréquence relative augmente légèrement en passant de -7.5×10^{-4} (6000 heures) à -3.7×10^{-4} (7000 heures). Cette augmentation est probablement due à une variation de l'environnement (température ou tension). Cependant, les mesures de la température et de la tension interne du FPGA ne révèlent pas de dérive. Alors, cela nous laisse penser que l'instrumentation dans le SYSMON qui mesure la tension et la température interne du FPGA à légèrement dérivée.

Pour 65°C et 115°C nous obtenons des exposants au temps, respectivement de 0.23 et 0.3, proches de ceux obtenus précédemment à 100°C. Cela nous laisse à penser qu'encore une fois, dans ces conditions de stress en température, le BTI est le principal mécanisme de dégradation. Cependant, à 25°C nous obtenons un exposant au temps de 0.86 soit bien supérieur à celui de

Tension	Effet tension			Effet température	
	V_{nom}	$1.1V_{nom}$	$1.2V_{nom}$	$1.15V_{nom}$	
Température	100°C			65°C	115°C
$a(\times 10^{-4})$	2.4	3.7	4.9	2.9	8.1
b	0.29				
$\Delta f_R(10000h)$	0.0035	0.0053	0.0071	0.0042	0.0117

TABLE 3.2 – Coefficients du modèle de puissance pour quantifier l’impact de la tension et de la température de stress sur le vieillissement

BTI. Dans le modèle de l’électron chanceux du HCI (voir section 1.5.3) l’exposant au temps se situe entre 0.5 et 1. De plus une des conditions favorables pour observer du HCI est une température basse ce qui est notre cas, cela suggère que nous observons du HCI.

Afin de quantifier l’impact de la tension et de la température sur le vieillissement, nous avons modélisé de nouveau les dégradations avec une loi de puissance mais en fixant l’exposant à $b = 0.29$ pour les trois conditions de la figure 3.12 (100°C) et pour les conditions à 65°C et 115°C de la figure 3.13. En effet, nous considérons que le mécanisme de dégradation est le BTI et qu’il est commun pour les températures allant de 65°C à 115°C. En fixant l’exposant au temps, nous pouvons comparer le facteur au temps a que nous notons sur le tableau 3.2.

Pour une température de stress à 100°C nous obtenons un facteur d’accélération du vieillissement de 2 entre les vieillissements avec une tension de stress V_{nom} et à $1.2V_{nom}$. Pour une tension de stress à $1.15V_{nom}$ nous obtenons un facteur d’accélération du vieillissement de 2.5 entre les vieillissements avec une température de stress à 65°C et à 115°C.

3.4.4 Effet de la fréquence de stress

Le HCI est un mécanisme de dégradation dépendant de la fréquence de stress. En effet, le HCI est généré lorsque du courant passe dans le transistor. Dans une LUT, les transistors conduisent du courant lorsque le niveau logique en sortie change. Donc, dans une LUT le HCI sera plus important lorsque l’entrée de la LUT sera connectée à un signal de stress de haute fréquence. Au contraire, on peut considérer que les dégradations dues au HCI sont presque nulles lorsqu’il y a un stress statique à l’entrée de la LUT puisque le courant devient négligeable.

Le BTI n’est pas dépendant à la conduite de courant mais au niveau de tension aux bornes du transistor. En considérant qu’une LUT est faite de portes de transmission comme présenté sur la figure 1.19, lorsque le niveau logique d’une entrée change, la porte de transmission qui était passante devient bloquée et une autre porte de transmission qui était bloquée devient passante. Dans le cas d’un stress dynamique, les deux portes de transmission alternent entre un état stressé et non stressé. Lorsqu’un transistor n’est pas dans des conditions de stress du BTI, sa tension de seuil revient vers sa tension initiale, c’est le mécanisme de récupération. On peut s’attendre alors à des dégradations moins importantes dues au BTI dans le cas d’un stress dynamique comparé à un stress statique pour lequel le transistor est en permanence dans un état de stress.

Dans cette partie nous allons étudier l’effet de la fréquence de stress sur les dégradations. Afin de séparer l’effet de la fréquence de stress sur le BTI de celui sur le HCI, nous présenterons les résultats pour des FPGAs avec des stress de températures chaudes ($T_{stress} \geq 65^\circ\text{C}$) un stress de température froide ($T_{stress} = 25^\circ$). Sur la figure 3.14, nous avons tracé les dérivées fréquentielles des

ROs d'architecture L1 B stressés par plusieurs fréquences. Lorsque le stress n'est pas statique, le rapport cyclique est de 0.5. Nous avons répété le tracé pour plusieurs conditions de température et de tension.

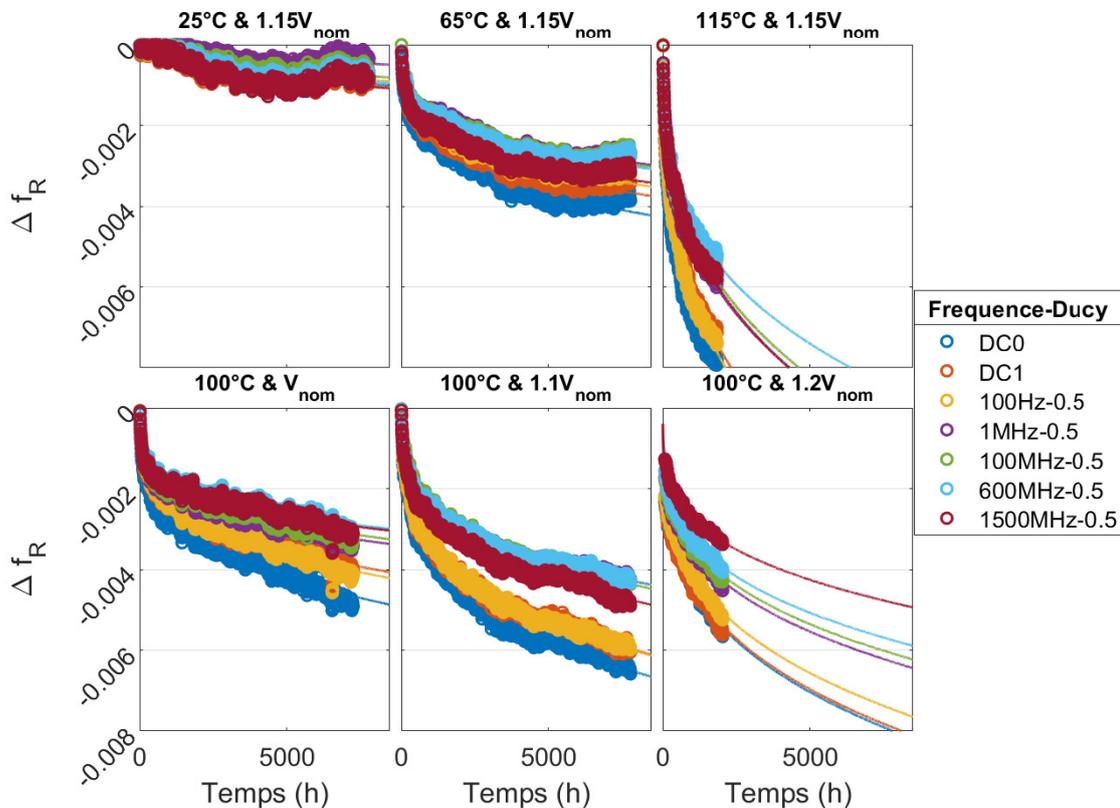


FIGURE 3.14 – Effet de la fréquence de stress sur la variation de la dérive fréquentielle relative des ROs pour différentes conditions de stress de températures et de tensions

Température chaude

Lorsque $T_{stress} \geq 65^\circ\text{C}$, la figure 3.14 révèle une amplitude des dégradations plus importante des ROs pour un stress statique et, plus particulièrement pour un stress statique DC0. La différence d'amplitude de dégradation entre les stress DC0 et DC1 est certainement liée aux transistors PMOS et NMOS qui sont vieillissent. En effet, un stress DC1 correspond à un signal de niveau haut en entrée de la LUT qui est transmis par un nombre de NMOS et PMOS différents à celui de la transmission d'un niveau bas, c'est à dire DC0. Donc, la différence d'amplitude de dégradation entre un stress DC0 et DC1 correspond à la prépondérance du NBTI dans un cas ou du PBTI dans l'autre cas. Nous verrons dans la section 3.5 comment extraire les dégradations du PBTI et du NBTI.

Nous avons expliqué dans la section 1.5.2 que le BTI est composé de deux mécanismes lorsque le stress est dynamique : dégradation et récupération (voir figure 1.26). Le fait que nous obtenions sur la figure 3.14 des dégradations moins importantes avec un stress dynamique suggère la présence du mécanisme de récupération. Par ailleurs, pour les températures 100°C et 115°C nous observons des dégradations plus importantes pour un stress dynamique avec une fréquence à 100Hz que pour les plus hautes fréquences.

Dans la section 1.5.2 nous avons présenté les résultats de l'étude [65, chapitre 21.3.3] consistant à montrer l'effet de la fréquence de stress sur les dégradations dues au BTI à partir de mesures réalisées sur des transistors isolés. Cette étude montre que l'amplitude des dégradations

est plus importante et est dépendante à la fréquence de stress lorsque $f_{stress} < 100\text{Hz}$ et l'amplitude des dégradations est moins importante et est indépendante à la fréquence de stress lorsque $f_{stress} > 100\text{Hz}$ (voir figure 1.27). L'amplitude des dégradations est indépendante à la fréquence de stress lorsque cette même fréquence est largement supérieure à la dynamique des dégradations (τ_θ) et des récupérations du BTI (τ_c). Donc, en observant des dégradations différentes sur la figure 3.15 entre $f_{stress} = 100\text{Hz}$ et $f_{stress} \geq 1\text{MHz}$, cela peut s'interpréter par, dans notre cas : $100\text{Hz} < [1/\tau_\theta, 1/\tau_c] < 1\text{MHz}$.

Finalement, pour les FPGAs stressés à haute température nous observons les symptômes du BTI suivants :

- Dégradations plus importantes pour un stress statique que pour un stress dynamique du au phénomène de récupération ;
- Dépendance des dégradations à la fréquence de stress en basse fréquence et indépendance à la fréquence de stress pour des stress en haute fréquence comme décrit dans [65, chapitre 21.3.3] et [66] ;

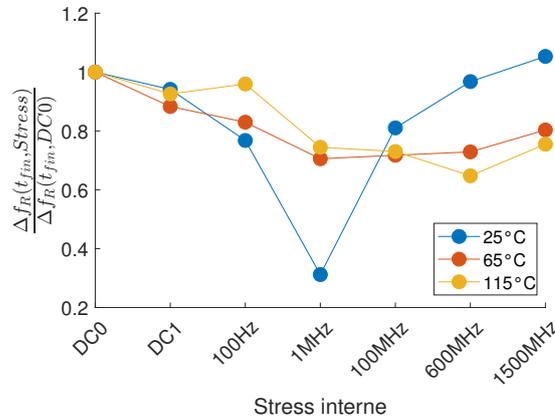


FIGURE 3.15 – Dégradation relative finale par rapport à la dégradation finale pour un stress statique DC0 en fonction de la fréquence de stress, pour plusieurs T_{stress} et $V_{stress} = 1.15V_{nom}$

Température froide

Sur la figure 3.14 lorsque $T_{stress} = 25^\circ\text{C}$ on constate que les dégradations les plus importantes apparaissent lorsque $f_{stress} = 1500\text{MHz}$ et les dégradations les moins importantes lorsque $f_{stress} = 1\text{MHz}$. Par soucis d'échelle, sur la figure 3.14 il est difficile d'observer l'impact des conditions intermédiaires sur les dégradations. Ainsi, nous avons tracé sur la figure 3.15 la dégradation relative finale, par rapport à la dégradation pour DC0, en fonction de la fréquence de stress. À $T_{stress} = 25^\circ\text{C}$, on constate que l'amplitude des dégradations augmente lorsque la fréquence augmente au dessus de 1MHz. Tandis que pour les températures supérieures les dégradations sont relativement stable en 1MHz et 1500MHz. Le comportement à basse fréquence est différent, certainement à cause de la dynamique de récupération et de dégradation du BTI qui devient proche de la fréquence de stress comme déjà expliqué pour les températures chaudes.

Dans la section 1.5.3, nous avons noté que le HCI se produit lorsque du courant circule, c'est à dire que le HCI est dépendant du nombre de commutation et donc de la fréquence. La dépendance des dégradations à la fréquence de stress nous conforte sur l'hypothèse d'observer du HCI.

3.4.5 Effet du rapport cyclique

Sur la figure 3.16 nous avons tracé la dérive fréquentielle en fonction du temps pour des RO L1 B et ce avec plusieurs rapports cycliques de stress. La fréquence de stress est de 100MHz qui

correspond à la fréquence de l'horloge de la carte du FPGA.

Dans les différentes conditions de stress de température et de tension on n'observe pas d'impact significatif du rapport cyclique de stress sur les dégradations de la fréquence relative. Sachant que les dégradations sont plus importantes pour un stress DC0 que pour un stress DC1 on pourrait s'attendre à obtenir des dégradations plus importantes pour un rapport cyclique de 0.1 que pour un rapport cyclique de 0.9 par exemple. Cependant, quelque soit le rapport cyclique on constate que les dégradations sont assez similaires. Nous verrons et modéliserons dans la section 4.3 l'impact du rapport cyclique de stress sur la dégradation du temps de propagation d'un front montant et descendant dans un étage du RO.

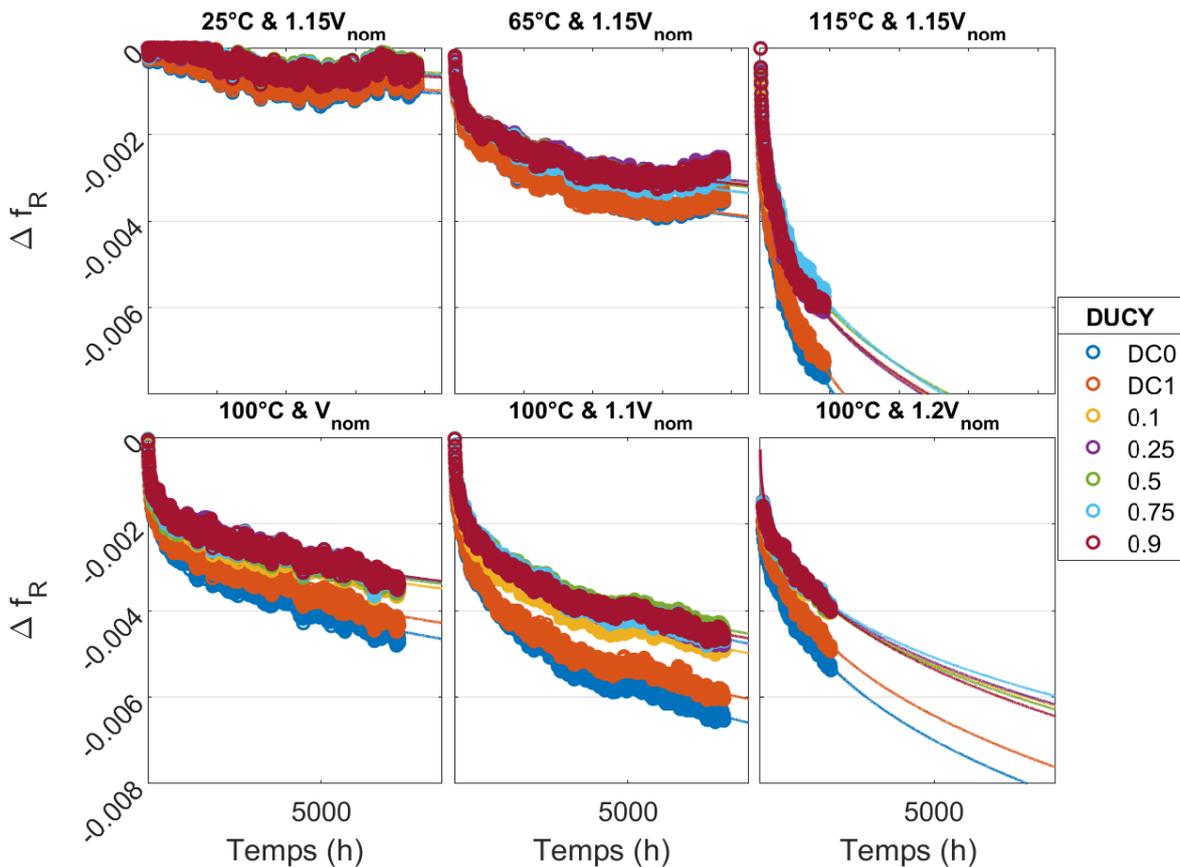


FIGURE 3.16 – Effet du rapport cyclique de stress sur la variation de la dérive fréquentielle relative des ROs pour différentes conditions de températures et de tension et avec une $f_{stress} = 100\text{MHz}$ pour les stress dynamiques.

3.4.6 Effet de l'architecture du RO

Un FPGA est un circuit numérique dont l'architecture interne est configurée pour assurer la fonction requise. Ainsi, un grand nombre de configurations du FPGA existe. Afin d'étudier l'effet du vieillissement sur différentes configurations du FPGA, nous avons implémenté dans le celui-ci différentes architectures de ROs (voir section 2.3) dont nous mesurons les dégradations. Dans cette partie nous analysons l'influence de l'architecture du RO sur le vieillissement. Pour cela nous étudions les quatre points suivants :

- Le type de ressource dans le RO : logique (RO Short Path) ou routage (RO Long Path) ;
- Le nombre d'entrées utilisées de la LUT ;
- Le numéro d'entrée utilisé de la LUT ;
- La configuration de la LUT (fonction) ;

Le type de ressource : RO Short Path (SP) et RO Long Path (LP)

Un RO SP est constitué de 13 LUTs et le placement du RO dans le FPGA est contraint pour que celui-ci soit implémenté dans maximum deux slices voisines. Un RO LP est constitué de deux LUTs et le placement du RO dans le FPGA est contraint pour que celui-ci soit implémenté dans deux slices l'une à l'opposée de l'autre. Ainsi, un RO SP est constitué principalement de ressources logiques (LUT) et un RO LP est constitué principalement de ressources de routage.

Sur la figure 3.17 nous avons tracé les dérives de fréquence relative de RO SP et de RO LP avec $T_{stress} = 100^{\circ}\text{C}$ et $V_{stress} = 1.1V_{nom}$. Huit sous catégories de RO SP (L1 I, L1 B...) et une seule catégorie de RO LP sont implémentées. Chaque courbe correspond à une moyenne des mesures de trois RO. Nous traçons les résultats pour un stress statique DC0 et DC1 car ce sont les conditions de vieillissement les plus aggravantes. La figure 3.17 montre que l'amplitude des dégradations est plus importante pour tous les RO SP par rapport aux RO LP. C'est aussi ce qui est observé dans [156]. Ce constat révèle que dans un FPGA, la dégradation des ressources logiques (LUT) est plus sensible au vieillissement que la dégradation des ressources de routage. Par ailleurs, on constate que, contrairement aux RO SP la dégradation des RO LP est plus sensible à un stress DC1 qu'à un stress DC0. Ne connaissant pas l'architecture interne du routage dans le FPGA, il est difficile d'interpréter ce résultat. Cependant, on peut encore penser que cette différence est liée au type de transistor (NMOS et PMOS) vieillis et donc à la prédominance du PBTI ou du NBTI.

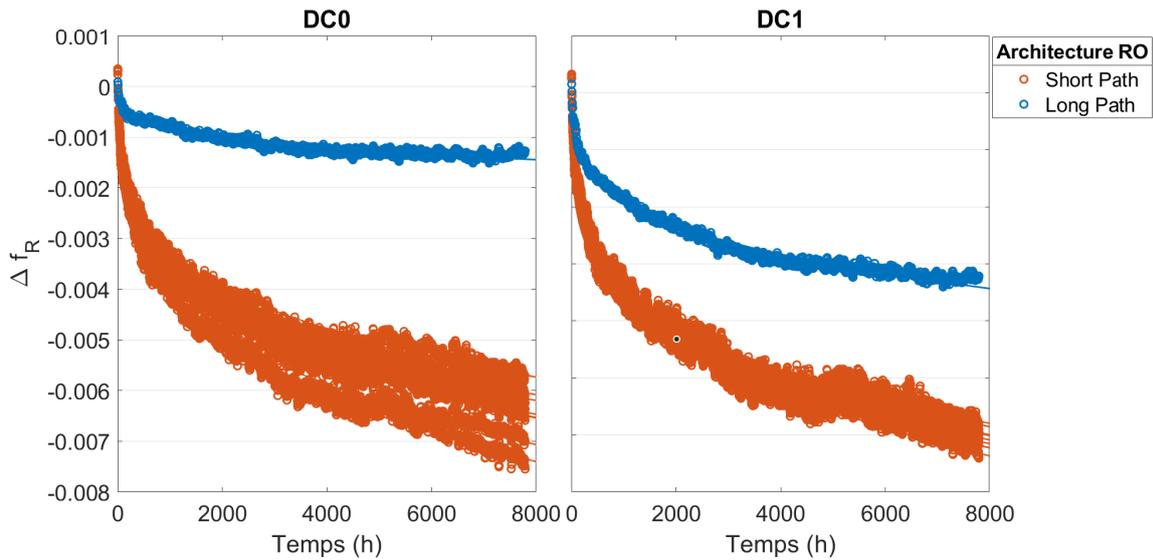


FIGURE 3.17 – Influence du type de ressource dans le RO, routage (LP) ou logique (SP), sur la dérive fréquentielle relative des ROs en stress statique DC0 et DC1

Le nombre d'entrées de la LUT utilisées

Le FPGA que nous vieillissons contient des LUTs à six entrées, cependant, suivant la conception souhaitée par l'utilisateur toutes les entrées des LUTs ne sont pas utilisées. Dans cette partie nous étudions l'effet du nombre d'entrées utilisées sur le vieillissement d'une LUT. Pour cela nous étudions l'effet du nombre d'entrées lorsque la LUT reçoit un stress statique et lorsqu'elle reçoit un stress dynamique.

Lorsque la LUT reçoit un stress statique en entrée, à priori, le nombre d'entrées utilisées n'a pas d'effet sur le vieillissement. Pour expliquer cela, nous introduisons la figure 3.18 qui représente le schéma simplifié d'une LUT à 4 entrées. Sur ce schéma on considère qu'une porte de transfert est passante si elle est commandée par un '1'. Sur la figure 3.18a, les 4 entrées sont utilisées. Lorsque le RO est en mode mesure, les entrées I1, I2 et I4 sont fixées à '0' pour que la LUT fonctionne comme

un buffer. En mode stress, les entrées I1, I2 et I4 sont stressées par un '0' statique. Alors le signal du RO se propage dans 6 portes de transmissions et 3 inverseurs qui ont été dégradé par le signal de stress. Sur la figure 3.18b nous considérons seulement 2 entrées de la LUT utilisées (I3 et I4). Les autres entrées sont fixées par défaut à '1'. En réalité dans le FPGA nous ne savons pas si les entrées non utilisées sont fixées à '1' ou '0', ce qui ne change rien dans le raisonnement de notre exemple. On constate que le signal du RO ne se propage plus dans le même chemin. Cependant, les entrées non utilisées sont fixes, donc, les portes de transmissions stressées sont les mêmes que celle qui transmettent le signal du RO. Finalement, on retrouve le même nombre de portes logiques dégradées par un stress statique quel que soit le nombre d'entrées utilisées dans la LUT.

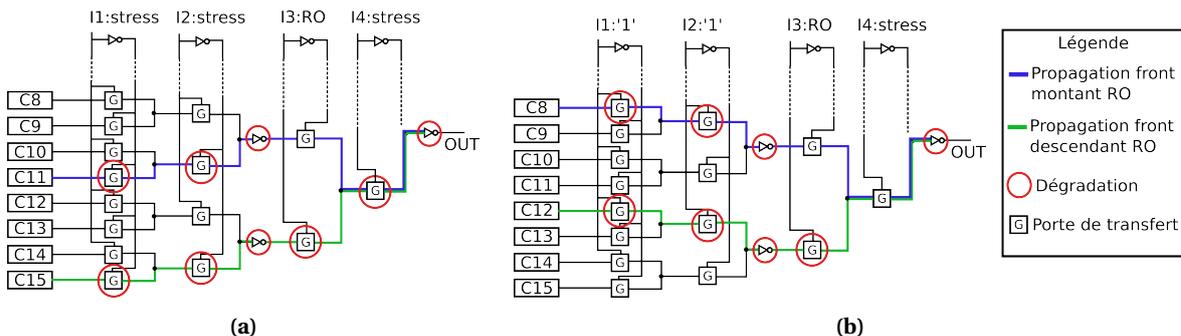


FIGURE 3.18 – Schéma simplifié d'une LUT à 4 entrées avec l'identification du chemin de propagation d'un niveau haut (bleu) et bas (vert) et des portes logiques vieilles lorsque (a) les 4 entrées de la LUT sont utilisées (b) 2 entrées de la LUT sont utilisées et les entrées non utilisées sont fixées par défaut à un niveau logique haut. Lorsque la LUT est en mode mesure, les entrées "stress" sont fixées à '0'. Lorsque la LUT est en mode stress, les entrées "stress" sont stressées par DC0.

Lorsque la LUT reçoit un stress dynamique en entrée, on pourrait s'attendre à des dégradations plus importantes lorsque le nombre d'entrées utilisées augmente. En effet, une entrée avec un stress dynamique génère une consommation plus importante de courant dans la LUT contrairement à une entrée statique. Si le nombre d'entrées recevant un signal dynamique augmente alors la consommation de la LUT augmente et par conséquent on peut observer une augmentation locale de la température qui accélère le vieillissement du au BTI.

Nous avons tracé, sur la figure 3.19, la dérive de fréquence relative finale des RO L2 XB avec différents nombres d'entrées de la LUT stressées : 2 entrées utilisées (L2 XB), 3 entrées utilisées (L3 XB), 5 entrées utilisées (L5 XB). Chaque point correspond à la moyenne des 10 dernières heures de mesures réalisées sur trois ROs. Pour un stress statique (DC0 et DC1) on remarque que le nombre d'entrées de la LUT stressées n'est pas significatif sur les dégradations. Par ailleurs, ce n'est pas toujours la même architecture qui provoque l'amplitude de dégradation la plus importante, par exemple : pour DC0 à 65°C et 1.15V_{nom} les dégradations sont maximales pour le RO L5 XB tandis qu'elles sont maximales pour le RO L3 XB à 65°C et 1.3V_{nom}. Ce résultat confirme notre précédente hypothèse sur l'effet du nombre d'entrées utilisées lorsque la LUT reçoit un stress statique.

La figure 3.19 montre que le nombre d'entrées n'a pas d'impact sur les dégradations même pour un stress dynamique. Cela montre que l'augmentation de la température due à un stress dynamique a un effet négligeable sur les dégradations.

Le numéro d'entrée utilisé

Sur la figure 3.20 nous avons tracé la dérive de fréquence relative finale pour des RO L3 utilisant les entrées suivantes de la LUT : XB entrées choisit par Vivado, XB I1 entrée I1 imposée, XB I5 entrée I5 imposées. L'entrée I1 correspond au chemin le plus long dans la LUT tandis que l'entrée I5, proche de la sortie, correspond au chemin le plus court dans la LUT. On constate, pour DC0 et 100MHz, des dégradations plus importantes lorsque l'entrée I1 de la LUT est utilisée, tandis que

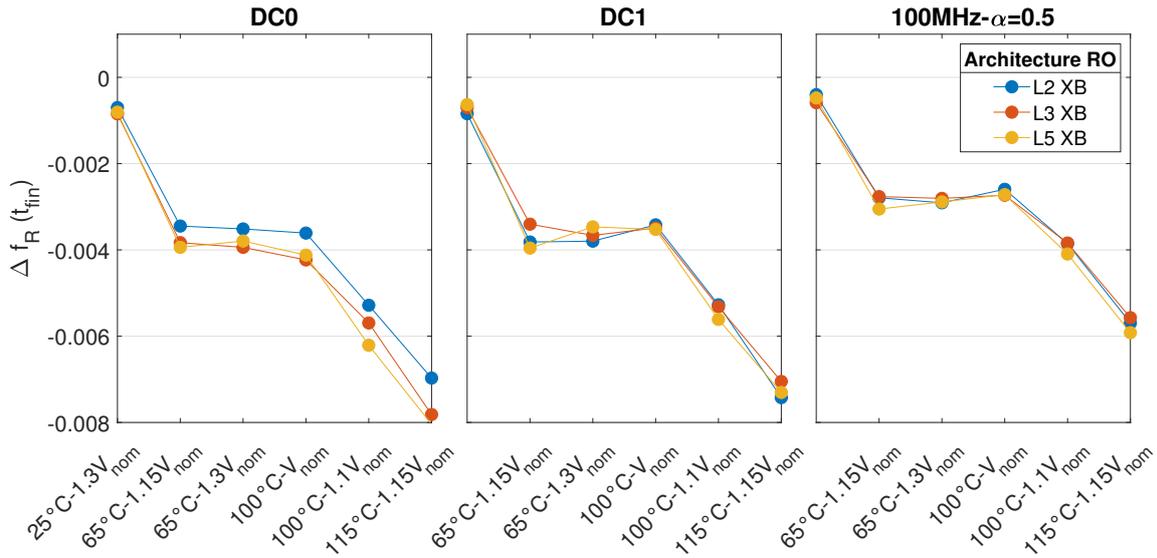


FIGURE 3.19 – Effet du nombre d'entrées de la LUT utilisées sur la dérive fréquentielle relative finale des ROs en fonction des conditions de stress pour un stress statique DC0 (gauche), DC1 (milieu) et dynamique 100MHz (droite)

les dégradations sont minimales lorsque l'entrée I5 est utilisée et cela quelle que soit la condition de stress en température et tension. À priori, la proportion de ressources logiques dans le RO L3 XB I1 est plus importante que dans le RO L3 XB I5. Ce résultat est fortement lié à l'observation déjà faite sur la figure 3.17 : les dégradations sont plus importantes dans les ressources logiques (LUT) que dans les ressources de routages.

Afin de mettre en évidence cette interprétation nous avons extrait depuis Vivado les entrées de LUT utilisées ainsi que les ressources de routage pour chaque RO. En utilisant les données de cette extraction nous calculons le rapport entre le nombre de ressources logiques multiplié par un nombre associé au numéro d'entrée et le nombre de ressources de routage tel que :

$$\text{Ratio LR} = \frac{\text{nombre ressources logiques pondere}}{\text{nombre ressources routage}} \quad (3.10)$$

avec :

$$\text{nombre ressources logiques} = \begin{bmatrix} nI1 & nI2 & nI3 & nI4 & nI5 & nI6 \end{bmatrix} \cdot \begin{bmatrix} 6 & 5 & 4 & 3 & 2 & 1 \end{bmatrix}^T \quad (3.11)$$

où $nI1$ est le nombre de fois que l'entrée I1 d'une LUT est utilisée dans un RO et $n_{\text{ROUT,RO}}$ est le nombre de ressources de routages dans le RO. Le résultat des rapports est inscrit sur la figure 3.20. De fait, plus le rapport est important et plus les dégradations le sont. Cela confirme que la différence de dégradation en fonction de l'entrée de la LUT utilisée dépend de la différence de dégradation entre les ressources logiques et de routage.

Concernant le stress DC1 on observe une amplitude des dégradations équivalentes. En effet, sur la 3.17 nous avons observé pour DC1 un écart des dégradations entre RO LP et RO SP moins important. C'est aussi ce qu'on observe pour DC1 sur la figure 3.20.

L'influence de la fonction de la LUT

Une LUT à six entrées peut implémenter 2^6 fonctions. Cependant, l'influence sur le vieillissement de la fonction d'une LUT peut être réduit à deux cas : soit le bit de configuration transmis

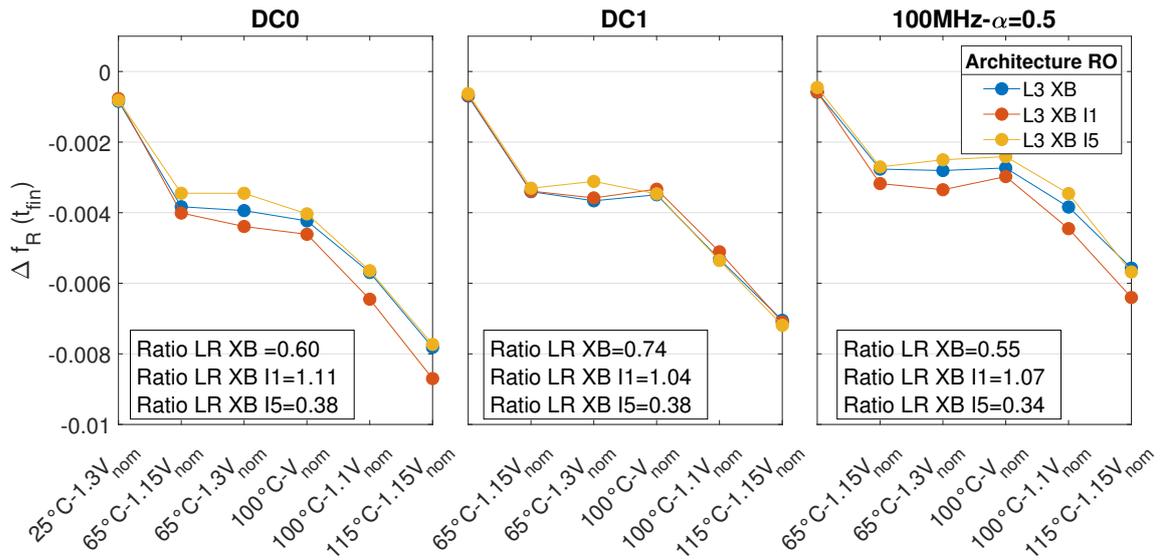


FIGURE 3.20 – Influence de l'entrée de la LUT utilisée sur la dérive fréquentielle relative finale des ROs en fonction des conditions de stress

vers la sortie de la LUT vaut '0' soit il vaut '1'. Donc, c'est la différence entre le NBTI et le PBTI qui marque la différence de dégradation.

Pour confirmer cette dernière remarque, nous traçons sur la figure 3.21 la dérive de fréquence relative finale pour des RO L1 utilisant des LUT de fonctions différentes : inverseur RO L1 I et buffer RO L1 B. À l'inverse d'une LUT buffer, le niveau logique en sortie d'une LUT inverseur est un '0' lorsque l'entrée est un '1' et le niveau logique en sortie est un '1' lorsque l'entrée est un '0'. Dans un RO L1 B toutes les LUTs subissent le même stress (DC0 ou DC1) puisque le niveau logique du signal en entrée de chaque LUT est le même de celui en sortie. Dans un RO L1 I, lorsque le stress est DC0 alors la moitié et une LUT sont stressées par DC0 et les autres LUT par DC1 puisque dans chaque LUT le signal est inversé. Inversement, lorsque le stress est DC1 alors la moitié et une LUT sont stressées par DC1 et les autres LUT par DC0.

Sur la figure 3.21 on observe un comportement inversé de l'amplitude des dégradations entre

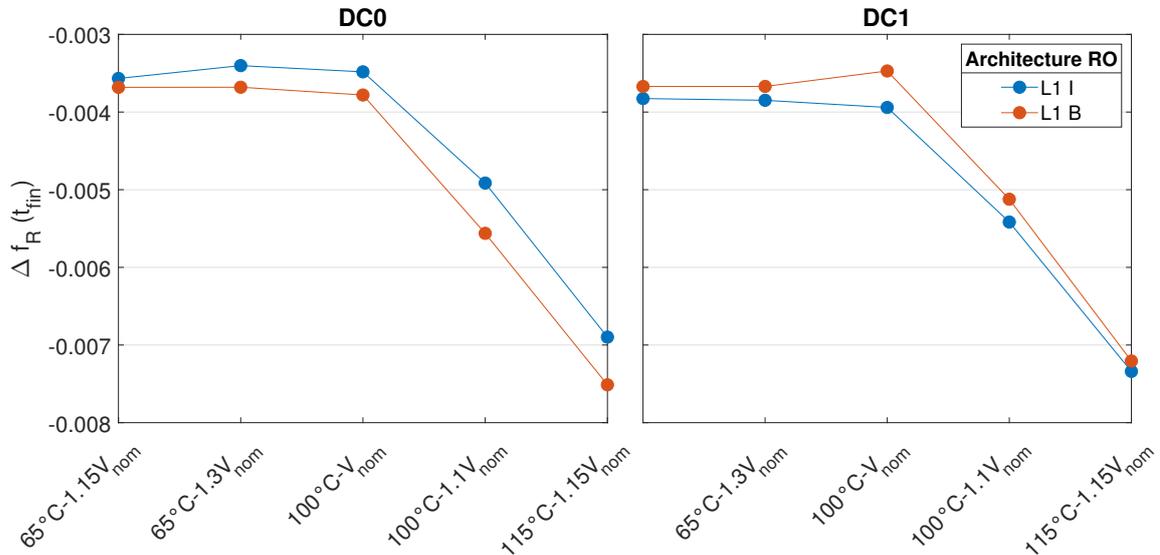


FIGURE 3.21 – Influence de la fonction de la LUT sur la dérive fréquentielle relative finale des ROs en fonction des conditions de stress

RO L1 B et le RO L1 I pour un stress DC0 et DC1. Cela confirme le fonctionnement inversé entre une LUT buffer et une LUT inverseur : DC0 stress la majorité des ressources avec un niveau logique bas dans la LUT tandis que DC0 stress la majorité des ressources avec un niveau logique haut dans la LUT. À ce niveau de l'étude, n'ayant pas connaissance de l'architecture interne d'une LUT il n'est pas possible d'évaluer si c'est le NBTI ou le PBTI qui induit le plus de dégradations.

3.4.7 Conclusion

Dans cette partie nous avons montré que le mode de défaillance de tous les ROs est une dérive paramétrique, ce qui nous oriente vers du BTI et du HCI. Voici les caractéristiques du BTI et du HCI que nous avons observé :

- BTI
 - Modélisation des dégradations avec une loi puissance révélant un exposant au temps de l'ordre de 0.28 pour $T_{stress} \geq 65^{\circ}\text{C}$;
 - Dégradation plus importante pour un stress statique ce qui met en évidence le phénomène de récupération;
 - Non dépendance à la fréquence de stress au dessus de 100Hz;
- HCI
 - Modélisation des dégradations avec une loi puissance révélant un exposant au temps de l'ordre de 0.8 pour $T_{stress} = 25^{\circ}\text{C}$;
 - Dépendance des dégradations à la fréquence de stress avec des dégradations plus importantes à haute fréquence;

Les dégradations sont bien plus importantes à température et tension élevée. Nous obtenons des dégradations maximales de l'ordre du 1% après 10000 heures de vieillissement et dans les pires conditions de stress, résultat rassurant quant à la fiabilité de la technologie.

Concernant l'influence de la configuration du RO, nous notons que les ressources logiques se dégradent plus vite que les ressources de routage dans le FPGA. Le nombre d'entrées utilisées dans la LUT n'a aucun effet sur les dégradations car lorsqu'une entrée n'est pas utilisée, elle est fixée par défaut à un niveau logique qui n'est rien d'autre qu'un stress statique. Le numéro de l'entrée de la LUT utilisée impacte le rapport entre le nombre de ressource de routage et logique et par conséquent, on observe plus de dégradation lorsque l'entrée correspondant au chemin le plus court de la LUT est utilisée. Finalement, en fonction du niveau logique en entrée de la LUT, la fonction de la LUT définit le niveau logique du signal qui se propage dans la LUT et par conséquent favorise les dégradations par du NBTI ou du PBTI.

3.5 Extraction des dégradations par ressources

3.5.1 Introduction

Nous avons étudié dans la section 3.4 l'effet du vieillissement en fonction de l'architecture du RO que l'on peut diviser en quatre parties : le type de ressources utilisées pour implémenter le RO (LUT ou routage), le nombre d'entrées de la LUT utilisées, le numéro de l'entrée de LUT utilisé et la fonction de la LUT. Dans un premier temps, nous avons observé une plus grande amplitude des dégradations pour les ROs constitués principalement de ressources logiques (RO Short Path) que pour les ROs constitués principalement de ressources de routage (RO Long Path). Cette différence de dégradations est certainement liée à la différence de l'architecture interne d'une LUT et de l'architecture de routage. À partir du brevet [44], on peut en effet supposer qu'une LUT est principalement composée de portes de transferts et d'inverseurs, tandis que, l'architecture du routage reste non documentée mais est principalement constituée de portes de transfert. Dans un second temps, nous avons observé une dépendance de l'amplitude des dégradations au numéro de l'entrée de LUT utilisé. En fonction de l'entrée de LUT utilisée le nombre de PMOS et NMOS vieillissent

change et par conséquent la proportion des dégradations due au NBTI et au PBTI varie.

Dans cette section, nous présenterons une méthode consistant à extraire le temps de propagation à t_0 des ressources de routage et de LUT à partir d'une régression linéaire entre la période des ROs et les ressources (routage et LUT) utilisées pour chaque RO. Nous nous appuyerons sur cette méthode pour en extraire les dégradations des ressources de routage et de LUT. Finalement, en supposant une architecture interne d'une LUT basée sur l'étude du brevet [44] ainsi que sur nos résultats, nous extrairons les dégradations du PMOS et NMOS pour un stress statique afin d'observer le NBTI et le PBTI séparément.

3.5.2 Extraction à t_0 du temps de propagation des ressources de LUT et de routage

Nous ne connaissons pas l'architecture interne du FPGA au niveau du transistor, cependant, Vivado fournit des informations détaillées des ressources utilisées dans le FPGA après implémentation. En principe ces informations sont utilisées pour faire de la conception FPGA très bas niveau ou bien encore pour vérifier la conception dans le cas d'une architecture critique. Dans cette partie nous utiliserons ces informations de ressources utilisées dans les ROs afin d'en extraire leurs temps de propagation à partir de la période des ROs.

Extraction des ressources dans Vivado

Avant d'effectuer l'extraction des temps de propagation, dans un premier temps, nous faisons un point sur la terminologie des ressources utilisées par Vivado. Cela nous permettra de mieux comprendre le rôle de chacune des ressources et l'architecture interne de notre FPGA.

Au niveau le plus bas du FPGA, l'élément est un *BEL* (Basic Element of Logic). Les *BEL* sont les composants les plus petits, indivisibles et représentables dans le tissu d'un FPGA. Il existe deux types de *BEL* : Les *BEL* logiques (LUT, bascule-D, additionneur, etc) identifiés en magenta sur la figure 3.22 et les *BEL* de routage (multiplexeur de routage) identifiés en vert sur la figure 3.22.

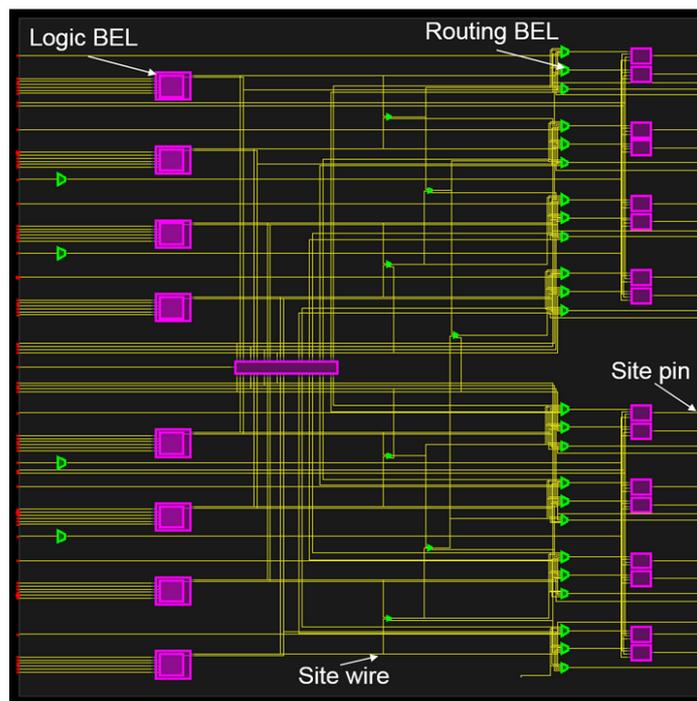


FIGURE 3.22 – Site logique (SLICEL), où les *BEL* logiques sont magenta, les *BEL* de routage sont vertes, les *site pins* sont rouges et les *site wires* sont jaunes [1]

Un groupe de *BEL* associés avec leurs connexions est appelé un *site*. Dans un *site* on retrouve trois principales catégories de ressources : *BEL* (logique et de routage), *site pin* (entrée et sortie du *site*) et *site wire* (réseau de connection dans le *site*). Il existe deux types de *site* : Les *sites* logiques tel que représenté sur la figure 3.22 et les *sites* d'interconnexion. Les *sites* ont aussi, comme les *BEL*, des pins de connection en entrée et sortie qui sont identifiés en rouge sur la figure 3.22.

La figure 3.23 présente les différentes catégories de ressources de routage entre un *site* d'interconnexion et une LUT située dans un *site* logique. Lorsqu'il y a une connection entre deux *BEL* cette connection est un *node*. Le *node* "IMUX_E10" connecte le *PIP* (Programmable Interconnect Point) qui est un *BEL* de routage au *BEL pin* "A1" qui est l'entrée d'une LUT. Un *node* est divisé en plusieurs tronçons de *wires* qui connectent deux pins, par exemple sur la figure 3.23 un *wire* connecte un *site pin* à un *BEL pin*. Dans cette thèse le niveau de détail du *wire* n'est pas nécessaire, c'est pourquoi nous utiliserons le *node*. Finalement, un *PIP*, comme évoqué dans la section 1.4.4, connecte deux *nodes* entre eux. Sur la figure 3.23 le *PIP* "IMUX_16_INT_OUT0->>IMUX_E10" connecte le *node* "IMUX_16_INT_OUT0" avec le *node* "IMUX_E10". Le *nets*, qui n'est pas représenté sur la figure 3.23 est une connection logique contrairement au *PIP*, *node* et *wire* qui sont des connections physique. Un *nets* commence en sortie d'un *BEL pin* logique est finit à l'entrée d'un *BEL pin* logique. Ce qui veut dire qu'un *nets* a un seul niveau logique qui dépend de la *BEL* logique de départ. Pour être connecté entre les deux *BEL pins*, le *nets* utilise les ressources physiques d'interconnexion que sont des *PIPs*, *nodes* et *wires*.

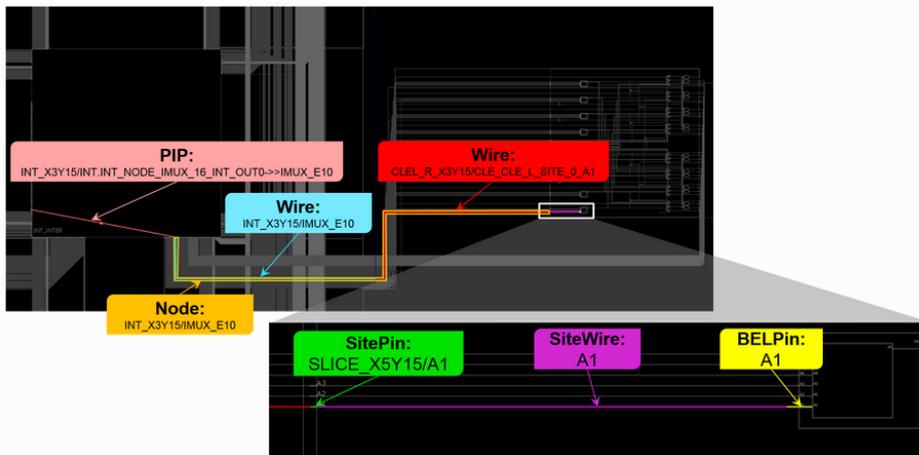
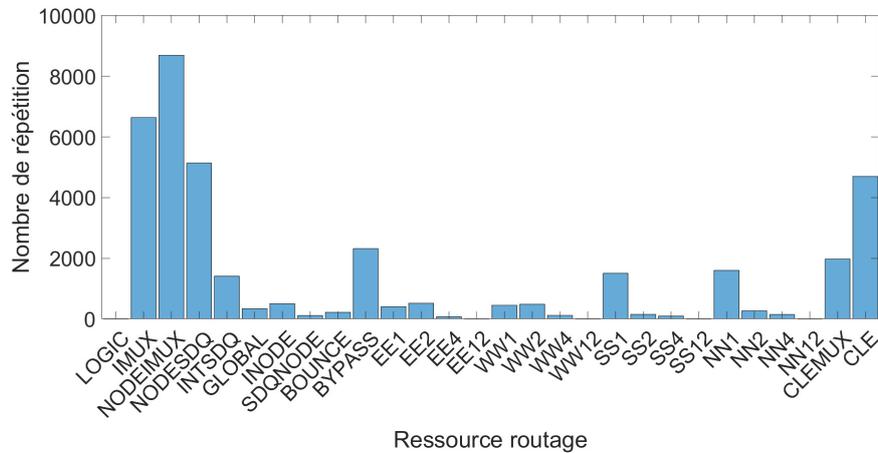


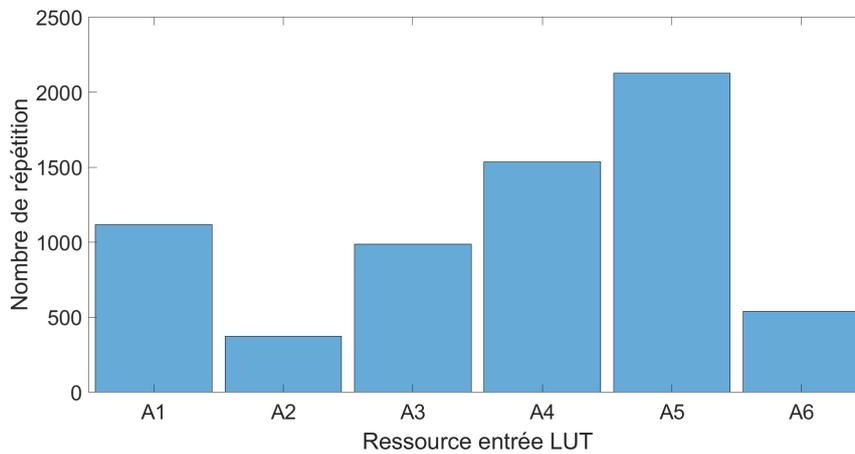
FIGURE 3.23 – Différentes ressources de routage dans un FPGA Xilinx [1]

Dans un premier temps, notre objectif consiste à extraire depuis Vivado les informations de ressources de routage (*PIPs* et *nodes*) et logiques (*BEL* et *BEL pins*) utilisées dans le FPGA pour implémenter chacun des 567 oscillateurs en anneaux. Nous avons écrit un script TCL pour extraire ces informations en post-implémentation sous Vivado. Pour ce faire, nous avons utilisé les fonctions suivantes : *get_bel_pins* [157, p.581], *get_nodes* [157, p.827] et *get_pips* [157, p.856].

La figure 3.24a rapporte le nombre de *nodes*, par catégorie, utilisée pour implémenter tous les ROs. Au total 37869 ressources de *nodes* sont utilisées pour implémenter les 567 ROs. La figure 3.24a dévoile qu'il existe 28 catégories de *nodes* au total, cependant, dans notre implémentation seules 23 catégories de *nodes* sont utilisées. Le nom du *nodes* a une signification suivant sa direction et son étendue dans le réseau de routage. Les *nodes* EE, WW, SS et NN sont des connections dont la direction respective est l'est, l'ouest, le sud ou le nord dans le FPGA. L'étendue du signal peut être de 1 *site*, 2 *sites*, 4 *sites* ou 12 *sites*. Les autres catégories de *nodes* sont des connections locales entre un *site* logique et d'interconnexion qui sont voisins comme cela est représenté sur la figure 1.21. Sachant qu'un *PIP* connecte deux *nodes*, il existe 28² catégories de *PIPs*, raison pour laquelle nous n'avons pas tracé le nombre de *PIPs* utilisés par catégories.



(a)



(b)

FIGURE 3.24 – Histogrammes des ressources utilisées pour implémenter les 567 ROs dans un FPGA classées par catégories (a) *nodes* (b) entrées de LUT

La figure 3.24b montre le nombre d'entrée de LUT par catégories utilisées pour implémenter tous les ROs. Au total 6678 entrées de LUT sont utilisées. En effet nous avons implémenté 504 RO SP (Short Path) composées de 13 LUTs et 63 RO LP (Long Path) composées de 2 LUTs. L'entrée de LUT A1 correspond au chemin le plus long dans la LUT (le plus à gauche sur la figure 1.19) et l'entrée A6 correspond au chemin le plus court (le plus à droite sur la figure 1.19).

Présentation et validation de la méthode : Régression linéaire multiple

Dans cette étude nous considérons que le temps de propagation d'un signal dans un RO correspond à la somme des temps de propagations dans chacune des ressources de routages et logiques qui constituent le RO. À présent que nous connaissons les ressources de routages et logiques utilisées pour implémenter chaque RO ainsi que la période mesuré du RO (voir figure 3.4),

nous pouvons considérer la relation suivante :

$$\begin{bmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \dots & R_{nm} \end{bmatrix} \cdot \begin{bmatrix} D_1(t_0) \\ D_2(t_0) \\ \vdots \\ D_m(t_0) \end{bmatrix} = \begin{bmatrix} T_1(t_0) \\ T_2(t_0) \\ \vdots \\ T_n(t_0) \end{bmatrix} \quad (3.12)$$

où T_i contient le i^{eme} temps de propagation du RO donné par Vivado (régression 1) ou mesuré (régression 2&3), D_j contient le temps de propagation extrait pour la j^{eme} catégorie de ressource et R_{ij} contient le nombre de ressources de la j^{eme} catégorie de ressource utilisé pour implémenter le i^{th} RO. L'indice n correspond au nombre total de RO (soit 567 pour la régression 1&2 et 5103 pour la régression 3) et l'indice m correspond au nombre total de catégories de ressources. Nous avons utilisé la solution classique de la méthode des moindres carrés ordinaire ($D = (R^t R)^{-1} R^t T$) pour obtenir la matrice D du système (3.12) en minimisant les résidus ($T - R \cdot D$). Si la matrice D possède des valeurs négatives, alors ces mêmes valeurs sont fixées à zéro et une nouvelle régression est réalisée. Notons que la matrice R contient les ressources de routage (*nodes*) et les ressources logiques (entrée LUT). Afin de simplifier le nombre d'inconnus dans la matrice R nous n'avons pas utilisé les *PIPs* qui, rappelons-le, sont divisées en 28² catégories différentes, cette simplification étant proposée par [158]. Donc, l'indice m est de 29 catégories, soit 23 catégories de *nodes* et 6 catégories d'entrée de LUT présentées sur la figure 3.24.

Nous avons réalisé trois régressions linéaires :

- Régression 1 : Avec le temps de propagation *slow min* des 567 ROs donné par Vivado (voir figure 3.5a) ;
- Régression 2 : Avec le temps de propagation des 567 ROs mesuré dans un FPGA à $T = 25^\circ C$ et $V = V_{nom}$;
- Régression 3 : Avec le temps de propagation initial des 5103 ROs mesuré dans les neuf FPGA à $T = 25^\circ C$ et $V = V_{nom}$;

Ainsi nous pourrons comparer l'extraction obtenue à partir des données de Vivado avec les mesures.

Pour confirmer la qualité de la régression linéaire, la figure 3.25 présente la prédiction du temps de propagation dans les ROs ($R \cdot D$) en fonction du temps de propagation (T) des ROs données par Vivado (gauche) ou mesurées dans un FPGA (milieu) ou neuf FPGAs (droite). Le résidu calculé sur chacune des figures correspond à l'écart relatif moyen entre la prédiction et Vivado ou

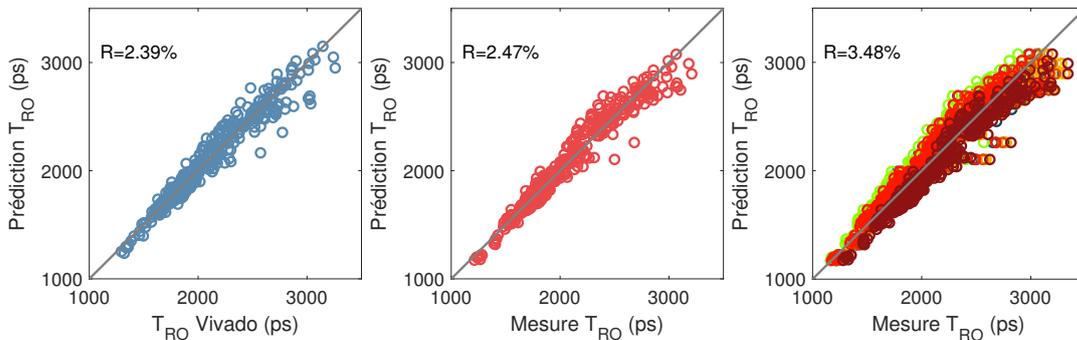


FIGURE 3.25 – Validation des régressions 1 (gauche) 2 (milieu) 3 (droite) en exprimant $R \cdot D$ en fonction du temps de propagation du RO. Sur le graphique de droite chaque couleur correspond à la période des ROs dans un FPGA.

la mesure. On obtient un faible résidu relatif de 2.5% pour tous les ROs d'un seul FPGA. Le léger bruit de la prédiction est probablement dû à deux principaux facteurs : la simplification lors de la catégorisation des ressources (on n'inclut pas les *PIPs* dans la matrice des ressources) et le bruit de mesure. Nous obtenons un résidu légèrement plus important pour la prédiction des neufs FPGAs. Ce bruit supplémentaire provient de la dispersion des performances inter FPGA comme évoqué dans la section 3.3.2.

Étude des résultats d'extraction

Dans la précédente partie, nous avons présenté et confirmé la qualité de la méthode pour reconstruire le temps de propagation d'un RO comme étant la somme du temps de propagation des ressources qui le compose. À présent, nous allons étudier les résultats obtenus. La figure 3.26 expose les résultats contenus dans la matrice D qui est obtenue en résolvant le système (3.12). On constate une cohérence des résultats entre la régression faite à partir des données Vivado et les mesures, ce qui conforte la cohérence de la méthode d'extraction.

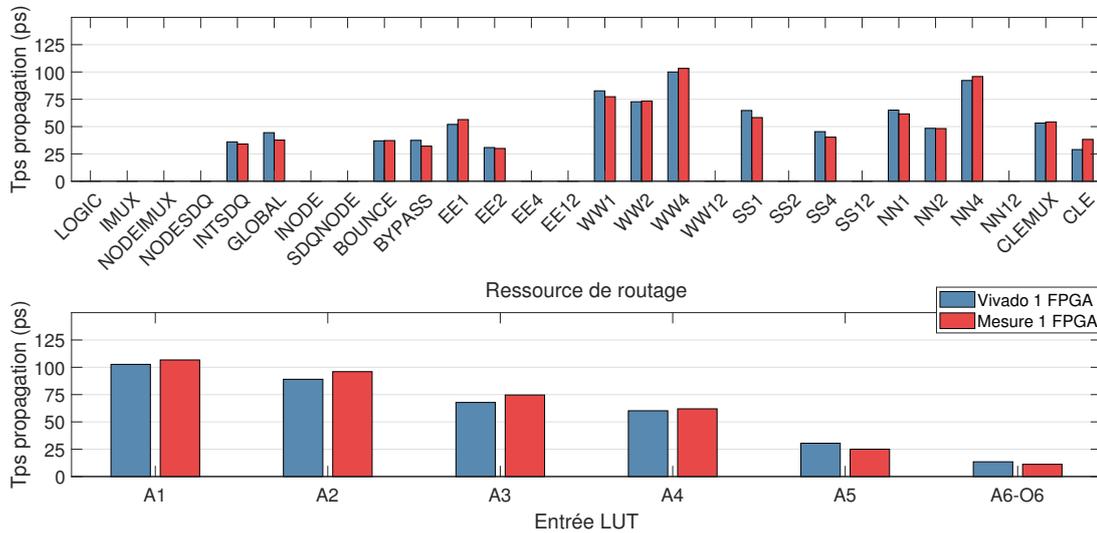


FIGURE 3.26 – Temps de propagations des ressources de routage (haut) et d'entrée de LUT (bas), correspondant à la matrice D obtenue en résolvant le système (3.12) avec les données Vivado (bleu) et les mesures faites sur 1 FPGA (rouge)

On remarque que certaines catégories de ressources de routage ont un résultat nul. En effet, sur la figure 3.24, certaines catégories ne sont pas utilisées pour implémenter les ROs. Par ailleurs, parce que les ressources utilisées ne sont pas optimisées pour résoudre le système (3.12), certains résultats négatifs ont été fixés à zéro.

La ressource de routage "CLEMUX" et "CLE" sont des *nodes* qui connectent la sortie de la LUT vers l'entrée du *site* d'interconnexion. Cependant, la ressource "CLEMUX" traverse un multiplexeur supplémentaire dans le *site* logique comparé au "CLE". On constate que le temps de propagation obtenu pour "CLEMUX" est supérieur à celui de la ressource "CLE", ce qui confirme la cohérence de la méthode d'extraction.

Concernant les résultats du temps de propagation des ressources de LUT on constate une parfaite cohérence avec l'architecture d'une LUT (voir figure 1.19) proposé dans le brevet [44]. Plus le chemin dans la LUT est long et plus le temps de propagation est important (voir table 3.3). De plus, en faisant la différence entre deux entrées voisines on obtient le temps de propagation dans un étage de la LUT (voir table 3.3). On constate alors que le temps de propagation $etage2 \approx etage4$ et qu'il est supérieur au temps de propagation dans les étages $etage1 \approx etage3 \approx etage5 \approx etage6$.

Cela nous confirme la présence d'inverseur entre l'étage 2 et 3 ainsi qu'entre l'étage 4 et 5 dans la LUT. Non seulement, ce résultat est cohérent avec le brevet de l'architecture interne d'une LUT, mais par ailleurs il nous conforte sur la supposition de l'architecture interne de la LUT dans notre FPGA proposée sur la figure 1.19.

Dans l'équation (3.12), les matrices R et $D(t_0)$ se décomposent en deux sous matrices propres aux ressources logiques et de routage :

$$\left[\begin{array}{c|c} R_{\text{ROUT}} & R_{\text{ILUT}} \end{array} \right] \cdot \left[\begin{array}{c} D_{\text{ROUT}}(t_0) \\ D_{\text{ILUT}}(t_0) \end{array} \right] = T(t_0) \quad (3.13)$$

où $D_{\text{ROUT}}(t_0)$ et $D_{\text{ILUT}}(t_0)$ sont les matrices contenant respectivement le temps de propagation initial pour chaque catégorie des ressources de routage et logique, R_{ROUT} et R_{ILUT} sont les matrices contenant respectivement le nombre de ressources de routage et logiques par catégorie :

$$R_{\text{ROUT}} = \begin{bmatrix} R_{\text{IMUX}_1} & R_{\text{NODEIMUX}_1} & \cdots & R_{\text{CLEMUX}_1} & R_{\text{CLE}_1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ R_{\text{IMUX}_n} & R_{\text{NODEIMUX}_n} & \cdots & R_{\text{CLEMUX}_n} & R_{\text{CLE}_n} \end{bmatrix} \quad (3.14)$$

$$R_{\text{ILUT}} = \begin{bmatrix} R_{\text{ILUT}_{A1,1}} & R_{\text{ILUT}_{A2,1}} & R_{\text{ILUT}_{A3,1}} & R_{\text{ILUT}_{A4,1}} & R_{\text{ILUT}_{A5,1}} & R_{\text{ILUT}_{A6,1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ R_{\text{ILUT}_{A1,n}} & R_{\text{ILUT}_{A2,n}} & R_{\text{ILUT}_{A3,n}} & R_{\text{ILUT}_{A4,n}} & R_{\text{ILUT}_{A5,n}} & R_{\text{ILUT}_{A6,n}} \end{bmatrix} \quad (3.15)$$

Les catégories des ressources de la matrice R_{ROUT} sont présentées sur la figure 3.24a et les catégories des ressources de la matrice R_{ILUT} sont présentées sur la figure 3.24b. Ainsi, nous pouvons calculer le temps de propagation initial dans les ressources de routage ($T_{\text{ROUT}}(t_0)$) et dans les ressources logiques ($T_{\text{ILUT}}(t_0)$) :

$$R_{\text{ROUT}} \cdot D_{\text{ROUT}}(t_0) = T_{\text{ROUT}}(t_0) \quad (3.16)$$

$$R_{\text{ILUT}} \cdot D_{\text{ILUT}}(t_0) = T_{\text{ILUT}}(t_0) \quad (3.17)$$

Sur la figure 3.27 nous avons tracé le rapport du temps de propagation dans les ressources de LUT par le temps de propagation total dans le RO. Chaque graphique de la figure est un histogramme de ce rapport pour tous les ROs, c'est à dire 567 ROs (stress interne et neuf FPGA confondus), d'architecture logique identique.

De manière générale, la figure 3.27 révèle que la majorité du temps de propagation du RO n'est pas générée par les ressources de LUT mais par les ressources de routage, même pour les RO SP (voir section 2.3) dont l'implémentation est contrainte dans une ou deux slices afin de limiter le nombre de ressource de routage. Et, encore plus étonnant, on observe pour le RO L3 XB I1 un temps de propagation équivalent entre les ressources de routages et de LUT alors que ce type de

TABLE 3.3 – Temps de propagation en fonction de l'entrée de la LUT extrait par la régression linéaire (3.12)

A1	A2	A3	A4	A5	A6
104ps	91ps	69ps	62ps	31ps	14ps
étage 1 : A1-A2	étage 2 : A2-A3	étage 3 : A3-A4	étage 4 : A4-A5	étage 5 : A5-A6	étage 6 : A6
13ps	22ps	7ps	31ps	17ps	14ps

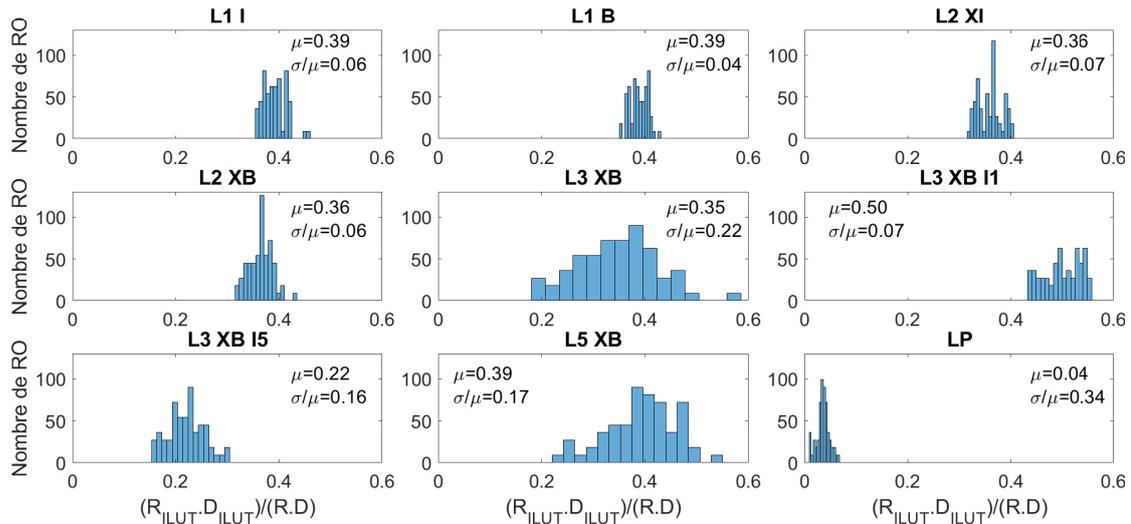


FIGURE 3.27 – Proportion du temps de propagation initial généré par les ressources de LUT par rapport au temps de propagation total dans un RO pour différentes architectures de RO

RO utilise uniquement l'entrée A1 de la LUT correspondant au chemin logique le plus long. Par ailleurs, on obtient un rapport du temps de propagation moyen dans les ressources de LUT de seulement 4% dans les RO LP composés de 2 LUTs ce qui est cohérent.

Aussi, la figure 3.27 dévoile jusqu'à 34% de dispersion relative pour les ROs de même architecture logique. Cela révèle une dispersion des ressources de routages et logiques utilisées lors de l'implémentation du RO dans le FPGA. La figure 2.4 montre que les ROs de même architecture mais de stress différents sont répartis dans le FPGA. Ceci laisse à penser qu'en fonction du placement dans le FPGA l'algorithme de placement&routage de Xilinx n'utilise pas les mêmes catégories de ressources. Comme nous avons vu sur figure 3.17, les ressources de routages et logiques vieillissent différemment. Le risque d'avoir de la dispersion du rapport du nombre de ressource de routage et logique utilisé pour des ROs de même architecture est de masquer l'effet du stress interne sur le vieillissement. C'est une raison suffisamment importante pour dissocier les dégradations des ressources de routage et logiques.

Dans cette partie nous avons introduit une méthode pour extraire le temps de propagation des ressources de routages et logiques pour chaque RO. Nous avons montré la justesse de cette méthode avec un résidu relatif moyen entre la prédiction du temps de propagation des ROs et la mesure inférieur à 2.5%. De plus, nous avons montré la cohérence des résultats du temps de propagation des différentes ressources de routages et entrées de LUT en comparant avec l'architecture interne du FPGA. Finalement, nous avons vu que les ressources de routages contribuent majoritairement au temps de propagation d'un RO contrairement aux ressources logiques qui sont moins importantes. La dispersion du rapport du temps de propagation dans les LUT et le routage entre architecture et pour une même architecture de RO révèle l'importance de dissocier ces deux contributeurs pour en étudier leurs vieillissements respectifs.

3.5.3 Extraction des dégradations de ressources de LUT et de routage

Dans la partie précédente nous avons extrait du temps de propagation des ROs la contribution du temps de propagation des ressources de routages et logiques à t_0 . Dans la continuité de cette méthode, nous allons extraire à chaque instant t la contribution du routage et des LUTs afin de connaître individuellement les dégradations dans chacune des ressources. Pour ce faire, nous itérons la résolution du système 3.12 à chaque instant de mesure. Ainsi nous obtenons la matrice D_{mo} qui contient le temps de propagation de la m^{eme} catégorie de ressource au o^{eme} temps de

vieillessement. Parce que les FPGAs se trouvent dans des conditions de vieillissement différentes, nous réalisons cette manipulation indépendamment pour les 9 FPGAs, ainsi, nous obtenons 9 matrices D_{mo} .

La figure 3.28a montre la qualité de l'extraction de la régression. On constate un important résidu relatif de 10%. La figure 3.28b présente la moyenne de la différence des dégradations entre la mesure et la prédiction en fonction du stress interne et pour toutes les architectures de RO. On constate que l'écart varie en fonction du stress interne considéré. En effet, lors de la régression, nous avons utilisé les mesures de tous les ROs, stress internes et architectures confondus. Cela signifie que la matrice D_{mo} représente une moyenne des dégradations de tous les stress internes (fréquence et rapport cyclique). Or, nous avons vu sur la figure 3.14 que le stress interne a un effet important sur les dégradations. Sachant que les dégradations sont plus importantes pour un stress statique et en basse fréquence, cela a pour effet de sous-estimer la prédiction de ces dégradations, contrairement à la prédiction des dégradations en hautes fréquences.

Pour obtenir une prédiction des dégradations de qualité, il est primordial de considérer individuellement chacun des stress internes lors de la régression. Cependant, cela revient à un système de régression (équation (3.12)) de 27 lignes : 3×9 architectures de RO par stress interne, pour 29 inconnues : temps de propagation dans les 6 catégories de LUTs et les 23 catégories de *nodes*. Pour résoudre le système il est nécessaire de diminuer le nombre d'inconnues.

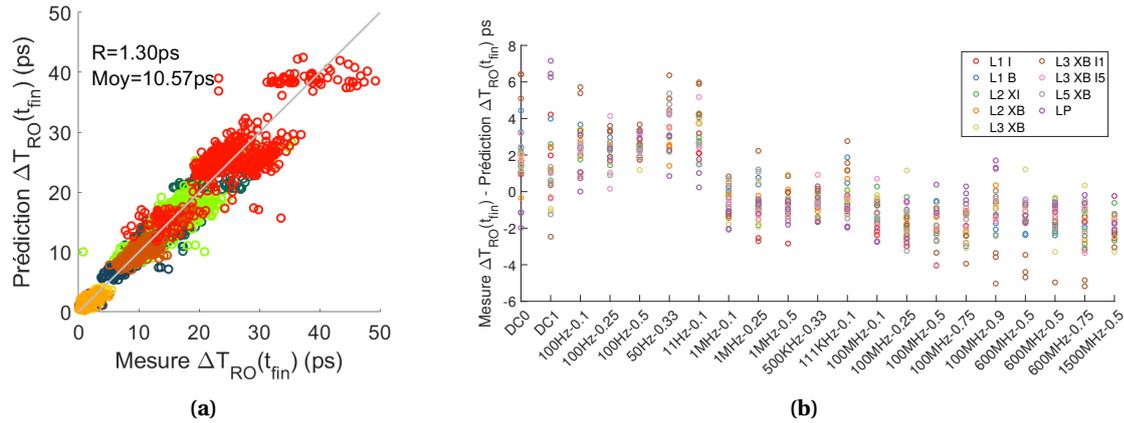


FIGURE 3.28 – Régression non simplifiée (a) Comparaison de la prédiction des dégradations de la période des ROs à t_{fin} prédites et mesurées dans les différents FPGA (b) Différence entre la dégradation à t_{fin} mesurée et prédite en fonction du stress interne, pour toutes les architectures de RO et un même FPGA ($T_{stress} = 100^\circ\text{C}$ et $V_{stress} = V_{nom}$)

Une ressource de routage est un ensemble de fils et de transistors passants. Alors, nous pouvons considérer que l'effet du vieillissement dans les différentes ressources de routage est proportionnel au temps de propagation initial. Ainsi, nous remplaçons les catégories de *nodes* qui sont 23 paramètres dans la matrice R par le temps de propagation dans les ressources de routage à t_0 extrait dans la partie 3.5.2 :

$$\begin{bmatrix} R_{IMUX_1} & R_{NODEIMUX_1} & \cdots & R_{CLEMUX_1} & R_{CLE_1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ R_{IMUX_n} & R_{NODEIMUX_n} & \cdots & R_{CLEMUX_n} & R_{CLE_n} \end{bmatrix} \cdot \begin{bmatrix} D_{IMUX}(t) \\ D_{NODEIMUX}(t) \\ \vdots \\ D_{CLEMUX}(t) \\ D_{CLE}(t) \end{bmatrix} \Rightarrow \begin{bmatrix} T_{ROUT,1}(t_0) \\ \vdots \\ T_{ROUT,n}(t_0) \end{bmatrix} \cdot \delta_{ROUT}(t) \quad (3.18)$$

où $\delta_{\text{ROUT}}(t)$ est le facteur de dégradation du routage qui est égale à 1 à t_0 . Ainsi, le système suivant permet d'identifier l'évolution de la période d'un RO et de dissocier le temps de propagation généré dans les ressources de routage ($T_{\text{ROUT}}(t)$) et logiques ($T_{\text{ILUT}}(t)$) :

$$\begin{aligned}
 & \left[\begin{array}{c|cccccc} T_{\text{ROUT},1}(t_0) & R_{\text{ILUT}_{A1,1}} & R_{\text{ILUT}_{A2,1}} & R_{\text{ILUT}_{A3,1}} & R_{\text{ILUT}_{A4,1}} & R_{\text{ILUT}_{A5,1}} & R_{\text{ILUT}_{A6,1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \\ T_{\text{ROUT},n}(t_0) & R_{\text{ILUT}_{A1,n}} & R_{\text{ILUT}_{A2,n}} & R_{\text{ILUT}_{A3,n}} & R_{\text{ILUT}_{A4,n}} & R_{\text{ILUT}_{A5,n}} & R_{\text{ILUT}_{A6,n}} \end{array} \right] \cdot \begin{bmatrix} \delta_{\text{ROUT}}(t) \\ D_{\text{ILUT}_{A1}}(t) \\ D_{\text{ILUT}_{A2}}(t) \\ D_{\text{ILUT}_{A3}}(t) \\ D_{\text{ILUT}_{A4}}(t) \\ D_{\text{ILUT}_{A5}}(t) \\ D_{\text{ILUT}_{A6}}(t) \end{bmatrix} \\
 & = \left[\begin{array}{c|c} T_{\text{ROUT}_1}(t) & T_{\text{ILUT}_1}(t) \\ \vdots & \vdots \\ T_{\text{ROUT}_n}(t) & T_{\text{ILUT}_n}(t) \end{array} \right] = \begin{bmatrix} T_1(t) \\ \vdots \\ T_n(t_0) \end{bmatrix} \quad (3.19)
 \end{aligned}$$

La figure 3.29a confirme qu'avec cette simplification des ressources de routage dans le système de régression nous obtenons une prédiction des dégradations deux fois plus précise que la précédente avec un résidu relatif inférieur de 5%. De plus avec cette nouvelle méthode nous pouvons prédire les dégradations individuellement pour chaque stress interne comme la figure 3.29b le montre.

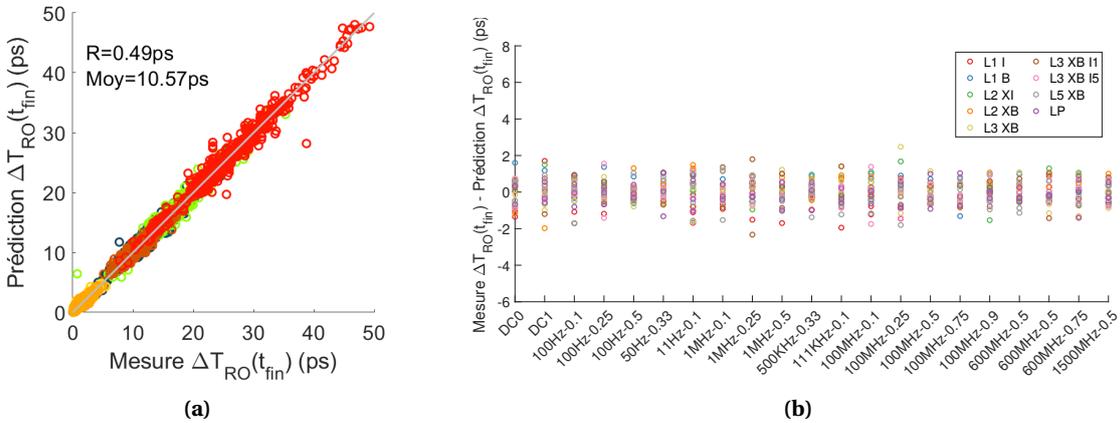


FIGURE 3.29 – Régression simplifiée (a) Comparaison de la prédiction des dégradations de la période des ROs à t_{fin} prédites et mesurées dans les différents FPGA (b) Différence entre la dégradation à t_{fin} mesurée et prédite en fonction du stress interne, pour toutes les architectures de RO et un même FPGA ($T_{\text{stress}} = 100^\circ\text{C}$ et $V_{\text{stress}} = V_{\text{nom}}$)

Après avoir montré la précision de la régression pour prédire la dégradation du RO à partir des ressources qui le constituent, nous allons étudier la cohérence du résultat. La figure 3.30 montre les dégradations respectives d'un RO L3 XB I1 correspondant à l'architecture avec le chemin logique (dans la LUT) le plus long et du RO LP avec le chemin de routage le plus long. Tout d'abord, on remarque la ressemblance entre la dérive du temps de propagation du RO mesuré et prédite. De plus, on observe que les dégradations de T_{ILUT} sont plus importantes que les dégradations de T_{ROUT} pour le RO L3 XB I1 et inversement pour le RO LP. Ce résultat confirme la cohérence de la méthode d'extraction des dégradations des ressources de routage et de LUT.

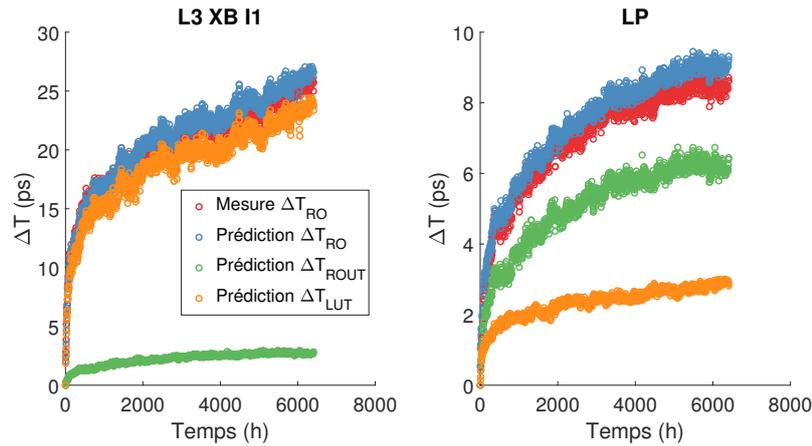


FIGURE 3.30 – Comparaison de la dégradation de la période d’un RO L3 XB I1 et LP mesurée et prédite ainsi que la décomposition de la dégradation du RO en une dégradation de routage et de LUT extraits à partir de la prédiction

La figure 3.31 présente sous forme d’histogramme le proportion des dégradations à t_{fin} due aux ressources de LUT pour chaque architecture de RO. Alors que la figure 3.27 met en évidence que le temps propagation d’un RO est en majorité généré par le retard accumulé dans les ressources de routage, on constate ici, que les dégradations dans les ressources de LUT sont plus importantes pour toutes les architectures de RO, mis à part le LP. Ceci rejoint l’observation faite sur la figure 3.17 qui montre que les dégradations relatives pour les ressources logiques sont plus importantes que pour les ressources de routages.

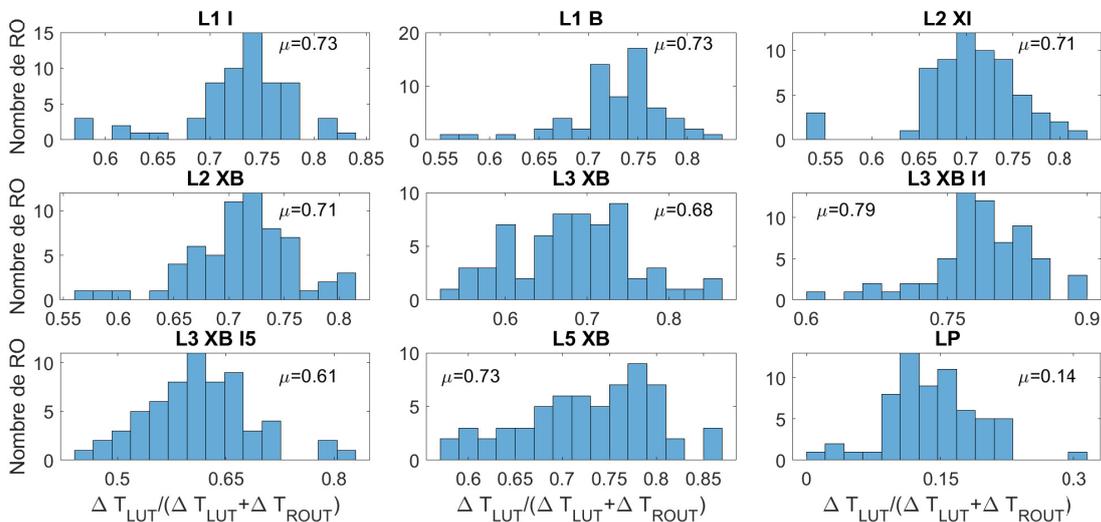


FIGURE 3.31 – Rapport des dégradations à t_{fin} dans les ressources de LUT classé par architecture et pour un FPGA ($T_{stress} = 100^\circ\text{C}$ et $V = V_{nom}$)

3.5.4 Extraction du temps de propagation en montée et descente

Jusqu’à présent, nous avons extrait les temps de propagations des ressources de routages et logiques à partir des périodes de RO. La période d’un RO se décompose elle même en un temps de propagation d’un front montant (τ_{rise}) et d’un front descendant (τ_{fall}). Dans la section 2.4.2 nous voyons que les transistors passants pour propager un front montant dans une LUT sont différents de ceux pour propager un front descendant. Nous avons vu comment dissocier le temps de propagation d’un front montant et descendant à partir de la mesure de fréquence et de rapport

cyclique du RO. Ainsi, nous pouvons observer l'effet du vieillissement sur les transistors passant pour un front montant et ceux passant pour un front descendant.

C'est pourquoi, à partir de la matrice du temps de propagation dans les ressources de LUT (D_{ILUT}) et de routages (D_{ROUT}) ainsi que du rapport cyclique du RO mesuré, nous avons calculé les matrices des temps de propagation d'un front descendant et montant dans les ressources de LUT ($D_{fall,ILUT}$ et $D_{rise,ILUT}$) et dans les ressources de routage ($D_{fall,ROUT}$ et $D_{rise,ROUT}$) tel que :

$$D_{fall,ILUT} = D_{ILUT} \times \alpha_{RO} \quad (3.20)$$

$$D_{fall,ROUT} = D_{ROUT} \times \alpha_{RO} \quad (3.21)$$

$$D_{rise,ILUT} = D_{ILUT} \times (1 - \alpha_{RO}) \quad (3.22)$$

$$D_{rise,ROUT} = D_{ROUT} \times (1 - \alpha_{RO}) \quad (3.23)$$

où α_{RO} est le rapport cyclique du signal du RO mesuré. Les matrices telles qu'elles sont définies dans (3.23) signifient que nous considérons la dégradation d'un front montant et descendant équilibrée au niveau des ressources de routages et de LUT. En utilisant cette méthode, nous obtenons un résidu relatif inférieur à 1% entre la prédiction et la mesure du temps de propagation d'un front montant et descendant.

La figure 3.32 présente les dégradations, à t_{fin} , de temps de propagations de fronts montants et descendant dans les RO. Pour la mesure de dégradation du temps de propagation de fronts montants, nous avons utilisé des ROs stressés par DC1 (étoiles sur la figure 3.32), tandis que, pour la mesure de dégradation du temps de propagation de fronts descendants, nous avons utilisé des RO stressés par DC0 (rond sur la figure 3.32). Dans les deux cas, les transistors passants dans le RO pendant le mode mesure et stress sont identiques. Dans ces conditions, l'amplitude des dégradations devrait dépendre du temps de propagation à t_0 (τ_{t_0}), temps qui dépend du nombre de ressources vieilles. Or, la figure 3.32 montre que l'amplitude des dégradations, à l'échelle du RO, est indépendante du temps de propagation à t_0 . En effet, comme nous avons vu sur la figure 3.27, chaque RO est composé d'un rapport de ressources de LUT et de routages différent, or ces deux types de ressources vieillissent différemment. Lorsque l'on étudie la dégradation du temps de propagation dans un RO, son amplitude ne dépend pas seulement du nombre de ressources vieilles mais aussi du type (routage ou LUT), c'est la raison pour laquelle on n'observe pas une forte corrélation entre l'amplitude des dégradations au niveau du RO et de son temps de propagation à t_0

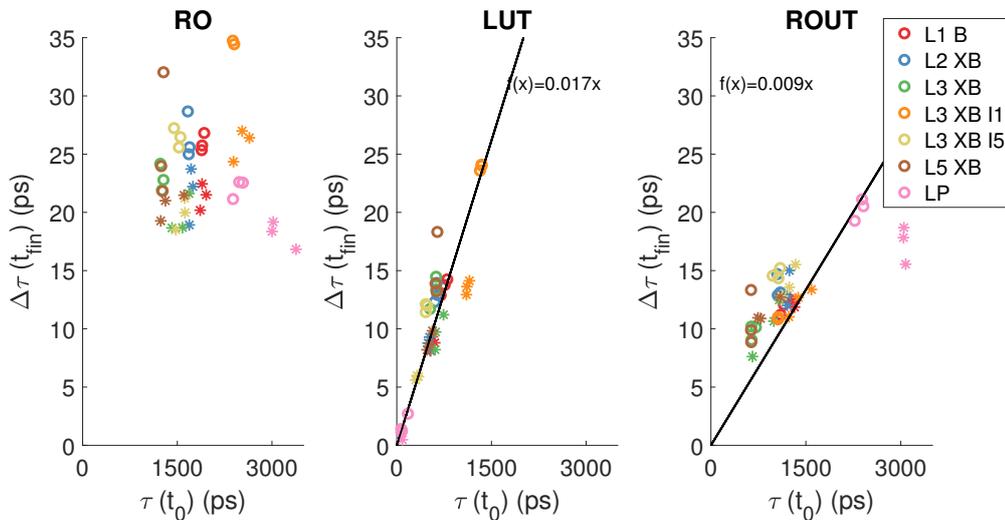


FIGURE 3.32 – Dégradation du temps de propagation de fronts montants pour un stress DC1 (étoile) et de fronts descendants pour un stress DC0 (cercle) au niveau du RO (gauche) au niveau des LUT (milieu) et au niveau des ressource de routage (droite) pour un FPGA ($T_{stress} = 100^\circ$ et $V_{stress} = V_{nom}$)

(voir à gauche de la figure 3.32). Cependant, lorsque l'on sépare les dégradations dans la LUT (voir au milieu de la figure 3.32) et dans le routage (voir à droite de la figure 3.32), on observe une forte corrélation linéaire et ceci quelle que soit l'architecture du RO. Donc, en séparant les dégradations de LUT et routage, pour un stress interne et externe (tension et température) fixe, l'amplitude des dégradations ne dépend plus que du temps de propagation à t_0 quelle que soit l'architecture du RO. Ceci justifie l'utilisation de la mesure relative, c'est à dire le rapport de dégradation par la mesure à t_0 , pour modéliser le vieillissement dans le chapitre 4.

Dans cette partie, nous avons présenté la méthode d'extraction, de qualité, des dégradation des ressources de routages et de LUT. À partir de cette extraction, nous avons montré que la majorité des dégradations d'un RO se produisent dans les ressources de LUT, en moyenne 72% dans les RO SP. Finalement, nous avons montré qu'après extraction des dégradations dans les LUT et le routage, l'amplitude des dégradations est proportionnelle au temps de propagation mesuré à t_0 . Ceci justifie, dans le chapitre 4, la modélisation du vieillissement d'un RO en dissociant les dégradations dans les ressources de LUT et de routages et l'utilisation de la mesure relative.

3.5.5 Extraction des dégradations des PMOS et NMOS pour un stress statique

Dans la partie précédente nous avons montré comment extraire, à partir de la mesure de dégradation de la période du RO, la dégradation du temps de propagation d'un front montant et descendant dans les ressources de LUT et de routage. Dans cette partie, nous allons extraire les dégradations au niveau du transistor. Dans un premier temps nous extrayons le temps de propagation d'un transistor avant le vieillissement puis nous extrayons les dégradations dans un PMOS et dans un NMOS.

La mesure du temps de propagation du front montant (τ_{rise}) et descendant (τ_{fall}) nous permet d'observer l'impact direct de la dégradation des transistors passants respectivement lors d'un stress statique DC1 et DC0. Nous verrons dans le chapitre 4 qu'il existe des effets secondaires de la dégradation des transistors au niveau circuit, tel que l'amélioration du temps de propagation, lorsque le niveau du front se propageant dans le RO est différent pendant les temps de stress et de mesure, par exemple pour un stress DC0 et la mesure de τ_{rise} , un stress DC1 et la mesure de τ_{fall} ou simplement pour un stress dynamique lors de la mesure de τ_{rise} et τ_{fall} . Afin d'éviter cet effet qui complexifie l'étude des dégradations au niveau du transistor, nous nous contenterons, dans cette partie, d'extraire les dégradations du PMOS et du NMOS dans le cas d'un stress statique, en utilisant la dégradation de τ_{rise} pour un stress DC1 et de τ_{fall} pour un stress DC0.

La table 3.4 présente le temps de propagation d'un front montant et descendant dans un RO (calculé à partir des mesures), dans les ressources de LUT et de routage (extrait dans la partie précédente) pour différentes conditions de stress. Nous avons observé sur la figure 3.8 que le rapport cyclique du RO, à $T = 25^\circ\text{C}$ et $V = V_{nom}$, est en moyenne égale à 0.5 c'est à dire que les temps de propagation d'un front montant et descendant sont équilibrés. Sur la 3.4 on remarque aussi que les temps de propagation d'un front montant et descendant sont équilibrés aussi bien dans le RO, que dans les ressources de LUT ou de routages. Ceci prouve qu'à t_0 les temps de propagation dans les NMOS et PMOS sont équilibrés. De plus, conformément à la caractéristique de fonctionnement du RO en fonction de la température et de la tension (voir figure 2.23) on constate que le temps de propagation est plus sensible à la tension qu'à la température. Par exemple, le temps de propagation est plus faible pour $T = 65^\circ\text{C}$ et $V = 1.3V_{nom}$ que pour $T = 100^\circ\text{C}$ et $V = V_{nom}$.

Dans un premier temps, nous extrayons le temps de propagation d'un transistor à t_0 . Parce qu'à t_0 le temps de propagation d'un NMOS et PMOS est équilibré, nous considérons le temps de propagation dans un PMOS égal à celui dans un NMOS. Pour l'extraction du temps de propagation d'un transistor, nous utilisons le temps de propagation dans les ressources de LUT. En effet, nous avons une meilleure connaissance de l'architecture interne d'une LUT basée sur l'étude du brevet

TABLE 3.4 – Temps de propagation dans différentes ressources à t_0 pour un RO L1 B

Ressource	Tps propa- gation (ps)	25°C V_{nom}	25°C $1.15V_{nom}$	25°C $1.3V_{nom}$	65°C $1.15V_{nom}$	65°C $1.3V_{nom}$	100°C V_{nom}	100°C $1.15V_{nom}$	115°C $1.15V_{nom}$
RO	tfall	1891	1670	1522	1646	1618	1889	1744	1883
	trise	1917	1620	1510	1662	1554	1894	1714	1830
LUT	tfall	847	702	591	642	577	785	682	756
	trise	864	669	592	645	568	786	673	740
ROUT	tfall	1044	969	931	1003	1042	1104	1062	1127
	trise	1053	951	918	1017	987	1108	1041	1090
Transistor	τ_{fall} et τ_{rise}	14.2	11.1	9.6	10.5	9.3	12.9	11.1	12.1

[44] mais aussi sur nos résultats obtenus et présentés sur la figure 3.26 qui nous confortent sur la position des inverseurs entre l'étage 2 et 3 ainsi qu'entre l'étage 4 et 5 tel que représenté sur la figure 1.19. Ainsi, nous pouvons retracer le nombre de transistors passants en fonction de l'entrée de la LUT utilisée. À t_0 , le temps de propagation d'un front montant ou descendant correspond alors à la somme des temps de propagation des transistors constituant le chemin de propagation :

$$\begin{bmatrix} R_{ILUT_{A1,1}} & R_{ILUT_{A2,1}} & R_{ILUT_{A3,1}} & R_{ILUT_{A4,1}} & R_{ILUT_{A5,1}} & R_{ILUT_{A6,1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ R_{ILUT_{A1,n}} & R_{ILUT_{A2,n}} & R_{ILUT_{A3,n}} & R_{ILUT_{A4,n}} & R_{ILUT_{A5,n}} & R_{ILUT_{A6,n}} \end{bmatrix} \cdot \begin{bmatrix} N_{Tr_{A1}} \\ N_{Tr_{A2}} \\ N_{Tr_{A3}} \\ N_{Tr_{A4}} \\ N_{Tr_{A5}} \\ N_{Tr_{A6}} \end{bmatrix} \cdot D_{Tr}(t_0) = \begin{bmatrix} \tau_{fall_1}(t_0) \\ \tau_{rise_1}(t_0) \\ \vdots \\ \tau_{fall_n}(t_0) \\ \tau_{rise_n}(t_0) \end{bmatrix} \quad (3.24)$$

où la matrice R_{ILUT} contient le nombre de fois où les entrées de LUT (A1, A2, etc) sont utilisées pour le n^{eme} RO, la matrice N_{Tr} contient le nombre de transistors passant dans une LUT en fonction de l'entrée utilisée, D_{Tr} est le temps de propagation d'un transistor à t_0 et la matrice τ contient les temps de propagation d'un front descendant et montant du n^{eme} RO. Conformément à l'architecture de LUT considérée (voir figure 1.19) la matrice N_{Tr} vaut :

$$\begin{bmatrix} 8 \\ 7 \\ 5 \\ 4 \\ 2 \\ 1 \end{bmatrix} \quad (3.25)$$

Pour résoudre le système (3.24) et obtenir le temps de propagation d'un transistor, nous avons utilisé les données de tous les ROs de type buffer (voir liste 2.3.2) stressés par DC0 et DC1, soit 42 ROs. La figure 3.33 présente la comparaison entre la reconstruction de τ_{fall} et τ_{rise} à partir du temps de propagation du transistor extrait ($R_{ILUT} \cdot N_{Tr} \cdot D_{Tr}$) avec la matrice τ . Les résultats de l'extraction sont donnés dans la dernière ligne de la table 3.4. À $T = 25^\circ\text{C}$ et $V = V_{nom}$ le temps de propagation d'un transistor est de 14.2ps.

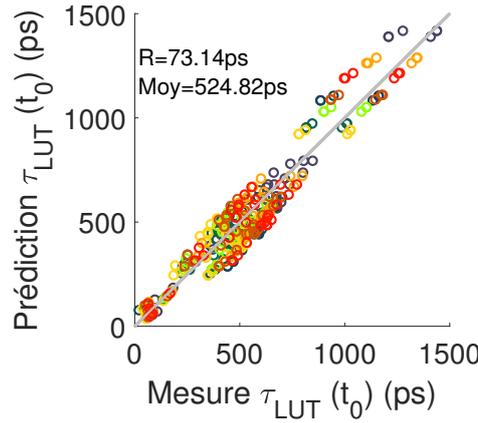


FIGURE 3.33 – Qualité de l'extraction au niveau transistor du temps de propagation de τ_{rise} et τ_{fall} à t_0 . Chaque couleur correspond à un FPGA

Afin de connaître la fiabilité séparément du PMOS et du NMOS, nous extrayons la dégradation dans le PMOS et le NMOS. Comme expliqué dans la section 1.3, nous considérons que lorsqu'un niveau logique haut est en entrée de la porte de transfert, c'est le PMOS qui est passant, alors que lorsque c'est un niveau logique bas c'est le NMOS qui est passant. Contrairement à l'inverseur, pour qui le PMOS est passant lorsqu'un niveau logique bas est en entrée et le NMOS est passant lorsqu'un niveau logique haut est en entrée. Alors, le système (3.24) devient :

$$\begin{bmatrix} R_{ILUT_{A1,1}} & R_{ILUT_{A2,1}} & R_{ILUT_{A3,1}} & R_{ILUT_{A4,1}} & R_{ILUT_{A5,1}} & R_{ILUT_{A6,1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ R_{ILUT_{A1,n}} & R_{ILUT_{A2,n}} & R_{ILUT_{A3,n}} & R_{ILUT_{A4,n}} & R_{ILUT_{A5,n}} & R_{ILUT_{A6,n}} \end{bmatrix} \cdot \begin{bmatrix} N_{PMOS_{A1}} & N_{NMOS_{A1}} \\ N_{PMOS_{A2}} & N_{NMOS_{A2}} \\ N_{PMOS_{A3}} & N_{NMOS_{A3}} \\ N_{PMOS_{A4}} & N_{NMOS_{A4}} \\ N_{PMOS_{A5}} & N_{NMOS_{A5}} \\ N_{PMOS_{A6}} & N_{NMOS_{A6}} \end{bmatrix} \cdot \begin{bmatrix} D_{PMOS}(t) \\ D_{NMOS}(t) \end{bmatrix} = \begin{bmatrix} \tau_1(t) \\ \vdots \\ \tau_n(t) \end{bmatrix} \quad (3.26)$$

où la matrice $N_{PMOS|NMOS}$ correspond au nombre de PMOS et NMOS passant en fonction de l'entrée utilisée, la matrice D correspond au temps de propagation d'un PMOS et NMOS à l'instant t , et la matrice $\tau(t)$ correspond au temps de propagation d'un front montant ou descendant à l'instant t . Nous utilisons les mêmes ROs que pour la régression à t_0 mais en distinguant τ_{fall} et τ_{rise} . Lorsque que le front en entrée de la LUT est un front montant alors la matrice $N_{PMOS|NMOS}$ correspond à :

$$\begin{bmatrix} 5 & 3 \\ 4 & 3 \\ 3 & 2 \\ 3 & 1 \\ 2 & 0 \\ 1 & 0 \end{bmatrix} \quad (3.27)$$

dans le cas d'un front descendant la matrice $N_{PMOS|NMOS}$ devient :

$$\begin{bmatrix} 3 & 5 \\ 3 & 4 \\ 2 & 3 \\ 1 & 3 \\ 0 & 2 \\ 0 & 1 \end{bmatrix} \quad (3.28)$$

Les figures 3.34a et 3.34b révèlent une prédiction bruitée de la dégradation du temps de propagation dans les LUTs à partir de la dégradation des PMOS et NMOS. Ce bruit provient du fait que notre extraction se base sur des valeurs qui sont issues elles-mêmes de l'extraction LUT/routages, du bruit de mesure ainsi que de la non connaissance exacte de l'architecture interne.

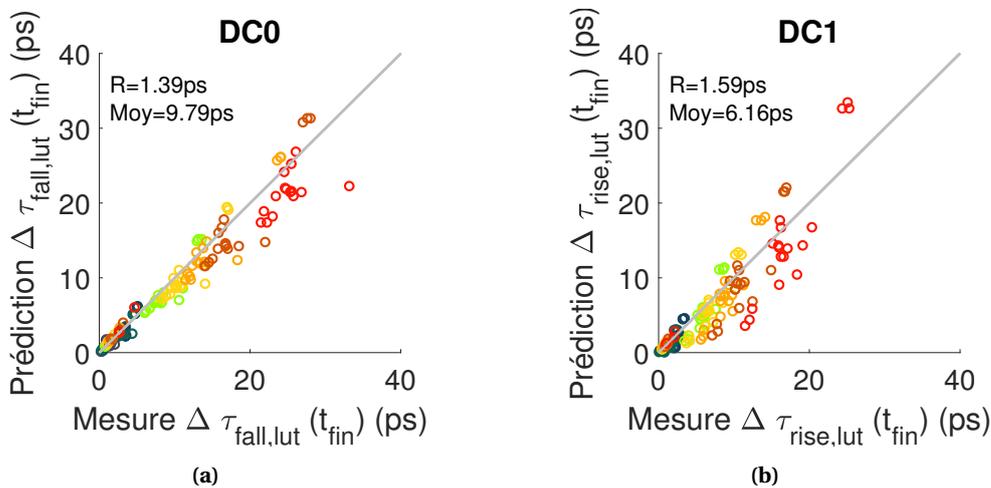


FIGURE 3.34 – Qualité de l'extraction de la dégradation au niveau transistor (a) τ_{fall} avec un stress DC0 (b) de τ_{rise} avec un stress DC1

La figure 3.35 présente la dégradation du temps de propagation du PMOS et NMOS, étant donné les conditions de stress (haute température, haute tension et stress statique) nous supposons qu'il s'agit de NBTI et de PBTI. Nous avons tracé la dégradation relative par rapport au temps de propagation à t_0 . De manière générale on constate que le PBTI est plus important que le NBTI. Ce résultat est cohérent avec ce qui est observé dans [2] dans une étude similaire à la nôtre. Cependant, comme évoqué dans la section 1.5.2, les fondeurs de FinFET, TSMC [28] et Intel [21] montrent que le PBTI devient négligeable comparé au NBTI. La contradiction entre [21, 28] et notre résultat remet en question la méthode d'extraction des dégradations à l'échelle du transistor. En effet, nous ne connaissons pas avec certitude l'architecture interne de la LUT. Notre interprétation de l'architecture interne de LUT est solide puisqu'elle se base sur le brevet [44] ainsi que sur les résultats que nous avons présentés sur la figure 3.26, cependant, il suffit que la LUT possède un inverseur de plus ou de moins par rapport à l'architecture que nous avons proposée (voir figure 1.19) et c'est le NBTI qui devient plus important que le PBTI.

Lorsqu'une étude de vieillissement est réalisée sur des échantillons de transistors, comme c'est le cas dans [21, 28], les tests sont de courte durée (quelques heures) et les conditions de stress en tension et en température sont élevées afin d'observer rapidement des dégradations. Nous avons réalisé une étude de vieillissement sur quelques milliers d'heures et dans des conditions de stress

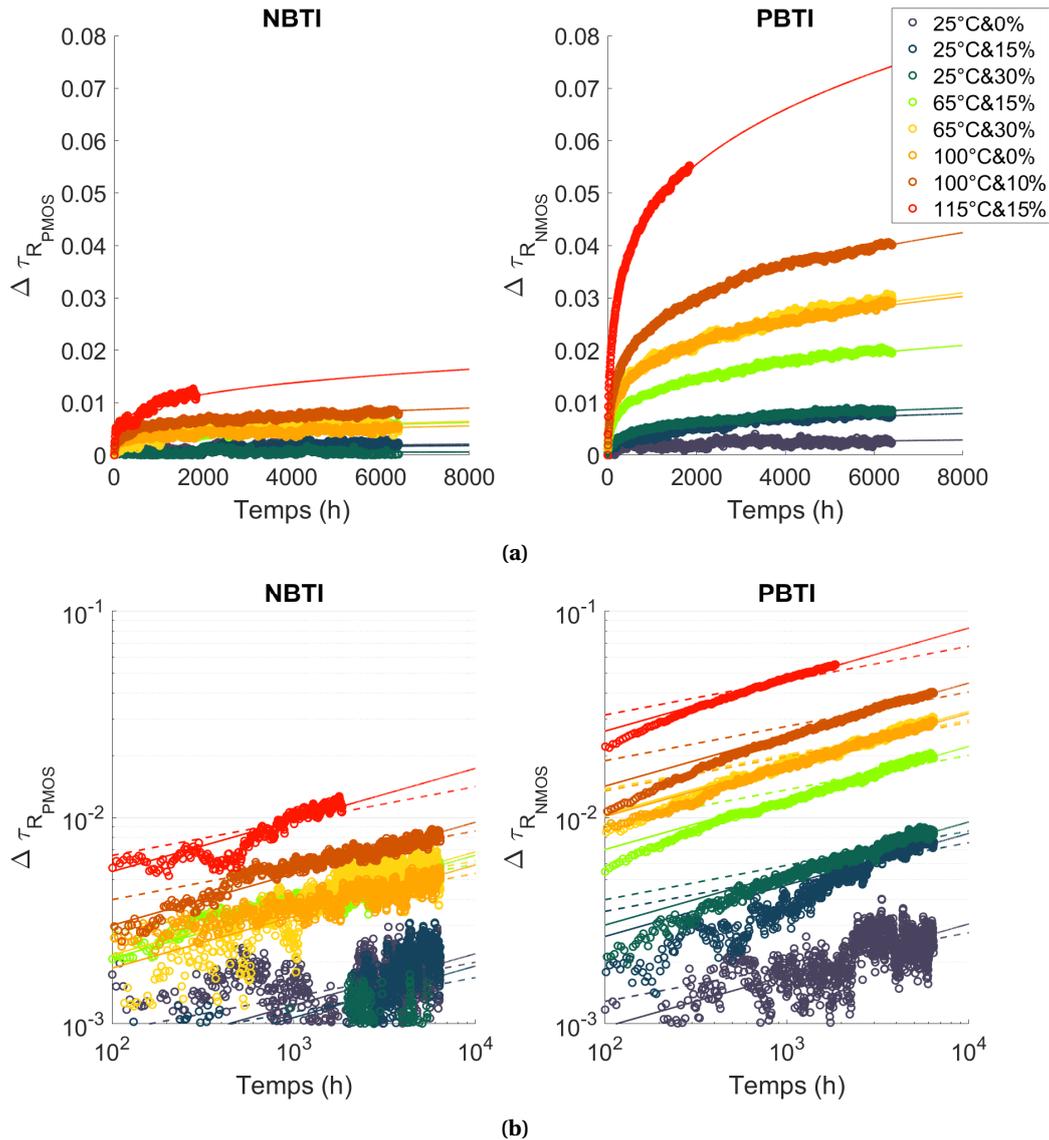


FIGURE 3.35 – Dégradation du temps de propagation d'un PMOS et du NMOS (a) en échelle linéaire et modèle puissance avec un exposant au temps fixé à 1/4 (b) en échelle loglog et modèle puissance avec un exposant au temps fixé à 1/4 (traits pleins) et 1/6 (pointillés)

proches des conditions d'utilisation. Si notre extraction est juste, le résultat obtenu signifie que le PBTI est plus important que le NBTI sur des longues durées, et à l'inverse, le NBTI est plus important que le PBTI pour des courtes durées [21, 28]. Dans la section 1.5.2, la littérature montre que le NBTI, tout comme le PBTI, est en réalité divisé en deux sous mécanismes : à court terme (<10 heures) c'est la diffusion d'hydrogène qui est dominante et à plus long terme c'est la diffusion de dihydrogène. Selon le modèle de réaction-diffusion, la diffusion d'hydrogène génère des défauts d'interface selon une loi puissance avec un exposant au temps de 1/4, tandis que, la diffusion de dihydrogène suit une loi puissance avec un exposant au temps de 1/6. La figure 3.35b montre qu'après 1000 heures de vieillissement les dégradations semblent suivre une loi puissance avec un exposant au temps de 1/4 ce qui révèle la génération de défauts d'interface dus à la diffusion d'hydrogène. Pour le NBTI, les mesures sont trop bruitées en dessous de 1000 heures. Pour le PBTI, les dégradations pour un vieillissement inférieur à 1000 heures on constate que l'exposant au temps est proche de 0.5, ce qui potentiellement dévoile un deuxième mécanisme avant 1000 heures de vieillissement.

En conclusion, nous avons montré dans cette partie comment dissocier les dégradations du

PBTI et du NBTI en nous basant sur les travaux de [2]. Les résultats obtenus montrent que les dégradations du PBTI sont plus importantes que celles du NBTI. Après 1000 heures de vieillissement, les dégradations suivent une loi puissance avec un exposant au temps de 1/4 ce qui peut correspondre à la génération de défauts d'interfaces dus à la diffusion d'hydrogène. Parce que l'extraction des dégradations à l'échelle du transistor montre un résidu assez important (voir figure 3.34), nous n'utiliserons plus les dégradations à l'échelle du transistor dans la suite de cette thèse.

3.5.6 Conclusion

Dans cette section nous avons montré une méthode permettant d'extraire le temps de propagation à t_0 des ressources de routages et de LUT constituant chacun des RO. Ainsi, nous avons vu que la majorité du temps de propagation d'un RO est induit par les ressources de routage, et cela même pour les RO contraints à être implémenté dans une seule et même slice. Dans cette continuité, nous avons développé une méthode permettant d'extraire les dégradations des ressources de routages et de LUT individuellement pour toutes les conditions de stress : température, tension et stress interne. Bien que la majorité du temps de propagation dans un RO soit générée par les ressources de routages, on constate que la majorité des dégradations du temps de propagation d'un RO est générée par les ressources de LUT. Nous avons identifié la contribution des ressources de routages et de LUT pour la dégradation du temps de propagation d'un front montant et descendant. L'amplitude des dégradations du temps de propagation dans les ressources de LUT et de routages sont proportionnelles au temps de propagation initiale, contrairement à la dégradation du RO qui contient les dégradations du routage et des LUTs. Finalement, nous avons extrait les dégradations du PMOS et du NMOS pour un stress statique afin d'identifier séparément le PBTI et le NBTI. On obtient des dégradations plus importantes pour le PBTI que pour le NBTI.

3.6 Conclusion

Dans ce chapitre nous avons introduit le plan d'expérience et le banc de test que nous utilisons pour mesurer la fiabilité des FPGAs. Le banc de test est conçu pour vieillir et mesurer les dégradations de neuf FPGAs Zynq UltraScale+ 16nm pendant plus de 10000 heures. L'architecture implémenté dans les neuf FPGAs est identique et correspond à celle présentée dans le chapitre 2. Chacun des neuf FPGA est soumis à un stress de température et de tension différent compris entre $25^{\circ}\text{C} \leq T_{stress} \leq 115^{\circ}\text{C}$ et $V_{nom} \leq V_{stress} \leq 1.3V_{nom}$. Nous choisissons le niveau de stress de température et de tension pour favoriser l'observation du BTI/TDDDB (haute température et haute tension) ou du HCI (basse température et haute tension).

Avant de vieillir les FPGAs, nous avons mesuré la fréquence d'oscillation de tous les ROs. Nous avons observé des variations importantes (plus du double) de la période d'oscillation de ROs de même architecture logique (même code VHDL), ce qui soulève l'importance de l'architecture physique de chacun des ROs dans le FPGA. Nous avons comparé les périodes des ROs calculées à partir de la mesure avec les périodes des ROs données par Vivado. Cette comparaison montre la cohérence entre nos mesures et les données de Vivado, ce qui prouve la capacité de notre banc de test à mesurer la fréquence des ROs. De plus, avec cette comparaison nous avons situé l'état de performance initial de nos FPGAs qui est entre 50% et 70% des meilleures performances attendues par Vivado. Aussi, la mesure initiale du rapport cyclique d'oscillation de tous les ROs montre que celui-ci est en moyenne de tous les ROs et FPGA confondus de 49.98% pour une température et une tension nominale. Ce qui veut dire qu'avant vieillissement le temps de propagation dans un PMOS et dans un NMOS est équilibré. Toujours avant vieillissement, nous avons caractérisé la fréquence et le rapport cyclique d'oscillation de tous les ROs dans tous les FPGAs en fonction de la température et de la tension. Nous avons montré l'intérêt d'utiliser un modèle propre à chaque RO pour la compensation post-mesure des variations résiduelles non compensées par les systèmes de régulation de température et de tension.

Après avoir vieillis 5103 ROs dans neuf FPGAs différents pendant 8000 heures, nous avons fait une analyse préliminaire des dégradations. L'effet du vieillissement sur tous les ROs est la dérive de la fréquence relative. Nous n'avons pas observé de conséquence du TDDDB sur le comportement d'un RO. Nous avons mis en évidence l'effet de la tension et de la température qui accroît les dégradations. Par exemple, la dégradation des ROs après 10000 heures de fonctionnement est 2.5 fois plus importante lorsque le FPGA est utilisé à 115°C au lieu de 65°C. Nous avons montré que les dégradations sont plus importantes pour un stress statique à haute température et pour un stress dynamique à basse température, ce qui nous laisse à penser que nous observons respectivement du BTI et du HCI. Nous obtenons des dégradations maximales de l'ordre du 1% après 8000 heures de vieillissement et dans les pires conditions de stress (haute température, haute tension et stress statique) ce qui nous conforte quant à la fiabilité de la technologie.

En comparant les mesures de dégradations dans les ROs SP (Short Path) et LP (Long Path), nous avons montré que les dégradations de la fréquence d'oscillation du RO sont plus importantes lorsqu'il est constitué principalement de ressources de LUT que lorsqu'il est constitué principalement de ressources de routage. D'ailleurs, nous observons aussi que la dégradation d'un RO L3 XB I1 qui utilise des LUTs dont l'entrée utilisée correspond au chemin le plus long (A1) est plus importante que la dégradation d'un RO L3 XB I5 qui utilise des LUTs dont l'entrée utilisée correspond au chemin le plus court (A5), et nous montrons que la proportion de ressources de LUT utilisée est plus importante dans le RO L3 XB I1 que dans le RO L3 XB I5. En revanche, nous avons montré que le nombre d'entrées utilisées dans la LUT n'a pas d'effet sur l'amplitude des dégradations, que ce soit pour un stress statique ou dynamique.

Parce que nous avons observé que l'effet du vieillissement n'a pas la même conséquence sur la dégradation des ressources de LUTs et de routage, nous avons développé une méthode pour dissocier les dégradations dans les deux types de ressources.

Dans un premier temps nous avons identifié l'accumulation du temps de propagation dans les ressources de LUT et de routage avant le vieillissement. Pour faire cela nous avons appliqué une régression linéaire multiple entre la période de chacun des ROs et les ressources de LUT et de routage extrait de Vivado pour chacun des ROs. Nous obtenons ainsi le temps de propagation dans chacune des ressources de LUT et de routage. À partir des résultats de cette première régression linéaire nous pouvons calculer la période initiale de tous les ROs à partir des ressources de LUT et de routage qu'ils utilisent avec un résidu moyen de 2.5%. Aussi, à partir des données de la régression, on constate que la majorité du temps de propagation d'un RO est dû aux ressources de routage aussi bien pour les RO SP et encore plus pour les RO LP.

Dans un deuxième temps, nous avons réduit le nombre de paramètres dans le système de régression linéaire en remplaçant les ressources de routage par le temps de propagation initial dans les ressources de routage. Ainsi, nous pouvons appliquer la régression linéaire individuellement pour chacun des 21 stress interne et pour chacun des neuf FPGAs. En réitérant la régression linéaire à chaque instant de mesure, nous obtenons ainsi la dégradation du temps de propagation dans les ressources de LUT et de routage. Ainsi, nous pouvons à partir des ressources de LUT et le temps de propagation initial dans les ressources de routage prédire les dégradations de la période de tous les ROs avec un résidu moyen de 5%. À partir des données de cette régression on constate que la majorité de la dégradation de la période d'un RO SP est due à la dégradation des ressources de LUT.

Finalement, en utilisant la dégradation du temps de propagation identifié dans les ressources de LUT et à partir d'une interprétation de l'architecture d'une LUT nous avons identifié le temps de propagation initial d'un transistor qui est de 14.2ps pour un PMOS et un NMOS à $T = 25^\circ\text{C}$ et $V = V_{nom}$. Ensuite, nous avons identifié en fonction du stress interne (DC0 ou DC1) le type de transistor vieillit et nous avons extrait les dégradations pour un PMOS et un NMOS dans les conditions de stress favorisant le BTI. Nous obtenons des dégradations plus importante pour le PBTI que pour le NBTI. Ce résultat n'est pas une certitude puisqu'il dépend de notre interpréta-

tion de l'architecture interne d'une LUT. Dans la littérature, des mesures de fiabilité faites dans des conditions de stress importantes et pendant quelques heures sur des 16nm FinFET montrent que le comportement inverse est obtenu : plus de NBTI et moins de PBTI [21, 28]. Si notre interprétation de l'architecture interne de la LUT est bonne, notre résultat révèle l'importance de faire des tests de vieillissement de longue durée et dans des conditions de stress proches des conditions d'utilisation comme nous avons fait.

Chapitre 4

Modélisation et prévision de la fiabilité des FPGAs

4.1 Introduction

Dans le chapitre précédent, nous avons présenté la dégradation des ROs mesurée par notre banc de test. Nous avons également mis en évidence les effets de la température, de la tension et du signal de stress interne sur les dégradations. Afin de prévoir la fiabilité du FPGA, il est nécessaire de développer des modèles en tenant compte de l'effet des différents stress. Aussi nous avons montré, parce que les dégradations dans les ressources de LUT et de routage sont différentes, comment extraire les dégradations dans les deux ressources séparément à partir des dégradations mesurées dans les ROs. Ainsi, dans ce chapitre nous construirons le modèle de dégradation d'un RO à partir de la modélisation dans les ressources de LUT et de routage séparément.

Concernant les dégradations dans les ressources de LUT et de routage, nous diviserons la modélisation en trois sections :

- section 4.2 : Nous modéliserons les dégradations dans les ressources de LUT et de routage avec des stress statiques DC0 et DC1. Nous modéliserons l'effet de la température de stress avec une loi d'Arrhenius et l'effet de la tension de stress avec une loi exponentielle. Nous identifierons un coefficient d'amplitude de dégradation propre à chaque entrée de LUT. Aussi, nous montrerons une relation proportionnelle entre l'amplitude des dégradations relatives et le temps de propagation dans les ressources de routage.
- section 4.3 : En reconstruisant les dégradations pour un stress dynamique à partir de notre modèle de dégradation pour un stress statique, nous montrerons clairement l'effet du mécanisme de récupération du BTI. Ainsi, nous modéliserons le mécanisme de récupération pour différents rapport cyclique de stress. En assemblant les modèles de dégradations des ressources de LUT et de routage nous montrerons la justesse de notre modèle à prédire les dégradations des ROs ;
- section 4.4 : Nous analyserons l'effet de la fréquence de stress sur les dégradations afin de mettre en avant l'effet du HCI ;

Finalement, en combinant les modèles développés dans chacune des sections, nous obtiendrons un modèle unique qui prédit la dégradation du temps de propagation d'un front montant et descendant dans un RO en fonction du temps de fonctionnement, de la température d'utilisation, de la tension d'utilisation, du type de signal utilisé (fréquence et rapport cyclique) et des ressources physiques utilisées pour implémenter le RO.

Afin de situer la fiabilité de notre FPGA par rapport à un FPGA conçu à partir d'une technologie de transistor planaire, dans la section 4.5, nous comparerons les dégradations des ROs que nous avons mesurés sur un FPGA Zynq UltraScale+ 16nm FinFET avec les modèles semi-

empiriques obtenus par Julien Coutet dans [2] pour la dégradations des ROs dans un FPGA Artix 28nm HKMG.

En considérant un seuil de défaillance, nous modéliserons les défaillances des ROs dans la section 4.6. Cela nous permettra de comparer les paramètres obtenus dans la modélisation des dégradations de ceux obtenus dans la modélisation des défaillances. De plus, nous présenterons des résultats de fiabilité directement exploitables par un industriel tel que le MTTF ou encore le taux de défaillance.

4.2 Modélisation des dégradations pour un stress statique

4.2.1 Introduction

Dans cette section, nous présentons la modélisation de l'effet du vieillissement pour un stress statique en détaillant les étapes de la modélisation, ce afin de valider la justesse du modèle et de confirmer la cohérence avec les mesures. Tout d'abord nous présentons les dégradations mesurées au niveau du RO, ensuite, parce que l'effet du vieillissement sur la dégradation des ressources logiques (LUT) et de routage n'est pas identique, nous modéliserons séparément les dégradations dans les deux types de ressources.

Pour modéliser les dégradations nous utilisons le temps de propagation d'un front montant (τ_{rise}) et descendant (τ_{fall}). En effet, pour l'analyse temporelle des circuits, le temps de propagation à prendre en considération est le plus grand des deux ($\max(\tau_{rise}, \tau_{fall})$). La période d'oscillation d'un RO ne donnant que la moyenne $\frac{\tau_{rise} + \tau_{fall}}{2}$, elle tend à minorer la valeur apparente de la dégradation, en particulier dans les cas où un des deux temps de propagation s'améliore comme nous montrerons dans le cas d'un stress statique.

En séparant les dégradations dans les ressources de LUT et de routage, la figure 3.32 montre qu'il y a une relation de proportionnalité entre l'amplitude des dégradations et le temps de propagation initial. Donc, dans des mêmes conditions de stress, l'amplitude des dégradations, quelle que soit l'architecture du RO, ne dépend que de son temps de propagation initial. C'est pourquoi nous utiliserons la dérive relative du temps de propagation d'un front descendant ($\Delta\tau_{fall_R}$) et montant ($\Delta\tau_{rise_R}$) pour modéliser les dégradations, tel que :

$$\Delta\tau_{front_R}(t) = \frac{\tau_{front}(t) - \tau_{front}(t_0)}{\tau_{front}(t_0)} \quad (4.1)$$

où *front* peut être montant (*rise*) ou descendant (*fall*).

4.2.2 Dégradations du temps de propagation dans un RO

Avant de modéliser les dégradations d'un RO en séparant les dégradations dans les ressources de LUT et de routage, nous allons présenter les dégradations mesurées dans un RO que l'on retrouve sur la figure 4.1. Cette figure montre que les dégradations du temps de propagation d'un front montant et descendant lorsque le RO subit un stress statique DC0 et DC1 et pour différentes conditions de tensions et de températures. Sur les figures 4.1a et 4.1d, c'est à dire lorsque le niveau logique après le front qui se propage pendant la mesure est identique de celui du stress, on observe une augmentation (dégradation) du temps de propagation. Sur les figures 4.1b et 4.1c, c'est à dire lorsque le niveau logique après le front qui se propage pendant la mesure est différent de celui du stress, on observe une diminution (amélioration) du temps de propagation.

Notre interprétation des résultats s'appuie sur la figure 4.2 qui représente le schéma simplifié d'une LUT configurée en buffer pour la propagation d'un front descendant en sortie (figure 4.2a) et d'un front montant en sortie (figure 4.2b). Aussi nous avons entouré en rouge les transistors vieilliss

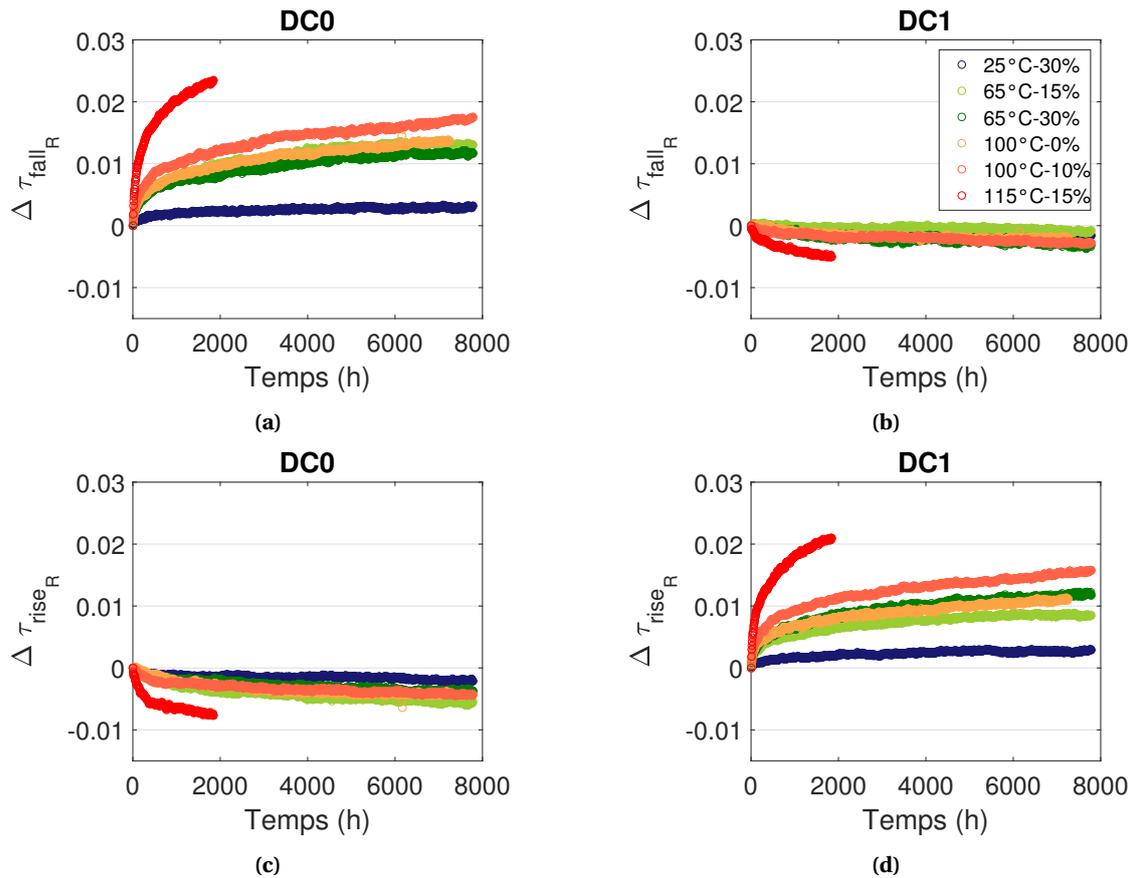


FIGURE 4.1 – Évolution de (a)(b) $\Delta \tau_{fall_R}$ (c)(d) $\Delta \tau_{rise_R}$ dans un RO L1 B pour différentes température et tension de stress et un signal de stress interne (a)(c) DC0 et (b)(d) DC1

dans la cas d'un stress statique DC0. Sur la figure 4.2a on constate que les transistors passant pour la propagation d'un front descendant sont les mêmes que les transistors vieillis par un stress DC0, tandis que, sur la figure 4.2b on constate que les transistors passant pour la propagation d'un front montant ne sont pas les mêmes que les transistors vieillis par un stress DC0.

De manière générale, un stress statique augmente la tension de seuil (V_{th}) du transistor passant. Dans un inverseur les conséquences dépendent du signal de stress :

- Entrées niveau bas : seul le PMOS est stressé; parce que le PMOS devient passant plus lentement et bloqué plus rapidement, à la sortie, le temps de propagation d'un front montant augmente et le temps de propagation d'un front descendant diminue,
- Entrées niveau haut : seul le NMOS est stressé; parce que le NMOS devient passant plus lentement et bloqué plus rapidement, à la sortie, le temps de propagation d'un front descendant augmente et le temps de propagation d'un front montant diminue.

Pour la porte de transfert, les conséquences dépendent aussi du signal de stress :

- Entrées niveau bas : seul le NMOS est stressé; parce que le NMOS devient passant plus lentement, à la sortie, le temps de propagation d'un front descendant augmente. Cependant, le temps de propagation d'un front montant ne change pas significativement puisque soit le signal passe par une autre porte de transfert non vieilli, soit le signal passe par la même porte de transfert le PMOS qui transmet le signal n'est pas vieilli et le temps de commutation du NMOS vieilli n'impacte pas la transmission d'un front montant (voir explication sur le fonctionnement d'une porte de transfert dans la section 1.3.2),
- Entrées niveau haut : seul le PMOS est stressé; parce que le PMOS devient passant plus lentement, à la sortie, le temps de propagation d'un front montant augmente. Cependant, le temps de propagation d'un front descendant ne change pas significativement.

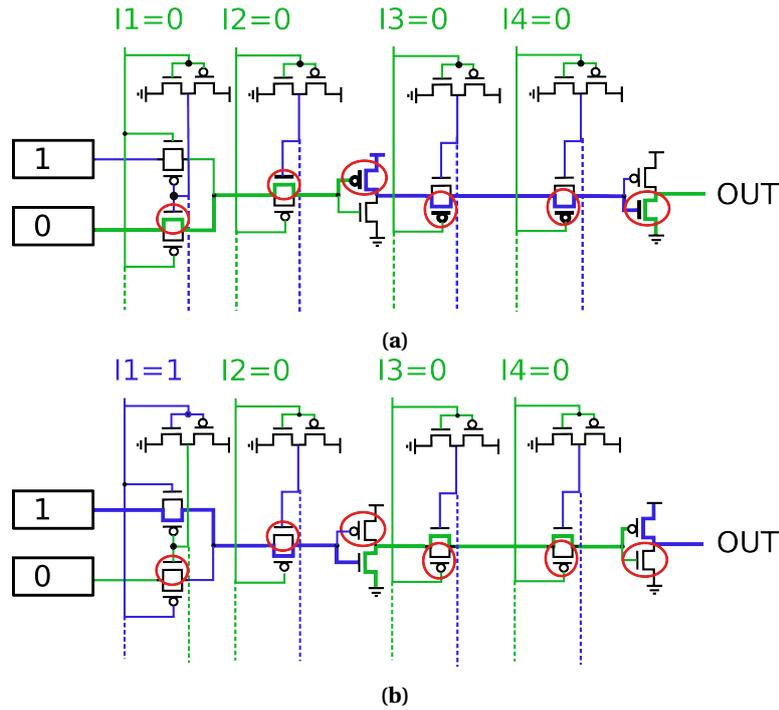


FIGURE 4.2 – Schéma d’une LUT à 4 entrées configurée en buffer où seule l’entrée I1 est utilisée les autres entrées sont fixées à ‘0’. Seuls les transistors utilisés dans cette configuration sont représentés. Identification du chemin (traits épais) de propagation pour conduire le signal entre les bits de configuration et la sortie (a) front descendant en sortie et d’un (b) front montant en sortie et identification des transistors dégradés par un stress DC0 (cercle rouge). Nous identifions un état haut en bleu et un état bas en vert.

Les effets de la dégradation du transistor sur le temps de propagation au niveau de la porte logique pour l’inverseur sont résumés sur la figure 4.3a et pour la porte de transmission sur la figure 4.3b.

Signal stress	Transistor stressé	Signal Mesuré	Invers.
'0'	PMOS	Descendant	
		Montant	
'1'	NMOS	Descendant	
		Montant	

(a)

Signal stress	Transistor stressé	Signal Mesuré : Front	Porte Transmis.
'0'	NMOS	Descendant	
		Montant	
'1'	PMOS	Descendant	
		Montant	

(b)

FIGURE 4.3 – Effet sur le temps de propagation d’un front montant et descendant en sortie d’une LUT en fonction du stress en entrée d’une LUT pour (a) un inverseur et (b) une porte de transfert

Parce qu’un RO est composé d’inverseurs et de portes de transmission, nous pouvons interpréter les résultats présentés sur la figure 4.4 à partir de la figure 4.3. Lorsque les transistors passant pour le signal mesuré sont identiques aux transistors stressés par le signal de stress, front descendant en sortie et DC0 en entrée ou front montant en sortie et DC1 en entrée, le temps de propagation augmente dans les inverseurs et dans les portes de transmission, par conséquent il augmente au niveau du RO (voir figure 4.4a et 4.4d).

Lorsque les transistors passant pour le signal mesuré sont différents aux transistors stressés par le signal de stress, front montant en sortie et DC0 en entrée ou front descendant en sortie et DC1 en entrée, alors le temps de propagation diminue dans les inverseurs et, ne change pas dans les portes de transmission, en conséquence de quoi il peut diminuer au niveau du RO (voir figure 4.4b et 4.4c).

4.2.3 Dégradation dans les ressources logiques

À présent nous utilisons les données extraites avec la méthode présentée dans la section 3.5 pour modéliser les dégradations dans les ressources logiques qui sont ici des ressources de LUT.

Observations et interprétations

La figure 4.4 expose l'évolution des temps de propagation de fronts montants et descendants pour des stress statiques DC0 et DC1. Sur les figures 4.4a et 4.4d on observe une augmentation (dégradation) du temps de propagation relatif plus importante que dans les ROs (voir figure 4.1). En effet, avec la figure 3.17 nous avons montré que la dégradation relative est plus importante dans les RO SP (avec beaucoup de LUT et moins de routage) que dans les RO LP (avec beaucoup de routage et moins de LUT). Ainsi, il est tout à fait cohérent en extrayant les dégradations dans les ressources de LUT uniquement que les dégradations relatives soient plus importante que dans le RO.

Sur les figures 4.4b et 4.4c, on constate que les améliorations sont plus faibles dans les ressources de LUT que dans un RO (voir figure 4.1). Nous avons montré précédemment (voir la table 4.3) que les améliorations du temps de propagation sont générées uniquement dans un inverseur. Or, dans une LUT le nombre d'inverseurs par rapport au nombre de portes de transferts est faible. Ceci explique les faibles améliorations du temps de propagation (inférieurs à 0.1%).

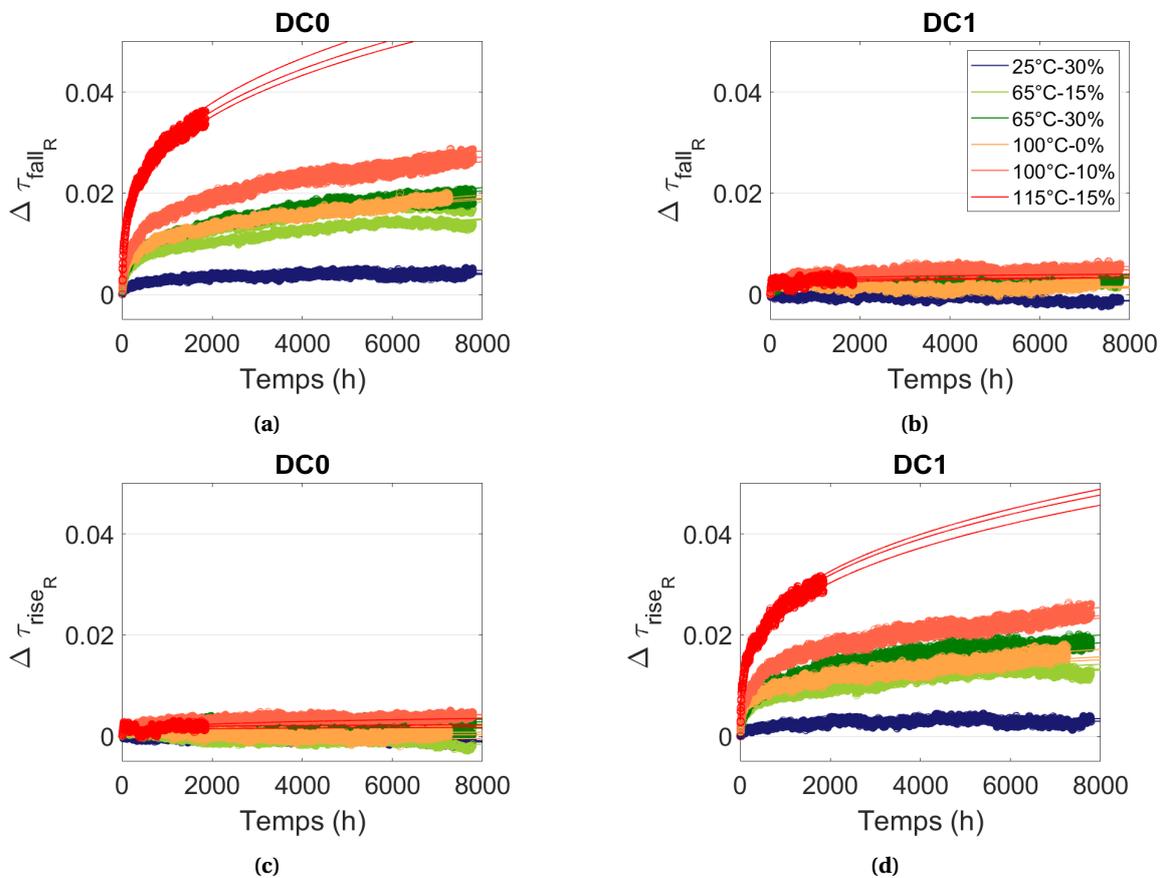


FIGURE 4.4 – Évolution de (a)(b) $\Delta \tau_{fall_R}$ (c)(d) $\Delta \tau_{rise_R}$ dans les ressources de LUT pour différentes température et tension de stress et un signal de stress interne (a)(c) DC0 et (b)(d) DC1 et pour trois RO L1 B

Modélisation

Modèle 1 : loi puissance

Afin de modéliser le vieillissement des LUTs, nous avons appliqué une 1^{ère} régression non linéaire sans changement de variable et en minimisant l'erreur entre les temps de propagations mesurés et une loi puissance en fonction du temps : $\Delta\tau = a \times t^b$ (voir traits pleins sur la figure 4.4) pour toutes les architectures de ROs de type buffer dans les six conditions de températures et de tensions que l'on retrouve sur la figure 4.4. Nous avons tracé le résultat des pré-facteurs a sur la figure 4.5a et des exposants b sur la figure 4.5b avec un intervalle de confiance à 95%. La figure 4.5a manifeste une dépendance du pré-facteur a aux conditions de stress de température et de tension. La figure 4.5b montre que l'exposant b est stable (entre 0.2 et 0.28) pour les dégradations. Pour les améliorations, parce que les amplitudes sont dans le bruit de mesure, l'intervalle de confiance est plus important et l'exposant b plus dispersé.

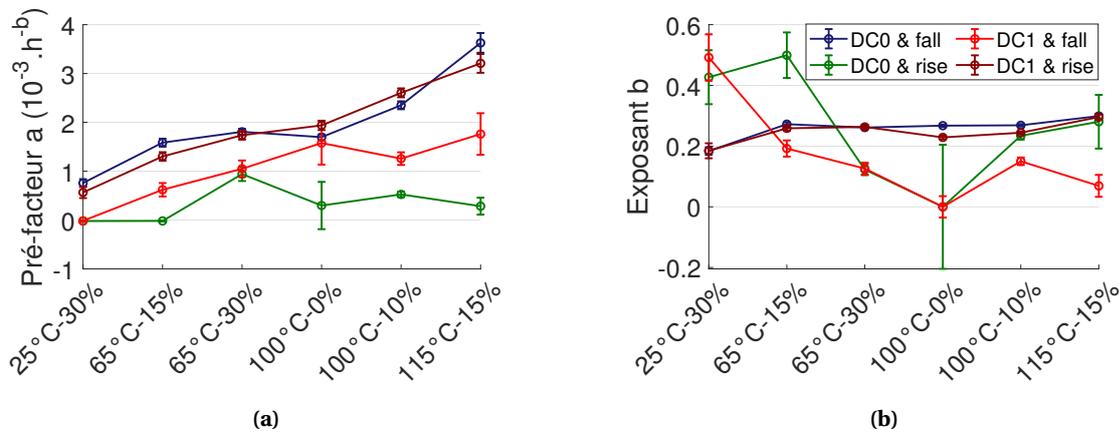


FIGURE 4.5 – Pré-facteur a et exposant b du modèle $\Delta\tau = a \times t^b$ en fonction de la condition de stress en température et tension pour un RO L1 B

Dans la loi de puissance, le pré-facteur du temps a détermine l'amplitude tandis que l'exposant du temps b détermine la dynamique. Dans les composants numériques, les mécanismes tels que le BTI ou le HCI définissent la dynamique des dégradations. Parce nous modélisons dans cette section les dégradations pour un stress statique on peut considérer que le mécanisme de dégradation est le BTI car le HCI n'intervient que pendant les commutations. Donc, la dynamique des dégradations est identique pour toutes les conditions de stress de température et de tension. Ainsi, nous fixons l'exposant au temps (b) de la loi puissance en prenant la moyenne des exposants b obtenus à toutes les conditions de stress. Nous obtenons un exposant $b = 0.25$ pour les dégradations et $b = 0.22$ pour les améliorations. L'exposant au temps (b) obtenu pour les dégradations et les améliorations est proche de celui qui caractérise le modèle de Reaction-Diffusion [159] ($b = 0.25$).

Modèle 2 : avec modélisation de l'effet de la température et de la tension

Après avoir fixé le coefficient b , indépendamment pour chaque architecture de RO, nous faisons une 2^{ème} régression non linéaire. Dans ce 2^{ème} modèle, nous introduisons le facteur d'accélération en tension (AF_V) qui suit une loi exponentielle ainsi que le facteur d'accélération en température (AF_T) qui suit une loi d'Arrhenius, tel que :

$$\Delta\tau_R = at^b = (A \times AF_V \times AF_T) t^b = A \times e^{\gamma(V_{stress} - V_{op})} e^{\frac{E_a}{k_B} \left(\frac{1}{T_{op}} - \frac{1}{T_{stress}} \right)} t^b \quad (4.2)$$

où A est le facteur d'amplitude, γ est le paramètre d'accélération électrique, V_{op} est la tension nominale : 0.85V, V_{stress} est la tension de stress, E_a est l'énergie d'activation, k_B la constante de Boltzmann, T_{op} est la température nominale : 298K and T_{stress} est la température de stress.

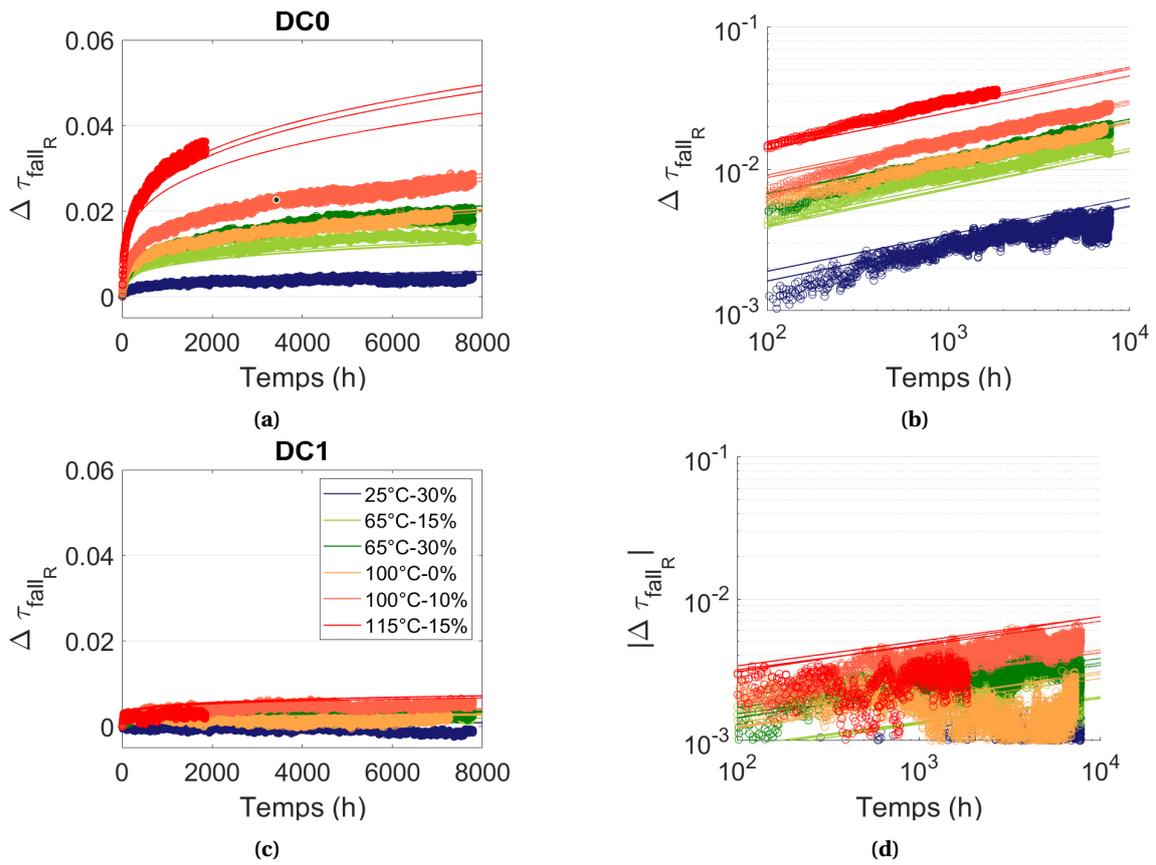


FIGURE 4.6 – Évolution de $\Delta \tau_{fallR}$ en (a)(c) échelle linéaire et (b)(d) log-log pour différentes température et tension de stress et un signal de stress interne (a)(b) DC0 et (c)(d) DC1 et pour trois RO L1 B

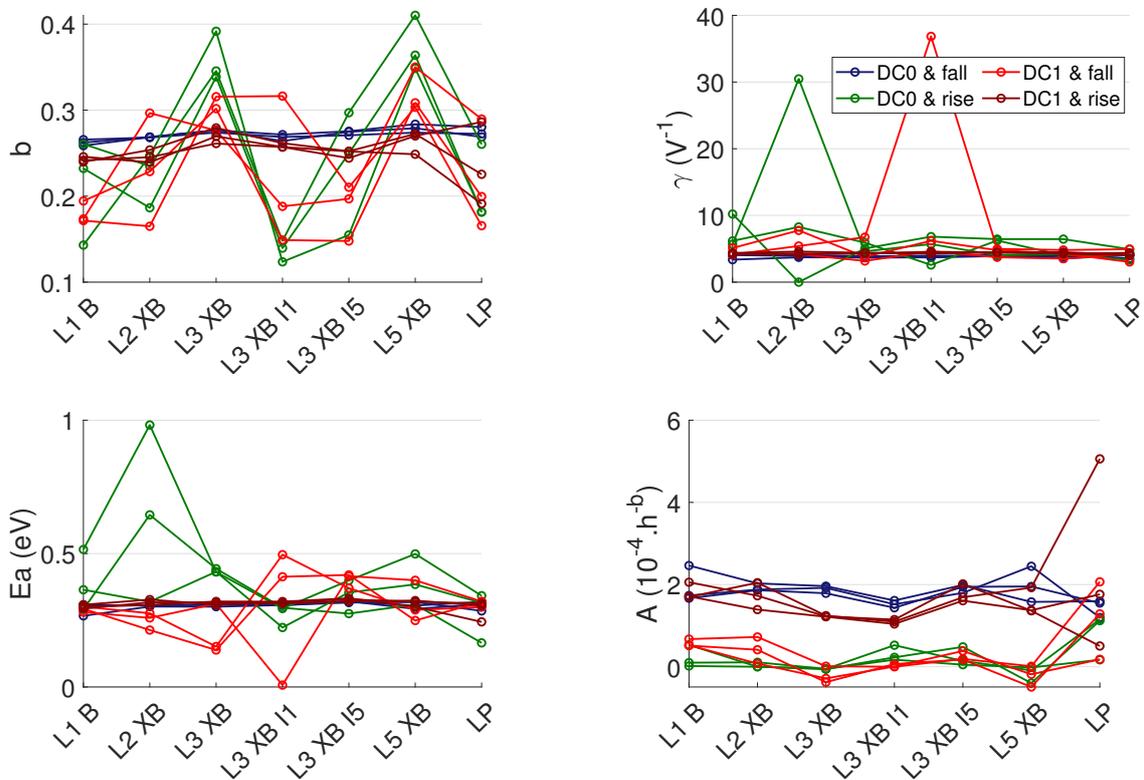


FIGURE 4.7 – Coefficients b , γ , Ea et A du modèle (4.2) en fonction de l'architecture du RO

La figure 4.6 confirme la cohérence entre les mesures et le modèle de l'équation (4.2), aussi, nous traçons sur la figure 4.7 les coefficients obtenus lors de la modélisation pour différentes architectures. Pour les améliorations, on note une dispersion des paramètres assez importante qui complique la modélisation. Pour les dégradations on remarque que les paramètres b , γ et Ea sont assez stables suivant les architectures. Ce qui veut dire que le même mécanisme de dégradation est observé dans chacune des architectures de RO et que la dépendance de l'amplitude des dégradations à la tension et à la température de stress est identique.

Modèle 3 : avec modélisation de l'architecture du RO

Suite à cette dernière remarque, nous avons réalisé une 3^{ème} régression linéaire à partir du modèle (4.2) mais cette fois-ci en prenant les paramètres b , γ et Ea communs à toutes les architectures de ROs. À cette étape de la modélisation, nous avons un modèle par RO avec comme seule différence entre chaque modèle le coefficient d'amplitude A . Sur la figure 4.8a nous avons rapporté l'amplitude du coefficient A de tous les ROs en fonction du temps de propagation à t_0 et aux conditions nominales. Sachant que nous modélisons la dégradation relative des ROs, le coefficient A devrait être constant. Or, sur cette figure nous voyons que ce n'est pas le cas. En effet, nous avons observé dans la section 3.5.5 que le nombre de PMOS et NMOS stressés dépend du numéro de l'entrée utilisée et que l'amplitude des dégradations est différente suivant que le transistor subisse du NBTI pour le PMOS ou du PBTI pour le NMOS. Par conséquent, l'amplitude des dégradations ne dépend pas que du temps de propagation initial dans la LUT mais dépend aussi de l'entrée de la LUT utilisée.

Ainsi, nous considérons que chaque entrée de LUT possède un coefficient d'amplitude A_{ILUT} et que la somme de ce coefficient pour chaque entrée utilisée correspond au coefficient d'amplitude A du RO tel que :

$$\begin{aligned}
 R_{ILUT}A_{ILUT} &= \begin{bmatrix} R_{ILUT_{A1,1}} & R_{ILUT_{A2,1}} & R_{ILUT_{A3,1}} & R_{ILUT_{A4,1}} & R_{ILUT_{A5,1}} & R_{ILUT_{A6,1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ R_{ILUT_{A1,n}} & R_{ILUT_{A2,n}} & R_{ILUT_{A3,n}} & R_{ILUT_{A4,n}} & R_{ILUT_{A5,n}} & R_{ILUT_{A6,n}} \end{bmatrix} \cdot \begin{bmatrix} C_{ILUT_{A1}} \\ C_{ILUT_{A2}} \\ C_{ILUT_{A3}} \\ C_{ILUT_{A4}} \\ C_{ILUT_{A5}} \\ C_{ILUT_{A6}} \end{bmatrix} \\
 &= \begin{bmatrix} A_1 \times \tau_{t01} \\ \vdots \\ A_n \times \tau_{t0n} \end{bmatrix} = \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix} \quad (4.3)
 \end{aligned}$$

où $R_{ILUT_{A1,n}}$ correspond au nombre de fois où l'entrée $A1$ d'une LUT est utilisée dans le $n^{ème}$ RO, $C_{ILUT_{A1}}$ est le coefficient d'amplitude absolu propre à l'entrée de LUT, A_n est le coefficient d'amplitude relatif du $n^{ième}$ RO et C_n est le coefficient d'amplitude absolu du $n^{ième}$ RO. En faisant une régression linéaire par la méthode des moindres carrés ordinaire, nous obtenons la matrice C_{ILUT} . Nous utilisons C_n le coefficient absolu d'amplitude du $n^{ième}$ RO ($C_n = A_n \times \tau_{t0n}$) pour que la régression soit fondée sur une relation physique. La figure 4.8b confirme la qualité de la prédiction pour obtenir le coefficient d'amplitude du RO à partir des coefficients d'amplitudes propres à chaque entrée de LUT. Nous avons rapporté dans la table 4.1 l'amplitude absolue des coefficients propres à chaque entrée de LUT. Globalement, pour les dégradations, on remarque que le coefficient C_{ILUT} est minimum pour l'entrée $A6$ et qu'il est irrégulièrement croissant jusqu'à l'entrée $A1$. De plus, pour les dégradations tous les coefficients sont positifs, tandis que certains coefficients sont négatifs pour les améliorations, confirmant ainsi l'amélioration du temps de propagation.

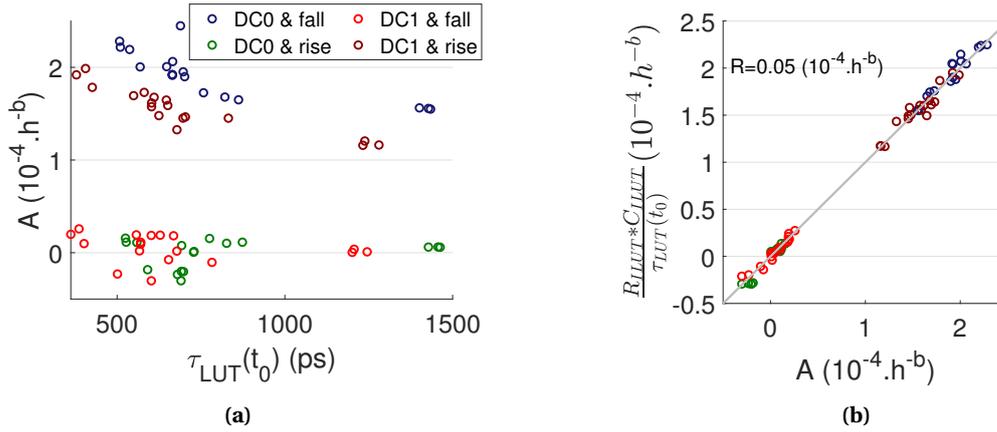


FIGURE 4.8 – (a) Paramètre d’amplitude A en fonction du temps de propagation à t_0 et dans les conditions nominales ($T = 25^\circ\text{C}$ et $V = V_{nom}$) (b) Qualité de la régression pour prédire l’amplitude du coefficient A à partir des entrées de LUT utilisées dans le RO

À ce niveau de la modélisation il est difficile d’interpréter plus en profondeur les résultats de la table 4.1 car on se situe déjà à plusieurs niveaux d’extraction et de modélisation. Cependant, nous disposons à présent de quatre modèles de dégradation, communs à tous les RO pour prédire la dégradation du temps de propagation relatif d’un front F (montant ou descendant) en fonction du stress statique S (DC1 ou DC0) dans les ressources (Res) de LUT tel que :

$$\Delta\tau_{ES,Res} = a_{ES,Res} t^{b_{ES,Res}} = \frac{R_{ILUT} \times C_{ILUT_{ES,Res}}}{\tau_{LUT}(t_0)} \times e^{Y_{ES,Res}(V_{stress}-V_{op})} e^{\frac{Ea_{ES,Res}}{k_B} \left(\frac{1}{T_{op}} - \frac{1}{T_{stress}} \right)} t^{b_{ES,Res}} \quad (4.4)$$

La table 4.1 rapporte les valeurs numériques des coefficients du modèle. Les coefficients R_{ILUT} et $\tau_{LUT}(t_0)$ sont des constantes propres à l’architecture physique de chaque RO.

TABLE 4.1 – Coefficients du modèles (4.4) pour prédire l’effet du vieillissement pour un stress statique sur les LUT

Effet vieillissement	F,S,Res	Coefficients modèle (4.4)				
		A_{ILUT} ($10^{-3} ps.h^{-b}$)	γ (V^{-1})	Ea (eV)	b	
Dégradation	fall,DC0,LUT	A1	15.4	4.21	0.31	0.27
		A2	17.1			
		A3	11.4			
		A4	12.2			
		A5	7.6			
		A6	5.1			
	rise,DC1,LUT	A1	10.0			
		A2	10.6			
		A3	2.9			
		A4	12.8			
		A5	4.9			
		A6	1.4			
Amélioration	fall,DC1,LUT	A1	0.4	5.98	0.31	0.24
		A2	5.9			
		A3	4.4			
		A4	-4.8			
		A5	0.6			
		A6	-2.1			
	rise,DC0,LUT	A1	1.1			
		A2	-1.2			
		A3	-2.1			
		A4	-1.8			
		A5	0.7			
		A6	-3.9			

Validation modèle final

La figure 4.9 confirme la cohérence entre le modèle et la mesure pour trois RO L1 B et la figure 4.10 montre la justesse de la prédiction du modèle. Sur cette dernière figure, nous avons tracé $\Delta\tau_R$, mesuré et prédit, au dernier instant de mesure, pour le front montant et descendant de tous les RO stressés par DC0 et DC1. Cette figure confirme la capacité de notre modèle à prédire l'effet du vieillissement, pour un stress statique et ce quelle que soit l'architecture du RO.

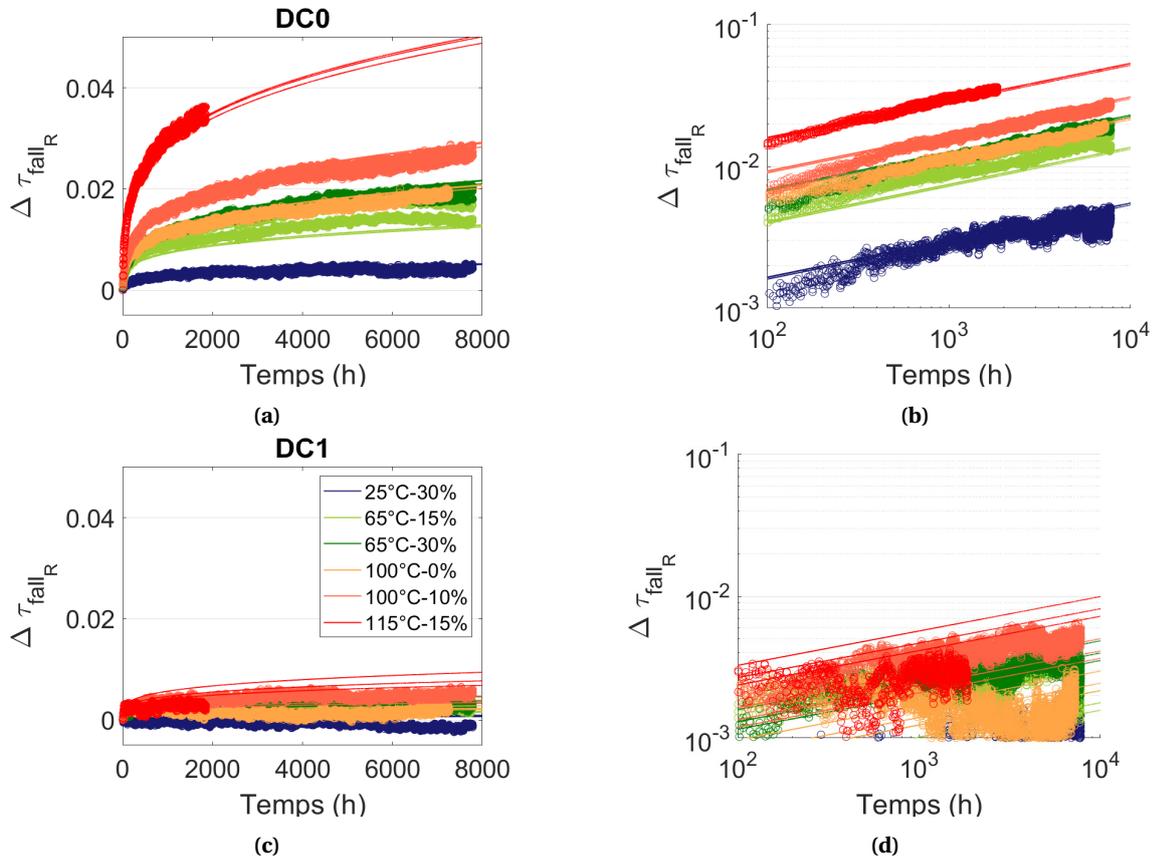


FIGURE 4.9 – Mesure et modèle (4.4) de $\Delta\tau_{fall_R}$ dans les ressources de LUT en (a)(c) échelle linéaire et (b)(d) log-log pour différentes température et tension de stress et un signal de stress interne (a)(b) DC0 et (c)(d) DC1 et pour trois RO L1 B

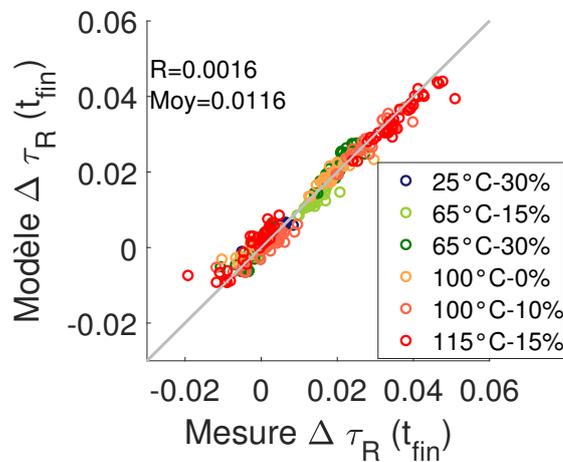


FIGURE 4.10 – Comparaison entre la prédiction à partir du modèle (4.4) et la mesure de l'effet du vieillissement, à t_{fin} , du temps de propagation d'un front montant et descendant pour un stress DC0 et DC1 dans les ressources de LUT

4.2.4 Ressources de routage

À présent, en utilisant les résultats issus de la méthode d'extraction des dégradations dans les ressources de routage présentée dans la section 3.5, nous modélisons les dégradations dans les ressources de routage.

Observations et interprétations

Comme pour les ressources de LUT, la figure 4.11 révèle l'augmentation du temps de propagation (dégradation) d'un front montant et descendant respectivement pour un stress DC1 et DC0 et la diminution du temps de propagation (amélioration) d'un front montant et descendant respectivement pour un stress DC0 et DC1. Cependant, en comparant avec l'effet du vieillissement sur les ressources de LUT présentées sur la figure 4.4 on constate que :

- L'amplitude des dégradations est moins importante : l'amplitude maximum des améliorations dans les ressources de LUT sont de $\approx 3.8\%$ et dans les ressources de routage elles sont de $\approx 1.8\%$. Dans les ressources de routage, la majorité du temps de propagation est induit par les fils de connections (voir section 1.3.3). L'observation des dégradations relatives plus faibles dans les ressources de routages par rapport aux ressources de LUT signifie que la dégradation des fils d'interconnexions est moins importante que les transistors dans les LUTs.
- L'amplitude des améliorations est plus importante : l'amplitude maximum des dégradations dans les ressources de LUT sont de $\approx 0.1\%$ et dans les ressources de routage elles sont de $\approx 1.5\%$. En considérant que les effets d'améliorations du temps de propagation sont présents dans les inverseurs (voir figure 4.3), cette dernière remarque laisse à penser que, dans les ressources de routages, le rapport entre le nombre d'inverseurs et de portes de transferts est plus important que dans les ressources de LUT.

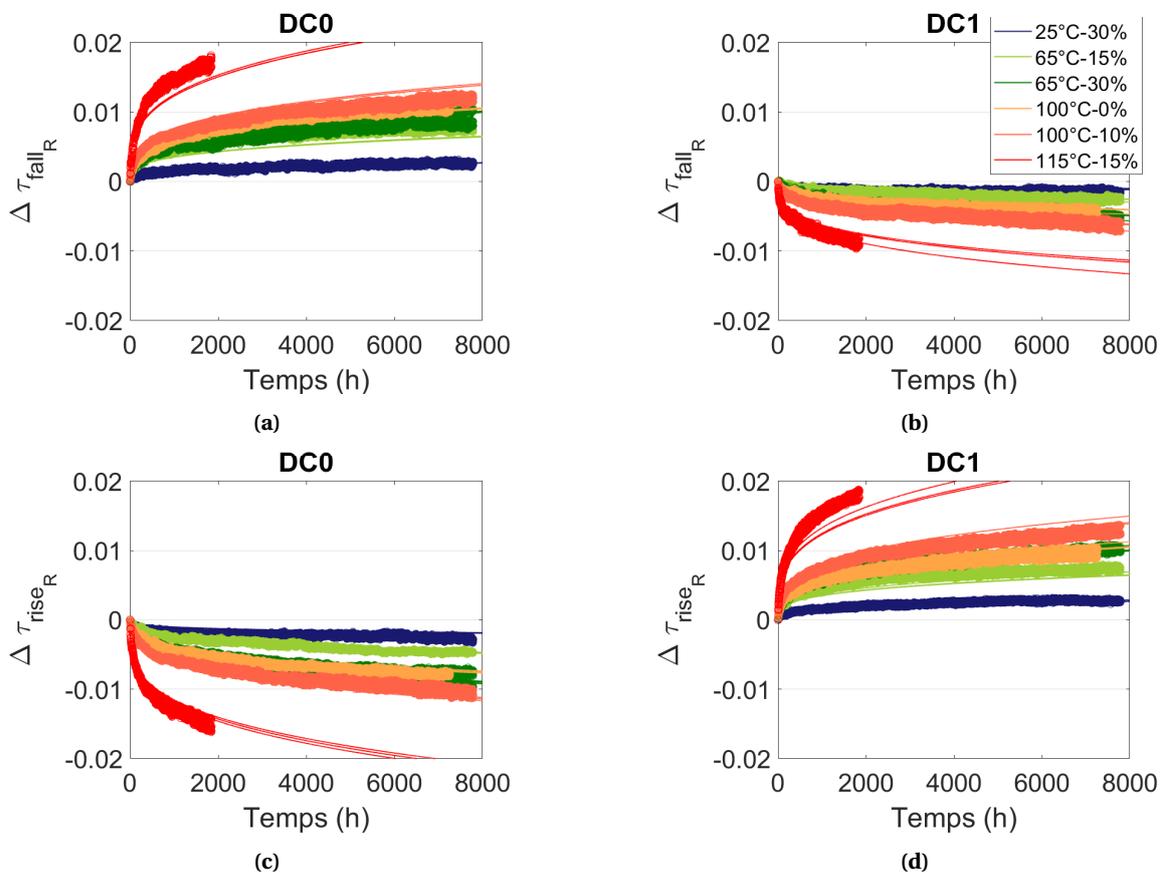


FIGURE 4.11 – Évolution de (a)(b) $\Delta\tau_{fall_R}$ (c)(d) $\Delta\tau_{rise_R}$ dans les ressources de routage pour différentes températures et tensions de stress et un signal de stress interne (a)(c) DC0 et (b)(d) DC1 et pour 3 RO L1 B

Modélisation

Concernant la modélisation des dégradations des ressources de routage, nous avons procédé comme pour les ressources de LUT :

- Modèle 1 : Loi puissance
- Modèle 2 : Modélisation de l'effet de la température et de la tension

Comme pour la modélisation des dégradations dans les ressources de LUT, les coefficients b , γ et Ea du modèle (4.2) sont communs à tous les ROs et seul le coefficient d'amplitude A est différent. La figure 4.11 compare les mesures avec ce dernier modèle. Sur la figure 4.12a nous avons tracé la valeur du coefficient d'amplitude A de chaque RO en fonction du temps de propagation du front dans ce même RO à t_0 et aux conditions nominales. Alors que l'amplitude des dégradations est normalement relative au temps de propagation, la figure 4.12a manifeste une dépendance, à priori linéaire, du coefficient A au temps de propagation.

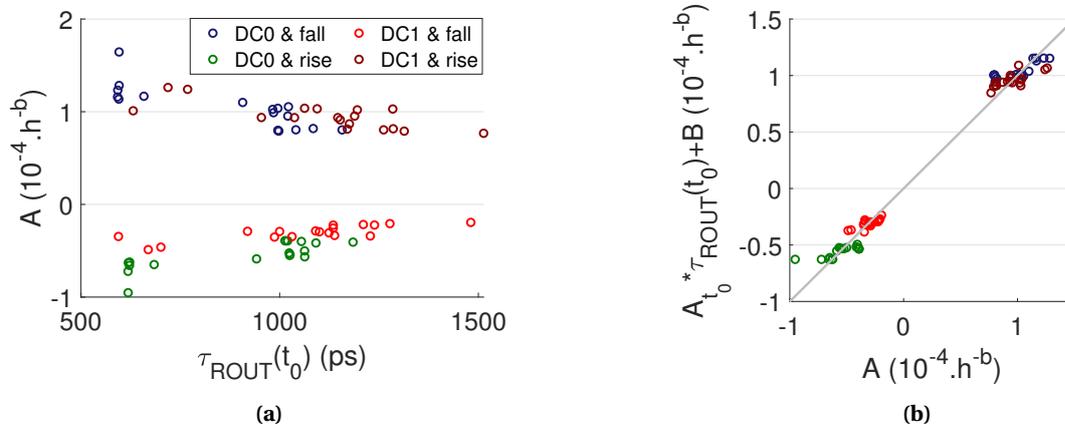


FIGURE 4.12 – (a) Coefficient d'amplitude A en fonction du temps de propagation à t_0 et dans les ressources de routages et pour les conditions nominales ($T = 25^\circ\text{C}$ et $V = V_{nom}$) (b) Qualité du modèle (4.5) pour prédire l'amplitude du coefficient A à partir du temps de propagation initial dans les ressources de routages

Voici notre interprétation de la dépendance du coefficient d'amplitude A au temps de propagation initial. Une ressource de routage est composée de fils et de PIPs (Programmable Interconnect Point) qui sont très certainement des portes de transfert et des inverseurs. Un chemin court (resp. long) correspond à une ressource de routage connectant deux ressources logiques proches (resp. éloignées). Dans un chemin court, le temps de propagation dans le fil de connexion est faible puisque la connexion est courte, cependant, pour assurer l'interconnexion la ressource de routage possède un minimum de portes de transferts et d'inverseurs. Dans un chemin long, le temps de propagation dans le fil de connexion est important puisque la connexion est longue, cependant, le nombre de portes de transferts et d'inverseurs n'augmente pas en proportion. En d'autres termes, le nombre de portes de transferts n'est pas proportionnel à la longueur de l'interconnexion. Parce que la dégradation dans les fils de connexion est moins importante que dans les ressources logiques, la dégradation relative dans les chemins courts de routage est plus importante que dans les chemins longs.

Nous avons modélisé cette dépendance avec un polynôme du premier degré tel que :

$$A = A_{t_0} \times \tau(t_0) + B \quad (4.5)$$

La figure 4.12b présente la qualité de cette modélisation. Ainsi, nous avons à présent quatre modèles de dégradation communs à tous les RO pour prédire la dégradation du temps de propagation relatif d'un front F (montant ou descendant) en fonction du stress statique S (DC1 ou DC0) dans

les ressources (Res) de routages tel que :

$$\Delta\tau_{R_{FS,Res}} = a_{FS,Res} t^{b_{FS,Res}} = (A_{t0_{FS,Res}} \tau(t_0) + B_{FS,Res}) \times e^{Y_{FS,Res}(V_{stress}-V_{op})} e^{\frac{Ea_{FS,Res}}{k_B} \left(\frac{1}{T_{op}} - \frac{1}{T_{stress}} \right)} t^{b_{FS,Res}} \quad (4.6)$$

La table 4.2 rapporte les valeurs numériques des coefficients du modèle.

TABLE 4.2 – Coefficients du modèles (4.6) pour prédire l’effet du vieillissement pour un stress statique sur les LUT

Effet vieillissement	F,S,Res	Coefficients modèle (4.6)				
		$A_{t_0}(10^{-9}.ps^{-1}h^{-b})$	$B(10^{-5}.h^{-b})$	$\gamma(V^{-1})$	$Ea(eV)$	b
Dégradation	fall,DC0,ROUT	-16.6	12.4	5.17	0.29	0.29
	rise,DC1,ROUT	-12.5	11.4			
Amélioration	fall,DC1,ROUT	7.8	-4.5	3.61	0.34	0.29
	rise,DC0,ROUT	10.9	-7.1			

Validation du modèle

La figure 4.13 présente la comparaison entre la mesure et le modèle de prédiction. On constate une surestimation du modèle liée à une mauvaise modélisation des facteurs d’accélération de

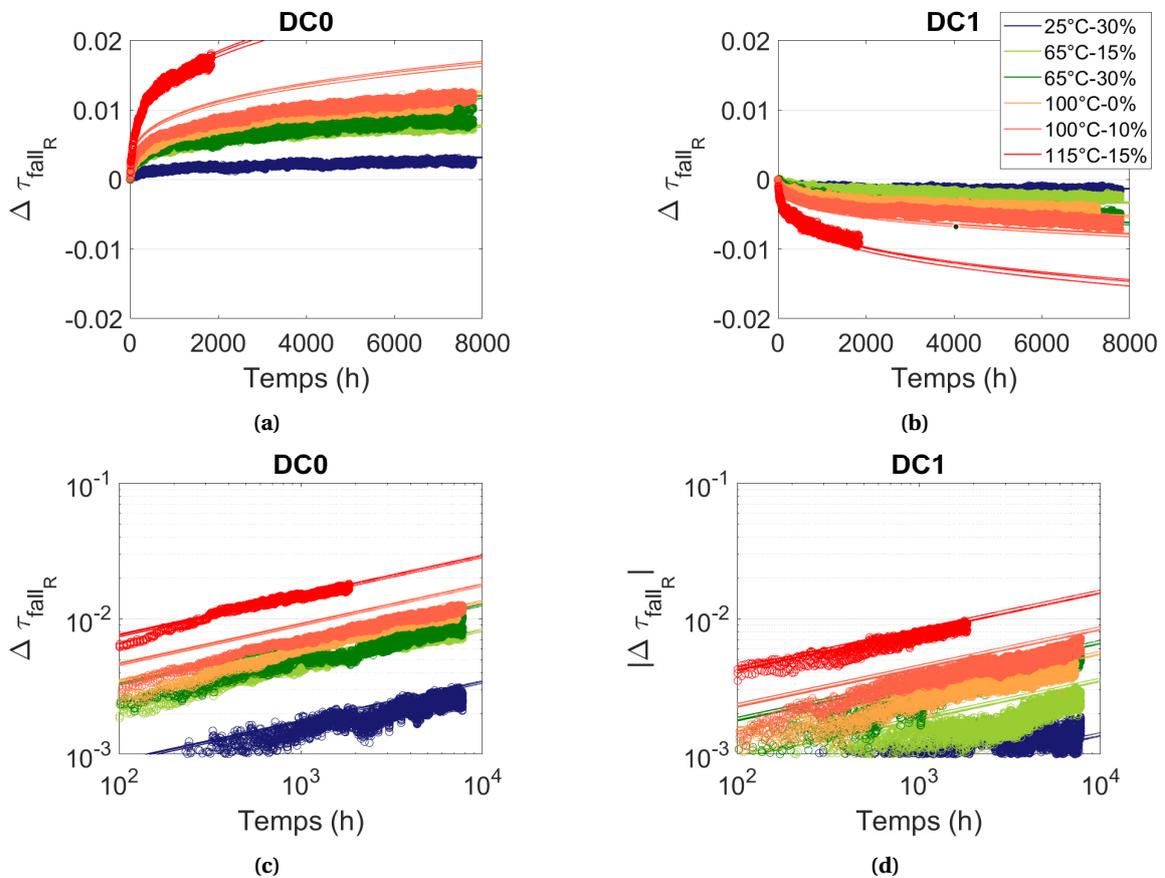


FIGURE 4.13 – Mesure et prédiction à partir du modèle (4.6) de (a)(b) $\Delta\tau_{fall_R}$ (c)(d) $\Delta\tau_{rise_R}$ dans les ressources de routages pour différentes température et tension de stress et un signal de stress interne (a)(c) DC0 et (b)(d) DC1 et pour trois RO L1 B

température et de tension. Cependant, comme le montre la figure 4.14, avec le modèle (4.6) nous pouvons prédire les dégradations finales, c'est à dire au dernier instant de mesure pour chaque FPGA, et ce avec une erreur relative moyenne de 10%.

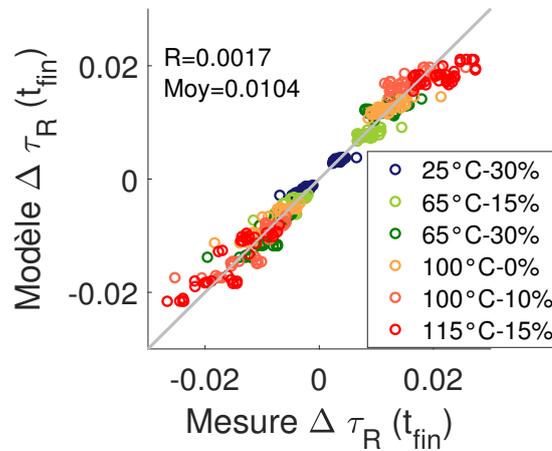


FIGURE 4.14 – Comparaison entre la prédiction à partir du modèle (4.6) et la mesure de l’effet du vieillissement, à t_{fin} , du temps de propagation d’un front montant et descendant pour un stress DC0 et DC1 dans les ressources de routages

Il est certain que notre modèle n’est pas parfait, mais il reste cohérent et surtout il constitue un modèle unique qui permet de prédire les dégradations dans les ressources de routages de toutes les architectures de ROs.

4.2.5 Conclusion

Dans cette section nous avons modélisé l’effet du vieillissement sur le temps de propagation d’un front montant et descendant dans les ressources de LUT d’une part, et de routages d’autre part, soumises à un stress statique. En fonction du niveau de stress et du niveau du front mesuré, deux effets du vieillissement sont possibles : dégradation (augmentation) du temps de propagation ou amélioration (diminution) du temps de propagation. Ainsi, pour les ressources de LUT et de routages, nous avons développé un modèle permettant de prédire :

- La dégradation avec un stress DC0;
- La dégradation avec un stress DC1;
- L’amélioration avec un stress DC0;
- L’amélioration avec un stress DC1;

Le modèle (4.4) est utilisé pour les ressources de LUT et les valeurs numériques sont précisées dans la table 4.1. Le modèle (4.6) est utilisé pour les ressources de routages et les valeurs numériques sont incluses dans la table 4.2. Le modèle montre une capacité à prédire l’effet du vieillissement à t_{fin} , c’est à dire au dernier instant de mesure, et ce avec une erreur de 10% environ pour les ressources de LUT et de routages.

4.3 Modélisation des dégradations pour un stress dynamique : rapport cyclique

Après avoir modélisé l’effet du vieillissement sur le temps de propagation d’un front pour un stress statique, nous modélisons à présent l’effet du rapport cyclique de stress (α_{stress}). Dans un premier temps nous modélisons les dégradations dynamique comme étant la somme des dégradations statique pour le temps passé en DC0 et le temps passé en DC1. Ce premier modèle montrera clairement l’effet du mécanisme de récupération du BTI. Ainsi, nous proposons de modéliser

le mécanisme de récupération afin d'avoir un modèle complet prévoyant les dégradations pour un stress dynamique. Nous utiliserons le même méthode de modélisation pour les dégradations dans les ressources de LUT et de routages contrairement à la modélisation pour un stress statique présentée dans la précédente section.

4.3.1 Effet du rapport cyclique de stress sur les dégradations

La figure 4.15 présente la dégradation relative, par rapport à la dégradation pour un stress DC0 ou DC1, d'un front montant ou descendant dans les ressources de LUT et de routages à t_{fin} , c'est à dire au dernier instant de mesure. Chaque point correspond à la moyenne des dégradations finales de tous les ROs et les barres à l'écart type. Aussi bien pour les ressources de LUT que de routages, on observe une dépendance importante de l'amplitude des dégradations au rapport cyclique de stress. Pour le routage par exemple, avec $\alpha_{stress} = 0.1$ la dégradation relative de $\Delta\tau_{fall_R}$ est de 0.5 puis elle est de -0.03 lorsque $\alpha_{stress} = 0.9$.

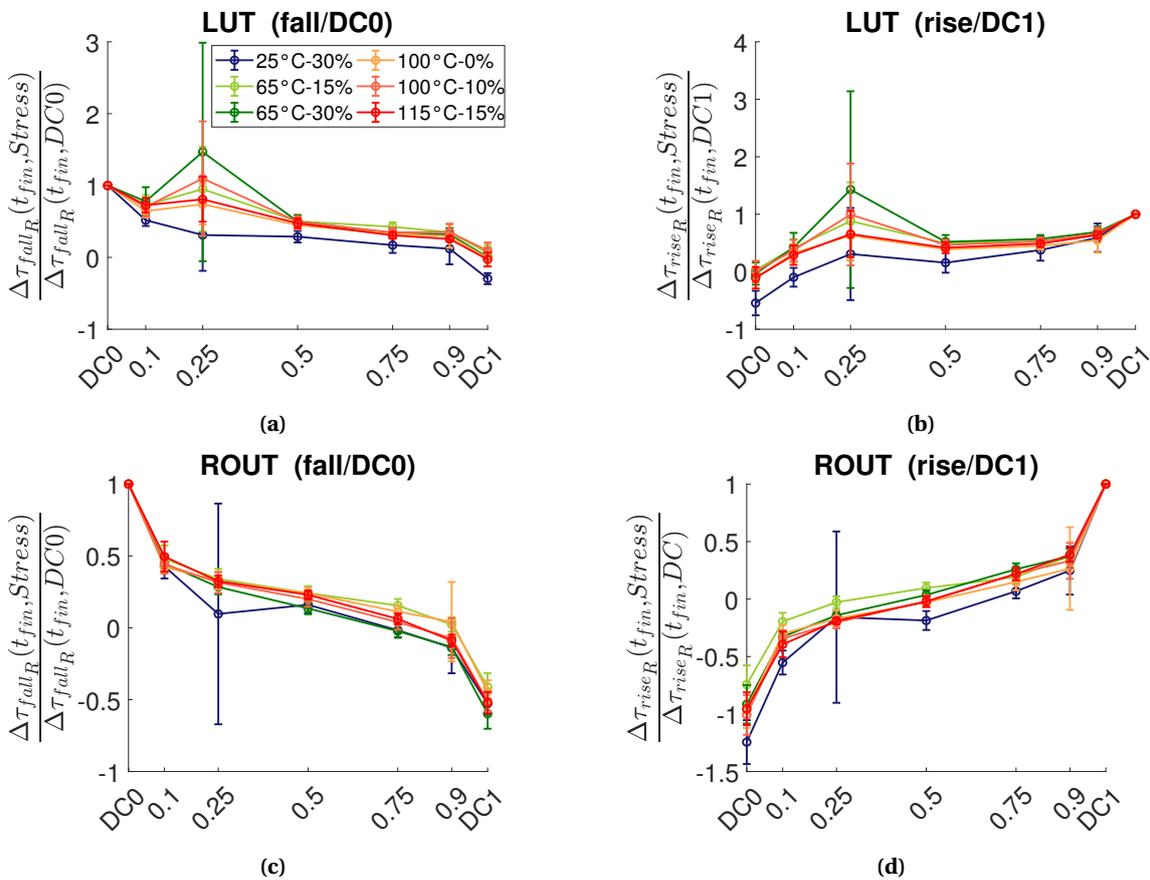


FIGURE 4.15 – Amplitude relative (par rapport au stress statique) des dégradations dans les ressources de (a)(b) LUT et de (c)(d) routage au dernier instant de mesure en fonction du rapport cyclique du stress interne et pour différentes conditions de stress de température et de tension (a)(c) τ_{fall_R} (b)(d) τ_{rise_R}

4.3.2 Modèle dynamique à partir des modèles statiques

Aussi bien dans les ressources de LUT que de routages, la figure 4.15 manifeste la diminution de l'amplitude des dégradations de $\Delta\tau_{fall_R}$ lorsque le rapport cyclique de stress augmente et inversement, l'augmentation de l'amplitude des dégradations de $\Delta\tau_{rise_R}$ lorsque le rapport cyclique de stress augmente. En effet, lorsque $\alpha_{stress} = 0.1$, alors le signal de stress vieillit 90% du temps les transistors passant lors d'un niveau bas et vieillit 10% du temps les transistors passant lors d'un niveau haut. Ainsi, il est normal d'observer des dégradations plus importantes de $\Delta\tau_{fall_R}$ pour un

signal de stress avec $\alpha_{stress} = 0.1$ que lorsque $\alpha_{stress} = 0.9$. Réciproquement, il est normal d'observer des dégradations plus importantes de $\Delta\tau_{rise_R}$ pour un signal de stress avec $\alpha_{stress} = 0.9$ que lorsque $\alpha_{stress} = 0.1$.

Afin d'élargir le modèle statique précédemment développé à des stress dynamiques, dans un premier temps nous essayons de modéliser le vieillissement directement à partir du modèle statique. Rappelons que la forme générale du modèle pour prédire la dégradation d'un front (F) montant ou descendant, pour un stress statique (S) DC0 ou DC1 aussi bien dans les ressources (Res) de LUT (équation (4.4)) que de routages (équation (4.6)) est :

$$\Delta\tau_{R_{FS},Res} = a_{FS,Res} t^{b_{FS,Res}} \quad (4.7)$$

Nous considérons que la dégradation avec un stress dynamique correspond à la somme des dégradations obtenues avec les stress statiques DC0 et DC1 en fonction du temps passé par le signal de stress au niveau haut et bas, donc, en fonction du rapport cyclique du signal de stress dynamique. Avec ce premier modèle dynamique nous exprimons $\Delta\tau_{R_F}$ tel que :

$$\Delta\tau_{R_{FRes}}(t, \alpha_{stress}) = a_{FDC1,Res} (\alpha_{stress} t)^{b_{FDC1,Res}} + a_{FDC0,Res} ((1 - \alpha_{stress}) t)^{b_{FDC0,Res}} \quad (4.8)$$

Le terme $((1 - \alpha_{stress}) \times t)$ correspond au temps de vieillissement avec un stress DC0 et $(\alpha_{stress} \times t)$ correspond au temps de vieillissement pour des dégradations avec un stress DC1. Si $\alpha_{stress} = 0$ alors la partie gauche de l'équation (4.8) est nulle, donc $\Delta\tau_{R_{FRes}}$ dépend uniquement du modèle statique DC0. Au contraire, si $\alpha_{stress} = 1$, alors $\Delta\tau_{R_{FRes}}$ dépend uniquement du modèle statique DC1.

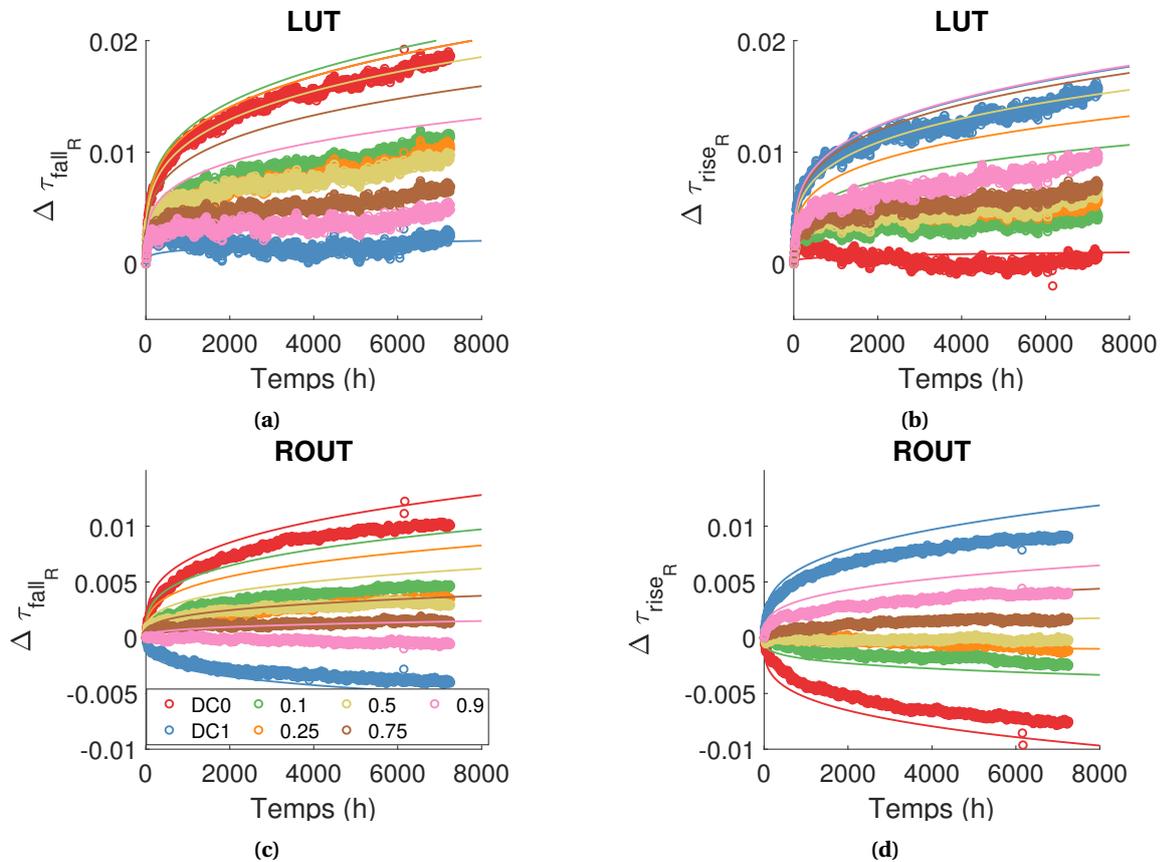


FIGURE 4.16 – Comparaison de la mesure et du modèle (4.8) pour prédire l'effet du vieillissement sur le temps de propagation d'un front descendant (a)(c) et montant (b)(d) dans les ressources de (a)(b) LUT et de (c)(d) routages pour différents rapports cycliques de stress pour un RO L1 B à $T_{stress} = 100^\circ\text{C}$ et $V = V_{nom}$

La figure 4.16 compare le modèle défini dans l'équation (4.8) avec les mesures. Pour les stress statiques DC0 et DC1, c'est directement le modèle statique défini dans la section précédente qui prédit l'effet du vieillissement. Pour les stress dynamiques on constate que le modèle sur-estime les dégradations. La différence entre les mesures et le modèle est probablement liée à l'effet de récupération du BTI que nous n'avons pas intégré dans ce premier modèle. En effet, comme nous l'avons détaillé dans la section 1.5.2, un transistor est affecté par le BTI lorsqu'il est passant, alors sa tension de seuil se dégrade. Lorsque le transistor n'est plus passant, sa tension de seuil revient vers sa tension de seuil initiale, c'est ce qu'on appelle le mécanisme de récupération. Dans une LUT, en fonction du niveau logique à transmettre, certains transistors sont passant et d'autres sont bloqués, donc, certains transistors sont dégradés et d'autres récupèrent. Dans le cas d'un stress dynamique, la dégradation que nous mesurons est en réalité la somme de deux phénomènes : la dégradation et la récupération lié au BTI.

4.3.3 Modélisation du mécanisme de récupération

À présent nous allons modéliser l'effet de récupération pour prédire correctement les dégradations avec un stress dynamique. L'effet de récupération du BTI pour des technologies high- κ et tri-gate est étudié dans [160] et [161]. Dans ces études les auteurs proposent un modèle empirique où l'effet de récupération est modélisé comme une fraction de récupération (RF) :

$$RF = \frac{1}{1 + B \left(\frac{t_{recovery}}{t_{stress}} \right)^\beta} \quad (4.9)$$

où B and β sont des paramètres d'ajustement. Dans notre cas, pour un stress de niveau '0' : $t_{recovery} = t_{high}$ et $t_{stress} = t_{low}$ alors que, pour un stress de niveau '1' : $t_{recovery} = t_{low}$ et $t_{stress} = t_{high}$. En considérant que $\alpha_{stress} = \frac{t_{high}}{t_{high} + t_{low}}$ nous pouvons écrire :

$$RF_0 = \frac{1}{1 + B_0 \left(\frac{\alpha_{stress}}{1 - \alpha_{stress}} \right)^{\beta_0}} \quad (4.10)$$

$$RF_1 = \frac{1}{1 + B_1 \left(\frac{1 - \alpha_{stress}}{\alpha_{stress}} \right)^{\beta_1}} \quad (4.11)$$

Les coefficients RF_0 et RF_1 représentent une fraction de récupération $\left(\frac{\text{dégradation avec recuperation}}{\text{dégradation sans recuperation}} \right)$ respectivement dans le cas d'une dégradation avec un stress de niveau 0 et de niveau 1. Afin de mieux comprendre le comportement de RF, nous avons tracé sur la figure 4.17 l'évolution de RF_0

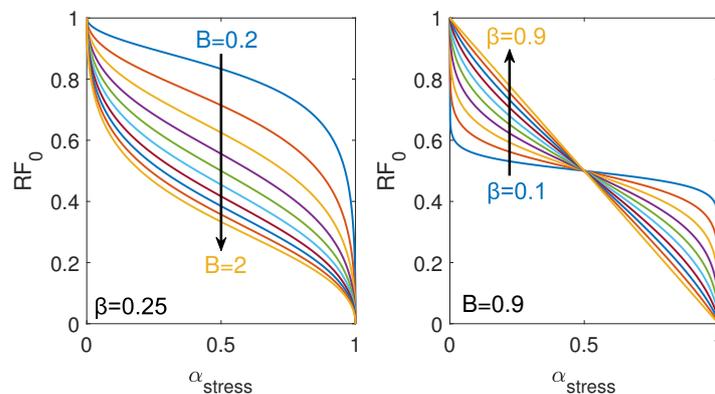


FIGURE 4.17 – Évolution de la fraction de récupération (RF_0) en fonction du rapport cyclique de stress en faisant varier (gauche) avec un pas de 0.1 le coefficient B et avec $\beta = 0.25$ (droite) avec un pas de 0.05 le coefficient β et avec $B = 0.9$

en fonction du rapport cyclique du signal de stress et en faisant évoluer le coefficient B puis β . Le coefficient RF_0 est compris entre 0 et 1. Si $\alpha_{stress} = 0$ alors $RF_0 = 1$ puisque cela veut dire qu'il n'y a pas de récupération (stress statique par exemple). Si $\alpha_{stress} = 1$ alors $RF = 0$, cela veut dire qu'il y a 100% de récupération (composant non stressé). On a le comportement inverse avec RF_1 . Lorsque $\alpha_{stress} = 0.5$, RF est indépendant du coefficient β et son amplitude dépend de B.

Finalement, en intégrant les équations (4.10) et (4.11) dans l'équation (4.8) on obtient le modèle dynamique incluant l'effet de récupération :

$$\Delta\tau_{R,Res}(t, \alpha_{stress}) = \left(a_{F,DC1,Res} (\alpha_{stress} t)^{b_{F,DC1,Res}} \right) RF_1 + \left(a_{F,DC0,Res} ((1 - \alpha_{stress}) t)^{b_{F,DC0,Res}} \right) RF_0 \quad (4.12)$$

Nous avons appliqué une régression non linéaire entre le modèle (4.12) et les mesures de dégradation pour différentes fréquences et rapport cyclique de stress. Sur la figure 4.18 nous observons une bonne cohérence entre le modèle précédemment défini et les mesures de dégradation avec des stress dynamiques et statiques. Cette figure montre que, quelque soit le stress dynamique ou statique et quel que soit le rapport cyclique du stress, nous pouvons prédire les dégradations avec un même modèle.

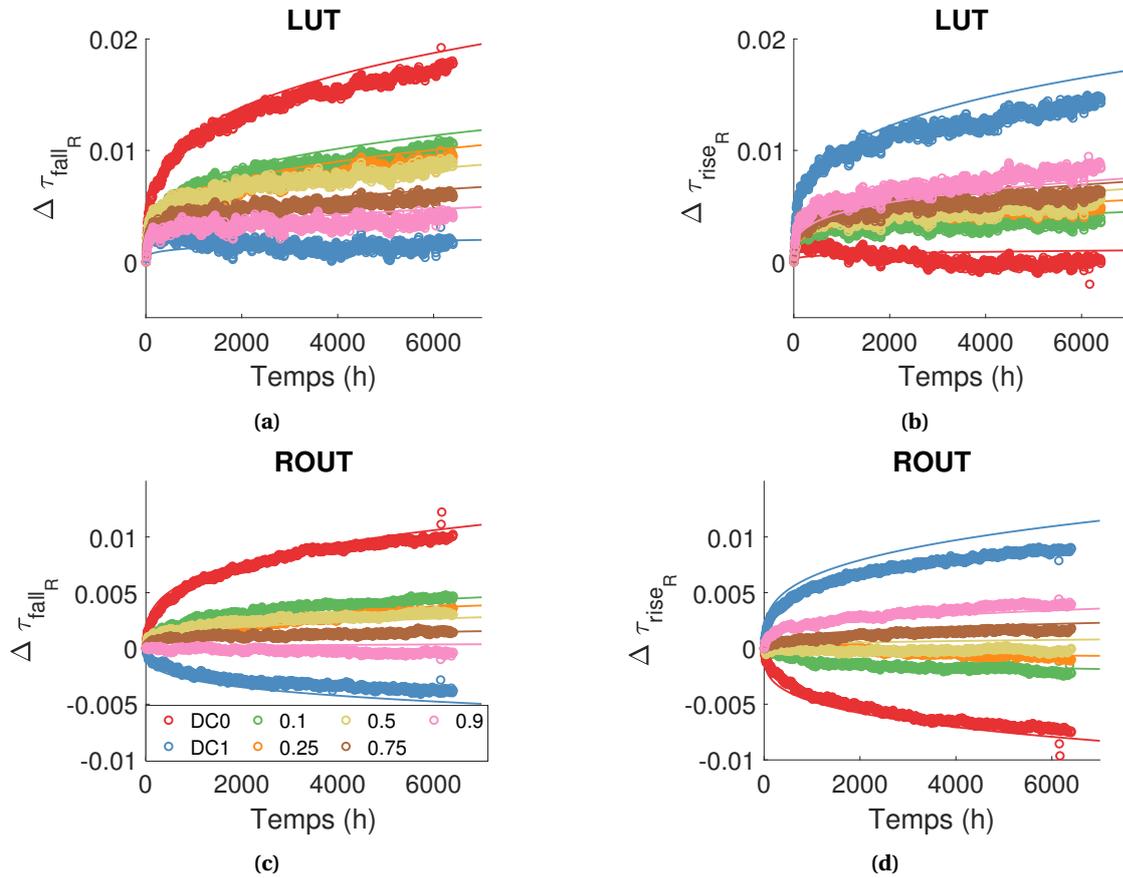


FIGURE 4.18 – Comparaison de la mesure et du modèle (4.12) pour prédire l'effet du vieillissement sur le temps de propagation d'un front descendant (a)(c) et montant (b)(d) dans les ressources de (a)(b) LUT et de (c)(d) routages pour différents rapports cycliques de stress pour un RO L1 B à $T_{stress} = 100^\circ C$ et $V = V_{nom}$

À ce niveau de la modélisation, nous avons un modèle par architecture de RO et par condition de stress de température et de tension. L'objectif est d'avoir un modèle commun à tous les RO dans toutes les conditions de stress. C'est pourquoi, pour toutes les architectures de RO, nous avons pris des coefficients B_0 , B_1 , β_0 et β_1 identiques correspondant à la moyenne des coefficients précédemment obtenus pour chacune des architectures de RO. La figure 4.19 confirme une dépendance des coefficients B et β à la tension et à la température. Dans [161] l'auteur montre que le coefficient

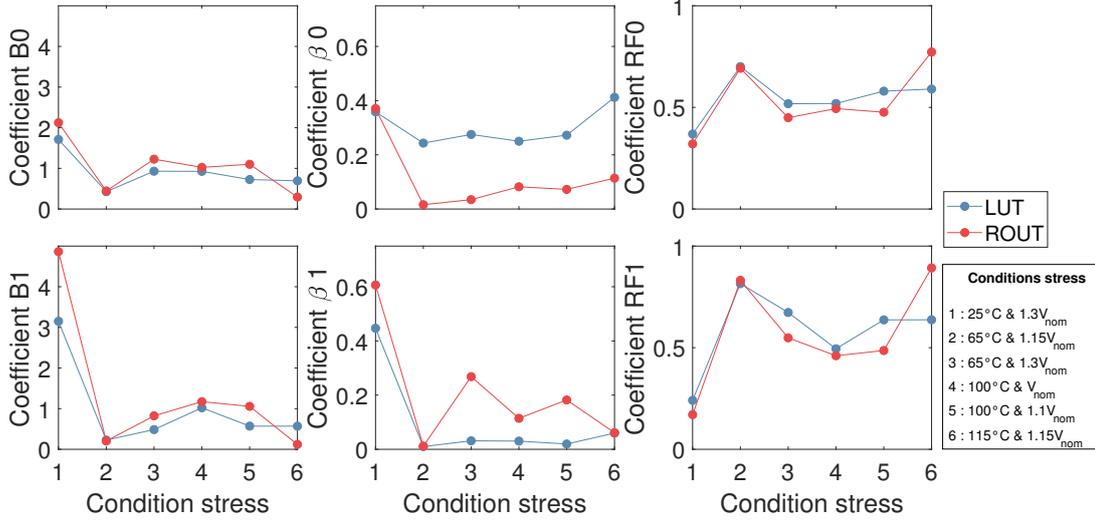


FIGURE 4.19 – Paramètres moyens pour toutes les architectures de RO définis dans les équations (4.10) et (4.11) en fonction du stress de température et de tension. RF_0 et RF_1 sont obtenus avec $\alpha_{stress} = 0.5$

B dépend de la tension et la fraction de récupération dépend de la température. L’auteur utilise un modèle empirique pour la dépendance de B à la tension et une loi d’Arrhenius pour la température. Contrairement, à l’étude réalisée dans [161] qui mesure les dégradations directement sur le composant, dans notre cas nous mesurons des dégradations sur un circuit compliquant la modélisation semi-empirique de la fraction de récupération (RF) avec une loi d’Arrhenius. Ainsi, nous modélisons empiriquement la dépendance des coefficients B et β avec un polynôme à deux variables, température et tension, du second ordre pour les deux variables :

$$\beta(T, V) = p_{00} + p_{10}(T - T_{nom}) + p_{00}(V - V_{nom}) + p_{11}(T - T_{nom})(V - V_{nom}) + p_{20}(T - T_{nom})^2 + p_{02}(V - V_{nom})^2 \quad (4.13)$$

$$B(T, V) = q_{00} + q_{10}(T - T_{nom}) + q_{00}(V - V_{nom}) + q_{11}(T - T_{nom})(V - V_{nom}) + q_{20}(T - T_{nom})^2 + q_{02}(V - V_{nom})^2 \quad (4.14)$$

La table 4.3 contient les paramètres, strictement empiriques, des modèles de β_0 , β_1 , B_0 et B_1 identifiés.

Finalement, en utilisant le modèle défini dans l’équation (4.12) et avec les coefficients de fraction de récupération (RF) obtenus en utilisant les modèles (4.14) et (4.13), nous avons calculé la dégradation au dernier instant de mesure (t_{fin}). La figure 4.20 compare la mesure et la prédiction de la dégradation du temps de propagation d’un front dans des ressources de LUT et de routages. La moyenne des dégradations relatives pour les ressources de LUT étant de 0.02 et pour les ressources de routages de 0.01, notre modèle prédit les dégradations avec un résidu moyen de 10%.

À partir du modèle (4.12) valable aussi bien dans les ressources de routages que de LUT, nous reconstruisons le modèle de dégradation dans le RO qui est le circuit directement mesuré par le banc de test :

$$\Delta\tau_{RERO}(t, \alpha_{stress}) = \Delta\tau_{RELUT}(t, \alpha_{stress}) + \Delta\tau_{REROUT}(t, \alpha_{stress}) \quad (4.15)$$

Avec un résidu relatif moyen d’environ 10%, la figure 4.21 confirme la capacité de notre modèle à prédire les dégradations dans les ROs à partir des modèles définis pour les dégradations dans les ressources de routages et de LUT.

TABLE 4.3 – Coefficients pour les modèles de β_0 , β_1 , B_0 et B_1

	β_0					
	$p00(\times 10^{-2})$	$p10(\times 10^{-3}K^{-1})$	$p01(V^{-1})$	$p11(\times 10^{-2}K^{-1}V^{-1})$	$p20(\times 10^{-6}K^{-2})$	$p02(V^{-2})$
LUT	76.1	-14.9	-2.6	3.3	108.0	3.9
Routage	-12.7	-2.1	2.8	-3.5	65.8	-3.3
	β_1					
	LUT	1.9	-6.2	2.3	-3.0	84.3
Routage	-229.6	43.3	15.9	-18.1	-148.8	-17.3
	B_0					
	LUT	-132.3	29.8	12.4	-19.6	2.7
Routage	-915.0	197.0	62.1	-73.8	-819.6	-69.5
	B_1					
	LUT	-97.8	-1.0	19.4	-31.8	370.1
Routage	-1143.6	204.3	92.8	-112.9	-484.0	-112.2

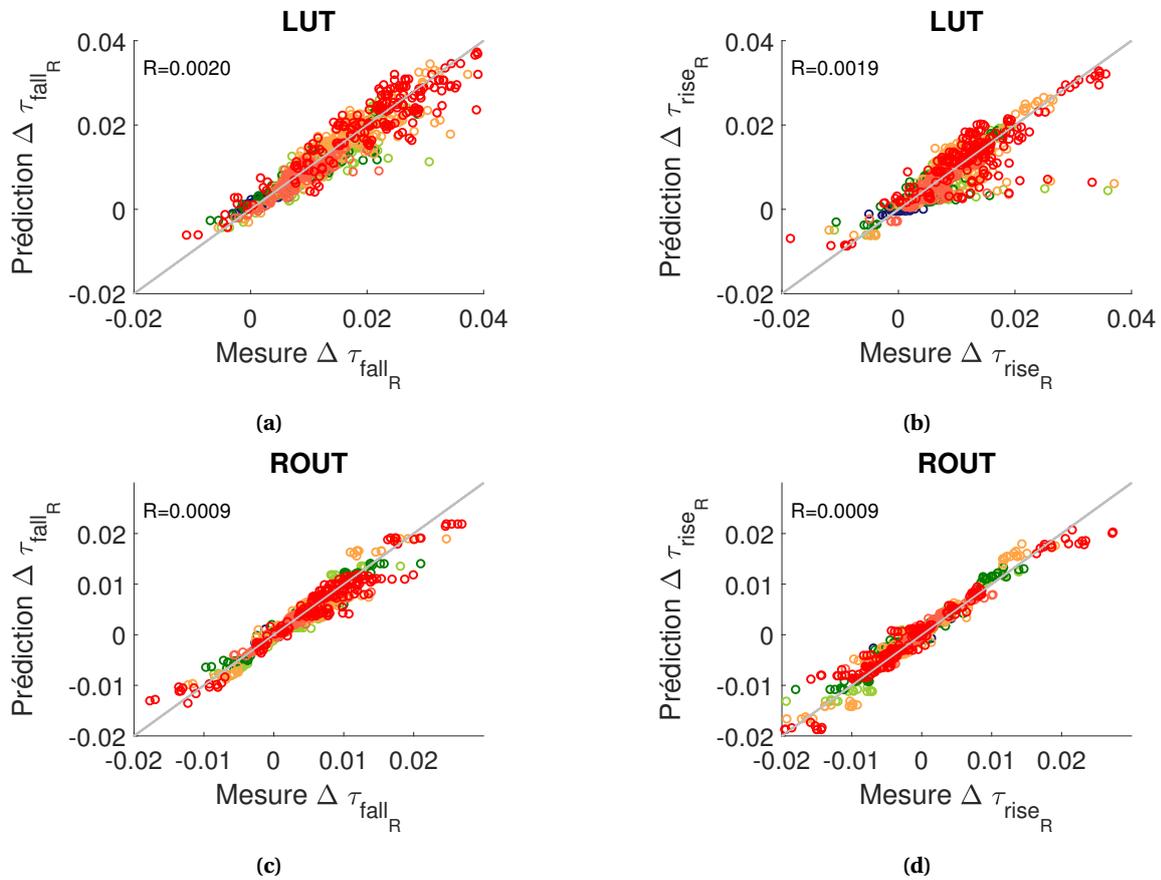


FIGURE 4.20 – Comparaison entre la mesure et la prédiction avec le modèle (4.12) des dégradations du temps de propagation, à t_{fin} , d'un front (a)(c) descendant et (b)(d) montant dans les ressources de (a)(b) LUT et de (c)(d) routages

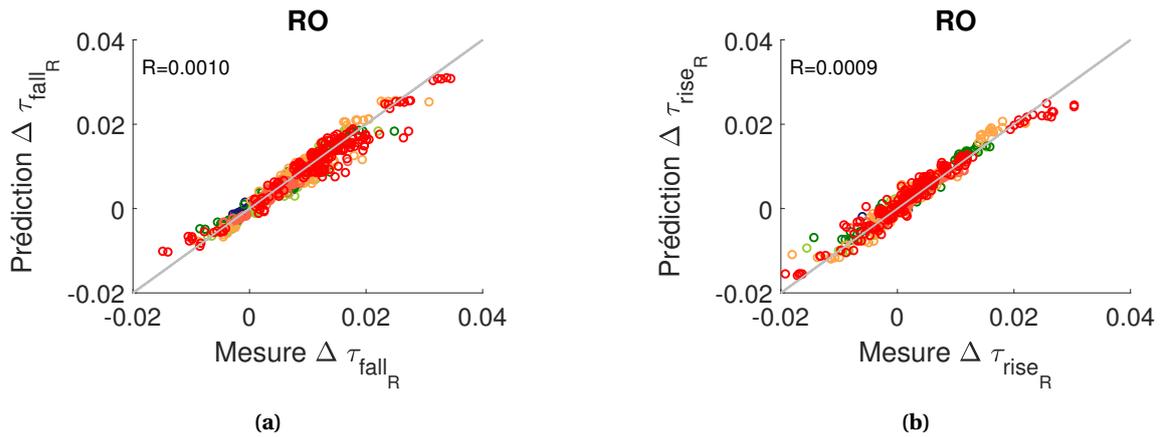


FIGURE 4.21 – Comparaison entre la mesure et la prédiction avec le modèle (4.15) des dégradations du temps de propagation, à t_{fin} , d'un front (a) descendant et (b) montant dans les ROs

4.3.4 Conclusion

Dans un premier temps, nous avons reconstruit les dégradations pour un stress dynamique directement à partir du modèle statique. Nous avons observé une sur-estimation du modèle, certainement à cause du mécanisme de récupération du BTI jusqu'ici non pris en compte dans notre modèle. Ainsi, nous avons modélisé le mécanisme de récupération en fonction du rapport cyclique, de la température et de la tension de stress. Ensuite, nous obtenons un modèle commun à toutes les architectures de ROs qui modélise l'effet du vieillissement sur le temps de propagation dans les ressources de LUT et de routages soumises à un stress statique ou dynamique avec différents rapports cycliques de stress. Finalement, nous avons sommé les modèles du routage et des LUT pour obtenir le modèle du RO. Nous obtenons une cohérence entre la prédiction du modèle et la mesure avec un résidu relatif moyen de 10% ce qui est satisfaisant.

4.4 Modélisation des dégradations pour un stress dynamique : fréquence

Dans la précédente section, nous avons modélisé l'effet du rapport cyclique de stress sur les dégradations. Dans cette section nous étudions l'effet de la fréquence de stress sur le vieillissement du temps de propagation d'un front montant et descendant dans les ressources de LUT et de routages. Tout d'abord, la figure 4.22 montre l'effet de la fréquence de stress sur la dégradation du temps de propagation mesuré dans un RO. Contrairement à la figure 4.18 où l'on voyait clairement l'effet du rapport cyclique de stress sur l'amplitude des dégradations, sur la figure 4.22 on constate que l'effet de la fréquence de stress est faible. Bien qu'au niveau du RO l'effet de la fréquence de stress sur les dégradations est négligeable, nous allons regarder si cela est toujours vrai dans les ressources de LUT et de routage en utilisant les données issues de la méthode d'extraction dans la section 3.5.

La figure 4.23 présente les dégradations des temps de propagations relatifs, par rapport à DC0 pour τ_{fall_R} et DC1 pour τ_{rise_R} , en fonction de la fréquence de stress et dans les ressources de LUT. Chaque point correspond à la moyenne de toutes les architectures et chaque barre son écart type.

Les figures 4.23a et 4.23b présentent l'effet du vieillissement sur les ressources de LUT. On observe, pour un stress DC1 (resp. DC0), une amélioration du temps de propagation d'un front descendant (resp. montant) comme observé dans la précédente section. Aussi bien pour τ_{fall_R} que pour τ_{rise_R} , on remarque que l'amplitude des dégradations est la plus importante à 100Hz pour les stress dynamiques; ce phénomène est sans doute lié à la dynamique de dégradation et de récupération pour le mécanisme de BTI (voir section 1.5.2). Cependant, nous n'avons pas de mesures de dégradation avec des stress de basses fréquence, mis à part 100Hz. Ainsi, nous ne pouvons pas

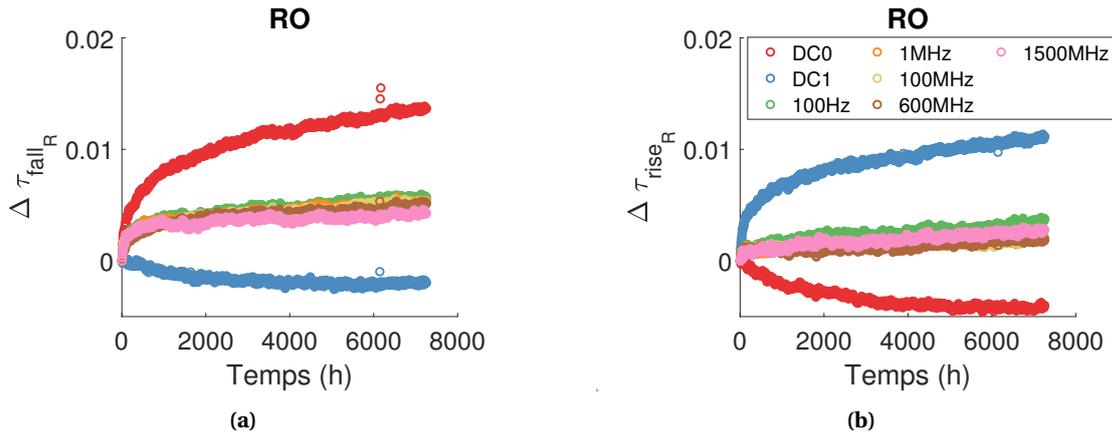


FIGURE 4.22 – Effet de la fréquence de stress sur la dégradation du temps de propagation dans un RO L1 B (a) τ_{fall_R} (b) τ_{rise_R} . $T_{stress} = 100^\circ\text{C}$, $V_{stress} = V_{nom}$ et $\alpha_{stress} = 0.5$.

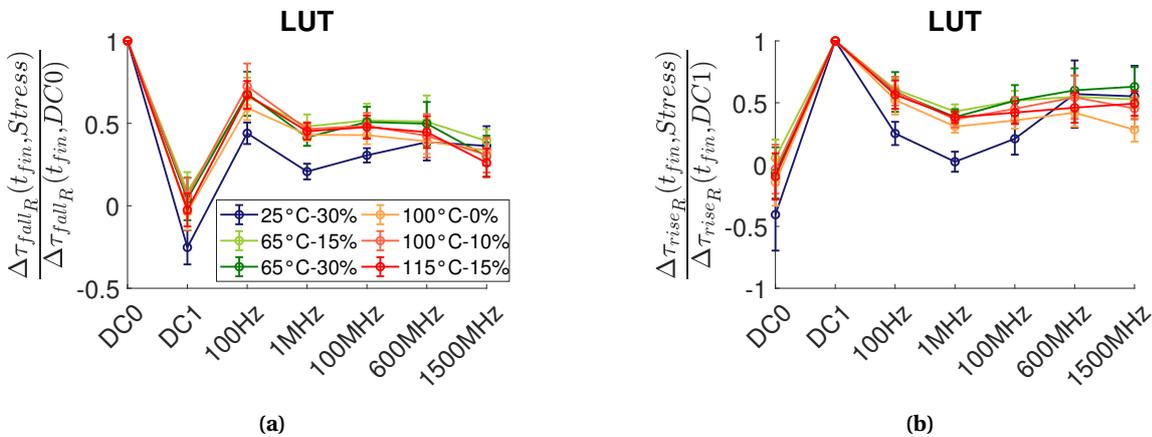


FIGURE 4.23 – Amplitude relative (par rapport au stress statique) des dégradations dans les ressources de LUT au dernier instant de mesure en fonction du stress interne et pour différentes conditions de stress de température et de tension (a) τ_{fall_R} (b) τ_{rise_R} . $\alpha_{stress} = 0.5$.

modéliser ce mécanisme qui correspond à la différence de dynamique entre les dégradations et les récupérations du BTI comme expliqué et modélisé dans [65, chapitre 21.3.3]. Pour les stress de hautes fréquences, nous constatons que l'amplitude est globalement stable dans les conditions de températures supérieures à 25°C .

Au contraire, pour $T_{stress} = 25^\circ\text{C}$ l'amplitude des dégradations augmente lorsque la fréquence de stress augmente. À priori, ce que nous observons à basse température peut être lié au HCI qui se produit à chaque fois que du courant passe dans le transistor. Donc, l'amplitude des dégradations doit augmenter avec la fréquence de stress sous l'effet du HCI.

La figure 4.24a présente les dégradations en fonction du temps pour plusieurs fréquences de stress et à $T_{stress} = 25^\circ\text{C}$. Nous avons modélisé les dégradations avec une loi puissance dont l'exposant au temps est de 0.5 comme défini dans le modèle de l'électron chanceux 1.5.3. La figure 4.24a confirme la cohérence entre les mesures et une loi puissance avec un exposant 0.5. Aussi, observe-t-on une amplitude croissante lorsque la fréquence de stress augmente. En réalité, même à basse température, les dégradations que l'on observe sur la figure 4.24a sont dues à l'effet du HCI et du BTI combinés. Dans [162, eq. 7], l'auteur propose de soustraire les dégradations obtenues aux hautes fréquences avec les dégradations obtenues à 25Hz. En effet, l'auteur considère que pour un stress de 25Hz le HCI est négligeable, ainsi, le BTI est prédominant. En considérant nos résultats, on constate que les dégradations à 100Hz sont plus importantes que les dégradations à hautes fréquences, ce qui rend cette méthode inapplicable. Cependant, le BTI est un mécanisme

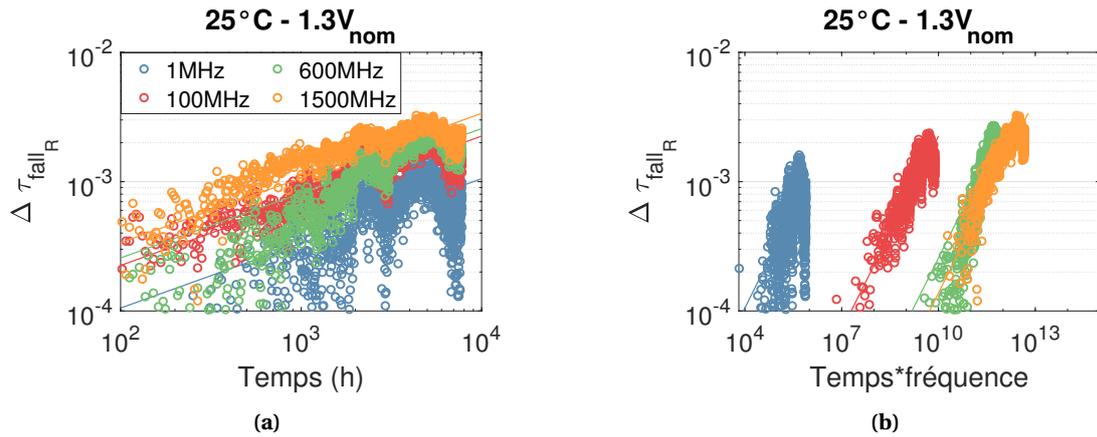


FIGURE 4.24 – Évolution de la dégradation de τ_{fallR} dans les ressources de LUT et pour un RO L1 XB pour différents stress dynamiques de hautes fréquences (a) en fonction du temps (b) en fonction du produit du temps et de la fréquence de stress. $\alpha_{stress} = 0.5$.

de dégradation indépendant de la fréquence de stress à haute fréquence, c'est d'ailleurs ce que l'on observe à haute température sur la figure 4.23. Par conséquent, on peut considérer que le BTI est identique pour un stress à 1MHz, 100MHz, 600MHz et 1500MHz, en considérant un rapport cyclique identique, et qu'il ne perturbe pas l'observation de l'effet du HCI dans ces conditions de stress.

Ainsi, pour aller plus loin dans l'étude du HCI, la figure 4.24b présente la dégradation de τ_{fallR} en fonction du produit du temps est de la fréquence de stress, c'est à dire en fonction du nombre de commutations. Comme le HCI dépend du nombre de commutations, on devrait obtenir sur la figure 4.24b un alignement des droites comme dans [162, fig. 11]. Hormis pour une fréquence de stress de 600MHz et 1500MHz, on n'observe pas d'alignement des dégradations, ce qui prouve qu'à 25°C les dégradations, sous l'effet du HCI, sont faibles voir négligeables devant le BTI.

Les figures 4.25a et 4.25b présentent l'effet du vieillissement sur les ressources de routages. On remarque que la dégradation de τ_{fallR} diminue légèrement et passe de 0.3 pour 100Hz à 0.1 pour 1500MHz. Nous n'avons pas d'interprétation pour ce résultat. Aussi, on constate que la dégradation de τ_{fallR} est relativement stable en fonction de la fréquence de stress.

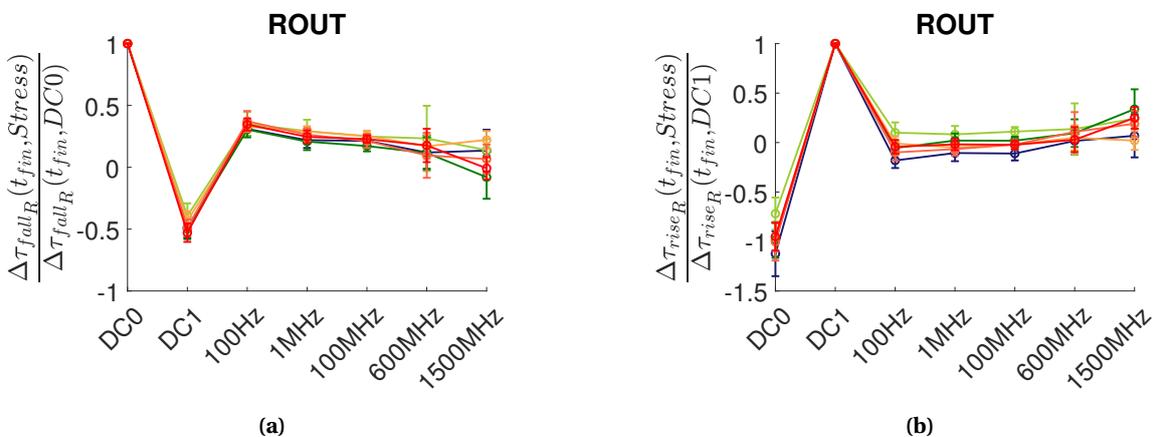


FIGURE 4.25 – Amplitude relative (par rapport au stress statique) des dégradations dans les ressources de routages au dernier instant de mesure en fonction du stress interne et pour différentes conditions de stress de température et de tension (a) τ_{fallR} (b) τ_{riseR} . $\alpha_{stress} = 0.5$

En conclusion de cette section et, concernant les ressources de LUT, nous avons observé quelques

caractéristiques du HCI, cependant, celui-ci reste négligeable par rapport au BTI même à basse température et pour un stress de haute fréquence. Globalement, aussi bien pour les ressources de LUT que de routages, l'effet de la fréquence de stress est faible sur les dégradations. De plus, les dégradations sont environs trois fois plus faibles pour un stress dynamique que pour un stress statique. Par conséquent, nous ne modéliserons pas l'effet de la fréquence de stress qui reste négligeable par rapport à l'effet du rapport cyclique, de la température et de la tension de stress.

4.5 Comparaison des dégradations dans un FPGA Artix 28nm HKMG et Zynq UltraScale+ 16nm FinFET

Dans cette thèse nous mesurons la fiabilité de FPGA de la dernière génération. Par dernière génération nous entendons des FPGA conçu avec une technologie FinFET. Dans la section 1.2 nous avons montré la différence d'architecture entre le FinFET et un transistor planaire. Dans la section 1.5 nous avons montré que cette évolution d'architecture génère un auto-échauffement plus important dans le transistor ce qui pouvait impacter les mécanismes de dégradation. Alors la question se pose : comment va évoluer la fiabilité entre un FPGA conçu avec une technologie planaire et un FPGA avec du FinFET?

Dans cette section, nous comparons les dérives des dégradations que nous avons observées pour des FPGAs Zynq UltraScale+ 16nm FinFET avec celles observées par Julien Coutet dans un FPGA Artix 28 nm HKMG [2]. Julien Coutet modélise les dégradations de la dérive fréquentielle des ROs pour un stress statique DC0 et DC1 avec le modèle suivant :

$$\Delta f_R(t, V, T) = a(V, T)t^b = A \cdot e^{\gamma \cdot V} e^{-\frac{E_a}{k_B T}} \cdot t^b \quad (4.16)$$

Les figures [2, fig. 121] et [2, fig. 120] montrent les mesures et le modèle des dérives fréquentielles pour un stress DC0 et DC1 dans un FPGA Artix. Les tables [2, tab. 27] et [2, fig. 26] présentent les paramètres du modèle (4.16) extraits pour un stress DC0 et DC1.

À partir de nos mesures, nous modélisons les dérives fréquentielles des ROs dans le FPGA Zynq UltraScale+ à partir du même modèle utilisé par Julien Coutet (4.16). Dans notre modélisation, nous utilisons les RO L1 B et RO L2 XB qui sont des architectures logiques identiques de celles utilisées dans [2]. Aussi, nous utilisons les mesures pour des conditions de stress de température ($25^\circ\text{C} \leq T_{stress} \leq 115^\circ\text{C}$) et de tension ($V_{nom} \leq V_{stress} \leq 1.3V_{nom}$) proches de celles utilisées dans [2] ($55^\circ\text{C} \leq T_{stress} \leq 115^\circ\text{C}$ et $V_{nom} \leq V_{stress} \leq 1.5V_{nom}$).

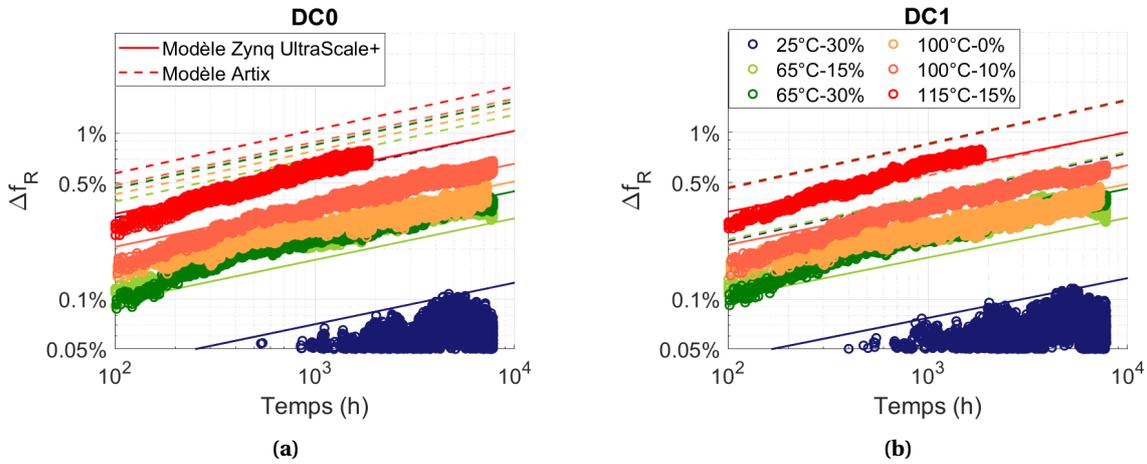
La table 4.4 regroupe les paramètres obtenus pour notre modélisation et ceux obtenus dans [2]. La figure 4.26 présente nos mesures de fréquences relatives, notre modèle identifié (traits pleins) ainsi que le modèle de [2] (pointillés) tracé pour les mêmes conditions de températures et de tension que les nôtres.

Sur la table 4.4, les coefficients du facteur d'accélération de température (E_a) et de tension (γ) sont plus important pour le FPGA Zynq UltraScale+ que pour l'Artix. Ceci dénote une plus grande sensibilité de l'amplitude des dégradations à la température et à la tension pour le Zynq UltraScale+. Cette sensibilité est visible sur la figure 4.26 où l'on constate les dégradations pour différentes conditions de température et de tension bien plus espacées pour le modèle du Zynq UltraScale+ que pour le modèle de l'Artix.

Pour l'Artix, les coefficients des facteurs d'accélération sont différents pour un stress DC0 et DC1. Julien Coutet [2, §3.4.1.1.1 et §3.4.1.1.2] justifie cette différence par l'explication suivante : en considérant que, pour un stress DC0, les types de transistors majoritairement stressés sont des PMOS, alors le NBTI constitue la principale cause des dégradations du RO. À l'inverse, pour un stress DC1, le PBTI est la principale cause des dégradations. Ainsi, la dépendance du NBTI et du PBTI à la température et à la tension de stress est différente. Cependant, dans notre cas, les coefficients des facteurs d'accélération que nous obtenons pour DC0 et DC1 sont presque identiques.

TABLE 4.4 – Paramètres du modèles (4.16) extraits pour un FPGA Zynq UltraScale+ 16nm FinFET et un FPGA Artix 28nm HKMG

	Zynq UltraScale+		Artix	
	DC0	DC1	DC0	DC1
$A (\times 10^2 \cdot h^{-b})$	-21.3	-15.3	-0.592	-0.066
b	0.251	0.240	0.262	0.265
E_a (eV)	0.275	0.268	0.089	0.160
γ (V^{-1})	2.957	3.156	1.231	4.804
$a(V_{nom}, 25^\circ C) (\times 10^3 \cdot h^{-b})$	-5.9	-6.6	-64.0	-15.6
$a(1.2V_{nom}, 115^\circ C) (\times 10^3 \cdot h^{-b})$	-117	-127	-182	-174


FIGURE 4.26 – Comparaison de la dégradation des dérives fréquentielles des ROs dans un FPGA Zynq UltraScale+ 16nm FinFET et Artix 28nm HKMG [2] obtenue dans les mêmes conditions de température et de tension

Donc dans une technologie FinFET, le NBTI et le PBTI semblent avoir la même dépendance à la température et la tension de stress.

La figure 4.26 montre que, pour une condition de température et de tension donnée, les dégradations dans l'Artix sont toujours supérieures aux dégradations dans le Zynq UltraScale+.

Les deux dernières lignes de la table 4.4 montrent le résultat du paramètre d'amplitude ($a(V, T)$) en tenant compte des facteurs d'accélération de température et de tension. On constate que le paramètre d'amplitude du modèle de l'Artix est supérieur à celui du Zynq UltraScale+, ce qui explique pourquoi nous observons des dégradations plus importantes pour l'Artix sur la figure 4.26. Parce que les dégradations dans le Zynq UltraScale+ sont plus sensibles à la température, on constate que pour $V = V_{nom}$ et $T = 25^\circ C$ les dégradations sont dix fois plus faibles que dans un Artix. Pour $V = 1.2V_{nom}$ et $T = 115^\circ C$, les dégradations dans le Zynq UltraScale+ sont 1.5 fois plus faibles que dans l'Artix.

Finalement, l'exposant au temps (b) pour le modèle de l'Artix est sensiblement égal à celui que nous avons mesuré dans le Zynq UltraScale+ ($0.240 \approx 0.265$).

Dans cette section, nous avons comparé les dégradations que nous avons modélisées dans un FPGA Zynq UltraScale+ 16nm avec la modélisation des dégradations dans un FPGA Artix 28nm

HKMG. Cette comparaison révèle, qu'en fonction de la condition de température et de tension, le facteur d'amplitude des dégradations ($a(V,T)$) est 1.5 à 10 fois plus importante pour l'Artix. L'exposant au temps est sensiblement similaire entre l'Artix et le Zynq UltraScale+. Donc, le FPGA UltraScale+ semble plus robuste aux dégradations du BTI pour des températures comprises entre 25°C et 115°C et des tensions comprises entre V_{nom} et $1.3V_{nom}$.

4.6 Modélisation des défaillances

4.6.1 Introduction

"La fiabilité est l'aptitude d'un dispositif à accomplir une fonction requise dans des conditions données pour une période de temps donnée". Dans un FPGA, le temps de propagation d'un signal dans un chemin combinatoire fixe la fréquence de fonctionnement c'est à dire la vitesse pour qu'une fonction soit exécutée. De plus, le lien entre la logique combinatoire et séquentielle est la synchronisation des signaux qui est fortement dépendante, là aussi, du temps de propagation du signal dans un chemin combinatoire. Dans notre étude nous mesurons le temps de propagation d'un signal dans un chemin combinatoire en implémentant des ROs. Dans notre cas on considère que la fonction requise est l'oscillation d'un RO au-dessus d'une certaine fréquence de seuil. Le vieillissement du RO a pour conséquence l'augmentation du temps de propagation à l'échelle du transistor, soit la diminution de la fréquence d'oscillation du RO. Lorsque la fréquence du RO devient plus faible que la fonction de seuil, alors on considère que le RO est défaillant et on observe alors une panne catalectique.

Dans cette section, nous nous proposons de modéliser la fiabilité des FPGAs à partir d'une approche statistique de la défaillance des ROs. Pour cela, nous extrapolons nos mesures de dégradations jusqu'à un critère de défaillance préalablement fixé et nous notons le temps de défaillance (TTF : Time To Fail). Ainsi, nous modélisons la fonction de répartition des TTF à partir d'une loi de distribution log-normale tout en identifiant les facteurs d'accélération de température et tension. Cette étude statistique sert à confirmer les résultats de la précédente modélisation. Aussi, les résultats issus d'une étude statistique de la fiabilité, tel que le MTTF (Mean Time To Failure), le taux de défaillance, sont directement exploitable par l'industriel avec beaucoup de précautions.

4.6.2 Seuil de défaillance

Pour connaître l'instant de défaillance du RO, il faut d'abord fixer un seuil de défaillance à partir duquel on considère le RO défaillant. Dans la continuité de la modélisation des dégradations, nous utiliserons le temps de propagation d'un front montant ou descendant calculé à partir de la mesure de fréquence et du rapport cyclique :

$$\begin{aligned}\tau_{fall,ro} &= \frac{\alpha_{ro}}{f_{ro}} \\ \tau_{rise,ro} &= \frac{1 - \alpha_{ro}}{f_{ro}}\end{aligned}\tag{4.17}$$

Dans un premier temps nous choisirons comme seuil de défaillance celui fixé par Vivado. Comme expliqué dans la section 3.3.2, pour chaque RO nous avons extrait le temps de propagation prédit par Vivado. Vivado retourne quatre temps de propagation : Fast Max, Fast Min, Slow Max et Slow Min. Fast et Slow correspondent au FPGA le plus rapide et le plus lent en sortie de fabrication tandis que Max et Min correspondent aux ressources dans le FPGA avec le maximum et le minimum de temps de propagation. Le seuil de défaillance extrait de Vivado correspond au temps de propagation maximum dans un FPGA avec des performances moyennes. Pour chaque RO on calcule :

$$\tau_{vivado} = \frac{\tau_{fast,max} + \tau_{slow,max}}{2}\tag{4.18}$$

Pour différents ROs, la figure 4.27a compare le temps de propagation limite donné par Vivado, que l'on peut considérer comme étant le seuil de dégradation, avec le temps de propagation d'un front descendant mesuré aux conditions nominales avant vieillissement. En moyenne, le seuil de défaillance fixé par Vivado est 1.21 fois plus grand que le temps de propagation mesuré. Cela signifie que pour considérer un RO défaillant, il faut que sa dérive relative de la dégradation du temps de propagation d'un front montant soit de 21%. La figure 4.27b compare le seuil de dégradation de Vivado avec la mesure après vieillissement, c'est à dire au dernier instant de mesure. La figure 4.27b révèle la faible dégradation des ROs par rapport au seuil de défaillance. Même pour la condition de stress la plus importante ($T_{stress} = 115^{\circ}\text{C}$ et $V_{stress} = 1.15V_{nom}$) en comparant avec les mesures obtenues avant le vieillissement, on constate une dégradation de 1.3% ce qui faible devant les 21% du seuil de dégradation de Vivado.

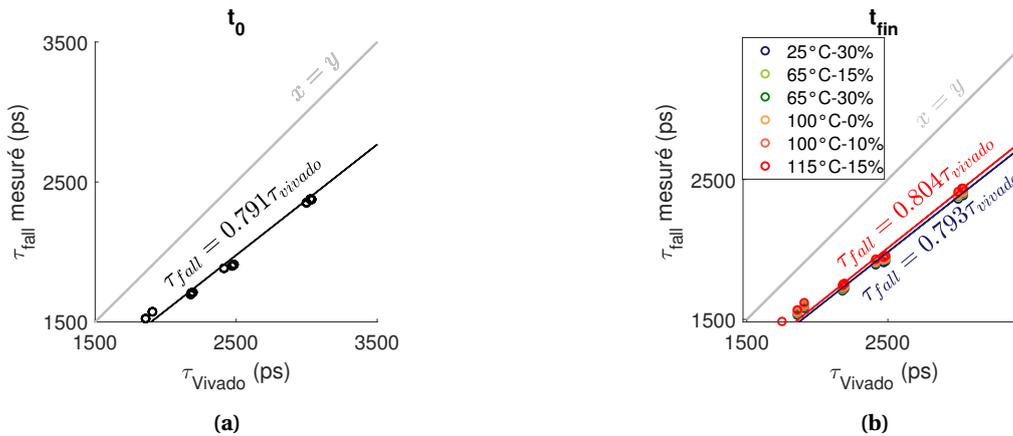


FIGURE 4.27 – Comparaison du seuil de défaillance extrait de Vivado (4.18) avec la mesure du temps de propagation d'un front descendant (a) avant vieillissement (b) après vieillissement pour un stress DC0. Chaque point correspond à un RO.

Les dégradations relatives maximum observées sur notre banc de test sont d'environ 2% pour la dégradation d'un front descendant, avec un stress DC0 et $T_{stress} = 115^{\circ}\text{C}$, $V_{stress} = 1.15V_{nom}$. Si nous souhaitons observer des ROs défaillants, ce qui est nécessaire dans le cadre de notre étude statistique il est nécessaire d'extrapoler nos mesures.

Afin de comparer nos résultats avec ceux de Julien Coutet [2], en plus du seuil de défaillance de Vivado, nous prendrons un seuil de défaillance relative de 10%.

4.6.3 Extrapolation des dérives

Nous souhaitons extrapoler nos mesures de dégradation jusqu'au seuil de dégradation. Pour cela nous avons modélisé la dégradation relative de chaque RO avec le modèle suivant :

$$\Delta\tau_R = \frac{\tau(t) - \tau_{t_0}}{\tau_{t_0}} = at^b \quad (4.19)$$

On fixe pour tous les ROs l'exposant au temps $b = 0.27$ qui correspond au modèle de dégradation établi dans la section 4.2 en considérant que le BTI est le principal mécanisme de dégradation observé.

À partir du modèle nous extrapolons la dérive jusqu'à atteindre le seuil de défaillance tel que :

$$at_{\text{TTF}}^b = \frac{\tau_{\text{seuil}}}{\tau_{t_0}} - 1$$

$$t_{\text{TTF}} = \left(\frac{\tau_{\text{seuil}}}{\tau_{t_0}} - 1 \right)^{1/b} \frac{1}{a} \quad (4.20)$$

Si le seuil de défaillance est fixé par Vivado alors $\tau_{seuil} = \tau_{vivado}$. Si le seuil de défaillance est de 10% alors $\tau_{seuil} = 1.1\tau_{t_0}$.

4.6.4 Modélisation statistique de la durée de vie

Afin de modéliser la durée de vie des FPGAs, comme pour la modélisation des dégradations, nous utilisons les données des six FPGAs avec un stress de température et de tension suivant : 25°C et 1.3V_{nom}, 65°C et 1.15V_{nom}, 65°C et 1.3V_{nom}, 100°C et V_{nom}, 100°C et 1.1V_{nom}, 115°C et 1.15V_{nom}. Pour que la durée de vie modélisé couvre toutes les configuration présentes dans notre banc de test, nous utilisons les ROs pour lesquels nous observons le plus de dégradation : architecture Short Path de type buffer (voir section 2.3.2), avec un stress DC0. Afin d'éviter les phénomènes d'améliorations observés par exemple sur la figure 4.11, nous utilisons pour un stress DC0 la dégradation d'un front descendant dans le RO. Nous avons 18 ROs par FPGA soit au total 108 ROs pour faire l'étude statistique. Pour chaque RO, nous extrapolons les dérives jusqu'au seuil de défaillance.

Comme dans [2], nous utilisons une loi de distribution log-normale, avec pour fonction de répartition :

$$F_{TTF}(t) = \frac{1}{2} \left(1 + erf \left(\frac{\ln(t) - \mu}{\sqrt{2}\sigma} \right) \right) \quad (4.21)$$

Pour une loi log-normale, le logarithme de TTF suit une loi normale. Ce qui à pour conséquence que la *mediane* = e^μ , l'*esperance* = $e^{\mu+\sigma^2/2}$ et la *variance* = $(e^{\sigma^2} - 1) e^{2\mu+\sigma^2}$. La fonction *erf* est la fonction d'erreur (1.16).

Pour chaque FPGA nous appliquons une régression non linéaire avec le modèle (4.21). La figure 4.28 compare le modèle que nous avons extrait de la régression avec les données expérimentales. Tout d'abord on constate que les temps de défaillances sont extrêmement grands, nous avons d'ailleurs exprimé l'axe des abscisses en années. Pour le seuil de défaillance Vivado, nous obtenons : $13.7 \leq \mu \leq 21.4$ et $0.96 \leq \sigma \leq 1.06$. Pour le seuil de défaillance Vivado, nous obtenons : $10.6 \leq \mu \leq 18.3$ et $0.55 \leq \sigma \leq 0.69$. On observe, sur la figure 4.28a, une discontinuité qui est très probablement liée aux différences de dégradation entre deux ROs d'architectures différentes. Finalement, l'alignement des mesures confirme que la distribution des défaillances suit une loi log-normale.

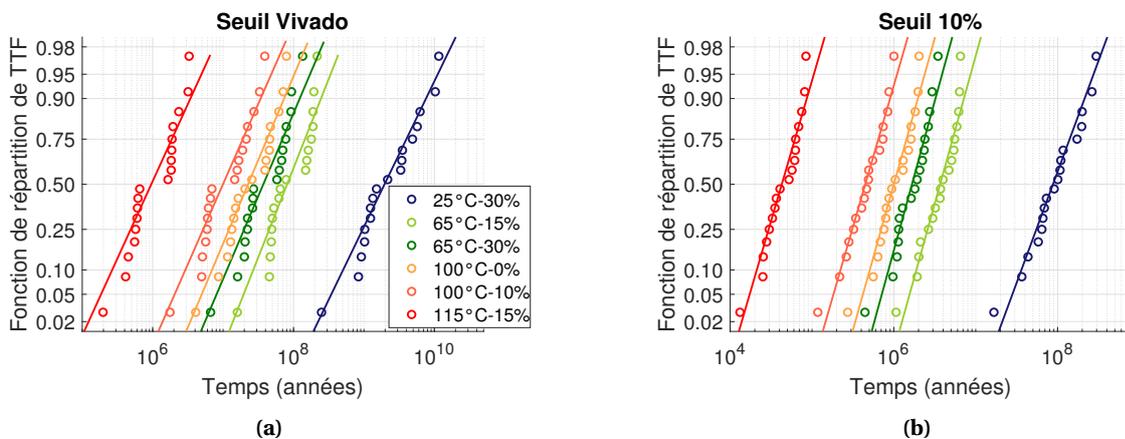


FIGURE 4.28 – Comparaison entre les données d'expérimentation et le modèle (4.21) pour la fonction de répartition de TTF (a) pour le seuil de défaillance Vivado (b) pour le seuil de défaillance 10%

À présent nous allons identifier le facteur d'accélération de température et de tension avec une loi d'Arrhenius et une loi exponentielle comme pour la modélisation des dégradations. Ainsi nous pouvons exprimer de nouveau la fonction de répartition :

$$F_{TTF}(t, V_{stress}, T_{stress}) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\ln(t e^{\gamma_{def} (V_{stress} - V_{op})} e^{\frac{Ea_{def}}{k_B} (\frac{1}{T_{op}} - \frac{1}{T_{stress}})}) - \mu}{\sqrt{2}\sigma}} \right) \right) \quad (4.22)$$

$$F_{TTF}(t, V_{stress}, T_{stress}) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\ln(t) + \gamma_{def} (V_{stress} - V_{op}) + \frac{Ea_{def}}{k_B} (\frac{1}{T_{op}} - \frac{1}{T_{stress}}) - \mu}{\sqrt{2}\sigma}} \right) \right)$$

Nous notons γ_{def} et Ea_{def} les paramètres extraits avec la modélisation de la répartition des temps de défaillance, tandis que, nous notons γ_{deg} et Ea_{deg} les paramètres extraits avec la modélisation de la dégradation du temps de propagation dans les sections précédentes.

À partir du modèle (4.22) nous avons extrait les coefficients γ_{def} et Ea_{def} des facteurs d'accélération que l'on retrouve dans la table 4.5. Ainsi, nous pouvons calculer le temps de défaillance (TTF) des ROs pour les conditions de fonctionnement opérationnel de température (T_{op}) et de tension (V_{op}), tel que :

$$TTF_{V_{op}, T_{op}} = TTF_{V_{stress}, T_{stress}} \times e^{\gamma(V_{stress} - V_{op})} \times e^{\frac{Ea}{k_B} (\frac{1}{T_{op}} - \frac{1}{T_{stress}})} \quad (4.23)$$

Sur la figure 4.29a nous avons tracé les temps de défaillances ramenés aux conditions opérationnelles $V_{op} = V_{nom}$ et $T_{op} = 25^\circ\text{C}$ pour le seuil de défaillance Vivado. Sur la figure 4.29b nous avons tracé les temps de défaillances ramenés aux conditions opérationnelles $V_{op} = 1.05V_{nom}$ et $T_{op} = 100^\circ\text{C}$ pour le seuil de défaillance à 10%, afin de comparer les résultats avec [2].

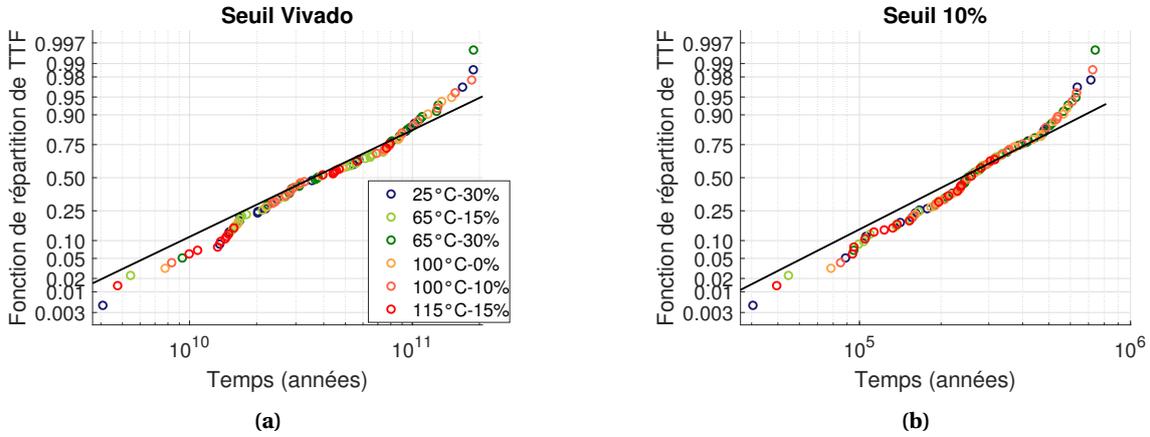


FIGURE 4.29 – Comparaison entre les données d'expérimentation et le modèle (trait plein) (4.22) pour la fonction de répartition de t_{TTF} avec (a) pour le seuil de défaillance Vivado, ramené à $T_{op} = 25^\circ\text{C}$ et $V_{op} = V_{nom}$ (b) pour le seuil de défaillance 10%, ramené à $T_{op} = 100^\circ\text{C}$ et $V_{op} = 1.05V_{nom}$.

Sur la table 4.5 on remarque une différence importante du coefficient μ entre le Zynq UltraScale+ et l'Artix, ce qui a une conséquence directe sur le MTTF qui est de 3.3×10^5 années pour le Zynq UltraScale+ et de 1.6×10^3 années pour l'Artix. Un tel MTTF n'a pas de sens, mais cela montre que la fiabilité du FPGA n'est que très peu impactée par les dégradations liées au BTI. Cela montre aussi, que le Zynq UltraScale+ 16nm FinFET est plus robuste au BTI que l'Artix 28nm HKMG, ce qui rejoint la comparaison des dégradations faite dans la section 4.5. La table 4.5 montre que le paramètre de forme σ ainsi que les coefficients des facteurs d'accélération γ et Ea sont proches, cela conforte la solidité de notre étude et met en évidence que le mécanisme de dégradation observé sur l'Artix est identique de celui que nous observons.

Pour faire le rapprochement entre les coefficients des facteurs d'accélération obtenus pour la modélisation des dégradations et des défaillances, nous devons appliquer la relation suivante

TABLE 4.5 – Paramètres extraits à partir du modèle (4.22) dans le cadre de notre étude sur le Zynq UltraScale+ 16nm FinFET et ceux obtenus dans [2] pour l'Artix 28nm HKMG

Paramètres	Zynq UltraScale+ 16nm FinFET						Artix 28nm HKMG		
	Seuil Vivado			Seuil 10%			Seuil 10%		
	95%	standard	105%	95%	standard	105%	95%	standard	105%
μ	24	24.3	24.6	12.3	12.38	12.5	1.5	3.45	5.39
σ	0.95	1.06	1.17	0.7	0.82	0.94	0.73	0.80	0.87
$\gamma_{def}(V^{-1})$	9.96	11.02	12.1	10.2	11.18	12.2	11.2	11.88	12.6
$Ea_{def}(eV)$	0.93	0.96	0.99	0.94	0.97	1	0.75	0.81	0.87

démontrée dans [2, eq. 111] :

$$\begin{aligned} \gamma_{deg} &= \gamma_{def} \times b \\ Ea_{deg} &= Ea_{def} \times b \end{aligned} \quad (4.24)$$

$$\begin{aligned} \gamma_{deg} &= 11.18 \times 0.3 = 3.35 \text{ V}^{-1} \\ Ea_{deg} &= 0.97 \times 0.3 = 0.29 \text{ eV} \end{aligned}$$

où b est l'exposant au temps du modèle de puissance. Pour la modélisation des dégradations nous obtenons : $\gamma_{deg} = 4.21 \text{ V}^{-1}$ et $Ea_{deg} = 0.31 \text{ eV}$ (voir table 4.1). La modélisation des défaillances effectuée dans cette section confirme la justesse des coefficients du facteur d'accélération de température et de tension obtenus dans la modélisation des dégradations.

À présent nous introduisons le taux de défaillance puisque c'est une donnée facilement exploitable par un industriel. Le taux de défaillance permet de connaître pour un temps de fonctionnement donné, la probabilité qu'un échantillon soit en panne pour une heure d'utilisation. Le taux de défaillance d'une loi log-normale s'écrit :

$$\lambda_{\text{TTF}}(t, V_{\text{stress}}, T_{\text{stress}}) = \frac{f_{\text{TTF}}(t, V_{\text{stress}}, T_{\text{stress}})}{1 - F_{\text{TTF}}(t, V_{\text{stress}}, T_{\text{stress}})} = \frac{\frac{1}{t \cdot \sigma \cdot \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln(t \cdot A_V \cdot A_T) - \mu}{\sigma} \right)^2}}{1 - \frac{1}{2} \left(1 + \text{erf} \left(\frac{\ln(t \cdot A_V \cdot A_T) - \mu}{\sqrt{2}\sigma} \right) \right)} \quad (4.25)$$

La figure 4.30 présente l'évolution du taux de défaillance exprimé en FIT (Failure In Time). Dans notre cas, le FIT correspond au nombre de RO défaillants dans un lot de 10^9 ROs pendant une heure de fonctionnement. Le FIT est souvent utilisé en électronique pour compenser la grande quantité de composants dans un circuit intégré. Le FPGA que nous utilisons contient autour de 10^6 cellules logiques composées de huit LUT (Look-Up Table). Un lot de 10^9 ROs correspond alors à 1000 FPGAs implémentant chacun 10^6 ROs de huit étages. La figure 4.30 montre que le taux de défaillance est croissant, ce qui confirme l'observation des défaillances causées par le vieillissement (partie de droite de la courbe en baignoire sur la figure 1.38). En considérant le seuil de défaillance le plus critique (10%) ainsi que la température la plus haute, après 50 ans le taux de défaillance passe la barre des 10^{-9} FIT, ce qui est acceptable pour l'industrie. Pour compléter l'analyse du taux de défaillance, la figure 4.31 présente l'évolution du taux de défaillance en fonction du temps de fonctionnement et du seuil de défaillance choisit pour $V_{op} = 1.05 V_{nom}$ et $T_{op} = 125^\circ \text{C}$. Pour certaines applications, il est possible qu'une baisse de performance inférieure à 10% ne soit pas acceptable par l'utilisateur. Pour un seuil de défaillance au-dessus de 10% le taux de défaillance devient dérisoire, c'est pourquoi nous l'avons limité à 10% sur la figure 4.31.

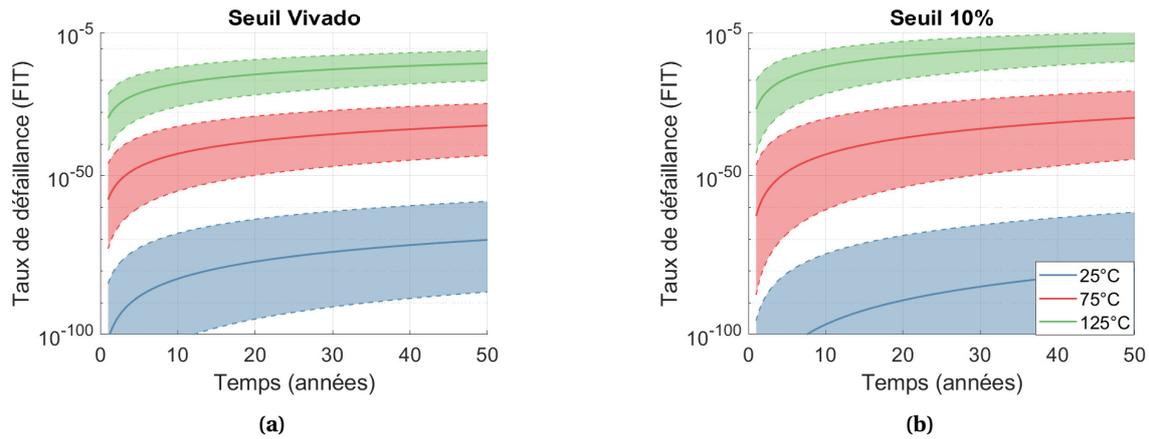


FIGURE 4.30 – Taux de défaillance avec une enveloppe à 10% pour une tension opérationnelle $V_{op} = 1.05V_{nom}$ et pour différentes température avec comme seuil de défaillance (a) Vivado et (b) 10%

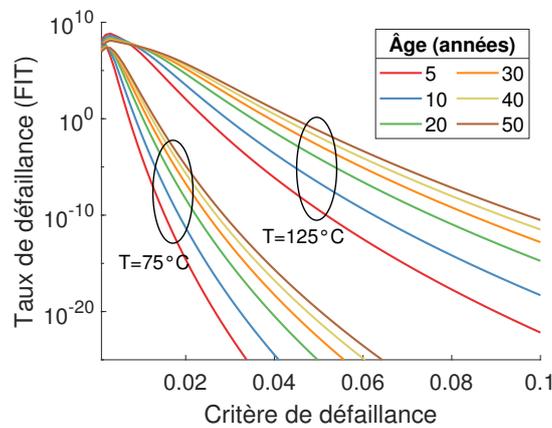


FIGURE 4.31 – Évolution du taux de défaillance en fonction du temps de fonctionnement et du seuil de défaillance pour $V_{op} = 1.05V_{nom}$ et deux températures de fonctionnement $T_{op} = 75^{\circ}\text{C}$ et $T_{op} = 125^{\circ}\text{C}$.

4.6.5 Conclusion

Dans cette section nous avons modélisé les défaillances des ROs en considérant un seuil de défaillance extrait de Vivado, en moyenne à 26.4%, ainsi qu'un seuil de défaillance de 10%. Nous utilisons les ROs pour lesquelles nous avons observé le plus de dégradation, soit pour un stress DC0. Nous avons utilisé une loi log-normale pour modéliser les temps de défaillance, puis, nous avons introduit les facteurs d'accélération de température et de tension dans ce même modèle. Les coefficients des facteurs d'accélération que nous obtenons sont très proches de ceux obtenus pour la modélisation des dégradations ($E_{a_{def}} = 0.97 \times 0.3 = 0.29\text{eV} \approx 0.31\text{eV} = E_{a_{deg}}$ et $\gamma_{def} = 11.18 \times 0.3 = 3.35\text{V}^{-1} \approx 4.21\text{V}^{-1} = \gamma_{deg}$), ce qui confirme la justesse de nos modélisations. Dans les conditions d'utilisation les plus critiques, nous obtenons $\text{MTTF} = 3.3 \times 10^5 \text{annees}$ qui est 100 fois plus grand que celui obtenu dans [2] pour l'Artix 28nm HKMG. Le MTTF obtenu prouve que le FPGA est très robuste face aux défaillances provoquées par le BTI. Pour l'étude du taux de défaillance nous obtenons $\lambda = 10^{-9}\text{FIT}$ après 50 ans de fonctionnement avec un seuil de défaillance à 10%, ce qui est largement acceptable par l'industrie. Finalement, nous proposons un abaque qui permet à l'utilisateur de connaître le taux de défaillance en fonction du seuil de défaillance et du temps de fonctionnement.

4.7 Conclusion

En séparant les dégradations dans les ressources de LUT et de routages, dans ce chapitre nous avons modélisé la dégradation des ROs mesurée par notre banc de test. Nous avons d'abord proposé un modèle semi-empirique pour les dégradations sous un stress statique. Ensuite, nous avons modélisé le mécanisme de récupération du BTI pour un stress dynamique que nous avons combiné avec le modèle statique pour obtenir un seul modèle qui prédit les dégradations sous un stress statique et dynamique. À haute température nous avons observé un effet de la fréquence de stress sur les dégradations qui est négligeable dans notre modélisation. À basse température, nous avons cherché du HCI mais les résultats ne mettent pas en évidence un effet significatif de ce mécanisme de dégradation qui est négligeable à 25°C devant le BTI. Finalement, en associant le modèle pour prédire les dégradations dans les ressources de LUT et de routages, nous obtenons un unique modèle qui prédit les dégradations dans un RO en fonction du temps de fonctionnement, de la tension de stress, de la température de stress, du rapport cyclique de stress, de l'architecture du RO caractérisé par son temps de propagation initiale (τ_{t_0}) et des entrées de LUT utilisées (R_{ILUT}) :

$$\begin{aligned}
 \Delta\tau_{R_{ERO}}(t, V_{stress}, T_{stress}, \alpha_{stress}, R_{ILUT}, \tau_{t_0}) = & \\
 & \left[\frac{R_{ILUT} \cdot A_{ILUT_{EDC1,LUT}}}{\tau_{t_0}} \cdot AF_{V_{EDC1,LUT}}(V_{stress}) \cdot AF_{T_{EDC1,LUT}}(T_{stress}) \cdot (\alpha_{stress} \cdot t)^{b_{EDC1,LUT}} \right] \\
 & \quad \times RF_{1_{LUT}}(V_{stress}, T_{stress}) \\
 + & \left[\frac{R_{ILUT} \cdot A_{ILUT_{EDC0,LUT}}}{\tau_{t_0}} \cdot AF_{V_{EDC0,LUT}}(V_{stress}) \cdot AF_{T_{EDC0,LUT}}(T_{stress}) \cdot (\alpha_{stress} \cdot t)^{b_{EDC0,LUT}} \right] \\
 & \quad \times RF_{0_{LUT}}(V_{stress}, T_{stress}) \\
 + & \left[(A_{t_{EDC1,ROUT}} \cdot \tau_{t_0} + B_{EDC1,ROUT}) \cdot AF_{V_{EDC1,ROUT}}(V_{stress}) \cdot AF_{T_{EDC1,ROUT}}(T_{stress}) \cdot (\alpha_{stress} \cdot t)^{b_{EDC1,ROUT}} \right] \\
 & \quad \times RF_{1_{ROUT}}(V_{stress}, T_{stress}) \\
 + & \left[(A_{t_{EDC0,ROUT}} \cdot \tau_{t_0} + B_{EDC0,ROUT}) \cdot AF_{V_{EDC0,ROUT}}(V_{stress}) \cdot AF_{T_{EDC0,ROUT}}(T_{stress}) \cdot (\alpha_{stress} \cdot t)^{b_{EDC0,ROUT}} \right] \\
 & \quad \times RF_{0_{ROUT}}(V_{stress}, T_{stress})
 \end{aligned} \tag{4.26}$$

où $AF_{V_{ES,Res}}(V_{stress})$ et $AF_{T_{ES,Res}}(T_{stress})$ sont les facteurs d'accélération en tension et température extraits pour un stress (S) DC0 et DC1 et pour les ressources (Res) de LUT et de routage (ROUT), tel que :

$$\begin{aligned}
 AF_{V_{ES,Res}}(V_{stress}) &= e^{Y_{ES,Res}(V_{stress}-V_{op})} \\
 AF_{T_{ES,Res}}(T_{stress}) &= e^{\frac{Ea_{ES,Res}}{k_B} \left(\frac{1}{T_{op}} - \frac{1}{T_{stress}} \right)}
 \end{aligned} \tag{4.27}$$

où $RF_{1_{Res}}(V_{stress}, T_{stress})$ et $RF_{0_{Res}}(V_{stress}, T_{stress})$ sont les fractions de récupération pour les transistors stressés par un niveau '1' et '0' dans les ressources (Res) de LUT et de routage (ROUT), tel que :

$$\begin{aligned}
 RF_{1_{Res}}(V_{stress}, T_{stress}) &= \frac{1}{1 + B_{1_{Res}}(V_{stress}, T_{stress}) \left(\frac{(1-\alpha_{stress})}{\alpha_{stress}} \right)^{\beta_{1_{Res}}(V_{stress}, T_{stress})}} \\
 RF_{0_{Res}}(V_{stress}, T_{stress}) &= \frac{1}{1 + B_{0_{Res}}(V_{stress}, T_{stress}) \left(\frac{\alpha_{stress}}{(1-\alpha_{stress})} \right)^{\beta_{0_{Res}}(V_{stress}, T_{stress})}}
 \end{aligned} \tag{4.28}$$

Dans l'équation (4.26) les deux premiers termes qui s'additionnent modélisent les dégradations dans les ressources de LUT et les deux derniers termes qui s'additionnent modélisent les dégradations dans les ressources de routages. Les valeurs numériques des coefficients du modèle (4.26) se trouvent dans les tables 4.1, 4.2, 4.3. Le modèle contient 76 paramètres :

- $A_{ILUT_{FDC1,LUT}}$ et $A_{ILUT_{FDC0,LUT}}$: matrices contenant chacune six paramètres d'amplitudes propres à chaque entrée de LUT;
- $A_{t0_{FDC1,ROUT}}$, $A_{t0_{FDC0,ROUT}}$, $B_{FDC1,ROUT}$ et $B_{FDC0,ROUT}$: quatre paramètres d'ajustements de l'amplitude des dégradations dans les ressources de routage;
- $AF_{V_{FDC1,LUT}}$, $AF_{V_{FDC0,LUT}}$, $AF_{V_{FDC1,ROUT}}$ et $AF_{V_{FDC0,ROUT}}$: quatre paramètres pour les facteurs d'accélération en tension;
- $AF_{T_{FDC1,LUT}}$, $AF_{T_{FDC0,LUT}}$, $AF_{T_{FDC1,ROUT}}$, $AF_{T_{FDC0,ROUT}}$: quatre paramètres pour les facteurs d'accélération en température;
- $b_{FDC1,LUT}$, $b_{FDC0,LUT}$, $b_{FDC1,ROUT}$ et $b_{FDC0,ROUT}$: quatre paramètres d'exposant au temps;
- RF_{LUT} , $RF_{0,LUT}$, $RF_{1,ROUT}$ et $RF_{0,ROUT}$: quatre *Recovery Fraction* (RF) pour modéliser le mécanisme de récupération. Dans l'expression de RF il y a deux sous paramètres (B et β) qui sont dépendant de la température et de la tension et que nous avons modélisé avec un polynôme du deuxième degré à deux variables. Donc chaque RF contient 12 paramètres soit un total de 48 paramètres pour tous les RF.

Le modèle contient six données d'entrées :

- t : le temps de vieillissement;
- V_{stress} : la tension de stress;
- T_{stress} : la température de stress;
- α_{stress} : le rapport cyclique de stress;
- R_{ILUT} : les entrées de LUT utilisées par le RO;
- τ_{t0} : le temps de propagation initial dans le RO.

Avec ce modèle nous avons prédit la dégradation de tous les ROs avec une erreur moyenne de 10%.

Nos résultats de dégradations mesurés et modélisés présentent, en comparaison avec l'étude de Julien Coutet sur un FPGA Artix 28nm HKMG [2], des dégradations 1.5 fois plus faible pour $T_{stress} = 115^\circ\text{C}$ et $V_{stress} = 1.2V_{nom}$. La proximité de l'exposant au temps, obtenu dans notre étude ($b = 0.25$) et celle de Julien Coutet ($b = 0.26$), montre que le même mécanisme de défaillance (BTI) est observé. Globalement, on peut confirmer que le FPGA Zynq UltraScale+ 16nm étudié dans cette thèse est plus robuste aux dégradations du temps de propagation dû au BTI que le FPGA Artix 28nm HKMG.

Finalement, en fixant un seuil de défaillance des ROs, nous avons modélisé les défaillances du FPGA. En utilisant un seuil de défaillance extrait de Vivado, nous avons obtenu un MTTF de l'ordre de 10^{10} *années* qui est extrêmement supérieur au temps d'utilisation de toutes applications envisageables. Ensuite, nous avons choisi un seuil de défaillance de 10% qui est identique à celui fixé par Julien Coutet [2]. Nous avons obtenu un MTTF de 10^5 *années*, 100 fois plus grand que celui dans [2] $MTTF = 10^3$ *années*. Ce dernier résultat est cohérent avec la comparaison des dégradations entre les deux technologies. De plus, les coefficients du facteurs d'accélération de température et de tension obtenus pour la modélisation des défaillances ($Ea_{def} = 0.97 \times 0.3 = 0.29\text{eV}$ et $\gamma_{def} = 11.18 \times 0.3 = 3.35\text{V}^{-1}$) sont relativement identiques à ceux obtenus pour la modélisation des dégradations ($Ea_{deg} = 0.31\text{eV}$ et $\gamma_{deg} = 4.21\text{V}^{-1}$), ce qui confirme la justesse de nos deux modélisations.

Conclusion et perspectives

Conclusion générale

Les travaux réalisés durant cette thèse visaient à mesurer et à prévoir la fiabilité sous l'effet du vieillissement des circuits numériques conçus avec une technologie FinFET. La fiabilité d'un circuit numérique est impactée par trois mécanismes de défaillance qui agissent au niveau du transistor : le Bias Temperature Instability (BTI), le Hot Carrier Injection (HCI) et le Time Dependant Dielectric Breakdown (TDDB). Ces trois mécanismes ont comme mode de défaillance commun la dégradation de la tension de seuil d'un transistor dont la conséquence est l'augmentation du temps de propagation d'un signal dans un circuit numérique. Le FPGA est un circuit numérique flexible et reconfigurable qui rend possible la conception d'une architecture de test pour mesurer les dégradations de temps de propagation d'un signal. En présentant l'état de l'art des tests de vieillissement visant à mesurer les dégradations dans des FPGAs, nous avons mis en avant l'importance de la précision de mesure qui permet ainsi d'observer des dégradations dans des conditions de vieillissement proches des conditions d'utilisation.

Le banc de test que nous avons développé se base sur l'implémentation d'oscillateurs en anneau (RO) dans un FPGA dont nous mesurons la fréquence d'oscillation. Nous avons implémenté neuf architectures différentes de RO stressés par un des 21 stress internes de différentes fréquences et rapports cycliques. Au total nous avons implémenté 567 ROs dans chacun des FPGAs. Avec la mesure de la fréquence d'oscillation du RO, nous mesurons également son rapport cyclique, ce qui nous permet de calculer le temps de propagation d'un front montant et d'un front descendant dans un RO. Dans le circuit de mesure, nous avons utilisé une référence de temps fournie par un récepteur GPS que nous évaluons 100 fois plus précise qu'une référence de temps générée par un oscillateur à quartz généralement utilisé. Parce que le RO est sensible à la température et à la tension, nous avons développé des systèmes de régulation. Pour la température nous avons utilisé un module Peltier que nous plaçons en contact avec le FPGA, nous permettant ainsi de réguler la température interne du FPGA de 15°C à 115°C avec une précision de 0.05°C. Pour la tension nous avons utilisé le Power Management Integrated Circuit (PMIC) déjà présent sur la carte de développement pour réguler la tension entre $V_{nom} = 850mV$ et $1.3V_{nom}$ avec une précision de 3.9mV. Nous avons mesuré l'évolution de la fréquence et du rapport cyclique d'oscillation d'un RO pour différentes températures et tension. Avec ces mesures, nous avons construit un modèle qui nous permet de compenser l'effet des erreurs résiduelles de régulation de la tension et de la température. Finalement, nous avons réalisé un premier test de vieillissement, pendant 1000 heures sur un seul FPGA où $T_{stress} = 100^\circ C$ et $V_{stress} = 1.25V_{nom}$, visant à montrer les performances du banc de test. Ainsi, nous avons mis en évidence que le banc de test est capable de mesurer les fréquences avec une précision relative de 9×10^{-5} , ce qui semble placer notre banc de test comme le plus précis de l'état de l'art.

Pour avoir une bonne prévision de la fiabilité des FPGAs, nous avons déployé le banc de test dans neuf FPGAs chacun avec une condition de température et de tension différentes qui sont dans les gammes 25°C – 115°C et V_{nom} – $1.3V_{nom}$. Avant de vieillir les FPGAs, nous avons mesuré les fréquences et rapports cycliques de tous les ROs et, comparant les temps de propagation de

tous les ROs issus de nos mesures avec les données de simulation extraits de Vivado nous avons montré que dans les performances des FPGAs se situent entre 50% et 70% des meilleures performances attendues par Vivado. Aussi, nous avons montré que la dispersion intra (0.6%) et inter (10%) FPGA mesurée est bien plus faible que celle prévue par Vivado (intra : 37% et inter : 93%). Après 8000 heures de vieillissement de 5103 ROs répartis dans neuf FPGAs, nous avons observé des dérives paramétriques, dues potentiellement au BTI et au HCI, sur tous les ROs et n'avons pas observé de comportement provoqué par du TDDDB. Nous avons obtenu des dégradations maximales de la fréquence relative d'oscillation de tous les ROs de l'ordre de 1%, ce qui en soi représente déjà un résultat rassurant quant à la fiabilité des FPGAs avec du FinFET. Nous avons montré que les dégradations sont plus importantes dans les ROs composés principalement de ressources logiques que dans les ROs composés principalement de ressources de routage. Ainsi, nous avons développé une méthode visant à identifier et à séparer les dégradations dans les ressources logiques et de celles dans les ressources de routage.

En nous appuyant sur la précédente méthode d'extraction, nous avons modélisé la dégradation des ROs en séparant la modélisation dans les ressources logiques et de routage. Le modèle que nous avons développé permet de prévoir la dégradation du temps de propagation d'un RO en fonction de : la température, la tension, le rapport cyclique de stress et l'architecture du RO, avec une précision de 10% par rapport aux mesures. À haute température, haute tension et pour un stress statique, nous avons observé des dégradations suivant une loi puissance avec un exposant au temps proche de 0.25 ce qui suggère que nous avons observé du BTI. De plus, les dégradations avec un stress dynamique montrent que l'amplitude dépend du rapport cyclique de stress, ce que nous interprétons comme étant la conséquence du mécanisme de récupération du BTI. En revanche, les dégradations sont relativement constantes suivant la fréquence de stress, ce qui montre que le HCI est négligeable dans nos conditions de test. Nous avons comparé nos mesures avec le modèle de dégradations provenant d'une étude similaire sur un FPGA Artix conçu avec des transistors planaires 28 nanomètres HKMG. Cette comparaison met en évidence que les dégradations dans les conditions nominales suivent dans les deux cas une loi puissance avec un exposant au temps identique mais un pré-facteur dix fois plus important pour l'Artix. Donc, la fiabilité sous l'effet du vieillissement est améliorée dans les FPGAs avec du FinFET en comparaison avec des FPGAs avec du MOSFET. Finalement, en fixant un seuil de défaillance relatif de 10%, qui est plus de deux fois plus faible à celui fixé par Vivado, nous avons modélisé la répartition des défaillances avec une loi log-normale et nous obtenons un Mean Time To Failure (MTTF) de 10^5 années. Ce MTTF prouve encore que l'impact des mécanismes de défaillance sous l'effet du vieillissement dans les transistors est mineur au regard de la fiabilité des circuits intégrés numériques à base de FinFET.

Perspectives

Évidemment les premières perspectives me venant à l'esprit touchent au banc de test. Dans les oscillateurs en anneaux (ROs) que nous avons implémentés, nous ne contrôlons pas les ressources de routage (nodes et PIPs) qui sont utilisées, et seulement pour certains oscillateurs en anneaux nous contrôlons l'entrée de la LUT utilisée. Lorsque nous mesurons les dégradations de deux ROs avec la même architecture fonctionnelle, en réalité nous mesurons deux ROs avec une architecture physique différente, compliquant ainsi l'interprétation de l'effet des stress sur les dégradations. Il serait donc intéressant d'implémenter des ROs en contrôlant toutes les ressources dans le FPGA afin d'avoir des ROs de même architecture physique. De plus, cela permettra d'avoir des données plus adéquates pour la méthode d'extraction des dégradations dans les ressources logiques et dans les ressources de routage que nous avons présentée.

Pour que notre modèle de prévision de la fiabilité des FPGAs soit plus complet il faudrait mesurer les dégradations pour des températures négatives, ce qui permettrait d'étudier l'effet du HCI, notamment à de faibles températures. Aussi, nous avons montré que les dégradations n'étaient

pas constantes en basse fréquence (100Hz), certainement à cause de la dynamique du mécanisme de récupération pour le BTI. Pour modéliser ce comportement il serait intéressant de mesurer des dégradations avec plusieurs stress de basses fréquences (0.1Hz, 1Hz, 10Hz, 100Hz ...).

La précision du système de régulation de la tension du FPGA que nous avons développé est limitée par l'utilisation d'un PMIC (Power Management Integrated Circuit) dont le pas de résolution de commande est de $3.9mV$. Pour améliorer la précision et la stabilité du banc de test il faudrait utiliser un PMIC avec une meilleure résolution.

À partir de l'extraction des dégradations dans les ressources logiques (LUT) et en connaissant l'architecture interne d'une LUT, nous avons montré comment extraire les dégradations à l'échelle du transistor. Cependant, cette méthode repose sur la supposition de l'architecture interne de la LUT. Pour consolider l'hypothèse de l'architecture interne de la LUT, il pourrait être intéressant de faire des simulations de l'architecture interne d'une LUT conçue avec des transistors 16 nanomètres FinFET de TSMC, ce qui correspond à la technologie présente dans le FPGA à l'étude dans cette thèse. Il sera alors peut-être nécessaire de développer une architecture dans le FPGA afin de comparer les résultats expérimentaux avec les résultats de simulation, de modifier l'architecture interne de la LUT jusqu'à la convergence des résultats de simulation.

Dans cette thèse nous avons montré que la mesure de la dégradation de ROs pendant un test de vieillissement nous permet de prévoir la fiabilité d'un FPGA. Je pense qu'il est possible et intéressant d'étendre cette méthode à une intégration dans un FPGA fonctionnel afin de générer en temps réel des données sur l'état de vie du FPGA. Nous pouvons développer un IP (Intellectual Property) intégrant quelques ROs et un circuit de mesure occupant moins de 1% des ressources totales dans un FPGA. Cet IP mesurerait la dégradation de la fréquence d'oscillation du RO et fournirait à l'utilisateur des données de dégradations relatives permettant de connaître l'état de vie du FPGA par rapport à son état initial. Nous avons montré dans cette thèse la forte dépendance de la fréquence du RO à la température et à la tension. Dans la plupart des utilisations, le FPGA n'est pas régulé en température ni en tension. Cependant, nous pouvons intégrer dans l'IP un modèle de compensation comme nous l'avons fait, qui se baserait sur une pré-caractérisation réalisée sur chacun des FPGAs pour lesquels nous proposerions l'IP. L'IP pourrait ainsi déclencher des alarmes indiquant le dépassement de certains seuils de dégradation (5%, 10%...) ou pourrait être lue lors d'une opération de maintenance par exemple ou tout simplement affichée en continu.

Publications

Articles

- Sobas, J., Airimitoiaie, T. B., & Marc, F. (2022). Development of a high accuracy and stability test bench for ageing measurement of 16 nm FinFETs based FPGA. *Microelectronics Reliability*, 138, 114698.
- Sobas, J., & Marc, F. (2023). Aging measurement and empirical modeling of BTI on FPGA using 16 nm FinFETs for static and dynamic stresses. *Microelectronics Reliability*, 115065.
- Travail en cours : Sobas, J., & Marc, F. Degradation measurement and modelling under ageing in a 16nm FinFET FPGA

Colloques

- Sobas, J., & Marc, F. Measurement and prediction of the reliability of latest generation digital components under ageing. Présenté au GDR SOC², Strasbourg, France, juin-2022.
- Sobas, J., & Marc, F. Modélisation du vieillissement dans un FPGA et comparaison des performances entre mesures et Vivado. Présenté au GDR SOC², Lyon, France, juin-2023.

Bibliographie

- [1] RapidWright. https://www.rapidwright.io/docs/Xilinx_Architecture.html. *RapidWright website*, 2023. xi, 52, 113, 114
- [2] Julien Coutet. Étude de la fiabilité et des mécanismes de dégradation dans les composants numériques de dernière génération. *thesis*, 2020-10-07. xv, 27, 53, 58, 68, 72, 98, 127, 129, 134, 156, 157, 159, 160, 161, 162, 163, 165
- [3] European Commission. Commission staff working document : A chips act for europe. <https://digital-strategy.ec.europa.eu/en/library/european-chips-act-communication-regulation-joint-undertaking-and-recommendation>, 1(147), 2022. 1
- [4] Apple. Communiqué de presse : Apple dévoile la puce m1 ultra, la plus puissante jamais intégrée à un ordinateur personnel. <https://www.apple.com/fr/newsroom/2022/03/apple-unveils-m1-ultra-the-worlds-most-powerful-chip-for-a-personal-computer/>, 2022. 1
- [5] Planet Scope. Informations publiées dans le monde sur le net. <https://www.planetoscope.com/Internet-/1523-informations-publiees-dans-le-monde-sur-le-net-en-gigaoctets-.html>, 2023. 1
- [6] <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022DC0045>. Communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions : A chips act for europe. *European Commission*, (45), 2022. 1
- [7] Eugene R Hnatek. *Integrated circuit quality and reliability*. CRC Press, 2018. 2
- [8] Gordon E. Moore. Cramming more components onto integrated circuits. *Electronics*, 38(8) :114, 1965. 5, 6
- [9] Gordon E. Moore. Progress in digital integrated electronics. *International Electron Devices Meeting*, 11(3) :11–13, 1975. 6
- [10] Gordon E. Moore. Lithography and the future of Moore's law. *Integrated Circuit Metrology, Inspection, and Process Control IX*, 2439 :2–17, 1995. 6
- [11] Hassan N. Khan, David A. Hounshell, and Erica R. H. Fuchs. Science and research policy at the end of moore's law. *Nature Electronics*, 1(1) :14–21, 2018-01-08. 6
- [12] Mart Graef. More than moore white paper. In *2021 IEEE International Roadmap for Devices and Systems Outbriefs*, pages 1–47. IEEE, 2021-11. 6, 7
- [13] SIA. Semiconductor industry association. www.semiconductors.org, 2022. 6, 7
- [14] IRDS. International roadmap for devices and systems. <https://irds.ieee.org/>, 2021. 6, 7

- [15] WSTS. World semiconductor trade statistics. <https://www.wsts.org/>, 2023. 7
- [16] Sah Chih-Tang. Evolution of the MOS transistor-from conception to vlsi. *Proceedings of the IEEE*, 76(10) :1280–1326, 1988. 10
- [17] R.H. Dennard, F.H. Gaensslen, Hwa-Nien Yu, V.L. Rideout, E. Bassous, and A.R. LeBlanc. Design of ion-implanted MOSFET's with very small physical dimensions. *IEEE Journal of Solid-State Circuits*, 9(5) :256–268, 1974. 10
- [18] Yuan Taur and Tak H. Ning. *Fundamentals of Modern VLSI Devices*. Cambridge University Press, 3 edition, 2021-12-02. 10
- [19] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson, and M. Bohr. A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon CMOS transistors. In *IEEE International Electron Devices Meeting 2003*, pages 11.6.1–11.6.3. IEEE, 2003. 10
- [20] K. Mistry, R. Chau, C.-H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, C. Allen, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, C. Auth, H. Liu, J. Maiz, B. McIntyre, P. Moon, J. Neiryck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, B. Beattie, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, J. Seiple, D. Simon, S. Sivakumar, P. Smith, D. Bergstrom, C. Thomas, T. Troeger, P. Vandervoorn, S. Williams, K. Zawadzki, M. Bost, M. Brazier, M. Buehler, and A. Cappellani. A 45nm logic technology with high-k+metal gate transistors, strained silicon, 9 cu interconnect layers, 193nm dry patterning, and 100% pb-free packaging. In *2007 IEEE International Electron Devices Meeting*, pages 247–250. IEEE, 2007-12. 10
- [21] C. Auth, C. Allen, A. Blattner, D. Bergstrom, M. Brazier, M. Bost, M. Buehler, V. Chikarmane, T. Ghani, T. Glassman, R. Grover, W. Han, D. Hanken, M. Hattendorf, P. Hentges, R. Heussner, J. Hicks, D. Ingerly, P. Jain, S. Jaloviar, R. James, D. Jones, J. Jopling, S. Joshi, C. Kenyon, H. Liu, R. McFadden, B. McIntyre, J. Neiryck, C. Parker, L. Pipes, I. Post, S. Pradhan, M. Prince, S. Ramey, T. Reynolds, J. Roesler, J. Sandford, J. Seiple, P. Smith, C. Thomas, D. Towner, T. Troeger, C. Weber, P. Yashar, K. Zawadzki, and K. Mistry. A 22nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors. In *2012 Symposium on VLSI Technology (VLSIT)*, pages 131–132. IEEE, 2012-06. 10, 30, 127, 128, 131
- [22] Sayeef Salahuddin, Kai Ni, and Suman Datta. The era of hyper-scaling in electronics. *Nature Electronics*, 1(8) :442–450, 2018-08-13. 10, 11
- [23] IRDS. International roadmap for devices and systems : Executive summary. <https://irds.ieee.org/>, 2022. 11
- [24] Henry H. Radamson, Huilong Zhu, Zhenhua Wu, Xiaobin He, Hongxiao Lin, Jinbiao Liu, Jinjuan Xiang, Zhenzhen Kong, Wenjuan Xiong, Junjie Li, Hushan Cui, Jianfeng Gao, Hong Yang, Yong Du, Buqing Xu, Ben Li, Xuwei Zhao, Jiahan Yu, Yan Dong, and Guilei Wang. State of the art and future perspectives in advanced CMOS technology. *Nanomaterials*, 10(8) :1555, 2020-08-07. 11
- [25] IRDS. International roadmap for devices and systems : Beyond CMOS and emerging materials integration. <https://irds.ieee.org/>, 2022. 11
- [26] E.J. Nowak, I. Aller, T. Ludwig, K. Kim, R.V. Joshi, Ching-Te Chuang, K. Bernstein, and R. Puri. Turning silicon on its edge [double gate CMOS/finfet technology]. *IEEE Circuits and Devices Magazine*, 20(1) :20–31, 2004. 11, 12

- [27] H.-S.P. Wong, D.J. Frank, and P.M. Solomon. Device design considerations for double-gate, ground-plane, and single-gated ultra-thin soi MOSFET's at the 25 nm channel length generation. In *International Electron Devices Meeting 1998. Technical Digest (Cat. No.98CH36217)*, pages 407–410, 1998. [11](#)
- [28] Shien-Yang Wu, C. Y. Lin, M. C. Chiang, J. J. Liaw, J. Y. Cheng, S. H. Yang, M. Liang, T. Miyashita, C. H. Tsai, B. C. Hsu, H. Y. Chen, T. Yamamoto, S. Y. Chang, V. S. Chang, C. H. Chang, J. H. Chen, H. F. Chen, K. C. Ting, Y. K. Wu, K. H. Pan, R. F. Tsui, C. H. Yao, P. R. Chang, H. M. Lien, T. L. Lee, H. M. Lee, W. Chang, T. Chang, R. Chen, M. Yeh, C. C. Chen, Y. H. Chiu, Y. H. Chen, H. C. Huang, Y. C. Lu, C. W. Chang, M. H. Tsai, C. C. Liu, K. S. Chen, C. C. Kuo, H. T. Lin, S. M. Jang, and Y. Ku. A 16nm finfet CMOS technology for mobile soc and computing applications. In *2013 IEEE International Electron Devices Meeting*, pages 9.1.1–9.1.4, 2013. [12](#), [30](#), [31](#), [127](#), [128](#), [131](#)
- [29] Syed Muhammad Yasser Sherazi, Chaitanya Jha, Dimitrios Rodopoulos, Peter Debacker, Bharani Chava, L Matti, MG Bardon, Pieter Schuddinck, Praveen Raghavan, V Gerousis, et al. Low track height standard cell design in in7 using scaling boosters. In *Design-Process-Technology Co-optimization for Manufacturability XI*, volume 10148, pages 279–286. SPIE, 2017. [12](#)
- [30] Mark T. Bohr. Logic technology scaling to continue moore's law. In *2018 IEEE 2nd Electron Devices Technology and Manufacturing Conference (EDTM)*, pages 1–3. IEEE, 2018-03. [12](#)
- [31] J.R Burns. Switching response of complementary-symmetry MOS transistor logic circuits. *RCA Review*, 25 :627, 1964. [13](#)
- [32] N. Hedenstierna and K.O. Jeppson. CMOS circuit speed and buffer optimization. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 6(2) :270–281, 1987-03. [14](#)
- [33] IRDS. International roadmap for devices and systems : More moore. <https://irds.ieee.org/>, 2022. [15](#)
- [34] Munehiro Tada, Naoya Inoue, and Yoshihiro Hayashi. Performance modeling of low-k Cu interconnects for 32-nm-node and beyond. *IEEE Transactions on Electron Devices*, 56(9) :1852–1861, 2009-09. [16](#)
- [35] Pawan Kapur and James P McVittie. Technology and reliability constrained future copper interconnects—part i : Resistance modeling. *IEEE TRANSACTIONS ON ELECTRON DEVICES*, 49(4), 2002.
- [36] Pawan Kapur, Gaurav Chandra, and James P McVittie. Technology and reliability constrained future copper interconnects—part II : Performance implications. *IEEE TRANSACTIONS ON ELECTRON DEVICES*, 49(4), 2002. [16](#)
- [37] Alfred Grill, Stephen M. Gates, Todd E. Ryan, Son V. Nguyen, and Deepika Priyadarshini. Progress in the development and understanding of advanced low k and ultralow k dielectrics for very large-scale integrated interconnects—state of the art. *Applied Physics Reviews*, 1(1) :011306, 2014-03. [16](#)
- [38] Alexander Barkalov, Larysa Titarenko, Alexander Barkalov, and Larysa Titarenko. Evolution of programmable logic. *Logic synthesis for FSM-based control units*, pages 53–75, 2009. [17](#)
- [39] Clive Maxfield. *The design warrior's guide to FPGAs : devices, tools and flows*. Elsevier, 2004. [17](#), [18](#)
- [40] Dave Galli, Mike Seither, and Peter Alfke. 15 years of innovation. *The quarterly journal for programmable logic users*, 1999. [18](#)

- [41] Xilinx. Zynq UltraScale+ MPSoC data sheet : Overview (DS891). *DataSheet*, 2021. [18](#)
- [42] Xilinx. UltraScale architecture configurable logic block (UG574). *DataSheet*, 2017. [19](#), [22](#)
- [43] Tao Pi and Patrick J. Crotty. FPGA lookup table with transmission gate structure for reliable low-voltage operation,6809552b1. *US patent*, 2004-10-26. [21](#), [22](#)
- [44] Steven P. Young, Venu M. Kondapalli, and Ramakrishna K. Tanikella. Six-input look-up table for use in a field programmable gate array,7061271b1. *US patent*, 2006-06-13. [22](#), [93](#), [112](#), [113](#), [117](#), [125](#), [127](#)
- [45] Manoj Chirania and Venu M. Kondapalli. Lookup table circuit optionally configurable as two or more smaller lookup tables with independent inputs,6998872b1. *US Patent*, 2006-02-14. [21](#), [22](#)
- [46] Steven P. Young, Kamal Chaudhary, and Trevor J. Bauer. Interconnect structure for a programmable logic device,6448808b2. *US patent*, 2002-09-10. [23](#)
- [47] Steven P. Young. Integrated circuit providing direct access to multi-directional interconnect lines in a general interconnect structure,7253658b1. *US patent*, 2007-08-07. [24](#)
- [48] Chin Hua Tan, Shankar Lakka, Ronald L. Cline, James B. Anderson, and Wayne E. Wennekamp. Integrated circuit with a selectable interconnect circuit for low power or high performance operation,7893712b1. *US patent*, 2011-02-22. [24](#)
- [49] C. Stroud, J. Nall, M. Lashinsky, and M. Abramovici. BIST-based diagnosis of FPGA interconnect. In *Proceedings. International Test Conference*, pages 618–627. IEEE, 2002. [24](#)
- [50] Aiwu Ruan, Junhao Yang, Li Wan, Bairui Jie, and Zhiqiang Tian. Insight into a generic interconnect resource model for xilinx virtex and spartan series FPGAs. *IEEE Transactions on Circuits and Systems II : Express Briefs*, 60(11) :801–805, 2013-11. [24](#)
- [51] Juan J. Rodriguez-Andina, Maria D. Valdes-Pena, and Maria J. Moure. Advanced features and industrial applications of FPGAs—a review. *IEEE Transactions on Industrial Informatics*, 11(4) :853–864, 2015-08. [24](#)
- [52] Vamsi Boppana, Ilya Ganusov, Vinod Kathail, and Vidya Rajagopalan. UltraScale+ MPSoC and FPGA families. *Xilinx*, 2015. [24](#), [25](#)
- [53] A. Goetzberger, A. D. Lopez, and R. J. Strain. On the formation of surface states during stress aging of thermalsi - sio2 interfaces. *Journal of The Electrochemical Society*, 120(1) :90, jan 1973. [26](#)
- [54] James H. Stathis, Souvik Mahapatra, and Tibor Grasser. Controversial issues in negative bias temperature instability. *Microelectronics Reliability*, 81 :244–251, 2018. [26](#), [29](#)
- [55] S. Mahapatra, N. Goel, S. Desai, S. Gupta, B. Jose, S. Mukhopadhyay, K. Joshi, A. Jain, A. E. Islam, and M. A. Alam. A comparative study of different physics-based NBTI models. *IEEE Transactions on Electron Devices*, 60(3) :901–916, 2013-03. [26](#)
- [56] S. Mahapatra, K. Ahmed, D. Varghese, A. E. Islam, G. Gupta, L. Madhav, D. Saha, and M. A. Alam. On the physical mechanism of NBTI in silicon oxynitride p-MOSFETs : Can differences in insulator processing conditions resolve the interface trap generation versus hole trapping controversy? In *2007 IEEE International Reliability Physics Symposium Proceedings. 45th Annual*, pages 1–9. IEEE, 2007-04. [26](#), [57](#)
- [57] S. Mahapatra, P. BharathKumar, and M.A. Alam. Investigation and modeling of interface and bulk trap generation during negative bias temperature instability of p-MOSFETs. *IEEE Transactions on Electron Devices*, 51(9) :1371–1379, 2004-09.

- [58] V. Huard, M. Denais, F. Perrier, N. Revil, C. Parthasarathy, A. Bravaix, and E. Vincent. A thorough investigation of MOSFETs NBTI degradation. *Microelectronics Reliability*, 45(1) :83–98, 2005-01. [26](#), [28](#), [57](#)
- [59] Chia Lin Chen, Y.M. Lin, C. Wang, and K. Wu. A new finding on NBTI lifetime model and an investigation on NBTI degradation characteristic for 1.2nm ultra thin oxide [MOSFETs]. In *2005 IEEE International Reliability Physics Symposium, 2005. Proceedings. 43rd Annual.*, pages 704–705, 2005. [27](#)
- [60] Jin Qin and Joseph B. Bernstein. Non-arrhenius temperature acceleration and stress-dependent voltage acceleration for semiconductor device involving multiple failure mechanisms. In *2006 IEEE International Integrated Reliability Workshop Final Report*, pages 93–97, 2006. [27](#)
- [61] Kaustubh Joshi, Subhadeep Mukhopadhyay, Nilesh Goel, and Souvik Mahapatra. A consistent physical framework for n and p BTI in HKMG MOSFETs. In *2012 IEEE International Reliability Physics Symposium (IRPS)*. IEEE, 2012-04. [27](#), [57](#)
- [62] D. Varghese, D. Saha, S. Mahapatra, K. Ahmed, F. Nouri, and M. Alam. On the dispersive versus arrhenius temperature activation of NBTI time evolution in plasma nitrated gate oxides : measurements, theory, and implications. In *IEEE International Electron Devices Meeting, 2005. IEDM Technical Digest.*, pages 684–687. IEEE, 2005. [27](#), [57](#)
- [63] S. Rangan, N. Mielke, and E.C.C. Yeh. Universal recovery behavior of negative bias temperature instability [PMOSFETs]. In *IEEE International Electron Devices Meeting 2003*, pages 14.3.1–14.3.4, 2003. [27](#)
- [64] Hans Reisinger, Ulrich Brunner, Wolfgang Heinrigs, Wolfgang Gustin, and Christian Schlunder. A comparison of fast methods for measuring NBTI degradation. *IEEE Transactions on Device and Materials Reliability*, 7(4) :531–539, 2007-12. [28](#)
- [65] Thomas Aichinger, Gregor Pobegen, and Michael Nelhiebel. Bias temperature instability for devices and circuits. *Springer New York*, 2014. [28](#), [29](#), [105](#), [106](#), [154](#)
- [66] Hans Reisinger, Tibor Grasser, Wolfgang Gustin, and Christian Schlünder. The statistical analysis of individual defects constituting NBTI and its implications for modeling dc- and ac-stress. In *2010 IEEE International Reliability Physics Symposium*, pages 7–15, 2010. [29](#), [106](#)
- [67] Tibor Grasser, Ben Kaczer, Wolfgang Goes, Hans Reisinger, Thomas Aichinger, Philipp Henberger, Paul-Jürgen Wagner, Franz Schanovsky, Jacopo Franco, María Toledano Toledano Luque, and Michael Nelhiebel. The paradigm shift in understanding the bias temperature instability : From reaction–diffusion to switching oxide traps. *IEEE Transactions on Electron Devices*, 58(11) :3652–3666, 2011. [29](#)
- [68] V. Huard, M. Denais, and C. Parthasarathy. NBTI degradation : From physical mechanisms to modelling. *Microelectronics Reliability*, 46(1) :1–23, 2006-01. [29](#)
- [69] Kjell O Jeppson and Christer M Svensson. Negative bias stress of MOS devices at high electric fields and degradation of mnos devices. *Journal of Applied Physics*, 48(5) :2004–2014, 1977. [29](#)
- [70] Dieter K Schroder and Jeff A Babcock. Negative bias temperature instability : Road to cross in deep submicron silicon semiconductor manufacturing. *Journal of applied Physics*, 94(1) :1–18, 2003.
- [71] Dieter K Schroder. Negative bias temperature instability : What do we understand? *Microelectronics Reliability*, 47(6) :841–852, 2007. [29](#)

- [72] J. H. Stathis and S. Zafar. The negative bias temperature instability in MOS devices : A review. *Microelectronics Reliability*, 46(2) :270–286, 2006. [29](#)
- [73] M. A. Alam and S. Mahapatra. A comprehensive model of PMOS NBTI degradation. *Microelectronics Reliability*, 45(1) :71–81, 2005-01-01. [29](#)
- [74] M.A. Alam, H. Kuflluoglu, D. Varghese, and S. Mahapatra. A comprehensive model for PMOS NBTI degradation : Recent progress. *Microelectronics Reliability*, 47(6) :853–862, 2007-06. [29](#), [86](#)
- [75] Subhadeep Mukhopadhyay, Nilesh Goel, and Souvik Mahapatra. A comparative study of NBTI and PBTI using different experimental techniques. *IEEE Transactions on Electron Devices*, 63(10) :4038–4045, 2016. [30](#)
- [76] Dawei Heh, Chadwin D. Young, and Gennadi Bersuker. Experimental evidence of the fast and slow charge trapping/detrapping processes in high- k dielectrics subjected to pbt stress. *IEEE Electron Device Letters*, 29(2) :180–182, 2008. [30](#)
- [77] S. Ramey, A. Ashutosh, C. Auth, J. Clifford, M. Hattendorf, J. Hicks, R. James, A. Rahman, V. Sharma, A. St. Amour, and C. Wiegand. Intrinsic transistor reliability improvements from 22nm tri-gate technology. In *2013 IEEE International Reliability Physics Symposium (IRPS)*, pages 4C.5.1–4C.5.5. IEEE, 2013-04. [30](#)
- [78] C.D. Young, K. Akarvardar, M.O. Baykan, K. Matthews, I. Ok, T. Ngai, K.-W. Ang, J. Pater, C.E. Smith, M.M. Hussain, P. Majhi, and C. Hobbs. (110) and (100) sidewall-oriented FinFETs : A performance and reliability investigation. *Solid-State Electronics*, 78 :2–10, 2012-12. [30](#), [31](#)
- [79] Kyong Taek Lee, Wonchang Kang, Eun-Ae Chung, Gunrae Kim, Hyewon Shim, Hyunwoo Lee, Hyejin Kim, Minhyeok Choe, Nae-In Lee, Anuj Patel, et al. Technology scaling on high-k & metal-gate finfet bti reliability. In *2013 IEEE International Reliability Physics Symposium (IRPS)*, pages 2D–1. IEEE, 2013. [31](#)
- [80] Sang-Yun Kim, Tai-su Park, Jae-Sung Lee, Donggun Park, Ki-Nam Kim, and Jong-Ho Lee. Negative bias temperature instability (NBTI) of bulk finfets. In *2005 IEEE International Reliability Physics Symposium, 2005. Proceedings. 43rd Annual.*, pages 538–540. IEEE, 2005. [31](#)
- [81] Mohamed Mounir Mahmoud and Norhayati Soin. A comparative study of lifetime reliability of planar MOSFET and FinFET due to BTI for the 16 nm CMOS technology node based on reaction-diffusion model. *Microelectronics Reliability*, 97 :53–65, 2019-06-01. Publisher : Elsevier Ltd. [30](#), [31](#)
- [82] R.B. Fair and R.C. Sun. Threshold-voltage instability in MOSFET's due to channel hot-hole emission. *IEEE Transactions on Electron Devices*, 28(1) :83–94, 1981. [31](#), [32](#), [57](#)
- [83] E Takeda and N Suzuki. An empirical model for device degradation due to hot-carrier injection. *IEEE electron device letters*, 4(4) :111–113, 1983. [31](#)
- [84] Yimin Wang, Yun Li, Yanbin Yang, and Wenchao Chen. Hot carrier injection reliability in nanoscale field effect transistors : Modeling and simulation methods. *Electronics*, 11(21) :3601, 2022-11-04. [31](#)
- [85] Souvik Mahapatra and Uma Sharma. A review of hot carrier degradation in n-channel MOSFETs—part i : Physical mechanism. *IEEE Transactions on Electron Devices*, 67(7) :2660–2671, 2020-07. [32](#), [33](#), [34](#)
- [86] L Brunet, X Garros, A Bravaix, A Subirats, F Andrieu, O Weber, P Scheiblin, M Rafik, E Vincent, and G Reimbold. Impact of backside interface on hot carriers degradation of thin film fd-soi nmosfets. In *2012 IEEE International Reliability Physics Symposium (IRPS)*, pages 3B–2. IEEE, 2012. [32](#)

- [87] F.-C. Hsu and K.-Y. Chiu. Temperature dependence of hot-electron-induced degradation in MOSFET's. *IEEE Electron Device Letters*, 5(5) :148–150, 1984-05. [32](#), [33](#), [57](#)
- [88] Yiming Qu, Xi Lin, Junkang Li, Ran Cheng, Xiao Yu, Zejie Zheng, Jiwu Lu, Bing Chen, and Yi Zhao. Ultra fast (< 1 ns) electrical characterization of self-heating effect and its impact on hot carrier injection in 14nm finfets. In *2017 IEEE International Electron Devices Meeting (IEDM)*, pages 39–2. IEEE, 2017. [32](#), [33](#), [57](#)
- [89] Alain Bravaix, Chloé Guérin, Vincent Huard, David Roy, Julien-Marc Roux, and Emmanuel Vincent. Hot-carrier acceleration factors for low power management in dc-ac stressed 40nm nmos node at high temperature. In *2009 IEEE International Reliability Physics Symposium*, pages 531–548. IEEE, 2009. [33](#)
- [90] Tibor Grasser, editor. *Hot Carrier Degradation in Semiconductor Devices*. Springer International Publishing, 2015. [33](#), [34](#)
- [91] K. Mistry and B. Doyle. The role of electron trap creation in enhanced hot-carrier degradation during ac stress (n-channel MOSFET). *IEEE Electron Device Letters*, 11(6) :267–269, 1990. [33](#)
- [92] Alain Bravaix, M Saliva, Florian Cacho, X Federspiel, Cheikh Ndiaye, Souhir Mhira, Edith Kussener, E Pauly, and Vincent Huard. Hot-carrier and bti damage distinction for high performance digital application in 28nm fdsoi and 28nm lp CMOS nodes. In *2016 IEEE 22nd International Symposium on On-Line Testing and Robust System Design (IOLTS)*, pages 43–46. IEEE, 2016. [33](#)
- [93] Mohammad Naouss and François Marc. FPGA LUT delay degradation due to HCI : Experiment and simulation results. *Microelectronics Reliability*, 64 :31–35, 2016. [33](#)
- [94] Chenming Hu, Simon C. Tam, Fu-Chieh Hsu, Ping-Keung Ko, Tung-Yi Chan, and K.W. Terrill. Hot-electron-induced MOSFET degradation - model,monitor,and improvement. *IEEE Journal of Solid-State Circuits*, 20(1) :295–305, 1985. [33](#)
- [95] S. Mahapatra, D. Saha, D. Varghese, and P.B. Kumar. On the generation and recovery of interface traps in MOSFETs subjected to NBTI, FN, and HCI stress. *IEEE Transactions on Electron Devices*, 53(7) :1583–1592, 2006. [33](#)
- [96] Hai Jiang, Xiaoyan Liu, Nuo Xu, Yandong He, Gang Du, and Xing Zhang. Investigation of self-heating effect on hot carrier degradation in multiple-fin SOI FinFETs. *IEEE Electron Device Letters*, 36(12) :1258–1260, 2015-12. [34](#), [35](#)
- [97] Wen-Teng Chang, Li-Gong Cin, and Wen-Kuan Yeh. Impact of fin width and back bias under hot carrier injection on double-gate FinFETs. *IEEE Transactions on Device and Materials Reliability*, 15(1) :86–89, 2015-03. [35](#)
- [98] S. Chabukswar, D. Maji, C.R. Manoj, K.G. Anil, V. Ramgopal Rao, F. Crupi, P. Magnone, G. Giusi, C. Pace, and N. Collaert. Implications of fin width scaling on variability and reliability of high-k metal gate FinFETs. *Microelectronic Engineering*, 87(10) :1963–1967, 2010-10. [35](#)
- [99] Hai Jiang, SangHoon Shin, Xiaoyan Liu, Xing Zhang, and Muhammad Ashraful Alam. Characterization of self-heating leads to universal scaling of HCI degradation of multi-fin SOI FinFETs. In *2016 IEEE International Reliability Physics Symposium (IRPS)*, pages 2A–3–1–2A–3–7. IEEE, 2016-04. [35](#)
- [100] K.F. Schuegraf and C. Hu. Hole injection oxide breakdown model for very low voltage lifetime extrapolation. In *31st Annual Proceedings Reliability Physics 1993*, pages 7–12. IEEE, 1993. [35](#)

- [101] J. W. McPherson and H. C. Mogul. Underlying physics of the thermochemical model in describing low-field time-dependent dielectric breakdown in SiO₂ thin films. *Journal of Applied Physics*, 84(3) :1513–1523, 1998-08. [35](#)
- [102] D. J. DiMaria and E. Cartier. Mechanism for stress-induced leakage currents in thin silicon dioxide films. *Journal of Applied Physics*, 78(6) :3883–3894, 1995-09-15. [35](#)
- [103] John Keane, Xiaofei Wang, Devin Persaud, and Chris H. Kim. An all-in-one silicon odometer for separately monitoring HCI,BTI,and TDDB. *IEEE Journal of Solid-State Circuits*, 45(4) :817–829, 2010-04. [36](#)
- [104] Mihir Choudhury, Vikas Chandra, Kartik Mohanram, and Robert Aitken. Analytical model for TDDB-based performance degradation in combinational logic. In *2010 Design, Automation & Test in Europe Conference & Exhibition (DATE 2010)*, pages 423–428. IEEE, 2010-03. [36](#)
- [105] K. Joshi, S. W. Chang, D. S. Huang, P. J. Liao, and Y.-H. Lee. Study of dynamic TDDB in scaled FinFET technologies. In *2018 IEEE International Reliability Physics Symposium (IRPS)*, pages P-GD.5–1–P-GD.5–5. IEEE, 2018-03. [36](#)
- [106] Hyunjin Kim, Minjung Jin, Hyunchul Sagong, Jinju Kim, Ukjin Jung, Minhyuck Choi, Junekyun Park, Sangchul Shin, and Sangwoo Pae. A systematic study of gate dielectric TDDB in FinFET technology. In *2018 IEEE International Reliability Physics Symposium (IRPS)*, pages 4A.4–1–4A.4–4. IEEE, 2018-03. [36](#)
- [107] Z. Chbili and A. Kerber. Self-heating impact on TDDB in bulk FinFET devices : Uniform vs non-uniform stress. In *2016 IEEE International Integrated Reliability Workshop (IIRW)*, pages 45–48. IEEE, 2016. [36](#)
- [108] Minjung Jin, Changze Liu, Jinju Kim, Jungin Kim, Hyewon Shim, Kangjung Kim, Gunrae Kim, Soonyoung Lee, Taiki Uemura, Man Chang, Taehyun An, Junekyun Park, and Sangwoo Pae. Reliability characterization of 10nm FinFET technology with multi-VT gate stack for low power and high performance. In *2016 IEEE International Electron Devices Meeting (IEDM)*, pages 15.1.1–15.1.4. IEEE, 2016-12. [36](#)
- [109] Adamantios Mettas. Modeling and analysis for multiple stress-type accelerated life data. In *Annual Reliability and Maintainability Symposium. 2000 Proceedings. International Symposium on Product Quality and Integrity (Cat. No. 00CH37055)*, pages 138–143. IEEE, 2000. [42](#)
- [110] Philippe Pougnet, Franck Bayle, Hichame Maanane, and Pierre Richard Dahoo. Reliability prediction of embedded electronic systems : the FIDES guide. *Embedded Mechatronic Systems*, pages 189–216, 2019. [42](#)
- [111] Justin SJ Wong, Pete Sedcole, and Peter YK Cheung. A transition probability based delay measurement method for arbitrary circuits on FPGAs. In *2008 International Conference on Field-Programmable Technology*, pages 105–112. IEEE, 2008. [43](#), [44](#), [45](#)
- [112] Pete Sedcole and Peter Y. K. Cheung. Within-die delay variability in 90nm FPGAs and beyond. In *2006 IEEE International Conference on Field Programmable Technology*, pages 97–104, 2006. [45](#), [93](#)
- [113] Aghiles Douadi, Giorgio Di Natale, Paolo Maistri, Elena-Ioana Vatajelu, and Vincent Beroulle. A study of high temperature effects on ring oscillator based physical unclonable functions. In *2023 IEEE 29th International Symposium on On-Line Testing and Robust System Design (IOLTS)*, pages 1–7. IEEE, 2023. [45](#)

- [114] Edward A Stott, Justin SJ Wong, Pete Sedcole, and Peter YK Cheung. Degradation in FPGAs : measurement and modelling. In *Proceedings of the 18th annual ACM/SIGDA international symposium on Field programmable gate arrays*, pages 229–238, 2010. [45](#), [48](#), [52](#), [53](#), [54](#), [68](#)
- [115] Florent Bruguier, Pascal Benoit, Philippe Maurine, and Lionel Torres. A new process characterization method for fpgas based on electromagnetic analysis. In *2011 21st International Conference on Field Programmable Logic and Applications*, pages 20–23. IEEE, 2011. [45](#), [46](#)
- [116] Petr Pfeifer and Zdenek Pliva. On measurement of impact of the metallization and FPGA design to the changes of slice parameters and generation of delay faults. In *22nd International Conference on Field Programmable Logic and Applications (FPL)*, pages 743–746. IEEE, 2012. [47](#)
- [117] Mohammad Naouss and François Marc. Design and implementation of a low cost test bench to assess the reliability of FPGA. *Microelectronics Reliability*, 55(9-10) :1341–1345, 2015. [45](#), [46](#), [51](#), [52](#), [53](#), [54](#), [63](#), [68](#)
- [118] Petr Pfeifer, Ben Kaczer, and Zdenek Pliva. A reliability lab-on-chip using programmable arrays. In *2014 IEEE International Reliability Physics Symposium*, pages CA.6.1–CA.6.8, 2014. [45](#), [47](#)
- [119] Petr Pfeifer and Zdenek Pliva. A new method for in situ measurement of parameters and degradation processes in modern nanoscale programmable devices. *Microprocessors and Microsystems*, 38(6) :605–619, 2014. [47](#), [48](#), [49](#), [53](#), [54](#)
- [120] Abdulazim Amouri, Florent Bruguier, Saman Kiamehr, Pascal Benoit, Lionel Torres, and Mehdi Tahoori. Aging effects in FPGAs : An experimental analysis. In *2014 24th international conference on Field Programmable Logic and Applications (FPL)*, pages 1–4. IEEE, 2014. [48](#), [49](#), [53](#), [54](#)
- [121] Julien Coutet, François Marc, and Jean-Claude Clément. Investigation of CMOS reliability in 28 nm through BTI and HCI extraction. *Microelectronics Reliability*, 146 :115007, 2023. [49](#), [50](#), [51](#), [53](#)
- [122] Joseph B Bernstein, Alain Bensoussan, and Emmanuel Bender. Reliability prediction with MTOL. *Microelectronics Reliability*, 68 :91–97, 2017. [50](#), [51](#), [53](#), [54](#)
- [123] E Bender, JB Bernstein, and A Bensoussan. Reliability prediction of finfet FPGAs by MTOL. *Microelectronics Reliability*, 114 :113809, 2020. [50](#), [51](#), [53](#), [54](#), [68](#)
- [124] Mohammad Naouss and François Marc. Modelling delay degradation due to NBTI in FPGA look-up tables. In *2016 26th International Conference on Field Programmable Logic and Applications (FPL)*, pages 1–4. IEEE, 2016. [52](#), [53](#), [58](#)
- [125] Mohammad Naouss and François Marc. FPGA LUT delay degradation due to HCI : Experiment and simulation results. *Microelectronics Reliability*, 64 :31–35, 2016. [52](#), [53](#), [58](#)
- [126] Edward Stott, Justin SJ Wong, and Peter YK Cheung. Degradation analysis and mitigation in FPGAs. In *2010 International Conference on Field Programmable Logic and Applications*, pages 428–433. IEEE, 2010. [52](#), [53](#)
- [127] Tanner Gaskin, Hayden Cook, Wesley Stirk, Robert Lucas, Jeffrey Goeders, and Brad Hutchings. Using novel configuration techniques for accelerated FPGA aging. In *2020 30th International Conference on Field-Programmable Logic and Applications (FPL)*, pages 169–175. IEEE, 2020. [52](#), [53](#), [54](#)
- [128] Mariem Slimani, Karim Benkalaia, and L Naviner. Analysis of ageing effects on artix7 xilinx FPGA. *Microelectronics Reliability*, 76 :168–173, 2017. [53](#), [54](#)

- [129] J Sobas, T-B Airimitoiaie, and F Marc. Development of a high accuracy and stability test bench for ageing measurement of 16 nm finfets based FPGA. *Microelectronics Reliability*, 138 :114698, 2022. 53
- [130] Mohammad Naouss. Conception et exploitation d'un banc d'auto-caractérisation pour la prévision de la fiabilité des circuits numériques programmables. *thesis*, 201610-07. 58, 60
- [131] Avnet. Datasheet ultra96-v2. *DataSheet*, 2020. 58
- [132] Xilinx. DS891, zynq UltraScale+ MPSoC data sheet : Overview. *DataSheet*, 2021. 58
- [133] Renesas. Versaclock 6e programmable clock with internal crystal (5p49v6975). *DataSheet*, 2023. 68
- [134] Digilent. Datasheet pmod gps. *DataSheet*, 2016. 68
- [135] Xilinx. Datasheet ultrascale architecture system monitor (ug580). *DataSheet*, 2021. 73
- [136] Xilinx. System management wizard v1.3, logicore ip product guide (pg185). *DataSheet*, 2019. 73
- [137] Sébastien Menezier and Eric Collard. Les modules peltier : A - théorie et applications. *Union des professeurs de physique et de chimie*, 995 :741–758, 06 2017. 74
- [138] BJ Huang and CL Duang. System dynamic model and temperature control of a thermoelectric cooler. *International journal of refrigeration*, 23(3) :197–207, 2000. 75, 77, 78
- [139] Adaptive Thermal Management. Peltier cooler module aph-071-10-08-s. *DataSheet*, 2021. 77
- [140] R. Palacios, A. Arenas, R.R. Pecharromán, and E.L. Pagola. Analytical procedure to obtain internal parameters from performance curves of commercial thermoelectric modules. *Applied Thermal Engineering*, 29(17) :3501–3505, 2009. 77
- [141] I.D. Landau. The r-s-t digital controller design and applications. *Control Engineering Practice*, 6(2) :155–165, 1998. 78
- [142] I.D. Landau TB Airimitoiaie. Mircod software for identification, robust control design, and controller order reduction. <http://tudor-bogdan.airimitoiaie.name/mircod.html>, 2020. 78
- [143] M.J. Nagy and S.J. Roman. The effect of pulse width modulation (pwm) frequency on the reliability of thermoelectric modules. In *Eighteenth International Conference on Thermoelectrics. Proceedings, ICT'99 (Cat. No.99TH8407)*, pages 123–125, 1999. 78
- [144] Infineon. Irps5401 pmic. *DataSheet*, 2020. 80
- [145] Avnet. Ultra96 sbc v2 (sch-us1sbc). *DataSheet*, 2020. 80
- [146] Infineon. Irps5401 command set. *DataSheet*, 2020. 80
- [147] Leo Liu and Nachiket Kapre. Timing-aware routing in the rapidwright framework. In *2019 29th International Conference on Field Programmable Logic and Applications (FPL)*, pages 24–30, 2019. 93
- [148] Pongstorn Maidee, Chris Neely, Alireza Kaviani, and Chris Lavin. An open-source lightweight timing model for rapidwright. In *2019 International Conference on Field-Programmable Technology (ICFPT)*, pages 171–178, 2019.

- [149] Gojman Benjamin, Nalmela Sirisha, Mehta Nikil, Howarth Nicholas, and Dehon Andre. Grok-lab : Generating real on-chip knowledge for intra-cluster delays using timing extraction. *ACM Trans. Reconfigurable Technol. Syst.*, 7(4), dec 2015.
- [150] Gojman Benjamin and DeHon André. Grok-int : Generating real on-chip knowledge for interconnect delays using timing extraction. In *2014 IEEE 22nd Annual International Symposium on Field-Programmable Custom Computing Machines*, pages 88–95, 2014. [93](#)
- [151] Tim Tuan, Austin Lesea, Chris Kingsley, and Steve Trimmer. Analysis of within-die process variation in 65nm FPGAs. In *2011 12th International Symposium on Quality Electronic Design*, pages 1–5, 2011. [93](#)
- [152] Konstantinos Maragos, Endri Taka, George Lentaris, Ioannis Stratakos, and Dimitrios Soudris. Analysis of performance variation in 16nm finfet FPGA devices. In *2019 29th International Conference on Field Programmable Logic and Applications (FPL)*, pages 38–44, 2019. [93](#), [96](#)
- [153] Etienne Bergeron, Marc Feeley, Marc-Andre Daigneault, and Jean Pierre David. Using dynamic reconfiguration to implement high-resolution programmable delays on an FPGA. In *2008 Joint 6th International IEEE Northeast Workshop on Circuits and Systems and TAISA Conference*, pages 265–268, 2008. [93](#)
- [154] Marc-Andre Daigneault and Jean Pierre David. A high-resolution time-to-digital converter on FPGA using dynamic reconfiguration. *IEEE Transactions on Instrumentation and Measurement*, 60(6) :2070–2079, 2011.
- [155] Safa Berrima, Yves Blaqui re, and Yvon Savaria. Sub-ps resolution programmable delays implemented in a xilinx FPGA. In *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pages 918–921, 2017. [93](#)
- [156] Edward .A Stott, Justin Wong, and Peter Cheung. Degradation in FPGAs :measurement and modelling. *ACM/SIGDA international symposium on Field programmable gate array*, pages 229–238, 2010. [108](#)
- [157] Xilinx. Vivado design suite tcl command reference guide (ug835). *User Guide Vivado*, 2018. [114](#)
- [158] Leo Liu and Nachiket Kapre. Timing-aware routing in the rapidwright framework. In *2019 29th International Conference on Field Programmable Logic and Applications (FPL)*, pages 24–30, 2019. [116](#)
- [159] Kjell O. Jeppson and Christer M. Svensson. Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices. *Journal of Applied Physics*, 48(5) :2004–2014, 1977. Publisher : American Institute of Physics. [138](#)
- [160] Stephen Ramey, Chetan Prasad, Marty Agostinelli, Sangwoo Pae, Steven Walstra, Satrajit Gupta, and Jeffrey Hicks. Frequency and recovery effects in high-k bti degradation. In *2009 IEEE International Reliability Physics Symposium*, pages 1023–1027, 2009. [149](#)
- [161] S. Ramey, A. Ashutosh, C. Auth, J. Clifford, M. Hattendorf, J. Hicks, R. James, A. Rahman, V. Sharma, A. St Amour, and C. Wiegand. Intrinsic transistor reliability improvements from 22nm tri-gate technology. In *2013 IEEE International Reliability Physics Symposium (IRPS)*, pages 4C.5.1–4C.5.5, 2013. [149](#), [150](#), [151](#)
- [162] Julien Coutet, Fran ois Marc, and Jean-Claude Cl ment. Investigation of CMOS reliability in 28nm through BTI and HCI extraction. *Microelectronics Reliability*, 146 :115007, 2023. [154](#), [155](#)

Titre : Mesure et prévision de la fiabilité des composants numériques de dernière génération sous l'effet du vieillissement

Résumé (FR) : Dans les années 2010, l'évolution de l'architecture d'un transistor planaire (MOSFET) a repoussé les limites de dimensionnement et a donné naissance au FinFET. Le FinFET, ce transistor avec un canal en forme d'aileron sortant du plan, est largement employé dans les circuits numériques de dernière génération avec des noeuds allant jusqu'à cinq nanomètres. Aujourd'hui, le FinFET a fait ses preuves en terme de performance, de faible consommation et de coût, mais qu'en est-il de sa fiabilité? L'évolution architecturale du FinFET a-t-elle ouvert les portes à l'amélioration des circuits numériques au prix d'une moins bonne fiabilité?

Dans cette thèse, nous vous proposons de découvrir le banc de test que nous avons développé visant à vieillir et mesurer les dégradations dans des FPGAs conçus avec des FinFET 16 nanomètres. En implémentant des oscillateurs en anneau dont nous mesurons l'évolution de la fréquence d'oscillation, qui dépend de la tension de seuil des transistors, nous observons l'effet des mécanismes de défaillances (BTI, HCI et TDDB) à l'échelle du transistor. Notre banc de test semble se classer dans l'état de l'art comme étant le plus précis pour mesurer des dégradations, lui offrant ainsi, la faculté de mesurer des dégradations pour des conditions de vieillissement proches des conditions d'utilisation rendant ainsi nos observations plus réalistes.

Après 8000 heures de vieillissement, nous avons observé les dégradations mesurées dans 5103 oscillateurs en anneau répartis dans neuf FPGAs. Nous présentons une méthode pour dissocier les dégradations dans les ressources logiques et dans les ressources de routage. Nous modélisons les dégradations et les défaillances en fonction de la température, de la tension, du rapport cyclique de stress et des ressources utilisées dans le FPGA. Finalement, nous comparons nos résultats de dégradations mesurées sur un FPGA 16 nanomètres FinFET avec un FPGA 28 nanomètres MOSFET afin de révéler l'évolution de la fiabilité en passant du MOSFET au FinFET dans les FPGAs.

Mots Clefs : BTI; Contraintes de temps; Fiabilité; FPGA; FinFET; HCI; Oscillateur en anneau; Récupération; Vivado

Title : Measure and prediction of latest-generation digital circuits' reliability under ageing effect

Abstract (EN) : In the 2010s, evolution of planar transistor (MOSFET) architecture has led to the extension of design limits, resulting in the FinFET. The FinFET, a transistor with a fin-shaped channel rising out of its surface, is widely used in the latest generation of digital circuits with nodes down to five nanometers. So far, FinFET has proved its high performance, low power consumption and low cost, but what about its reliability? Has the architectural evolution of FinFET allowed digital circuits to be improved at the cost of reduced reliability?

In this thesis, we present the test bench we have developed to age and measure degradation in FPGAs based on 16-nanometre FinFETs. By implementing ring oscillators and measuring the evolution of the oscillation frequency, which depends on the threshold voltage of the transistors, we observe the effect of the failure mechanisms (BTI, HCI and TDDB) at the transistor scale. Our test bench seems to be the most accurate in the state of the art for measuring degradation, giving it the capacity to measure degradation for ageing conditions close to operating conditions, making our observations more lifelike.

After 8000 hours of ageing, we have observed the degradations measured in 5103 ring oscillators distributed into nine FPGAs. We present a method to distinguish degradations in logic resources from degradations in routing resources. We model degradation and failure as a function of temperature, voltage, stress duty cycle and the resources used in the FPGA. Finally, we compare our degradation measured in a 16 nm FinFET FPGA with a 28 nm MOSFET FPGA in order to show the evolution of reliability between MOSFET and FinFET in FPGAs.

Keywords : BTI; FinFET; FPGA; HCI; Recovery; Reliability; Ring Oscillator; Timing; Vivado
