



HAL
open science

Gestion de la consommation basée sur l'adaptation dynamique de la tension, fréquence et body bias sur les systèmes sur puce en technologie FD-SOI

Yeter Akgul

► **To cite this version:**

Yeter Akgul. Gestion de la consommation basée sur l'adaptation dynamique de la tension, fréquence et body bias sur les systèmes sur puce en technologie FD-SOI. Micro et nanotechnologies/Microélectronique. Université Montpellier II - Sciences et Techniques du Languedoc, 2014. Français. NNT : 2014MON20132 . tel-01654475

HAL Id: tel-01654475

<https://theses.hal.science/tel-01654475>

Submitted on 4 Dec 2017

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

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

THÈSE

Pour obtenir le grade de
Docteur

Délivré par l'Université de Montpellier II

Préparée au sein de l'école doctorale I2S
Et de l'unité de recherche LIRMM

Spécialité: **Systems Automatiques et Microélectroniques**

Présentée par **Yeter AKGUL**

**Gestion de la consommation
basée sur l'adaptation
dynamique de la tension,
fréquence et body bias sur les
systèmes sur puce en
technologie FD-SOI**

Soutenue le 09 Décembre 2014 devant le jury composé de

M. François PÊCHEUX	Pr	LIP6, UPMC	Rapporteur
M. Jean-Philippe DIGUET	DR CNRS	LAB-STICC, UBS	Rapporteur
M. Frédéric PÉTROT	Pr	TIMA, INPG	Examinateur
M. Gilles SASSATELLI	DR CNRS	LIRMM	Examinateur
Mme Suzanne LESECQ	Ing., HDR	CEA-LETI	Examinateur
M. Sylvain CLERC	Ing.	STMicroelectronics	Invité
M. Lionel TORRES	Pr	LIRMM, UM2	Directeur de thèse
M. Pascal BENOIT	MCf	LIRMM, UM2	Examinateur
M. Diego PUSCHINI	Ing., Dr.	CEA-LETI	Examinateur

Remerciements

Cette thèse a été effectuée en collaboration entre le Département Architectures Conception Logiciels Embarqués (DACLE) du CEA-Leti, Grenoble, France ; et le Département de Microélectronique du Laboratoire d'Informatique, Robotique et Microélectronique de Montpellier (LIRMM), France. Le laboratoire d'accueil du CEA est le Laboratoire Infrastructure et Atelier Logiciel pour Puces (LIALP). Ces travaux ont été réalisés en collaboration avec le Laboratoire Intégration Silicium et Architecture Numérique (LISAN).

Je tiens tout d'abord à témoigner ma reconnaissance aux membres du Jury d'avoir accepté d'évaluer mes travaux de thèse. Je remercie les rapporteurs, M. François Pêcheux, professeur à l'Université Pierre et Marie Curie ; et M. Jean-Philippe Diguët, directeur de recherche CNRS à l'Université Bretagne-Sud, pour leurs analyses de mon manuscrit et leurs retours. Merci au Président du Jury de thèse, M. Frédéric Pétrot, professeur à Grenoble INP. J'exprime également toute ma gratitude aux examinateurs, M. Gilles Sassatelli, directeur de recherche CNRS au Laboratoire d'Informatique, Robotique et Microélectronique de Montpellier (LIRMM) ; et Mme Suzanne Lesecq, ingénieur au CEA-Leti ; et à M. Sylvain Clerc, ingénieur à STMicroelectronics Crolles, membre invité du Jury. Merci pour tout l'intérêt porté à mes travaux.

Je tiens à remercier M. Thierry Colette, chef du département DACLE, M. Marc Belleville, directeur scientifique du DACLE, et M. Vincent Olive, chef du LIALP pour leur soutien.

Je remercie mon directeur de thèse M. Lionel Torres et mon encadrant universitaire M. Pascal Benoit pour leur disponibilité et leurs remarques pertinentes qui ont permis de faire avancer mes travaux de thèse. Je tiens à remercier également, tout particulièrement, mes encadrants CEA, Mme Suzanne Lesecq (encadrante de stage CEA) et M. Diego Puschini (encadrant de thèse CEA), pour leur disponibilité, leur patience, leur expertise et leur amitié. Merci de votre implication.

Je remercie également tous les membres du laboratoire LISAN avec qui j'ai pu travailler et qui m'ont aidé pendant ces quatre années, particulièrement Edith, Ivan M.P., Olivier B., Yvain, Antoine, Maxime D., Pascal et David. Un grand

merci également à Gaël P. de son aide précieuse sur les convertisseurs DC-DC. Je remercie Julien, Jean-Pierre, Warody, Mathieu B., Mathieu G., François, Anca, Maxime L., Wijdene, Caroline, Rémi et Isabel pour les discussions chemises haïïennes, ski, escalade ou gourmandises. Je remercie aussi les " thésards " du laboratoire, Safietou, Molka, Victor, Ozan, Mai, Nicolas, Fernando, Tiana, Ivan L. et Olesia pour leurs amitiés. J'ai passé de beaux moments à vos côtés.

Un grand merci à Juliana, je suis désolée d'avoir beaucoup fait travailler ton mari. Je remercie Betul, Jun-Young, Paul-Antoine, Carlos, Yacine avec qui on a passé de bonnes soirées. Merci aussi à mes amis de Montpellier, Cécile, Leticia, Bruno, Antho, Alexandre, Laurent, Damien, Granola, Bubu et Jérem.

Pour finir, je remercie tous mes amis et ma famille qui m'ont soutenu. Une pensée toute particulière va à Hava, Ibrahim, Mehpare et Sophie. Un grand merci à mes parents, mon frère et mes sœurs pour leur soutien. Enfin, je remercie affectueusement Lionel de sa patience, de sa sérénité, de ses remarques pertinentes sur mes travaux et de son soutien absolu.

À mes parents, à ma famille.

Table des matières

Remerciements	i
Table des matières	iii
1 Introduction	1
2 Évolution de la gestion de la consommation	7
2.1 Solution au niveau technologique	8
2.1.1 Technologie CMOS BULK	8
2.1.2 FinFET	10
2.1.3 SOI	10
2.2 Gestion de la consommation au niveau circuit et système	12
2.2.1 Modèles de consommation de circuits synchrones	13
2.2.2 Gestion en mode standby	13
2.2.3 Gestion en cours d'exécution	16
2.2.4 Synthèse	22
2.3 Actionneurs pour agir sur F , V_{dd} , V_{bb}	22
2.3.1 Génération d'horloge variable intégrée	22
2.3.2 Tension d'alimentation variable	25
2.3.3 Tensions d'alimentation discrètes	30
2.3.4 Tension de body bias	31
2.3.5 Conclusion	33
2.4 Synthèse	33
3 Placement optimal de Power Mode	37
3.1 Gestion de la consommation : DVFS	38
3.2 Hypothèses de travail et définitions	39
3.3 Gestion d'un circuit possédant trois PM	42
3.3.1 Contexte	43
3.3.2 Gain en consommation lié à l'ajout d'un PM intermédiaire	45

3.3.3	Gain en consommation moyen	50
3.3.4	Méthode de placement du PM intermédiaire	53
3.3.5	Synthèse	64
3.4	Gestion d'un système à N PM	65
3.4.1	Synthèse	66
3.5	Conclusion	66
4	Adaptation dynamique de (F, V_{dd}, V_{bb})	69
4.1	Synthèse de la gestion d'un circuit possédant 3 DoF	71
4.2	Hypothèses de travail et définitions	72
4.3	Gestion d'un circuit possédant trois DoF discrets	75
4.3.1	Propriété d'une fonction discrètement convexe	77
4.3.2	Méthode proposée pour gérer le compromis consommation-performance	80
4.3.3	Conclusion	84
4.4	Gestion d'un circuit possédant trois DoF dont un discret	85
4.4.1	Phase d'identification	90
4.4.2	Identification du sous-ensemble croissant, convexe par morceaux	95
4.4.3	Exécution des tâches	98
4.4.4	Politique de gestion du hopping	103
4.5	Conclusion	105
5	Application et résultats	107
5.1	Application à un circuit possédant un nombre fini de PM	108
5.1.1	Cas d'étude : simulation d'un oscillateur en anneau	108
5.1.2	Cas d'étude : mesures d'un circuit	116
5.1.3	Conclusion	121
5.2	Application à un circuit composé d'un nombre infini de PM	121
5.2.1	Identification du PWCS	122
5.2.2	Évaluation du gain en consommation lors de l'exécution d'une tâche	126
5.2.3	Optimisation du hopping : politique de gestion du V_{bb} hopping	133
5.2.4	Prise en compte du coût de commutation des actionneurs .	141
5.3	Variations de procédés de fabrication	144
5.4	Conclusion	146
6	Bilan et perspectives	151
	Bibliographie	159
	Bibliographie Personnelle	167
	Publications internationales	167

Brevet	167
Table des figures	169
Liste des tableaux	175
Liste des acronymes	177

Chapitre 1

Introduction

Contexte et motivations

Depuis la création du premier transistor, la taille des transistors n'a cessé de diminuer. En 1965, Gordon Moore, l'un des trois fondateurs d'INTEL, a proposé un modèle économique qui prévoyait de doubler tous les ans la complexité des circuits intégrés. En 1975, Gordon Moore a réajusté sa proposition et a affirmé que « le nombre de composants, et donc la puissance des microprocesseurs doublera désormais tous les deux ans ». Cette affirmation, connue sous le nom de loi de Moore et basée sur des expérimentations dans le domaine de la micro-électronique, a conditionné l'évolution des circuits intégrés ces dernières décennies. À titre de comparaison, si la loi de Moore était appliquée au secteur de l'automobile, les voitures pourraient atteindre une vitesse maximale à peine cent fois moins rapide que la vitesse de la lumière.

Grâce à cette course à la miniaturisation, la taille et le coût de production des circuits intégrés ont diminué conjointement. Cela permet la mise en place de circuits à moindre coût dans la vie de tous les jours. Cependant, les limites physiques de la course à la miniaturisation sont atteintes en technologie CMOS BULK. En effet, cette limite, attendue par les technologues, est observable en technologie CMOS BULK 28 nm [1]. Cette limite est due à la consommation électrique des transistors qui ne peut plus être diminuée en réduisant leur taille. Ainsi la performance des circuits intégrés se trouve aujourd'hui limitée par la consommation énergétique.

Pour remédier aux problèmes de miniaturisation et de consommation, de nouvelles pistes ont été explorées. En 1992, le Silicium sur Isolant (SOI) a été créé. En 1999, STMicroelectronics et le CEA-Leti ont fabriqué la technologie SON (Silicon on Nothing) pour démontrer l'intérêt des films minces. Par la suite, la technologie FD-SOI (Fully Depleted SOI) à BOX mince (ou Ultra Thin Body and Box FD-

SOI) a été créée. Cette dernière est disponible pour la production depuis 2012. Une autre technologie est le FinFET (Fin-shaped Field Effect Transistor) qui a été développée par INTEL. Ces technologies permettent de poursuivre la miniaturisation, en consommant moins que la technologie CMOS BULK à performance et nœud technologique identiques.

Besoin en performance et gestion de la consommation

Le besoin en performance des circuits intégrés était satisfait par l'augmentation du nombre de transistors par cœur en premier lieu. Cependant, la performance n'était plus suffisante. Dans un second temps, le nombre de cœurs par circuits intégrés a augmenté afin de satisfaire le besoin en performance. La complexité des circuits a augmenté et la consommation aussi. La Figure 1.1 illustre l'évolution du nombre de cœurs de calcul par systèmes sur puce ou SoC (System on Chip) prévue en 2011 par l'ITRS (International Technology Roadmap for Semiconductors) pour la prochaine décennie. Le nombre de cœurs de calcul intégrés au sein d'un seul SoC ne cessent d'augmenter. Malgré la diminution de la taille des transistors, la surface d'un circuit tend de nouveau à augmenter car le nombre de cœurs sur un même circuit augmente exponentiellement. Selon la performance requise, le cœur peut être arrêté s'il n'est pas utilisé. La gestion, à grain fin, cœur par cœur ou d'un groupement de cœurs, de la consommation des circuits devient donc cruciale. Pour cela, il faut adapter les paramètres de fonctionnement de chaque domaine de consommation dynamiquement. Ceci devient un enjeu majeur car la consommation doit être réduite à tous les niveaux.

Problématique

Comment réduire la consommation des circuits intégrés ? Dans la littérature ce problème a tout d'abord été adressé par l'adaptation dynamique de la fréquence (DFS). Le DFS permet de réduire la consommation mais n'est pas la meilleure solution d'un point de vue efficacité énergétique ; l'énergie consommée est constante. En effet, en modifiant la fréquence, le temps nécessaire pour effectuer le travail est modifié. Une seconde technique est l'adaptation dynamique de la tension et de la fréquence (DVFS) qui permet de diminuer l'énergie consommée. L'efficacité de cette technique dépend du mécanisme choisi pour appliquer la tension. Par exemple, si le régulateur de tension est intégré, son rendement peut varier entre 50% et 90% selon la tension de sortie. Cette technique n'est donc pas optimale d'un point de vue énergétique pour un régulateur intégré. Comment améliorer la consommation en utilisant la technique du DVFS avec un convertisseur intégré ? Une technique connue mais très peu répandue est le DVFS discret qui est composé de 2, 3 ou 4 points de fonctionnement. En général, l'actionneur discret utilisé

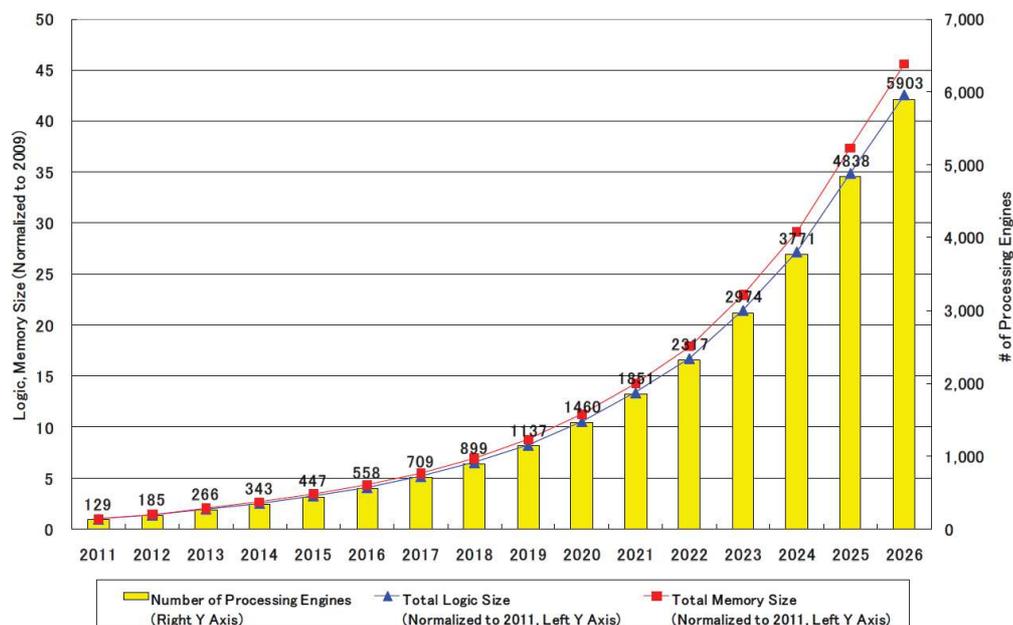


FIGURE 1.1: Tendence des futures évolutions du nombre de cœurs par SoC (source : [2]).

ne permet d'appliquer au circuit que 2 points de fonctionnement (un actionneur est un dispositif fournissant un paramètre physique). Comment définir et placer les points de fonctionnement d'un DVFS discret ? Le gain en consommation dépend-il du choix des points de fonctionnement ? De nouvelles technologies sont développées telles que le FD-SOI, elles offrent de nouvelles opportunités comme l'utilisation d'un degré de liberté supplémentaire qui est la tension de body bias. Quel gain supplémentaire apporte ce nouveau degré de liberté ? Est-il possible de coordonner la gestion de la fréquence, la tension d'alimentation et la tension de body bias ?

Objectifs et contributions de la thèse

L'objectif de cette thèse est de fournir des solutions de gestion de la consommation lorsque l'actionneur de tension d'alimentation est discret tout en considérant les possibilités supplémentaires apportées par les technologies telles que le FD-SOI. Des techniques de gestion de la consommation dynamique et de placement de points de fonctionnement seront proposées dans ce contexte.

Les contributions principales de ce manuscrit sont :

- l'évaluation de l'ajout de points de fonctionnement supplémentaires sur un circuit ;

- l'évaluation du gain en consommation associé au choix du placement des points de fonctionnement supplémentaires ;
- la proposition d'une méthode permettant de placer les points de fonctionnement afin de maximiser le gain en consommation ;
- l'exploration des nouvelles opportunités technologiques afin de proposer des solutions innovantes de gestion dynamique de la consommation. La tension d'alimentation, la fréquence d'horloge et la tension de body bias sont adaptées afin d'optimiser la consommation énergétique des circuits complexes en considérant les contraintes d'implémentation des actionneurs.

Contexte des travaux

Cette thèse a été effectuée en collaboration entre le Département Architectures Conception Logiciels Embarqués (DACLE) du CEA-Leti, Grenoble, France ; et le Département de Microélectronique du Laboratoire d'Informatique, Robotique et Microélectronique de Montpellier (LIRMM), France. Le laboratoire d'accueil du CEA est le Laboratoire Infrastructure et Atelier Logiciel pour Puces (LIALP) où les domaines de recherche incluent la gestion de la consommation électrique des systèmes embarqués au niveau logiciel afin de mieux exploiter les opportunités matérielles.

Organisation du manuscrit

Le chapitre 2 présente les différentes solutions technologiques existantes qui permettent de pallier le problème de miniaturisation que rencontre la technologie CMOS BULK. Les différentes solutions de gestion du compromis consommation-performance qui ont été développées au cours du temps afin d'améliorer l'efficacité énergétique des circuits sont exposées. Puis, les actionneurs permettant de modifier la fréquence F , la tension d'alimentation V_{dd} et la tension de body bias V_{bb} sont illustrés. Le choix de différents actionneurs est alors ensuite justifié. Enfin les problèmes que l'on souhaite résoudre sont exposés.

Dans le chapitre 3, nous évaluons si l'ajout de points de fonctionnement supplémentaires permet de réduire la consommation d'un circuit. Nous quantifions le gain en consommation de l'ajout de points de fonctionnement sur deux circuits différents. Nous proposons une solution analytique qui permet de déterminer la fréquence d'horloge des points de fonctionnement supplémentaires permettant de maximiser le gain en consommation sur toute la gamme de fréquence d'horloge disponible sur le circuit.

Le chapitre 4 expose deux méthodes de gestion de la consommation *via* la gestion dynamique de (F, V_{dd}, V_{bb}) . Ces deux méthodes reposent sur des concepts théoriques pour prouver l'optimalité d'un point de vue consommation tout en

garantissant la performance. La première méthode consiste à gérer un circuit composé de trois actionneurs discrets et la seconde consiste à gérer un circuit composé de trois actionneurs dont un discret.

Le chapitre 5 évalue les gains en consommation apportés par les méthodes proposées dans le chapitre 4. La validation est réalisée sur des profils de consommation en fonction de la fréquence issus dans un premier temps de simulation et puis de mesures sur un circuit réel implémenté en technologie FD-SOI 28 nm.

Enfin le chapitre 6 conclut ce manuscrit et propose les perspectives à ces travaux.

Chapitre 2

Évolution de la gestion de la consommation

La capacité d'intégration des systèmes sur puce augmente, cette croissance suit encore la loi de Moore [3]. Le principe de Dennard [4] mentionne qu'à chaque miniaturisation la densité de puissance est constante. Cela revient à dire que si la taille d'un transistor est divisée par 2, sa consommation est réduite par 4. Cependant, en réduisant la taille des transistors, il n'est plus possible de réduire la tension et le courant dont ont besoin les transistors pour fonctionner de manière fiable. Cela annonce ainsi la fin de l'ère Dennard. La consommation est devenue l'une des raisons principales de la limitation en performance. La performance des systèmes intégrés a tendance à stagner. La réduction de la consommation n'est plus possible car la consommation statique augmente. Cette augmentation de la consommation statique modifie particulièrement le rapport entre la consommation statique et dynamique. La consommation affecte la température du circuit intégré qui, par effet boule de neige, augmente la consommation statique qui dépend de la température.

Des techniques de gestion sont étudiées à différents niveaux afin de réduire la consommation totale des systèmes. En effet, la gestion de la consommation des circuits peut être réalisée au niveau système d'exploitation, au niveau compilateur, au niveau circuit et système et voire même au niveau technologique.

Il existe d'innombrables techniques au niveau logiciel. Nous allons évoquer quelques exemples.

La gestion au niveau système d'exploitation peut être réalisée en déterminant la quantité de travail pour exécuter une tâche. Par la suite le système d'exploitation décide de la fréquence à appliquer afin d'exécuter la tâche (principe de la

technique du DFS) [5].

La gestion de la consommation au niveau architecture et logiciel peut être réalisée en combinant des cœurs hétérogènes. Par exemple, l'architecture big.LITTLE [6] de ARM est une architecture hétérogène combinant deux clusters, un avec des cœurs basse consommation (Cortex-A7) et l'autre avec des cœurs haute performance (Cortex-A15). Chaque cluster possède sa propre gestion de la consommation-performance, basée sur la technique DVFS. Selon la performance requise par une tâche, elle peut être migrée d'un cœur à un autre cœur afin de réduire la consommation. La migration de tâche permet soit d'améliorer l'expérience de l'utilisateur soit de respecter la performance de la tâche, *i.e.* que la tâche soit finie à temps.

Ce chapitre présente l'évolution des solutions permettant de gérer le compromis consommation-performance des dispositifs. Nous allons tout d'abord exposer les solutions technologiques permettant d'agir sur ce compromis directement ou indirectement.

De nouvelles technologies ont été développées afin de pallier les limites physiques de la miniaturisation atteinte par la technologie CMOS BULK. Nous allons ensuite nous intéresser plus particulièrement à la gestion de la consommation au niveau circuit et système. Par la suite, nous allons présenter les solutions existantes pour gérer le compromis consommation-performance au niveau logiciel. Puis, des méthodes de gestion du compromis consommation-performance seront présentées.

2.1 Solution au niveau technologique

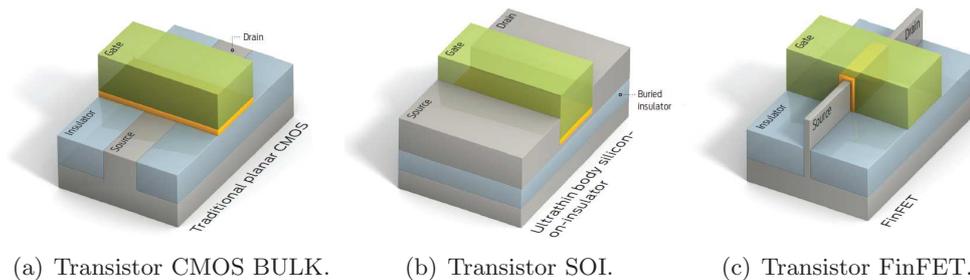


FIGURE 2.1: Différentes technologies de transistors (source : [7]).

2.1.1 Technologie CMOS BULK

Au fil des années, les transistors MOSFET (Metal Oxide Semiconductor Field-Effect Transistor) de type N ou P ont beaucoup évolué. Ces transistors sont

utilisés pour former des portes logiques CMOS (Complementary Metal Oxide Semiconductor), voir la Figure 2.1(a).

En CMOS BULK, la gestion du compromis consommation-performance peut être réalisée en modifiant la tension d'alimentation, la fréquence ou la tension du substrat V_{bb} , voir la Figure 2.2. La modification de la tension du substrat entraîne un changement de la tension de seuil V_{th} des transistors. V_{th} est définie comme étant la tension entre la grille et la source pour laquelle la zone d'inversion apparaît ; un canal de conduction se crée entre la source et le drain. Lorsque la tension grille-source est inférieure à la tension de seuil, le transistor est bloqué ; il n'est pas conducteur. Sinon, le transistor est passant ; il conduit le courant entre la source et le drain. V_{th} peut être fixée à la calibration pour compenser les variations de procédés de fabrication ou selon la technologie, la tension V_{th} peut être modifiée en cours d'exécution afin d'agir sur la performance ou la consommation. Cependant, en CMOS BULK, il est nécessaire d'implanter des caissons profonds (*triple-well*) pour pouvoir modifier dynamiquement la tension de substrat. De plus, les transistors partagent le même caisson, si l'on souhaite appliquer une technique de *Body Biasing*, alors les transistors d'un même caisson sont polarisés au même potentiel.

Une technique souvent utilisée au niveau conception post-routage est la technique Multi-Vth. L'objectif est de réduire la consommation statique des chemins où la vitesse n'est pas requise. Certaines cellules LVT (low Vth) sont remplacées par des cellules HVT (high Vth) afin d'augmenter la tension de seuil de ces chemins. Cependant, il est nécessaire de vérifier l'intégrité du signal, par exemple les retards de couplage.

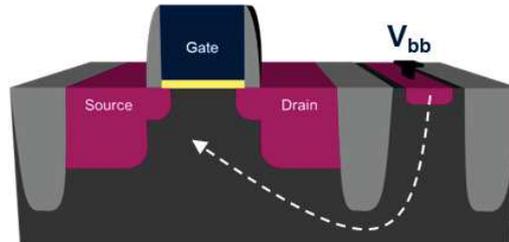


FIGURE 2.2: Tension de polarisation du substrat en technologie BULK (source : [8]).

De plus, la course à la miniaturisation en technologie CMOS BULK, a atteint ses limites, physiques et économiques. En effet, le coût de développement d'une technologie augmente par nœud technologique. Afin de satisfaire le marché des systèmes intégrés, des technologies ont été développées afin de produire des circuits plus performants. Cela est dû au fait qu'ils embarquent plus de transistors pour une même surface. Le développement de nouvelles technologies est donc un enjeu stratégique pour le marché du semi-conducteur. Actuellement deux choix

technologiques se démarquent.

2.1.2 FinFET

La technologie FinFET (Fin-shaped Field Effect Transistor) a été développée par INTEL. Il s'agit d'un transistor basé sur un transistor classique MOSFET dont la grille entoure le canal de tous les côtés car le drain et la source forment une « ailette » (Fin en anglais) surélevée. Ce transistor est décrit comme un transistor 3D ou *tri-gate*, voir la Figure 2.1(c). Pour un même nœud technologique qu'en CMOS BULK, les transistors consomment moins à performance identique. Cette réduction de la consommation résulte de la réduction des courants de fuite par le contrôle électrostatique du canal qui a été amélioré [9]. De même, pour une même consommation qu'en CMOS BULK, la fréquence appliquée aux transistors en technologie FinFET peut être augmentée. Cependant, cette technologie nécessite un investissement économique pour développer de nouvelles lignes de production et des techniques pour les concepteurs. De plus, le procédé de fabrication de cette technologie est très complexe.

Nous ne nous intéresserons pas à cette technologie car elle n'offre pas l'opportunité de modifier la tension de body bias afin d'agir sur le compromis consommation-performance. En effet, un de nos objectifs est d'évaluer l'apport de la technique de *Body Biasing*.

2.1.3 SOI

La technologie SOI (Silicon On Insulator) a été développée par STMicroelectronics et le CEA-Leti, voir sur la Figure 2.1(b). Deux technologies ont été produites le Partially Depleted SOI (PD-SOI) puis le Fully Depleted SOI (plus communément appelé FD-SOI) [10]. Le PD-SOI permet d'augmenter la performance des circuits par rapport au CMOS BULK mais rencontre les mêmes limites de miniaturisation que le BULK.

Le FD-SOI permet à la fois de créer des circuits haute performance et basse consommation. Par ailleurs, cette technologie permet de réduire la taille des transistors en dessous de 22nm. Elle a donc les mêmes avantages que la technologie FinFET en termes de performance et de consommation. En effet, un circuit en FD-SOI consommera moins qu'en CMOS BULK. Cette réduction de la consommation est aussi liée au meilleur contrôle électrostatique du canal du transistor. Un autre avantage s'ajoute au développement de cette technologie, les layouts et les lignes de production de la technologie CMOS BULK peuvent être réutilisés.

En SOI, le body de chaque transistor est électriquement indépendant et isolé des autres transistors. Cette spécificité permet de faire varier indépendamment la tension de seuil V_{th} des transistors et ainsi de découper le circuit en plusieurs îlots

indépendants. Par ailleurs, la technologie SOI, comparée au CMOS BULK, permet de gagner en surface pour un même nœud et de supprimer le coût additionnel apporté par une technique de gestion de la polarisation du substrat [11].

Il existe différentes technologies SOI mais nous portons un intérêt particulier à la technologie FD-SOI à oxyde enterré mince (UTBB pour Ultra Thin Body and Box en anglais). Cette technologie présente de nombreux avantages comme la gamme de modification de la tension de body bias qui est plus importante qu'en CMOS BULK et PD-SOI. Afin d'alléger la lecture, l'acronyme FD-SOI sera utilisé dans la suite du manuscrit au lieu de FD-SOI à oxyde enterré mince et de UTBB FD-SOI.

En technologie FD-SOI, une couche de face arrière ou Back-Plane (BP), connue aussi sous le nom de Ground-Plane, (GP) est formée grâce au fort niveau de concentration du dopage du silicium sous le BOX (Buried Oxide). La tension de seuil V_{th} dépend du type de dopage de la face arrière et de la valeur de la tension entre le body et la source V_{bs} appliquée. Ainsi, la tension de seuil V_{th} peut être modifiée dynamiquement en fonction de la tension V_{bs} appliquée. Sur la Figure 2.3, nous pouvons observer qu'une nouvelle entrée permet d'accéder au body en technologie FD-SOI. Cette tension permet de modifier indirectement la tension de seuil V_{th} . La tension de seuil est donnée par l'équation suivante :

$$V_{th} = V_{th0} - \beta_{body} \cdot \sqrt{V_{bs}} \quad (2.1)$$

avec V_{th0} est la tension de seuil de référence quand aucune charge n'est accumulée dans le body (*i.e.* quand $V_{bs} = 0$ V), β_{body} est le coefficient d'effet de body et V_{bs} est la tension entre la source et le body.

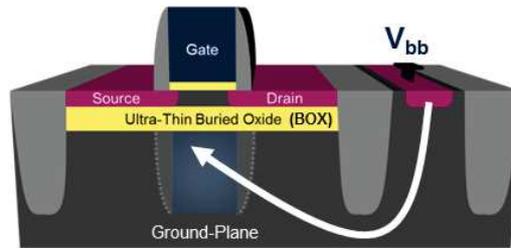


FIGURE 2.3: Nouvelle entrée permettant d'agir sur la tension de body bias en technologie FD-SOI (source : [8]).

L'avantage principal de la technologie FD-SOI est que l'on peut ajuster dynamiquement la tension de body bias V_{bb} pour augmenter les performances ou réduire la consommation. Notamment, la consommation statique dépend exponentiellement de V_{th} . Par ailleurs, la gamme de tension de body bias applicable est plus importante en FD-SOI qu'en CMOS BULK. Cette large gamme de tension de body bias est en grande partie due à l'isolation par le BOX [12].

Notons qu'en technologie FD-SOI la technique Multi-Vth n'est pas utilisée car il faut un espacement minimum entre les différentes cellules. Cela implique donc une surface plus grande pour le circuit.

On peut également remarquer que le choix d'une technologie ou de l'autre peut influencer le nombre d'actionneurs que l'on va implémenter. Un actionneur est un dispositif qui transforme la grandeur qui lui est fournie en une autre grandeur. Dans cette thèse, la technologie FD-SOI est retenue pour développer le travail, mais il est à noter que ce qui va être présenté reste valable en utilisant d'autres technologies. Nous allons maintenant présenter les différentes techniques de gestion de la consommation et de la performance au niveau circuit et système.

2.2 Gestion de la consommation au niveau circuit et système

La gestion de la consommation des systèmes intégrés est devenue une priorité depuis plusieurs années. Ceci est dû au fait que le besoin en performance de ces systèmes augmente. Leur consommation est liée à la performance, si on augmente donc la performance, on augmente aussi la consommation. Cependant, cette relation n'est pas linéaire. Nous portons un intérêt particulier à la gestion de la consommation et de la performance au niveau circuit et système, car c'est une contrainte de plus en plus forte pour les circuits intégrés, notamment pour les systèmes embarqués.

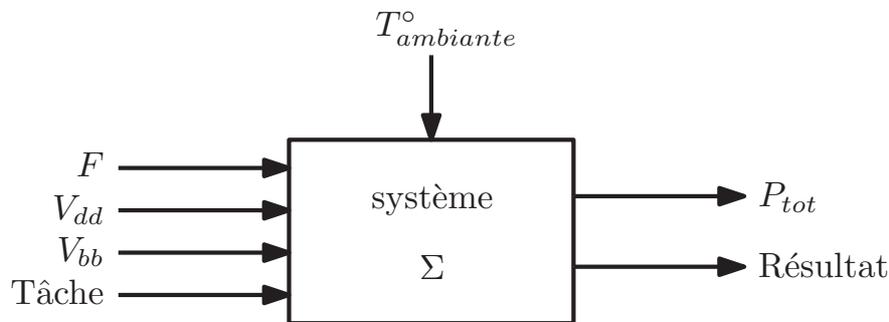


FIGURE 2.4: Représentation du système par un modèle entrées-sorties.

La gestion de la consommation à ce niveau se fait à l'aide de techniques matérielles, permettant d'agir sur la consommation et/ou la performance. Nous considérons que les dispositifs étudiés peuvent être représentés par un modèle possédant plusieurs entrées (F , V_{dd} , V_{bb}), et plusieurs sorties (par exemple la consommation totale P_{tot} du circuit) comme sur la Figure 2.4. Nous allons tout d'abord analyser les modèles de consommation dynamique et statique. La consommation d'un circuit peut être réduite en cours de conception (*design time*), en cours d'exé-

cution (*run time*), en *standby* lorsque le circuit n'est pas en activité. Le *standby* est un cas particulier du *run time*. La gestion de la consommation au niveau circuit et système regroupe les techniques de gestion lorsque le circuit n'est pas en activité (*standby*) ou s'il est en cours d'exécution. Nous allons présenter les différentes techniques existantes à ce jour pour les deux modes de fonctionnement.

2.2.1 Modèles de consommation de circuits synchrones

La consommation des circuits est le facteur limitant l'augmentation de la performance. Des modèles pour la consommation dynamique et la consommation statique pour les circuits synchrones, en technologie MOSFET, ont été proposés [13].

La consommation totale P_{tot} du circuit est composée de la consommation dynamique P_{dyn} et de la consommation statique P_{stat} . La consommation dynamique est composée de la consommation court-circuit et de la consommation durant les commutations des fils.

La consommation dynamique P_{dyn} représente la consommation relevant des changements d'états des transistors du circuit. Elle dépend linéairement de la fréquence d'horloge F , de l'activité A et quadratiquement de la tension d'alimentation V_{dd} (voir l'équation (2.2)).

$$P_{dyn} \propto A \cdot F \cdot V_{dd}^2 \quad (2.2)$$

La consommation statique P_{stat} est due aux fuites de courants lorsque le circuit est alimenté. Elle dépend de la tension d'alimentation V_{dd} , de la tension de seuil V_{th} et de la température T° du circuit, comme le montre l'équation (2.3). La dépendance de la consommation statique à la tension d'alimentation V_{dd} croît avec la diminution de la taille des transistors. Si la tension de seuil V_{th} augmente, la consommation statique diminue. Alors que si la température du cœur augmente, la consommation statique augmente aussi.

$$P_{stat} \propto \gamma \cdot V_{dd}^\epsilon \cdot e^{-\kappa \cdot \frac{V_{th}}{T^\circ}} \quad (2.3)$$

Le paramètre A dépend des caractéristiques de la tâche exécutée et du circuit permettant l'exécution des tâches alors que γ , ϵ et κ sont dépendants des caractéristiques du circuit et de la technologie employée.

Ces modèles de consommation vont être utilisés pour expliquer l'effet que chaque entrée (F, V_{dd}, V_{bb}) du système peut avoir sur la consommation totale en utilisant les différentes techniques au niveau circuit et système existantes à ce jour.

2.2.2 Gestion en mode standby

Dans la littérature, il existe des techniques permettant d'arrêter soit le signal d'horloge F soit l'alimentation V_{dd} .

Technique d'arrêt de l'horloge : clock gating

La source principale de la consommation dynamique des circuits synchrones en *standby* est l'arbre d'horloge. La technique d'arrêt de l'horloge (clock gating) consiste à couper le signal d'horloge des parties d'un circuit synchrone qui sont inactives [14]. L'horloge étant coupée, la consommation dynamique des parties concernées est nulle car $F = 0$ Hz, comme décrit à l'équation (2.2). Par ailleurs, le fait d'arrêter le signal d'horloge d'une partie du circuit permet de maintenir l'état des valeurs à la sortie des bascules arrêtées. En effet, si la valeur à l'entrée d'une des bascules arrêtées change, la sortie ne sera pas modifiée car la bascule est inactive. Cela permet d'éviter de commuter et de perdre de l'énergie inutilement. La technique d'arrêt d'horloge est un bon moyen de réduire la consommation dynamique lorsqu'une partie du circuit n'est pas en activité [15]. Notons que cette technique ne permet pas d'agir sur la consommation statique d'un circuit qui est la partie dominante de la consommation des circuits en *standby*. Néanmoins, cette technique a un faible impact sur la conception et l'implémentation du circuit. En effet, l'insertion du clock-gating est faite au niveau de la synthèse par les outils. La Figure 2.5 illustre un circuit de clock gating basé sur une bascule Latch.

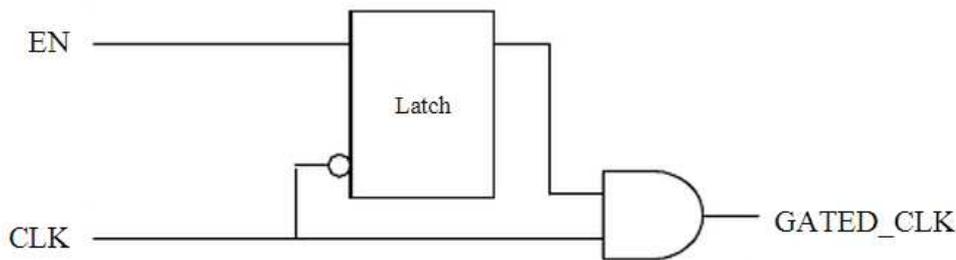


FIGURE 2.5: Circuit de clock gating basé sur une bascule Latch (source : [16]).

Technique d'extinction de l'alimentation : power gating

Afin d'éliminer les courants de fuite, il faut idéalement déconnecter la partie que l'on souhaite éteindre de l'alimentation. Le power gating consiste à arrêter d'alimenter les blocs du circuit qui ne sont pas utilisés [17]. Ainsi, la consommation statique des blocs éteints est nulle car elle dépend de la tension d'alimentation. Cependant cette technique introduit des bruits nommés Power Gating Noise (PGN) comprenant les variations de tensions résistives (IR -drop) et les variations de tensions inductives (Ldi/dt). Ces variations sont dues à des chutes de potentiel par rapport à la valeur nominale de tension. Elles sont provoquées par des appels de courant. Les PGN sont dus aux grands appels de courants ou aux décharges de courant durant les phases de transition (passage d'éteint à allumer ou *vice-versa*). Ils affectent les grilles d'alimentation et de masse. Notons que la remise

en route de l'alimentation revient à une perte de temps et d'énergie importante. C'est pourquoi il faut bien prendre en compte la durée d'extinction de l'alimentation pour évaluer si cela en vaut la peine ou non. De plus, l'impact de cette technique sur la conception et sur l'implémentation est important. En effet, la conception de cette technique n'est pas automatisée et elle a un surcoût matériel. Cependant, elle permet une réduction de la consommation très conséquente du fait de la dépendance P_{dyn} et P_{stat} à V_{dd} .

Dynamic Body Biasing

Une autre manière de réduire les courants de fuite est d'augmenter la tension de seuil V_{th} des transistors. La polarisation du substrat est une technique permettant de modifier la tension de seuil V_{th} des transistors en polarisant le caisson en BULK [18] ou en polarisant le body d'un transistor en FD-SOI. La modification de la polarisation du substrat en *standby* est connue sous le nom de Dynamic Body Biasing (DBB). Pour un transistor NMOS, si on polarise le substrat négativement par rapport à la source du transistor ($V_{bs} < 0$), la tension de seuil du transistor augmente :

$$V_{th_n} = V_{th_{n0}} + \gamma \cdot (\sqrt{V_{sb} + |2\phi_f|} - |2\phi_f|) \quad (2.4)$$

avec V_{th_n} la tension de seuil d'un transistor NMOS, $V_{th_{n0}}$ la tension de seuil d'un transistor NMOS lorsque la tension de polarisation du substrat est nulle, ϕ_f le potentiel de travail du transistor qui dépend de la densité d'électrons et de trous et γ un facteur dépendant du dopage, de la permittivité du vide, et de la valeur de la capacité d'oxyde.

Pour un transistor PMOS, si l'on polarise le substrat positivement par rapport à la source du transistor ($V_{bs} > 0$), la tension de seuil du transistor augmente en valeur absolue :

$$V_{th_p} = V_{th_{p0}} + \gamma \cdot (\sqrt{V_{bs} + |2\phi_f|} - |2\phi_f|) \quad (2.5)$$

avec V_{th_p} la tension de seuil d'un transistor PMOS, $V_{th_{p0}}$ la tension de seuil d'un transistor PMOS lorsque la tension de polarisation du substrat est nulle.

Le DBB utilise la technique permettant de polariser le substrat en mode inverse qui est communément appelée RBB. Le RBB permet de réduire les courants de fuite en augmentant V_{th} mais il augmente l'effet canal-court¹, les variations de la tension de seuil à travers la puce ainsi que l'effet tunnel¹ (BTBT). Rappelons qu'une modification linéaire en tension de seuil revient à agir de manière exponentielle sur la consommation statique (voir l'équation (2.3)). Le principal inconvénient de cette technique est son coût énergétique lors de la charge et de la

1. Il s'agit d'effets que l'on souhaite éviter.

décharge de la capacité du substrat. D'autres inconvénients existent selon la technologie utilisée. Par exemple, en BULK, il faut implanter des caissons profonds (*triple-well*), causant ainsi un surcoût en termes de surface.

Cette technique a un fort impact sur la conception et sur l'implémentation du circuit.

Nous avons vu différentes techniques de gestion de la consommation au niveau circuit et système lorsque le circuit n'est pas en activité. Nous allons maintenant nous intéresser à la gestion de la consommation en cours d'exécution.

2.2.3 Gestion en cours d'exécution

La gestion de la consommation au niveau circuit et système lorsque le circuit est actif se fait en agissant sur la fréquence F , la tension d'alimentation V_{dd} et la tension de body bias V_{bb} (voir la Figure 2.4). Ces grandeurs sont adaptées en cours d'exécution aux contraintes de consommation ou de performance pour réaliser un compromis. La Figure 2.6 illustre l'allure du profil de la consommation en fonction de la fréquence pour les différentes techniques que nous allons introduire.

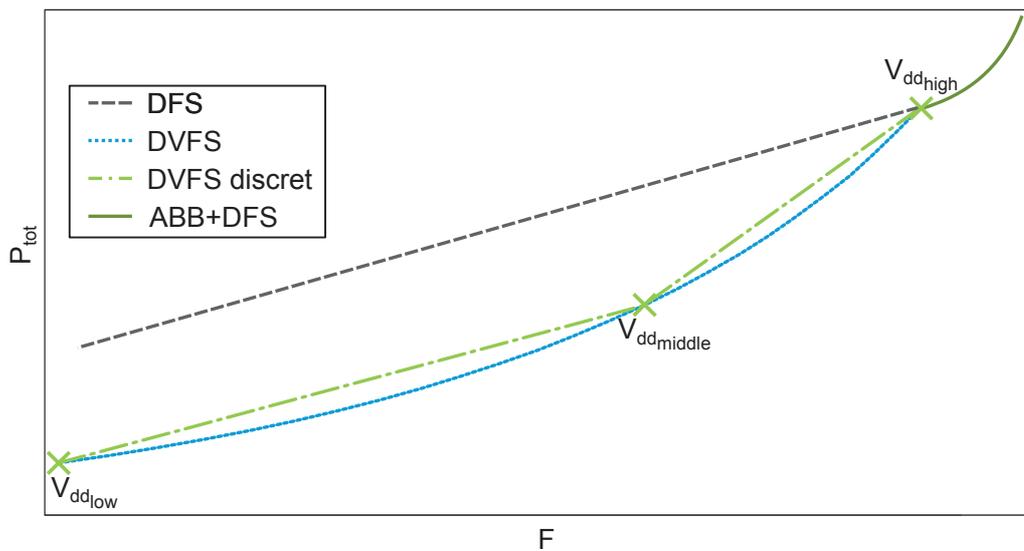


FIGURE 2.6: Gestion de la consommation et de la performance avec les techniques au niveau circuit et système.

Technique de DFS

Les premières techniques de gestion dynamique consistaient à ajuster seulement la fréquence d'horloge. Cette technique est connue sous le nom de DFS (Dynamic Frequency Scaling). Cette technique permet d'agir sur la performance et la consommation. En réduisant la fréquence, on diminue linéairement la consom-

tion dynamique P_{dyn} , comme le montre l'équation (2.2), voir la Figure 2.6. Cependant, on n'agit pas sur la consommation statique P_{stat} , voir l'équation (2.3). Cette technique permet de réduire la consommation dans le temps imparti pour exécuter la tâche d'un circuit mais ne permet pas de réduire l'énergie dissipée durant le temps imparti. En effet, si la fréquence requise est inférieure à la fréquence maximale atteignable pour la tension d'alimentation appliquée, dans ce cas le DFS n'est pas efficace d'un point de vue énergétique. Dans [19], les auteurs proposent d'utiliser la technique de DFS permettant de modifier le compromis gain en consommation et surcoût en consommation apporté par l'architecture implémentée. Chaque cœur de calcul a son propre DFS afin d'adapter dynamiquement la fréquence en fonction de la charge de calculs et de la charge de communications. L'objectif de cette publication est de satisfaire la contrainte de performance de l'application. La performance des cœurs de calculs non utilisés ou sous-exploités par l'application est dégradée afin de réduire la consommation du circuit.

L'impact du DFS sur la conception et l'implémentation est important. Il faut tout d'abord concevoir l'actionneur de fréquence et le contrôleur puis les implémenter. Le DFS a un rapport coût d'implémentation, réduction de la consommation moins intéressante que d'autres méthodes (voir la Figure 2.6).

Adaptation de la tension et la fréquence

Si l'on agit sur la fréquence, il est nécessaire d'adapter la tension d'alimentation pour améliorer l'efficacité énergétique d'un circuit [20] et pour éviter de produire des erreurs temporelles. Une erreur temporelle (communément appelé *timing fault*) est engendrée lorsque la fréquence d'horloge appliquée est supérieure à la fréquence maximale atteignable par le circuit. Afin de ne pas produire d'erreur temporelle dans le cas où les entrées ne sont pas modifiées conjointement [21] [22] [23], il est préconisé :

- pour augmenter la vitesse d'exécution, il faut tout d'abord augmenter la tension d'alimentation V_{dd} puis augmenter la fréquence F (voir la Figure 2.7(a));
- pour diminuer la vitesse d'exécution, il faut tout d'abord diminuer la fréquence F puis réduire la tension d'alimentation V_{dd} (voir la Figure 2.7(b)).

Dans [21], les auteurs proposent de suivre un chemin $F(V_{dd})$ prédéfini afin d'éviter de perdre de l'énergie lors de la commutation, voir la Figure 2.7(c). Le contrôle conjoint de F et V_{dd} permet de rester toujours au plus proche de la fréquence maximale atteignable F_{max} pour une tension d'alimentation.

La technique DVFS (Dynamic Voltage and Frequency Scaling), connue aussi sous l'appellation DVS (Dynamic Voltage Scaling) permet de modifier les deux grandeurs F et V_{dd} . Il s'agit d'une des techniques les plus étudiées à ce jour. Pour une tension, la fréquence est adaptée afin d'être proche de F_{max} , cela permet d'augmenter l'efficacité énergétique en minimisant la consommation dynamique.

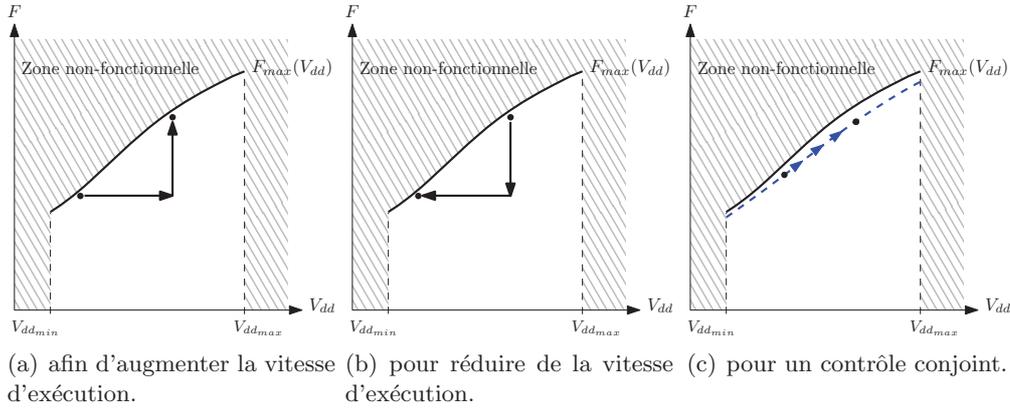


FIGURE 2.7: Séquence de modification de la fréquence F et de la tension d'alimentation V_{dd}

De plus, en modifiant la tension d'alimentation, on joue à la fois sur la consommation dynamique et statique, comme le montrent les équations (2.2) et (2.3) et la Figure 2.6. Cela permet de réduire la consommation totale P_{tot} .

Si la performance requise par le système intégré est réduite, la fréquence va être réduite et la tension d'alimentation aussi. Ainsi, la consommation dynamique sera réduite quadratiquement avec la tension d'alimentation et linéairement avec la fréquence. De plus, la consommation statique sera réduite linéairement avec V_{dd} .

Les auteurs de [24] suggèrent de découper le circuit en plusieurs domaines d'horloge (Multiple Clock Domains ou MCD) dans lesquels une adaptation de la tension d'alimentation et de la fréquence peut être réalisée indépendamment des autres domaines d'horloge.

Dans [25], un DVFS est utilisé pour exécuter des applications en temps non-réel sur un circuit intégré. L'objectif est de minimiser la dissipation énergétique en dégradant la performance tout en prenant en compte en cours d'exécution les accès mémoires externes. Le DVFS repose sur un modèle de régression permettant à l'unité de calcul centrale ou CPU de calculer la charge de travail et le temps mort nécessaire pour le prochain intervalle de temps. Cela permet d'adapter la tension d'alimentation V_{dd} et la fréquence F du circuit pour économiser de l'énergie tout en respectant la contrainte temporelle faible.

Les auteurs de [26] proposent des DVFS qui prennent en compte les variations de procédés de fabrication. Durant une phase de test, les variations de procédés de fabrication de chaque îlot de tension-fréquence ou VFI (de l'anglais Voltage/Frequency Island) sont caractérisées. La mesure des courants de fuite d'un VFI permet de déterminer le paramètre de variation de procédé de fabrication effective. Selon la valeur de ce paramètre, les niveaux de tension et fréquence du VFI vont être augmentés ou diminués selon les courants de fuite du VFI.

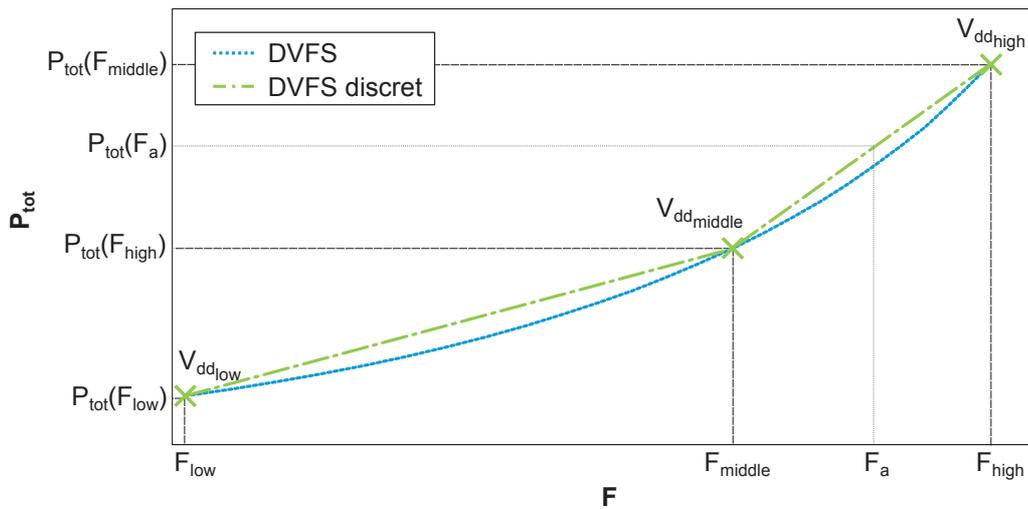


FIGURE 2.8: Profil consommation-performance DVFS et DVFS discret.

La plupart des DVFS nécessitent que la tension d'alimentation soit ajustée de manière adaptative et en continu². Cela peut introduire un surcoût important en fonction du système dans lequel le processeur est embarqué. La plupart des circuits composés de microprocesseur intègrent des actionneurs de tension d'alimentation complexes permettant d'appliquer une large gamme de valeurs. Cependant, dans certains circuits le coût d'un tel actionneur peut être immense. Heureusement, les avantages du DVFS peuvent être obtenus même avec un petit nombre de valeurs discrètes. En combinant un actionneur de tension d'alimentation discret³ et un actionneur de fréquence discret, on peut obtenir un DVFS discret [27]. Dans ce cas, la consommation équivalente $P_{tot}(F_a)$ pour une fréquence F_a correspond à l'interpolation linéaire entre les différents points de fonctionnement tension-fréquence sur le profil consommation-fréquence (voir la Figure 2.8). La performance est atteinte en ajustant le ratio de temps passé à chaque point de fonctionnement tension-fréquence. La performance est donc atteinte en mode *dithering* lorsque la performance requise ne correspond pas à la fréquence d'un point de fonctionnement. Le *dithering* revient à exécuter une tâche en sautant d'un point de fonctionnement vers un autre point de fonctionnement. Par exemple du point de fonctionnement $(F_{low}, V_{dd_{low}})$ vers le point de fonctionnement $(F_{middle}, V_{dd_{middle}})$. Dans la suite du manuscrit, le mode *dithering* sera communément appelé le mode *hopping*.

Les auteurs de [28] proposent de coupler l'actionneur de tension d'alimentation et celui de fréquence à un système de surveillance de la variabilité environnemen-

2. Un actionneur continu peut fournir un ensemble continu de valeurs (ou *a minima* un grand nombre de valeurs discrètes) à sa sortie.

3. Un actionneur discret peut fournir qu'un nombre limité de valeurs à sa sortie, par exemple 2, 3, 4 ou 5.

tale et à un contrôleur qui décidera des valeurs à appliquer aux actionneurs. Cela revient à ajouter des capteurs de température afin de pouvoir s'approcher au plus près de la fréquence maximale atteignable F_{max} pour la tension d'alimentation appliquée et la température du système [29]. L'architecture locale est nommée AVFS (Adaptive Voltage and Frequency Scaling). L'AVFS peut être implémenté avec des actionneurs continus ou avec des actionneurs discrets.

La technique DVFS a une grande influence sur la conception et l'implémentation des actionneurs sur le circuit. Cependant, elle permet de réduire de manière importante la consommation. P_{dyn} est liée quadratiquement à V_{dd} et linéairement à F et P_{stat} dépend de V_{dd} (voir la Figure 2.6).

Polarisation du substrat

Nous avons vu précédemment des techniques de gestion de la consommation avec l'adaptation dynamique de la tension d'alimentation V_{dd} . En cohérence avec les techniques présentées en *standby*, il nous paraît logique de considérer l'adaptation de la tension de seuil V_{th} en cours d'exécution pour réduire la consommation. Cette technique est communément appelée Adaptive Body Biasing ou ABB. L'ABB est l'équivalent du DBB en cours d'exécution. L'ABB est une technique intéressante car elle permet d'agir de manière dynamique sur la consommation statique. Cette technique est généralement utilisée pour compenser les variations de température, de procédés de fabrication ou l'effet du vieillissement.

Dans l'article [30], les auteurs proposent de modifier la tension de polarisation du substrat (indirectement, la tension de seuil) afin de compenser les variations de procédé et de température. Leur objectif est d'éviter les erreurs temporelles et de réduire les courants de fuite. Ainsi, il est possible de maintenir les performances du circuit.

Les articles [31], [32], [33], [34] proposent l'utilisation de l'ABB pour modifier la tension de body bias afin d'agir sur la tension de seuil des transistors. La technique d'ABB permet de générer et de contrôler la tension body bias ou tension de polarisation du substrat pour lutter contre les erreurs temporelles dues aux variations de température, de tension, à l'instabilité de la température de polarisation (Bias Temperature Instability, BTI) et pour augmenter les performances du circuit quand il est en fin de vie.

L'ABB utilise le Forward Body Biasing (FBB). Le FBB permet de diminuer la tension de seuil V_{th} du circuit afin d'en améliorer les performances. On applique une tension $V_{bs} > 0$ pour les NMOS et $V_{bs} < 0$ pour les PMOS. Par contre si l'on augmente la tension de seuil V_{th} , les courants de fuite sont réduits mais le temps de traversée des portes est augmenté tandis qu'en la diminuant le temps de propagation est réduit (i.e. la vitesse augmente) mais les courants de fuite sont plus importants [35].

Le DBB utilise le RBB qui est moins efficace avec la diminution de la tech-

nologie [36]. Pour éviter d'augmenter l'effet canal-court, nous pouvons utiliser le FBB qui est particulièrement adapté aux transistors lents c'est-à-dire qui ont de grandes tensions de seuil.

Il est préférable d'utiliser une combinaison des deux techniques, c'est-à-dire RBB et FBB pour fonctionner dans une large gamme de V_{th} . Cela permet ainsi de pouvoir agir sur la consommation à la fois en *standby* et en cours d'exécution.

Dans la plupart des articles cités, la tension V_{bs} est modifiée selon les performances souhaitées, les valeurs optimales sont précalculées puis stockées, ces dernières étant obtenues par simulation. Notons que les valeurs optimales obtenues en simulation peuvent être différentes des valeurs optimales réelles.

Rappelons que le *Body Biasing* a un effet sur l'implémentation de la technique sur le circuit et sur la conception du circuit. Notons que cette technique permet d'agir sur la consommation statique.

Adaptation des tensions d'alimentation et de body bias

Nous avons vu les avantages des techniques DVFS et ABB. La combinaison des deux techniques est une bonne solution car elle permet d'agir, encore plus que le DVFS classique, à la fois sur la consommation dynamique et sur la consommation statique. L'action sur V_{dd} et V_{bb} a un plus grand impact sur la consommation statique P_{stat} par rapport à n'agir que sur V_{dd} . En effet, P_{stat} est réduit linéairement avec V_{dd} et de manière exponentielle avec la tension de body bias V_{bb} .

L'article [37] fut l'un des premiers articles à proposer une solution permettant de minimiser l'énergie dissipée par un processeur en combinant les techniques DVFS et ABB. Ils proposent une expression analytique permettant de trouver le compromis optimal entre la tension d'alimentation et la tension de body bias pour une fréquence donnée. Ils considèrent des actionneurs continus.

Les auteurs de [34] proposent de combiner les techniques DVFS et ABB pour diminuer la consommation tout en augmentant les performances du circuit. Ils cherchent la tension d'alimentation permettant de réduire l'énergie dissipée par l'exécution de la tâche pour une fréquence donnée. Dans un second temps, ils obtiennent la tension de body bias à appliquer pour la fréquence et la tension d'alimentation minimisant l'énergie dissipée. Des actionneurs continus sont considérés dans ces travaux mais les contraintes d'implémentation physique ne sont pas identiques à notre contexte (coût surface, coût production, rendement des actionneurs, etc). Par ailleurs, la procédure exposée n'est pas adaptative face aux variations paramétriques telles que la température.

L'article [38] propose une analyse en phase de test du circuit afin d'ajuster la tension d'alimentation et la tension de body bias. L'objectif est d'optimiser la consommation pour une performance cible en phase de test, dynamiquement, mais une seule fois pour toute la durée de vie du système. La procédure proposée

n'est pas adaptative face aux variations paramétriques telles que la température.

2.2.4 Synthèse

Le Tableau 2.1 [39] résume les avantages et les inconvénients des différentes techniques permettant d'agir sur la fréquence F , la tension d'alimentation V_{dd} et la tension de body bias V_{bb} en *standby* ou en cours d'exécution que nous avons présentées.

La consommation des circuits complexes est importante car ils sont composés de plusieurs îlots. Afin de gérer à grain fin la consommation des circuits complexes, il est nécessaire de dupliquer les actionneurs dans chaque îlot. Cependant, le coût de la duplication des actionneurs est à prendre en compte. Pour cela, des concessions au niveau des actionneurs peuvent être réalisées. Par exemple, un des actionneurs peut être discrétisé. Le Tableau 2.2 résume les avantages et les inconvénients sur la gestion du compromis consommation-performance en agissant sur les entrées F , V_{dd} , V_{bb} .

Pour pouvoir modifier la valeur des entrées (F, V_{dd}, V_{bb}) , nous avons besoin d'actionneurs. Dans notre contexte, on a besoin de pouvoir les intégrer sur la puce. La section suivante présente les différents actionneurs existants.

2.3 Actionneurs pour agir sur F , V_{dd} , V_{bb}

Le choix des actionneurs utilisés est essentiel car ils permettent de gérer dynamiquement les entrées du circuit. De plus, ce choix influence aussi les pertes en consommation du circuit car le rendement des actionneurs, en particulier celui de la tension d'alimentation V_{dd} , n'est pas parfait. Notons que les actionneurs sont choisis de manière à être intégrés.

2.3.1 Génération d'horloge variable intégrée

La génération d'horloge dans les circuits intégrés peut être réalisée avec différentes solutions. Nous allons en présenter deux, la boucle à verrouillage de phase (PLL) et la boucle à verrouillage de fréquence (FLL). En fait les deux solutions sont largement utilisées. Une boucle à verrouillage de phase ou de fréquence est composée de trois éléments, un comparateur, un filtre et un oscillateur contrôlé en tension (VCO), voir la Figure 2.9.

Boucle à verrouillage de phase

Une boucle à verrouillage de phase ou PLL (Phase Locked-Loop) est un système permettant d'asservir une fréquence de sortie sur un multiple de la fréquence d'entrée. Gupta résume dans [40], les implémentations les plus courantes pour une PLL.

TABLEAU 2.1: Synthèse des techniques de gestion du compromis consommation-performance (source : [39]).

Technique	F	V_{dd}	V_{bb}	Influence sur la consommation	Pénalité temporelle	Pénalité en surface	Impact sur la conception	Impact sur l'implémentation
Clock gating	On/Off	Fixe	Fixe	Moyen - Agit sur P_{dyn}	Très faible	Très faible	Faible	Faible
Power gating	Fixe	On/Off	Fixe	Énorme - Agit sur P_{dyn} et P_{stat}	Faible	Faible	Important	Important
DBB	Fixe	Fixe	Continu	Important - Agit sur P_{stat}	Faible	Faible	Important	Important
DFS	Continu	Fixe	Fixe	Moyen - Agit sur P_{dyn}	Faible	Faible	Important	Important
DVFS (DVS)	Continu	Continu	Fixe	Important - Agit sur P_{dyn}	Faible	Faible	Important	Important
DVFS discret	Continu	Discret	Fixe	Important - Agit sur P_{dyn} et P_{stat}	Faible	Faible	Important	Important
ABB	Fixe	Fixe	Continu /Dynamique	Important - Agit sur P_{stat}	Faible	Faible	Important	Important
DVFS + ABB	Continu	Continu	Continu	Énorme - Agit sur P_{dyn} et encore plus sur P_{stat}	Faible	Faible	Important	Important

TABEAU 2.2: Synthèse des avantages et des inconvénients sur la gestion du compromis consommation-performance en agissant sur les entrées F , V_{dd} , V_{bb} .

F	V_{dd}	V_{bb}	Avantage(s)	Inconvénient(s)
continue	continue	/	Efficacité énergétique	Gestion de la consommation à gros grain
continue	discrète	/	Gestion de la consommation à grain fin	Minimum Local vs Minimum Global
continue	continue	continue	V_{bb} permet d'agir sur l' efficacité énergétique	Gestion de la consommation à gros grain
continue	discrète	continue	Gestion de la consommation à grain fin V_{bb} permet d'agir sur l' efficacité énergétique	Consommation ?

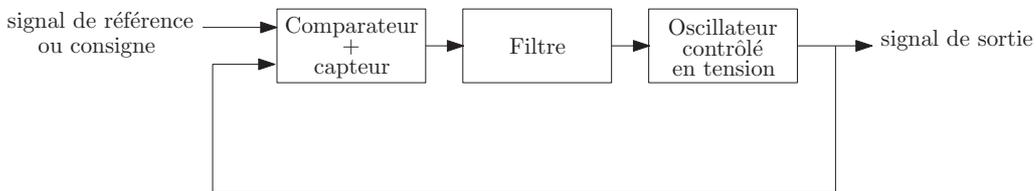


FIGURE 2.9: Structure de base d'une FLL et d'une PLL.

Une PLL complètement analogique conçue en technologie CMOS 45nm est proposée dans [41]. Ce papier s'inscrit dans un contexte où la fiabilité et la réduction de la densité de puissance étaient primordiales. Cette PLL est utilisée pour générer l'horloge dans les circuits intégrés numériques (CIN) et pour la synthèse en fréquence dans les émetteurs-récepteurs sans fil.

Une PLL a un coût de développement important car les cellules utilisées pour la conception ne sont pas standards. Le comparateur de phase est souvent analogique : il nécessite dans ce cas d'être adapté lors de la conception pour chaque nouveau nœud technologique, et son coût en surface est important, ce qui peut causer des problèmes d'intégration. Une PLL a un temps de démarrage non négligeable mais une fois démarrée, elle est précise.

Boucle à verrouillage de fréquence

Une boucle à verrouillage de fréquence ou FLL (Frequency Locked-Loop) est un système bouclé dont la variable de sortie à asservir est la fréquence (voir la Figure 2.9). Elle compare la fréquence en sortie de l'oscillateur contrôlé (par exemples en tension, en courant) à la fréquence de référence puis effectue les modifications requises afin que la fréquence de l'oscillateur soit réglée sur la fréquence de référence.

Les auteurs de [42] proposent une implémentation de FLL qui peut être dynamiquement contrôlée, qui est robuste aux variations de procédés de fabrication et de tension et de température.

L'avantage principal d'une FLL est qu'elle peut être entièrement conçue avec des cellules standards. Cela permet de réduire considérablement la conception pour un nouveau nœud technologique. Une FLL peut être conçue en grande partie en numérique. Néanmoins, le convertisseur numérique/analogique et le VCO sont analogiques. Cette conception permet de réduire la surface et facilite l'intégration de ce générateur à un circuit.

Pour un même nœud technologique, une FLL a une surface réduite par rapport à une PLL, permettant ainsi de réduire la consommation du générateur.

Comparaison des générateurs d'horloge

Le Tableau 2.3 résume les avantages et les inconvénients des générateurs d'horloge intégrés que nous avons considérés. Selon les contraintes d'implémentation physique, l'une ou l'autre solution peut être considérée. La FLL se distingue par son intégration et sa surface. En effet, une FLL peut être conçue en utilisant des cellules standards.

TABLEAU 2.3: Comparaison entre une PLL et une FLL

Actionneur	Entrée	Sortie	Avantages	Inconvénients
PLL	Phase	Phase	<ul style="list-style-type: none"> - Rapidité de mesure d'une différence de phase - Haute fréquence - Précision 	<ul style="list-style-type: none"> - Intégration - Temps de démarrage - Temps de conception - Surface
FLL	Fréquence	Fréquence	<ul style="list-style-type: none"> - Intégration - Consommation - Surface 	<ul style="list-style-type: none"> - Temps nécessaire pour évaluer une différence de fréquence

2.3.2 Tension d'alimentation variable

Un actionneur de tension est constitué d'un convertisseur DC-DC transformant une tension d'entrée en une tension régulée de sortie. Nous allons présenter trois exemples de générateur de tension parmi d'autres possibilités.

Régulateurs linéaires

Un régulateur linéaire est un système de régulation de tension basé sur un composant actif travaillant dans sa zone linéaire [43]. Le régulateur linéaire est utilisé afin de maintenir une tension constante. Il se comporte comme une résistance variable afin de maintenir la tension de sortie constante quel que soit

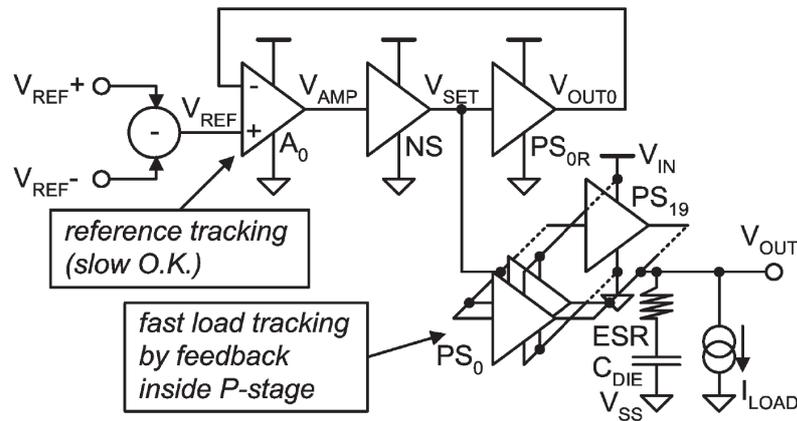


FIGURE 2.10: Schéma bloc du régulateur linéaire rapide (source : [43]).

le courant traversant la charge. La résistance équivalente dissipe de l'énergie, le rendement du régulateur est donc lié à la dissipation du dispositif. Le rendement η dépend du ratio de conversion $\frac{V_{out}}{V_{in}}$, avec V_{out} la tension de sortie du régulateur et V_{in} la tension d'entrée du régulateur qui est fixe. Le rendement dépend quasiment linéairement du ratio de conversion. Si l'on souhaite diminuer la tension de sortie V_{out} alors le rendement diminue lui aussi. Dans [43], un régulateur linéaire permettant d'appliquer plusieurs valeurs de tensions d'alimentation est proposé. La Figure 2.10 représente le régulateur linéaire proposée dans [43].

Convertisseurs inductifs

Un convertisseur inductif est une alimentation à découpage convertissant une tension d'entrée continue en une tension de sortie continue. Il est basé sur des composants actifs et d'une inductance permettant de stocker ou de transférer de l'énergie. Selon le convertisseur considéré, on peut baisser la valeur de la tension d'entrée (convertisseur Buck [44]) ou augmenter la valeur de la tension d'entrée (convertisseur Boost [45]) ou les deux (convertisseur Buck-Boost [46]). La Figure 2.11 illustre le convertisseur Buck proposé dans [44].

Selon les composants utilisés, le rendement d'un convertisseur inductif peut dépasser les 95%. De plus, le rendement de ce type d'actionneur ne dépend pas du ratio de conversion. Cependant, l'intégration de ce type de convertisseur reste complexe à cause de ses éléments passifs.

Convertisseurs capacitifs

Un convertisseur capacitif connu aussi sous le nom de pompe de charge, est un convertisseur de tension composé de composants actifs et de capacités pour stocker et transférer l'énergie [47]. Les capacités sont reliées en série ou en paral-

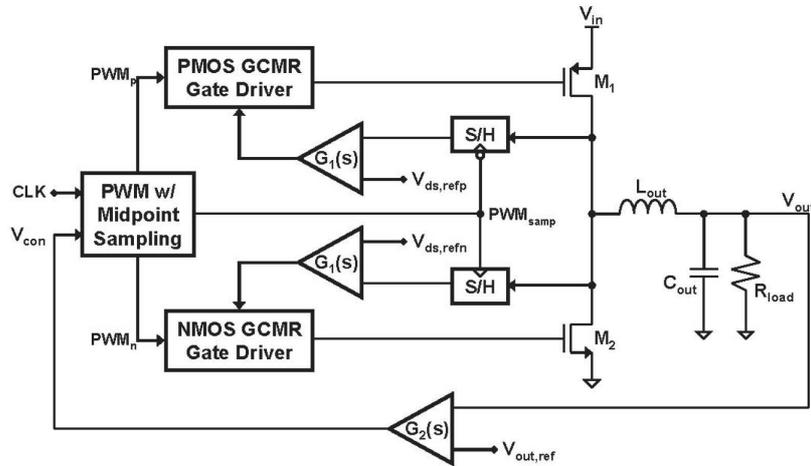


FIGURE 2.11: Schéma bloc d'un convertisseur Buck (source : [44]).

lèle. Selon le nombre de capacité et la configuration choisie (série ou parallèle), la tension d'entrée V_{in} sera multipliée ou divisée par un coefficient m pendant la phase de charge et de décharge des capacités. La sortie V_{out} est donc fonction de l'entrée avec $V_{out} = m \cdot V_{in}$ ou $V_{out} = \frac{1}{m} \cdot V_{in} = n \cdot V_{in}$. Cependant, la valeur de sortie V_{out} peut être inférieure à la valeur applicable. Le rendement η de l'actionneur étant dégradé par les pertes dues aux commutations des transistors et aux pertes par conduction car les transistors sont résistifs. Par ailleurs, ces pertes sont liées à la fréquence de commutation des capacités. Néanmoins, en augmentant la fréquence de commutation, des plus petites capacités peuvent être utilisés afin de réduire la surface du convertisseur pour un même ratio de conversion $\frac{V_{out}}{V_{in}}$. Notons que différents ratio de conversion sont disponibles sur un même convertisseur. Un compromis entre la fréquence de commutation et la taille des capacités permet de déterminer les valeurs optimales pour le convertisseur souhaité.

En réalité, on a $V_{out} = \eta \cdot m \cdot V_{in}$ ou $V_{out} = \eta \cdot n \cdot V_{in}$

La Figure 2.12 illustre l'allure typique du rendement d'un convertisseur capacitif en fonction de la tension d'alimentation en sortie de l'actionneur pour les différents ratios de conversion disponible. Si l'on considère une large gamme de tension d'alimentation applicables en sortie de l'actionneur pour un ratio de conversion, par exemple $[V_{dd_{i-1}}, V_{dd_{i+1}}]$. On peut observer sur la Figure 2.12 que sur la gamme $[V_{dd_{i-1}}, V_{dd_{i+1}}]$ le rendement du convertisseur varie énormément. En effet, ce type de convertisseur n'a un rendement élevé ($> 80\%$) que pour une gamme de tension d'alimentation réduite. Pour le reste de la gamme de tension en sortie du convertisseur, cet actionneur a un moins bon rendement.

Le Tableau 2.4 extrait de l'article [48] compare plusieurs convertisseurs capacitifs. Nous pouvons observer que le rendement maximal pour les travaux com-

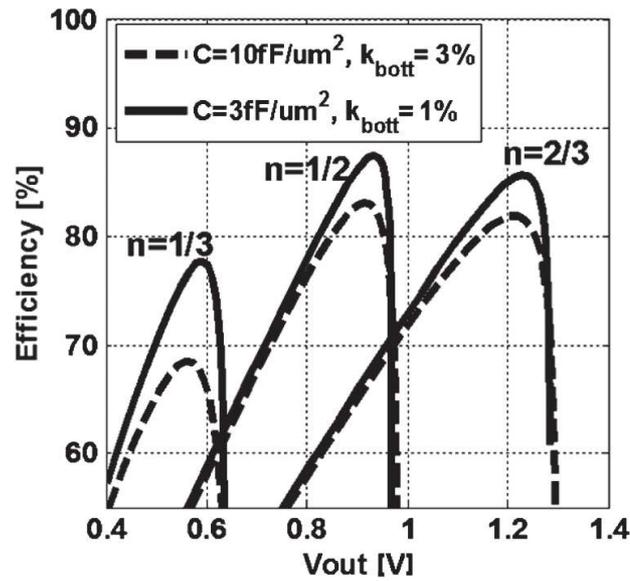


FIGURE 2.12: Rendement η attendu pour un convertisseur capacitif en fonction de la tension en sortie de l'actionneur V_{out} disposant de différents ratio de conversion n , pour différentes capacités (source : [48]).

TABLEAU 2.4: Comparaison de convertisseurs capacitifs continu-continu (*Switched-Capacitor DC-DC Converter*) (source : [48]).

Work	[49]	[50]	[51]	[48]
Technology	130nm Bulk	32nm Bulk	45nm SOI	32nm SOI
Topology	2/1 step-up	2/1 step-up	1/2 step-down	2/3, 1/2, 1/3 step-down
Capacitor	MIM	Metal Finger	Deep trench	CMOS oxide
Interleaved Phases	16	32	1	32
C_{out}	400pF ($=C_{fly}$)	0	Yes	0
Converter Area	2.25 mm^2	$6678 \mu\text{m}^2$	$1200 \mu\text{m}^2$	0.378 mm^2 (1.4% used for load)
Quoted Efficiency η	82%	60%	90%	79.16% (in 1/2 step-down)
Power density @ η	0.67 mW/mm^2	1.123 W/mm^2	2.185 W/mm^2	0.86 W/mm^2
Conductance density @ η	0.2 mS/mm^2	0.5 S/mm^2	2.421 S/mm^2	1.11 S/mm^2

parés est situé dans la gamme [60 ; 90]%. Dans [52], un convertisseur capacitif en technologie SOI 28nm ayant un rendement de plus de 90% pour une surface de $0,36 \text{ mm}^2$ est proposé. Ce convertisseur est illustré sur la Figure 2.13.

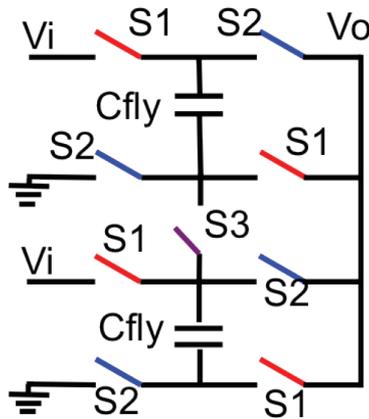


FIGURE 2.13: Schéma bloc d'un convertisseur capacitif (source : [52]).

Résumé des actionneurs de tension

Le Tableau 2.5 regroupe les avantages et les inconvénients des convertisseurs de tension. Le choix du convertisseur peut être dirigé par le compromis entre le rendement et la surface du convertisseur. Par ailleurs, le choix du convertisseur est fortement orienté selon s'il doit être intégré ou non. La vitesse de commutation de l'actionneur, la conception du circuit (basse consommation ou haute performance), la densité de puissance sont des critères parmi d'autres.

En effet, si le convertisseur est externe, le choix de l'actionneur s'orientera vers un convertisseur inductif car ce type de convertisseur a un très bon rendement. Si le convertisseur doit être intégré, il faudra choisir un convertisseur capacitif car l'intégration est compatible avec la technologie CMOS. Cependant, ce type d'actionneur a une mauvaise densité de puissance. En général, il peut fournir 1 W/mm^2 . Notons que le convertisseur linéaire n'est pas préconisé si l'on souhaite une grande gamme de tension d'alimentation car le rendement dépend linéairement du ratio de conversion $\frac{V_{out}}{V_{in}}$.

TABLEAU 2.5: Comparaison de convertisseurs de tension

Convertisseur	Avantages	Inconvénients
Linéaire	- Implémentation facile	- Rendement dépend du ratio de conversion
Inductif	- Rendement indépendant du ratio de conversion - Rendement élevé (> 90%)	- Intégration difficile due aux composants passifs non intégrables
Capacitif	- Intégration compatible CMOS sans composant externes	- Densité de puissance - Rendement dépend du ratio de conversion - Rendement (70 – 90%)

Pour pallier les problèmes de rendement des actionneurs continus et de surface lors de la duplication des actionneurs, on peut par exemple utiliser un actionneur de tension discret ou communément appelé mécanisme de V_{dd} -Hopping.

2.3.3 Tensions d'alimentation discrètes

Nous avons vu que le rendement des convertisseurs continus et en particulier celui des convertisseurs capacitifs est le facteur limitant pour le choix de tel convertisseur dans un contexte où les actionneurs doivent être intégrés au dispositif. Une partie de la consommation est due aux rendements des actionneurs. Afin de réduire la consommation, le choix des actionneurs est crucial.

Un convertisseur DC-DC continu est composé d'interrupteurs et d'un filtre qui permet de lisser la tension en sortie du convertisseur. La tension en sortie atteint la valeur désirée en moyenne.

Un mécanisme de V_{dd} -Hopping avec m niveaux de tension d'alimentation discrets peut être composé d'un seul convertisseur DC-DC capacitif disposant de m ratio de conversion. Cependant, cette solution n'est envisageable que si l'on considère un seul îlot de tension-fréquence ou VFI. En effet, si l'on dispose que d'un seul DC-DC, la gestion de la consommation ne peut être réalisée à grain fin.

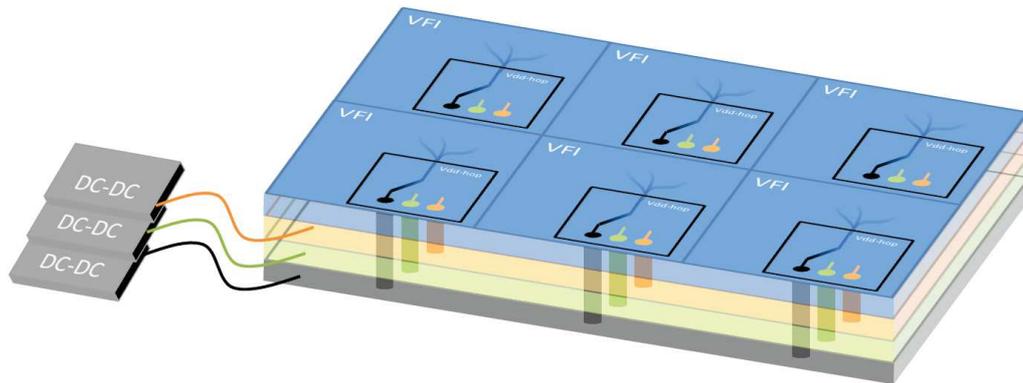


FIGURE 2.14: Schéma bloc de l'alimentation d'une architecture divisée en plusieurs îlots avec un mécanisme de V_{dd} -Hopping.

Un mécanisme de V_{dd} -Hopping disposant de n niveaux de tension d'alimentation discrets peut être composé de n convertisseurs DC-DC et d'interrupteurs pour chaque îlot tension-fréquence. Chaque convertisseur DC-DC permet d'appliquer un niveau de tension d'alimentation fixé. Les interrupteurs intégrés dans chaque îlot permettent de commuter d'une grille d'alimentation à une autre. Le fait de commuter d'une grille à une autre permet de modifier la valeur de la tension d'alimentation V_{dd} appliquée à l'îlot. Au lieu d'atteindre une tension d'alimentation en moyenne, un saut en tension permet de modifier la tension d'alimentation V_{dd} appliquée. Cependant, un circuit alimenté par un mécanisme de V_{dd} -Hopping

a besoin de $n + 1$ grilles d'alimentation, soit n grilles pour niveaux de tension d'alimentation et une grille de masse.

Les convertisseurs DC-DC pourraient être conçus de manière à maximiser leur rendement pour les tensions d'alimentation souhaitées. Sinon, le choix des valeurs de tension d'alimentation à implémenter pour un mécanisme de saut de tension ou V_{dd} -*Hopping* pourrait être orienté par l'allure du rendement du convertisseur en fonction de la tension de sortie V_{out} . Cela revient à fixer la valeur de la tension d'alimentation pour chaque convertisseur de manière à maximiser le rendement. On aurait ainsi pour la sortie de l'actionneur V_{dd} un ensemble fini de valeurs possibles avec un rendement proche de 1 pour chacune de ces valeurs au lieu d'avoir un intervalle de valeurs possibles dont le rendement varie. Si l'on considère le convertisseur capacitif de la Figure 2.12 dupliqué 3 fois, on aurait choisi d'implémenter les valeurs suivantes, $V_{dd_1} = 0,58$ V, $V_{dd_2} = 0,94$ V et $V_{dd_3} = 1,24$ V car elles correspondent aux pics de rendement des convertisseurs DC-DC décrits.

De plus, dans le cas où l'actionneur de tension est intégré sur puce (*on-chip*), la surface des actionneurs devient un critère prépondérant et ce critère est accentué si les actionneurs doivent être dupliqués. Si l'on considère une architecture divisée en plusieurs îlots, les contraintes d'implémentation physique (coût surface, coût production, rendement des actionneurs, *etc*), dues à la duplication des actionneurs dans chaque îlot tension-fréquence ou VFI, doivent être pris en compte. Si un actionneur de tension discret est implémenté, le surcoût de l'actionneur dans chaque VFI se limite à l'ajout d'un interrupteur avec différentes positions et de la partie contrôle. Dupliquer un actionneur continu dans chaque VFI n'est pas réaliste du fait de la surface nécessaire. Cette solution est non seulement moins coûteuse en termes de rendement mais aussi en termes de surface si le convertisseur est dupliqué.

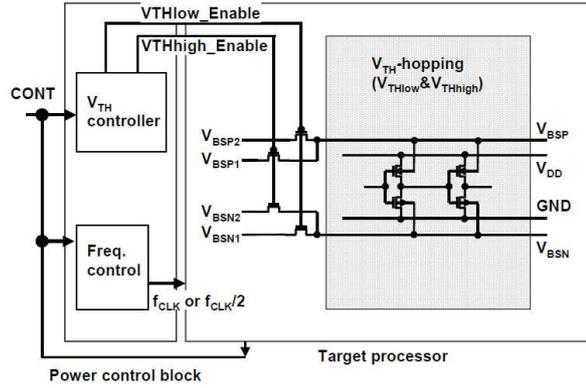
Les auteurs de [53] prouvent l'efficacité d'un DVFS discret avec 2 ou 3 niveaux de tension d'alimentation. Ils observent que le gain apporté par l'ajout de niveaux supplémentaires n'est pas significatif pour leur cas d'étude. Le nombre de niveaux de tensions d'alimentation souhaité est dirigé par le compromis consommation-performance. Notons que la gamme de fréquence dont ils disposent est relativement faible (centaines de MHz). Selon la gamme de fréquence désirée, 3 ou 4 niveaux de tensions d'alimentation peuvent être implémentés pour le mécanisme de V_{dd} -*Hopping*.

2.3.4 Tension de body bias

L'actionneur de tension de body bias V_{bb} peut être un actionneur discret comme dans [54] ou un convertisseur continu comme dans [31], [55].

Les auteurs de [54] proposent l'utilisation d'un mécanisme de V_{th} -*Hopping* tel que celui de la Figure 2.15 extraite de [54].

La table de vérité du circuit de la Figure 2.15 est donnée dans le Tableau 2.6.

FIGURE 2.15: Schéma bloc du V_{th} -Hopping (source : [54]).

Les tensions V_{BSN1} , V_{BSN2} , V_{BSP1} , V_{BSP2} correspondent dans l'ordre aux valeurs de tension de polarisation du substrat appliquées aux transistors T_{N1} , T_{N2} , T_{P1} , T_{P2} . Selon la tension V_{th} désirée, le contrôleur va mettre à 1 l'une des deux sorties (VTHhigh_Enable et VTHlow_Enable). Si VTHlow_Enable est à 1, les transistors T_{P1} , T_{N1} seront passants, la tension de seuil diminuera en valeur absolue. Par contre si VTHhigh_Enable vaut 1, les transistors T_{P2} , T_{N2} seront passants, la tension de seuil augmentera en conséquence.

TABLEAU 2.6: Table de vérité du schéma bloc de la Figure 2.15.

VTHhigh_Enable	VTHlow_Enable	V_{BSP1}	V_{BSP2}	V_{BSN1}	V_{BSN2}
0	1	1	0	1	0
1	0	0	1	0	1

Les auteurs de [56] proposent une pompe de charge dont la gamme de tension en sortie est de $[0 ; 1,3]$ V. Cette pompe de charge a un temps de réponse de l'ordre de quelques μs .

Les auteurs de [55] proposent eux une pompe de charge continue afin de compenser l'effet des variations des paramètres pour des transistors issus d'une même puce (variations With-In-Die) ou issus de différentes puces (variations Die-to-Die). Pour une fréquence cible, la tension de body bias appliquée aux transistors NMOS et PMOS est choisie afin de minimiser les courants de fuite du bloc d'exécution.

Notons que minimiser les courants de fuite pour une fréquence cible ne mène pas toujours à minimiser la consommation totale. En particulier si l'on peut agir à la fois sur la tension d'alimentation V_{dd} et la tension de body bias V_{bb} . En effet, pour une fréquence cible, le couple $(V_{dd2}, V_{bb_{min}})$ peut réduire les courants de fuite par rapport à appliquer $(V_{dd1}, V_{bb_{max}})$, avec $V_{dd1} < V_{dd2}$ et $V_{bb_{min}} < V_{bb_{max}}$. Cependant, comme $V_{dd1} < V_{dd2}$, la consommation dynamique du couple $(V_{dd2}, V_{bb_{min}})$

est supérieure à la consommation dynamique du couple $(V_{dd_1}, V_{bb_{max}})$ pour une même fréquence. Ainsi, la consommation totale du couple $(V_{dd_1}, V_{bb_{max}})$ peut être inférieure à celle du couple $(V_{dd_2}, V_{bb_{min}})$.

2.3.5 Conclusion

Indépendamment de son implémentation matérielle, les actionneurs V_{bb} et F peuvent être discrets ou continus. Les actionneurs V_{bb} et F sont considérés comme continus s'ils sont capables de produire un grand nombre de valeurs en sortie, comparé au nombre de valeurs de sortie pour l'actionneur de la tension d'alimentation V_{dd} . Par exemple, la génération d'un signal d'horloge grâce à une boucle à verrouillage de fréquence (FLL) [42] ou une boucle à verrouillage de phase (PLL) [57] est possible tandis que V_{bb} peut être fourni par un actionneur discret comme dans [54] ou par un convertisseur continu comme dans [55]. Par la suite, au moins un actionneur sera considéré comme discret car cela permet de réduire le coût d'implémentation. En effet, le surcoût en surface doit être minimisé et le rendement des actionneurs doit être maximisé. Cet aspect est encore plus crucial dans le cas d'une architecture divisée en plusieurs îlots pour lequel les actionneurs sont répliqués dans chaque îlot.

Nous avons étudié les différents actionneurs permettant de modifier les entrées du système. Nous avons aussi introduit l'actionneur de la tension de body bias. Cet actionneur devient intéressant avec les nouvelles technologies comme la technologie FD-SOI qui a une gamme de modification importante.

Nous avons vu les techniques de gestion du compromis consommation-performance au niveau circuit et système. Nous avons également étudié les avantages et inconvénients de chaque technique.

2.4 Synthèse

Les nouveaux systèmes intégrés comprennent plusieurs PE au lieu d'un seul PE. Cette combinaison de plusieurs PE nécessite une gestion de la consommation complexe. En effet, par effet de bord un PE va influencer un autre PE.

L'optimisation d'un circuit composé de plusieurs îlots tension-fréquence passe par une gestion énergétique efficace. La gestion du compromis consommation-performance à grain fin semble être essentielle. Pour cela, les actionneurs de tensions et de fréquence doivent être disponibles sur chaque îlot. Cependant, la duplication des actionneurs a un coût matériel. En utilisant un actionneur de tension discret, le surcoût peut être réduit. En effet, répliquer un mécanisme de *V_{dd}-Hopping*, revient à dupliquer la partie contrôle et l'interrupteur dans chaque îlot.

Ce chapitre a permis de donner le contexte et les contraintes d'implémentation physique de ce travail. Contrairement aux travaux précédents, les méthodes

proposées permettent de prendre en compte les contraintes d'implémentation physique liées à ce contexte. En particulier, le fait que la sortie de l'actionneur de tension d'alimentation est discrète. Le choix des valeurs en sortie de l'actionneur discret est important.

Dans [53], les auteurs ont prouvé que 2 ou 3 niveaux de tension permettaient d'avoir une bonne efficacité énergétique pour leur cas d'étude. Cependant, ils ne donnent pas de solution pour placer les valeurs de ces tensions afin de maximiser le gain en consommation. Comment le choix des valeurs des points de fonctionnement peut-il influencer la consommation du circuit ? Le placement des valeurs des points de fonctionnement peut-il être défini de manière à maximiser le gain en consommation ?

Nous avons évoqué les articles suivants [34], [37] et [38] dans la section 2.2.3. Ces articles proposent de combiner les techniques DVFS et ABB. Dans [34] et [37], les auteurs proposent la combinaison des deux techniques afin de gérer le compromis consommation-performance d'un système.

Dans [36] et [38], les auteurs proposent de combiner les techniques DVFS et ABB. L'ABB va permettre de compenser la variabilité du circuit. La combinaison des deux techniques permet de réduire la consommation mais la tension de body bias n'est pas adaptée dynamiquement. En effet, elle n'est pas modifiée en cours d'exécution mais seulement en phase de test. Il ne s'agit donc pas d'ABB à proprement parler.

Les auteurs de [58] proposent de combiner les techniques DVFS et ABB afin de minimiser l'effet de la variabilité sur les performances et la consommation d'un microprocesseur.

Les auteurs de [59] analysent la combinaison du DVFS avec une variation de la tension de body bias afin d'effectuer une optimisation multi-critère. Leur étude cherche à améliorer la fiabilité du circuit en agissant sur le « Soft Error Rate » (SER) et à réaliser un compromis entre la consommation et la performance. Cette optimisation multi-critère est difficile à résoudre. En effet, leur solution est élégante mais nécessite beaucoup de ressources pour trouver les valeurs de tension d'alimentation, de tension de body bias et de fréquence à appliquer.

Dans tous les articles présentés dans cette section, les auteurs considèrent des actionneurs continus. Les contraintes d'implémentation physique (coût surface, coût production, rendement des actionneurs, *etc*) ne sont pas identiques à notre contexte. Nous allons proposer des méthodes permettant de gérer dynamiquement les actionneurs en tenant compte des contraintes d'implémentation physique.

Le contexte technologique permet d'intégrer l'actionneur de la tension de body bias. Celui-ci permet d'augmenter la gamme de fréquence disponible sur le circuit. Dans [56], les 3 actionneurs (F, V_{dd}, V_{bb}) sont implémentés sur un même circuit. Les actionneurs choisis ont des sorties continues. Les auteurs de [53] ont prouvé que 2 ou 3 niveaux de tension étaient amplement suffisant pour leur exemple.

Ce nombre de valeur permettait d'être efficace du point de vu énergie sans un surcôté matériel important. Cependant, le contexte technologique offre une large gamme de fréquence. Dans ce cas, 3 ou 4 niveaux de tension d'alimentation discrets peuvent être nécessaires pour atteindre une bonne efficacité énergétique. Comment exploiter les opportunités de gestion de la consommation offerte par le contexte technologique ? Que se passe-t-il si l'on considère un actionneur de tension discret ? Comment peut-on exploiter (F, V_{dd}, V_{bb}) afin de maximiser le gain en consommation ?

Chapitre 3

Placement optimal de Power Mode

Ce chapitre a pour objectif de donner une méthode permettant de placer les points de fonctionnement ou Power Mode (PM) afin de maximiser le gain en consommation. Un PM est défini comme un couple fréquence-puissance (F, P_{tot}) avec F la fréquence d'horloge et P_{tot} la puissance totale consommée associée à F . Cette notion va nous permettre de nous affranchir des degrés de liberté (DoF, *Degree of Freedom* en anglais) sur lesquels nous allons agir. Un DoF est défini comme une entrée du système sur lequel on peut agir pour modifier le comportement de ce système. Par exemple, il peut s'agir de la tension d'alimentation V_{dd} à laquelle est alimentée le circuit, de la fréquence d'horloge F , de la tension de polarisation du body V_{bb} . La tension de polarisation du body V_{bb} , plus connue sous le nom de tension de body bias, correspond à la tension sur laquelle on doit agir si l'on veut modifier la tension de seuil des transistors, V_{th} .

Chaque transistor possède des caractéristiques typiques comme sa concentration en porteurs (dopage), sa longueur, sa largeur, sa tension de seuil, etc. La tension de seuil est définie comme étant la tension entre la grille et la source pour laquelle la zone d'inversion apparaît, c'est-à-dire la création du canal de conduction entre la source et le drain. Cette tension se note V_{th} , th étant l'abréviation de *threshold* en anglais (seuil). Lorsque la tension grille-source est inférieure à la tension de seuil, on dit que le transistor est bloqué, il ne conduit pas. Dans le cas contraire, on dit qu'il est passant, il conduit le courant entre le drain et la source. Notons que cette tension dépend de la technologie mais elle peut être légèrement différente à cause de la variabilité. Par exemple, les variations de procédés de fabrication pour des transistors issus d'une même puce (variations With-In-Die) ou issus de différentes puces (variations Die-to-Die) ont une influence sur V_{th} .

Dans ce chapitre, les dispositifs étudiés seront équipés de 2 degrés de liberté

ou DoF, voir la Figure 3.1. Cependant, l'étude reste valable quel que soit le nombre de DoF. L'objectif est de choisir les PM afin de maximiser le gain en consommation du circuit et plus particulièrement du processeur ou du cœur de calcul (Processing Element PE). Les combinaisons existantes à ce jour de 2 degrés de liberté permettant de gérer le compromis consommation-performance d'un circuit sont (F, V_{dd}) [60] [61], (F, V_{bb}) [62] [63] et (V_{dd}, V_{bb}) [64] [65]. Nous allons tout d'abord présenter un récapitulatif de l'existant dans le cas où les 2 degrés de liberté considérés sont continus. Nous allons ensuite spécifier l'hypothèse de travail principale et les définitions nécessaires à la compréhension du travail qui va être exposé. Plus particulièrement, un dispositif présentant les 2 DoF discrétisés (F et V_{dd}) sera notre cas d'étude. Nous allons évaluer dans ce cas si l'addition d'un PM supplémentaire représente une valeur ajoutée lorsque deux PM sont déjà disponibles. Ici, valeur ajoutée signifie que le gain en consommation justifie le surcoût matériel et la complexité du contrôle associée. Cette approche sera étendue à un nombre supérieur de PM. Une conclusion terminera ce chapitre.

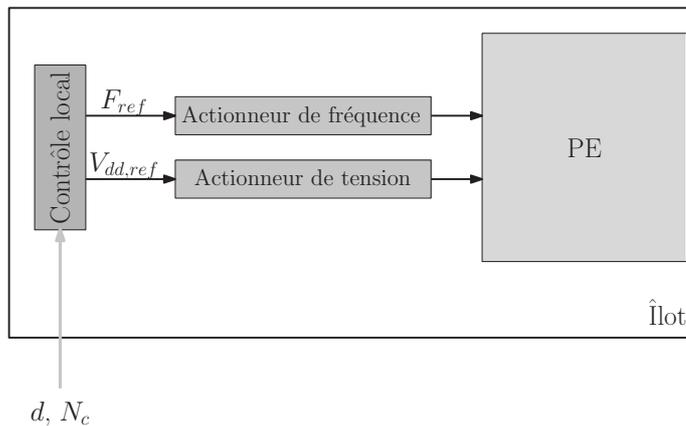


FIGURE 3.1: Schéma d'un îlot avec 2 actionneurs. d est le temps imparti pour exécuter la tâche et N_c correspond au nombre de cycles d'horloge nécessaires pour exécuter la tâche.

Ce chapitre a donné lieu à une publication scientifique internationale dans le domaine de l'électronique FTFC [66].

Nous allons maintenant présenter quelques exemples issus de l'état de l'art de gestion de la consommation dans le cas où 2 actionneurs continus sont considérés. Les entrées du système permettant de gérer la consommation sont la tension d'alimentation V_{dd} et la fréquence d'horloge F du circuit.

3.1 Gestion de la consommation : DVFS

L'article [67] utilise la Q-fonction pour optimiser la consommation d'un dispositif pour une performance donnée. Cette fonction correspond à un ratio de

puissance sur la vitesse. Leur cas d'étude est un disque dur. Les auteurs de [67] prouvent que le profil optimal de vitesse est atteint quand la relation Q-fonction - vitesse est convexe. De plus, les auteurs affirment que les dispositifs satisfont par défaut la propriété de convexité. Cette affirmation est correcte si le circuit est géré par 2 actionneurs (*e.g.* (F, V_{dd}) ou (F, V_{bb}) ou (V_{dd}, V_{bb})). Néanmoins, pour les technologies avancées, lorsque 3 actionneurs sont combinés (dont au moins un discret), cette affirmation n'est pas toujours valable. De plus, il est prouvé dans [67] que la consommation optimale est atteinte lorsque les fréquences voisines de la fréquence cible sont utilisées pour exécuter la tâche si les points de la Q-fonction(vitesse) forment un ensemble convexe.

Les auteurs de [26] proposent une technique de DVFS qui prend en compte les variations de procédés de fabrication. Lors de la phase de test, le procédé de fabrication de chaque îlot de tension-fréquence ou VFI (Voltage/Frequency Island) est caractérisé *via* la mesure des courants de fuite. Si les courants de fuite du VFI sont élevés, la tension d'alimentation et la fréquence de ce VFI vont être diminuées.

Les auteurs de [20] ont conçu un DVFS pour améliorer l'efficacité énergétique d'un circuit. L'idée sous-jacente est de fournir le plus grand pic de fréquence possible pour exécuter les tâches ayant de fortes contraintes d'exécution puis la fréquence et la tension d'alimentation sont réduites afin de maximiser la durée de vie de la batterie. Notons que la tension d'alimentation est adaptée à la fréquence appliquée.

Les articles présentés sont composés d'un actionneur de fréquence F et d'un actionneur de tension d'alimentation V_{dd} continu. Nous portons un intérêt particulier à la gestion de la consommation en utilisant un actionneur de tension d'alimentation discret. Nous allons donner le contexte nécessaire à notre étude.

3.2 Hypothèses de travail et définitions

Un PM est obtenu en modifiant la valeur des DoF. En modifiant soit la tension d'alimentation V_{dd} , soit la tension de body bias V_{bb} , on agit alors sur la fréquence d'horloge maximale F_{max} qui peut être appliquée sans produire d'erreur temporelle. Une erreur temporelle est occasionnée lorsque la fréquence d'horloge appliquée est supérieure à la fréquence maximale atteignable. Dans ce cas, le signal d'entrée d'une ou plusieurs bascules est affecté à la bonne valeur mais l'événement se produit au mauvais moment, c'est-à-dire en dehors de l'intervalle de temps spécifié (période d'horloge). Dans ce cas, il faudra ré-exécuter les instructions depuis le début, correspondant à une perte à la fois en termes de performance et de consommation. Cela signifie que l'efficacité énergétique du dispositif est en baisse.

Dans les exemples qui vont suivre, la tension d'alimentation V_{dd} sera modifiée

afin d'agir sur la fréquence maximale atteignable F_{max} .

Dans la suite de ce travail, la fréquence maximale qui peut être appliquée au système sans produire d'erreur temporelle est supposée être atteinte pour la tension d'alimentation V_{dd} appliquée. Notons que le critère à minimiser est la puissance consommée par le dispositif, ce qui nous amène à poser l'hypothèse suivante.

Hypothèse 1. *La consommation P_{tot} d'un circuit synchrone est fonction des entrées du système F et V_{dd} . Pour une fréquence cible fixée, il existe une méthode pour déterminer la valeur de la tension d'alimentation V_{dd} permettant de minimiser P_{tot} pour le couple (F, V_{dd}) . En d'autres termes, pour une entrée donnée, on sait comment appliquer l'autre entrée de manière à minimiser la consommation P_{tot} .*

Cette première hypothèse assure que pour une fréquence F_{don} donnée, il existe une technique permettant de déterminer la tension d'alimentation optimale $V_{dd_{nom}}$ qui minimise la consommation P_{tot} du dispositif pour ce couple d'entrées tout en garantissant la fonctionnalité.

Notons que cette fréquence F_{don} peut être atteinte en appliquant une tension d'alimentation $V_{dd_{sup}}$, avec $V_{dd_{sup}} > V_{dd_{nom}}$. Cependant, l'efficacité énergétique est moindre car la fréquence F_{don} est inférieure à la fréquence maximale atteignable sans produire d'erreur temporelle pour la tension d'alimentation $V_{dd_{sup}}$ appliquée. Rappelons que la consommation dynamique P_{dyn} dépend linéairement de la fréquence et quadratiquement de la tension d'alimentation. Cela revient à réduire la performance pour cette tension d'alimentation et donc à une perte de performance et de consommation. Le dispositif perd ainsi en efficacité énergétique. Pour une fréquence F_{don} donnée, le but est d'appliquer la plus petite tension d'alimentation V_{dd} permettant d'atteindre cette fréquence.

Il faut appliquer la tension d'alimentation minimale V_{dd} permettant d'atteindre la fréquence F_{max} . Cependant, une marge de sécurité est préconisée dans le choix de F_{max} pour éviter de produire des erreurs temporelles si une des entrées du système ou un des paramètres (perturbations) varie. En effet, les variations de procédés de fabrication, de tension et de température ou PVT (Process Voltage Temperature) vont affecter la fréquence maximale F_{max} que l'on peut appliquer pour une tension V_{dd} . Si la fréquence F_{max} atteignable diminue à l'instant t à cause de variations PVT, la fréquence F appliquée à l'instant t est supérieure à la fréquence F_{max} atteignable après les variations PVT. Ainsi, des erreurs temporelles peuvent apparaître si la fréquence F appliquée correspond à la fréquence F_{max} atteignable avant ces variations PVT. Les variations de procédés de fabrication sont considérées comme statiques car la dynamique de variation du vieillissement est de l'ordre de l'année. Par contre, les variations de température et de tension sont considérées dynamiques. Les variations de tension sont dues à des chutes de potentiel par rapport à la valeur nominale de tension provoquées

par des appels de courant. Ce phénomène de chutes de tension est plus connu sous le nom d'*IR-drop*.

Des capteurs de fonctionnalité, par exemple les capteurs TSM (Timing Slack Monitor) [68] ou les capteurs de type Razor [69], permettent de vérifier si des erreurs temporelles sont produites. Ces capteurs peuvent permettre de vérifier l'hypothèse 1. Dans le cas où un capteur TSM est disponible sur le dispositif, on peut fixer V_{dd} , puis augmenter la fréquence F jusqu'à F_{max} sans produire d'erreur temporelle. L'application de F_{max} garantit la minimisation de la consommation P_{tot} en maximisant la performance pour l'état actuel du circuit, *i.e.* la tension V_{dd} appliquée, la température T° et l'activité A subies.

Par la suite on considérera que l'activité A représente le nombre de transistors qui commutent dans un intervalle de temps, par exemple un cycle d'horloge. La fréquence qu'il faut appliquer au dispositif pour finir la tâche dans le temps imparti est appelée la fréquence cible F_{target} . Il s'agit de la fréquence qui satisfait la performance requise pour exécuter la tâche T dans le temps imparti. Dans le cas où une exécution est constituée de 2 parties avec 2 activités moyennes différentes, on considère ce travail comme 2 tâches distinctes. Cette décomposition de tâche va nous permettre d'exécuter une tâche en appliquant au maximum deux PM différents. Cela nous permet de réduire le coût de commutation des actionneurs par rapport à appliquer plus de deux PM pour exécuter la tâche.

Le choix des actionneurs utilisés permettant de gérer dynamiquement les entrées du circuit est essentiel. En effet, les pertes en consommation du circuit sont aussi liées au rendement des actionneurs, en particulier celui de la tension d'alimentation V_{dd} , tel que nous l'avons vu dans le chapitre 2.

Pour des contraintes d'implémentation physique, le choix des actionneurs s'oriente vers des actionneurs discrets, en particulier celui de la tension d'alimentation V_{dd} (par exemple un mécanisme de *V_{dd}-Hopping* [28]).

Ces aspects permettent de formaliser l'hypothèse de travail numéro 2 concernant le choix des actionneurs considérés discrets car cela permet de réduire le coût d'implémentation physique. En effet, le surcoût en surface doit être minimisé tandis que le rendement des actionneurs doit être maximisé.

Hypothèse 2 (a). *La tension d'alimentation V_{dd} et la fréquence d'horloge F sont fournies par des actionneurs qui sont capables de fournir en sortie des valeurs discrètes. Le nombre possible de valeurs de sortie est limité (*i.e.* quelques unités).*

L'hypothèse de travail numéro 3 concerne la connaissance *a priori* de la consommation totale P_{tot} du dispositif. La consommation P_{tot} d'un tel dispositif est fonction des valeurs F et V_{dd} délivrées par les actionneurs mais aussi de la température du circuit T° , de l'activité A et d'autres paramètres physiques. On considère que la mesure ou l'estimation de la consommation totale P_{tot} a été réalisée au préalable. Par exemple, la consommation peut être obtenue en phase

de conception par simulation ou elle peut être estimée ou mesurée par l'intermédiaire de capteurs en cours d'exécution. On peut par exemple coupler des capteurs de tension [70] à un capteur de courant [71] pour déterminer la consommation du circuit.

Hypothèse 3. *La consommation totale P_{tot} du dispositif peut être déterminée.*

On suppose que le circuit exécute une tâche T et qu'il connaît les exigences en performance, ce qui est le cas pour les systèmes temps-réel. Pour d'autres systèmes, les informations relatives aux exigences d'exécution (performance requise) peuvent être fournies par un système de plus haut niveau (par exemple [72]).

D'un point de vue applicatif, on suppose que le circuit exécute une tâche déterminée et que cette tâche a une contrainte de performance, correspondant au temps imparti (d) pour exécuter la tâche. Pour que la tâche soit finie à temps, la performance moyenne doit correspondre à la performance du circuit cadencé avec une horloge à la fréquence cible F_{target} . La fréquence cible F_{target} correspond à la fréquence d'horloge qui devrait être appliquée au circuit pour exécuter la tâche avec la performance requise. On considère que des tâches se succèdent et que la fréquence F_{target} varie en fonction de la tâche considérée.

Hypothèse 4. *La tâche T requiert N_C cycles d'horloge pour être exécutée dans le temps imparti d . Ainsi, la tâche doit être exécutée à une fréquence équivalente à F_{target} . De plus, au moins deux paramètres parmi N_C , d et F_{target} sont connus.*

L'ensemble des hypothèses nécessaires à la mise en place du cadre du travail étant défini, nous pouvons désormais préciser le problème :

Problème. *On souhaite choisir la position du (des) PM intermédiaire(s) afin de maximiser le gain en consommation par rapport à une situation où l'on ne dispose que de deux PM pour exécuter les tâches T , en considérant les contraintes d'implémentation physique et les contraintes applicatives.*

3.3 Gestion d'un circuit possédant trois PM

Dans cette section, le choix des PM à implémenter va s'effectuer avec 2 DoF, à savoir F et V_{dd} . L'objectif est de choisir les PM à implémenter sur le dispositif afin de maximiser le gain en consommation sous contraintes de performance. On propose une stratégie de positionnement des PM à travers les deux entrées F et V_{dd} .

L'ajout d'un PM va être évalué dans le cas où 2 PM sont déjà disponibles sur le circuit. Pour l'exemple étudié, nous allons en particulier analyser si l'ajout d'un niveau de tension d'alimentation pour le mécanisme V_{dd} -Hopping est efficace d'un point de vue énergétique lorsque deux niveaux de tension d'alimentation sont déjà disponibles. Rappelons que, efficace signifie que le gain en consommation justifie

le surcoût matériel et la complexité du contrôle associée à l'ajout d'un niveau supplémentaire de tension. La même étude peut être menée en changeant les DoF (par exemple avec (F, V_{bb})), ou en ajoutant un DoF.

La Figure 3.2 présente un exemple de positionnement de PM dans le plan (F, P_{tot}) . On considère que PM_{low} et PM_{high} sont fixés durant la conception pour avoir une gamme de fonctionnement $[F_{low} ; F_{high}]$. PM_{low} (resp. PM_{high}) est obtenu en appliquant la plus petite (resp. grande) valeur de fréquence F_{low} (resp. F_{high}). Il faut placer le(s) PM intermédiaire(s) afin de maximiser le gain en consommation suivant certains critères. Les critères peuvent concerner par exemple des contraintes applicatives ou la connaissance de la densité de probabilité que la fréquence cible F_{target} soit demandée. Rappelons que la notion de PM correspond à un couple de fréquence-puissance et que les entrées (F, V_{dd}) permettant d'appliquer un PM sont choisies de manière à ne pas produire d'erreur temporelle.

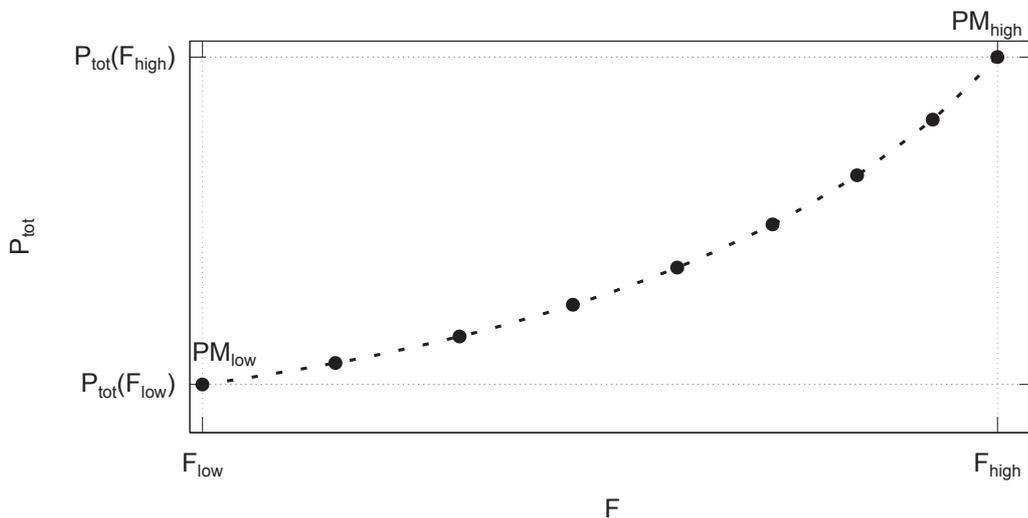


FIGURE 3.2: Exemple de positionnement des PM dans le plan (F, P_{tot}) .

Notons que le coût de commutation des actionneurs n'a pas été pris en compte dans le travail qui va suivre car il dépend de plusieurs facteurs. Le coût de commutation est lié au choix des actionneurs, c'est-à-dire à la dynamique et à la consommation des actionneurs.

3.3.1 Contexte

Le travail qui suit permet d'évaluer si l'ajout d'un PM intermédiaire (PM_{middle}), lorsque deux PM sont déjà disponibles sur le dispositif (PM_{low} et PM_{high}), permet de réduire la consommation. On propose de positionner le PM intermédiaire suivant certains critères. S'il s'agit d'un circuit dédié à une seule application ou bien d'un circuit générique qui peut être utilisé pour différentes applications, les

besoins en termes de performance ne seront pas identiques pour les deux circuits. En conséquence, le choix du PM intermédiaire ne sera sûrement pas identique pour ces deux dispositifs.

Dans le cas d'étude qui va suivre, les PM extrêmes, c'est-à-dire PM_{low} et PM_{high} sont fixés. Rappelons que PM_{low} correspond au PM ayant la fréquence la plus petite (la fréquence d'un PM correspond à la fréquence maximale qui peut être appliquée sans produire d'erreur temporelle). De même, PM_{high} correspond au PM dont la fréquence est la plus grande. Cette étude permet donc d'évaluer si l'ajout d'un PM intermédiaire est pertinent. Ce PM intermédiaire peut être obtenu par l'ajout d'un niveau de tension pour un mécanisme de *V_{dd}-Hopping* qui en possède déjà deux. Cette étude permet aussi de déterminer la valeur du PM intermédiaire qui permet de maximiser le gain en consommation sur le circuit considéré.

Le cas d'étude est le circuit LoCoMoTIV (Local Compensation of Modern Technology Induced Variability) [73]. Il s'agit d'un circuit implémenté en technologie CMOS BULK 32nm. Ce prototype d'architecture MPSoC disposant de mécanismes de gestion de la variabilité et de la consommation à grain fin est composé de 4 PE (STXP70). La moitié des PE intègrent leurs propres actionneurs de tension et de fréquence, la partie contrôle associée et des capteurs afin de gérer la consommation à grain fin avec la technique AVFS. L'autre moitié des PE exécutent les tâches avec PM_{high} .

Les valeurs pour ce circuit sont extraites de simulations STA (*Static Timing Analysis*) d'un seul PE. Les entrées de la simulation sont la tension d'alimentation V_{dd} , la température T° subie par le circuit, l'activité A et le temps de simulation. La simulation permet d'obtenir la consommation totale P_{tot} , la consommation statique P_{stat} et la fréquence maximale atteignable F_{max} pour V_{dd} et T° . Nous avons besoin de valeurs de consommation en fonction de la fréquence pour différentes tensions d'alimentation afin d'évaluer la contribution de l'ajout d'un PM intermédiaire à l'aide d'un outil d'analyse numérique, MATLAB.

Rappelons que nous considérons que les PM extrêmes sont fixés par des contraintes d'implémentation physique ou des contraintes d'exécution. Les tensions minimale et maximale que l'on peut appliquer au circuit sont définies à la conception du circuit. Cependant, les tensions choisies peuvent légèrement différer. En effet, la tension de PM_{low} peut être supérieure à la tension minimale à partir de laquelle le circuit fonctionne. De même, la tension de PM_{high} peut être inférieure à la tension maximale à partir de laquelle le circuit risque de subir des dommages ou un vieillissement prématuré. En outre, on peut essayer d'appliquer une tension plus basse que la tension minimale mais on ne peut pas dépasser la tension maximale du circuit, au risque de le rendre inutilisable.

Le profil de consommation en fonction de la fréquence est obtenu pour une activité A et une température T° constantes. Les PM extrêmes, soit PM_{high} et

PM_{low} , sont fixés aux valeurs suivantes, $PM_{high} = (929 \text{ MHz} , 191,92 \text{ mW})$ et $PM_{low} = (487 \text{ MHz} , 24,63 \text{ mW})$. Ces valeurs seront fixes pour tous les exemples suivants.

Lorsque la fréquence cible F_{target} n'est pas applicable directement sur le circuit c'est-à-dire $F_{target} \neq \{F_1 ; F_2 ; \dots\}$, la tâche est exécutée en mode *hopping* avec la fréquence supérieure et la fréquence inférieure à F_{target} . Sur la Figure 3.3, α représente le ratio de temps que l'on va passer à la fréquence F_2 (voir l'équation (3.1)), le reste du temps $(1 - \alpha) \cdot d$ la tâche étant exécutée à la fréquence F_1 , avec $F_1 < F_2$.

$$\alpha \cdot F_2 + (1 - \alpha) \cdot F_1 = F_{target} \Rightarrow \alpha = \frac{F_{target} - F_1}{F_2 - F_1} \quad (3.1)$$

Dans notre cas d'étude, si deux PM sont disponibles $F_1 = F_{low}$ et $F_2 = F_{high}$. Si trois PM sont disponibles $F_1 = F_{low}$, $F_2 = F_{middle}$ et $F_3 = F_{high}$.

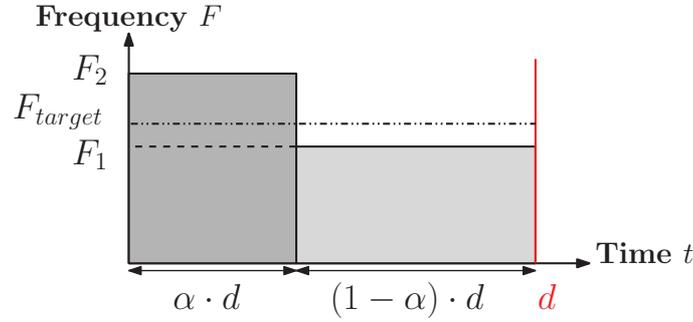


FIGURE 3.3: *Hopping* entre deux fréquences pour atteindre une performance équivalente à la fréquence F_{target} avec α le rapport cyclique et d le temps imparti connu pour la tâche considérée.

3.3.2 Gain en consommation lié à l'ajout d'un PM intermédiaire

Nous allons évaluer le gain en consommation obtenu par l'ajout d'un PM intermédiaire. La consommation d'une tâche T exécutée avec PM_{high} et PM_{low} va être comparée à l'exécution de la même tâche T , c'est-à-dire avec la même F_{target} avec deux PM parmi PM_{high} , PM_{middle} et PM_{low} . Rappelons, que le coût de commutation des actionneurs n'est pas pris en compte.

Le gain en consommation G_{3PM} pour F_{target} est obtenu avec l'équation suivante :

$$G_{3PM}(F_{target})(\%) = 100 \cdot \frac{P_{tot2PM}(F_{target}) - P_{tot3PM}(F_{target})}{P_{tot2PM}(F_{target})} \quad (3.2)$$

avec $G_{3PM}(F_{target})$ le gain en consommation en ajoutant un PM pour la fréquence cible F_{target} en pourcentage, $P_{tot2PM}(F_{target})$ la consommation associée à l'exécution d'une tâche avec F_{target} lorsque deux PM sont disponibles, $P_{tot3PM}(F_{target})$ la

consommation de l'exécution d'une tâche avec F_{target} lorsqu'un PM intermédiaire est ajouté, soit trois PM.

Lors de l'exécution d'une tâche, différentes situations peuvent être rencontrées. Le gain en consommation est alors évalué de la manière suivante :

- la fréquence F_{target} est comprise entre les fréquences de PM_{low} et PM_{middle} . La consommation en exécutant la tâche avec PM_{low} et PM_{middle} est comparée à la consommation en exécutant la tâche avec PM_{low} et PM_{high} . Les consommations $P_{tot_{2PM}}(F_{target})$ et $P_{tot_{3PM}}(F_{target})$ sont obtenues avec le système d'équations (3.3).
- F_{target} est comprise entre les fréquences de PM_{middle} et PM_{high} . La consommation de l'exécution de la tâche avec PM_{middle} et PM_{high} est comparée à l'exécution de la tâche avec PM_{low} et PM_{high} . $P_{tot_{2PM}}(F_{target})$ et $P_{tot_{3PM}}(F_{target})$ sont obtenues avec le système d'équations (3.4).
- la fréquence F_{target} correspond à la fréquence de PM_{middle} . La consommation de l'exécution de la tâche avec PM_{middle} est comparée à la consommation de l'exécution de la tâche avec PM_{low} et PM_{high} . $P_{tot_{2PM}}(F_{target})$ et $P_{tot_{3PM}}(F_{target})$ sont obtenues avec le système d'équations (3.5).

La détermination du rapport cyclique α correspondant au ratio de temps où la tâche est exécutée à la fréquence supérieure, est illustrée sur la Figure 3.3 et déterminée par l'équation (3.1).

$$\left\{ \begin{array}{l} F_{low} < F_{target} < F_{middle} \\ F_{target} = \alpha_1 \cdot F_{middle} + (1 - \alpha_1) \cdot F_{low} = \alpha_2 \cdot F_{high} + (1 - \alpha_2) \cdot F_{low} \\ P_{tot_{3PM}}(F_{target}) = \alpha_1 \cdot P_{tot}(F_{middle}) + (1 - \alpha_1) \cdot P_{tot}(F_{low}) \\ P_{tot_{2PM}}(F_{target}) = \alpha_2 \cdot P_{tot}(F_{high}) + (1 - \alpha_2) \cdot P_{tot}(F_{low}) \end{array} \right. \quad (3.3)$$

$$\left\{ \begin{array}{l} F_{middle} < F_{target} < F_{high} \\ F_{target} = \alpha_3 \cdot F_{high} + (1 - \alpha_3) \cdot F_{middle} = \alpha_4 \cdot F_{high} + (1 - \alpha_4) \cdot F_{low} \\ P_{tot_{3PM}}(F_{target}) = \alpha_3 \cdot P_{tot}(F_{high}) + (1 - \alpha_3) \cdot P_{tot}(F_{middle}) \\ P_{tot_{2PM}}(F_{target}) = \alpha_4 \cdot P_{tot}(F_{high}) + (1 - \alpha_4) \cdot P_{tot}(F_{low}) \end{array} \right. \quad (3.4)$$

$$\left\{ \begin{array}{l} F_{target} = F_{middle} = \alpha_5 \cdot F_{high} + (1 - \alpha_5) \cdot F_{low} \\ P_{tot_{3PM}}(F_{target}) = P_{tot}(F_{middle}) \\ P_{tot_{2PM}}(F_{target}) = \alpha_5 \cdot P_{tot}(F_{high}) + (1 - \alpha_5) \cdot P_{tot}(F_{low}) \end{array} \right. \quad (3.5)$$

Nous allons donner quelques exemples de PM intermédiaire et évaluer le gain en consommation apporté par l'ajout de ces PM intermédiaires.

Illustrations du gain en ajoutant un PM intermédiaire

Le PM intermédiaire choisi pour cet exemple est associé à une fréquence de 820 MHz. Les valeurs de la consommation totale du circuit P_{tot} , la fréquence d'horloge F et la tension d'alimentation V_{dd} associées à chaque PM, pour cet exemple, sont données dans le Tableau 3.1.

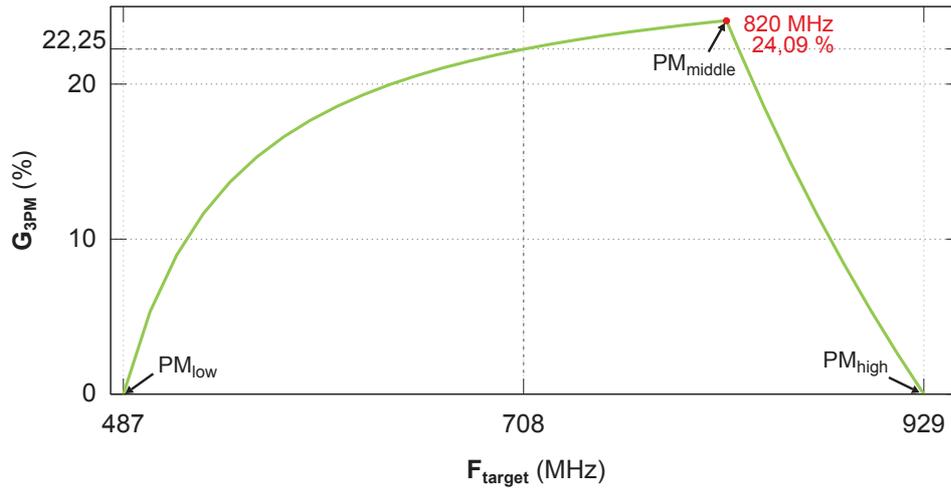


FIGURE 3.4: Gain en consommation G_{3PM} en ayant 3 PM disponibles par rapport à 2 PM disponibles, avec la fréquence de PM_{middle} choisie à 820 MHz.

TABLEAU 3.1: Valeurs de la consommation P_{tot} , de la fréquence F et de la tension d'alimentation V_{dd} associées à chaque PM, avec la fréquence de PM_{middle} choisie à 820 MHz.

PM	P_{tot} (mW)	F (MHz)	V_{dd} (mV)
PM _{high}	191,92	929	1 200
PM _{middle}	114,37	820	1 000
PM _{low}	24,63	487	600

La Figure 3.4 illustre le gain en consommation en ayant 3 PM disponibles par rapport à 2 PM disponibles sur le même circuit en fonction de la fréquence cible F_{target} . Par exemple, si F_{target} est égale à 708 MHz, la tâche est exécutée avec PM_{low} et PM_{middle}, si 3 PM sont disponibles. Lorsque 2 PM sont disponibles sur le circuit, la même tâche est exécutée avec PM_{low} et PM_{high} avec une fréquence moyenne égale à F_{target} . Le gain en consommation G_{3PM} en ajoutant le PM intermédiaire à 820 MHz est alors de 22,25% pour $F_{target} = 708$ MHz, voir la Figure 3.4.

Le maximum de gain en consommation est atteint pour la fréquence cible égale à la fréquence de PM_{middle} , soit 820 MHz. Dans ce cas, la consommation est réduite de 24,09%, voir la Figure 3.4.

Nous pouvons également lire la Figure 3.4 de la manière suivante : le gain en consommation en ayant 3 PM disponibles par rapport à 2 PM disponibles est d'au moins 15% pour F_{target} comprise entre [560 ; 840] MHz pour ce PM intermédiaire.

La Figure 3.5 illustre le gain en consommation pour un autre cas. Dans ce cas, PM_{middle} correspond à la fréquence de 652 MHz. Les valeurs de la consommation P_{tot} , la fréquence F et la tension d'alimentation V_{dd} associées à chaque PM, pour le second exemple, sont données dans le Tableau 3.2.

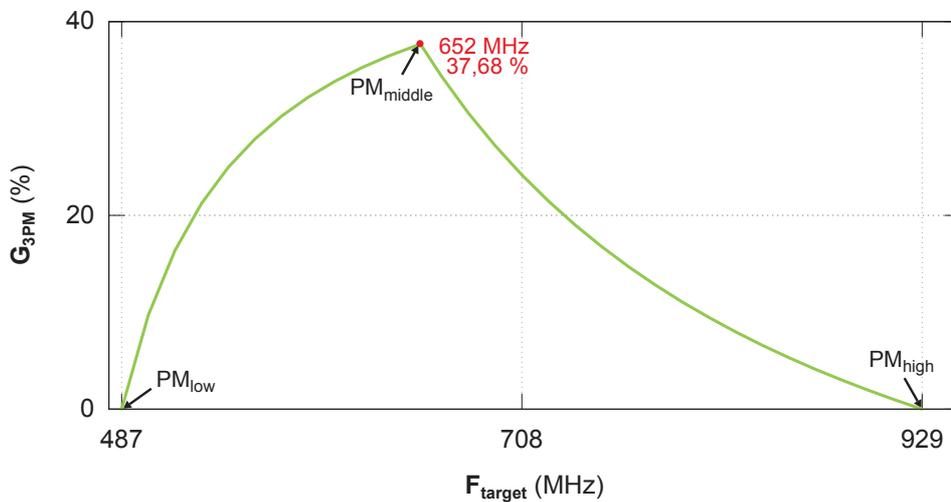


FIGURE 3.5: Gain en consommation G_{3PM} en ayant 3 PM disponibles par rapport à 2 PM disponibles, avec la fréquence de PM_{middle} choisie à 652 MHz.

TABLEAU 3.2: Valeurs de la consommation P_{tot} , de la fréquence F et de la tension d'alimentation V_{dd} associées à chaque PM, avec la fréquence de PM_{middle} choisie à 652 MHz.

PM	P_{tot} (mW)	F (MHz)	V_{dd} (mV)
PM_{high}	191,92	929	1 200
PM_{middle}	54,27	652	775
PM_{low}	24,63	487	600

Pour cet exemple, le maximum de gain en consommation est de 37,68%. Il est atteint pour la fréquence de PM_{middle} , soit 652 MHz. De plus, un gain d'au moins

15% peut être attendu pour F_{target} comprise entre [516 ; 766] MHz en ajoutant un PM intermédiaire ayant une fréquence de 652 MHz.

On observe que le gain est fortement dépendant de la position de PM_{middle} . Nous allons donc évaluer le gain pour toutes les possibilités de PM intermédiaire.

Évaluation du gain pour toutes les possibilités de PM intermédiaire

Ici, nous avons réalisé la même étude que pour les exemples précédents mais pour tous les PM intermédiaires compris entre PM_{low} et PM_{high} . La Figure 3.6 présente le gain max G_{max} correspondant à chacun des PM intermédiaires testés. Cela revient à obtenir le tracé du gain en consommation en fonction de la fréquence F_{middle} pour tous les PM intermédiaires possibles (voir les Figures 3.4 et 3.5 qui correspondent à des cas particuliers).

Nous observons que le maximum des gains maximaux est de 37,68% et qu'il est obtenu pour F_{middle} égale à 652 MHz. La fréquence du maximum, soit 652 MHz ne correspond pas au milieu de la gamme [F_{low} ; F_{high}], soit 708 MHz. La tension du maximum, soit $V_{dd_{middle}} = 775$ mV, ne correspond pas non plus à la moyenne de la gamme de tension d'alimentation, *i.e.* $\frac{V_{dd_{high}} - V_{dd_{low}}}{2} + V_{dd_{low}}$, soit 900 mV.

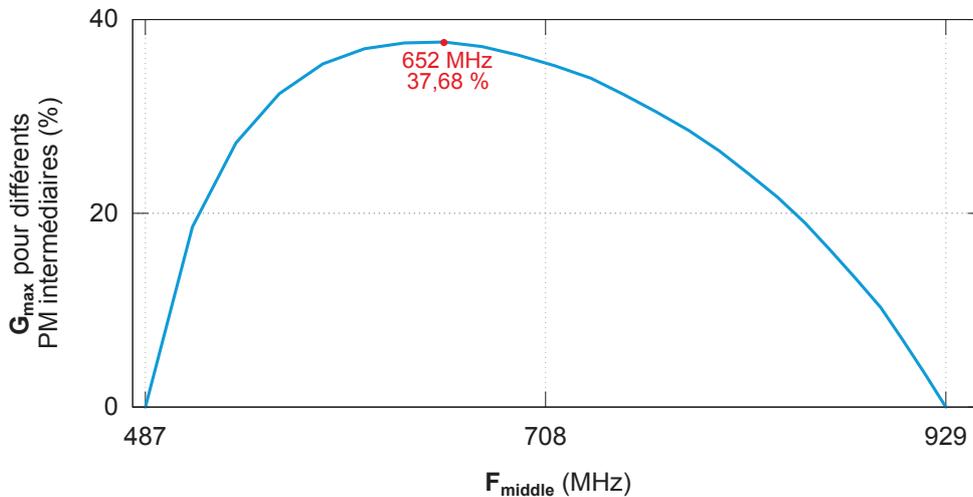


FIGURE 3.6: Gain en consommation maximal G_{max} obtenu par l'ajout d'un PM intermédiaire pour toutes les possibilités en fonction de la fréquence du PM intermédiaire F_{middle} .

Nous avons évalué le gain en consommation pour différents PM intermédiaires ainsi que le gain maximal apporté. Le profil de gain en fonction de la fréquence cible ne permet pas à lui seul de déterminer où le PM intermédiaire doit être placé

afin de maximiser le gain. Comment placer le PM intermédiaire afin de maximiser le gain en consommation ? Quelle métrique permet de maximiser le gain, est-ce le gain maximal ou le gain moyen ?

3.3.3 Gain en consommation moyen

Le gain en consommation et la moyenne de gain en consommation dépend de la position du PM intermédiaire.

Lors de la conception du circuit, si l'on sait que le circuit est dédié à l'exécution d'un certain nombre de tâches récurrentes, la gamme de fréquence nécessaire pour exécuter ces tâches peut être déterminée et on peut ainsi choisir le PM intermédiaire de manière à maximiser le gain en consommation pour cette gamme de fréquence. Cela revient à maximiser l'aire dans le plan gain en consommation avec 3 PM par rapport à 2 PM en fonction de la fréquence cible. Rappelons que les PM extrêmes sont fixés lors de la conception. Pour pré-sélectionner un certain nombre de PM en considérant la gamme de fréquence, on peut utiliser la Figure 3.7. Elle illustre le tracé de gain maximal G_{max} et le tracé de gain G_{3PM} pour chaque PM intermédiaire. Pour cela il faut tracer sur cette courbe la gamme de fréquence considérée et si possible, le gain minimal souhaité.

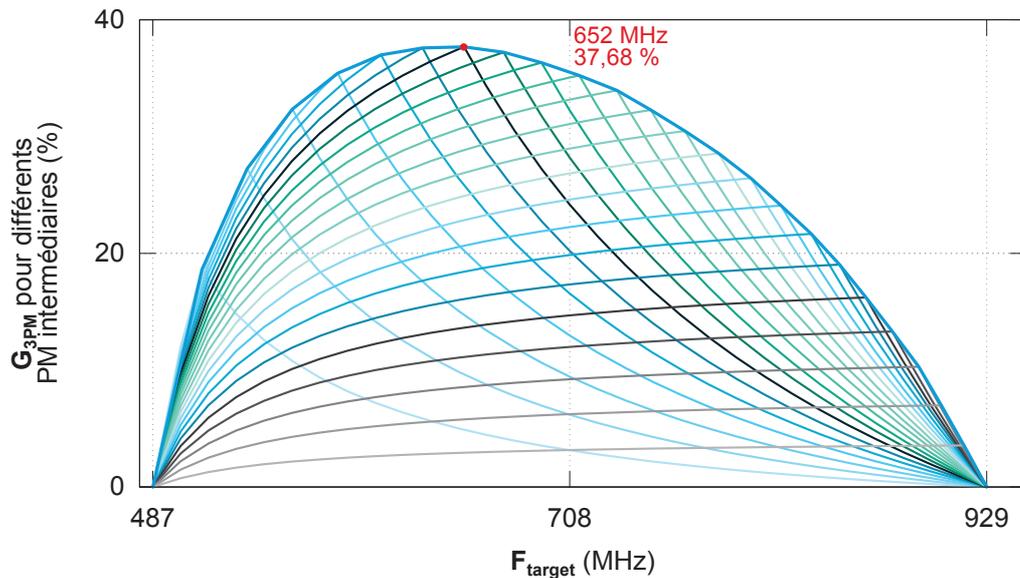


FIGURE 3.7: Contribution de l'ajout d'un PM intermédiaire sur le gain en consommation avec 3 PM par rapport à 2 PM disponibles, pour toutes les possibilités.

Par exemple, si la gamme de fréquence nécessaire pour exécuter les tâches est [580 ; 750] MHz, les PM intermédiaires à retenir sont les PM dont les fréquences sont comprises dans cette gamme, soit 8 PM pour cette gamme de fréquence (voir la Figure 3.8). Pour choisir le PM maximisant le gain, il faut calculer l'aire S_G

du gain en consommation dans la gamme de fréquence considérée. Pour cela, il faut calculer l'intégrale de la fonction dans la gamme considérée. Pour cela, nous avons besoin de calculer ou de modéliser la fonction. Même si cela ne pose pas de difficulté, nous proposons d'utiliser une méthode approchée afin de déterminer l'aire grâce à une méthode d'intégration approchée des intégrales définies.

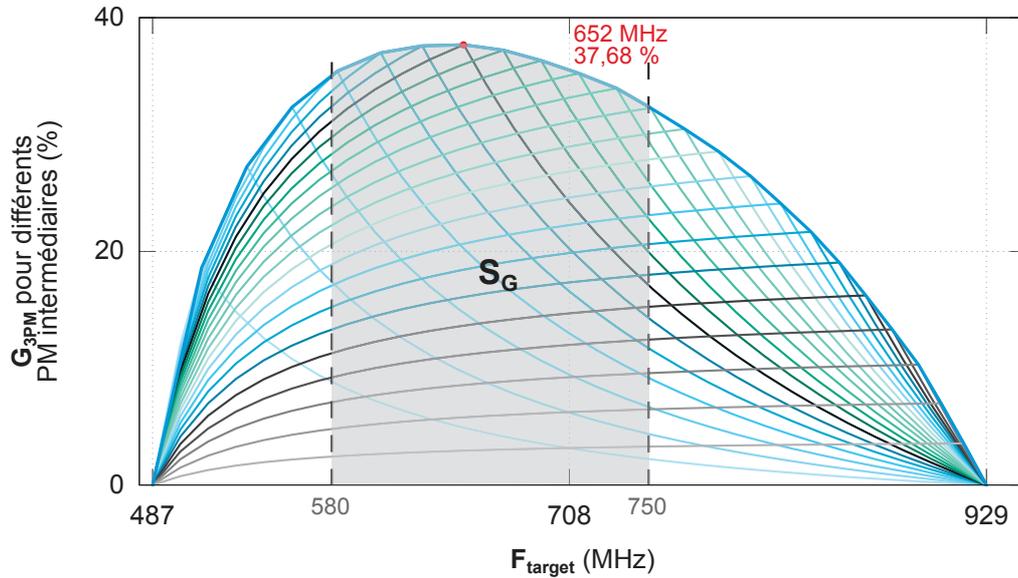


FIGURE 3.8: Contribution de l'ajout d'un PM intermédiaire sur le gain en consommation avec 3 PM par rapport à 2 PM disponibles, pour une gamme de fréquence.

Nous allons maintenant présenter comment nous pouvons estimer la moyenne du gain en consommation. Sachant que, la moyenne du gain en consommation $\overline{G_G}$ sur l'intervalle $[F_a ; F_b]$ est égale à :

$$\overline{G_G} = \frac{1}{\Delta F} \cdot S_G = \frac{1}{F_b - F_a} \cdot S_G \quad (3.6)$$

avec $\overline{G_G}$ le gain moyen en consommation sur la gamme de fréquence considérée, $\Delta F = F_b - F_a$ avec $[F_a ; F_b]$ la gamme de fréquence considérée, S_G la surface du gain en fonction de la fréquence sur la gamme de fréquence considérée, $S_G = \int_{F_a}^{F_b} G_{3PM}(F) dF$.

Afin d'obtenir le gain moyen il faut déterminer la surface S_G du gain en fonction de F_{target} sur la gamme de fréquence considérée. Comment peut-on le déterminer ? Il existe différentes méthodes de calcul approché permettant de déterminer l'aire d'une fonction. Nous évaluerons l'aire à l'aide d'une fonction permettant de calculer l'aire d'un polygone avec MATLAB. Notons que les méthodes permettant de calculer l'aire sont des outils utilisés pour obtenir le gain en consommation moyen pour le PM intermédiaire considéré.

Évaluation du gain en consommation moyen

Nous allons évaluer le gain moyen sur une gamme de fréquence considérée en utilisant la fonction permettant de calculer l'aire d'un polygone sous MATLAB. Cela revient à obtenir l'aire avec une méthode similaire à la méthode Newton-Cotes (ou méthode des trapèzes).

La gamme de fréquence considérée est [700 ; 900] MHz. Nous allons évaluer le gain moyen $\overline{G_G}$ à l'aide d'une fonction qui détermine la surface d'un polygone sous MATLAB, sachant que le gain a été déterminé avec un pas de $1 \cdot 10^{-4}$ MHz. On peut supposer que la surface est très proche de l'aire réelle. Le gain va être estimé pour différents PM intermédiaire appartenant à la gamme de fréquence considérée.

Dans le Tableau 3.3, le gain moyen avec la surface d'un polygone G_G sur l'intervalle [700 ; 900] MHz a été évalué pour différents PM intermédiaire. Nous pouvons observer que le maximum de gain moyen G_G sur l'intervalle [700 ; 900] MHz est de 28,10% et est atteint pour $F_{middle} = 787$ MHz.

TABLEAU 3.3: La moyenne du gain en consommation $\overline{G_G}$ obtenue par l'ajout du PM intermédiaire défini avec la consommation $P_{tot}(F_{middle})$ associée à la fréquence F_{middle} et à $V_{dd_{middle}}$ sur l'intervalle [700 ; 900] MHz.

PM _{middle}			$\overline{G_G}$ (%)
$P_{tot}(F_{middle})$ (mW)	F_{middle} (MHz)	$V_{dd_{middle}}$ (mV)	
71,36	713	850	23,0282
77,77	733	875	25,3074
84,36	751	900	26,6683
91,33	769	925	27,6085
98,72	787	950	28,0968
106,39	804	975	27,9148
114,37	820	1000	27,0437
122,77	836	1025	25,7145
131,48	851	1050	23,6971
140,48	865	1075	21,0140
149,97	879	1100	17,8129
159,95	893	1125	14,0745

Nous allons maintenant présenter une méthode plus formelle pour trouver quel PM permet de maximiser le gain moyen sur un intervalle.

3.3.4 Méthode de placement du PM intermédiaire

Afin de s'affranchir d'un calcul approché du gain en consommation, nous proposons d'utiliser le profil de la consommation en fonction de la fréquence du circuit et d'y tracer le triangle rejoignant les 3 PM, voir la Figure 3.9. Dans la suite du manuscrit, cette étude sera nommée la méthode du triangle. La réduction en consommation \overline{PR}_{Tri} apportée par l'ajout de ce PM intermédiaire sur toute la gamme de fréquence disponibles $[F_{low} ; F_{high}]$ est proportionnelle à l'aire de ce triangle S_{Tri} . En effet, le profil $P_{tot}(F)$ pour l'ajout d'un PM intermédiaire doit être le plus proche possible du profil $P_{tot}(F)$ continu. Après quelques manipulations, on obtient :

$$S_{Tri} = \frac{1}{2} \cdot \{(F_{middle} - F_{low}) \cdot (P_{tot}(F_{high}) - P_{tot}(F_{low})) - (F_{high} - F_{low}) \cdot (P_{tot}(F_{middle}) - P_{tot}(F_{low}))\} \quad (3.7)$$

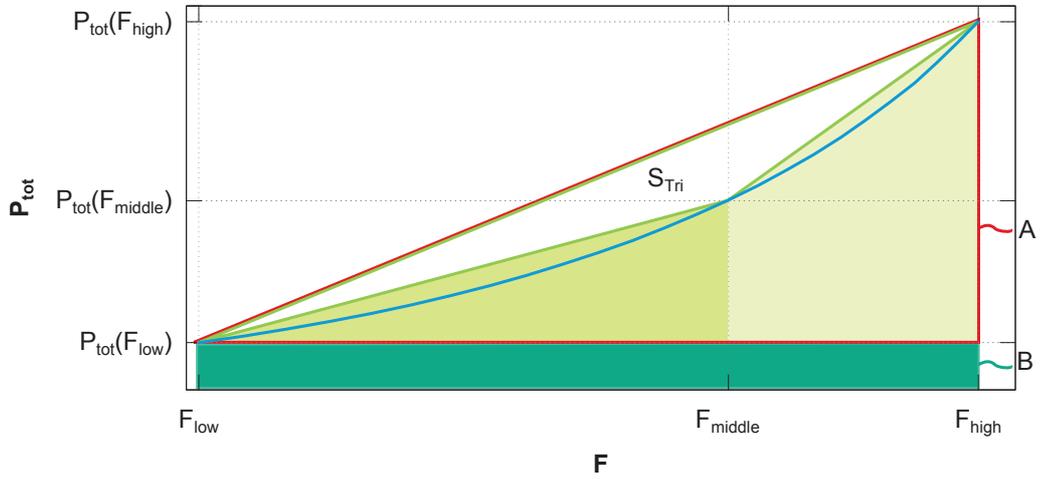


FIGURE 3.9: Consommation en fonction de la fréquence, avec l'apport d'un PM intermédiaire $(F_{middle}, P_{tot}(F_{middle}))$ sur toute la gamme de fréquence.

La moyenne de la réduction en consommation sur la gamme $[F_{low} ; F_{high}]$ est donnée par :

$$\begin{aligned} \overline{PR}_{Tri} &= 100 \cdot \frac{P_{tot_{2PM}} - P_{tot_{3PM}}}{P_{tot_{2PM}}} \\ &= 100 \cdot \frac{A + B - (A + B - S_{Tri})}{A + B} \\ &= 100 \cdot \frac{S_{Tri}}{\frac{(P_{tot}(F_{high}) - P_{tot}(F_{low})) \cdot (F_{high} - F_{low})}{2} + P_{tot}(F_{low}) \cdot (F_{high} - F_{low})} \end{aligned} \quad (3.8)$$

avec \overline{PR}_{Tri} la réduction en consommation en pourcentage, $P_{tot_{2PM}}$ la surface associée à la consommation avec 2 PM sur toute la gamme de fréquence disponible, $P_{tot_{3PM}}$ la surface associée à la consommation avec 3 PM sur toute la gamme de fréquence disponible, S_{Tri} la surface du triangle formée par les PM.

Lorsque la gamme de fréquence désirée correspond à toute la plage de fréquence, il faut maximiser l'aire du triangle S_{Tri} pour maximiser la réduction en consommation.

Cependant, si la gamme de fréquence d'exécution du circuit est l'intervalle $[F_a ; F_b]$ et que l'on souhaite garder les PM extrêmes pour un fonctionnement *Low Power* ou *High Performance*, avec $F_a < F_{middle} < F_b$, la réduction en consommation est proportionnelle à l'aire du pentagone S_{G1} colorée sur la Figure 3.10. S_{G1} est obtenue grâce à :

$$S_{G1} = S_{Tri} - S_1 - S_2 \quad (3.9)$$

avec l'aire du triangle S_1 :

$$\begin{aligned} S_1 = & \frac{1}{2} \cdot \left\{ (F_a - F_{low}) \cdot \left(\frac{P_{tot}(F_{high}) - P_{tot}(F_{low})}{F_{high} - F_{low}} \cdot (F_a - F_{low}) \right. \right. \\ & + P_{tot}(F_{low}) - P_{tot}(F_{low})) \\ & - (F_a - F_{low}) \cdot \left(\frac{P_{tot}(F_{middle}) - P_{tot}(F_{low})}{F_{middle} - F_{low}} \cdot (F_a - F_{low}) \right. \\ & \left. \left. + P_{tot}(F_{low}) - P_{tot}(F_{low})) \right\} \end{aligned} \quad (3.10)$$

et l'aire du triangle S_2 :

$$\begin{aligned} S_2 = & \frac{1}{2} \cdot \left\{ (F_b - F_{high}) \cdot \left(\frac{P_{tot}(F_{high}) - P_{tot}(F_{middle})}{F_{high} - F_{middle}} \cdot (F_b - F_{middle}) \right. \right. \\ & + P_{tot}(F_{middle}) - P_{tot}(F_{high})) \\ & - (F_b - F_{high}) \cdot \left(\frac{P_{tot}(F_{high}) - P_{tot}(F_{low})}{F_{high} - F_{low}} \cdot (F_b - F_{low}) \right. \\ & \left. \left. + P_{tot}(F_{low}) - P_{tot}(F_{high})) \right\} \end{aligned} \quad (3.11)$$

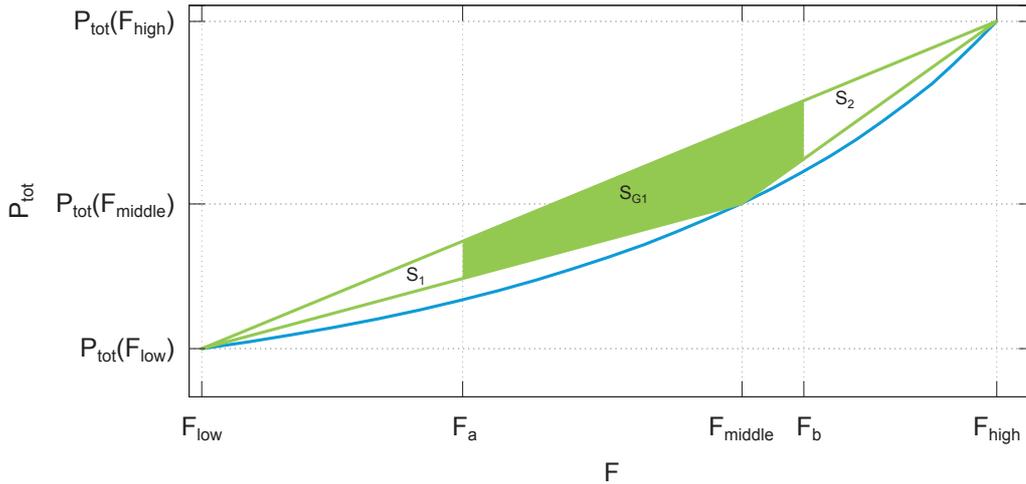


FIGURE 3.10: Consommation en fonction de la fréquence, avec l'apport d'un PM intermédiaire $(F_{middle}, P_{tot}(F_{middle}))$ sur la gamme de fréquence $[F_a ; F_b]$.

Pour maximiser la réduction en consommation \overline{PR} dans la gamme de fréquence considérée $[F_a ; F_b]$, il faut maximiser l'aire du pentagone S_{G1} qui est proportionnelle à \overline{PR} .

La moyenne de la réduction en consommation sur la gamme $[F_a ; F_b]$ est alors donnée par :

$$\overline{PR}_{G1} = \frac{S_{G1}}{\frac{1}{2} \cdot (F_b - F_a) \cdot \left(\frac{P_{tot}(F_{high}) - P_{tot}(F_{low})}{F_{high} - F_{low}} \cdot (F_a + F_b - 2 \cdot F_{low}) + 2 \cdot P_{tot}(F_{low}) \right)} \quad (3.12)$$

Si le circuit exécute des tâches dans la gamme de fréquence typique $[F_c ; F_d]$, avec $F_c < F_d < F_{middle}$, la réduction en consommation dans la gamme de fréquence considérée $[F_c ; F_d]$, est proportionnelle à l'aire du trapèze S_{G2} colorée sur la Figure 3.11. S_{G2} est égale à :

$$\begin{aligned} S_{G2} = & \frac{1}{2} \cdot \{ [F_c \cdot \left(\frac{P_{tot}(F_{middle}) - P_{tot}(F_{low})}{F_{middle} - F_{low}} \cdot (F_d - F_{low}) + P_{tot}(F_{low}) \right) \\ & - F_d \cdot \left(\frac{P_{tot}(F_{middle}) - P_{tot}(F_{low})}{F_{middle} - F_{low}} \cdot (F_c - F_{low}) + P_{tot}(F_{low}) \right)] \\ & + [F_d \cdot \left(\frac{P_{tot}(F_{high}) - P_{tot}(F_{low})}{F_{high} - F_{low}} \cdot (F_d - F_{low}) + P_{tot}(F_{low}) \right) \\ & - F_d \cdot \left(\frac{P_{tot}(F_{middle}) - P_{tot}(F_{low})}{F_{middle} - F_{low}} \cdot (F_d - F_{low}) + P_{tot}(F_{low}) \right)] \\ & + [F_d \cdot \left(\frac{P_{tot}(F_{high}) - P_{tot}(F_{low})}{F_{high} - F_{low}} \cdot (F_c - F_{low}) + P_{tot}(F_{low}) \right) \\ & - F_c \cdot \left(\frac{P_{tot}(F_{high}) - P_{tot}(F_{low})}{F_{high} - F_{low}} \cdot (F_d - F_{low}) + P_{tot}(F_{low}) \right)] \\ & + [F_c \cdot \left(\frac{P_{tot}(F_{middle}) - P_{tot}(F_{low})}{F_{middle} - F_{low}} \cdot (F_c - F_{low}) + P_{tot}(F_{low}) \right) \\ & - F_c \cdot \left(\frac{P_{tot}(F_{high}) - P_{tot}(F_{low})}{F_{high} - F_{low}} \cdot (F_c - F_{low}) + P_{tot}(F_{low}) \right)] \} \end{aligned} \quad (3.13)$$

Pour maximiser la réduction en consommation dans la gamme de fréquence considérée $[F_c ; F_d]$, il faut maximiser l'aire du trapèze S_{G2} .

La moyenne de la réduction en consommation sur la gamme $[F_c ; F_d]$ est égale à :

$$\overline{PR}_{G2} = \frac{S_{G2}}{\frac{1}{2} \cdot (F_d - F_c) \cdot \left(\frac{P_{tot}(F_{high}) - P_{tot}(F_{low})}{F_{high} - F_{low}} \cdot (F_c + F_d - 2 \cdot F_{low}) + 2 \cdot P_{tot}(F_{low}) \right)} \quad (3.14)$$

Si la gamme de fréquence désirée correspond à une seule fréquence, dans ce cas là, il faut déterminer le PM ayant la fréquence la plus proche de cette fréquence fixe.

Nous allons comparer la moyenne de réduction en consommation obtenue avec MATLAB *via* le profil de gain en fonction de la fréquence cible à la moyenne de

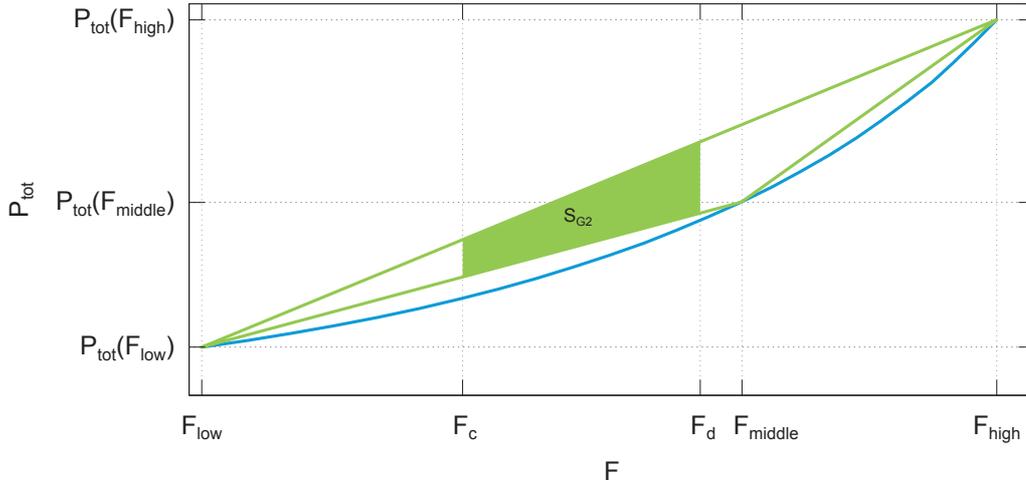


FIGURE 3.11: Consommation en fonction de la fréquence, avec l'apport d'un PM intermédiaire $(F_{middle}, P_{tot}(F_{middle}))$ sur la gamme de fréquence $[F_c ; F_d]$.

réduction en consommation obtenue *via* la surface sur profil de consommation en fonction de la fréquence.

Comparaison de la méthode de calcul approché à la méthode du triangle

Nous souhaitons comparer le gain en consommation moyen $\overline{G_G}$ obtenu avec MATLAB à la réduction en consommation obtenue avec la méthode du triangle. Cependant, les deux métriques sont différentes. Travailler sur le profil gain en consommation en fonction de la fréquence cible F_{target} est plus intuitif car c'est assez visuel mais le vrai apport du PM intermédiaire est obtenu sur le profil de consommation en fonction de la fréquence. Nous allons donc évaluer la réduction en consommation moyenne avec MATLAB $\overline{PR_M}$ afin de pouvoir comparer les deux méthodes. La moyenne de réduction en consommation $\overline{PR_M}$ sur la gamme $[F_{low} ; F_{high}]$ est obtenue à l'aide de l'équation ci-dessous :

$$\overline{PR_M} = 100 \cdot \frac{S_G}{P_{tot2PM}} = 100 \cdot \frac{S_G}{A + B} \quad (3.15)$$

avec $\overline{PR_M}$ la réduction en consommation en pourcentage obtenue avec MATLAB, S_G l'aire du gain en consommation dans la gamme de fréquence considérée obtenue (ici $[F_{low} ; F_{high}]$) avec MATLAB, P_{tot2PM} la surface associée à la consommation avec 2 PM sur toute la gamme de fréquence disponible $[F_{low} ; F_{high}]$.

$\overline{PR_M}$ correspond donc à la réduction en consommation normalisée. $\overline{PR_M}$ et $\overline{PR_{Tri}}$ sont proportionnelles à un gain en consommation.

Nous allons évaluer la réduction en consommation moyenne $\overline{PR_M}$ obtenue avec MATLAB pour toutes les possibilités de PM intermédiaire sur toute la plage de fréquence disponible. Nous allons aussi évaluer la moyenne de la réduction en

consommation $\overline{PR_{Tri}}$ obtenue grâce au profil de consommation en fonction de la fréquence pour tous les PM intermédiaires possibles sur toute la plage de fréquence disponible. Le Tableau 3.4 présente les valeurs de la consommation, de la fréquence, de la tension d'alimentation et des moyennes de réduction en consommation pour toutes les possibilités de PM intermédiaire, voir également la Figure 3.12. Nous remarquons que la moyenne de réduction en consommation est maximale pour le PM intermédiaire dont la fréquence est de 751 MHz. Le gain en consommation maximal avait été obtenu pour $F_{middle} = 652$ MHz, voir la Figure 3.6. Les deux métriques utilisées sont différentes, le choix du PM intermédiaire est donc différent.

La Figure 3.12 illustre la moyenne de réduction en consommation $\overline{PR_M}$ et la moyenne de réduction en consommation $\overline{PR_{Tri}}$ en fonction de la fréquence du PM intermédiaire F_{middle} . Nous pouvons observer sur la Figure 3.12 que le maximum pour le $\overline{PR_M}(F_{middle})$ et $\overline{PR_{Tri}}(F_{middle})$ est identique. Nous observons que les deux courbes se superposent. La précision obtenue avec les deux méthodes est identique mais la méthode du triangle est plus rapide en termes de calculs. En effet, $\overline{PR_M}$ est obtenue en calculant l'aire d'un polygone. Cela revient à calculer l'aire avec la méthode des trapèzes. La réduction en consommation $\overline{PR_{Tri}}$ obtenue avec la méthode du triangle correspond à la surface d'un triangle quelconque normalisée par la somme de la surface d'un triangle rectangle et la surface d'un rectangle.

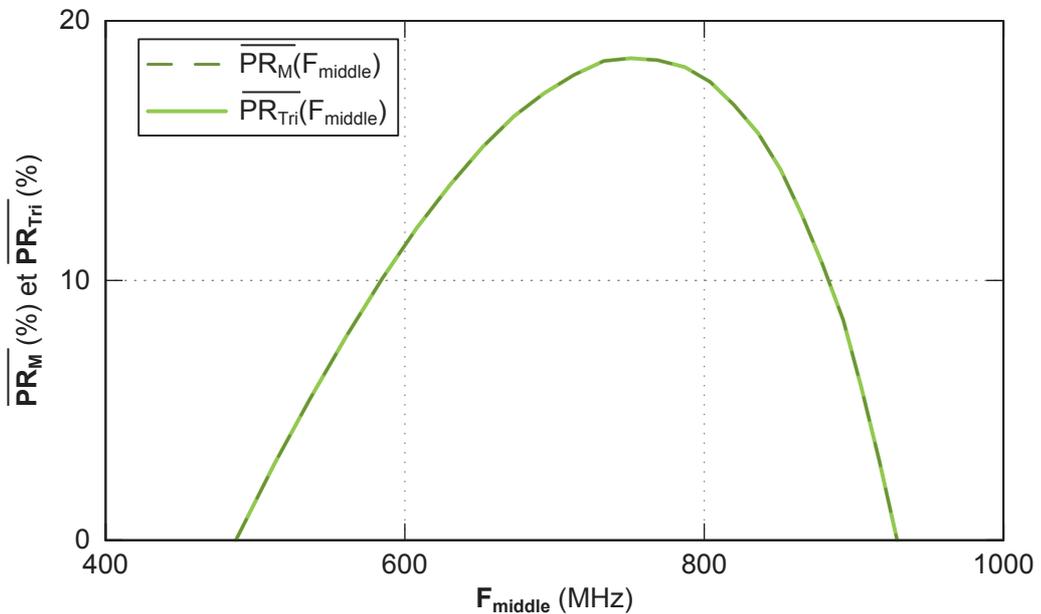


FIGURE 3.12: Réductions en consommation moyenne $\overline{PR_M}$ et $\overline{PR_{Tri}}$ obtenues par l'ajout d'un PM intermédiaire en fonction de la fréquence du PM intermédiaire F_{middle} , pour toutes les possibilités.

Une fois que le PM intermédiaire a été choisi et que les actionneurs ont été implémentés, dans le cas où les actionneurs sont embarqués, les PM ne pourront plus être modifiés. Dans ce cas, une solution sera de modifier la partie logicielle et donc la fréquence cible des applications de manière à s'approcher du maximum du gain en consommation.

TABLEAU 3.4: La réduction en consommation moyenne \overline{PR}_M avec MATLAB et la réduction en consommation moyenne \overline{PR}_{Tri} avec la méthode du triangle obtenues par l'ajout du PM intermédiaire défini avec la consommation $P_{tot}(F_{middle})$ associée à la fréquence F_{middle} et à $V_{dd_{middle}}$ sur toute la gamme de fréquence.

PM _{middle}			\overline{PR}_M (%)	\overline{PR}_{Tri} (%)
$P_{tot}(F_{middle})$ (mW)	F_{middle} (MHz)	$V_{dd_{middle}}$ (mV)		
28,06	513	625	2,9603	2,9603
31,68	537	650	5,4834	5,4834
35,61	561	675	7,8632	7,8632
39,86	585	700	10,0953	10,0953
44,38	608	725	12,0280	12,0280
49,15	630	750	13,6704	13,6704
54,27	652	775	15,1512	15,1512
59,67	673	800	16,3279	16,3279
65,33	693	825	17,2098	17,2098
71,36	713	850	17,9208	17,9208
77,77	733	875	18,4563	18,4563
84,36	751	900	18,5591	18,5591
91,33	769	925	18,4865	18,4865
98,72	787	950	18,2199	18,2199
106,39	804	975	17,6493	17,6493
114,37	820	1000	16,7607	16,7607
122,77	836	1025	15,6781	15,6781
131,48	851	1050	14,2776	14,2776
140,48	865	1075	12,5685	12,5685
149,97	879	1100	10,6330	10,6330
159,95	893	1125	8,4713	8,4713
170,07	905	1150	5,8953	5,8953
180,71	917	1175	3,0793	3,0793

Solution analytique

Le PM intermédiaire permettant de maximiser le gain en consommation moyen peut être déterminé par une recherche exhaustive. Cependant, il est plus intéres-

sant d'obtenir une solution analytique. La surface du triangle est donnée par l'équation suivante :

$$S_{Tri} = \frac{1}{2} \cdot \{(F_{middle} - F_{low}) \cdot (P_{tot}(F_{high}) - P_{tot}(F_{low})) - (F_{high} - F_{low}) \cdot (P_{tot}(F_{middle}) - P_{tot}(F_{low}))\} \quad (3.16)$$

L'objectif est de placer le PM intermédiaire de manière à maximiser le gain en consommation. Cela revient à maximiser l'aire du triangle S_{Tri} . Notre variable est donc le PM intermédiaire qui est défini par le couple $(F_{middle}, P_{tot}(F_{middle}))$, que l'on va nommer F . On suppose que la consommation $P_{tot}(F)$ peut être modélisée par une fonction cubique. On considère donc que $P_{tot}(F) = a \cdot F^3 + b \cdot F^2 + c \cdot F + e$, avec a, b, c, e les paramètres à déterminer. Cette équation est substituée dans l'équation (3.16).

$$S_{Tri}(F) = \frac{1}{2} \cdot \{(F - F_{low}) \cdot (P_{tot}(F_{high}) - P_{tot}(F_{low})) - (F_{high} - F_{low}) \cdot (a \cdot F^3 + b \cdot F^2 + c \cdot F + e - P_{tot}(F_{low}))\} \quad (3.17)$$

Afin de trouver le maximum local de la fonction, il faut que la dérivée s'annule en F^* , *i.e.* $\frac{dS_{Tri}(F^*)}{dF} = 0$ et que la dérivée seconde soit négative en F^* , *i.e.* $\frac{d^2S_{Tri}(F^*)}{dF^2} < 0$. Cela revient à dire que la fonction est concave en F^* .

La dérivée de $S_{Tri}(F)$ est donnée par :

$$\frac{dS_{Tri}(F)}{dF} = \frac{1}{2} \cdot \{(P_{tot}(F_{high}) - P_{tot}(F_{low})) - (F_{high} - F_{low}) \cdot (3 \cdot a \cdot F^2 + 2 \cdot b \cdot F + c)\} \quad (3.18)$$

Le discriminant de l'équation (3.18) est :

$$\Delta_S = 4 \cdot (F_{high} - F_{low})^2 \cdot (b^2 - 3 \cdot a \cdot c) + 12 \cdot a \cdot (F_{high} - F_{low}) \cdot (P_{tot}(F_{high}) - P_{tot}(F_{low})) \quad (3.19)$$

Si Δ_S est positive, il existe au moins une solution $F_S \in \mathbb{R}$. Les solutions de l'équation (3.18) sont :

$$F_{S(1,2)} = \frac{2 \cdot b \cdot (F_{high} - F_{low}) \pm \sqrt{\Delta_S}}{-6 \cdot a \cdot (F_{high} - F_{low})} \quad (3.20)$$

La dérivée seconde de $S_{Tri}(F)$ est donnée par :

$$\frac{d^2S_{Tri}(F)}{dF^2} = \frac{-(F_{high} - F_{low}) \cdot (6 \cdot a \cdot F + 2 \cdot b)}{2} \quad (3.21)$$

La dérivée seconde s'annule en :

$$F_{S_0} = \frac{-b}{3 \cdot a}$$

La dérivée seconde est négative pour :

- $F < F_{S_0}$, si $a < 0$;
- $F > F_{S_0}$, si $a > 0$.

On a :

$$F_{S_{(1,2)}} = F_{S_0} \pm \frac{\sqrt{\Delta_S}}{6 \cdot a \cdot (F_{high} - F_{low})}$$

Si a est différent de 0, la solution permettant de maximiser le gain est donc :

$$F_{middle}^{opt} = F_{S_2} = \frac{2 \cdot b \cdot (F_{high} - F_{low}) - \sqrt{\Delta_S}}{-6 \cdot a \cdot (F_{high} - F_{low})}$$

Le Tableau 3.5 donne le signe des dérivées en fonction des paramètres à trouver lors de la détermination du modèle cubique de $P_{tot}(F)$. Si le paramètre identifié a est strictement positif ou strictement négatif, alors la fréquence F_{middle}^{opt} permet de maximiser le gain en consommation.

TABLEAU 3.5: Tableau de signes de l'évolution de S_{Tri} .

F	$-\infty$	$\min(F_{S_{(1,2)}})$	F_{S_0}	$\max(F_{S_{(1,2)}})$	$+\infty$
$\frac{dS_{Tri}}{dF}$	signe(-a)		signe(a)		signe(-a)
$\frac{d^2S_{Tri}}{dF^2}$		signe(a)		signe(-a)	

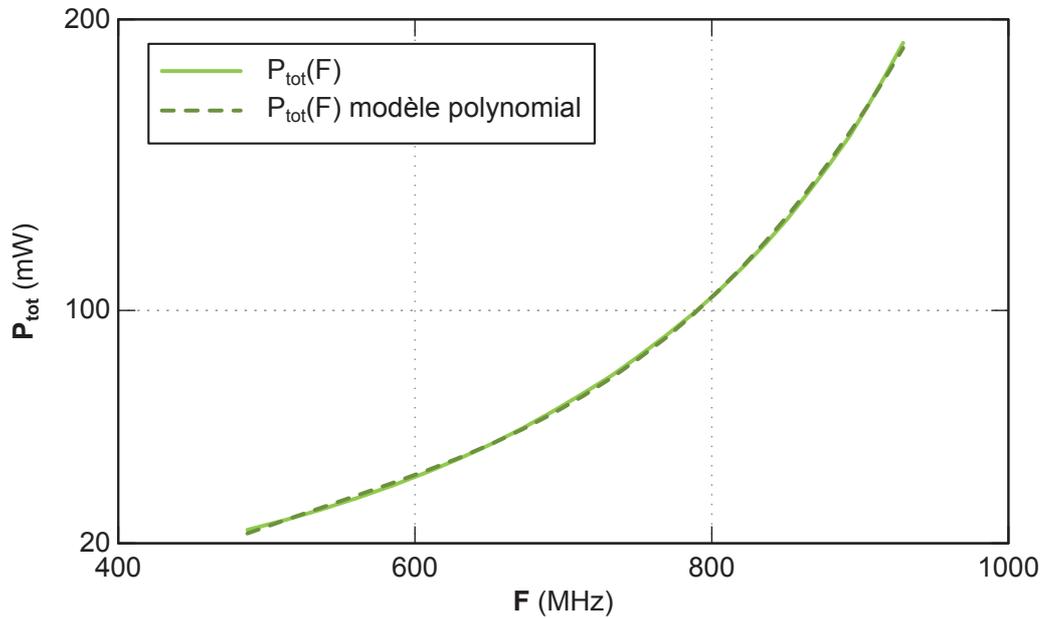


FIGURE 3.13: Profil de consommation en fonction de la fréquence $P_{tot}(F)$ et modèle cubique sur LoCoMoTIV.

La Figure 3.13 illustre le profil de consommation en fonction de la fréquence et le modèle cubique obtenu.

Le modèle de $P_{tot}(F)$ est donnée par :

$$P_{tot}(F) = 1,55249933426394 \cdot 10^{-6} \cdot F^3 - 0,00252396189728313 \cdot F^2 \\ + 1,54075527550773 \cdot F - 307,677127131594$$

a est positif, la solution est donc $F_{middle}^{opt} = 751,9662$ MHz.

La fréquence F_{middle} maximisant la surface du triangle, obtenue avec une recherche exhaustive était de 751 MHz (voir le Tableau 3.4). La solution analytique donne une fréquence F_{middle} égale à 751,9662 MHz. Les résultats sont donc cohérents et nous pouvons affirmer que la solution analytique proposée est relativement précise. La fréquence F_{middle} obtenue avec la solution analytique est légèrement différente de la solution obtenue avec la recherche exhaustive car nous n'avons que 23 valeurs de PM intermédiaire disponibles dans la gamme $[F_{low} ; F_{high}]$.

Résultats pour placement de PM intermédiaire sur un circuit en technologie FD-SOI

Notre second cas d'étude est un processeur de traitement du signal ou DSP (Digital Signal Processing) appelé Frisbee et implémenté en technologie STMicroelectronics FD-SOI 28nm [74]. Il s'agit d'un prototype permettant d'agir sur la tension d'alimentation V_{dd} , la tension de body bias V_{bb} et la fréquence F . La consommation du circuit a été mesurée sur le circuit pour différentes tensions d'alimentation V_{dd} , tensions de body bias V_{bb} et fréquences F appliquées. La fréquence appliquée pour un couple (V_{dd}, V_{bb}) correspond à la fréquence maximale F_{max} atteignable sans produire d'erreur temporelle. Il est à noter que même si la consommation totale du circuit a été mesurée, les tâches exécutées pour ces mesures ont peu d'accès mémoire afin de pouvoir considérer les mesures comme représentatives de la consommation du cœur de calcul.

Pour le cas d'étude considéré, la tension d'alimentation minimale est $V_{dd_{min}} = 0,7$ V, la tension d'alimentation maximale est $V_{dd_{max}} = 1,3$ V. La tension de body bias est comprise entre $V_{bb_{min}} = 0$ mV et $V_{bb_{max}} = 1500$ mV.

La Figure 3.14 illustre la réduction en consommation moyenne $\overline{PR_{Tri}}$ évaluée avec un outil d'analyse numérique, MATLAB. Ces réductions ont été déterminées à partir des mesures issues du circuit pour quatre valeurs de tension de body bias V_{bb} différentes. Le Tableau 3.6 contient la réduction en consommation moyenne $\overline{PR_{Tri}}$ maximale pour la fréquence intermédiaire optimale F_{middle}^{opt} . Ce tableau contient aussi $\overline{PR_{Tri}}$ pour deux fréquences autour de F_{middle} . Pour une tension $V_{bb} = 0$ mV, en plaçant correctement le PM intermédiaire, la consommation peut être réduite de 21,19% en moyenne. Nous pouvons observer qu'une variation de 100 MHz de F_{middle} n'influe au maximum que de 1% sur la réduction en consommation moyenne $\overline{PR_{Tri}}$.

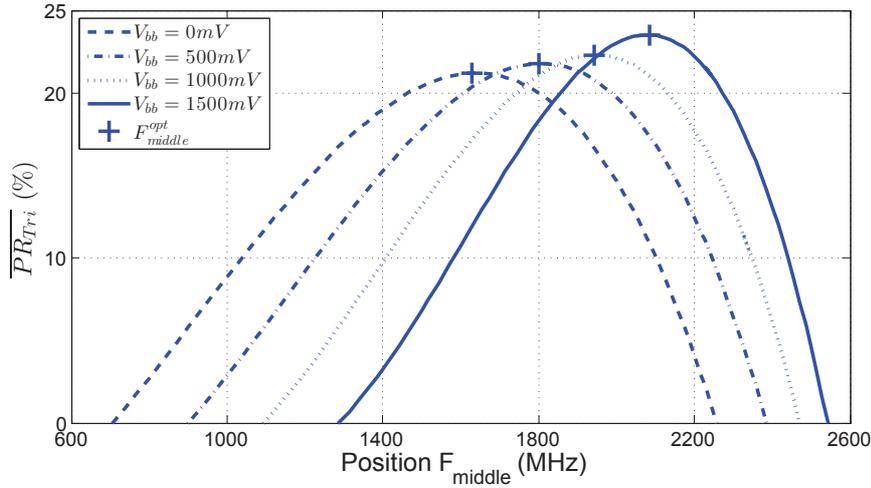


FIGURE 3.14: Réduction en consommation moyenne \overline{PR}_{Tri} obtenue par l'ajout d'un PM intermédiaire en fonction de la fréquence du PM intermédiaire F_{middle} , pour quatre V_{bb} .

TABLEAU 3.6: Variation de la réduction en consommation moyenne \overline{PR}_{Tri} lorsque F_{middle} varie.

V_{bb}	0 mV	500 mV	1000 mV	1500 mV
$F_{middle}^{opt} - 100MHz$	20,92%	21,38%	21,65%	22,83%
F_{middle}^{opt}	21,19%	21,79%	22,28%	23,52%
$F_{middle}^{opt} + 100MHz$	20,57%	21,18%	21,69%	22,53%

La Figure 3.15 illustre la réduction en consommation moyenne \overline{PR}_{Tri} lorsque les 3 actionneurs sont discrets. La réduction en consommation moyenne maximale est de 35% pour $F_{middle}^{opt} = 2019$ MHz. La Figure 3.16 illustre le profil de consommation en fonction de la fréquence, le polynôme approximé et le triangle reliant les 3 PM pour $F_{middle}^{opt} = 2019$ MHz correspondant au cas optimal. Le gain en consommation maximal est de 50,3% lorsque $F_{target} = 2019$ MHz. Le Tableau 3.7 présente les valeurs de F , V_{dd} et V_{bb} pour les 3 PM permettant de maximiser \overline{PR}_{Tri} .

TABLEAU 3.7: Meilleurs PM avec 3 actionneurs discrets.

PM	PM_{low}	PM_{middle}	PM_{high}
V_{dd}	0,7 V	1,1 V	1,3 V
V_{bb}	0 mV	1008 mV	1500 mV
F	704 MHz	2019 MHz	2542 MHz

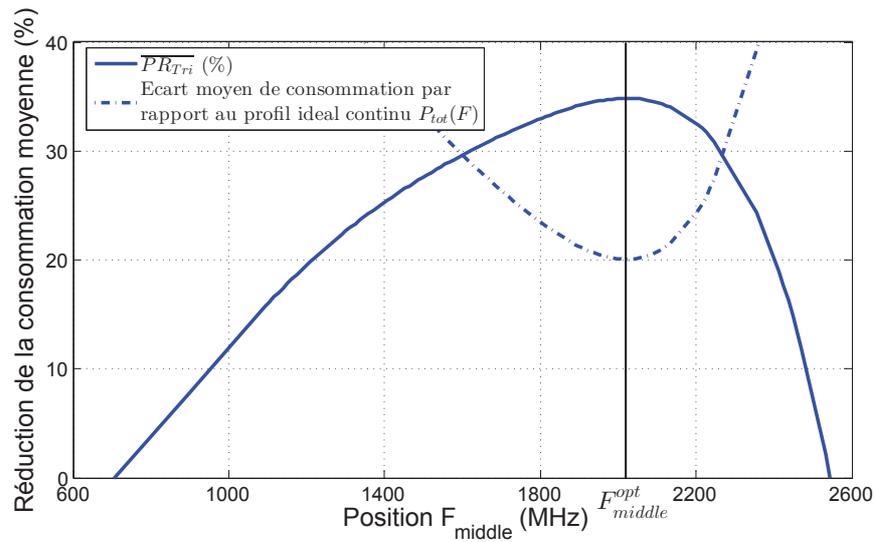


FIGURE 3.15: Réduction en consommation moyenne $\overline{PR_{Tri}}$ obtenue par l'ajout d'un PM intermédiaire en fonction de la fréquence du PM intermédiaire F_{middle} en modifiant (F, V_{dd}, V_{bb}) .

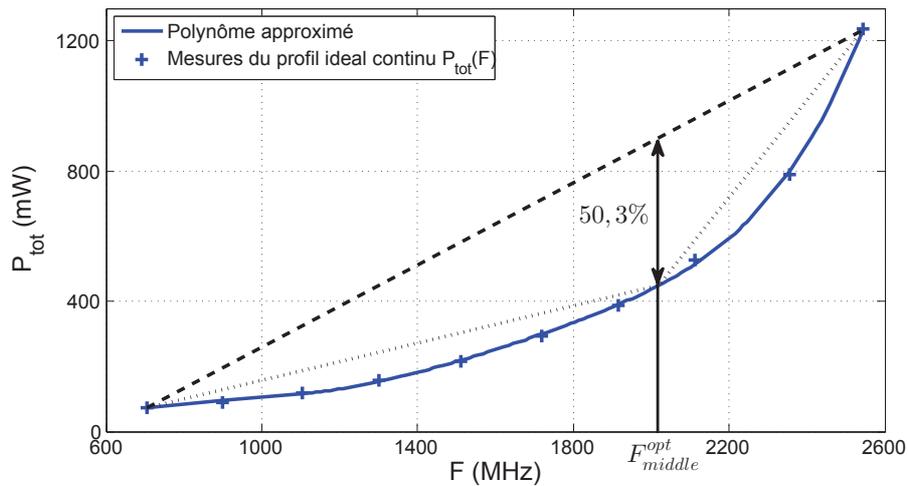


FIGURE 3.16: Profil de consommation en fonction de la fréquence pour la position optimale F_{middle}^{opt} en modifiant (F, V_{dd}, V_{bb}) .

Probabilité de gain

Si l'on souhaite estimer statistiquement la qualité d'approche, dans notre cas l'espérance de réduction de consommation en fonction de la fréquence $PR(F)$. On peut considérer que la variable aléatoire PR prend les valeurs pr_1, pr_2, \dots, pr_n . $PR_i = PR(F = F_i)$ correspond à l'espérance de réduction en consommation pour

la fréquence F_i . Si, la densité de probabilité (ou PDF) de la variable aléatoire F ($PDF(F)$) est connue, nous pouvons améliorer l'estimation de l'espérance de réduction en consommation en fonction de la fréquence et donc le choix du PM intermédiaire. Cela revient à connaître la probabilité p_i que la fréquence F_i soit demandée. L'idée sous-jacente est que la fréquence d'apparition f_i des valeurs de fréquence F_i est proche de la probabilité p_i . Sachant que la réduction en consommation pr_i dépend de la fréquence demandée F_i , la probabilité de réduction en consommation pr_i est donc égale à la probabilité p_i de la fréquence F_i . Pour obtenir l'espérance de réduction en consommation pour la fréquence F_i , on peut multiplier point par point les valeurs de gain pr_i par la probabilité que la fréquence F_i soit demandée p_i . On suppose donc que l'espérance de réduction en consommation PR_i pour la fréquence F_i est donnée par $\overline{PR_i} = pr_i \cdot p_i$. Ainsi, l'espérance de réduction en consommation en fonction de la fréquence $PR(F)$ est donnée par le vecteur suivant :

$$PR(F) = \begin{pmatrix} PR_1(F_1) \\ \cdot \\ \cdot \\ PR_i(F_i) \\ \cdot \\ \cdot \\ PR_n(F_n) \end{pmatrix} \quad (3.22)$$

Cela revient donc à multiplier point par point $\overline{PR_{Tri}(F_{middle})}$ (sur le Figure 3.12) par la densité de probabilité de la variable aléatoire F . Le PM intermédiaire permettant de maximiser le gain sachant la PDF correspond au maximum de $PR(F)$.

3.3.5 Synthèse

Nous avons vu que l'ajout d'un PM intermédiaire dans le cas où 2 PM sont déjà disponibles apporte un gain en consommation ou une réduction en consommation notable. La détermination de sa position est dictée par les contraintes applicatives.

Pour maximiser le gain en consommation apporté par ce PM intermédiaire, le choix du PM à implémenter est important. Pour cela, nous avons proposé différentes manières de calculer la moyenne du gain en consommation obtenu. La première approche consiste à évaluer, dans l'intervalle de fréquence considéré, le gain en consommation moyen *via* le calcul de l'aire du polygone sur le profil du gain en fonction de la fréquence F_{target} . La seconde approche consiste à évaluer la moyenne de la réduction en consommation en calculant la surface sur le profil de consommation en fonction de la fréquence (méthode du triangle). Cette seconde approche semble être plus rapide et élégante que la première approche propo-

sée. De plus, une solution analytique a été proposée pour la méthode du triangle. Cette solution permet de déterminer la fréquence du PM intermédiaire à implémenter afin de maximiser le gain en consommation. Par ailleurs, une solution a été proposée afin de maximiser le gain sachant la PDF.

Nous avons étudié le cas de 3 PM disponibles sur un dispositif. Nous allons maintenant étendre l'approche au cas de N PM disponibles sur le dispositif.

3.4 Gestion d'un système à N PM

Nous allons maintenant étendre l'étude précédente. Comment maximiser la réduction en consommation dans le cas où N PM sont disponibles sur le dispositif.

Pour minimiser la consommation du circuit, nous allons considérer qu'une tâche ayant une contrainte en performance F_{target} va être exécutée avec les 2 PM ayant les fréquences les plus proches de cette contrainte. De plus, nous supposons que la tâche est exécutée en effectuant une seule commutation entre les 2 PM. Lorsque la performance requise correspond à la fréquence associée à un PM, cette fréquence est appliquée pour exécuter la tâche.

Le polygone a pour sommets les points $PM_1, PM_2, \dots, PM_N, PM_{N+1} = PM_1$. Dans ce cas, l'aire du polygone est égale à :

$$S_{TP} = \frac{1}{2} \left| \sum_{i=1}^N (F_i \cdot P_{tot}(F_{i+1}) - F_{i+1} \cdot P_{tot}(F_i)) \right| \quad (3.23)$$

Si la gamme de fréquence requise correspond à toute la gamme de fréquence, pour maximiser la réduction en consommation, il faut maximiser l'aire du polygone S_{TP} (voir la Figure 3.17).

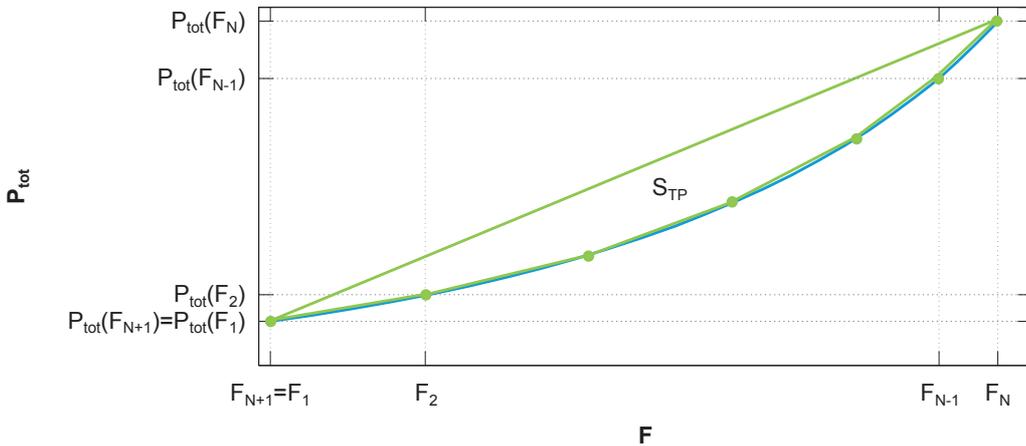


FIGURE 3.17: Consommation en fonction de la fréquence, avec l'apport de N PM sur toute la gamme de fréquence.

Lorsque la gamme de fréquence requise $[F_g, F_h]$ correspond à une partie de la gamme de fréquence disponible, il faut maximiser l'aire du polygone S_{GPP2}

délimité par les fréquences F_g et F_h , voir la Figure 3.18.

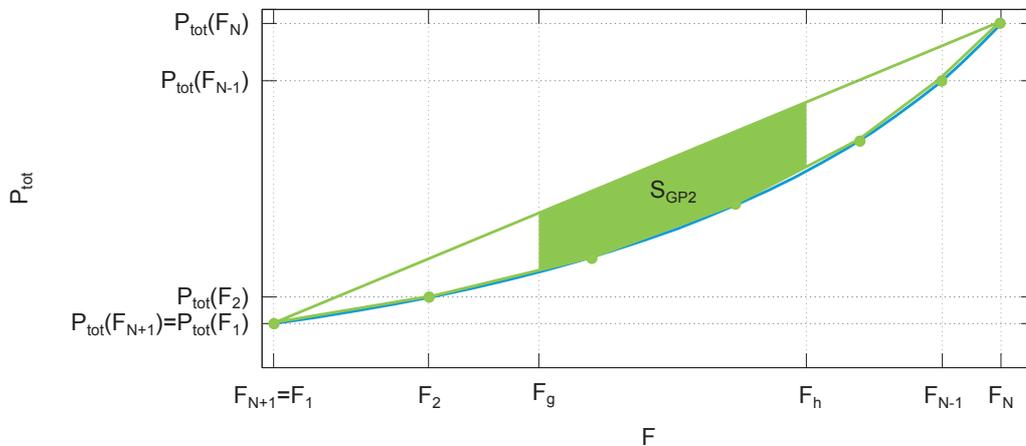


FIGURE 3.18: Consommation en fonction de la fréquence, avec l'apport de N PM sur la gamme de fréquence $[F_g ; F_h]$.

Lorsque l'on considère un système possédant N PM, l'idéal serait de placer les fréquences des PM proches des performances régulièrement demandées, cela revient à considérer les contraintes applicatives. L'approche devient intéressante lorsqu'un petit nombre N de valeurs discrètes est disponible. Si le nombre de PM N est grand (au moins une dizaine de PM), nous n'avons pas besoin de passer par une étape de maximisation de la surface et *in fine* de la réduction en consommation car la méthode est similaire à du DVFS classique car les actionneurs sont considérés comme continus (avec un pas de discrétisation fin).

3.4.1 Synthèse

Nous avons vu que dans le cas où N PM sont disponibles sur le système, il faut maximiser l'aire du polygone si la performance requise appartient à toute la gamme de fréquence.

Dans le cas où la performance requise par le dispositif se situe entre deux fréquences, le fait de maximiser l'aire du polygone délimité par ces fréquences permettra de maximiser la réduction en consommation.

Si le nombre de PM N est grand (au moins une dizaine de PM), cela revient à faire du DVFS classique car le pas de discrétisation des PM est fin.

3.5 Conclusion

Dans ce chapitre, nous avons proposé des méthodes permettant de maximiser le gain en consommation lorsqu'un PM intermédiaire est ajouté dans le cas où 2 PM sont déjà disponibles sur le système.

Nous avons exposé une première méthode pour calculer la moyenne de gain en consommation à partir de la surface du profil gain $G_{3PM}(F_{target})$. Le gain moyen est obtenu *via* MATLAB.

Nous avons aussi proposé une manière d'obtenir la moyenne de réduction en consommation $\overline{PR_{Tri}}$ à l'aide de la surface du triangle formée par les 3 PM ou de la surface du polygone délimité par l'intervalle de fréquence.

Nous avons montré que l'apport réel du PM intermédiaire est obtenue sur le profil de consommation en fonction de la fréquence même si le profil de gain en fonction de la fréquence cible F_{target} est plus attrayant. Ainsi, pour maximiser le gain apporter par le PM intermédiaire, il faut maximiser la surface du triangle ou du polygone délimité par la gamme de fréquence sur le profil de consommation en fonction de la fréquence.

Une solution analytique permettant de placer le PM intermédiaire afin de maximiser le gain en consommation a été fournie.

Si la densité de probabilité que la fréquence cible F_{target} soit demandée est connue, elle peut être intégrée au profil $\overline{PR_{Tri}(F)}$. Cela permet ainsi de maximiser l'espérance de gain.

Nous nous sommes intéressés au cas où le circuit est alimenté par deux actionneurs, à savoir celui de V_{dd} et celui de F . Notons que les méthodes proposées pour maximiser le gain en consommation peuvent être appliquées quels que soient les DoF lorsqu'un nombre fini de PM est disponible sur le circuit. Avec les nouvelles technologies, de nouvelles opportunités de gestion d'un 3^{ième} actionneur (la tension de body bias V_{bb}) s'offrent à nous. Ce nouveau DoF permet de gérer le compromis consommation-performance du circuit. En effet, en modifiant cette tension nous pouvons soit augmenter la performance du circuit pour un même budget consommation soit diminuer la consommation du circuit pour une même performance. Nous allons maintenant nous intéresser à la gestion de la consommation d'un circuit composé de trois DoF.

Chapitre 4

Adaptation dynamique de (F , V_{dd} , V_{bb})

Nous avons étudié dans le chapitre précédent le placement de points de fonctionnement ou PM (Power Mode) lorsque 3 PM sont disponibles sur le circuit. La solution proposée permet de placer le PM intermédiaire afin de maximiser le gain en consommation apporté par ce PM. L'apport a été évalué avec deux degrés de liberté ou DoF (*Degrees of Freedom* en anglais) mais l'étude reste valable en ayant un DoF de plus [75].

Ce chapitre va permettre d'évaluer l'apport de l'ajout d'un degré de liberté aux deux DoF étudiés dans le chapitre 3, c'est-à-dire la tension d'alimentation V_{dd} et la fréquence d'horloge F . Le travail effectué s'inscrit dans le contexte de circuits numériques synchrones. Ces circuits sont alimentés par une tension d'alimentation V_{dd} et cadencés par une horloge à une fréquence F . En fonction de la technologie employée (comme par exemple la technologie SOI), les circuits modernes permettent une modification de la tension de seuil des transistors V_{th} à travers la polarisation du substrat (*body*) des cellules logiques concernées. Une tension de « body bias » (V_{bb}) peut ainsi être appliquée afin de modifier la tension de seuil des transistors et modifier en conséquence la vitesse de commutation des cellules. Cela entraîne une modification de la puissance consommée.

Les entrées du PE F , V_{dd} et V_{bb} sont fortement interdépendantes. Rappelons que la fréquence d'horloge F appliquée ne doit pas dépasser une valeur maximale pour un couple de tension d'alimentation V_{dd} et tension de body bias V_{bb} donné. Autrement dit, les entrées du système V_{dd} et V_{bb} limitent la fréquence maximale F_{max} d'horloge que l'on peut appliquer sans générer d'erreur temporelle, plus connue sous le nom de *timing fault*. Une erreur temporelle se produit quand le signal d'entrée d'une ou plusieurs bascules (*flip-flops*) ne s'établit pas dans une pé-

riode d'horloge. Ainsi, une fréquence maximale F_{max} est associée à chaque couple possible de tension d'alimentation V_{dd} et de tension de body bias V_{bb} . Néanmoins, plusieurs combinaisons possibles de V_{dd} et V_{bb} peuvent permettre d'appliquer la même fréquence maximale d'horloge F_{max} . Aussi, si la tension d'alimentation du circuit est modifiée, la fréquence maximale qui peut être appliquée au circuit sans produire d'erreur temporelle est également influencée. De plus, la fréquence maximale dépend aussi de la tension de polarisation du body (V_{bb}) appliquée.

Dans ce chapitre, les 3 degrés de liberté qui vont être étudiés sont la tension d'alimentation V_{dd} , la fréquence d'horloge F et la tension de body bias V_{bb} , voir la Figure 4.1. L'objectif est de contrôler localement les DoF afin de gérer le compromis consommation-performance du PE. Cette étude permettra d'évaluer si l'exploitation (l'ajout ou l'utilisation) d'un DoF sur un dispositif existant et possédant déjà 2 DoF (F, V_{dd}), permet de réduire sa consommation. Nous allons tout d'abord dresser une synthèse de l'état de l'art dans le cas où les 3 DoF sont disponibles. Nous allons ensuite poser les hypothèses de travail et les définitions utiles pour formaliser la contribution présentée dans ce chapitre. Nous allons également évaluer l'apport que l'on peut avoir avec notre méthode dans le cas où le dispositif est équipé de 3 DoF discrets. L'apport de la méthode sera ensuite évalué avec l'exemple d'un système disposant d'un DoF discret. Nous terminerons ce chapitre par une conclusion.

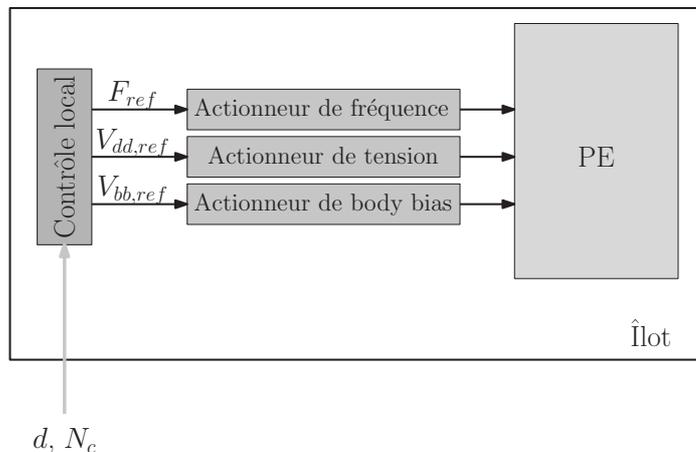


FIGURE 4.1: Schéma d'un îlot avec 3 actionneurs. d est le temps imparti et N_c correspond au nombre de cycles d'horloge.

Ce chapitre a donné lieu à plusieurs publications scientifiques internationales (DAC [76], PATMOS [77]). Un dépôt de brevet [78], montre l'intérêt industriel potentiel.

Nous allons maintenant présenter les principaux travaux de la littérature concernant la gestion des DoF (F, V_{dd}, V_{bb}).

4.1 Synthèse de la gestion d'un circuit possédant 3 DoF

Nous allons faire une synthèse des travaux sur la gestion d'un circuit alimenté par 3 DoF (F, V_{dd}, V_{bb}). À ce jour il n'existe aucun circuit industriel permettant de gérer les 3 DoF dynamiquement. Cependant, il existe des circuits prototypes tels que celui de STMicroelectronics qui propose un processeur composé des 3 DoF (F, V_{dd}, V_{bb}) dans [56], et celui de STMicroelectronics et du CEA permettant de gérer ces 3 DoF de manière dynamique [74]. Le prototype [74] est un processeur de traitement du signal ou DSP (Digital Signal Processing) implémenté en technologie STMicroelectronics FD-SOI 28nm.

Dans le cas où les 3 DoF sont continus, on peut choisir le couple (V_{dd}, V_{bb}) de manière à minimiser la consommation pour la fréquence cible F_{target} . Cela revient à chercher le PM permettant d'atteindre F_{target} et appartenant à l'ensemble convexe. Cet ensemble convexe correspond à la frontière de Pareto. En effet, si l'on souhaite réduire la consommation on doit dans ce cas réduire la fréquence (et la tension d'alimentation) et donc dégrader la performance. Réciproquement, si on veut augmenter la performance, l'augmentation de la consommation est le prix à payer. Afin d'atteindre cette optimum, des capteurs TSM peuvent être utilisés pour trouver le PM à consommation minimale pour une performance donnée.

Dans [38], une analyse en phase de test du circuit est proposée afin d'ajuster la tension d'alimentation V_{dd} et la tension de body bias V_{bb} . L'objectif est d'optimiser la consommation pour une performance cible en phase de test, mais une seule fois pour toute la durée de vie du système. En effet, la tension de body bias V_{bb} est choisie en phase de test puis elle est appliquée. Par la suite, cette tension V_{bb} ne peut plus être modifiée en phase de fonctionnement du circuit. Cette proposition correspond à une manière classique de compenser la variabilité du circuit.

Dans [59], les auteurs analysent la combinaison du DVFS avec une variation de la tension de body bias afin d'effectuer une optimisation multi-critère. Conjointement avec un compromis entre la consommation et la performance, l'étude cherche à améliorer la fiabilité à travers une métrique de « Soft Error Rate » (SER). Les variables contrôlées (tensions d'alimentation, tension de body bias et fréquence) sont considérées continues ou discrétisées avec un pas de discrétisation fin. La méthode est difficilement exploitable dans un contexte où au moins un des actionneurs est de type *hopping*, donc avec peu de valeurs applicables (typiquement 2 ou 3).

La tension de body bias V_{bb} est utilisée pour compenser les variations de procédés de fabrication en fixant la valeur de la tension V_{bb} de manière à améliorer le compromis consommation-performance. Cependant, cette méthode ne permet pas de gérer dynamiquement les 3 DoF. Une gestion dynamique des 3 DoF est proposée dans [59], mais cette solution n'est pas réaliste si les contraintes d'implémentation physique (coût surface, coût production, rendement des actionneurs, etc) sont identiques à notre contexte. Nous allons proposer des méthodes per-

mettant de gérer dynamiquement les 3 DoF tout en considérant les contraintes d'implémentation physique.

4.2 Hypothèses de travail et définitions

Les définitions et les hypothèses sont déjà formulées dans la section 3.2 du chapitre précédent. Rappelons qu'un point de fonctionnement ou PM (de l'anglais Power Mode) est défini comme le couple (F, P_{tot}) avec F la fréquence d'horloge et P_{tot} la puissance consommée associée à F .

Dans ce travail, nous supposons que la tension d'alimentation V_{dd} et la tension de polarisation du body V_{bb} sont choisies telles que la fréquence maximale peut être appliquée au système sans produire d'erreur temporelle. Notons que le critère à minimiser est la puissance consommée par le dispositif, ce qui conduit à redéfinir la première hypothèse de travail.

Hypothèse 1. *La consommation P_{tot} est fonction des entrées du système $(F, V_{dd}$ et $V_{bb})$. Pour deux sorties d'actionneurs données parmi ce triplet, il existe une méthode pour déterminer la valeur de sortie du troisième actionneur afin de minimiser P_{tot} .*

Rappelons que cette première hypothèse assure que la consommation est minimale pour les tensions V_{dd} et V_{bb} (voir la section 3.2). Cela revient à appliquer la fréquence maximale atteignable F_{max} pour le couple (V_{dd}, V_{bb}) . Avec des capteurs de fonctionnalité ou TSM (Timing Slack Monitor) [68], on peut obtenir des informations permettant de vérifier cette hypothèse. En effet, dans le cas où un capteur TSM est utilisé, on peut fixer V_{dd} et V_{bb} dans un premier temps. On peut ensuite augmenter progressivement la fréquence F jusqu'à ce que la fréquence F_{max} soit détectée par le capteur TSM, sans produire d'erreur temporelle. Les TSM permettent de minimiser la consommation P_{tot} pour le couple (V_{dd}, V_{bb}) , la température T° et l'activité A données et pour la fréquence appliquée F_{max} .

La sélection des PM peut s'effectuer avec 3 les DoF F, V_{dd} et V_{bb} . Le but est de minimiser la consommation sous contraintes de performance en proposant une gestion coordonnée des trois entrées du système F, V_{dd} et V_{bb} . Cette gestion doit garantir d'une part que la tâche est exécutée à une fréquence moyenne équivalente à la fréquence cible F_{target} , et d'autre part que les entrées du système F, V_{dd} et V_{bb} sont choisies afin de ne pas produire d'erreur temporelle tout en minimisant la consommation du circuit. De plus, comme on considère que plusieurs tâches, se succèdent, une variation dynamique des entrées du système (F, V_{dd}, V_{bb}) est proposée (i.e. une adaptation en cours d'exécution de l'application). Une telle gestion est réalisée grâce à un ensemble d'actionneurs, avec possiblement un actionneur pour chaque entrée du système à régler F, V_{dd} et V_{bb} .

Il existe dans l'état de l'art une grande variété d'actionneurs capables de gérer chacune de ces entrées indépendamment. Selon l'implémentation choisie, le DoF géré par l'actionneur peut être continu¹ ou continu par morceaux² ou discret³.

Le choix de l'actionneur de tension est primordial car les pertes en consommation du circuit sont aussi liées au rendement de cet actionneur.

Dans un contexte où l'actionneur de tension est embarqué sachant qu'il sera dupliqué dans chaque îlot d'une architecture MPSoC et, en considérant, les contraintes d'implémentation physique (coût surface, coût production, rendement des actionneurs, *etc*), il sera considéré comme discret, par exemple un mécanisme *V_{dd}-Hopping* [28].

Quelle que soit l'implémentation matérielle choisie (*hardware*), les actionneurs V_{bb} et F peuvent être discrets ou continus. Les actionneurs V_{bb} et F sont considérés comme continus s'ils sont capables de produire un grand nombre de valeurs en sortie (comparé au nombre de valeurs de sortie pour la tension d'alimentation V_{dd}). Par exemple, la génération d'un signal d'horloge grâce à une boucle à verrouillage de fréquence (FLL) [42] ou une boucle à verrouillage de phase (PLL) [57] peut être implémentée tandis que V_{bb} peut être fourni par un actionneur discret comme dans [54] ou par une pompe de charge comme dans [55]. Ces aspects permettent de formaliser la seconde hypothèse.

Hypothèse 2 (b). *La tension d'alimentation V_{dd} est fournie par un actionneur discret avec un nombre limité de valeurs de sortie possibles. La tension de body bias V_{bb} et la fréquence d'horloge F sont fournies par des actionneurs discrets ou continus avec un intervalle de valeurs disponible en sortie. Au moins un des actionneurs est considéré comme discret, c'est-à-dire que cet actionneur ne peut fournir qu'un petit (i.e. quelques unités) nombre de valeurs en sortie.*

La troisième hypothèse de travail concerne la connaissance de la consommation totale P_{tot} du dispositif. Il est connu de l'état de l'art que la consommation P_{tot} d'un tel dispositif est fonction des valeurs F , V_{dd} et V_{bb} , fournies par les actionneurs mais aussi de la température du circuit T° , de l'activité A et d'autres paramètres physiques. On considère que la mesure ou l'estimation de la consommation totale P_{tot} a été établie au préalable. Par exemple, P_{tot} peut être estimée

1. « Une fonction $f(x)$, définie dans un voisinage de $x = \epsilon$ et en ϵ lui-même, est dite continue en ϵ si pour tout $\epsilon > 0$ il existe un nombre $\delta(\epsilon)$ tel que $|f(x) - f(\epsilon)| < \epsilon$ pour tout x vérifiant $|x - \epsilon| < \delta(\epsilon)$ » [79].

2. « La fonction $f : [a, b[\rightarrow \mathbb{R}$ est continue par morceaux si et seulement s'il existe une subdivision (x_0, x_1, \dots, x_n) de $[a, b[$ telle que f est continue sur les intervalles $]x_{k-1}, x_k[$ et admet aux points x_k une limite à gauche et une limite à droite (excepté en a et en b où il existe une seule de ces deux limites) » [80].

3. « Une variable aléatoire est dite discrète si elle ne prend que des valeurs discontinues dans un intervalle donné (borné ou non borné) » [81]. Par exemple, si le DoF à gérer est la tension d'alimentation, et que l'actionneur utilisé est un convertisseur DC-DC alors, le DoF est considéré comme continu. Si l'actionneur utilisé est réalisé avec un mécanisme de saut en tension (*V_{dd}-Hopping*) [28], alors le DoF sera considéré comme discret.

comme dans [82] à l'aide de compteurs de performance. Des solutions sont proposées dans la littérature pour mesurer l'activité A à l'aide de compteurs de performance [83]. Des moniteurs de tension d'alimentation ainsi que de température [70] peuvent être ajoutés pour déterminer la consommation du dispositif.

Hypothèse 3. *La consommation totale P_{tot} du dispositif peut être déterminée.*

L'hypothèse suivante concerne la Figure 4.2 qui représente deux morceaux de courbes correspondant aux tensions V_{dd_i} et $V_{dd_{i+1}}$ dans le plan $P_{tot}(F)$ avec V_{dd_i} le $i^{\text{ème}}$ niveau de tension d'alimentation. Notons que la consommation dépend d'autres « perturbations » subies comme la température T° et l'activité A du dispositif.

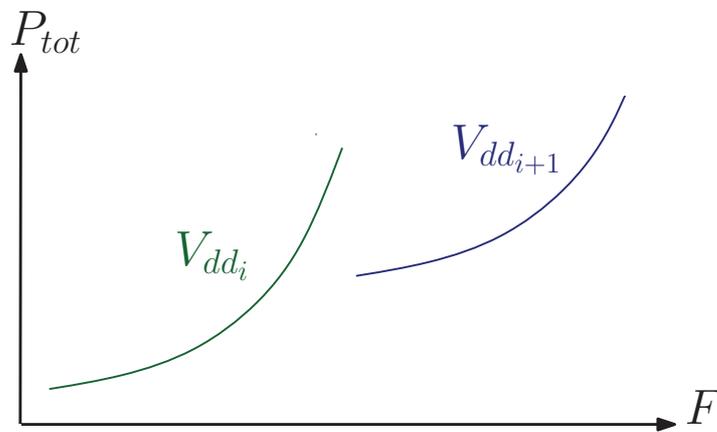


FIGURE 4.2: Ensemble de PM dans le plan (F, P_{tot}) .

On suppose que le dispositif exécute une tâche T et que les exigences en performance sont connues, ce qui est le cas pour les systèmes temps-réel. Par exemple, l'horizon d'application de F_{target} peut être fourni par un système haut niveau (voir [72]).

D'un point de vue applicatif, on suppose que le circuit exécute une tâche déterminée et que cette tâche a une contrainte de performance, correspondant à un temps imparti pour exécuter la tâche. Autrement dit, la tâche doit être complètement exécutée au plus tard avant une date limite. Pour pouvoir finir la tâche à temps, elle doit être exécutée à une performance moyenne équivalente à celle du circuit cadencé avec une fréquence d'horloge cible F_{target} . Cette dernière est la fréquence d'horloge qui devrait être appliquée au circuit pour exécuter la tâche avec la performance requise. On considère qu'une suite de tâches se succèdent et que la fréquence F_{target} n'est pas forcément identique pour toutes les tâches considérées.

Hypothèse 4. *Le circuit exécute une tâche T qui requiert N_C cycles d'horloge pour finir la tâche T avant une date limite d connue. Ainsi, la tâche doit être exécutée à une fréquence équivalente à F_{target} . De plus, au moins deux paramètres parmi N_C , d et F_{target} sont connus.*

Ayant défini l'ensemble des hypothèses nécessaires à la mise en place du cadre de travail, nous pouvons désormais définir la problématique à résoudre, à savoir, comment choisir l'ensemble des entrées du système (F, V_{dd}, V_{bb}) qui optimisent la consommation globale du dispositif. Il a été prouvé dans [67] que le compromis consommation-performance est une fonction convexe. Cette affirmation demeure exacte pour les techniques traditionnelles de gestion de la consommation comme par exemple l'adaptation dynamique de tension et de fréquence (Dynamic Voltage and Frequency Scaling ou DVFS) où les DoF sont F et V_{dd} . La possibilité de modifier la tension de polarisation du substrat ajoute un nouveau DoF. Cependant, quand on prend en compte les hypothèses définies dans ce chapitre, **le compromis consommation-performance n'est plus globalement convexe**. Ainsi le problème à résoudre est le suivant.

Problème. *On souhaite choisir le triplet (F, V_{dd}, V_{bb}) ayant la consommation minimale pour atteindre la fréquence requise F_{target} pour la tâche T en considérant les hypothèses définies dans ce chapitre, sachant que plusieurs combinaisons (V_{dd}, V_{bb}) peuvent être appliquées pour exécuter la tâche sans produire d'erreur temporelle en appliquant F .*

4.3 Gestion d'un circuit possédant trois DoF discrets

Les PM sont obtenus grâce à 3 actionneurs qui permettent de modifier la tension d'alimentation V_{dd} , la tension de polarisation du body V_{bb} et la fréquence d'horloge F d'un îlot (voir Figure 4.1). Dans les MPSoC, les capteurs et les actionneurs peuvent être intégrés dans chaque îlot. Afin de réduire le surcoût en surface, des mécanismes ayant des sorties discrètes seront privilégiés aux actionneurs continus. Ici, V_{dd} est considéré choisi dans un ensemble fini, fourni par exemple, par un mécanisme de V_{dd} -hopping. La fréquence d'horloge F est aussi supposée discrète avec un petit nombre de valeurs atteignables. Elle peut être fournie par exemple par une FLL [42]. La tension de polarisation du body V_{bb} prend ses valeurs dans un ensemble fini et discret. Elle peut être appliquée grâce à une technique de V_{bb} -hopping [54]. Ainsi, il existe un nombre fini N de PM $S = \{(F_i, P_{tot}(F_i)), i = 1 : N\}$, où $F_i \leq F_{i+1}$ pour tout $i = 1 : N - 1$. Cet ensemble n'est pas forcément discrètement convexe comme défini dans [84] et comme précisé dans [67].

On propose maintenant, une méthode pour sélectionner la distribution des PM. Cette méthode est basée sur la notion de convexité. Elle étend les résultats

de [67]. Notons que la méthode que l'on a proposée dans [66] pour choisir la distribution des PM peut être appliquée à l'ensemble de PM.

Définition 1. *Un ensemble est défini convexe sur un intervalle I si et seulement si (voir Figure 4.3) $\forall F_1, F_2 \in I, \forall \beta \in [0, 1]$,*

$$P_{tot}(\beta \cdot F_1 + (1 - \beta) \cdot F_2) \leq \beta \cdot P_{tot}(F_1) + (1 - \beta) \cdot P_{tot}(F_2) \quad (4.1)$$

avec $P_{tot}(F)$ la consommation correspondant à la fréquence F .

En d'autres termes, un ensemble est défini comme convexe si chaque PM_β de la courbe $P_{tot}(F)$ entre deux PM, PM_1 et PM_2 , se situe sous le segment reliant les PM_1 et PM_2 .

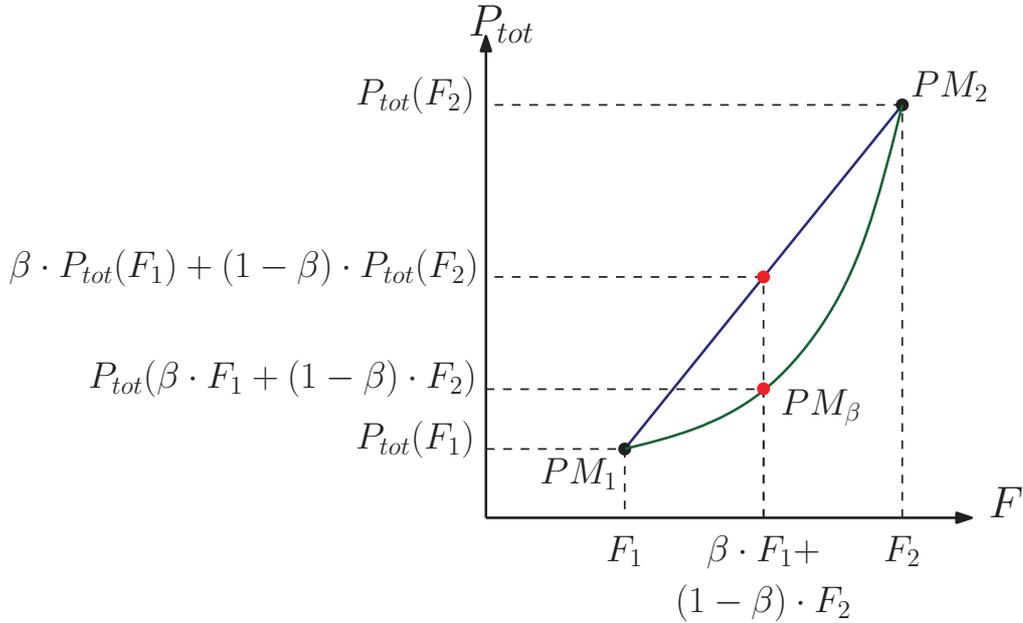


FIGURE 4.3: Exemple d'un ensemble convexe croissant $P_{tot}(F)$.

Des travaux précédents [67] ont présenté des résultats en considérant 2 degrés de liberté ou DoF (F, V_{dd}) pour gérer le dispositif. Les auteurs de [67] ont prouvé qu'exécuter une tâche avec les fréquences voisines de la fréquence cible permet de minimiser la consommation si le rapport de consommation sur la vitesse P_{tot}/F en fonction de la vitesse $1/F$ est convexe. Nous allons montrer que cette affirmation demeure exacte dans le cas d'un ensemble discrètement convexe croissant ou DCS de PM. Contrairement aux auteurs de [67], nous travaillerons directement sur la courbe $P_{tot}(F)$.

4.3.1 Propriété d'une fonction discrètement convexe

Soit \mathbb{F} l'ensemble discret des fréquences $F_n \in \mathbb{R}^+$ applicables au circuit et \mathbb{P} l'ensemble des consommations $P_{tot} \in \mathbb{R}^+$ tel que $P_{tot} : F_n \rightarrow P_{tot}(F_n), \forall n \in \mathbb{E}, \mathbb{E} \subset \mathbb{N}^* \setminus \{1\}$. On suppose que la fonction $P_{tot}(F)$ est croissante. Les fréquences $F_{n-1}, F_n, F_{n+1} \in \mathbb{F}$ sont définies telles que $F_{n-1} < F_n < F_{n+1}$. Comme \mathbb{P} est un ensemble croissant sur \mathbb{F} , on a $P_{tot}(F_{n-1}) \leq P_{tot}(F_n) \leq P_{tot}(F_{n+1})$.

On note $P_{tot}^{(i,j)}(F_n)$ (voir Figures 4.4 et 4.5), la valeur de puissance correspondant à la fréquence F_n qui se situe sur le segment de droite qui relie les PM_i et PM_j avec $i < j, \forall i, j \in \mathbb{E}$. La notation a_i^j (voir Figures 4.4 et 4.5) correspond à la pente du segment de droite reliant les PM_i et PM_j .

Prenons par exemple, la valeur de la consommation située sur le segment de droite reliant PM_3 et PM_6 pour la fréquence F_4 . Elle sera notée $P_{tot}^{(3,6)}(F_4)$. La pente du segment de droite reliant PM_3 et PM_6 est notée a_3^6 .

Preuve pour la fréquence supérieure à F_{target}

Soit $F_{target} \notin \mathbb{F}$ la fréquence à atteindre en moyenne grâce à des techniques de saut (*hopping*) avec $F_{n-1} < F_{target} < F_n$ tels que sur la Figure 4.4. Le but est de démontrer qu'une exécution en mode *hopping* avec F_{n-1} et F_n n'aura pas une consommation supérieure à une exécution en mode *hopping* avec F_{n-1} et F_{n+k} , $\forall k, 1 \leq k \leq N - n$, si \mathbb{P} forme un ensemble croissant et discrètement convexe sur \mathbb{F} . Cela revient à prouver que :

$$P_{tot}^{(n-1,n)}(F_{target}) \leq P_{tot}^{(n-1,n+k)}(F_{target})$$

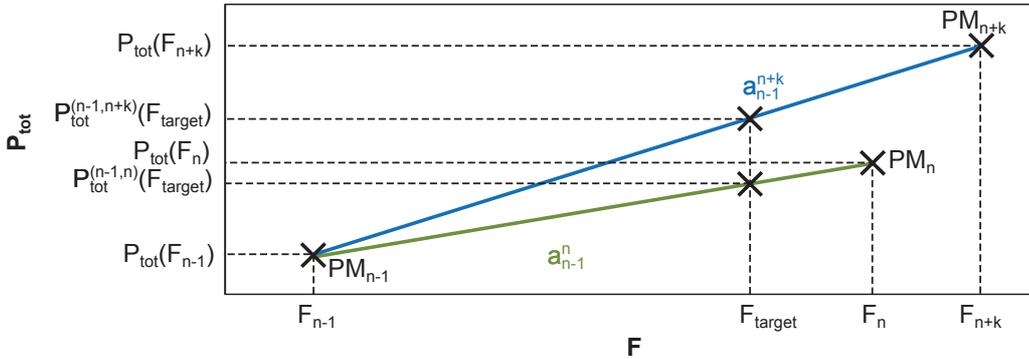


FIGURE 4.4: Exemple de 3 fréquences applicables (F_{n-1}, F_n, F_{n+k}) dans le plan (F, P_{tot}) .

Preuve 1. La consommation correspondant à F_{target} sur le segment reliant les PM_{n-1} et PM_{n+k} est donnée par :

$$P_{tot}^{(n-1,n+k)}(F_{target}) = a_{n-1}^{n+k} \cdot (F_{target} - F_{n-1}) + P_{tot}(F_{n-1}) \quad (4.2)$$

On calcule la valeur correspondant à F_{target} sur le segment qui relie les PM_{n-1} et PM_n :

$$P_{tot}^{(n-1,n)}(F_{target}) = a_{n-1}^n \cdot (F_{target} - F_{n-1}) + P_{tot}(F_{n-1}) \quad (4.3)$$

En soustrayant (4.2) et (4.3) membre à membre, il vient :

$$P_{tot}^{(n-1,n+k)}(F_{target}) - P_{tot}^{(n-1,n)}(F_{target}) = (a_{n-1}^{n+k} - a_{n-1}^n) \cdot (F_{target} - F_{n-1}) \quad (4.4)$$

Par hypothèse, on a :

$$F_{n-1} < F_{target} < F_n \Rightarrow F_{target} - F_{n-1} > 0$$

Par hypothèse, \mathbb{P} est un ensemble croissant discrètement convexe sur \mathbb{F} . Par conséquence, on a :

$$a_{n-1}^n \leq a_{n-1}^{n+k} \Leftrightarrow a_{n-1}^{n+k} - a_{n-1}^n \geq 0$$

On peut donc en déduire que :

$$P_{tot}^{(n-1,n+k)}(F_{target}) - P_{tot}^{(n-1,n)}(F_{target}) \geq 0 \Leftrightarrow P_{tot}^{(n-1,n+k)}(F_{target}) \geq P_{tot}^{(n-1,n)}(F_{target})$$

On a prouvé que l'utilisation de la fréquence immédiatement supérieure à F_{target} minimisera la consommation lorsque la tâche est exécutée en mode *hopping* si l'ensemble des consommations correspondant aux fréquences applicables forme un ensemble croissant discrètement convexe dans le plan (F, P_{tot}) .

Preuve pour la fréquence inférieure de F_{target}

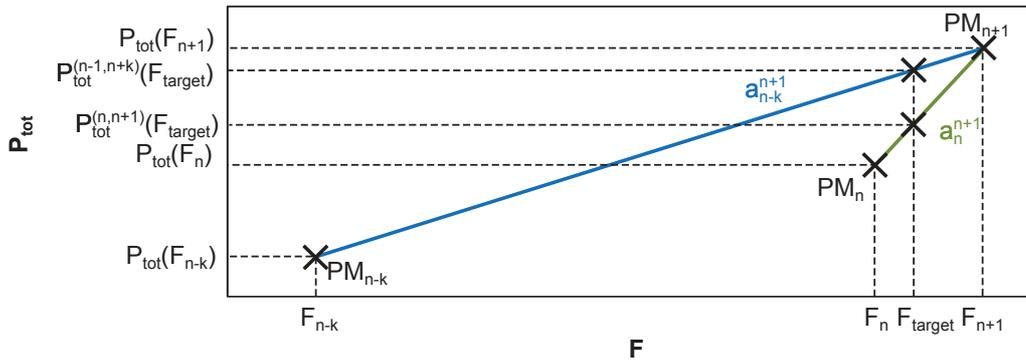


FIGURE 4.5: Exemple de 3 fréquences applicables (F_{n-k}, F_n, F_{n+1}) dans le plan (F, P_{tot}) .

Soit $F_{target} \notin \mathbb{F}$ la fréquence à atteindre par la technique dite de *hopping* avec $F_n < F_{target} < F_{n+1}$ tels que sur la Figure 4.5. On veut démontrer qu'une exécution en mode *hopping* avec F_n et F_{n+1} ne consommera pas plus qu'une

exécution en mode *hopping* avec F_{n-k} et F_{n+1} , $\forall k, 1 \leq k \leq n-1$, si \mathbb{P} est un ensemble croissant et discrètement convexe sur \mathbb{F} . Cela revient à démontrer que :

$$P_{tot}^{(n-k,n+1)}(F_{target}) \geq P_{tot}^{(n,n+1)}(F_{target})$$

Preuve 2. On calcule la valeur correspondant à F_{target} sur le segment de droite qui relie les PM_{n-k} et PM_{n+1} :

$$P_{tot}^{(n-k,n+1)}(F_{target}) = a_{n-k}^{n+1} \cdot (F_{target} - F_{n+1}) + P_{tot}(F_{n+1}) \quad (4.5)$$

La consommation correspondant à F_{target} sur le segment de droite reliant les PM_n et PM_{n+1} est :

$$P_{tot}^{(n,n+1)}(F_{target}) = a_n^{n+1} \cdot (F_{target} - F_{n+1}) + P_{tot}(F_{n+1}) \quad (4.6)$$

En soustrayant (4.5) et (4.6) membre à membre, on trouve l'équation :

$$P_{tot}^{(n-k,n+1)}(F_{target}) - P_{tot}^{(n,n+1)}(F_{target}) = (a_{n-k}^{n+1} - a_n^{n+1}) \cdot (F_{target} - F_{n+1}) \quad (4.7)$$

Par hypothèse, on a :

$$F_n < F_{target} < F_{n+1} \Leftrightarrow F_{target} - F_{n+1} < 0$$

Par hypothèse, \mathbb{P} est un ensemble croissant discrètement convexe sur \mathbb{F} . On obtient donc :

$$a_{n-k}^{n+1} \leq a_n^{n+1} \Leftrightarrow a_{n-k}^{n+1} - a_n^{n+1} \leq 0$$

On peut donc en déduire :

$$P_{tot}^{(n-k,n+1)}(F_{target}) - P_{tot}^{(n,n+1)}(F_{target}) \geq 0 \Leftrightarrow P_{tot}^{(n-k,n+1)}(F_{target}) \geq P_{tot}^{(n,n+1)}(F_{target})$$

On a prouvé que l'utilisation de la fréquence immédiatement inférieure à F_{target} minimisera la consommation lors de l'exécution de la tâche en mode *hopping* si l'ensemble des consommations correspondant aux fréquences que l'on peut appliquer forme un ensemble croissant discrètement convexe dans le plan (F, P_{tot}) .

Finalement, on peut en déduire que l'exécution de la tâche avec le couple de fréquences immédiatement supérieure et inférieure à F_{target} en mode *hopping* minimisera la consommation si l'ensemble des points (F, P_{tot}) disponibles forme un ensemble croissant et discrètement convexe dans le plan (F, P_{tot}) .

Notons que cette proposition s'étend à un ensemble convexe par morceaux ou à un ensemble convexe $P_{tot}(F)$.

Ce résultat est proche de celui de [67] où les auteurs prouvent que la vitesse optimale est atteinte pour une exécution de la tâche en combinant les deux fréquences voisines appartenant à l'ensemble convexe du rapport de consommation sur la vitesse en fonction de la vitesse. Ici, la preuve est conduite directement sur P_{tot} , et non sur le rapport de consommation sur la vitesse.

Ainsi, pour assurer une exécution optimale en mode *hopping*, un sous-ensemble discrètement convexe croissant ou DCS de PM doit être extrait de l'ensemble de PM. Cependant, les auteurs de [67] affirment que la fonction $P_{tot}(F)$ est convexe. Cette affirmation est exacte dans leur contexte car ils ne gèrent que 2 DoF. Cependant, lorsque 3 DoF doivent être gérés, la propriété de convexité ne peut plus être garantie.

4.3.2 Méthode proposée pour gérer le compromis consommation-performance

La méthode proposée est constituée de deux parties distinctes (voir les Figures 4.6 et 4.7) [77] [78]. La première consiste à trouver le sous-ensemble discrètement convexe croissant ou DCS de l'ensemble de PM car l'utilisation de 3 actionneurs dont 1 discret ne mène pas forcément à un ensemble de PM convexe (étapes 1 à 4 sur la Figure 4.6). La seconde partie de la méthode consiste à exécuter la tâche avec les deux fréquences qui encadrent la fréquence cible F_{target} dans le DCS (étapes 5 à 7 sur la Figure 4.7). L'objectif est d'atteindre en moyenne la fréquence cible afin de satisfaire les contraintes en performance, cette fréquence pouvant être atteinte en appliquant un seul PM ou une combinaison de PM. La méthode vise à réduire la consommation du dispositif quel que soit le ratio de la consommation statique sur la consommation dynamique.

Pour résumer, comme l'ensemble de PM ne remplit pas forcément la propriété de convexité, un sous-ensemble de PM doit être défini pour remplir l'hypothèse de convexité discrète. Ainsi, une exécution en mode *hopping* garantira la consommation minimum du dispositif pour la performance visée. Notons que le travail ci-présent traite directement des valeurs de consommation, sans utiliser une équation liant la consommation aux entrées du système et aux différents paramètres physiques.

Identification du sous-ensemble discrètement convexe croissant

Nous avons vu précédemment qu'il faut mettre en place une phase d'identification du DCS de l'ensemble de PM (étape 3 sur la Figure 4.6). Pour identifier le sous-ensemble des PM appartenant au DCS, nous avons besoin du profil $P_{tot}(F)$. Nous avons aussi besoin de calculer, $\forall n \in \mathbb{E}$, $\mathbb{E} \subset \mathbb{N}^* \setminus \{1\}$, la valeur des 3 pentes suivantes a_{n-1}^n , a_{n-1}^{n+1} et a_n^{n+1} . La pente du segment de droite reliant les PM_{n-1} et PM_n est donnée par :

$$a_{n-1}^n = \frac{P_{tot}(F_n) - P_{tot}(F_{n-1})}{F_n - F_{n-1}}. \quad (4.8)$$

De même, la pente du segment de droite qui relie les PM_{n-1} et PM_{n+1} est définie par :

$$a_{n-1}^{n+1} = \frac{P_{tot}(F_{n+1}) - P_{tot}(F_{n-1})}{F_{n+1} - F_{n-1}}. \quad (4.9)$$

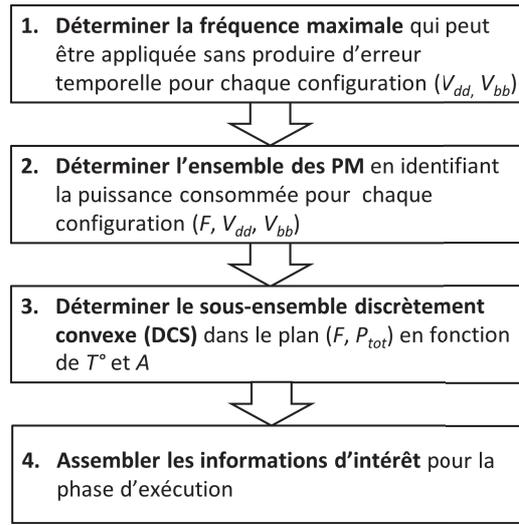


FIGURE 4.6: Phase d'identification du DCS permettant de déterminer les modes de fonctionnement possibles et les modes optimaux (étapes 1 à 4)

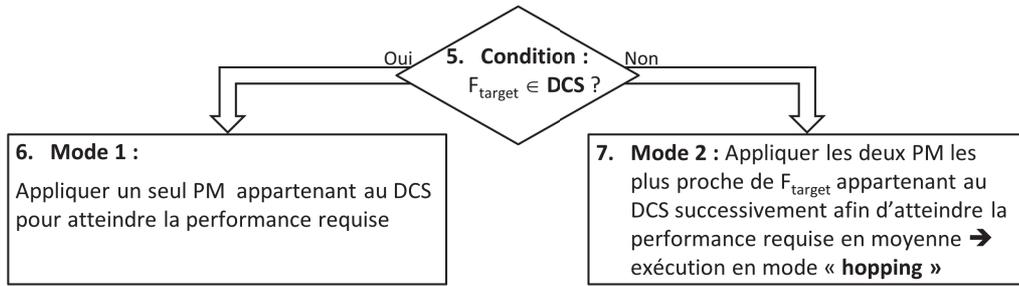


FIGURE 4.7: Phase de fonctionnement (exécution de la tâche) (étapes 5 à 7).

La valeur de la pente du segment reliant les PM_n et PM_{n+1} est calculée par :

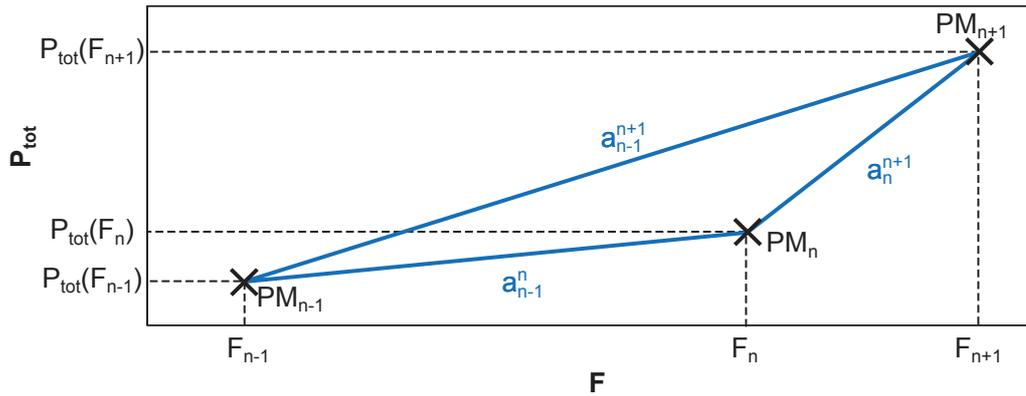
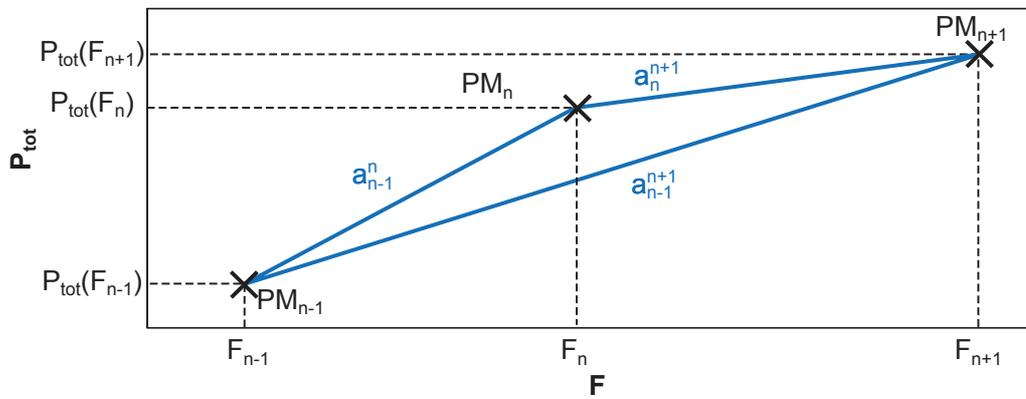
$$a_n^{n+1} = \frac{P_{tot}(F_{n+1}) - P_{tot}(F_n)}{F_{n+1} - F_n}. \quad (4.10)$$

Le PM dont la consommation est minimale appartient au DCS. Les PM dont la fréquence est inférieure à la fréquence de ce PM n'appartiennent pas au DCS.

Le PM dont la fréquence est maximale appartient au DCS. Les PM dont la consommation est supérieure à la consommation de ce PM n'appartiennent pas au DCS.

$\forall n \in \mathbb{E}, \mathbb{E} \subset \mathbb{N}^* \setminus \{1\}$, si $a_{n-1}^n \leq a_{n-1}^{n+1} \leq a_n^{n+1}$ alors le PM_n appartient au DCS (voir Figure 4.8). Si $a_{n-1}^n > a_{n-1}^{n+1} > a_n^{n+1}$, le PM_n n'appartient pas au DCS (voir Figure 4.9). Notons que les pentes doivent être positives.

$\forall n \in \mathbb{E}, \mathbb{E} \subset \mathbb{N}^* \setminus \{1\}$, si une des pentes a_{n-1}^n est négative alors que toutes les pentes précédentes étaient positives. Il est possible que les PM identifiés précédemment comme appartenant au DCS n'y appartiennent pas. Dans ce cas, il

FIGURE 4.8: $P_{tot}(F)$ avec PM_n appartenant au DCS.FIGURE 4.9: $P_{tot}(F)$ avec PM_n n'appartenant pas au DCS.

faut recalculer les pentes entre le PM_n et les PM identifiés comme appartenant au DCS, PM_{DCS_i} . Si la pente entre les PM_n et PM_{DCS_i} est positive alors le PM_{DCS_i} appartient DCS. Sinon, le PM_{DCS_i} n'appartient pas au DCS.

Ainsi, le sous-ensemble discrètement convexe croissant (DCS) est identifié.

Exécution des tâches

Après avoir identifié l'ensemble des PM appartenant au sous-ensemble discrètement convexe croissant (DCS), nous proposons d'utiliser uniquement ces PM lors de l'exécution de la tâche. La phase d'exécution correspond aux étapes 5 à 7 sur la Figure 4.7. Cette proposition permet de minimiser la consommation par rapport à une exécution de la tâche sans la phase de sélection des PM. Cette affirmation a été démontrée ci-avant.

Les Figures 4.10, 4.11 et 4.12 illustrent la consommation P_{tot} en fonction de la fréquence F avec les PM appartenant au DCS en vert et les PM n'appartenant pas au DCS en rouge.

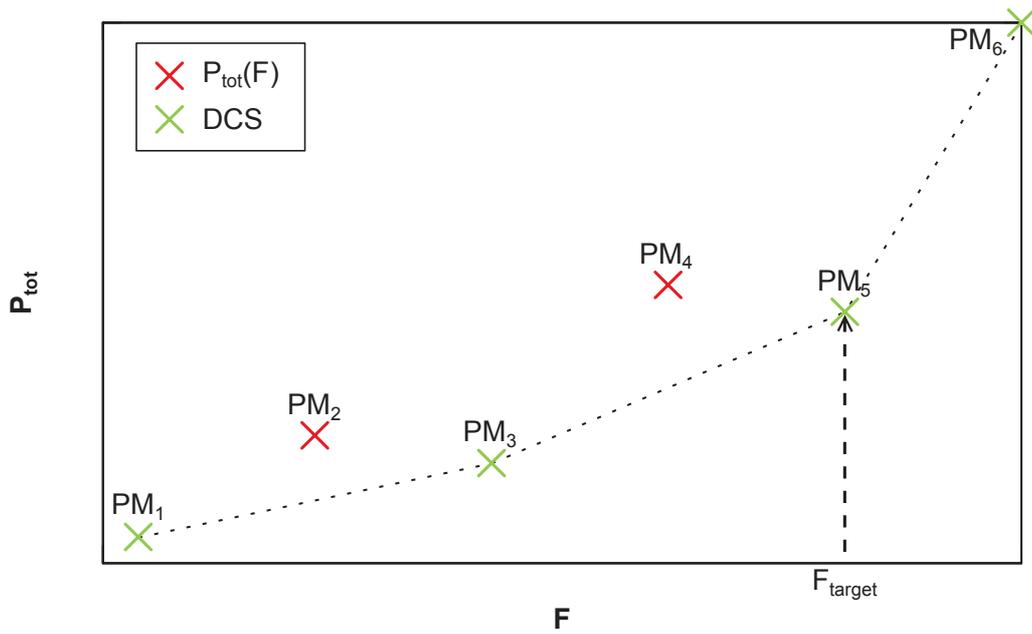


FIGURE 4.10: $P_{tot}(F)$ avec F_{target} correspondant à la fréquence d'un PM appartenant au DCS.

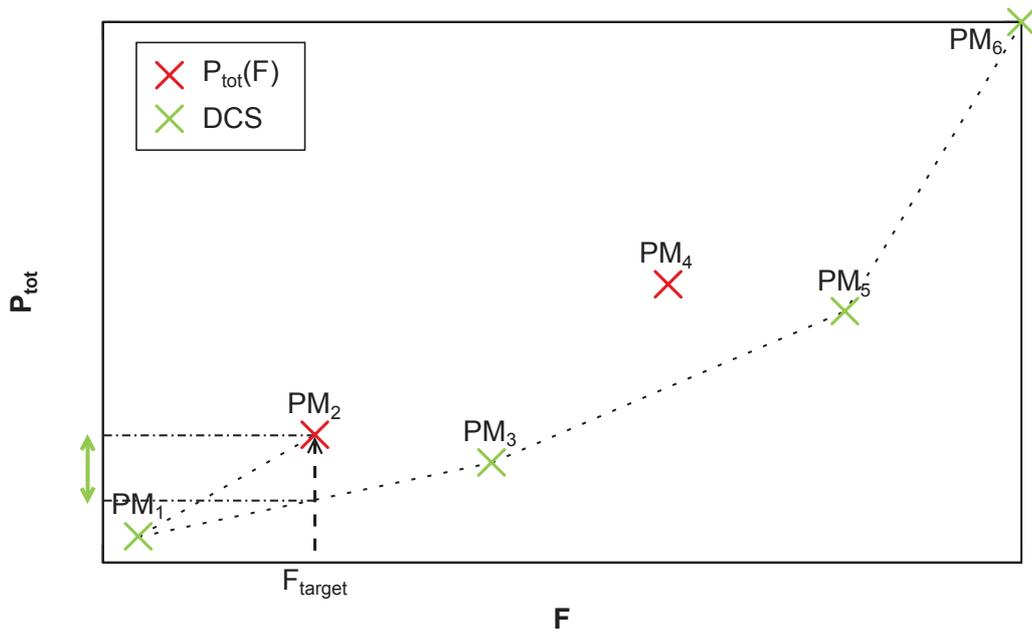


FIGURE 4.11: $P_{tot}(F)$ avec F_{target} correspondant à la fréquence d'un PM n'appartenant pas au DCS.

Différentes situations peuvent être rencontrées lors de l'exécution d'une tâche :

- si la fréquence cible F_{target} est disponible sur le dispositif et si la fréquence

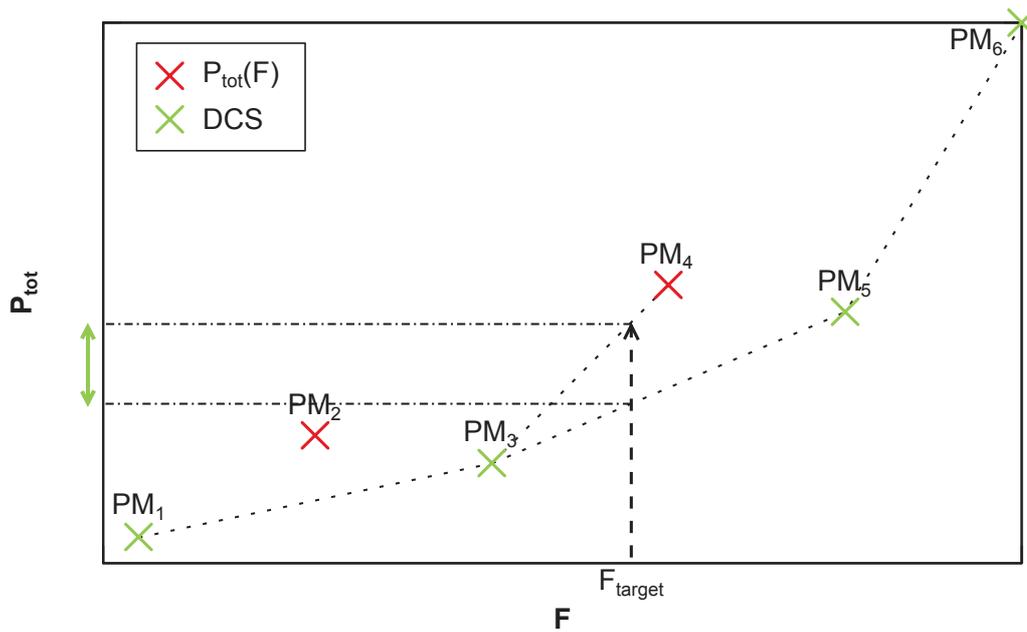


FIGURE 4.12: $P_{tot}(F)$ avec F_{target} ne correspondant à la fréquence d'aucun PM.

du PM correspondant à F_{target} appartient au DCS, le PM sera appliqué directement pour exécuter la tâche (voir la Figure 4.10) ;

- si la fréquence cible F_{target} est applicable et si la fréquence du PM correspondant à F_{target} n'appartient pas au DCS, les 2 PM appartenant au DCS encadrant F_{target} seront appliqués pour exécuter la tâche (voir la Figure 4.11) ;
- si la fréquence cible F_{target} n'existe pas sur le dispositif, les 2 PM dont les fréquences encadrent F_{target} et qui appartiennent au DCS seront appliqués pour exécuter la tâche (voir la Figure 4.12).

4.3.3 Conclusion

Le choix du (des) PM à appliquer pour minimiser la consommation sous contrainte de performance en agissant sur le triplet (F, V_{dd}, V_{bb}) n'est pas trivial lorsque les hypothèses 1 à 4 sont prises en compte. Nous avons vu qu'il fallait passer par une phase d'identification du sous-ensemble discrètement convexe croissant ou DCS quand le circuit est géré à travers 3 DoF discrets. Concernant la phase de fonctionnement, la tâche T sera exécutée en appliquant un (des) PM appartenant au DCS.

Nous verrons dans le chapitre 5 des exemples pour les différents scénarios d'exécution.

Nous avons étudié dans cette section le cas de 3 DoF discrets. Cependant, un des DoF peut être considéré continu, ou tel que défini dans notre modèle, avec un

grand nombre de valeurs possibles en sortie (en comparaison des quelques unités pour un actionneur discret). Dans la section ci-après, nous étudierons le cas de deux DoF continus associés à un DoF discret.

4.4 Gestion d'un circuit possédant trois DoF dont un discret

Les PM sont obtenus grâce à 3 DoF permettant de modifier la tension d'alimentation V_{dd} , la tension de polarisation du body V_{bb} et la fréquence d'horloge F , un seul de ces DoF étant considéré discret. On considère maintenant que l'actionneur V_{dd} fournit un ensemble fini de valeurs de sortie.

Ainsi, il existe un nombre infini de PM. **Cependant, cet ensemble n'est pas forcément convexe par morceaux comme défini dans [85].**

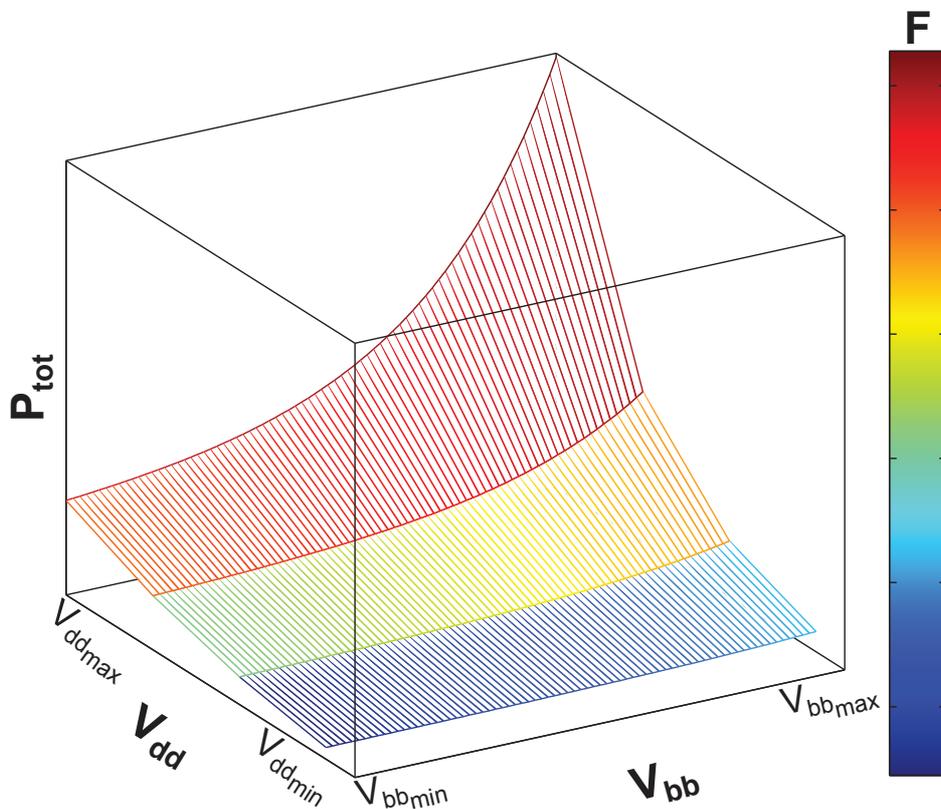


FIGURE 4.13: Consommation P_{tot} d'un oscillateur en anneau en fonction de $V_{bb} \in [V_{bb_{min}} ; V_{bb_{max}}]$ et de $V_{dd} \in \{V_{dd_{min}} ; V_{dd_{max}}\}$. La coloration correspond à la fréquence maximale (F_{max}) applicable pour chaque couple (V_{dd}, V_{bb}) .

La Figure 4.13 illustre la consommation P_{tot} d'un oscillateur en anneau en fonction de V_{dd} et de V_{bb} , avec $V_{dd} \in \{V_{dd_{min}} ; V_{dd_{max}}\}$ et $V_{bb} \in [V_{bb_{min}} ; V_{bb_{max}}]$. Pour chaque couple (V_{dd}, V_{bb}) , on voit la consommation et la gamme de fréquence

associées. En effet, la surface est colorée en fonction de la fréquence maximale (F_{max}) applicable sans produire d'erreur temporelle pour chaque couple (V_{dd}, V_{bb}). Étant donné les hypothèses précédentes et une fois que la fréquence équivalente F_{target} est connue, il faut choisir le(s) couple(s) (V_{dd}, V_{bb}) à appliquer pour pouvoir atteindre la performance équivalente à cette fréquence. Plusieurs cas sont possibles.

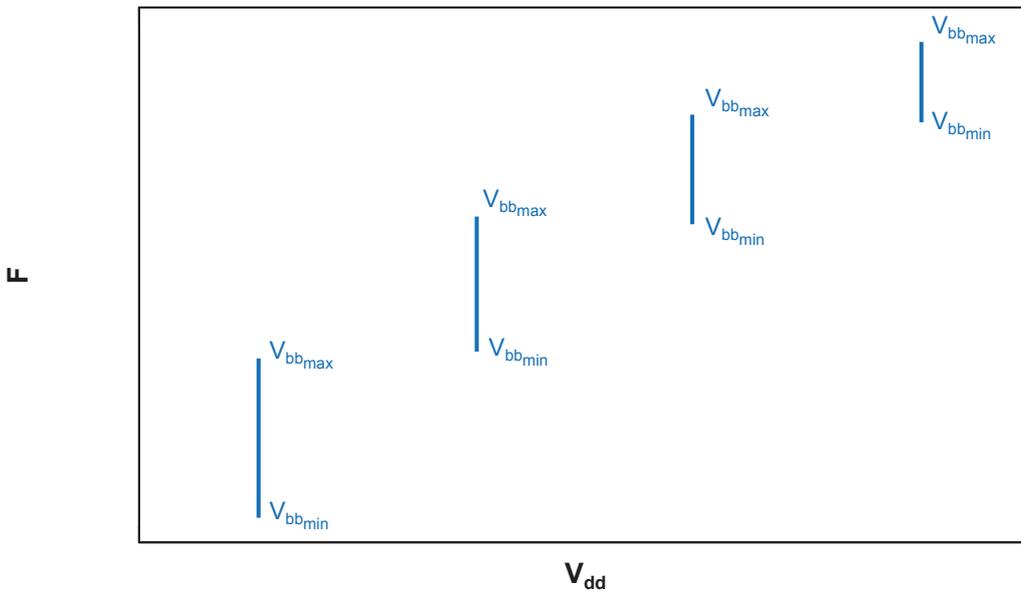


FIGURE 4.14: Fréquence maximale (F_{max}) applicable en fonction de $V_{bb} \in [V_{bb_{min}} ; V_{bb_{max}}]$ pour 4 valeurs de V_{dd} .

Dans l'exemple illustré par la Figure 4.14, on a tracé la fréquence maximale F_{max} que l'on peut appliquer au circuit sans avoir d'erreur temporelle en fonction des deux variables V_{dd} et V_{bb} dans le cas où la tension d'alimentation V_{dd} prend 4 valeurs différentes et la tension de body bias appartient à un intervalle $[V_{bb_{min}} ; V_{bb_{max}}]$. Cette figure nous permet d'obtenir des informations sur la tension d'alimentation à appliquer pour pouvoir atteindre une gamme de fréquence. Pour appliquer une fréquence cible, il faudra tout d'abord choisir la tension d'alimentation permettant d'atteindre cette fréquence puis ajuster la tension de body bias afin d'atteindre la fréquence maximale sans produire d'erreur temporelle. La fréquence maximale en fonction de V_{dd} et V_{bb} peut être obtenue par exemple grâce à un capteur TSM [68].

La Figure 4.15 représente la fréquence maximale applicable au circuit en fonction de V_{dd} et V_{bb} . On peut y apercevoir l'exemple d'une fréquence cible atteignable en appliquant un unique couple (V_{dd}, V_{bb}). Si la valeur F_{target} n'est accessible que par un couple (V_{dd}, V_{bb}), une solution possible est d'appliquer ce couple pour atteindre cette fréquence (voir la Figure 4.15 pour un exemple avec

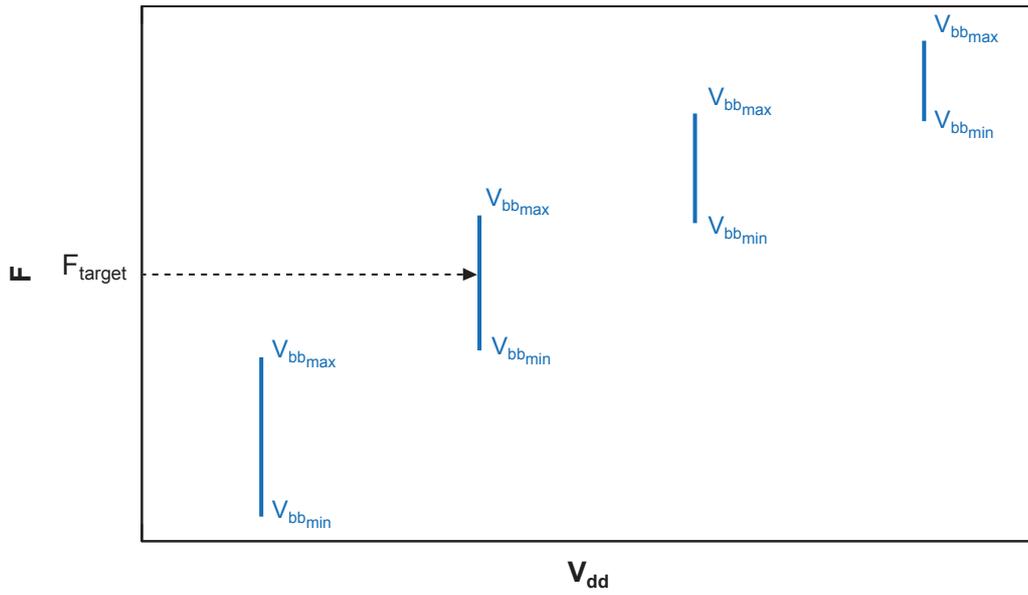


FIGURE 4.15: Exemple avec 4 valeurs de V_{dd} , V_{bb} continu et F_{target} atteignable par un seul couple (V_{dd}, V_{bb}) .

4 niveaux de tension V_{dd}).

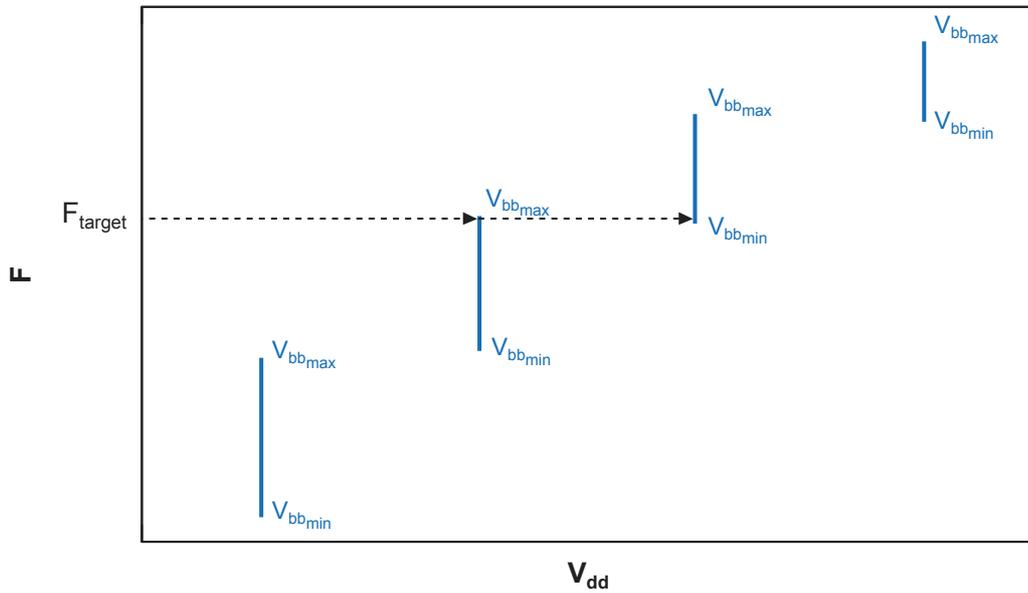


FIGURE 4.16: Exemple avec 4 valeurs de V_{dd} , V_{bb} continu et F_{target} atteignable par deux couples (V_{dd}, V_{bb}) .

Dans le cas où la fréquence F_{target} est atteignable par au moins deux couples (V_{dd}, V_{bb}) (voir Figure 4.16), le problème à résoudre est de choisir le couple à appliquer pour minimiser la puissance consommée par le circuit en garantissant

la performance requise.

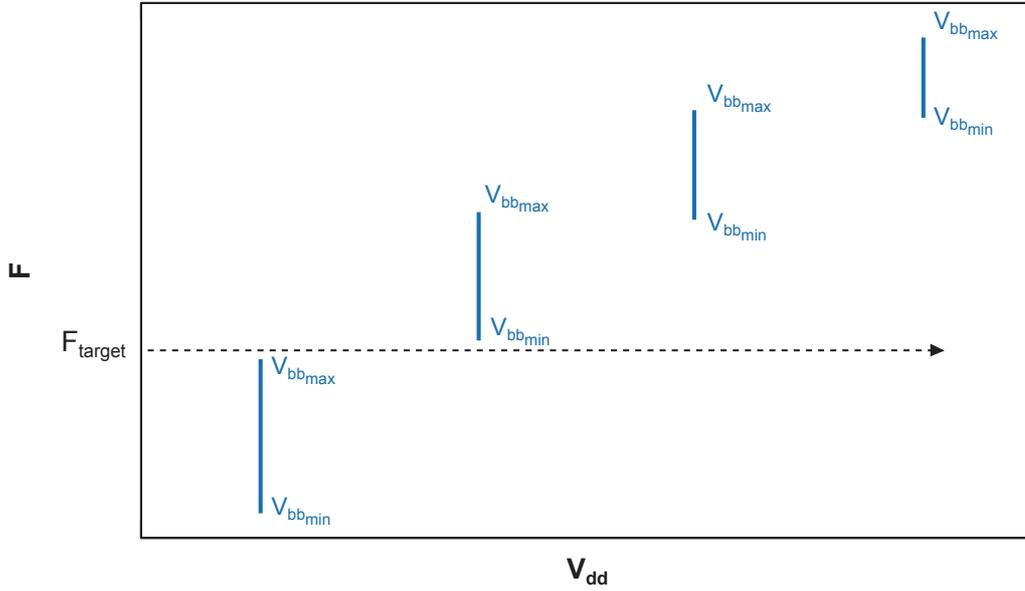


FIGURE 4.17: Exemple avec 4 valeurs de V_{dd} , V_{bb} continu et F_{target} non atteignable par un couple (V_{dd}, V_{bb}) .

Enfin, on peut se trouver dans la situation où la fréquence cible F_{target} ne correspond pas à une fréquence F_{max} atteignable par le circuit (voir Figure 4.17). Dans ce cas, une solution est d'appliquer un couple (V_{dd}, V_{bb}) correspondant à une fréquence $F_{max} > F_{target}$. Néanmoins, la consommation du circuit ne sera pas, *a priori*, minimale. Une deuxième solution possible est d'implémenter un système dit de *hopping* sur toutes les variables (F, V_{dd}, V_{bb}) pour pouvoir atteindre en moyenne une performance équivalente à F_{target} . Il faudra donc exécuter une partie de la tâche avec une fréquence F_2 supérieure à F_{target} et le reste de la tâche avec une fréquence F_1 inférieure à F_{target} . Pour pouvoir modifier la fréquence, on doit modifier la tension d'alimentation V_{dd} et/ou la tension de body bias V_{bb} . Sur la Figure 4.18, α représente le ratio de temps que l'on va passer à la fréquence F_2 , avec $\alpha = \frac{F_{target} - F_1}{F_2 - F_1}$, le reste du temps la tâche étant exécutée à la fréquence F_1 .

Le problème est alors le suivant : comment choisir V_{dd} et V_{bb} pour pouvoir appliquer une fréquence équivalente F_{target} et minimiser la consommation du circuit ?

Pour répondre à cette question, nous traçons sur la Figure 4.19 les informations données sur les Figures 4.13 et 4.14 dans le plan (F, P_{tot}) . Cela revient à couper en 4 points la surface $P_{tot}(V_{dd}, V_{bb})$ selon l'axe de tension d'alimentation et d'utiliser les informations concernant le tracé $F_{max}(V_{dd}, V_{bb})$. La Figure 4.21 permet d'avoir toutes les informations $(P_{tot}, F, V_{dd}$ et $V_{bb})$ sur une seule figure et

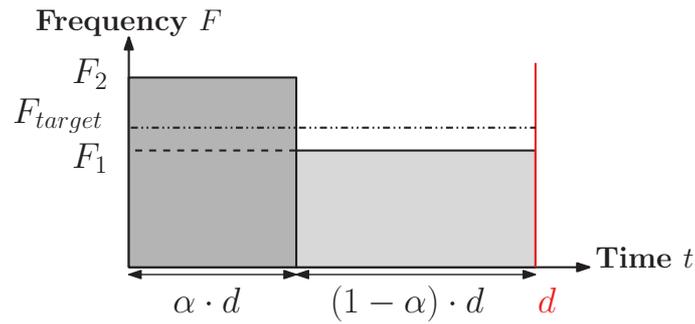


FIGURE 4.18: *Hopping* entre deux fréquences pour atteindre une performance équivalente à la fréquence F_{target} , avec α le rapport cyclique et d le temps imparti, connu pour la tâche considérée.

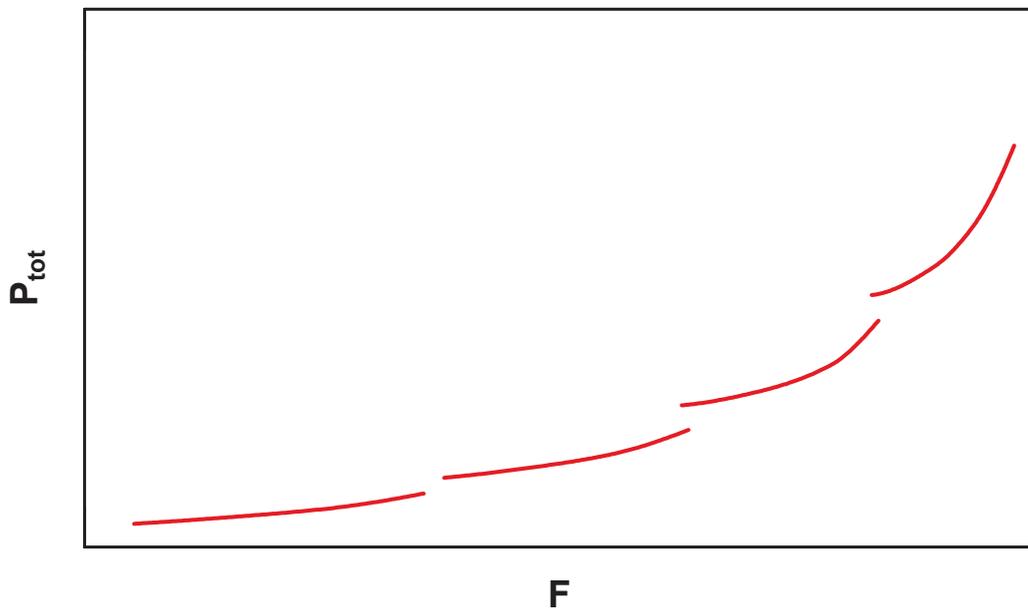


FIGURE 4.19: Puissance P_{tot} en fonction de la fréquence pour $4 V_{dd}$ et $V_{bb} \in [V_{bb_{min}} ; V_{bb_{max}}]$.

de pouvoir mener notre réflexion à partir de cette représentation.

La méthode proposée peut être divisée en deux phases [76] [78]. Dans la première phase, on détermine les solutions possibles et on identifie parmi ces solutions celles qui permettent d'optimiser la consommation du circuit tout en atteignant la performance requise (phase d'identification). Dans une deuxième phase, la configuration nécessaire pour atteindre la performance requise est choisie et appliquée (phase d'exécution de la tâche). Cependant pour pouvoir déterminer les PM permettant d'optimiser la consommation énergétique, il faut tout d'abord réaliser une phase d'identification qui consiste à déterminer la valeur de ces PM.

4.4.1 Phase d'identification

La phase d'identification est représentée sur la Figure 4.20. Elle peut être entièrement réalisée lors de la conception pendant la caractérisation post-fabrication ou en temps d'exécution, c'est-à-dire lors du fonctionnement du circuit. La solution retenue dépendra de l'implémentation visée. Cette phase d'identification se décompose en 5 étapes représentées sur la Figure 4.20.

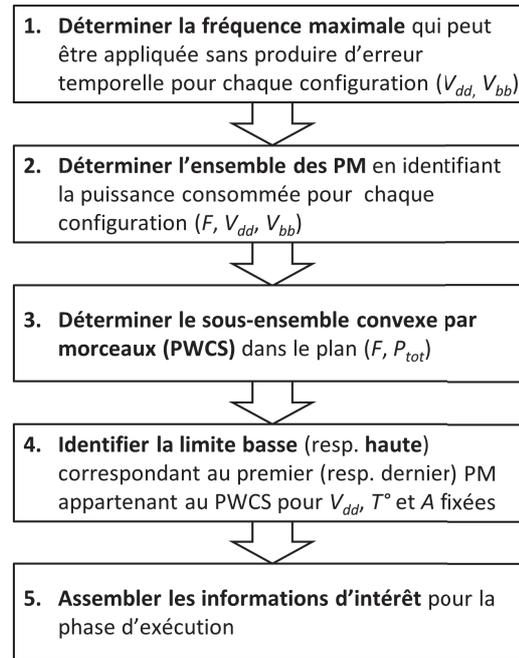


FIGURE 4.20: Phase d'identification du PWCS permettant de déterminer les modes de fonctionnement possibles et les modes optimaux (étapes 1 à 5).

1. Déterminer les fréquences maximales

Chaque valeur de fréquence F applicable au circuit est associée à une tension d'alimentation V_{dd} et à une tension de body bias V_{bb} . Si l'on considère qu'il existe n valeurs de tension d'alimentation V_{dd} , pour une température T° , F_{max_i} avec $i = 1 : n$, correspond à la fréquence maximale que l'on peut appliquer au système. Elle dépend de V_{dd_i} et de V_{bb} , avec la tension d'alimentation V_{dd_i} et la tension de body bias V_{bb} . F_{max_i} est donc fonction de ces deux entrées du système $F_{max_i}(V_{bb}, V_{dd_i})$.

On obtient F_{max_i} pour toutes les tensions d'alimentation possibles V_{dd_i} et pour tout V_{bb} appartenant à l'intervalle $[V_{bb_{min}} ; V_{bb_{max}}]$. Par exemple, avec $n = 4$, cette étape permettra de déterminer les relations représentées sur la Figure 4.14. La fréquence maximale pour un couple (V_{dd_i}, V_{bb}) peut être obtenue par l'intermédiaire de capteurs tels que les Timing Slack Monitors (TSM) [68].

2. Déterminer l'ensemble des PM

Une fois $F_{max_i}(V_{bb}, V_{dd_i})$ déterminée, on identifie la puissance consommée P_{tot} pour les variables V_{dd} , V_{bb} et F_{max_i} appliquées au circuit et pour les « perturbations » subies par le système, à savoir, la température T° et l'activité du circuit A . Rappelons que l'activité A est une représentation du nombre de portes logiques qui basculent dans un intervalle de temps défini et que les PM retenus ne produisent pas d'erreur temporelle.

Certaines fréquences F_{target} peuvent ne pas être atteignables en appliquant un des PM identifiés, c'est particulièrement le cas lorsqu'au moins une des variables est discrète.

Par exemple, avec $n = 4$, cette étape permettra de déterminer les courbes représentées sur la Figure 4.21 où chaque courbe correspond à l'ensemble des PM atteignables pour chaque valeur de tension d'alimentation V_{dd_i} disponible.

L'identification des PM peut s'effectuer avec des capteurs spécifiques en cours d'utilisation du circuit ou en phase de conception ou de test du circuit [38].

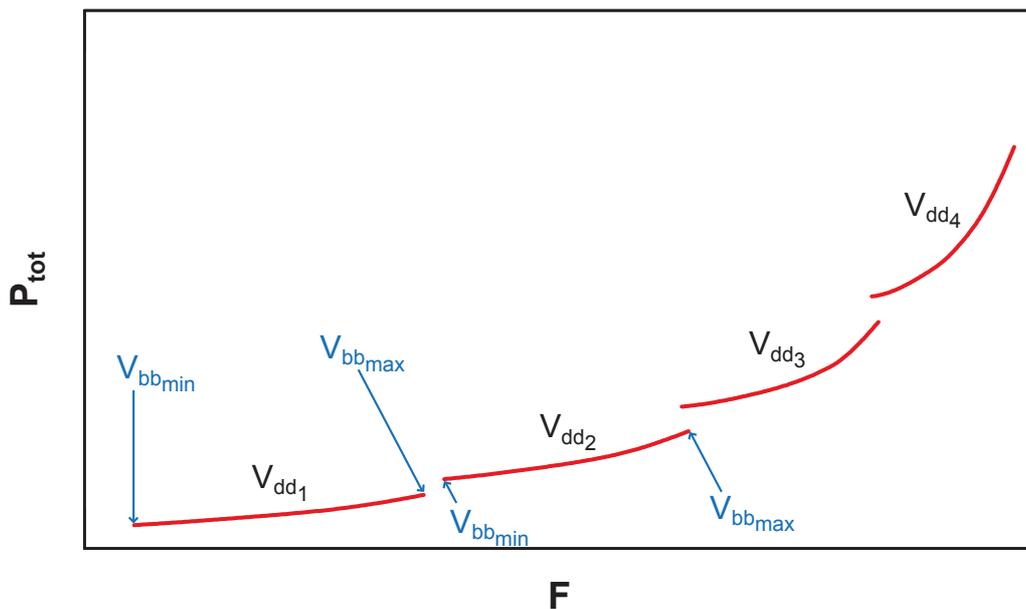


FIGURE 4.21: Puissance P_{tot} en fonction de la fréquence pour 4 V_{dd} différents et $V_{bb} \in [V_{bb_{min}} ; V_{bb_{max}}]$.

3. Déterminer le sous-ensemble croissant, convexe par morceaux ou PWCS

Dans cette étape, pour une température T° et une activité A données, on détermine le sous-ensemble croissant, convexe par morceaux issu des PM identifiés dans l'étape précédente.

Par exemple, pour $n = 4$, cette étape permettra de déterminer les morceaux des courbes $P_{tot}(F)$ représentés en vert sur la Figure 4.22. Ces morceaux appartiennent à un sous-ensemble croissant, convexe par morceaux comme défini dans [85].

La détermination du PWCS peut être réalisée, par exemple, à l'aide du parcours de Graham [86] ou de la marche de Jarvis [87]. Il s'agit d'algorithmes permettant de déterminer l'enveloppe convexe d'un ensemble fini de points. Pour cela, il faudra discrétiser la courbe avec un pas fin pour pouvoir déterminer le PWCS en ligne avec l'un des deux algorithmes. Nous donnerons le détail de la détermination du PWCS dans la section 4.4.2 lorsque la consommation P_{tot} peut être modélisée à l'aide d'une fonction quadratique en F . Rappelons que si $P_{tot}(F)$ contient uniquement des morceaux de courbes formant un ensemble globalement convexe et si une technique de *hopping* est applicable, exécuter une tâche avec les deux fréquences voisines les plus proches de la fréquence cible F_{target} lorsque cette dernière n'appartient pas au PWCS, permettra de minimiser la consommation du circuit (voir la Section 4.3.1 où la preuve a été faite pour un ensemble discret et peut être généralisée à un ensemble continu).

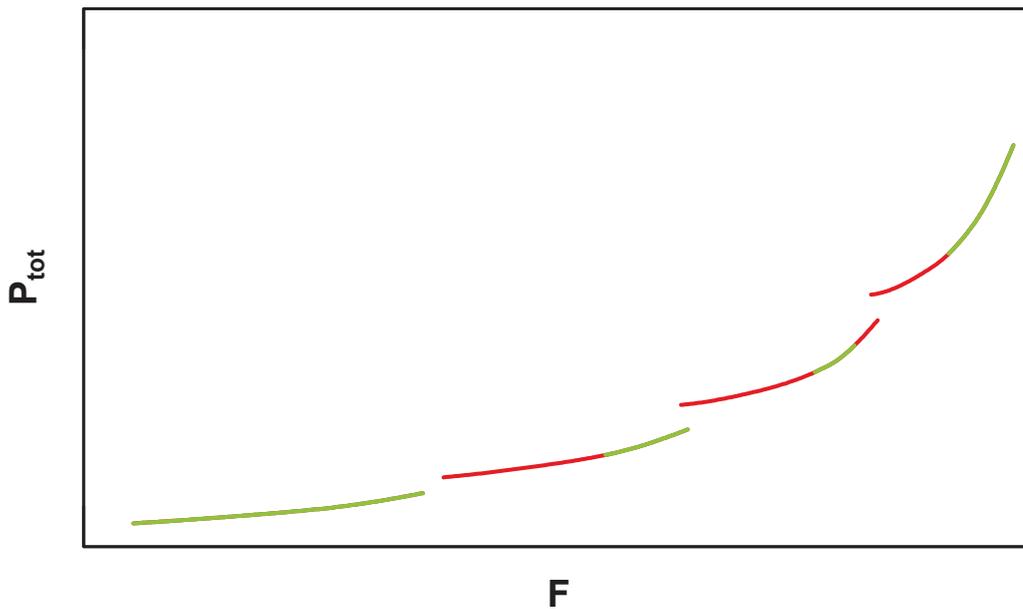


FIGURE 4.22: Identification des morceaux de courbe $P_{tot}(F)$ appartenant au PWCS.

4. Déterminer les limites basses et hautes

Pour une température T° et une activité A données, on cherche à déterminer les « limites basses » et « limites hautes » des morceaux de courbes $P_{tot}(F)$ formant le sous-ensemble croissant, convexe par morceaux (PWCS). Autre-

ment dit, on cherche à déterminer le premier et le dernier PM de chaque morceau de courbe appartenant au PWCS.

Pour T° , A et V_{dd_i} , on appelle « limite basse » le point dans le plan $P_{tot}(F)$ du PWCS qui a la plus petite valeur de tension de body bias V_{bb} . On note V_{bb_i-} la « limite basse » de V_{bb} associée à V_{dd_i} . Réciproquement, pour V_{dd_i} on appelle « limite haute » le point dans le plan $P_{tot}(F)$ du PWCS ayant la plus grande valeur de tension V_{bb} . On note V_{bb_i+} la « limite haute » de V_{bb} associée V_{dd_i} .

Pour $n = 4$, cette étape permet de déterminer les points limites du sous-ensemble convexe par morceaux de $P_{tot}(F)$ représentés en bleu sur la Figure 4.23.

L'identification des PM limites peut être effectuée lors de la conception, ou lors du test post-fabrication, ou bien encore, en cours d'exécution. Cette étape peut être réalisée conjointement avec l'étape précédente afin de mutualiser les moyens requis et d'optimiser l'implémentation dans le cas d'une détermination en cours d'exécution.

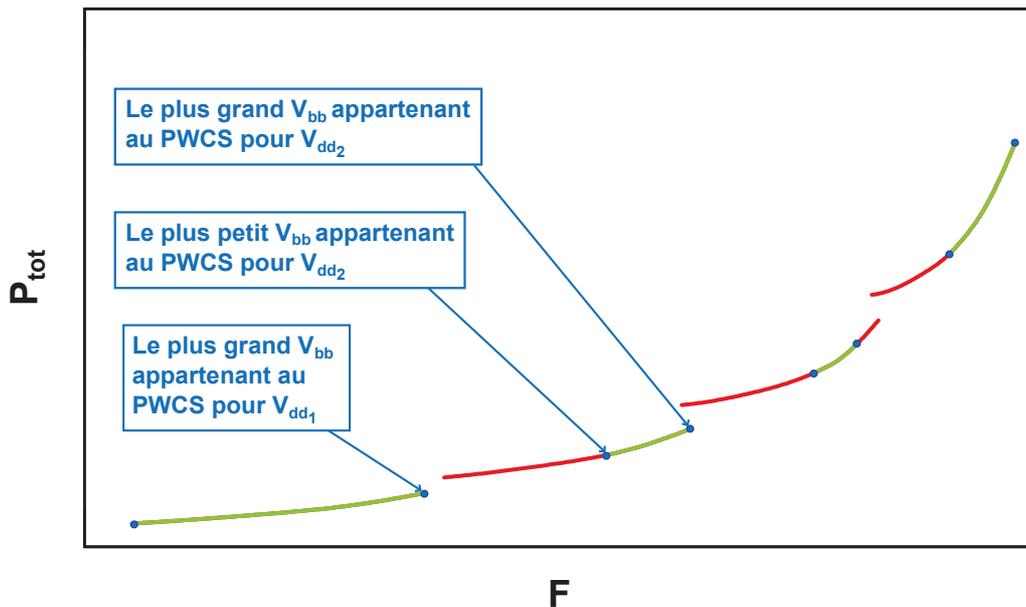


FIGURE 4.23: Détermination des limites basses et hautes pour T° et A données.

5. Assembler les informations d'intérêt pour la phase d'exécution

La sortie de cette phase est l'identification de tous les PM appartenant au sous-ensemble croissant, convexe par morceaux (PWCS) dans le plan (F, P_{tot}) en fonction des « perturbations » T° et A , et des entrées du système (tension d'alimentation V_{dd_i} , tension de body bias V_{bb} et fréquence F).

Une représentation possible pour l'ensemble des PM est celle du Tableau 4.1 où l'on voit les « limites basses » et « limites hautes ». Ainsi, pour une fréquence cible F_{target} , on peut vérifier si elle appartient à un intervalle de fréquence $[F_{i-}, F_{i+}]$ pour une tension d'alimentation unique V_{dd_i} . Puis si tel est le cas, on détermine la valeur de la tension de body à appliquer. Pour fixer les idées, prenons un exemple. Si la fréquence F_{target} est comprise entre F_{2-} et F_{2+} , il faut appliquer la tension d'alimentation V_{dd_2} pour atteindre cette performance et une tension de body bias V_{bb} comprise entre $V_{bb_{2-}}$ et $V_{bb_{2+}}$. La tension de body bias sera déterminée en fonction de la fréquence, de la température et de l'activité.

Notons que, la valeur de tension de body bias à appliquer V_{bb} peut être déterminée en fonction de la fréquence cible F_{target} , de la tension d'alimentation V_{dd} , de la température T° et de l'activité A pour les PM ne correspondant pas aux limites basses et hautes par l'intermédiaire d'un capteur TSM [68] ou d'une fonction à identifier.

TABLEAU 4.1: Limites pour des gammes de T° et A données.

$[T_1^\circ, T_2^\circ] [A_1, A_2]$		
V_{dd_i}	$V_{bb_{i-}}, F_{i-}$	$V_{bb_{i+}}, F_{i+}$
$V_{dd_{i+1}}$	$V_{bb_{(i+1)-}}, F_{(i+1)-}$	$V_{bb_{(i+1)+}}, F_{(i+1)+}$
...
$[T_1^\circ, T_2^\circ] [A_2, A_3]$		
V_{dd_i}	$V_{bb_{i-}}, F_{i-}$	$V_{bb_{i+}}, F_{i+}$
$V_{dd_{i+1}}$	$V_{bb_{(i+1)-}}, F_{(i+1)-}$	$V_{bb_{(i+1)+}}, F_{(i+1)+}$
...
$[T_2^\circ, T_3^\circ] [A_1, A_2]$		
V_{dd_i}	$V_{bb_{i-}}, F_{i-}$	$V_{bb_{i+}}, F_{i+}$
$V_{dd_{i+1}}$	$V_{bb_{(i+1)-}}, F_{(i+1)-}$	$V_{bb_{(i+1)+}}, F_{(i+1)+}$
...

Une implémentation typique consisterait à stocker l'ensemble des PM identifiés. Cependant, une implémentation optimisée pourra considérer le stockage des PM limites (basses et hautes) pour une température de référence T_{ref}° et une activité de référence A_{ref} . Ensuite, des fonctions de modification de ces PM pour les adapter en considérant des valeurs mesurées ou estimées de la température T° et l'activité A [70], [88] du circuit pourraient être appliquées. De même, la fonction pour calculer la tension de body bias V_{bb} peut être implémentée une seule fois et paramétrée en fonction de la température T° , de l'activité A et de la tension d'alimentation V_{dd} .

4.4.2 Identification du sous-ensemble croissant, convexe par morceaux

La phase d'identification revient à déterminer le sous-ensemble croissant, convexe par morceaux ou PWCS. Par souci de simplicité, seulement deux courbes $P_{tot}(F)$ seront considérées par la suite (voir les Figures 4.24 et 4.25), la méthode étant généralisable à n'importe quel nombre de courbes $P_{tot}(F)$.

Les Figures 4.24 et 4.25 montrent la consommation du circuit en fonction de la fréquence pour deux tensions d'alimentation différentes. L'identification du PWCS avec plus de 2 courbes sera réalisée en les considérant 2-à-2. **En première approximation, la fonction $P_{tot_i}(F)$ est modélisée à l'aide d'une fonction quadratique en F car P_{tot_i} est minimisée à travers la tension V_{bb} pour chaque F et pour la tension d'alimentation V_{dd_i} donnée** (Hypothèse 1). Cette modélisation va permettre d'identifier le PWCS. Un autre modèle plus complexe pourrait être utilisé mais par souci de simplicité la modélisation par une fonction quadratique permet d'avoir une solution analytique.

Nous allons présenter deux cas d'identification du PWCS, notés cas I et cas II dans la suite. Le cas I correspond à la situation où seulement une partie de chaque courbe appartient au PWCS. Dans le cas II, toute la courbe correspondant à la plus petite des deux tensions d'alimentation appartient au PWCS.

I. Cas I

L'objectif est de trouver la fréquence F_{i+} correspondant à la limite haute de $P_{tot_i}(F)$ pour la tension d'alimentation V_{dd_i} et la fréquence $F_{(i+1)-}$ correspondant à la limite basse $P_{tot_{i+1}}(F)$ pour la tension d'alimentation $V_{dd_{i+1}}$ de manière à ce que les parties restantes de $P_{tot_i}(F)$ et $P_{tot_{i+1}}(F)$ appartiennent au PWCS. En conséquence, les PM ayant une fréquence supérieure à F_{i+} pour la courbe correspondant à V_{dd_i} et les PM ayant une fréquence inférieure à $F_{(i+1)-}$ pour la courbe correspondant à $V_{dd_{i+1}}$ n'appartiennent pas au PWCS.

On suppose que :

- a) pour une tension d'alimentation V_{dd_i} , $P_{tot_i}(F)$ est une fonction croissante convexe telle que $P_{tot_i}(F) = a_i \cdot F^2 + b_i \cdot F + c_i$;
- b) pour chaque tension d'alimentation V_{dd_i} , la courbe $P_{tot_i}(F)$ a un sous-ensemble de points qui appartiennent au PWCS ;
- c) $P_{tot_{i+1}}(F_{(i+1)-}) > P_{tot_i}(F_{i+}) > 0$ et $F_{(i+1)-} > F_{i+} > 0$, où P_{tot} et F représentent respectivement la consommation totale du dispositif et la fréquence d'horloge.

On souhaite que la tangente à droite à la courbe $P_{tot_i}(F)$ au point F_{i+} soit égale à la tangente à gauche à la courbe $P_{tot_{i+1}}(F)$ au point $F_{(i+1)-}$.

À partir des modèles et des points précédents, avec F_{i+} , $F_{(i+1)-}$, $P_{tot_i}(F_{i+})$ et $P_{tot_{i+1}}(F_{(i+1)-})$ les paramètres à identifier, il vient :

$$\begin{cases} P_{tot_i}(F_{i+}) = a_1 \cdot F_{i+}^2 + b_1 \cdot F_{i+} + c_1 \\ P_{tot_{i+1}}(F_{(i+1)-}) = a_2 \cdot F_{(i+1)-}^2 + b_2 \cdot F_{(i+1)-} + c_2 \\ 2 \cdot a_1 \cdot F_{i+} + b_1 = 2 \cdot a_2 \cdot F_{(i+1)-} + b_2 \\ 2 \cdot a_1 \cdot F_{i+} + b_1 = \frac{P_{tot_{i+1}}(F_{(i+1)-}) - P_{tot_i}(F_{i+})}{F_{(i+1)-} - F_{i+}} \end{cases} \quad (4.11)$$

Après quelques manipulations mathématiques, on obtient :

$$F_{i+}^2 \cdot \frac{a_1 \cdot a_2 - a_1^2}{a_2} + F_{i+} \cdot \frac{a_1 \cdot b_2 - a_1 \cdot b_1}{a_2} - \frac{(b_1 - b_2)^2}{4 \cdot a_2} + c_2 - c_1 = 0 \quad (4.12)$$

Le discriminant de l'équation (4.12) est :

$$\Delta_1 = \frac{a_1 \cdot a_2 \cdot ((b_1 - b_2)^2 - 4 \cdot (c_2 - c_1) \cdot (a_2 - a_1))}{a_2^2} \quad (4.13)$$

Si Δ_1 est positif, alors il existe une solution $F_{i+} \in \mathbb{R}^+$. Ainsi les solutions de (4.12) sont :

$$F_{i+(1,2)} = \frac{a_1 \cdot b_1 - a_1 \cdot b_2 \pm a_2 \cdot \sqrt{\Delta_1}}{2 \cdot (a_1 \cdot a_2 - a_1^2)} \quad (4.14)$$

La valeur de F_{i+} correspond à la solution de $F_{i+(1,2)}$ positive, qui en outre est dans la gamme de fréquence associée à $P_{tot_i}(F)$. $F_{(i+1)-}$ est déduite de la troisième équation de (4.11) :

$$F_{(i+1)-} = \frac{2 \cdot a_1 \cdot F_{i+} + b_1 - b_2}{2 \cdot a_2} \quad (4.15)$$

$P_{tot_i}(F_{i+})$ et $P_{tot_{i+1}}(F_{(i+1)-})$ peuvent ensuite être calculées.

II. Cas II

Étant donné que toute la courbe $P_{tot_i}(F)$ correspondant à V_{dd_i} appartient au PWCS, l'objectif est de trouver la fréquence $F_{(i+1)-}$ correspondant à la limite basse $P_{tot_{i+1}}(F)$ pour la tension d'alimentation $V_{dd_{i+1}}$ de manière à assurer que $P_{tot_i}(F)$ et la partie restante $P_{tot_{i+1}}(F)$ sont dans le PWCS.

Considérons les points suivants :

- pour une tension d'alimentation V_{dd_i} , $P_{tot_i}(F)$ est une fonction croissante convexe avec $P_{tot_i}(F) = a_i \cdot F^2 + b_i \cdot F + c_i$;
- pour chaque tension d'alimentation V_{dd_i} , la courbe $P_{tot_i}(F)$ a un sous-ensemble qui appartient au PWCS ;

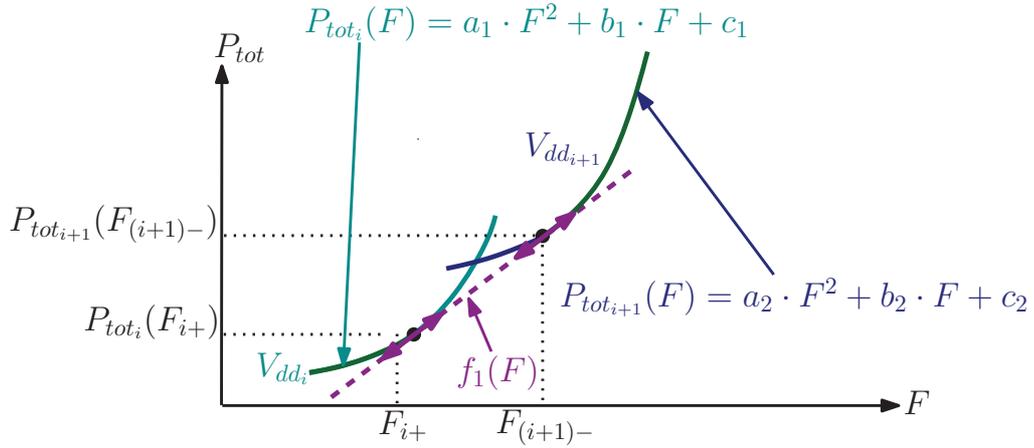


FIGURE 4.24: Modèles de consommation avec les PM appartenant au PWCS en vert pour le cas I.

- c) $P_{tot_{i+1}}(F_{(i+1)-}) > P_{tot_i}(F_{i+}) > 0$ et $F_{(i+1)-} > F_{i+} > 0$, avec P_{tot} et F correspondant respectivement à la consommation totale du dispositif et à la fréquence d'horloge.

On souhaite garantir que la tangente à gauche à la courbe $P_{tot_{i+1}}(F)$ au point $F_{(i+1)-}$ est égale à la pente de $f_1(F)$. f_1 correspond au segment reliant les deux tangentes.

À partir des modèles et des points ci-avant, où $F_{(i+1)-}$ et $P_{tot_{i+1}}(F_{(i+1)-})$ sont les paramètres à identifier, il en résulte que :

$$\begin{cases} P_{tot_i}(F_{i+}) = a_1 \cdot F_{i+}^2 + b_1 \cdot F_{i+} + c_1 \\ P_{tot_{i+1}}(F_{(i+1)-}) = a_2 \cdot F_{(i+1)-}^2 + b_2 \cdot F_{(i+1)-} + c_2 \\ 2 \cdot a_2 \cdot F_{(i+1)-} + b_2 = \frac{P_{tot_{i+1}}(F_{(i+1)-}) - P_{tot_i}(F_{i+})}{F_{(i+1)-} - F_{i+}} \end{cases} \quad (4.16)$$

À partir de quelques manipulations des équations de (4.16), on a :

$$a_2 \cdot F_{(i+1)-}^2 - 2 \cdot a_2 \cdot F_{i+} \cdot F_{(i+1)-} + c_1 - c_2 + F_{i+} \cdot (a_1 \cdot F_{i+} + b_1 - b_2) = 0 \quad (4.17)$$

Le discriminant de l'équation (4.17) est :

$$\Delta_2 = (-2 \cdot a_2 \cdot F_{i+})^2 - 4 \cdot a_2 \cdot (F_{i+} \cdot (a_1 \cdot F_{i+} + b_1 - b_2) + c_1 - c_2) \quad (4.18)$$

Si Δ_2 est positif, il existe une solution $F_{(i+1)-} \in \mathbb{R}^+$.

Les solutions de l'équation (4.17) sont :

$$F_{(i+1)-}^{(1,2)} = \frac{2 \cdot a_2 \cdot F_{i+} \pm \sqrt{\Delta_2}}{2 \cdot a_2} \quad (4.19)$$

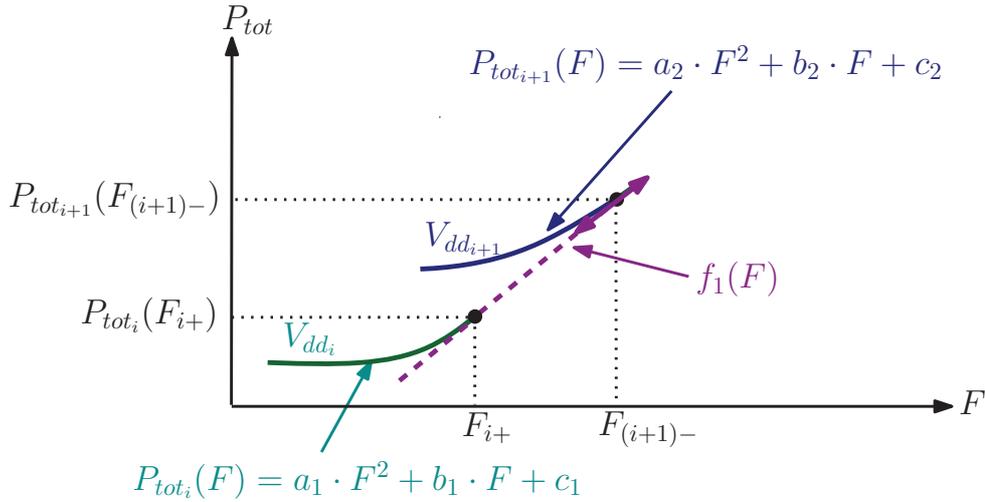


FIGURE 4.25: Modèles de consommation avec les PM appartenant au PWCS tracés en vert pour le cas II.

$F_{(i+1)-}$ correspond à la solution positive de $F_{(i+1)-}^{(1,2)}$. En outre, elle appartient à la gamme de fréquence associée à $P_{tot_{i+1}}(F)$.

Ensuite, l'inconnu $P_{tot_{i+1}}(F_{(i+1)-})$ peut être déduite de la première équation de (4.16).

Une fois l'identification du PWCS effectuée, la deuxième phase de la méthode proposée correspond à la phase d'exécution des tâches.

4.4.3 Exécution des tâches

La deuxième phase de la proposition consiste à appliquer un ou plusieurs PM afin d'atteindre la fréquence cible (F_{target}) requise pour exécuter une tâche avec une performance donnée. Si la fréquence F_{target} n'est pas atteignable, ou si la fréquence F_{target} ne correspond pas à la fréquence d'un PM appartenant au PWCS, 2 PM seront appliqués afin d'atteindre F_{target} en moyenne au lieu d'appliquer qu'un seul PM dont la fréquence est supérieure à F_{target} . Cette phase correspond aux étapes 1 à 5 de la méthode proposée, voir la Figure 4.26.

Le ou les PM retenus dépendront des valeurs de T° et A . Les PM utilisables ont été déterminés dans la phase d'identification. Ils appartiennent au PWCS dans le plan (F, P_{tot}) et permettent donc d'obtenir *in fine* une consommation minimale pour une performance à atteindre (i.e. une fréquence équivalente fixée F_{target}).

Un exemple de sous-ensemble de PM déterminés dans la phase d'identification est représenté sur la Figure 4.27 par les morceaux rouges de courbes pour un système avec $n = 4$ niveaux de V_{dd} . L'axe des abscisses de cette figure comporte

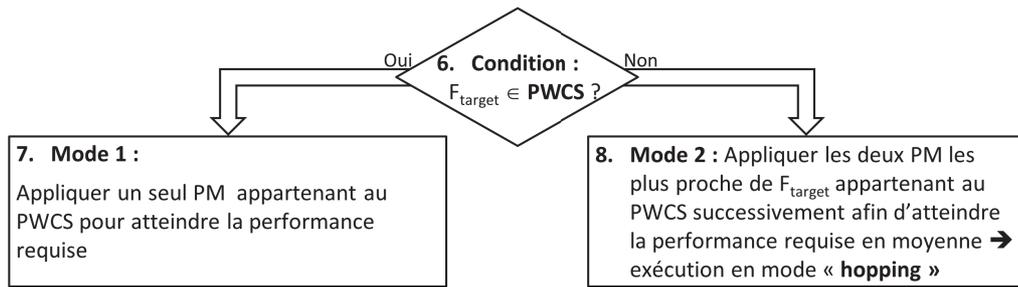
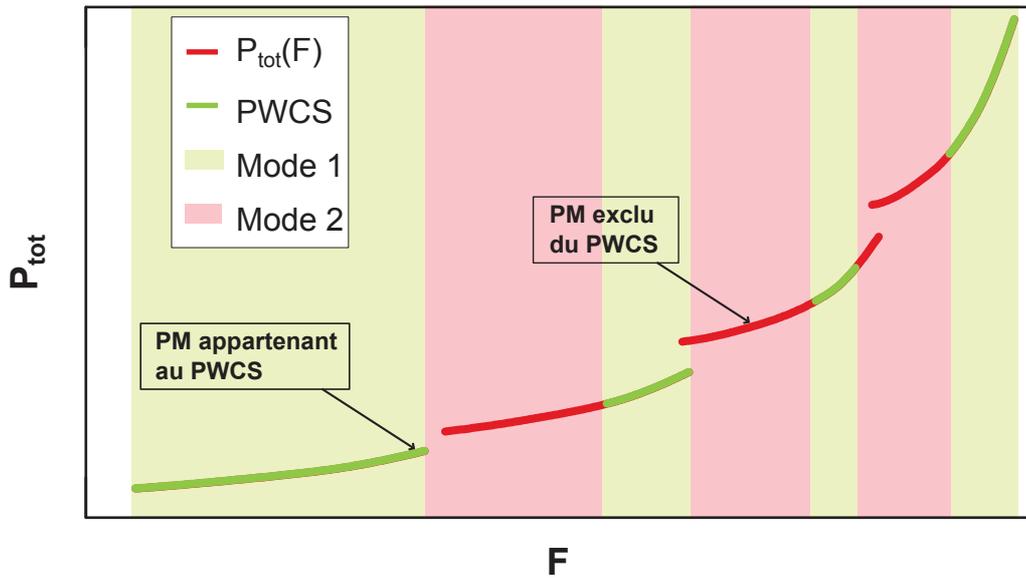


FIGURE 4.26: Phase de fonctionnement (exécution de la tâche) (étapes 6 à 8).

FIGURE 4.27: Ensemble de PM et modes d'application possibles. Mode 1 ou M1 : on applique un unique PM. Mode 2 ou M2 : on applique 2 PM en mode *hopping*.

des zones pour lesquelles les PM associés appartiennent au PWCS (identifiées en vert) et d'autres zones n'appartenant pas au PWCS (identifiées en rouge). Autrement dit, dans certaines plages de l'axe des abscisses (correspondant aux morceaux de courbes en vert sur la Figure 4.27), on peut appliquer les entrées du système V_{dd} et V_{bb} correspondant à la fréquence recherchée F_{target} . La consommation sera minimale car il n'existe pas d'autres moyens possibles d'atteindre la même F_{target} pour un coût énergétique inférieur. Dans ces zones, la fréquence F_{target} est appliquée directement.

En dehors de ces régions, *i.e.* dans les intervalles de l'axe des abscisses où les morceaux de courbes n'appartiennent pas au PWCS, l'application directe de F_{target} garantira la performance souhaitée mais ne garantira pas une consommation minimale. Un mode d'exécution en faisant du *hopping* entre 2 PM est alors possible. Il permet de garantir une consommation inférieure à celle obtenue avec

l'application directe de F_{target} , et une performance équivalente à celle offerte par l'application directe de la fréquence F_{target} .

On identifie donc deux modes possibles d'application des PM. Le Mode 1 ou M1 (voir la Figure 4.27) fait référence à une fréquence F_{target} souhaitée atteignable par un PM appartenant au PWCS. Dans ce cas, la fréquence F_{target} sera appliquée directement. Le Mode 2 ou M2 (voir la Figure 4.27) fait référence à une fréquence F_{target} non atteignable par un unique PM appartenant au PWCS. Le Mode 2 considère l'application successive de 2 PM appartenant au PWCS avec un ratio de temps passé dans chaque PM tel que la performance finale est équivalente en moyenne à la performance obtenue avec la fréquence F_{target} mais avec une consommation inférieure à celle que l'on aurait avec l'application directe de la fréquence F_{target} .

Après la phase d'identification (étapes 1 à 5, l'ensemble des PM a été déterminé et le PWCS a été identifié. La Figure 4.26 illustre la phase d'exécution, constituée de 3 étapes numérotées de 6 à 8. La sélection du mode de fonctionnement (M1, M2) à appliquer selon la fréquence équivalente F_{target} permettant d'atteindre la performance désirée y est résumée :

6. Condition

Les gammes de fréquences appartenant au PWCS sont délimitées par les limites basses et hautes déterminées dans l'étape 4 de la phase d'identification, en fonction de la température T° et de l'activité A du circuit. Afin de choisir le mode d'application du PM (Mode 1 ou Mode 2), on vérifie si la fréquence équivalente recherchée F_{target} est comprise entre la limite basse F_{i-} et la limite haute F_{i+} d'un niveau de tension d'alimentation V_{dd_i} (voir Figure 4.27).

Si cette condition s'avère vraie, on passe à l'étape 7 afin d'appliquer le Mode 1. Au contraire, si elle est fautive, on applique le Mode 2 de l'étape 8.

7. Mode 1 : Appliquer un PM

Le choix du Mode 1 est réalisé lorsque la fréquence F_{target} est atteignable par un seul PM appartenant au PWCS de $P_{tot}(F)$ (voir la Figure 4.28). Ce mode garantit la performance recherchée pour une consommation minimale. Pour l'appliquer, on se sert de la sortie de la phase d'identification (étape 5) ainsi que des conditions de fonctionnement du circuit (température T° et activité A si ces paramètres sont disponibles) afin d'identifier le couple (V_{dd} , V_{bb}) qui définit (avec F_{target}) le PM à appliquer.

8. Mode 2 : *Hopping* entre 2 PM

Le choix du Mode 2 est réalisé lorsque la fréquence F_{target} n'est pas atteignable par un seul PM appartenant au PWCS de $P_{tot}(F)$.

Ce mode d'application considère l'application des 2 PM les plus proches de F_{target} appartenant au PWCS.

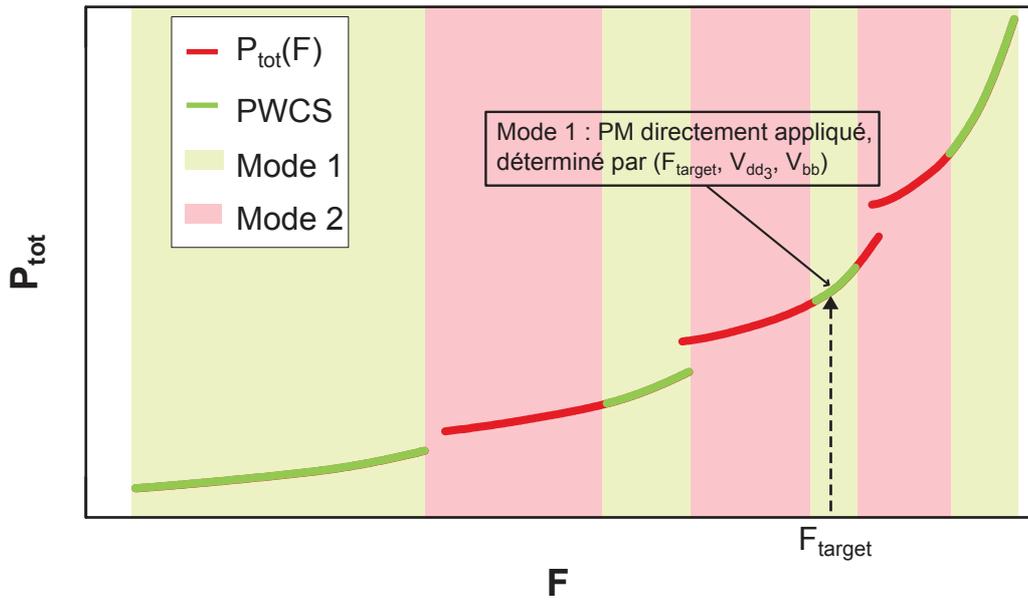


FIGURE 4.28: Exemple d'application du Mode 1.

Les PM identifiés seront appliqués successivement avec un ratio de temps passé dans chaque PM tel que la fréquence moyenne permette d'atteindre une performance équivalente à celle de F_{target} (principe du *hopping* de PM) [89]. Le ratio de temps passé à la fréquence $F_{(i+1)-}$ est obtenu par $\alpha = \frac{F_{target} - F_{i+}}{F_{(i+1)-} - F_{i+}}$, avec $i = 1 : n - 1$.

En pratique, les PM choisis correspondent aux points limites du PWCS qui encadrent la fréquence équivalente recherchée F_{target} . Dans l'exemple de la Figure 4.29, la fréquence F_{target} se situe entre les PM correspondant à $(F_{2+}, V_{dd2}, V_{bb2+})$ et $(F_{3-}, V_{dd3}, V_{bb3-})$. Ce sont ces deux PM qui devront être appliqués avec une technique de *hopping* afin d'optimiser la consommation. Ceci est représenté sur la Figure 4.29 par la ligne rouge en pointillée. La puissance consommée est alors $P_{tot}(F_{target}) = \alpha \cdot P_{tot}(F_{3-}) + (1 - \alpha) \cdot P_{tot}(F_{2+})$, avec $F_{target} = \alpha \cdot F_{3-} + (1 - \alpha) \cdot F_{2+}$.

Rappelons qu'après avoir déterminé le PWCS, les limites sont stockées afin de faciliter la phase d'exécution. De plus, dans le cas où aucune mesure ou estimation de la consommation n'est disponible, la connaissance des gammes de température et d'activité peut être utilisée pour choisir le mode d'exécution (Mode 1 ou Mode 2). Les données devant être stockées sont alors disponibles dans le Tableau 4.2. Ce dernier contient les valeurs des limites (basse et haute) et les fréquences associées aux limites pour chaque tension d'alimentation V_{dd_i} disponible, pour des gammes de température et d'activité données.

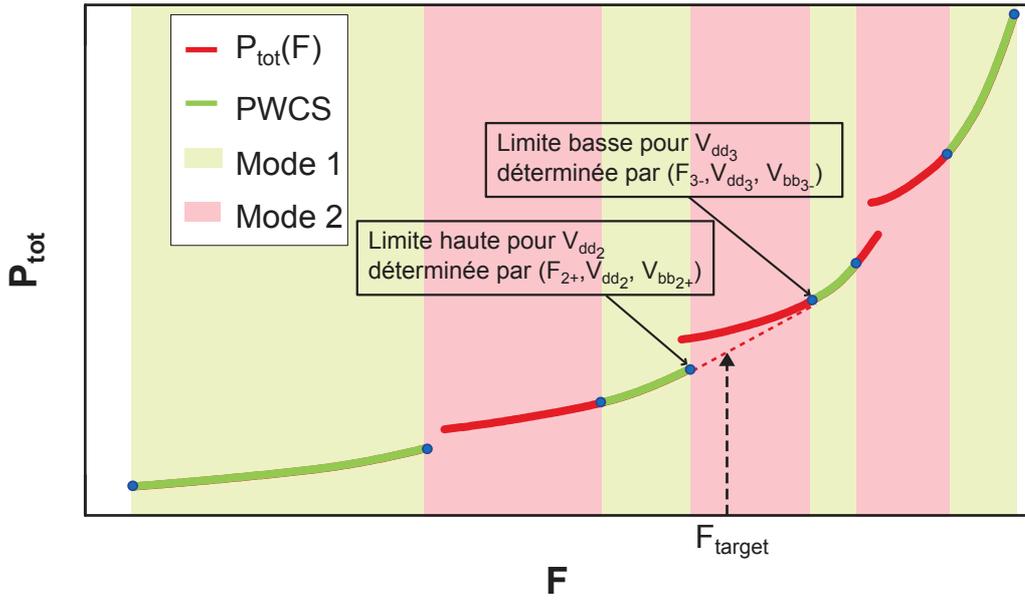


FIGURE 4.29: Exemple d'application du Mode 2.

Implémentation proposée

Pour une fréquence F_{target} donnée (Hypothèse 4), une température T° et une activité A mesurées ou estimées, 3 scénarios peuvent être rencontrés :

- F_{target} correspond à la fréquence d'un PM dans le PWCS. Dans ce cas, la tâche T est exécutée dans le Mode 1 avec le couple (V_{dd}, V_{bb}) approprié ;
- F_{target} est comprise entre deux fréquences dans le Tableau 4.2 et une unique tension d'alimentation V_{dd} est associée à ces deux fréquences. Dans ce cas, la tâche T sera exécutée en Mode 1. V_{dd} et F_{target} sont appliquées tandis que la tension de body bias V_{bb} devant être appliquée pour atteindre la fréquence maximale sans produire d'erreur temporelle peut être trouvée grâce à un capteur TSM [68] ;
- F_{target} se situe entre deux fréquences dans le Tableau 4.2, chaque fréquence est associée à une tension V_{dd} différente. Dans ce cas, la tâche T sera exécutée en Mode 2. Les valeurs de F , V_{dd} et V_{bb} associées aux deux limites seront appliquées successivement pour atteindre en moyenne F_{target} .

TABLEAU 4.2: Limites pour des gammes de température T° et activité A données.

$[T_1^\circ, T_2^\circ] [A_1, A_2]$		
V_{dd_i}	$V_{bb_{i-}}, F_{i-}$	$V_{bb_{i+}}, F_{i+}$
$V_{dd_{i+1}}$	$V_{bb_{(i+1)-}}, F_{(i+1)-}$	$V_{bb_{(i+1)+}}, F_{(i+1)+}$
...

Le nombre total de valeurs à stocker Sn peut être calculé avec :

$$Sn = (Vn \cdot VSn + TASN) \cdot Tn \cdot An \quad (4.20)$$

où Vn est le nombre de tensions d'alimentation, VSn est le nombre de valeurs à stocker par tension d'alimentation, $TASN$ est le nombre de valeurs de température et d'activité à stocker pour des gammes de température et d'activité données, Tn est le nombre de gammes de température et An est le nombre de gammes d'activité.

Par exemple, dans le cas où il y a 4 niveaux de tension d'alimentation V_{dd} ($Vn = 4$), 3 gammes de température T° ($Tn = 3$) et 2 gammes d'activité A ($An = 2$), 144 valeurs doivent être stockées afin de faciliter la phase d'exécution. En effet, il faut stocker 5 valeurs par tension d'alimentation V_{dd_i} qui sont V_{dd_i} , $V_{bb_{i-}}$, F_{i-} , $V_{bb_{i+}}$ et F_{i+} soit $VSn = 5$. Par ailleurs, il faut stocker T_i° , T_{i+1}° , A_i et A_{i+1} , $TASN$ est donc égal à 4. On obtient ainsi :

$$Sn = (4 \cdot 5 + 4) \cdot 3 \cdot 2 = 144 \quad (4.21)$$

Par ailleurs l'application du Mode 2 (*hopping* entre deux PM) peut être délicate en raison des limitations techniques et/ou technologiques de la mise en œuvre. Par exemple, l'implémentation peut ne pas permettre un saut « rapide » entre deux valeurs de V_{bb} . Dans ce cas précis, une solution sous-optimale peut être envisagée en appliquant une technique de *hopping* entre 2 PM définis pour une même tension V_{bb} mais avec des fréquences F et tensions d'alimentation V_{dd} différentes. La solution optimale est d'appliquer les PM les plus proches de la fréquence cible F_{target} appartenant au PWCS. Cependant ces deux PM ne possèdent pas en général la même tension V_{bb} .

4.4.4 Politique de gestion du hopping

L'exécution d'une tâche en Mode 1 est simple à réaliser et correspond à l'efficacité énergétique optimale. Par contre, si le Mode 2 doit être appliqué pour atteindre F_{target} en moyenne, la politique à associer à ce mode n'est pas triviale. En effet, une exécution en Mode 2 implique d'exécuter la tâche avec la limite haute $V_{bb_{i+}}$ associée à V_{dd_i} et la limite basse $V_{bb_{(i+1)-}}$ associée à $V_{dd_{i+1}}$ pour $F_{target} \in]F_{i+}, F_{(i+1)-}[$. Ainsi, deux triplets doivent être appliquées au dispositif, à savoir $(F_{i+}, V_{dd_i}, V_{bb_{i+}})$ et $(F_{(i+1)-}, V_{dd_{i+1}}, V_{bb_{(i+1)-}})$. Par conséquent, des mécanismes dit de *hopping* de $V_{dd_{i+1}}$ à V_{dd_i} , de $F_{(i+1)-}$ à F_{i+} et de $V_{bb_{(i+1)-}}$ à $V_{bb_{i+}}$ devraient être appliqués.

Cependant, une solution de *hopping* en V_{bb} rapide (quelques ns) est aujourd'hui difficilement réalisable si les contraintes d'implémentation physique des actionneurs sont considérées. Dans ce cas, on peut chercher à appliquer 2 PM avec

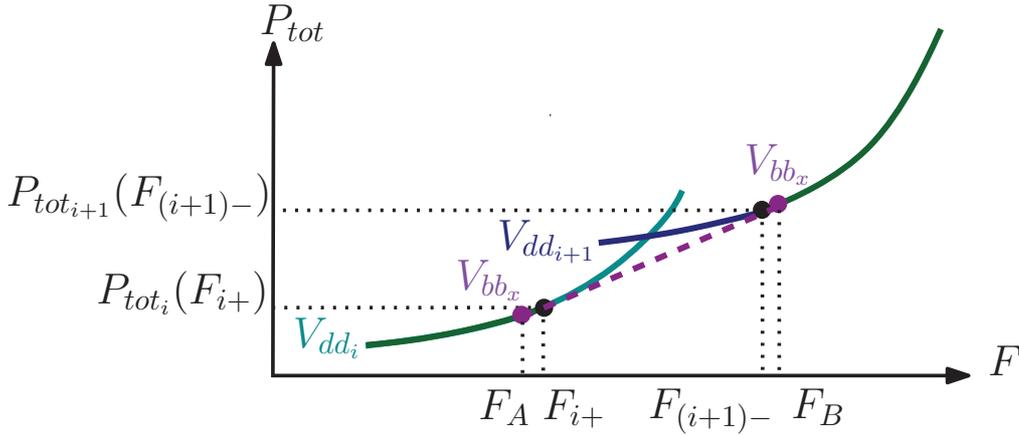


FIGURE 4.30: Mode 2 : Exécution en mode *hopping* avec un *hopping* en F et en V_{dd} , avec une tension de body bias constante V_{bb_x} .

une valeur de tension de body bias constante (e.g. V_{bb_x}) et un mécanisme de *hopping* pour V_{dd} de $V_{dd_{i+1}}$ à V_{dd_i} et pour F de F_B à F_A peut être appliqué afin d'atteindre en moyenne F_{target} . Cette approche est représentée sur la Figure 4.30. F_B est la fréquence appliquée avec le couple $(V_{dd_{i+1}}, V_{bb_x})$ et F_A est la fréquence associée au couple (V_{dd_i}, V_{bb_x}) .

Les fréquences appliquées F_B (pendant la période de temps $\alpha \cdot d$) et F_A et la puissance $P_{tot}(F_A, F_B, V_{bb_x})$ consommée pendant l'exécution en mode *hopping* avec une unique valeur de tension de body V_{bb_x} , satisfont les relations :

$$\begin{cases} F_{target} = \alpha \cdot F_B + (1 - \alpha) \cdot F_A \\ P_{tot}(F_A, F_B, V_{bb_x}) = \alpha \cdot P_{tot_{i+1}}(F_B, V_{bb_x}) + (1 - \alpha) \cdot P_{tot_i}(F_A, V_{bb_x}) \end{cases} \quad (4.22)$$

La consommation associée à F_{target} en appliquant un seul PM qui n'appartient pas au PWCS peut être comparée à la consommation donnée par le système (4.22). En fait, le gain en consommation ($G_{so}(F_{target})$) est donné par :

$$G_{so}(F_{target}) = 100 \cdot \frac{P_{tot_{th}}(F_{target}) - P_{tot}(F_A, F_B, V_{bb_x})}{P_{tot_{th}}(F_{target})} \quad (4.23)$$

avec $G_{so}(F_{target})$ le gain en consommation pour la fréquence cible F_{target} en pourcentage, $P_{tot_{th}}(F_{target})$ la consommation sans la méthode pour F_{target} atteignable en Mode 2, $P_{tot}(F_A, F_B, V_{bb_x})$ la consommation avec la solution sous-optimale pour F_{target} avec V_{bb_x} constant.

La tension de body bias V_{bb_x} sera choisie de manière à maximiser le gain G_{so} pour une F_{target} donnée en exécutant la tâche en Mode 2.

Cette solution sous-optimale permet d'éviter un *hopping* en V_{bb} . En effet, pour agir sur la tension V_{bb} , on utilise le principe de modulation dynamique de la tension de polarisation des caissons des transistors (Adaptive Body Biasing (ABB))

en anglais). Ces caissons étant modélisés par un condensateur, leur décharge n'est pas aisée à réaliser. En plus, cela revient à perdre de l'énergie. Or, notre méthode a été développée dans le but de minimiser la consommation.

4.5 Conclusion

L'objectif du travail est de minimiser la consommation d'un circuit sous contraintes de performance *via* une gestion coordonnée des trois entrées du système (F, V_{dd}, V_{bb}) . Une même performance peut être atteinte par l'application de différents couples (V_{dd}, V_{bb}) mais avec des consommations différentes. Il est donc nécessaire de développer une méthodologie de choix des tensions V_{dd} et V_{bb} permettant d'atteindre F_{target} pour garantir la performance tout en minimisant la consommation.

Dans ce chapitre, nous avons proposé une méthode permettant de minimiser la consommation pour une performance donnée quand 3 DoF discrets sont considérés. Nous avons montré qu'une phase d'identification du sous-ensemble discrètement convexe croissant ou DCS est indispensable. L'exécution de la tâche T lors de la phase de fonctionnement, se fera en appliquant un (des) PM appartenant au DCS.

Nous avons proposé une méthode permettant de minimiser la consommation en cours d'exécution d'un circuit constitué de 3 DoF dont un discret. Nous avons montré qu'une phase d'identification du sous-ensemble croissant, convexe par morceaux (PWCS) est nécessaire. La phase de fonctionnement est acheminée par l'application de PM appartenant au PWCS pour exécuter la tâche T avec la performance visée afin de garantir une consommation minimale. Lorsqu'une tâche T est exécutée en utilisant le Mode 2, une solution sous-optimale a été suggérée si l'implémentation de l'actionneur de tension de body bias V_{bb} ne permet pas de réaliser un saut « rapide » (quelques *ns*) entre deux valeurs de V_{bb} en prenant en compte les contraintes d'implémentation physique des actionneurs.

Dans le chapitre 5, le gain en consommation que la méthode proposée peut apporter sera évalué sur un circuit implémenté en technologie FD-SOI 28nm, en considérant un DoF discret et deux autres continus.

Chapitre 5

Application et résultats

Nous avons vu dans le chapitre précédent deux méthodes permettant de réduire la consommation d'un circuit pour une fréquence donnée. Les deux méthodes sont constituées de deux phases, la phase d'identification et la phase de fonctionnement.

La première méthode est valable pour un ensemble discret de PM obtenus *via* 3 degrés de liberté ou DoF discrets. La phase d'identification consiste à extraire un sous-ensemble discrètement convexe croissant ou DCS (Discretely Convex Subset) à partir d'un ensemble discret qui n'est pas globalement convexe. La seconde phase est la phase de fonctionnement consistant à exécuter les tâches avec les PM appartenant au DCS.

Le seconde méthode proposée est valide pour un ensemble continu par morceaux de PM obtenus par l'intermédiaire de 3 DoF dont un DoF discret. La phase d'identification permet de rendre globalement convexe un ensemble en identifiant les PM appartenant au sous-ensemble continu par morceaux croissant ou PWCS. La phase de fonctionnement a lieu en cours d'exécution, des PM appartenant au PWCS seront appliqués pour exécuter les tâches.

Ce chapitre va permettre d'évaluer le gain en consommation lorsque les méthodes proposées dans le chapitre 4 sont appliquées à un circuit possédant trois degrés de liberté pour gérer le compromis consommation-performance. Les trois DoF considérés dans ce chapitre sont la fréquence d'horloge F , la tension d'alimentation V_{dd} et la tension de body bias V_{bb} .

- Dans un premier temps, nous allons étudier le cas de 3 DoF discrets (nombre fini de PM) ;
- Ensuite sera exposé le cas où 2 DoF continus sont combinés à 1 DoF discret (nombre infini de PM) ;
- Puis, nous allons proposer une solution pour prendre en compte le coût de commutation ;

- L’impact des variations de procédés de fabrication sur la méthode proposée sera présenté lorsqu’un nombre infini de PM est disponible sur le circuit ;
- Le chapitre se finira sur une conclusion.

5.1 Application à un circuit possédant un nombre fini de PM

Nous allons tester la méthode présentée dans le chapitre 4 lorsqu’un nombre fini de PM est disponible sur le circuit. Les PM sont obtenus en ajustant (F, V_{dd}, V_{bb}) . Cette méthode permet de rendre globalement convexe un ensemble discret de PM. Rappelons, que la méthode est constituée de deux phases. La phase d’identification consiste à déterminer les PM appartenant au sous-ensemble discrètement convexe croissant ou DCS. La seconde phase de la méthode est la phase de fonctionnement qui consiste à appliquer les PM en cours d’exécution afin de minimiser la consommation pour une fréquence donnée.

Deux cas d’études vont être exposés. Le premier utilise le profil de consommation en fonction de la fréquence issue de simulations d’un oscillateur en anneau en technologie FD-SOI. Le second utilise le profil $P_{tot}(F)$ extrait de mesures effectuées sur un circuit réel en technologie FD-SOI.

5.1.1 Cas d’étude : simulation d’un oscillateur en anneau

Notre premier cas d’étude est un oscillateur en anneau ou RO (Ring-Oscillator) en technologie STMicroelectronics FD-SOI 28nm. Un oscillateur en anneau est constitué d’un nombre impair d’inverseurs et sa fréquence d’oscillation dépend entre autres de la tension d’alimentation et de la tension de body bias appliquées. Le RO considéré est constitué de 101 inverseurs. La polarisation du substrat est symétrique : la tension V_{bb} est appliquée au substrat des transistors NMOS et la tension $-V_{bb}$ est appliquée au substrat des transistors PMOS. Pour qu’un RO commence à osciller, il a besoin d’une perturbation en entrée. C’est pourquoi, dans les conditions initiales de simulations du RO, un inverseur instable est nécessaire. Cet inverseur va permettre de lancer les oscillations.

Les valeurs pour ce RO sont extraites de simulations électriques de type ELDO. Les entrées de la simulation sont la tension d’alimentation V_{dd} , la tension de body bias V_{bb} , la température T° et le temps de simulation. La simulation permet d’obtenir la consommation totale P_{tot} , la consommation statique P_{stat} et la fréquence d’oscillation F du RO. Le profil de la consommation en fonction de la fréquence d’un RO est représentatif, au premier ordre, de l’allure du profil de la consommation en fonction de la fréquence d’un circuit réel, implémenté dans la même technologie et fonctionnant à sa fréquence maximale. Notons que les valeurs absolues ne sont pas significatives, c’est pourquoi nous avons normalisé

les valeurs de consommation et de fréquence.

Nous allons comparer la consommation en utilisant la méthode proposée dans le cas d'un nombre fini de PM disponibles, à la consommation sans la méthode. Nous allons évaluer le gain en consommation que la méthode peut apporter en considérant que le profil de la consommation en fonction de la fréquence correspond à l'allure de consommation d'un circuit. Le gain en consommation est évalué avec un outil d'analyse numérique, MATLAB.

Nous ne considérons que 2 niveaux de tension de body bias V_{bb} et 3 niveaux de tension d'alimentation V_{dd} disponibles sur le dispositif, soit 6 PM. Pour une tension d'alimentation V_{dd} , la tension de body bias V_{bb} est augmentée, ce qui a pour effet d'augmenter la fréquence maximale F_{max} atteignable pour la tension d'alimentation appliquée. La fréquence est donc boostée. Nous allons considérer les valeurs suivantes pour les tensions $V_{dd_1} = 0,6$ Volt (V), $V_{dd_2} = 0,8$ V, $V_{dd_3} = 1,0$ V, $V_{bb_L} = 0$ V (niveau non-boosté) et $V_{bb_H} = 1,2$ V (niveau boosté). La tension de body bias maximale applicable avec la technologie FD-SOI est de 3 V, mais en réalité la tension de body maximale appliquée est de 1,5 V afin d'éviter un vieillissement prématuré du circuit. La température de simulation est de 100°C . Les simulations permettent d'obtenir la consommation statique P_{stat} et la consommation dynamique P_{dyn} , avec $P_{dyn} = P_{tot} - P_{stat}$. L'activité A d'un RO est constante et égale à 100%. Afin d'extrapoler le comportement d'un circuit complet en incluant l'activité A , on choisit $A = 30\%$ tel que la consommation totale P_{tot} soit réaliste :

$$P_{tot} = A \cdot P_{dyn} + P_{stat} \quad (5.1)$$

Cela revient à multiplier la consommation dynamique par l'activité A considérée pour la tension d'alimentation V_{dd} et la tension de body bias V_{bb} appliquée. La consommation statique ne dépend pas de l'activité mais elle dépend de V_{dd} et V_{bb} .

La Figure 5.1 représente la fréquence d'oscillation F en fonction de la tension d'alimentation V_{dd} pour V_{bb_L} le niveau non-boosté et V_{bb_H} le niveau boosté. Néanmoins, la méthode proposée étant développée pour minimiser la consommation P_{tot} , la représentation dans le plan (F, P_{tot}) sur la Figure 5.2 est préférée afin d'identifier les PM appartenant au sous-ensemble discrètement convexe croissant ou DCS tel que défini dans [84].

La Figure 5.2 illustre la courbe de la consommation P_{tot} en fonction de la fréquence F formée par les 6 PM précédemment cités. Les valeurs normalisées de la fréquence et de la consommation des PM sont données dans le Tableau 5.1 (voir la Figure 5.1 aussi). Nous pouvons observer sur la Figure 5.2, que l'ensemble des PM ne forme pas un ensemble globalement convexe et plus précisément un ensemble discrètement convexe croissant dans le plan (F, P_{tot}) . Ainsi un sous-ensemble de PM est identifié afin de former un DCS dans le plan (F, P_{tot}) . Comme

vu dans le chapitre 4, le sous-ensemble permettra de minimiser la consommation du circuit quelque soit la demande de performance.

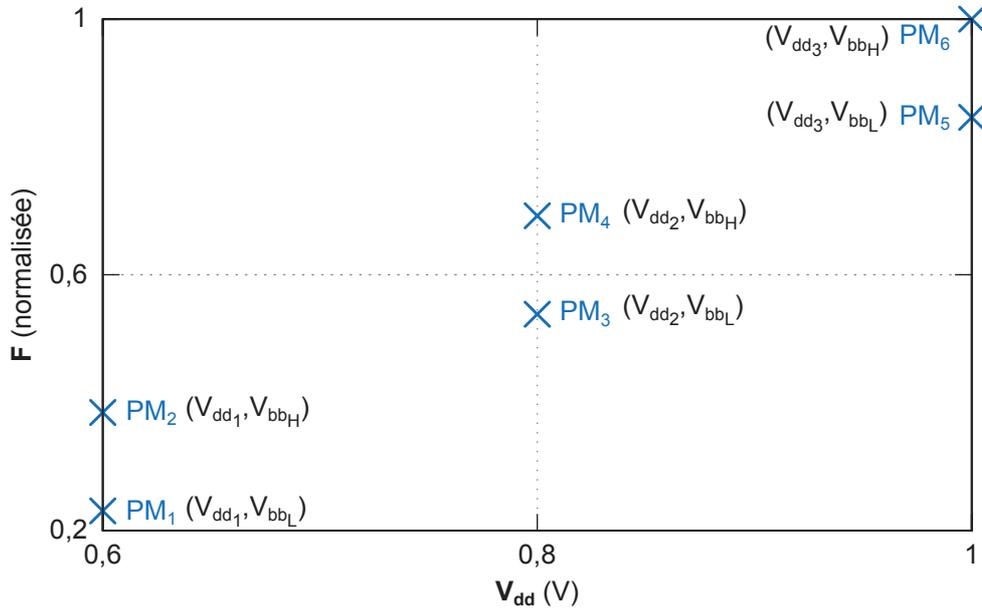


FIGURE 5.1: Ensemble de fréquences disponibles F en fonction de la tension d'alimentation V_{dd} : $F(V_{dd})$.

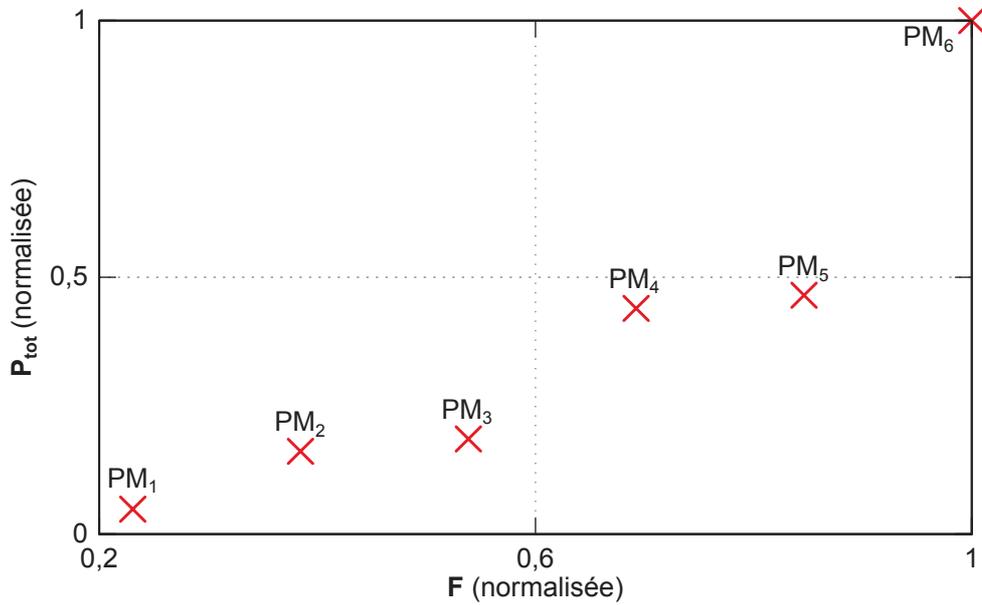


FIGURE 5.2: Ensemble de PM dans le plan (F, P_{tot}) avec P_{tot} la consommation totale du circuit et F la fréquence d'horloge.

TABLEAU 5.1: Valeurs de la tension d'alimentation V_{dd} , la tension de body bias V_{bb} , la fréquence normalisée F , la consommation normalisée P_{tot} pour chaque PM.

PM	V_{dd} (V)	V_{bb} (V)	F (normalisée)	P_{tot} (normalisée)
PM_1	0,6	0	0,2308	0,0484
PM_2	0,6	1,2	0,3846	0,1612
PM_3	0,8	0	0,5385	0,1852
PM_4	0,8	1,2	0,6923	0,4397
PM_5	1,0	0	0,8462	0,4651
PM_6	1,0	1,2	1,0000	1,0000

Identification du DCS

La première phase de la méthode est la phase d'identification du sous-ensemble discrètement convexe croissant DCS. Pour cela, nous allons déterminer les valeurs des pentes reliant les PM sur le profil de la consommation en fonction de la fréquence, voir la Figure 5.3. Rappelons que le PM_n est défini par le couple $(F_n, P_{tot}(F_n))$. Pour chaque PM_n , avec $1 < n < N$ et N le nombre de PM, nous allons comparer les pentes a_{n-1}^n , a_{n-1}^{n+1} et a_n^{n+1} reliant respectivement les PM_{n-1} et PM_n , les PM_{n-1} et PM_{n+1} et les PM_n et PM_{n+1} . Le Tableau 5.2 illustre la valeur des 3 pentes a_{n-1}^n , a_{n-1}^{n+1} et a_n^{n+1} pour les différents PM.

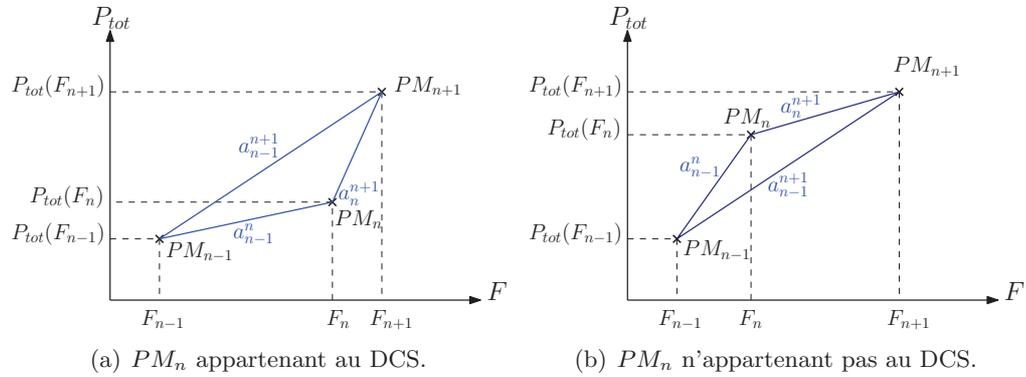


FIGURE 5.3: $P_{tot}(F)$ avec PM_n appartenant ou non au DCS croissant.

Le PM dont la consommation est la plus basse appartient au DCS. Les PM dont la fréquence est inférieure à la fréquence de ce PM n'appartiennent pas au DCS. Le PM_1 appartient donc au DCS.

Le PM dont la fréquence est la plus haute appartient au DCS. Les PM dont la consommation est supérieure à la consommation de ce PM n'appartiennent pas au DCS. Le PM_N , c'est-à-dire le PM_6 , appartient au DCS.

Pour les autres PM, si les pentes $a_{n-1}^n \leq a_{n-1}^{n+1} \leq a_n^{n+1}$, alors le PM_n appartient

au DCS. Les PM_3 et PM_5 appartiennent au DCS, voir le Tableau 5.2.

TABLEAU 5.2: Valeur des pentes afin d'identifier les PM appartenant au DCS, avec $1 < n < N$.

n	a_{n-1}^n	a_{n-1}^{n+1}	a_n^{n+1}	$a_{n-1}^n \leq a_{n-1}^{n+1} \leq a_n^{n+1}$?
2	0,7334	0,4446	0,1559	non
3	0,1559	0,9051	1,6547	oui
4	1,6547	0,9097	0,1650	non
5	0,1650	1,8209	3,4779	oui

À la fin de cette étape le DCS a été identifié. Les PM appartenant au DCS sont illustrés en vert sur la Figure 5.4 (soit les PM_1 , PM_3 , PM_5 , PM_6). Ces PM appartiennent au DCS. Dès lors, ce sont ces PM qui seront utilisés afin de minimiser la consommation du circuit. Pour satisfaire les contraintes de performance tout en minimisant la consommation, les tâches seront donc exécutées en appliquant seulement les PM identifiés par la méthode, présentée précédemment, possiblement en exécutant les tâches en mode *hopping*. Rappelons que l'on suppose que le profil de consommation en fonction de la fréquence correspond à celui d'un circuit réel.

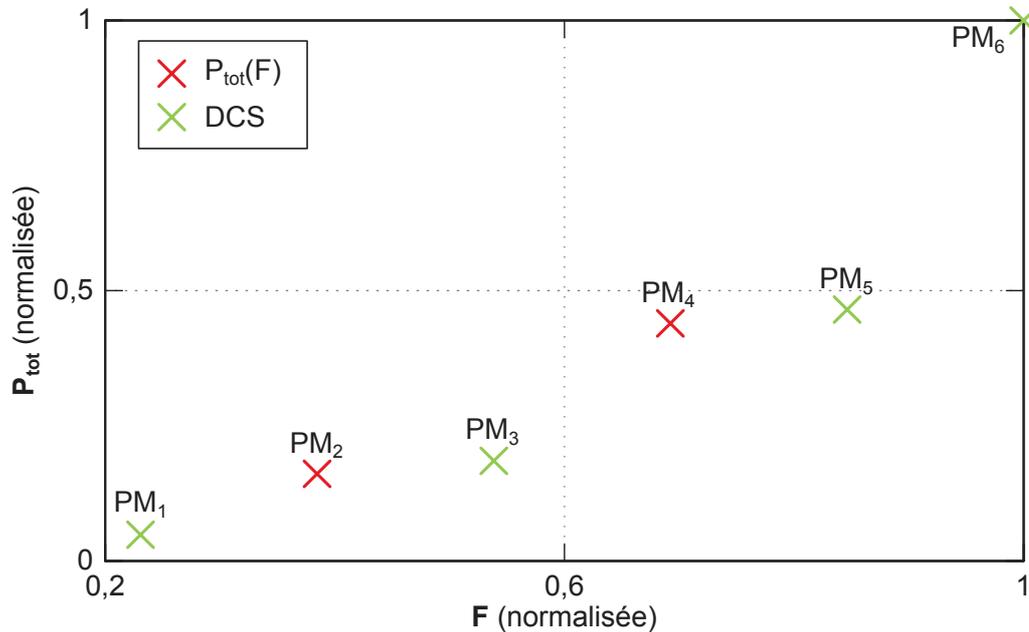


FIGURE 5.4: Ensemble de PM dans le plan (F, P_{tot}) avec P_{tot} la consommation totale et F la fréquence avec les PM appartenant au DCS en vert.

Évaluation du gain en consommation lors de l'exécution d'une tâche

Nous allons évaluer le gain en consommation que peut apporter la méthode proposée lorsqu'un nombre fini de PM est disponible sur le circuit. Lors de l'exécution d'une tâche, différentes situations peuvent être rencontrées. Le gain en consommation est alors évalué de la manière suivante :

- la fréquence cible F_{target} est disponible sur le dispositif par l'application d'un PM dont la fréquence correspond à F_{target} , et ce PM appartient au DCS. La consommation de l'exécution de la tâche avec le PM dont la fréquence correspond à F_{target} appartenant au DCS est comparée à la consommation en exécutant avec le PM dont la fréquence correspond à F_{target} . Comme le PM dont la fréquence correspond à F_{target} appartient déjà au DCS, le gain en consommation est nul ;
- la fréquence cible F_{target} est atteignable en appliquant un PM dont la fréquence correspond à F_{target} mais qui n'appartient pas au DCS. La consommation de l'exécution de la tâche avec les 2 PM appartenant au DCS dont les fréquences encadrent F_{target} est comparée à la consommation de l'exécution de la tâche avec le PM dont la fréquence correspond à F_{target} ;
- la fréquence cible F_{target} n'est pas applicable directement sur le dispositif car aucun PM n'y correspond. La consommation de l'exécution en appliquant les 2 PM dont les fréquences encadrent F_{target} et qui appartiennent au DCS est confrontée à la consommation en exécutant avec les 2 PM dont les fréquences sont les plus proches de F_{target} .

En résumé, nous allons comparer la consommation d'un dispositif en considérant une phase d'identification du DCS, à la consommation de ce même dispositif sans passer par une phase d'identification. Cela revient à évaluer le gain en consommation obtenu en utilisant la méthode proposée. Il est calculé avec la formule ci-dessous :

$$G(F_{target})(\%) = 100 \cdot \frac{P_{tot_{th}}(F_{target}) - P_{tot_{DCS}}(F_{target})}{P_{tot_{th}}(F_{target})} \quad (5.2)$$

avec $G(F_{target})$ le gain en consommation pour la fréquence cible F_{target} en pourcentage, $P_{tot_{th}}(F_{target})$ la consommation sans phase d'identification pour la fréquence cible F_{target} , $P_{tot_{DCS}}(F_{target})$ la consommation avec la méthode proposée pour la fréquence cible F_{target} .

La Figure 5.5 illustre le gain en consommation réalisé en utilisant la méthode proposée. Les PM appartenant au DCS ont été identifiés précédemment, les tâches sont maintenant exécutées en appliquant seulement les PM appartenant au DCS.

Pour les tâches dont la fréquence requise est comprise entre la fréquence du PM_1 et celle du PM_3 , on peut observer que le gain maximal correspond à la fréquence du PM_2 , soit $F_{target} = F_2$. L'exécution d'une tâche dont F_{target} correspond à la fréquence du PM_2 en appliquant les PM_1 et PM_3 au lieu du PM_2 permet de réduire la consommation de 27,55%. La moyenne de gain obtenue en

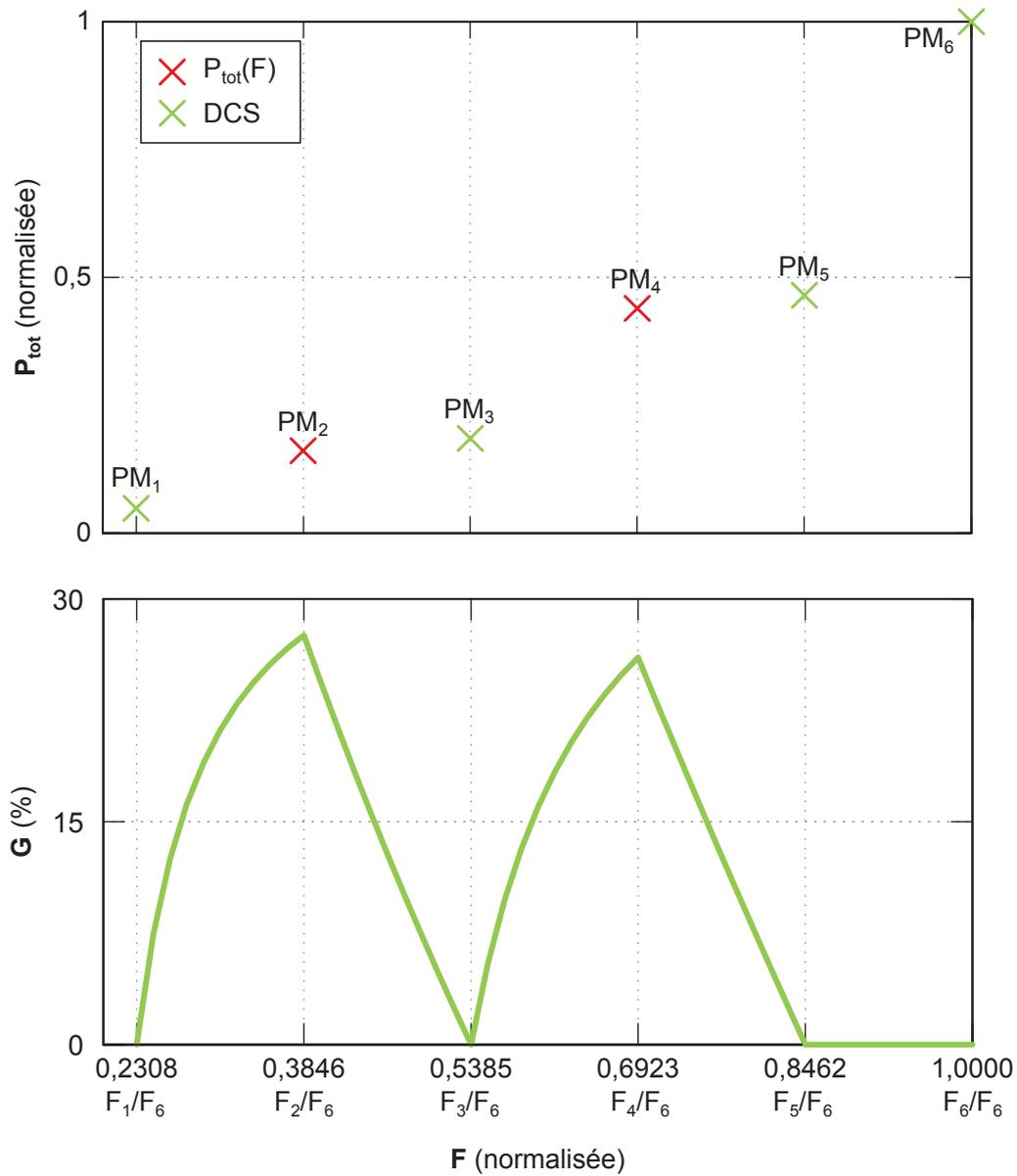


FIGURE 5.5: Consommation P_{tot} en fonction de la fréquence F (en haut). Gain en consommation G avec la méthode proposée en fonction de la fréquence F (en bas).

utilisant la méthode proposée est de 16,09% sur la gamme de fréquence $[F_1 ; F_3]$. Cette moyenne de gain en consommation a été évalué avec la méthode des trapèzes et à l'aide de l'équation (5.3).

$$\overline{G_{GTr}} = \frac{1}{\Delta F} \cdot S_{GTr} \quad (5.3)$$

avec $\overline{G_{GT_r}}$ le gain moyen en consommation obtenu avec la méthode des trapèzes sur la gamme de fréquence considérée, $\Delta F = F_b - F_a$ avec $[F_a ; F_b]$ la gamme de fréquence considérée, S_{GT_r} la surface du gain en fonction de la fréquence obtenu avec la méthode des trapèzes sur la gamme de fréquence considérée.

La consommation peut être réduite en n'appliquant pas le PM_4 pour exécuter les tâches dont la fréquence requise est comprise entre la fréquence du PM_3 et celle du PM_5 . Le gain maximal correspond à la fréquence du PM_4 , soit $F_{target} = F_4$, en exécutant une tâche dont la fréquence correspond à celle du PM_4 avec les PM_3 et PM_5 au lieu du PM_4 la consommation est réduite de 26,06%. La moyenne de gain en consommation obtenue est de 14,74% dans ce cas sur la gamme de fréquence $[F_3 ; F_5]$.

Le gain moyen obtenu avec cette méthode sur toute la gamme de fréquence est de 12,33%.

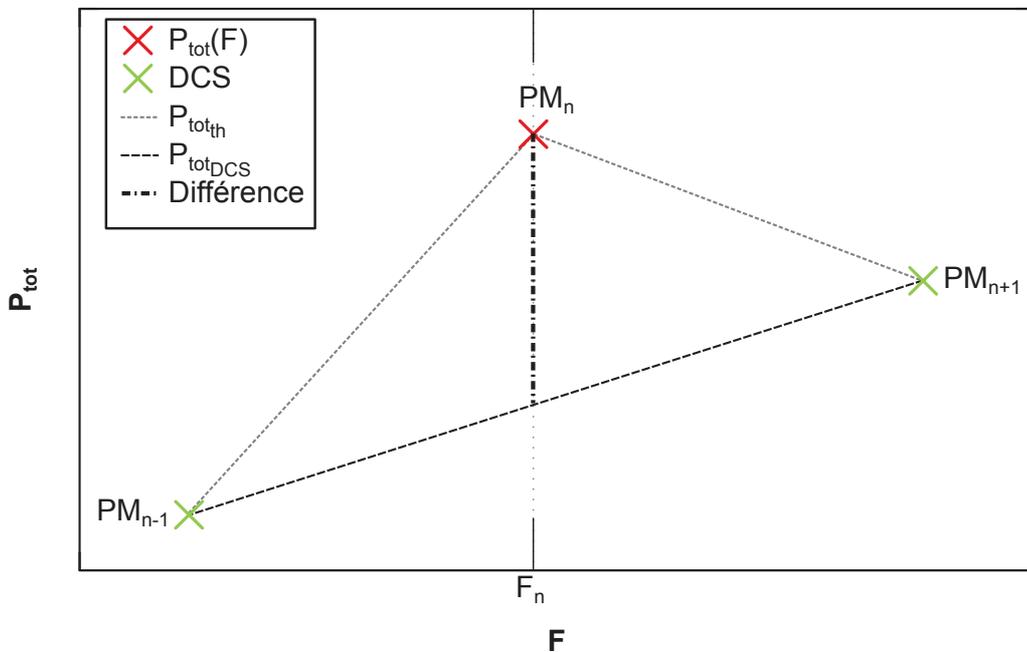


FIGURE 5.6: Lien entre le gain en consommation et le plan (F, P_{tot}) .

Nous observons que des pics de gain sont présents quand la fréquence cible est égale à la fréquence d'un PM n'appartenant pas au DCS. Cela est dû au fait que le gain en consommation est calculé comme la différence entre la consommation sans la méthode $P_{tot_{th}}$ et la consommation avec la méthode $P_{tot_{DCS}}$ normalisée par $P_{tot_{th}}$. Rappelons que la consommation pour chaque fréquence cible correspond à l'interpolation linéaire entre les différents points de fonctionnement tensions-fréquences. La Figure 5.6 illustre sur le plan (F, P_{tot}) , la différence en consommation permettant de calculer le gain pour la fréquence du PM_n , soit

$F_{target} = F_n$. La différence en consommation est maximale entre le PM_n et le segment de droites reliant les PM_{n-1} et PM_{n+1} .

Il en résulte de cette étude que l'utilisation du DCS pour exécuter les tâches permettrait de réduire la consommation jusqu'à 27,55% pour le cas considéré. Par ailleurs, le gain moyen obtenu est supérieure à 12,33% pour ce cas d'étude. Cette première validation permet d'affirmer que la méthode proposée est intéressante pour réduire la consommation d'un circuit. Dans la section suivante, nous allons évaluer le gain en consommation sur un circuit réel.

5.1.2 Cas d'étude : mesures d'un circuit

Notre second cas d'étude est un processeur de traitement du signal ou DSP (Digital Signal Processing) implémenté en technologie STMicroelectronics FD-SOI 28nm, appelé Frisbee [74]. Ce circuit est un prototype permettant d'agir sur les 3 DoF qui nous intéressent, soit la tension d'alimentation V_{dd} , la tension de body bias V_{bb} et la fréquence F . L'actionneur de fréquence est implémenté sur puce (*on-chip*). Les actionneurs V_{dd} et V_{bb} sont implémentés hors puce (*off-chip*). La consommation du circuit a été mesurée sur le circuit pour différentes tensions d'alimentation V_{dd} , tensions de body bias V_{bb} et fréquences F appliquées. La fréquence appliquée pour un couple (V_{dd}, V_{bb}) correspond à la fréquence maximale F_{max} atteignable sans produire d'erreur temporelle. La fréquence maximale atteignable est déterminée à l'aide de capteurs TSM [68].

Les mesures ont été obtenues en exécutant la même tâche pour différentes tensions d'alimentation, tension de body bias et fréquence appliquées. Étant donné que la même tâche a été exécutée pour toutes les mesures de consommation, l'activité moyenne A est donc presque constante. L'activité moyenne A appartient donc à une gamme d'activités.

Sachant que, le profil de consommation en fonction de la fréquence est fonction de l'activité A et de la température T° . La température des mesures correspond à la température ambiante. Afin d'assurer que le circuit revienne à la température ambiante, le circuit a été éteint entre chaque mesure. La température T° est donc quasi-constante. On peut supposer que T° appartient à une gamme de températures.

La tâche exécutée est constituée d'une série de calculs exécutés sur une partie du circuit puis le résultat obtenu est comparé à une valeur de référence correspondant au résultat attendu. Cette comparaison permet de vérifier une seconde fois que la fréquence appliquée n'est pas supérieure à la fréquence maximale atteignable. En effet, une erreur temporelle produira un résultat différent.

La consommation de l'exécution d'une tâche en utilisant la méthode proposée est comparée à la consommation en exécutant la même tâche lorsque la méthode n'est pas disponible. Le gain en consommation que la méthode peut apporter est évalué à l'aide d'un outil d'analyse numérique, MATLAB.

Nous ne considérons que 2 niveaux de tension de body bias V_{bb} et 4 niveaux de tension d'alimentation V_{dd} disponibles sur le circuit, soit 8 PM. Un ensemble discret de PM est donc considéré. Pour une tension d'alimentation V_{dd} , la tension de body bias V_{bb} est augmentée, ce qui permet de booster la fréquence. Nous allons considérer les valeurs suivantes pour les tensions $V_{dd_1} = 0,7$ Volt (V), $V_{dd_2} = 0,9$ V, $V_{dd_3} = 1,1$ V, $V_{dd_4} = 1,3$ V, $V_{bb_L} = 0$ V (niveau non-boosté) et $V_{bb_H} = 1,2$ V (niveau boosté).

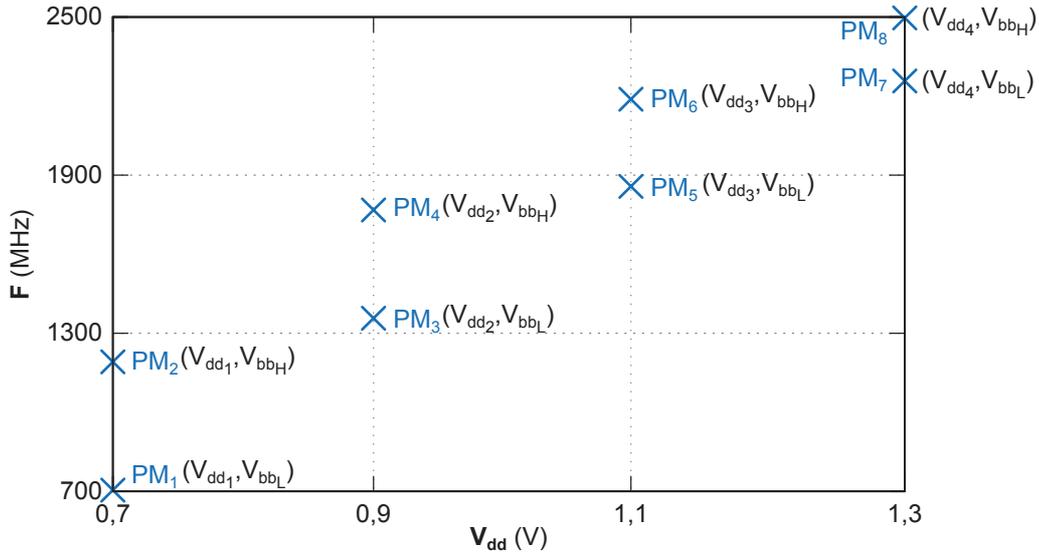


FIGURE 5.7: Ensemble de fréquences disponibles F en fonction de la tension d'alimentation V_{dd} de Frisbee.

TABEAU 5.3: Valeurs de la tension d'alimentation V_{dd} , la tension de body bias V_{bb} , la fréquence F , la consommation P_{tot} pour chaque PM de Frisbee.

PM	V_{dd} (V)	V_{bb} (V)	F (MHz)	P_{tot} (mW)
PM_1	0,7	0	705	73,91
PM_2	0,7	1,2	1191	137,29
PM_3	0,9	0	1357	219,30
PM_4	0,9	1,2	1768	315,89
PM_5	1,1	0	1857	448,14
PM_6	1,1	1,2	2188	590,83
PM_7	1,3	0	2257	796,47
PM_8	1,3	1,2	2497	1073,91

La Figure 5.7 représente la fréquence d'horloge F en fonction de la tension d'alimentation V_{dd} avec V_{bb_L} le niveau non-boosté et V_{bb_H} le niveau boosté. La Figure 5.8 illustre la courbe de la consommation P_{tot} en fonction de la fréquence

F formée par les 8 PM. Les valeurs de la fréquence et de la consommation des PM sont disponibles dans le Tableau 5.3. On peut observer sur la Figure 5.8 que l'ensemble des PM ne forme pas un ensemble globalement convexe et plus précisément un ensemble discrètement convexe dans le plan (F, P_{tot}) . La phase d'identification du DCS permet de réaliser cette sélection.

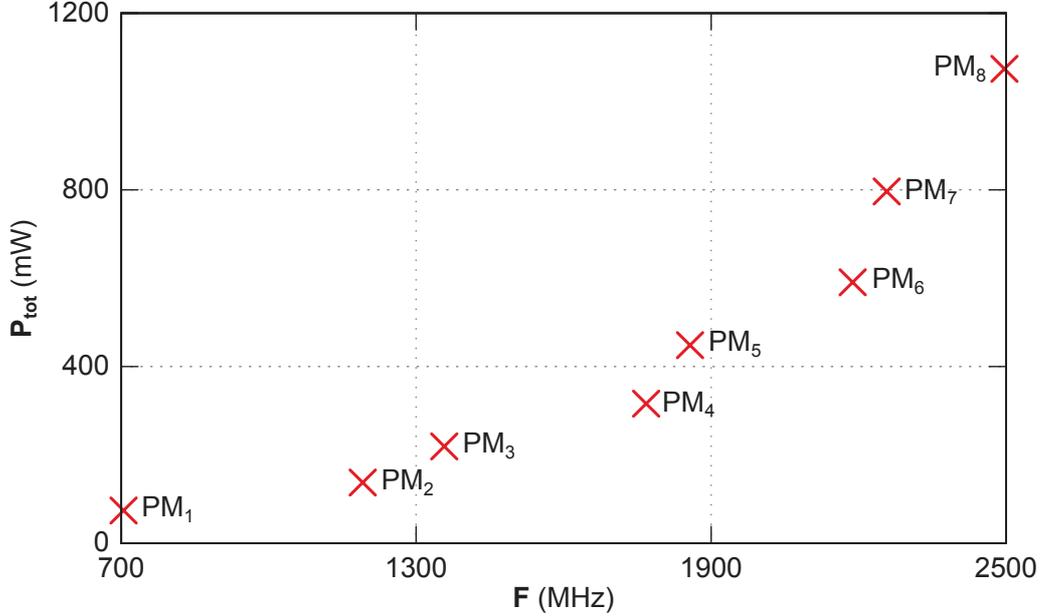


FIGURE 5.8: Ensemble de PM dans le plan (F, P_{tot}) avec P_{tot} la consommation totale du circuit et F la fréquence d'horloge de Frisbee.

Identification du DCS

Nous allons identifier le DCS. Afin de l'identifier, la méthode détermine les valeurs des pentes reliant les PM sur le profil de la consommation en fonction de la fréquence, voir la Figure 5.3. Le Tableau 5.4 illustre la valeur des 3 pentes a_{n-1}^n , a_{n-1}^{n+1} et a_n^{n+1} pour les différents PM, avec $1 < n < N$ et N le nombre de PM disponibles sur le circuit.

Le PM_1 appartient au DCS car le PM dont la consommation est la plus basse appartient au DCS. Les PM dont la fréquence est inférieure à celle de ce PM n'appartiennent pas au DCS.

Le PM_8 appartient au DCS car le PM dont la fréquence est la plus haute appartient au DCS. Les PM dont la consommation est supérieure à celle de ce PM n'appartiennent pas au DCS.

Les PM_1 , PM_2 , PM_4 , PM_6 , PM_8 appartiennent au DCS et sont illustrés en vert sur la Figure 5.9. Ces PM permettront de minimiser la consommation du circuit quelque soit F_{target} . Afin de satisfaire les contraintes de performance tout en

minimisant la consommation, les tâches seront exécutées en appliquant seulement les PM identifiés comme appartenant DCS, possiblement en mode *hopping*.

TABLEAU 5.4: Valeur des pentes afin d'identifier les PM appartenant au DCS pour Frisbee, avec $1 < n < N$.

n	a_{n-1}^n	a_{n-1}^{n+1}	a_n^{n+1}	$a_{n-1}^n \leq a_{n-1}^{n+1} \leq a_n^{n+1} ?$
2	0,1302	0,2229	0,4947	oui
3	0,4947	0,3099	0,2352	non
4	0,2352	0,4577	1,4789	oui
5	1,4789	0,6544	0,4314	non
6	0,4314	0,8708	2,9690	oui
7	2,9690	1,5646	1,1584	non

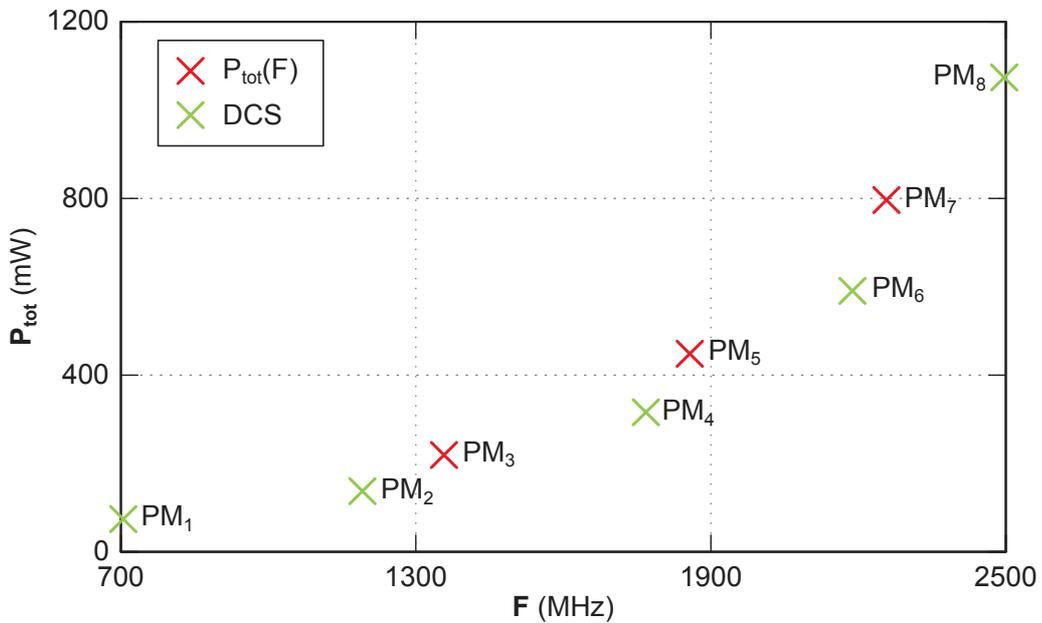


FIGURE 5.9: Ensemble de PM dans le plan (F, P_{tot}) avec P_{tot} la consommation totale et F la fréquence avec les PM appartenant au DCS en vert de Frisbee.

Évaluation du gain en consommation lors de l'exécution d'une tâche

Le gain en consommation en utilisant la méthode proposée va être évalué. Dans la section 5.1.1, la manière dont le gain est déterminé et les différentes situations d'exécution ont été présentées. La consommation de l'exécution d'une tâche en utilisant la méthode proposée est comparée à la consommation de l'exécution de la même tâche sans la méthode.

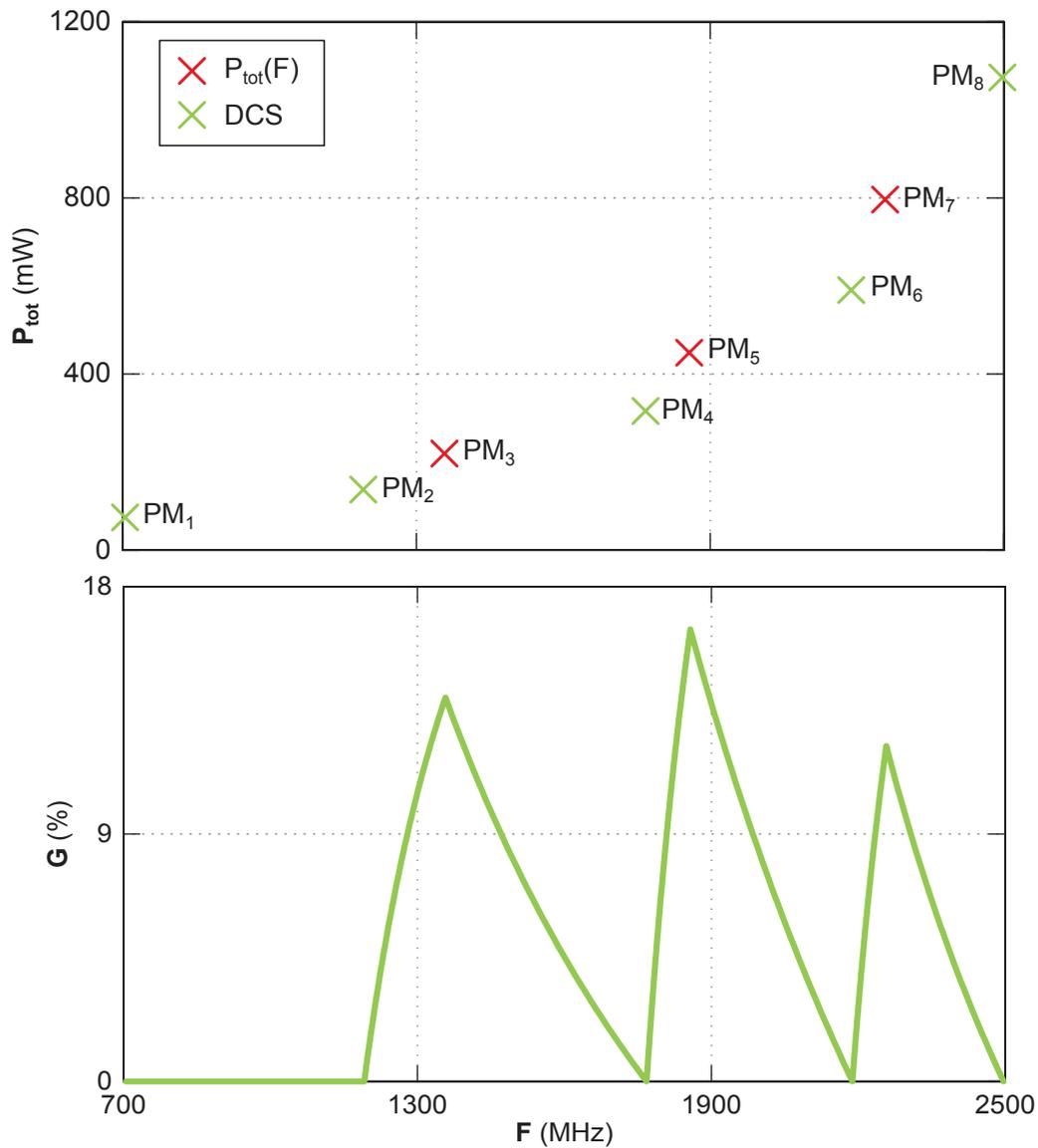


FIGURE 5.10: Consommation P_{tot} en fonction de la fréquence F du circuit Frisbee (en haut). Gain en consommation G avec la méthode proposée en fonction de la fréquence F sur Frisbee (en bas).

La Figure 5.10 montre le gain en consommation en utilisant la méthode proposée. Les PM appartenant au DCS sont identifiés afin d'être appliqués pour exécuter les tâches.

Le Tableau 5.5 illustre la valeur du gain moyen $\overline{G_{GT_r}}$ sur la gamme de fréquence $[F_a ; F_b]$ et la valeur du gain maximal G_{max} pour la fréquence F_i du PM_i sur la gamme de fréquence $[F_a ; F_b]$.

En n'appliquant pas le PM_5 pour exécuter les tâches dont F_{target} est comprise

TABLEAU 5.5: Valeur du gain moyen $\overline{G_{GT_r}}$ sur l'intervalle $[F_a ; F_b]$ considéré et valeur du gain en consommation maximal G_{max} pour la fréquence cible F_i du PM_i .

$[F_a ; F_b]$	$\overline{G_{GT_r}}$ (%)	F_i du PM_i	G_{max} (%)
$[F_2 ; F_4]$	6,70	F_3	13,97
$[F_4 ; F_6]$	7,84	F_5	16,45
$[F_6 ; F_8]$	5,77	F_7	12,21

entre la fréquence du PM_4 et la fréquence du PM_6 . Le gain maximal correspond à $F_{target} = F_4$. En exécutant une tâche dont la fréquence correspond à celle du PM_5 , en appliquant les PM_4 et PM_6 au lieu d'appliquer le PM_5 la consommation est réduite de 16,45%. La moyenne de gain en consommation obtenue sur la gamme de fréquence $[F_4 ; F_6]$ est de 7,84%.

Le gain moyen obtenu avec cette méthode sur toute la gamme de fréquence, soit $[F_1 ; F_8]$, est de 4,99%. Même si on ne gagne que 4,99% en moyenne, on a des pics de gains. Le gain maximal obtenu pour le circuit implémenté en technologie FD-SOI est de 16,45%. Ce gain paraît important. La méthode peut être implémentée si son surcoût est inférieur au gain en consommation qu'elle apporte.

5.1.3 Conclusion

Pour un circuit avec un nombre fini de PM, les résultats présentés dans cette section montrent que la méthode proposée permet de réduire la consommation. Le gain a été évalué à partir de simulations de consommation pour un RO et de mesures obtenues sur un circuit réel en technologie FD-SOI 28nm.

Nous venons d'étudier le cas où les actionneurs ont des sorties discrètes. Nous allons maintenant étudier la combinaison de deux actionneurs avec des sorties continues et un actionneur ayant une sortie discrétisée. L'actionneur discret sera l'actionneur de tension d'alimentation. Rappelons, que ce choix est orienté par la prise en compte des contraintes d'implémentation physique (coût, surface, rendement des actionneurs, ...).

5.2 Application à un circuit composé d'un nombre infini de PM

Dans cette section, nous allons valider la méthode proposée pour un système disposant de trois DoF, dont un est discret. Ainsi nous obtenons un nombre infini de PM pour notre cas d'étude. Nous allons évaluer le gain en consommation que

peut apporter la méthode présentée dans le chapitre 4 lorsqu'un nombre infini de PM est disponible sur le circuit. La méthode est composée de deux phases, la phase d'identification des PM appartenant au sous-ensemble convexe par morceaux croissant ou PWCS. La seconde phase de la méthode est la phase de fonctionnement qui consiste à appliquer les PM afin de minimiser la consommation en cours d'exécution.

Le circuit considéré est Frisbee [74], voir la section 5.1.2. La même tâche est exécutée en modifiant les valeurs (F, V_{dd}, V_{bb}) pour mesurer la consommation du circuit. Il est éteint entre chaque mesure. L'activité moyenne A est quasi-constante. De même pour la température.

Nous allons considérer que l'actionneur de tension d'alimentation est discret. Les valeurs en sorties des actionneurs F et V_{bb} sont discrétisées avec un pas très fin. Ils sont donc considérés comme continus. Nous allons évaluer l'apport de la méthode proposée dans la section 4.4 du chapitre 4 dans les cas où 2, 3 et 4 niveaux de tension d'alimentation sont disponibles.

La tension d'alimentation minimale $V_{dd_{min}}$ est de 0,7 V et la tension d'alimentation maximale $V_{dd_{max}}$ est de 1,3 V. Les valeurs des tensions d'alimentation pour les exemples sont choisies arbitrairement avec un pas entre chaque tension d'alimentation d'environ 0,2 V. Cela permet de couvrir une gamme de fréquence importante. Rappelons que la tension de body bias est réduite à l'intervalle [0 ; 1500] mV.

Rappelons que la méthode proposée consiste à rendre globalement convexe un ensemble continu par morceaux grâce à la phase d'identification du PWCS. L'identification du PWCS est réalisée pour une température T° et une activité A . La phase de fonctionnement consiste à exécuter les tâches avec les PM les plus adéquats appartenant au PWCS.

5.2.1 Identification du PWCS

L'identification du PWCS se fait à partir du profil de consommation du circuit en fonction de la fréquence d'horloge. La consommation dépend de la température T° et de l'activité A . Par soucis d'implémentation, on considère des gammes de T° et A car cela nous évite de stocker toutes les limites basses (premier PM appartenant au PWCS pour un V_{dd}) et hautes (dernier PM appartenant au PWCS pour un V_{dd}) pour chaque température et chaque activité. L'utilisation de gammes de températures et de gammes d'activités revient à discrétiser les modèles d'évolution de la limite basse et de la limite haute en fonction de T° et A pour une tension d'alimentation V_{dd_i} donnée. L'implémentation de modèles pour déterminer les limites est plus coûteuse en termes de calculs car à chaque modification d'une des entrées du système ou de la température ou de l'activité, les limites devront être recalculées.

Pour une tension d'alimentation V_{dd_i} , la consommation totale $P_{tot_i}(F)$, peut

être modélisée par un polynôme quadratique en fréquence. Cependant, cette modélisation n'est pas assez précise. Nous allons réaliser l'identification du PWCS grâce à un outil de calcul numérique (MATLAB) pour éviter d'avoir à modéliser pour chaque tension d'alimentation la courbe $P_{tot_i}(F)$ par une fonction cubique de la fréquence (voir la Figure 5.11 par exemple). Rappelons, qu'il existe des algorithmes permettant d'optimiser le nombre de calculs pour déterminer le PWCS pour un nombre fini de points. Par exemple, la marche de Jarvis [87] ou le parcours de Graham [86]. Dans ce cas, il faudrait discrétiser la fonction $P_{tot}(F)$ avec un pas fin pour pouvoir déterminer le PWCS avec l'un des deux algorithmes (marche de Jarvis ou parcours de Graham).

Exemple avec deux niveaux de tension V_{dd}

Le profil $P_{tot}(F)$ obtenu pour le circuit est visible sur la Figure 5.11 lorsque l'on considère 2 niveaux de tension d'alimentation pour l'actionneur de tension V_{dd} . L'ensemble des PM appartenant au PWCS est tracé en vert sur la Figure 5.11. Un PM est défini par le couple (F, P_{tot}) obtenu en agissant sur le triplet d'entrées du système (F, V_{dd}, V_{bb}) . Les tensions d'alimentation pour cet exemple ont été choisies arbitrairement avec $V_{dd_1} = 0,8$ V et $V_{dd_2} = 1,0$ V. Les valeurs minimales et maximales de V_{bb} et de F des PM appartenant au PWCS pour chaque V_{dd_i} sont données dans le Tableau 5.6. Rappelons que ces valeurs correspondent au premier et dernier PM appartenant au PWCS pour V_{dd_i} .

TABLEAU 5.6: Valeurs limites lorsque 2 niveaux de V_{dd} sont disponibles.

i	V_{dd_i} (V)	$V_{bb_{i-}}$ (mV)	$V_{bb_{i+}}$ (mV)	F_{i-} (MHz)	F_{i+} (MHz)
1	0,8	0	1500	1021	1600
2	1,0	958	1500	1930	2100

Nous pouvons remarquer que les gammes de tensions de body bias $[V_{bb_{i-}} ; V_{bb_{i+}}]$ sont différentes pour chaque tension d'alimentation V_{dd_i} . Rappelons que $V_{bb_{i-}}$ (resp $V_{bb_{i+}}$) correspond à la valeur de la tension de body bias associée à la limite basse (resp. haute) pour la tension d'alimentation V_{dd_i} . De même F_{i-} (resp F_{i+}) correspond à la valeur de la fréquence associée à la limite basse (resp. haute) pour V_{dd_i} . Le fait d'avoir différentes gammes de $[V_{bb_{i-}} ; V_{bb_{i+}}]$ pour chaque tension d'alimentation V_{dd_i} est totalement normal car le PWCS dépend de la courbure de chaque morceau de courbe $P_{tot_i}(F)$ de la courbe $P_{tot}(F)$. Le pourcentage de la courbe appartenant au PWCS pour cet exemple est de 71,33%.

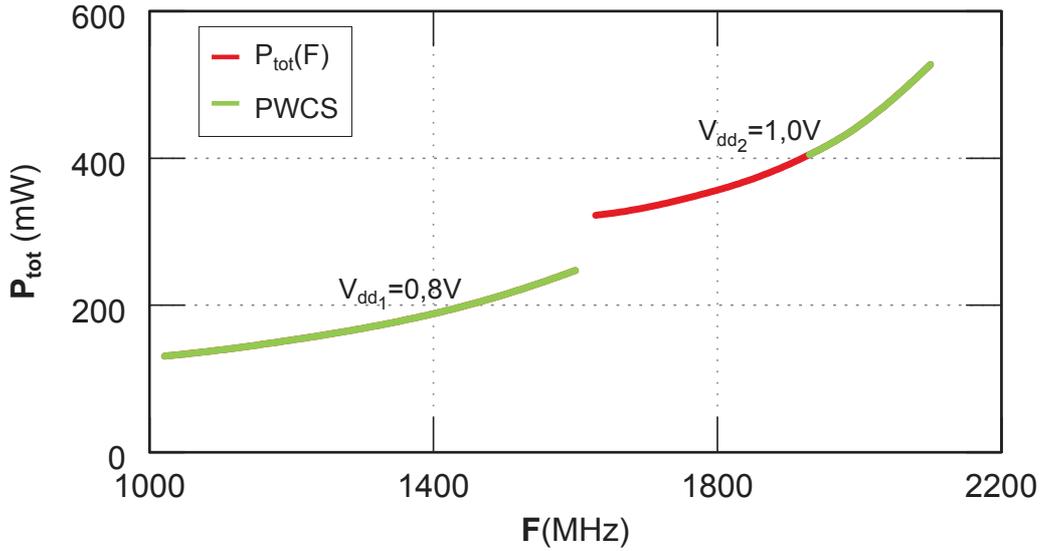


FIGURE 5.11: Ensemble de PM dans le plan (F, P_{tot}) pour 2 niveaux de V_{dd} avec les PM appartenant au PWCS en vert.

Exemple avec trois niveaux de tension V_{dd}

La Figure 5.12 montre la consommation P_{tot} en fonction de la fréquence F quand 3 niveaux de tension d'alimentation sont considérés pour l'actionneur de tension V_{dd} . Les valeurs de tensions d'alimentation ont été choisies de manière arbitraire. Elles sont différentes de l'exemple précédent. On a $V_{dd_1} = 0,72$ V, $V_{dd_2} = 0,92$ V et $V_{dd_3} = 1,11$ V.

TABLEAU 5.7: Valeurs limites lorsque 3 niveaux de V_{dd} sont disponibles.

i	V_{dd_i} (V)	$V_{bb_{i-}}$ (mV)	$V_{bb_{i+}}$ (mV)	F_{i-} (MHz)	F_{i+} (MHz)
1	0,72	0	1500	771	1386
2	0,92	1059	1500	1767	1914
3	1,11	1059	1500	2182	2300

L'ensemble des PM en vert sur la Figure 5.12 a été identifié comme appartenant au PWCS. Pour cet exemple, les valeurs limites pour chaque tension d'alimentation V_{dd_i} sont disponibles dans le Tableau 5.7. Deux gammes de tensions de body bias $[V_{bb_{i-}} ; V_{bb_{i+}}]$ pour la tension d'alimentation V_{dd_i} parmi trois sont identiques, ceci est fortuit. Les gammes de tensions $[V_{bb_{i-}} ; V_{bb_{i+}}]$ sont issues du PWCS dont la détermination va complètement dépendre de la courbure de chaque morceau de la courbe $P_{tot}(F)$. En effet, la gamme $[V_{bb_{i-}} ; V_{bb_{i+}}]$ ne sera pas forcément identique pour une tension V_{dd_i} différente. Ici, 58,05% de la courbe

$P_{tot}(F)$ appartient au PWCS pour cet exemple.

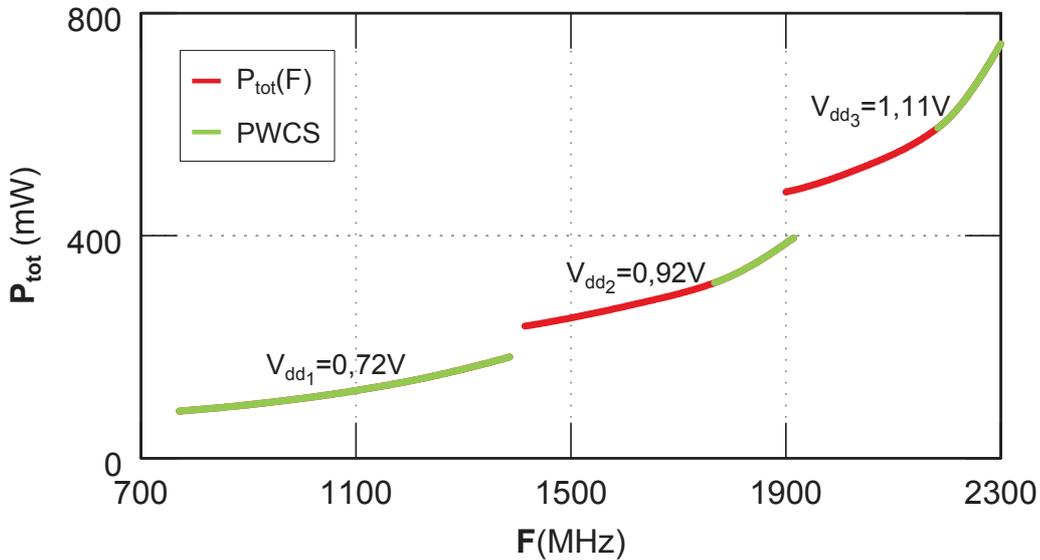


FIGURE 5.12: Ensemble de PM dans le plan (F, P_{tot}) pour 3 niveaux de V_{dd} avec les PM appartenant au PWCS en vert.

Exemple avec quatre niveaux de tension V_{dd}

Dans le dernier exemple, on considère 4 niveaux de tension d'alimentation. De nouveau, les valeurs de tension d'alimentation ont été choisies de manière à être différentes des deux exemples précédents et de manière arbitraire. Les valeurs sont $V_{dd_1} = 0,7$ V, $V_{dd_2} = 0,9$ V, $V_{dd_3} = 1,1$ V et $V_{dd_4} = 1,3$ V.

TABLEAU 5.8: Valeurs limites lorsque 4 niveaux de V_{dd} sont disponibles.

i	V_{dd_i} (V)	$V_{bb_{i-}}$ (mV)	$V_{bb_{i+}}$ (mV)	F_{i-} (MHz)	F_{i+} (MHz)
1	0,7	0	1500	705	1314
2	0,9	996	1500	1699	1871
3	1,1	1021	1324	2139	2223
4	1,3	819	1500	2421	2557

La Figure 5.13 illustre la consommation P_{tot} en fonction de la fréquence F pour 4 niveaux de V_{dd} . Le sous-ensemble des PM appartenant au PWCS est identifié en vert sur la Figure 5.13. Les valeurs limites pour chaque tension d'alimentation V_{dd} sont disponibles dans le Tableau 5.8 pour l'exemple avec 4 niveaux de V_{dd} . De la même manière, les gammes de tensions de body bias $[V_{bb_{i-}} ; V_{bb_{i+}}]$ sont différentes

pour chaque tension d'alimentation V_{dd_i} . Notons que seulement 54,52% de la courbe $P_{tot}(F)$ appartient au PWCS pour cet exemple.

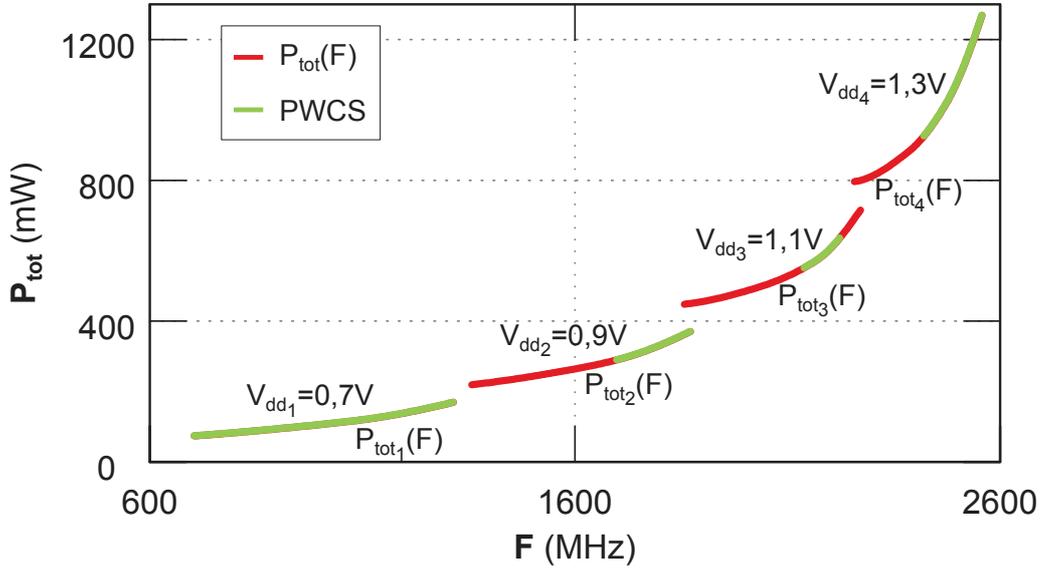


FIGURE 5.13: Ensemble de PM dans le plan (F, P_{tot}) pour 4 niveaux de V_{dd} avec les PM appartenant au PWCS en vert.

Remarque sur l'identification du PWCS

Nous pouvons remarquer que pour la plus petite tension d'alimentation V_{dd_1} , toute la courbe $P_{tot_1}(F)$ appartient au PWCS, et ce pour les trois exemples présentés ci-avant. Cela est dû au pas choisi entre V_{dd_1} et V_{dd_2} . Si, le pas entre V_{dd_1} et V_{dd_2} est plus petit, une partie de la courbe $P_{tot_1}(F)$ pour V_{dd_1} peut ne pas appartenir au PWCS car les courbes $P_{tot_1}(F)$ et $P_{tot_2}(F)$ se seraient croisées. En d'autres termes, elles auraient des fréquences communes. Ceci est due à la courbure de $P_{tot_i}(F)$ de chaque V_{dd_i} . Si, le pas entre V_{dd_1} et V_{dd_2} est plus grand, il se peut que toute la courbe $P_{tot_1}(F)$ appartienne au PWCS.

Le PWCS a été obtenu à l'aide d'un outil d'analyse numérique (correspondant à la phase d'identification de la méthode proposée) pour les 3 exemples. La seconde phase de la méthode est la phase d'exécution. Nous allons évaluer le gain en consommation que la méthode proposée peut apporter pour les 3 exemples précédemment vus.

5.2.2 Évaluation du gain en consommation lors de l'exécution d'une tâche

Dans cette section, nous allons évaluer l'apport de notre méthode grâce aux profils consommation-fréquence extraits du circuit Frisbee. Rappelons que le gain

en consommation est évalué avec un outil d'analyse numérique, MATLAB. Nous allons comparer la consommation associée à l'exécution d'une tâche avec la fréquence cible F_{target} en appliquant le PM dont la fréquence correspond à F_{target} sans prise en compte du PWCS, par rapport à la consommation lorsque cette tâche (c'est-à-dire même contrainte de performance) est exécutée avec le Mode 2. Rappelons que ce mode permet d'exécuter la tâche en mode *hopping* en utilisant les 2 PM appartenant au PWCS dont les fréquences sont les plus proches de la fréquence cible F_{target} . Notons que si F_{target} correspond à la fréquence de 2 PM, nous considérerons pour le calcul du gain le PM ayant la consommation minimale. Si F_{target} n'est pas directement disponible grâce à un PM, le gain sera considéré comme nul car aucun point de comparaison n'est disponible. En général, si F_{target} n'existe pas, le mode de fonctionnement adopté par les concepteurs est d'appliquer une fréquence supérieure à F_{target} puis le reste du temps le circuit n'exécute rien. La consommation d'une telle exécution n'est pas toujours optimale car elle dépend fortement du ratio entre la consommation statique et la consommation dynamique.

Le gain en consommation apporté par le Mode 2 (M2) sera illustré pour les 3 exemples précédents. Les PM en Mode 1 (M1) appartiennent au PWCS, c'est pourquoi le gain en consommation est nul pour ces PM. Pour les fréquences cibles non atteignables par un PM appartenant au PWCS, la consommation équivalente à appliquer directement F_{target} est comparée à la consommation équivalente à appliquer les deux PM dont les fréquences sont les plus proches de F_{target} , appartenant au PWCS, voir la formule (5.4).

$$\begin{aligned} G(F_{target})(\%) &= 100 \cdot \frac{P_{tot_{th}}(F_{target}) - P_{tot_{PWCS}}(F_{target})}{P_{tot_{th}}(F_{target})} \\ &= 100 \cdot \frac{P_{tot_{th}}(F_{target}) - (\alpha \cdot P_{tot_{i+1}}(F_{(i+1)-}) + (1-\alpha) \cdot P_{tot_i}(F_{i+}))}{P_{tot_{th}}(F_{target})} \end{aligned} \quad (5.4)$$

avec $G(F_{target})$ le gain en consommation pour la fréquence cible F_{target} en pourcentage, $P_{tot_{th}}(F_{target})$ la consommation associée à la fréquence cible F_{target} qui est applicable, $P_{tot_{PWCS}}(F_{target})$ la consommation pour la fréquence cible F_{target} en exécutant en Mode 2 et α représente le ratio de temps que l'on va passer à la fréquence $F_{(i+1)-}$ correspondant à la limite basse de $V_{dd_{i+1}}$.

Exemple avec deux niveaux de tension V_{dd}

La Figure 5.14 montre le gain en consommation que la méthode proposée peut apporter si elle est appliquée à ce circuit lorsque 2 niveaux de V_{dd} sont disponibles. Le gain maximal sur la gamme de fréquence où il est préférable d'utiliser le Mode 2, soit $]F_{1+}; F_{2-}[$, est de 19,07% pour une fréquence cible de 1629 MHz. Le Tableau 5.9 donne le gain pour différentes F_{target} . La moyenne de gain en consommation sur la gamme $]F_{1+}; F_{2-}[$ est de 6,55%. Cette moyenne de gain en consommation a été obtenue avec la méthode des trapèzes.

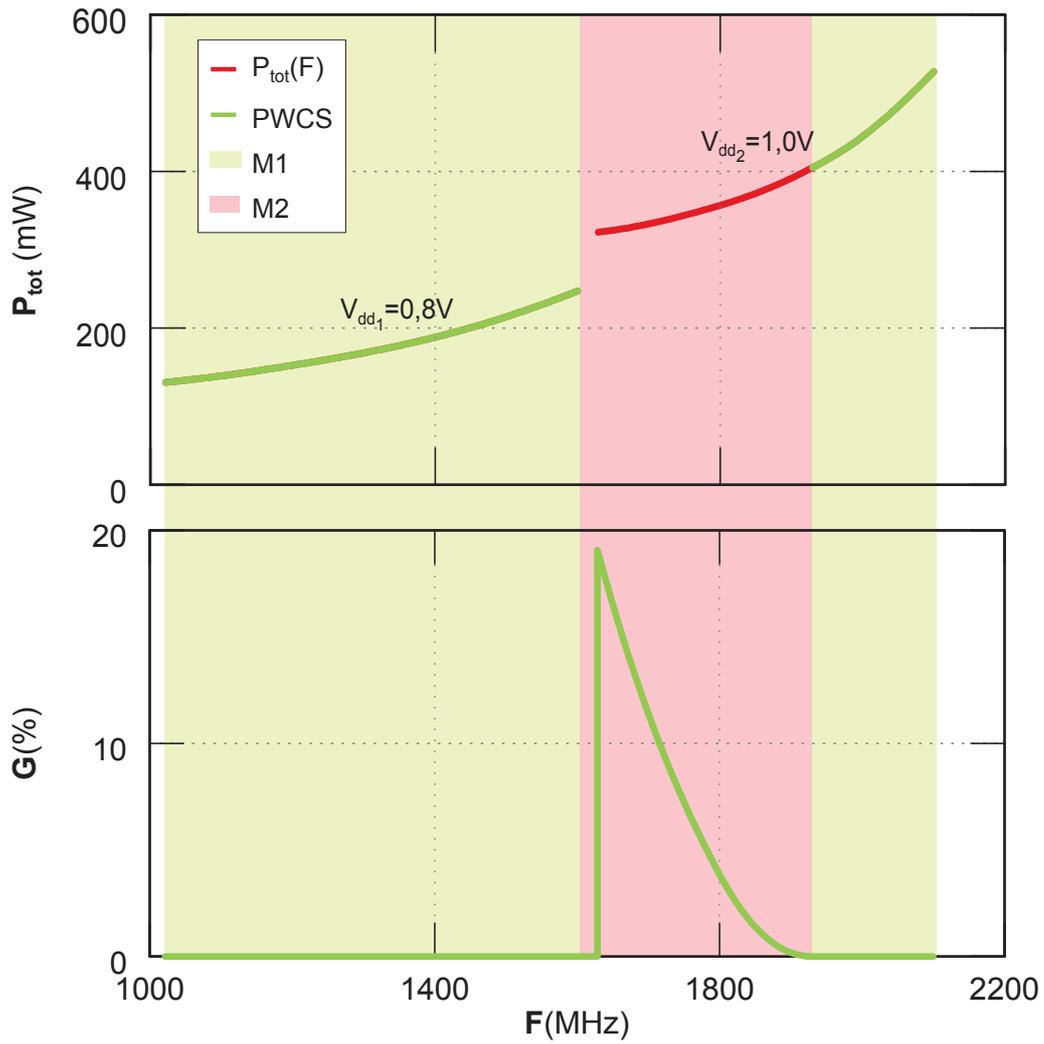


FIGURE 5.14: Consommation P_{tot} en fonction de la fréquence F (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 2 niveaux de V_{dd} (en bas).

TABEAU 5.9: Gain en consommation G en exécutant avec le Mode 2 pour 2 V_{dd} .

F_{target} (MHz)	i	$V_{bb_{i+}}$ (mV)	$V_{bb_{(i+1)-}}$ (mV)	F_{i+} (MHz)	$F_{(i+1)-}$ (MHz)	G (%)
1629	1	1500	958	1600	1930	19,07
1700	1	1500	958	1600	1930	11,41
1799	1	1500	958	1600	1930	3,91
1902	1	1500	958	1600	1930	0,16

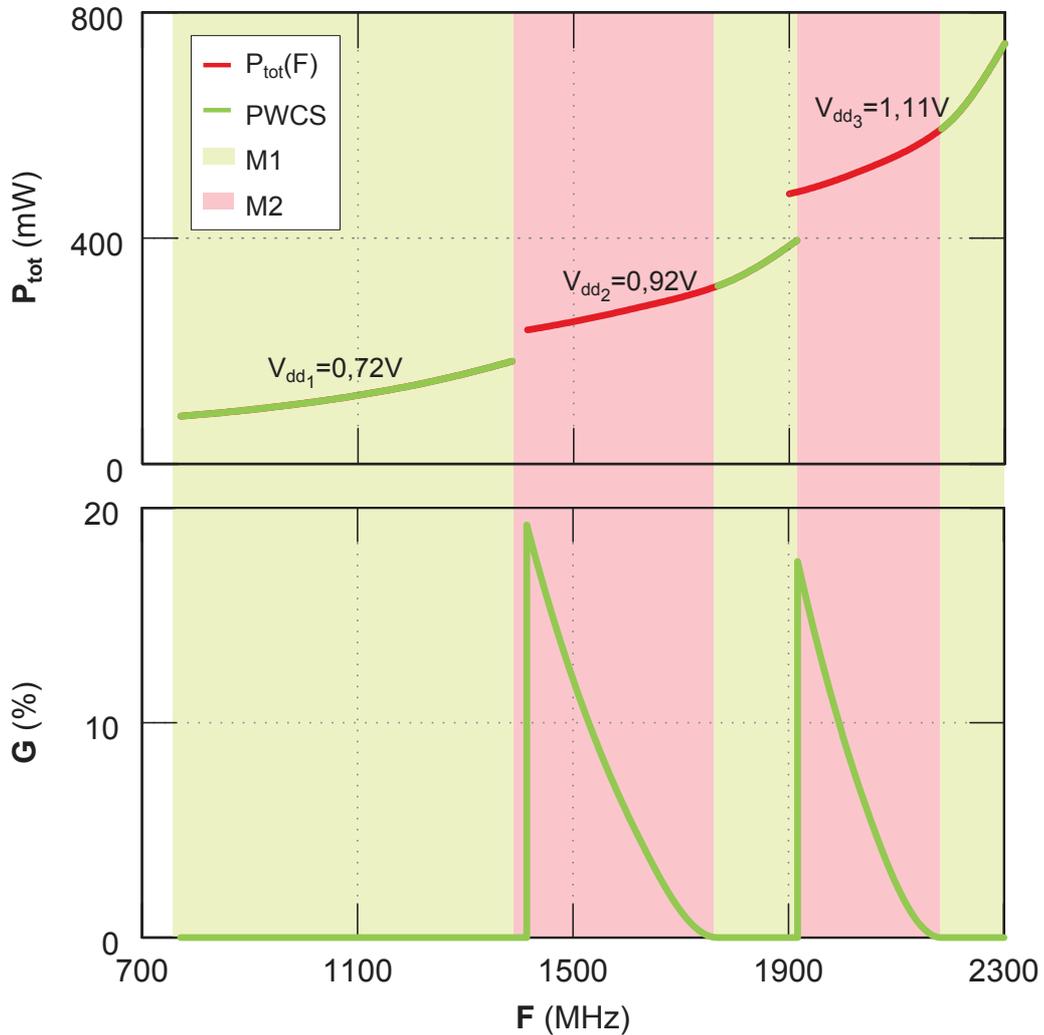
Exemple avec trois niveaux de tension V_{dd} 

FIGURE 5.15: Consommation P_{tot} en fonction de la fréquence F (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 3 niveaux de V_{dd} (en bas).

Le gain en consommation en fonction de F_{target} en utilisant la méthode proposée quand on a 3 niveaux de V_{dd} est donné sur la Figure 5.15. Le Tableau 5.10 montre le gain en consommation pour différentes F_{target} . Le gain maximal sur la première gamme de fréquence en Mode 2, soit $]F_{1+}; F_{2-}[$, est de 19,20% pour $F_{target} = 1414$ MHz, voir le Tableau 5.10 et la Figure 5.15. La moyenne de gain en consommation sur cette gamme de fréquence (*i.e.* $]F_{1+}; F_{2-}[$) est de 7,27%. Le gain maximal sur la seconde gamme de fréquence en Mode 2, soit $]F_{2+}; F_{3-}[$, est de 17,51% pour $F_{target} = 1917$ MHz, voir le Tableau 5.10 et la Figure 5.15. La moyenne de gain en consommation sur cette gamme de fréquence

(*i.e.* $]F_{2+}; F_{3-}[$) est de 6,60%. La moyenne de gain en consommation sur toutes les gammes de fréquence en Mode 2 (*i.e.* sur $]F_{1+}; F_{2-}[$ et sur $]F_{2+}; F_{3-}[$), apportée par la méthode proposée est de 6,98%.

TABLEAU 5.10: Gain en consommation G en exécutant avec le Mode 2 pour 3 V_{dd} .

F_{target} (MHz)	i	$V_{bb_{i+}}$ (mV)	$V_{bb_{(i+1)-}}$ (mV)	F_{i+} (MHz)	$F_{(i+1)-}$ (MHz)	G (%)
1414	1	1500	1059	1386	1767	19,20
1515	1	1500	1059	1386	1767	10,95
1616	1	1500	1059	1386	1767	4,98
1717	1	1500	1059	1386	1767	0,72
1917	2	1500	1059	1914	2182	17,51
2018	2	1500	1059	1914	2182	8,04
2119	2	1500	1059	1914	2182	1,56

Exemple avec quatre niveaux de tension V_{dd}

TABLEAU 5.11: Gain en consommation G en exécutant avec le Mode 2 pour 4 V_{dd} .

F_{target} (MHz)	i	$V_{bb_{i+}}$ (mV)	$V_{bb_{(i+1)-}}$ (mV)	F_{i+} (MHz)	$F_{(i+1)-}$ (MHz)	G (%)
1357	1	1500	996	1314	1699	16,61
1500	1	1500	996	1314	1699	6,43
1651	1	1500	996	1314	1699	0,51
1875	2	1500	1021	1871	2139	17,31
2000	2	1500	1021	1871	2139	5,67
2101	2	1500	1021	1871	2139	0,47
2233	3	1324	819	2223	2421	0,09
2271	3	1324	819	2223	2421	1,17
2272	3	1324	819	2223	2421	11,48
2350	3	1324	819	2223	2421	3,23
2401	3	1324	819	2223	2421	0,30

La Figure 5.16 montre le gain en consommation en fonction de F_{target} que la méthode proposée peut apporter si elle est appliquée à ce circuit pour le cas où 4 niveaux de V_{dd} sont disponibles. Le Tableau 5.11 illustre le gain en consommation pour différentes F_{target} . Le gain maximal sur la première gamme de fréquence en

Mode 2, soit $]F_{1+}; F_{2-}[$, est de 16,61% pour une fréquence cible égale à 1357 MHz, voir le Tableau 5.11 et la Figure 5.16. La moyenne de gain en consommation sur la première gamme de fréquence est de 6,01%.

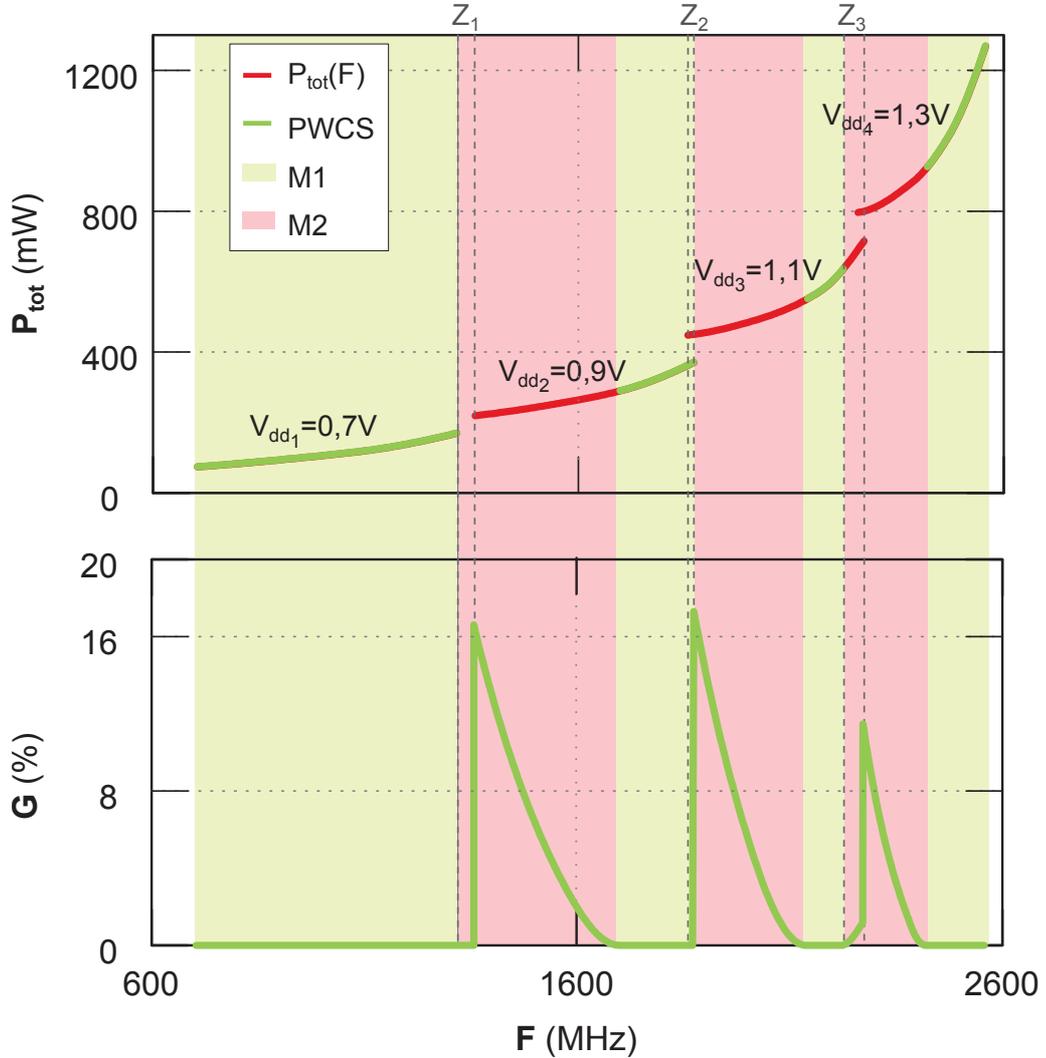


FIGURE 5.16: Consommation P_{tot} en fonction de la fréquence F (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 4 niveaux de V_{dd} (en bas).

Le gain maximal sur la seconde gamme de fréquence en Mode 2, soit $]F_{2+}; F_{3-}[$, est de 17,31% pour $F_{target} = 1875$ MHz, voir le Tableau 5.11 et la Figure 5.16. La moyenne de gain en consommation sur la seconde gamme de fréquence est de 6,24%. Le gain maximal sur la troisième et dernière gamme de fréquence en Mode 2, soit $]F_{3+}; F_{4-}[$, est de 11,48% pour $F_{target} = 2272$ MHz, voir le Tableau 5.11 et la Figure 5.16. La moyenne de gain en consommation sur cette dernière gamme de fréquence est de 3,27%. La moyenne de gain en consommation sur toutes les

gammes de fréquences en Mode 2 (*i.e.* $]F_{1+}; F_{2-}[$, $]F_{2+}; F_{3-}[$ et $]F_{3+}; F_{4-}[$), apportée par la méthode proposée en Mode 2 est de 5,41%.

Analyse du gain en consommation avec la méthode proposée

Nous allons analyser la Figure 5.16 mais l'analyse reste identique pour les Figures 5.14 et 5.15. Dans la zone Z_1 , identifiée sur la Figure 5.16, aucun PM permet d'atteindre les fréquences appartenant à cette zone. Dans cette zone, notre méthode permet de gagner en consommation par rapport à la non utilisation de la méthode mais nous ne pouvons pas évaluer le gain. Si nous avions évalué le gain, le résultat ne serait pas juste car nous aurions évalué le gain de deux manières différentes. Nous aurions comparé la consommation de l'exécution en Mode 2 dans cette zone à la consommation de l'exécution de la tâche en appliquant la fréquence la plus grande disponible sur le circuit puis le circuit sera mis en *standby*. En *standby*, la consommation est dominée par la consommation statique. Alors que le gain est évalué en comparant la consommation de l'exécution en Mode 2 à la consommation de l'exécution avec un PM dont la fréquence correspond à F_{target} et n'appartient pas au PWCS.

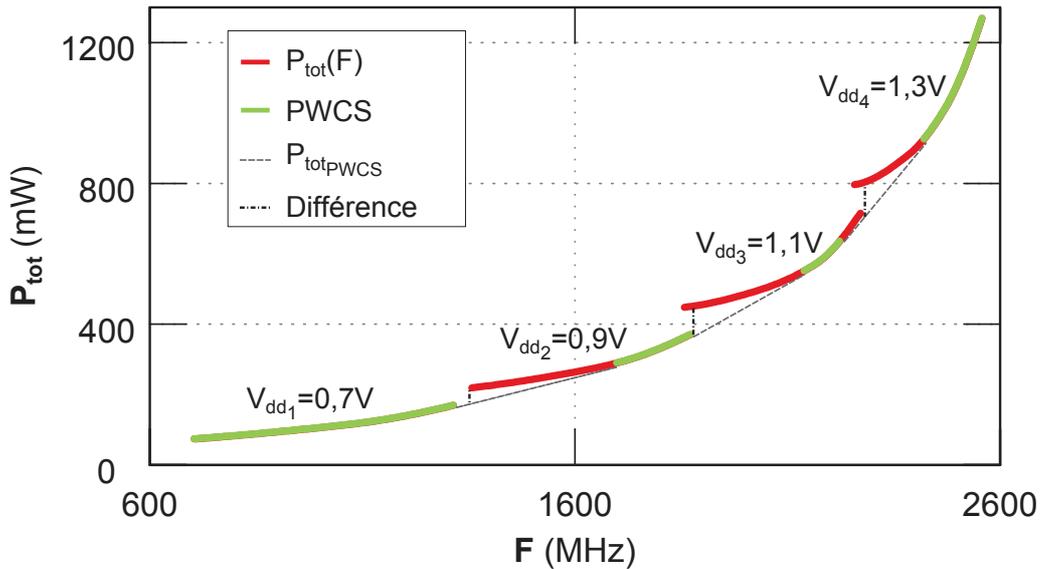


FIGURE 5.17: Lien entre le maximum de gain en consommation et le plan (F, P_{tot}) pour Frisbee.

À la fin de cette zone, nous pouvons observer un maxima de gain en consommation. Cette augmentation soudaine est due à la différence de consommation entre les PM. Si nous traçons un segment reliant la limite haute de V_{dd1} et la limite basse de V_{dd2} sur $P_{tot}(F)$, la différence de consommation entre le premier PM pour V_{dd2} et la consommation pour la même fréquence sur ce segment est supérieure aux différences de consommation que l'on peut obtenir avec les autres

PM dont les fréquences sont comprises dans l'intervalle $]F_{1+}; F_{2-}[$, voir la Figure 5.17.

Dans la zone Z_2 de la Figure 5.16, le gain en consommation est nul. Cela est dû au fait que les fréquences cibles comprises dans la zone Z_2 sont atteignables par deux PM. Sachant qu'il y a un des deux PM qui appartient au PWCS. Nous avons utilisé pour calculer le gain les PM ayant une consommation minimale parmi ces PM, soit les PM appartenant au PWCS.

Dans la zone Z_3 de la Figure 5.16, nous remarquons que le gain en consommation augmente avant d'atteindre un maxima. Les PM dont la fréquence est supérieure (resp. inférieure) à la fréquence F_{3+} (resp. F_{4-}) pour la tension V_{dd_3} (resp. V_{dd_4}) n'appartiennent pas au PWCS. Si nous traçons le segment reliant la limite haute de V_{dd_3} et la limite basse de V_{dd_4} sur $P_{tot}(F)$, en pointillé sur la Figure 5.17, les différences de consommation varient. La différence de consommation pour F_{target} correspond à la différence entre la consommation du PM correspondant à F_{target} et n'appartenant pas au PWCS et la consommation équivalente en mode *hopping* pour F_{target} . En effet, les différences de consommation pour les PM n'appartenant pas au PWCS pour V_{dd_3} sont inférieures aux différences de consommation que l'on peut obtenir pour les PM n'appartenant pas au PWCS pour V_{dd_4} . Sur la Figure 5.17, les différences de consommation sont tracées avec des lignes mixtes.

Pour l'estimation des gains en consommation présentés jusqu'ici, le coût de commutation des actionneurs n'a pas été pris en compte. Nous discuterons de cet aspect à la section 5.2.4. Par ailleurs, une exécution en Mode 2 implique d'exécuter la tâche avec deux triplets, à savoir $(F_{i+}, V_{dd_i}, V_{bb_{i+}})$ et $(F_{(i+1)-}, V_{dd_{i+1}}, V_{bb_{(i+1)-}})$. Par conséquent, des mécanismes dit de *hopping* de $V_{dd_{i+1}}$ à V_{dd_i} , de $F_{(i+1)-}$ à F_{i+} et de $V_{bb_{(i+1)-}}$ à $V_{bb_{i+}}$ devraient être appliqués. Or, une solution dite de *hopping* en V_{bb} est aujourd'hui difficilement réalisable car les temps de chargement et déchargement des capacités ne sont pas compatibles avec la dynamique nécessaire (centaines de ns). Nous allons donc maintenant discuter de la politique de gestion du body bias *hopping* que l'on peut réaliser afin de proposer une implémentation plus réaliste du mode d'exécution *hopping*.

5.2.3 Optimisation du hopping : politique de gestion du V_{bb} hopping

Étant donné les difficultés que représente une exécution en *hopping* en V_{bb} dans un contexte où la tâche doit être exécutée rapidement (quelques secondes), une solution sous-optimale de la gestion de l'exécution de la tâche a été proposée dans la section 4.4.4 du chapitre 4. Cette solution consiste à appliquer 2 PM avec une valeur de tension de body bias constante (*e.g.* V_{bb_x}) pour atteindre une fréquence cible en Mode 2. Un mécanisme de *hopping* pour V_{dd} , de $V_{dd_{i+1}}$ à V_{dd_i} , et pour F , de F_B à F_A , peut être appliqué afin d'atteindre en moyenne F_{target} , voir la Figure 5.18. Cette solution permet de s'affranchir de l'intégration d'un

mécanisme de *hopping* en V_{bb} . Elle peut être utilisée pour prendre en compte les contraintes d'implémentation physique des actionneurs (par exemple la surface, la consommation). Nous allons comparer la consommation de cette proposition à celle résultant de l'application d'un PM correspondant à F_{target} et n'appartenant pas au PWCS. Cela revient à évaluer l'apport d'une exécution en appliquant cette proposition sous-optimale par rapport à une exécution avec un PM non-optimal (*i.e.* n'appartenant pas au PWCS).

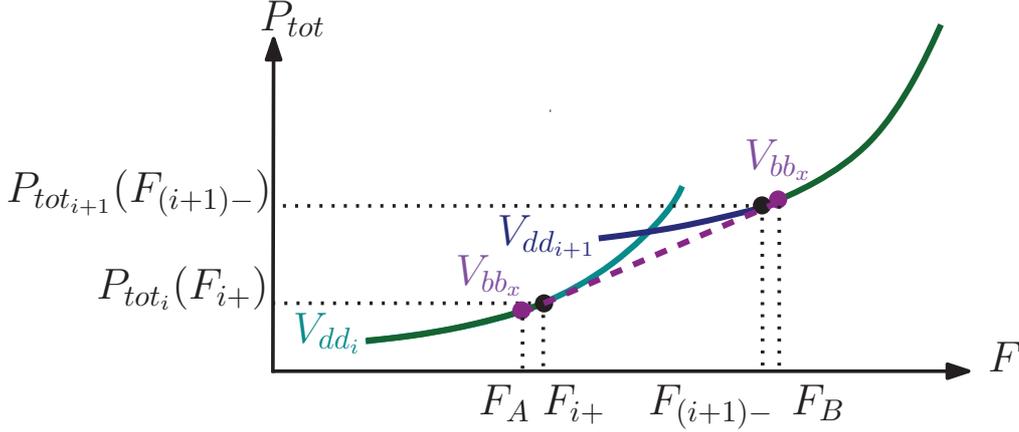


FIGURE 5.18: Mode 2 : Exécution en mode *hopping* avec un *hopping* en V_{dd} et en F , avec une tension de body bias constante V_{bb_x} .

Le circuit considéré est Frisbee [74], voir la section 5.1.2. Nous allons évaluer l'apport de la méthode sous-optimale dans le cas où 4 niveaux de tension d'alimentation sont disponibles sur le circuit. Nous allons considérer les mêmes tensions d'alimentations qui précèdent, *i.e.* $V_{dd_1} = 0,7$ V, $V_{dd_2} = 0,9$ V, $V_{dd_3} = 1,1$ V et $V_{dd_4} = 1,3$ V. La tension de body bias appartient à l'intervalle $[0 ; 1500]$ mV. Le gain en consommation est donné par :

$$\begin{aligned}
 G_{so}(F_{target})(\%) &= 100 \cdot \frac{P_{tot_{th}}(F_{target}) - P_{tot}(F_A, F_B, V_{bb_x})}{P_{tot_{th}}(F_{target})} \\
 &= 100 \cdot \frac{P_{tot_{th}}(F_{target}) - (\alpha \cdot P_{tot_{i+1}}(F_B, V_{bb_x}) + (1-\alpha) \cdot P_{tot_i}(F_A, V_{bb_x}))}{P_{tot_{th}}(F_{target})}
 \end{aligned} \tag{5.5}$$

avec $G_{so}(F_{target})$ le gain en consommation pour la fréquence cible F_{target} en pourcentage, $P_{tot_{th}}(F_{target})$ la consommation sans phase d'identification pour la fréquence cible F_{target} atteignable en Mode 2, $P_{tot}(F_A, F_B, V_{bb_x})$ la consommation avec la solution sous-optimale pour la fréquence cible F_{target} .

Notons que le coût de commutation des actionneurs n'a pas été pris en compte dans le calcul du gain en consommation en utilisant la solution sous-optimale proposée. Cependant, on peut supposer que pour la solution sous-optimale, il est inférieure au coût de commutation des actionneurs pour le Mode 2. En effet, le Mode 2 implique que les 3 sorties des actionneurs doivent être modifiées avec un

mécanisme dit de *hopping* alors que la solution sous-optimale consiste à modifier seulement 2 sorties des actionneurs avec un mécanisme dit de *hopping* ; les valeurs pour l'entrée V_{dd} est identique au Mode 2, la variation sur l'entrée F sera plus grande comparée au Mode 2 mais on garde une entrée constante pour V_{bb} .

Application de la solution sous-optimale : 3 pics de gains

La Figure 5.19 illustre le gain en consommation en fonction de la valeur de body bias V_{bb_x} constante appliquée pour exécuter les tâches dont la fréquence cible correspond aux 3 gains maximaux de la Figure 5.16. Ces F_{target} ont été choisies arbitrairement. Les tâches sont exécutées en appliquant 2 PM avec une tension de body bias constante V_{bb_x} . Le gain G_{so} en exécutant une tâche avec 2 PM appartenant au PWCS avec V_{bb_x} a été évalué et est tracé en ligne continue sur la Figure 5.19. G_{so} en exécutant une tâche avec 2 PM dont 1 PM appartient au PWCS est tracé en ligne mixte pour V_{bb_x} sur la Figure 5.19. Si la tâche est exécutée avec 2 PM dont aucun (0 PM) appartient au PWCS avec V_{bb_x} constant, G_{so} est tracé en ligne pointillée sur la Figure 5.19.

Pour $F_{target} = 1875$ MHz (resp. 2272 MHz), si la valeur de la tension de body bias V_{bb_x} appliquée est inférieure 996 mV (resp. 819 mV), alors les deux PM appliqués pour exécuter la tâche, n'appartiennent pas au PWCS. Le Tableau 5.12 liste le nombre de PM appartenant au PWCS selon la gamme de fréquence cible F_{target} , la gamme de tension de body bias et les tensions d'alimentation V_{dd_i} et $V_{dd_{i+1}}$ appliquées pour exécuter la tâche avec la solution sous-optimale.

Pour chaque F_{target} , le Tableau 5.13 (resp. le Tableau 5.14) affiche le gain G obtenu en appliquant le Mode 2, le gain minimum $G_{so_{min}}$ (resp. maximum $G_{so_{max}}$), la valeur de la tension de body bias constante V_{bb_x} associée à $G_{so_{min}}$ (resp. $G_{so_{max}}$) et le nombre PM appliqué qui appartiennent au PWCS. Si la valeur de gain est négative, la valeur absolue du gain correspond dans ce cas à une perte. Pour $F_{target} = 1357$ MHz, le gain maximal $G_{so_{max}}$ obtenu avec la méthode sous-optimale est de 15,58% pour $V_{bb_x} = 1500$ mV. Le gain G en appliquant le Mode 2 pour la même F_{target} est de 16,61%, voir le Tableau 5.14.

Nous pouvons déduire de la Figure 5.19, des Tableaux 5.13 et 5.12 qu'exécuter une tâche avec 0 PM appartenant au PWCS avec une tension V_{bb_x} constante peut mener à des pertes en consommation par rapport à appliquer directement la fréquence F_{target} dont le PM correspondant n'appartient pas au PWCS. Par exemple, le PM obtenu en appliquant le couple ($V_{dd_i} = 1,1$ V , $V_{bb_x} = 0$ mV) n'appartient pas au PWCS. Le PM obtenu en appliquant ($V_{dd_i} = 1,3$ V , $V_{bb_x} = 0$ mV) n'appartient pas non plus au PWCS. Si l'on souhaite atteindre une fréquence cible de 2272 MHz en appliquant les 2 PM cités dont la tension de body bias est égale à 0 mV, la perte de consommation en appliquant ces deux PM comparée à appliquer directement F_{target} est de 1,2%.

De plus, nous pouvons tirer de la Figure 5.19, des Tableaux 5.13 et 5.12, que le

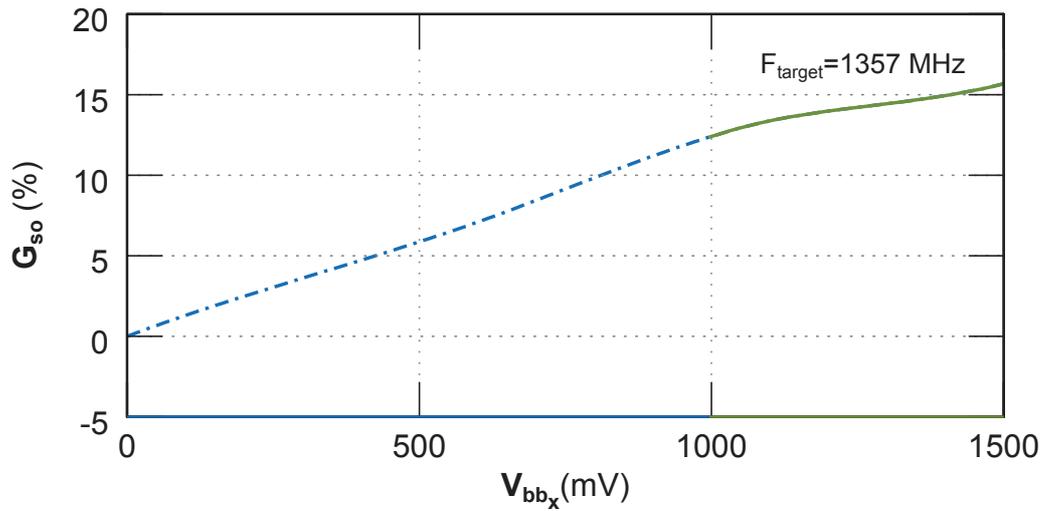
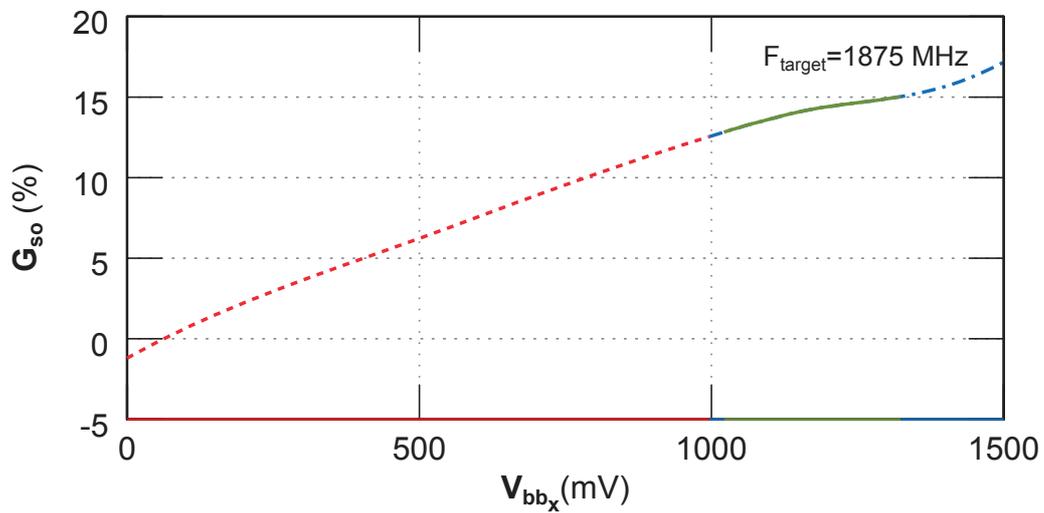
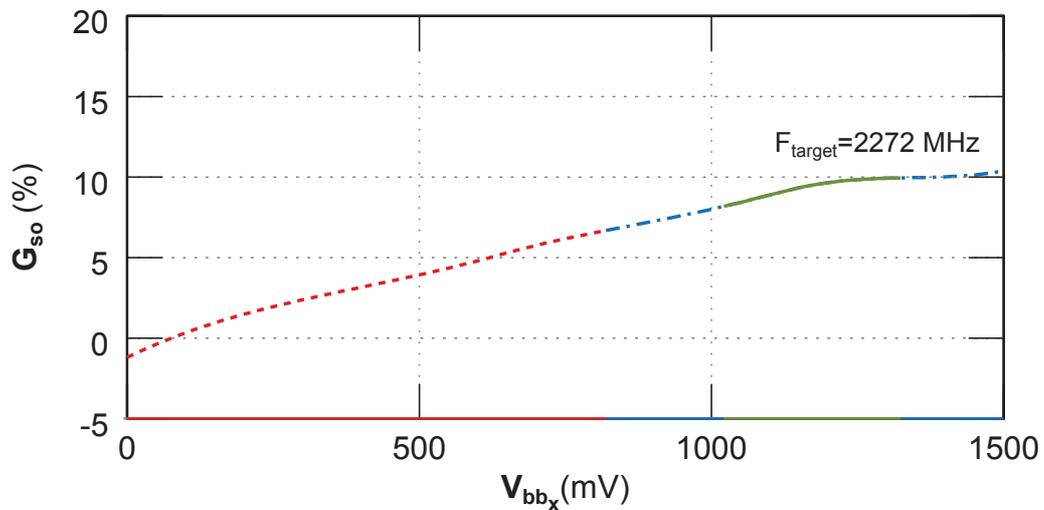
(a) $F_{target} = 1357$ MHz(b) $F_{target} = 1875$ MHz(c) $F_{target} = 2272$ MHz

FIGURE 5.19: Gain en consommation G_{so} en exécutant avec 2 PM (en ligne continue), 1 PM (en ligne mixte) et 0 PM (en ligne pointillée) appartenant au PWCS avec V_{bb_x} constant pour les 3 gains maximaux de la Figure 5.16.

TABLEAU 5.12: Nombre de PM appartenant au PWCS en exécutant avec les tensions V_{dd_i} et $V_{dd_{i+1}}$ pour la gamme V_{bb_x} et la gamme F_{target} .

Gamme F_{target}	Gamme V_{bb_x} (mV)	V_{dd_i}	$V_{dd_{i+1}}$	Nombre PM ∈ PWCS?
$]F_{1+} ; F_{2-}[$	$[0 ; 996[$	V_{dd_1}	V_{dd_2}	1
$]F_{1+} ; F_{2-}[$	$[996 ; 1500]$	V_{dd_1}	V_{dd_2}	2
$]F_{2+} ; F_{3-}[$	$[0 ; 996[$	V_{dd_2}	V_{dd_3}	0
$]F_{2+} ; F_{3-}[$	$[996 ; 1021[$	V_{dd_2}	V_{dd_3}	1
$]F_{2+} ; F_{3-}[$	$[1021 ; 1324]$	V_{dd_2}	V_{dd_3}	2
$]F_{2+} ; F_{3-}[$	$]1324 ; 1500]$	V_{dd_2}	V_{dd_3}	1
$]F_{3+} ; F_{4-}[$	$[0 ; 819[$	V_{dd_3}	V_{dd_4}	0
$]F_{3+} ; F_{4-}[$	$[819 ; 1021[$	V_{dd_3}	V_{dd_4}	1
$]F_{3+} ; F_{4-}[$	$[1021 ; 1324]$	V_{dd_3}	V_{dd_4}	2
$]F_{3+} ; F_{4-}[$	$]1324 ; 1500]$	V_{dd_3}	V_{dd_4}	1

TABLEAU 5.13: Gain G en exécutant avec le Mode 2 et gain minimal $G_{so_{min}}$ en exécutant avec 2 PM quelconques afin d'atteindre F_{target} en moyenne pour V_{bb_x} constant pour les 3 gains maximaux de la Figure 5.16.

F_{target} (MHz)	G en Mode 2 (%)	V_{bb_x} (mV)	$G_{so_{min}}$ (%)	Nombre de PM ∈ PWCS
1357	16,61	0	0,00	1
1875	17,31	0	-1,20	0
2272	11,48	0	-1,20	0

TABLEAU 5.14: Gain G en exécutant avec le Mode 2 et gain maximal $G_{so_{max}}$ en exécutant avec 2 PM quelconques afin d'atteindre F_{target} en moyenne pour V_{bb_x} constant pour les 3 gains maximaux de la Figure 5.16.

F_{target} (MHz)	G en Mode 2 (%)	V_{bb_x} (mV)	$G_{so_{max}}$ (%)	Nombre de PM ∈ PWCS
1357	16,61	1500	15,68	2
1875	17,31	1500	17,18	1
2272	11,48	1500	10,39	1

gain en consommation G_{so} en appliquant 0 PM appartenant au PWCS peut être inférieur au gain obtenu en exécutant la même tâche avec 2 PM dont au moins un appartient au PWCS. Pour les fréquences comprises entre $]F_{2+} ; F_{3-}[$, si l'on applique deux PM avec une tension de body bias V_{bb_x} constante dont la valeur

est comprise dans la gamme $[0mV ; 996mV[$, alors ces deux PM n'appartiennent pas au PWCS. De même, pour les fréquences comprises entre $]F_{3+} ; F_{4-}[$, si l'on applique deux PM avec une tension de body bias constante V_{bb_x} dont la valeur est comprise dans la gamme $[0mV ; 819mV[$, alors les deux PM appliquées n'appartiennent pas au PWCS. Pour $F_{target} = 1875$ MHz, si V_{bb_x} est inférieure à 996 mV, le gain en consommation pour cette valeur de tension de body bias est inférieur au gain en consommation pour V_{bb_x} supérieure à 996 mV pour cette fréquence F_{target} .

De plus, nous remarquons que les gains en consommation G_{so} en appliquant 2 PM quelconques ayant une même valeur de tension de body bias sont inférieurs par rapport aux gains en consommation G obtenus en exécutant avec le Mode 2 pour une même F_{target} . Par exemple, pour F_{target} égale à 1357 MHz, le gain G en exécutant avec le Mode 2 est de 16,61% alors que le gain maximal $G_{so_{max}}$ obtenu pour cette même fréquence est de 15,68% avec cette proposition sous-optimale, voir le Tableau 5.14.

Nous observons sur la Figure 5.19 (voir le Tableau 5.13) que le gain minimal $G_{so_{min}}$ pour les 3 F_{target} correspond à l'application de $V_{bb_x} = 0$ mV.

Par ailleurs, le gain maximal $G_{so_{max}}$ n'est pas obtenu en appliquant 2 PM appartenant au PWCS, comme on pourrait s'y attendre. $G_{so_{max}}$ est obtenu en appliquant au moins un PM appartenant au PWCS. Ce comportement est observable sur la Figure 5.19 et le Tableau 5.14. Par exemple, pour la fréquence cible F_{target} de 2272 MHz, le gain maximal est pour la tension de body bias V_{bb_x} de 1500 mV et un des PM appliqué appartient au PWCS, voir le Tableau 5.14.

En observant la Figure 5.19, nous aurions tendance à conjecturer que le gain maximal est atteint quelque soit la fréquence cible pour la valeur de tension de body bias maximale soit 1500 mV dans ce cas. Cependant, ce cas particulier ne reflète pas la réalité. Nous allons évaluer la solution sous-optimale pour différentes fréquences F_{target} .

Application de la solution sous-optimale : F_{target} quelconques

La Figure 5.20 représente le même tracé que sur la Figure 5.19 pour différentes F_{target} . Le gain G_{so} en exécutant une tâche avec 2 PM (ligne continue), 1 PM (ligne mixte) et 0 PM (ligne pointillée) appartenant au PWCS avec V_{bb_x} constant a été évalué. Il est disponible sur la Figure 5.19. Pour chaque F_{target} , le Tableau 5.15 (resp. le Tableau 5.16) contient le gain G obtenu en appliquant le Mode 2, le gain minimum $G_{so_{min}}$ (resp. maximum $G_{so_{max}}$), la valeur de la tension de body bias constante V_{bb_x} associée à $G_{so_{min}}$ (resp. $G_{so_{max}}$) et le nombre de PM appliqués appartenant au PWCS.

Nous pouvons observer sur la Figure 5.20 que les gains maximaux pour les 3 F_{target} sont atteints en appliquant une tension V_{bb_x} différente de 1500 mV. Ainsi,

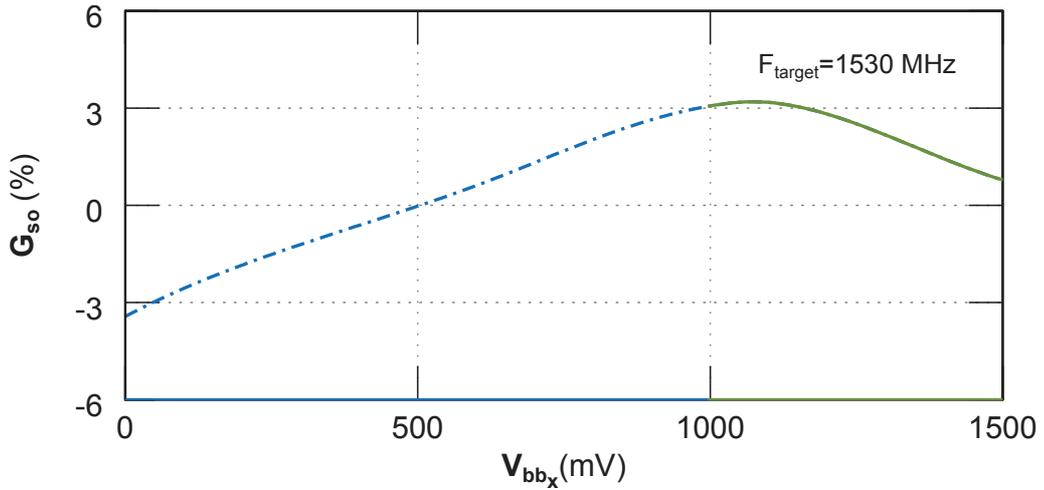
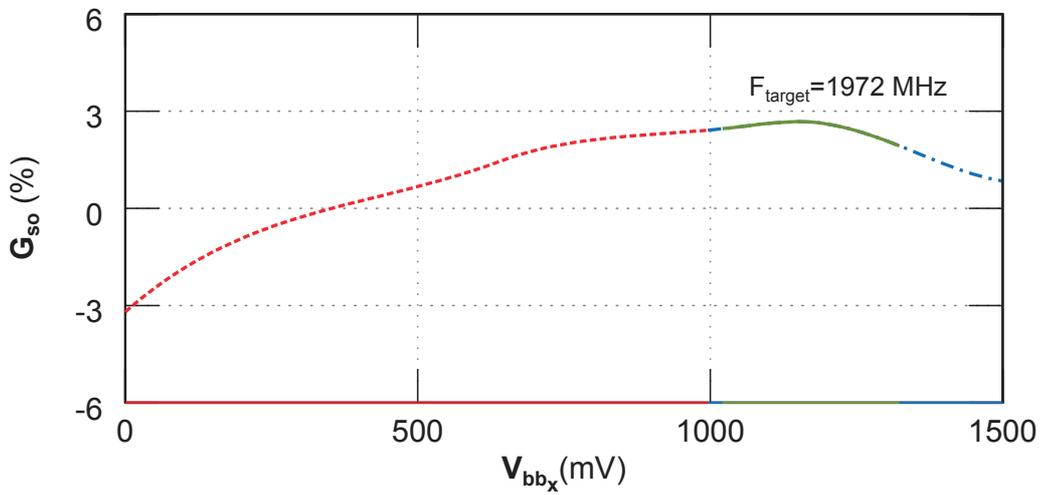
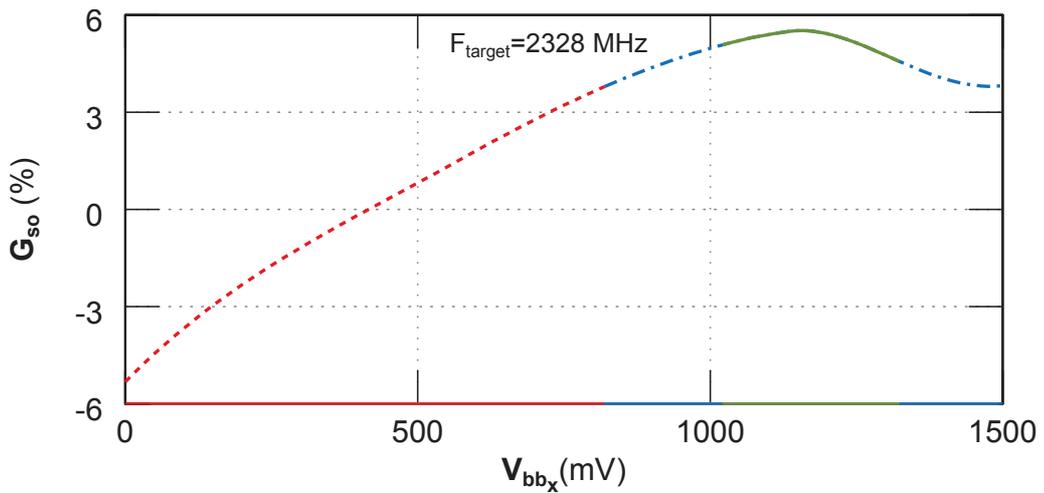

 (a) $F_{target} = 1530$ MHz

 (b) $F_{target} = 1972$ MHz

 (c) $F_{target} = 2328$ MHz

FIGURE 5.20: Gain en consommation G_{so} en exécutant avec 2 PM (en ligne continue), 1 PM (en ligne mixte) et 0 PM (en ligne pointillée) appartenant au PWCS avec V_{bb_x} constant pour 3 F_{target} choisies arbitrairement.

le gain maximal avec la solution sous-optimale n'est pas toujours atteignable en appliquant une tension V_{bb_x} de 1500 mV.

Nous remarquons que les gains maximaux sont atteints en appliquant 2 PM appartenant au PWCS, voir la Figure 5.20.

Sur la Figure 5.20, nous pouvons remarquer pour la fréquence cible de 1972 MHz que le gain en consommation en appliquant 0 PM appartenant au PWCS est parfois supérieur au gain obtenu en exécutant la même tâche avec 2 PM dont au moins un appartient au PWCS.

TABLEAU 5.15: Gain G en exécutant avec le Mode 2 et gain minimal $G_{so_{min}}$ en exécutant avec 2 PM quelconques afin d'atteindre F_{target} en moyenne pour V_{bb_x} constant pour 3 F_{target} choisies arbitrairement.

F_{target} (MHz)	G Mode 2 (%)	V_{bb_x} (mV)	$G_{so_{min}}$ min. (%)	Nombre de PM \in PWCS
1530	4,90	0	-3,43	1
1972	7,76	0	-5,32	0
2328	5,05	0	-3,20	0

TABLEAU 5.16: Gain G en exécutant avec le Mode 2 et gain maximal $G_{so_{max}}$ en exécutant avec 2 PM quelconques afin d'atteindre F_{target} en moyenne pour V_{bb_x} constant pour 3 F_{target} choisies arbitrairement.

F_{target} (MHz)	G en Mode 2 (%)	V_{bb_x} (mV)	$G_{so_{max}}$ max. (%)	Nombre de PM \in PWCS
1530	4,90	1084	3,19	2
1972	7,76	1160	5,52	2
2328	5,05	1147	2,68	2

Discussion

Afin, de maximiser le gain quelque soit F_{target} , il faudrait modéliser la consommation avec la solution sous-optimale. Ainsi, une solution analytique peut être trouvée quelque soit F_{target} . Cependant, comme nous n'avons pas modélisé la consommation de la solution sous-optimale, nous ne pouvons donner une solution pour trouver la valeur de la tension de body bias V_{bb_x} qui maximise le gain en consommation G_{so} avec cette solution sous-optimale. Nous pouvons juste proposer de choisir V_{bb_x} de manière à avoir au moins un des PM correspondant à V_{bb_x} qui appartient au PWCS. Néanmoins, cette proposition ne mène pas toujours à un gain supérieur par rapport à appliquer 0 PM appartenant au PWCS. Cependant,

nous avons constaté que les gains maximaux $G_{so_{max}}$ obtenus pour les différentes fréquences correspondaient toujours à exécuter la tâche avec 2 PM dont au moins un appartient au PWCS. Une autre solution serait de stocker pour chaque F_{target} la valeur de V_{bb_x} qui permet de maximiser le gain pour chaque gamme de température T° et d'activité A pour la solution sous-optimale. Mais cette solution semble être peu réaliste car coûteuse en surface (mémoire).

Jusqu'ici, nous n'avons pas pris en compte le coût de commutation des actionneurs dans le calcul des gains. Nous allons maintenant proposer une façon de prendre en compte le coût de commutation des actionneurs lorsque le Mode 2 est appliqué.

5.2.4 Prise en compte du coût de commutation des actionneurs

Nous avons évalué la consommation qui pouvait être économisée si la méthode proposée est appliquée pour le cas où deux actionneurs continus sont combinés à un actionneur discret. Nous n'avons pas pris en compte dans le calcul du gain en consommation le coût de commutation des actionneurs. Le coût de commutation dépend des actionneurs choisis, c'est-à-dire qu'il dépend de la dynamique (vitesse de commutation) et de la consommation des actionneurs.

Nous allons reprendre comme cas d'étude le circuit Frisbee [74], voir la section 5.1.2. Nous allons proposer une manière permettant de choisir si le Mode 1 ou le Mode 2 est le plus économique en énergie en prenant en compte le coût de commutation des actionneurs. Nous allons reprendre le dernier exemple avec 4 niveaux de tension d'alimentation. Rappelons que les tensions d'alimentations sont $V_{dd_1} = 0,7$ V, $V_{dd_2} = 0,9$ V, $V_{dd_3} = 1,1$ V et $V_{dd_4} = 1,3$ V. La tension de body bias appartient toujours à l'intervalle $[0 ; 1500]$ mV.

D'après [90], l'efficacité énergétique d'un mécanisme de V_{dd} -Hopping couplé à un actionneur de fréquence est de 95%. Le coût de commutation des actionneurs V_{dd} et F correspond donc environ à 5% du gain attendu avec la technique de AVFS.

Les PM dont le gain est inférieur au coût de commutation seront identifiés afin d'être intégrés dans un nouveau sous-ensemble que l'on va nommer PN pour *New Piece*. On définit PN comme le sous-ensemble des PM n'appartenant pas au PWCS et dont le gain en Mode 2 est inférieur au coût de commutation. On applique donc pour le sous-ensemble PN le Mode 1. Nous pouvons modéliser le coût de commutation de 2 façons : soit par une fonction de F ou par une constante dans le plan (F, G) .

Coût de commutation constant

Prenons pour exemple un coût de commutation constant de 5%. La Figure 5.21 illustre le gain en consommation en fonction de la fréquence et un coût de commu-

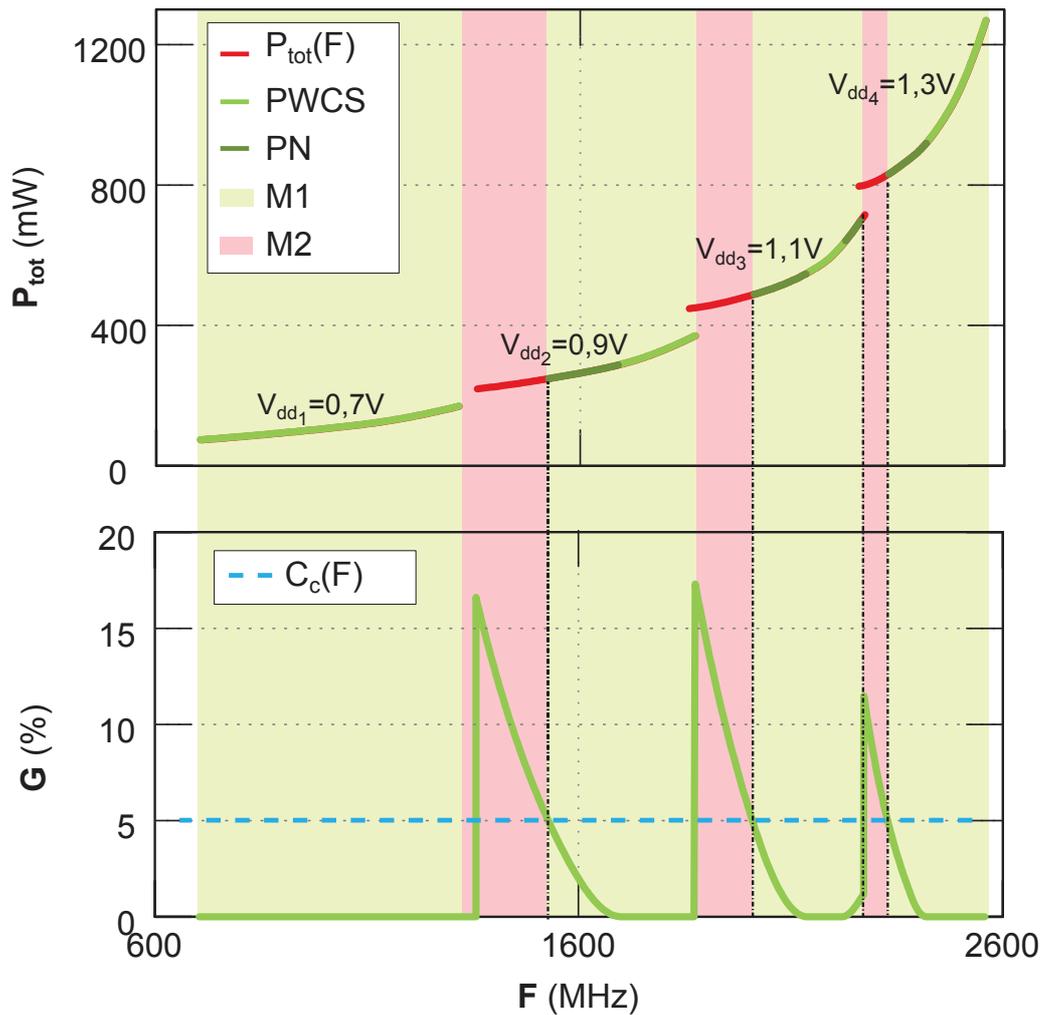


FIGURE 5.21: Consommation P_{tot} en fonction de la fréquence F avec les sous-ensembles PWCS et PN (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 4 niveaux de V_{dd} pour un coût de commutation constant (en bas).

tation des actionneurs constants (en pointillé). Les PM dont le gain en consommation est inférieur ou égal à 5% sont en vert foncé sur la Figure 5.21. On identifie un nouveau sous-ensemble PN de PM. En effet, si l'on prend en compte le coût de commutation pour les PM appartenant au nouveau sous-ensemble PN, il vaut mieux appliquer ces PM directement au lieu d'appliquer le Mode 2. Pour ces PM et seulement ceux-là, la méthode est modifiée. En effet, si F_{target} appartient à PN, on applique directement le PM pour exécuter une tâche au lieu d'exécuter en Mode 2 (voir la Figure 5.21). Pour les autres PM, la méthode reste inchangée; on exécute toujours les tâches avec le Mode 1 ou le Mode 2. Lorsque le Mode 2 est appliqué, il faut exécuter la tâche avec les 2 PM les plus proches de la fréquence

cible qui appartient au PWCS initial, c'est-à-dire celui qui a été identifié sans prendre en compte le coût de commutation des actionneurs.

Coût de commutation non constant

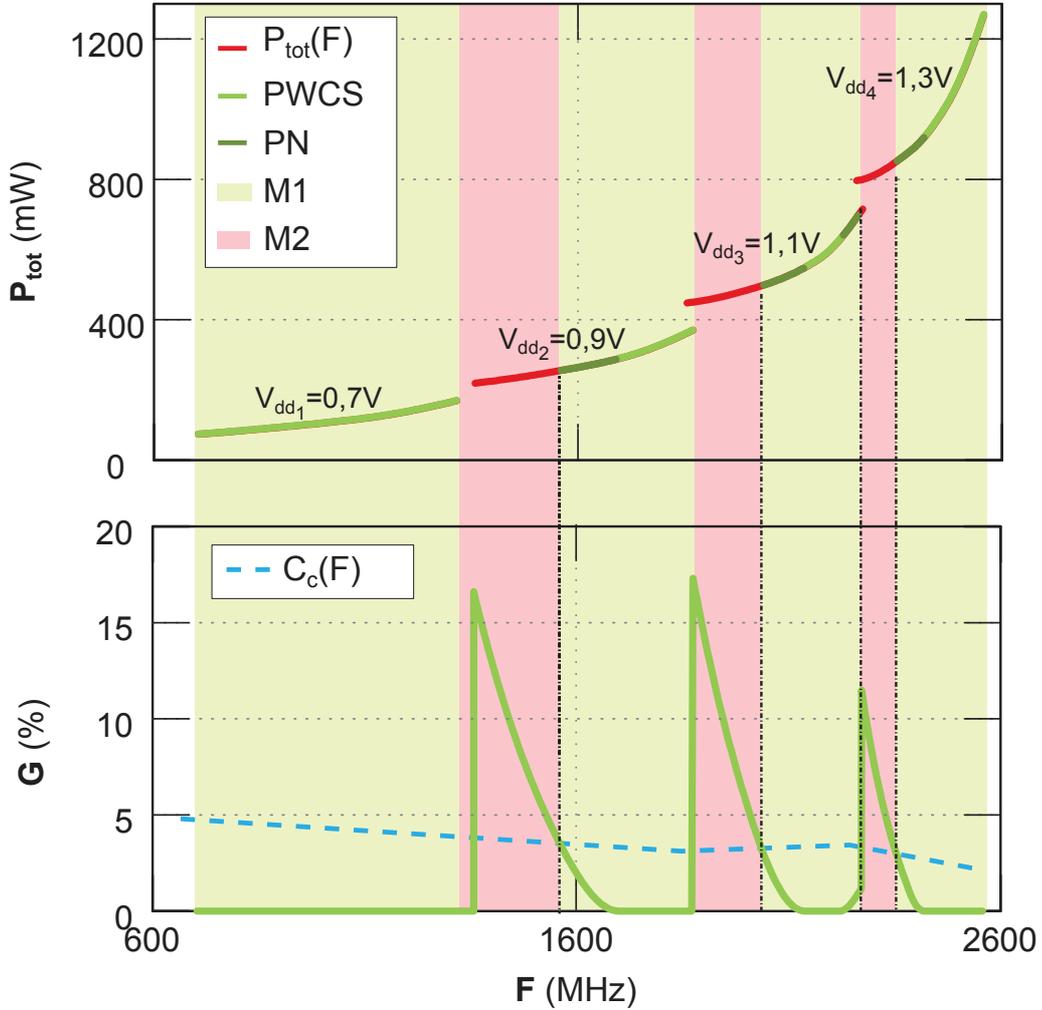


FIGURE 5.22: Consommation P_{tot} en fonction de la fréquence F avec les sous-ensembles $PWCS$ et PN (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 4 niveaux de V_{dd} pour un coût de commutation non constant (en bas).

La Figure 5.22 illustre le gain en consommation en fonction de la fréquence et le coût de commutation C_c en fonction de la fréquence non constant en pointillé. Le profil de $C_c(F)$ a été choisi arbitrairement. Il permet de donner une idée pour prendre en compte un coût de commutation non constant. Nous avons identifié le sous-ensemble PN de PM . Si F_{target} correspond à un PM appartenant à PN (ensemble des PM en vert foncé), le PM est appliqué afin d'atteindre F_{target} . Le

gain en Mode 2 pour les F_{target} correspondant au PM appartenant au PN est inférieur au coût de commutation des actionneurs.

Nous avons vu une manière de prendre en compte le coût de commutation des actionneurs. Nous allons maintenant voir l'impact des variations de procédés de fabrication sur la méthode proposée.

5.3 Variations de procédés de fabrication

Nous souhaitons évaluer quel est l'effet des variations de procédés de fabrication sur la méthode. Pour cela, nous étudierons l'impact des variations sur le résultat de l'identification du PWCS. Nous évaluerons l'impact des variations de procédés de fabrication sur le gain en consommation. Nous disposons d'un ensemble de 16 puces du même circuit provenant de 2 wafers différents et d'un même lot de fabrication.

Nous allons reprendre comme cas d'étude le circuit Frisbee [74], voir la section 5.1.2. Rappelons que les tensions d'alimentation sont $V_{dd_1} = 0,7$ V, $V_{dd_2} = 0,9$ V, $V_{dd_3} = 1,1$ V et $V_{dd_4} = 1,3$ V et la tension de body bias $V_{bb} \in [0; 1500]$ mV. Pour évaluer l'impact de variations de procédés de fabrication, nous disposons de mesures de consommation sur 16 puces du circuit Frisbee issues d'un même lot de fabrication. Pour notre évaluation, nous allons tout d'abord identifier le PWCS pour toutes les puces. Nous allons identifier les limites basses et hautes du PWCS pour chaque tension d'alimentation et le gain maximal obtenu avec la méthode proposée pour chaque puce.

La Figure 5.23 illustre le profil de consommation en fonction de la fréquence pour différentes puces. Pour chaque profil (F, P_{tot}), le PWCS est identifié en vert. Nous pouvons observer que les variations de procédés de fabrication ont un impact sur la fréquence maximale atteignable pour une tension d'alimentation V_{dd} et une tension de body bias V_{bb} fixées. Nous pouvons remarquer que les variations de procédés de fabrication jouent aussi sur la consommation liée au triplet (F, V_{dd}, V_{bb}). Par exemple, la consommation de la puce A est 6,6% supérieure à celle de la meilleure puce alors que le même triplet (1314 MHz, 0,7 V, 1500 mV) est appliquée. Rappelons que la fréquence appliquée pour chaque tension d'alimentation V_{dd} et tension de body bias V_{bb} correspond à la fréquence maximale atteignable pour le couple (V_{dd}, V_{bb}). De plus, pour une tension fixée et pour la même gamme de tension de body soit $[0 ; 1500]$ mV, nous remarquons que la gamme de fréquence atteignable pour la meilleure puce est plus grande que la gamme de fréquence atteignable pour les deux autres puces (puce A et puce B). Par ailleurs, pour la meilleure puce, nous pouvons atteindre une fréquence plus importante en général pour une même tension d'alimentation V_{dd} et une même tension de body bias V_{bb} tout en réduisant la consommation.

La Figure 5.24 illustre le profil de la consommation en fonction de la fréquence

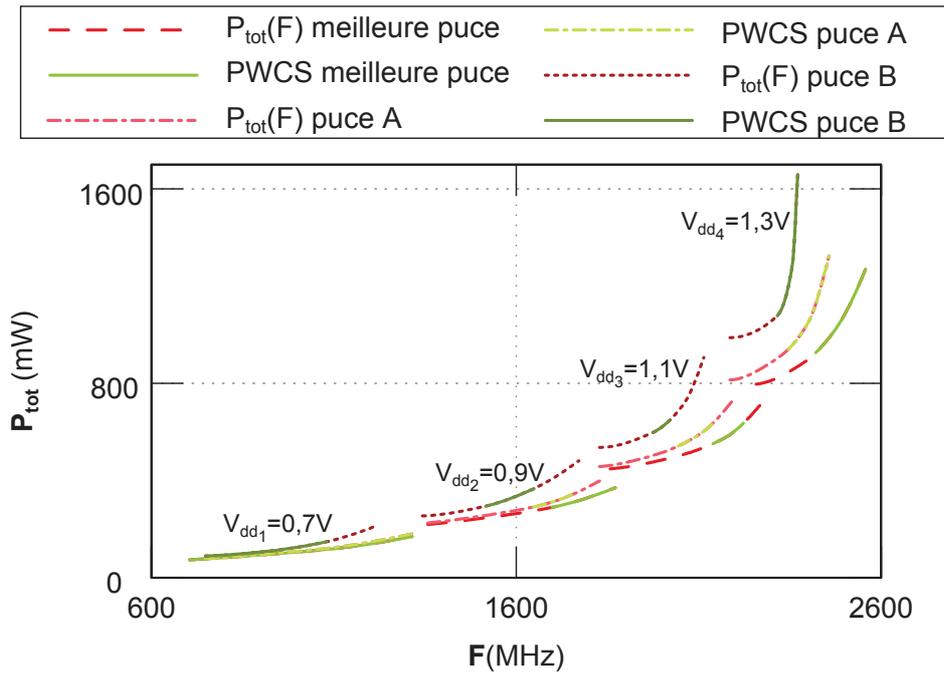


FIGURE 5.23: Ensemble de PM dans le plan (F, P_{tot}) pour 4 niveaux de V_{dd} et 3 puces différentes avec les PM appartenant au PWCS en vert.

pour les 16 puces disponibles. Le PWCS pour chaque profil est identifié en vert. Nous pouvons aussi apercevoir que les limites basses et les limites hautes du PWCS sont fortement dépendantes de la courbure du profil de la consommation en fonction la fréquence et donc des procédés de fabrication. Cependant, nous allons voir que l'on ne peut prévoir comment les limites vont évoluer en fonction des variations de procédés de fabrication.

Nous avons identifié les limites basses et hautes du PWCS pour chaque tension d'alimentation V_{dd} et pour chaque puce. Ces limites sont disponibles sur la Figure 5.25 pour différentes puces. La limite basse (resp. haute) pour la tension d'alimentation V_{dd_i} est notée F_{i-} (resp. F_{i+}). Nous observons que la distribution des limites est éparpillée sur le plan (F, P_{tot}) .

Le gain en consommation en Mode 2 dépend du profil de consommation et ainsi des variations de procédés de fabrication. Les pics de gain et les gains maximaux G_{max} pour chaque puce sont tracés sur la Figure 5.26. Le plus grand gain maximal obtenue en Mode 2 pour toutes les puces disponibles est de 26,66%. Le plus petit gain maximal obtenue en Mode 2 pour toutes les puces disponibles est de 12,55%. On remarque que le gain maximal ne correspond pas toujours au même pic de gain. Le pic de gain se situe parfois entre F_{2-} et F_{3+} et parfois entre F_{3-} et F_{4+} .

Nous pouvons conclure que la méthode proposée permet de gagner en consom-

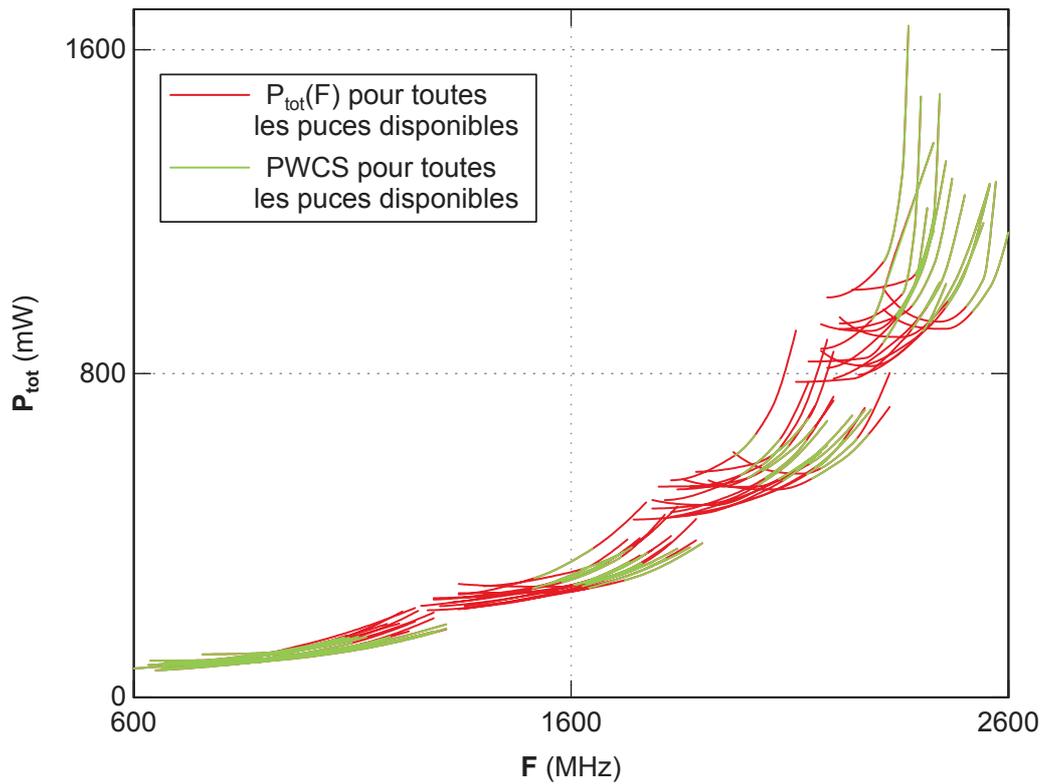


FIGURE 5.24: Ensemble de PM dans le plan (F, P_{tot}) pour 4 niveaux de V_{dd} pour 16 puces avec les PM appartenant au PWCS en vert.

mation. Le plus petit gain maximal est de 12,55% pour une fréquence cible. Nous pouvons attester de l'efficacité de la méthode proposée sur un circuit composé de deux actionneurs continus et un actionneur discret avec une large gamme de fréquence.

Nous avons évalué l'impact des variations de procédés de fabrication sur la méthode proposée. Nous avons vu que les limites basses et hautes, le gain maximal pour une puce dépend du procédé de fabrication. Cependant, étant donné qu'il ne peut être quantifié, nous ne pouvons modéliser les valeurs limites en fonction des variations de procédés de fabrication. Pour palier ce problème, nous proposons de calibrer la méthode pour chaque circuit.

5.4 Conclusion

Dans ce chapitre, nous avons présenté plusieurs exemples d'évaluation des gains de consommation des méthodes proposées dans le chapitre 4. Le premier exemple portait sur une méthode considérant un nombre fini de PM. Nous avons proposé d'extraire le DCS et d'exécuter les tâches avec les PM appartenant au DCS. Pour l'exemple de Frisbee, la méthode proposée permet de réduire la

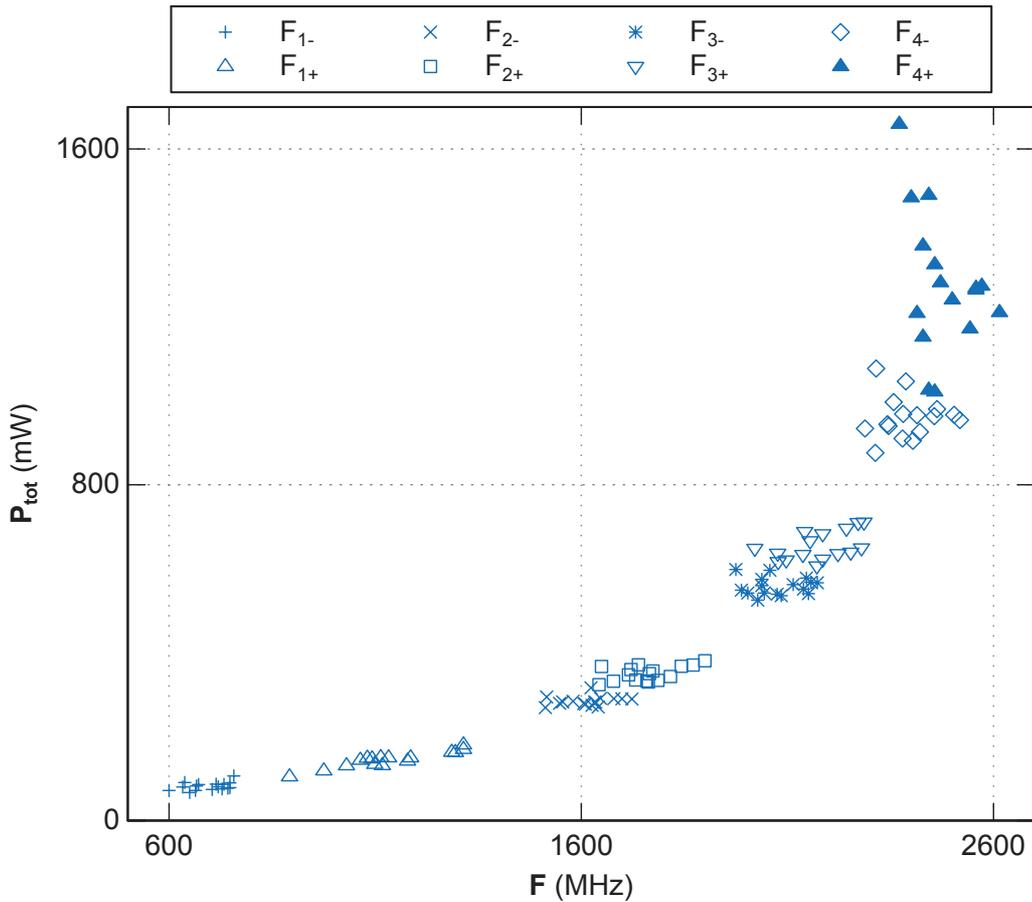


FIGURE 5.25: Ensemble des limites basses et hautes du PWCS dans le plan (F, P_{tot}) pour 4 niveaux de V_{dd} et 16 puces.

consommation jusqu'à 16,45% en exécutant les tâches seulement avec les PM appartenant au DCS.

Le second exemple reposait sur des mesures extraites d'un circuit réel composé d'un nombre infini de PM. Nous avons proposé d'extraire le PWCS et d'exécuter les tâches avec les PM appartenant au PWCS. Nous avons vu trois exemples d'applications avec différents nombres et différents niveaux de tension d'alimentation V_{dd} choisies arbitrairement. Nous avons montré que la méthode proposée permet de réduire dynamiquement la consommation d'un circuit composé d'un nombre infini de PM. Le gain maximal obtenu en Mode 2 est pour l'exemple avec 2 niveaux de V_{dd} et est égale à 19,07%.

Nous avons proposé une solution sous-optimale permettant d'éviter le saut en V_{bb} lorsque la tâche est exécutée en Mode 2. Cette méthode sous-optimale permet de réduire la consommation de 17,18% au maximum, comparée à appliquer directement un PM qui n'appartient pas au PWCS. Néanmoins, cette solution sous-optimale peut conduire à des pertes de consommation si l'on applique deux

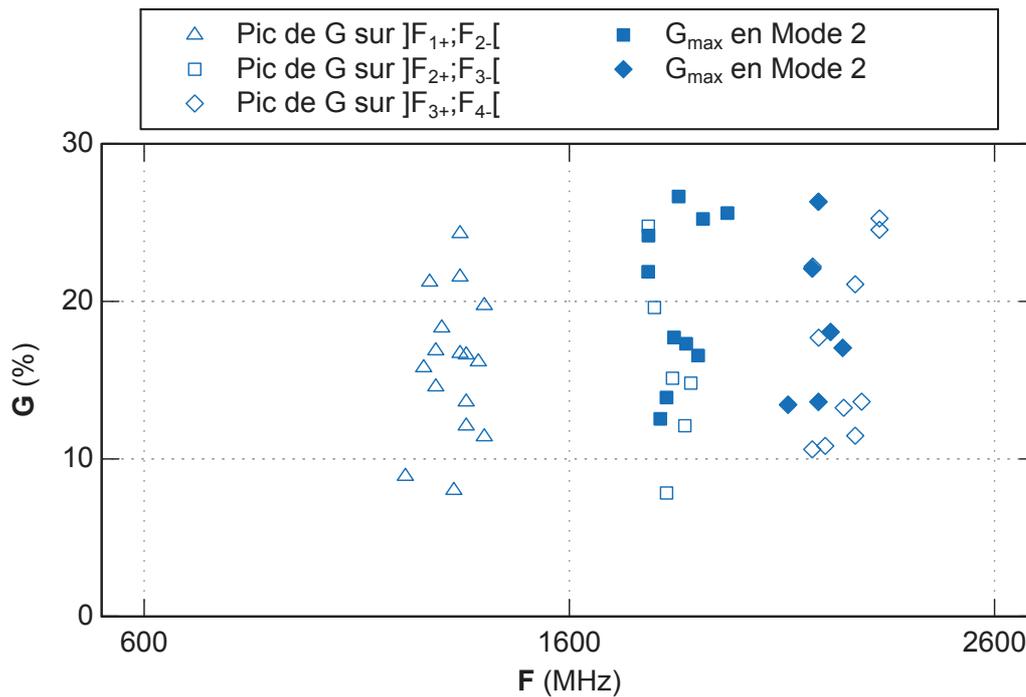


FIGURE 5.26: Pic de gain et gain maximal G_{max} pour chaque puce lorsque 4 niveaux de V_{dd} sont disponibles.

PM qui n'appartiennent pas au PWCS comparée à ne pas appliquer la méthode. Nous avons constaté que les gains maximaux obtenus avec cette proposition sous-optimale correspondaient à exécuter les tâches avec au moins un des deux PM appartenant au PWCS. C'est pourquoi, nous proposons pour cette solution sous-optimale, d'exécuter les tâches en appliquant au moins un PM appartenant au PWCS.

Nous avons également proposé une manière de prendre en compte le coût de commutation des actionneurs. Nous avons proposé d'identifier un nouveau sous-ensemble de PM PN qui prend en compte le coût de commutation. Pour les performances correspondant à un PM de PN , les tâches doivent être exécutées en appliquant directement le PM. En effet, lorsque le coût de commutation des actionneurs est pris en compte, pour ces PM, l'exécution en appliquant le Mode 2 est plus coûteuse en termes de consommation.

Enfin, nous avons étudié l'impact des variations de procédés de fabrication sur la méthode proposée. Nous avons observé que la méthode était fortement impactée par le procédés de fabrication. Nous avons déterminé les gains maximaux pour différentes puces de Frisbee. Le maximum de gain maximal pour toutes les puces disponibles est de 26,66%. Le minimum de gain maximal pour toutes les puces disponibles est de 12,55%.

Dans la suite, nous allons discuter des applications possibles de la méthode

proposée dans d'autres contexte puis conclure ce travail.

Chapitre 6

Bilan et perspectives

Synthèse

La loi de Moore prédit l'augmentation de la densité des transistors sur un circuit intégré. Les circuits électroniques sont de moins en moins coûteux et de plus en plus puissants. Cette évolution a intégré dans la vie de tous les jours beaucoup de dispositifs, le domaine de l'Internet des Objets (IoT) est un exemple. Cependant, la miniaturisation des transistors a atteint des limites physiques en CMOS BULK 28nm. En effet, à partir de ce nœud technologique, les circuits sont sensibles aux variations de procédés de fabrication, de tension et de température (PVT) et la consommation statique a augmenté d'une manière importante. La consommation devient donc le facteur limitant le développement de la performance des circuits intégrés. Deux approches se sont démarquées pour faire face à cette rupture technologique, le FinFET et le FD-SOI. Dans ces travaux, nous nous sommes en particulier intéressés à la seconde technologie. L'objectif de ces travaux a été de proposer des méthodes de gestion de la consommation en garantissant la performance du circuit.

La gestion de la consommation a été étudiée à différents niveaux. Dans ces travaux, le niveau d'abstraction est le niveau architecture/système, nous considérons un circuit composé de plusieurs domaines de consommation. La gestion de la consommation, en agissant juste sur la tension d'alimentation V_{dd} et la fréquence F , n'est pas toujours suffisante. Par ailleurs, si les systèmes sur puce (SoC) sont composés de plus en plus de cœurs, la décomposition du SoC en plusieurs domaines de consommation introduit le problème lié à la duplication des actionneurs dans chaque domaine de consommation. Les contraintes d'implémentation physique des actionneurs, c'est-à-dire le coût des actionneurs en termes de surface, de consommation et de rendement sont à prendre en compte. L'optimisation de la consommation en choisissant les points de fonctionnement ou PM (Power Mode)

avec 3 actionneurs n'est pas intuitive. Rappelons qu'un PM est défini comme un couple (F, P_{tot}) avec F la fréquence d'horloge et P_{tot} la consommation totale du circuit. Notons qu'une même fréquence peut être atteinte en appliquant différents couples (V_{dd}, V_{bb}) , avec V_{bb} la tension de body bias.

Les problèmes abordés dans ces travaux ont été tout d'abord d'évaluer l'apport de PM supplémentaires sur un circuit. Le gain en consommation atteignable en fonction du placement de PM a été estimé. Une solution permettant de placer les PM dans le plan $P_{tot}(F)$ afin de maximiser le gain en consommation énergétique a été proposée. La technologie FD-SOI offre l'opportunité d'agir sur la tension de body bias V_{bb} . Nous avons évalué l'apport de l'actionneur de V_{bb} . Par ailleurs, des méthodes de gestion de (F, V_{dd}, V_{bb}) de manière coordonnée, en cours d'exécution, ont été proposées. Elles permettent d'optimiser la consommation totale du circuit.

Dans le chapitre 2, nous avons présenté les différentes solutions technologiques qui ont permis de faire face aux problèmes de miniaturisation qu'a rencontrée la technologie CMOS BULK. Différentes solutions de gestion du compromis consommation-performance améliorant l'efficacité énergétique des circuits ont été exposées. Certaines de ces techniques nécessitent des technologies particulières, par exemple le *Body Biasing*. De plus, ces techniques nécessitent de pouvoir agir sur les paramètres *via* les actionneurs. Une synthèse des actionneurs agissant sur la fréquence F , la tension d'alimentation V_{dd} et la tension de body bias V_{bb} a été exposée. Le choix d'utiliser un actionneur de tension discret a été expliqué.

Le nombre de niveaux de tensions d'alimentation à implémenter dépend de la gamme de fréquence nécessaire et du profil d'exécution. Le convertisseur capacitif est l'un des rares générateurs de tension facilement intégrable. Néanmoins, ce type d'actionneur n'a un bon rendement ($> 80\%$) que sur une très petite gamme de tension d'alimentation. C'est pourquoi, nous nous sommes orientés vers un actionneur de tension discret. Cela permet d'intégrer des convertisseurs capacitifs de manière à maximiser le rendement pour les valeurs de tensions discrètes souhaitées. Si l'actionneur de tension n'a pas besoin d'être intégré au circuit alors dans ce cas notre choix s'orientera vers un convertisseur inductif qui possède un bon rendement ($> 90\%$) sur une large gamme de tension d'alimentation. Par ailleurs, l'évolution prévue par l'ITRS sur le nombre croissant de cœurs par SoC renforce le besoin de gestion de la consommation à grain fin. Ceci justifie le choix d'un actionneur de tension discret. La gestion de chaque cœur peut être réalisée avec un mécanisme de saut en tension (par exemple *V_{dd}-Hopping*) en dupliquant seulement l'interrupteur et la partie contrôle de l'actionneur dans chaque cœur.

Dans le chapitre 3, des méthodes de minimisation de la consommation pouvant être appliquées lors de la conception d'un circuit ou lors de sa calibration ont été proposées. Ces méthodes ont pour but de choisir le placement des PM du circuit afin de maximiser le gain en consommation. Nous avons proposé d'évaluer

le gain en consommation que l'ajout de PM supplémentaires apporte. Une solution analytique a été proposée. Cette solution a pour but d'approcher le profil de la consommation associé au mode *hopping* du profil idéal associé à l'utilisation d'actionneurs continus. Rappelons, qu'une exécution en mode *hopping* consiste à exécuter une partie de la tâche avec un PM, puis le reste du temps impartit un autre PM est appliqué. La solution analytique nécessite d'approximer le profil de consommation en fonction de la fréquence par un polynôme d'ordre 3. Un polynôme d'ordre 3 est en adéquation avec le profil $P_{tot}(F)$ si l'on agit dynamiquement sur F et V_{dd} . Si l'on modifie dynamiquement F , V_{dd} et V_{bb} le polynôme d'ordre 3 permet tout de même d'obtenir une solution analytique pour maximiser le gain en consommation moyen. En effet, la méthode permet d'obtenir une solution même si certaines variations ne sont pas prises en compte. Nous avons observé qu'une variation de 100 MHz sur la fréquence du PM intermédiaire a une influence maximale de 1% sur le gain maximal.

L'ajout de PM intermédiaire a été évalué à partir de valeurs de consommation issues d'un circuit en FD-SOI. Lorsque les PM sont obtenus en modifiant F et V_{dd} avec V_{bb} fixée à 0 mV, le gain en consommation peut atteindre 21% pour le circuit considéré pour des actionneurs discrets. Si une variation dynamique de la tension de body bias est considérée, alors la consommation peut être réduite d'environ 35% en ajoutant un PM intermédiaire. Cela revient à intégrer un nouvel actionneur discret qui est le saut en tension de body bias.

Dans le chapitre précédent, la tension de body bias V_{bb} était soit fixée à une valeur, soit elle était gérée dynamiquement avec des valeurs discrètes. Cependant, l'actionneur de tension de body bias peut être discret ou continu. Dans le chapitre 4, l'objectif est de minimiser la consommation d'un circuit, en cours d'exécution, en agissant sur F , V_{dd} et V_{bb} tout en garantissant la performance. Des méthodes de minimisation de la consommation ont été exposées. Ces méthodes, basées sur des concepts théoriques (convexité), ont pour objectifs d'être appliquées à des circuits pouvant fonctionner avec un grand nombre de PM. Le but est de choisir, en cours de fonctionnement, les PM qui minimisent la consommation du circuit. Avec ces trois degrés de liberté, une même fréquence peut être atteinte en appliquant différents couples (V_{dd}, V_{bb}) . Cependant, ces couples induisent des consommations différentes. Des méthodes doivent être développées pour déterminer les tensions V_{dd} et V_{bb} qui permettent d'atteindre la fréquence cible afin de garantir la performance tout en minimisant la consommation. Les méthodes proposées consistent à rendre globalement convexe, c'est-à-dire discrètement convexe ou convexe par morceaux, l'ensemble des PM disponibles représentés dans le plan $P_{tot}(F)$. Cette phase correspond à la phase d'identification. La seconde phase est la phase de fonctionnement qui consiste à exécuter les tâches avec les PM appartenant au sous-ensemble discrètement convexe (DCS) ou au sous-ensemble convexe par morceaux (PWCS).

Nous nous sommes en particulier intéressés au cas où $P_{tot}(F)$ est continue par morceaux. Les actionneurs de fréquence et de tension de body bias sont considérés continus et l'actionneur de tension d'alimentation est considéré discret. Nous avons étudié 3 exemples différents, chacun des exemples ayant un nombre de valeurs de tension d'alimentation différent. Nous avons évalué le gain en consommation de la méthode proposée. Les résultats du chapitre 5 ont montré qu'un gain maximal d'environ 19%, a été obtenu pour le circuit considéré, pour 2 niveaux de V_{dd} . Pour les autres exemples, le gain maximal atteint 17% et le gain moyen maximal est de 7%. Notons que le gain est obtenu en considérant des PM qui sont des minimums locaux, c'est-à-dire que pour chaque PM, la fréquence associée tend vers la fréquence maximale atteignable pour le couple (V_{dd}, V_{bb}) associé.

Nous avons mis en évidence que l'ajout de l'actionneur de body bias réduit, sur le circuit FD-SOI considéré, la consommation d'environ 14% si l'on ajoute un PM intermédiaire aux deux déjà disponibles. Par ailleurs, l'identification du sous-ensemble convexe par morceaux permet de gagner en moyenne entre 5% et 7% en consommation par rapport à un ensemble non convexe.

L'intérêt d'implémenter la méthode proposée dépend de la granularité spatiale de contrôle souhaitée et de la granularité temporelle des tâches considérées. Elle va aussi dépendre du profil d'exécution du circuit (la densité de probabilité de la fréquence F), de la gamme de fréquence considérée et du surcoût de la méthode.

La granularité temporelle des tâches est également à prendre en compte car selon le temps imparti et le coût de commutation d'un PM vers un autre PM, il est parfois préférable d'appliquer un PM non optimal. En effet, si le coût de commutation est supérieur au gain en consommation apporté par la méthode en mode *hopping*, le choix d'un PM non optimal est préférable.

Si la gamme de fréquence requise par le circuit est réduite, une gestion dynamique de F et V_{bb} peut être suffisante. En effet, en fixant la tension d'alimentation et en modifiant F et V_{bb} , une gamme de fréquence raisonnable peut être obtenue. Dans les exemples étudiés, nous avons observé que la gamme de fréquence atteignable en ajustant F et V_{bb} est de quelques centaines de MHz.

L'application de la méthode nécessite l'obtention du profil de consommation en fonction de la fréquence. Est-ce que ce profil peut être obtenu en cours d'exécution ? Afin de l'obtenir des capteurs de consommation sont indispensables. Le sous-ensemble convexe par morceaux doit-il être évalué en ligne ? Des ressources de calcul sont alors essentielles afin de déterminer le sous-ensemble convexe par morceaux. Est-ce que l'implémentation de toute la méthode en cours d'exécution est réaliste ? Une implémentation possible est de modéliser la consommation, puis de calibrer le circuit à l'aide d'une mesure de consommation à température et activité constantes, pour un triplet (F, V_{dd}, V_{bb}) .

Les méthodes proposées dans ces travaux sont devenues d'autant plus intéres-

sante à développer car la technologie FD-SOI permet de réaliser du *Body Biasing* pour un surcoût matériel limité. Par ailleurs, la gamme de modification de la tension de body bias en FD-SOI est importante (environ 1,5 V), ce qui élargit la gamme de fréquence atteignable par le circuit.

Grâce aux méthodes proposées et aux opportunités offertes par le FD-SOI, des circuits moins énergivores peuvent être conçus. En effet, des gains substantiels peuvent être obtenus avec les méthodes proposées.

Perspectives

Dans notre étude, en cours d'exécution, la valeur des actionneurs peut être modifiée dynamiquement. Ceci nécessite de commuter la valeur de sortie des actionneurs. Mais à l'aide de quelle métrique évalue-t-on le coût de commutation des actionneurs? Il serait intéressant de l'inclure aux méthodes proposées. En effet, le temps et le coût de commutation dépend complètement de l'actionneur choisi. Communément, il est estimé à 5% du gain en consommation attendu pour les trois actionneurs. Ce gain est en-dessous des gains obtenus par l'ajout de l'actionneur de tension de body bias en appliquant les méthodes proposées.

Comment choisir les actionneurs à implémenter afin de minimiser la consommation d'un circuit? Quelles sont les métriques que nous devons considérer? Le gain en consommation dépend du rendement de l'actionneur de tension, du coût d'implémentation (surface par exemple) des actionneurs et de la densité de consommation des actionneurs.

Nous avons étudié l'impact des variations de procédés de fabrication sur la méthode proposée (PWCS). Cependant, comme nous ne pouvons quantifier le procédé de fabrication, nous ne pouvons inclure ce paramètre dans notre méthode. L'impact des variations de température n'a pas été évalué sur cette même méthode. Une étude importante serait de modéliser ces impacts sur la méthode consistant à rendre convexe par morceaux un ensemble de PM dans le plan $P_{tot}(F)$.

Au niveau technologique, il est possible de concevoir des circuits Multi-Vth en CMOS BULK. L'objectif est de réduire la consommation statique des chemins où la vitesse n'est pas requise en remplaçant certains types de cellules standards par d'autres types de cellules standards. Néanmoins, cette technique nécessite de vérifier l'intégrité du signal, par exemple les retards de couplage. En technologie FD-SOI, cette technique n'est pas encore concevable car elle nécessite un espacement important entre les différents types de cellules.

Si un jour la technologie FD-SOI offre la possibilité de créer des circuits Multi-Vth, en réduisant l'espacement entre les cellules LVT (low Vth) et les cellules RVT (regular Vth), des cœurs Multi-Vth pourraient être conçus. Dans ce cas, de nouvelles possibilités de gestion sont envisageables. La première est de gérer

les deux îlots formés respectivement par les cellules LVT et RVT de manière différente. Cela revient à considérer deux îlots tensions-fréquence pour un même cœur. Cela implique donc une gestion coordonnée afin que le signal soit intègre. Cette solution paraît assez complexe d'un point de vue contrôle. Une seconde approche est de considérer un seul « îlot » tensions-fréquence afin de minimiser la consommation totale du cœur. Cela revient à choisir le triplet $(F, V_{dd}, V_{bb_{virtuelle}})$ à appliquer en considérant le sous-ensemble convexe par morceaux. Néanmoins, cela nécessite deux actionneurs de tension de body bias différents, un pour les cellules LVT et un pour les cellules RVT, c'est pourquoi il est nécessaire d'introduire la notion de $V_{bb_{virtuelle}}$, avec $V_{bb_{virtuelle}} = (V_{bb_{LVT}}, V_{bb_{RVT}})$.

Application au niveau architecture

La méthode consistant à identifier le sous-ensemble convexe par morceaux d'un ensemble de PM peut être appliquée quels que soient les degrés de liberté considérés. Il suffit que le profil de consommation en fonction de la fréquence (ou énergie en fonction de la vitesse) ne soit pas globalement convexe. Par exemple, la méthode peut être appliquée à des architectures hétérogènes, c'est-à-dire combinant des cœurs de calcul ou des clusters différents. Il est possible de combiner des cœurs basse consommation à des cœurs haute performance. L'architecture big.LITTLE [6] de ARM et l'architecture HeteroMates [91] de INTEL sont des exemples d'architectures hétérogènes. Dans ce cas les degrés de liberté sont la fréquence d'horloge F , la tension d'alimentation V_{dd} et le choix du cœur pour l'exécution de la tâche.

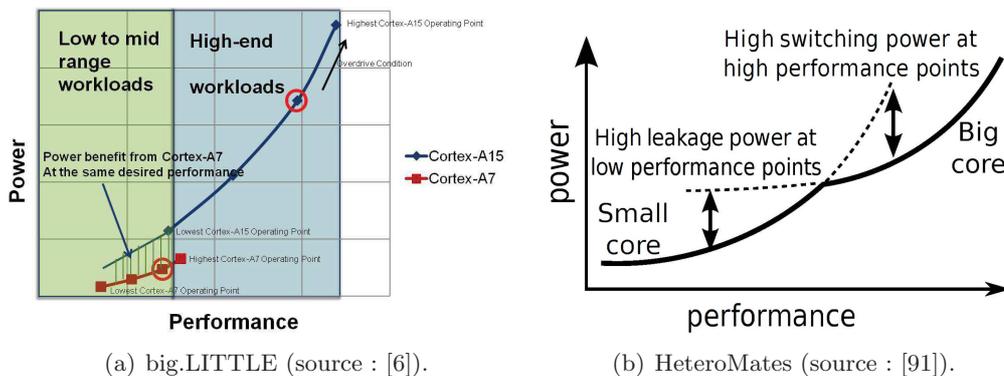


FIGURE 6.1: Profil consommation en fonction de la performance pour des architectures hétérogènes.

Les Figures 6.1(a) et 6.1(b) illustrent respectivement l'allure du profil de consommation en fonction de la performance des deux architectures hétérogènes. Nous remarquons que le profil de consommation en fonction de la performance n'est pas globalement convexe. Nous pouvons identifier le sous-ensemble convexe

par morceaux (PWCS). Après identification des sous-ensembles des courbes appartenant à l'enveloppe convexe, les modes opérationnels définis dans ces travaux (Mode 1 ou Mode 2) peuvent être appliqués. Si la performance requise correspond à un PM du sous-ensemble sélectionné, il est directement appliqué, sinon une exécution en mode *hopping* est effectuée, correspondant ici à une migration de tâches.

Un autre exemple est la généralisation de la méthode proposée à un système multi-cœurs. Un contrôleur haut niveau peut distribuer les tâches en fonction de la performance requise. Dans ce cadre, le choix du cœur d'exécution se fera en tenant compte des profils consommation-performance de chaque cœur. La tâche sera exécutée sur le cœur dont la consommation est minimale pour la performance demandée.

À long terme

Les méthodes proposées dans ces travaux sont applicables à des technologies non planaires de type FinFET, mais aussi aux nouvelles technologies mémoires et voire même à différents domaines (par exemple le domaine IoT).

Le FD-SOI va sûrement rencontrer les mêmes problèmes (consommation, miniaturisation) que la technologie CMOS BULK aujourd'hui. Des nouvelles technologies offriront peut-être de nouvelles opportunités comme l'a fait le FD-SOI. Le problème rencontré risque de se répéter et l'approche proposée de minimisation de la consommation restera valide. De plus, la promesse en performances des circuits futurs, nous laisse entrevoir que notre méthode restera intéressante en permettant des gains en consommations substantiels.

Bibliographie

- [1] Accessed October 05, 2014. [Online]. Available : http://www.st.com/web/en/about_st/moore_law.html [cited at p. 1]
- [2] Accessed September 03, 2013. [Online]. Available : <http://www.itrs.net/Links/2011ITRS/2011Chapters/2011SysDrivers.pdf> [cited at p. 3, 169]
- [3] G. Moore, “Cramming more components onto integrated circuits,” *Proceedings of the IEEE*, vol. 86, no. 1, pp. 82–85, Jan 1998. [cited at p. 7]
- [4] R. Dennard, F. Gaensslen, V. Rideout, E. Bassous, and A. LeBlanc, “Design of ion-implanted mosfet’s with very small physical dimensions,” *Solid-State Circuits, IEEE Journal of*, vol. 9, no. 5, pp. 256–268, Oct 1974. [cited at p. 7]
- [5] P. Venkatesh and S. Alexey, “The ondemand governor,” in *Proc. Linux Symposium*, vol. 2, 2006, pp. 223–238. [cited at p. 8]
- [6] P. Greenhalgh, “Big. little processing with arm cortex-a15 & cortex-a7,” *ARM White Paper*, 2011. [cited at p. 8, 156]
- [7] K. Ahmed and K. Schuegraf, “Transistor wars,” *Spectrum, IEEE*, vol. 48, pp. 50–66, 2011. [cited at p. 8, 169]
- [8] G. Cesana, “The fd-soi technology for energy efficient socs,” in *2014 FD-SOI FORUM*, 2014. [Online]. Available : <http://www.soiconsortium.org/fully-depleted-soi/presentations/september-2014-fd-soi-forum/TP2521%20The%20FD-SOI%20technology%20for%20energy%20eff.pdf> [cited at p. 9, 11, 169]
- [9] T. Chiarella, L. Witters, A. Mercha, C. Kerner, M. Rakowski, C. Ortolland, L.-A. Ragnarsson, B. Parvais, A. D. Keersgieter, S. Kubicek, A. Redolfi, C. Vrancken, S. Brus, A. Lauwers, P. Absil, S. Biesemans, and T. Hoffmann, “Benchmarking {SOI} and bulk finfet alternatives for {PLANAR} {CMOS} scaling succession,” *Solid-State Electronics*, vol. 54, no. 9, pp. 855 – 860, 2010, selected Papers from the {ESSDERC} 2009 Conference. [Online]. Available : <http://www.sciencedirect.com/science/article/pii/S0038110110001279> [cited at p. 10]
- [10] C. Fenouillet-Beranger, P. Perreau, L. Tosti, O. Thomas, J. Noel, O. Weber, F. Andrieu, M. Casse, X. Garros, T. Benoist, S. Haendler, A. Bajolet, F. Boeuf, K. Bourdelle, F. Boedt, and O. Faynot, “Low power utbox and back plane (bp) fdsoi technology for 32nm node and below,” in *IC Design Technology (ICICDT), 2011 IEEE International Conference on*, May 2011, pp. 1–4. [cited at p. 10]

- [11] J.-L. Pelloie, “Physical ip for soi design infrastructure,” in *SOI Conference, 2007 IEEE International*, Oct 2007, pp. 129–130. [cited at p. 11]
- [12] C. Fenouillet-Beranger, O. Thomas, P. Perreau, J.-P. Noel, A. Bajolet, S. Haendler, L. Tosti, S. Barnola, R. Beneyton, C. Perrot, C. de Buttet, F. Abbate, F. Baron, B. Pernet, Y. Campidelli, L. Pinzelli, P. Gouraud, M. Cassé, C. Borowiak, O. Weber, F. Andrieu, S. Denorme, F. Boeuf, O. Faynot, T. Skotnicki, K. Bourdelle, B. Nguyen, and F. Boedt, “Efficient multi-vt fdsoi technology with utbox for low power circuit design,” in *VLSI Technology (VLSIT), 2010 Symposium on*, June 2010, pp. 65–66. [cited at p. 11]
- [13] C. Schuster, “Leakage aware digital design optimization for minimal total power consumption in nanometer cmos technologies,” Ph.D. dissertation, University of Neuchâtel, march 2007. [cited at p. 13]
- [14] Q. Wu, M. Pedram, and X. Wu, “Clock-gating and its application to low power design of sequential circuits,” *Circuits and Systems I : Fundamental Theory and Applications, IEEE Transactions on*, vol. 47, no. 3, pp. 415–420, Mar 2000. [cited at p. 14]
- [15] M. Ohashi, T. Hashimoto, S. Kuromaru, M. Matsuo, T. Mori-iwa, M. Hamada, Y. Sugisawa, M. Arita, H. Tomita, M. Hoshino, H. Miyajima, T. Nakamura, K. Ishida, T. Kimura, Y. Kohashi, T. Kondo, A. Inoue, H. Fujimoto, K. Watada, T. Fukunaga, T. Nishi, H. Ito, and J. Michiyama, “A 27 mhz 11.1 mw mpeg-4 video decoder lsi for mobile application,” in *Solid-State Circuits Conference, 2002. Digest of Technical Papers. ISSCC. 2002 IEEE International*, vol. 1, Feb 2002, pp. 366–474 vol.1. [cited at p. 14]
- [16] F. Emmett and M. Biegel, “Power reduction through rtl clock gating,” *SNUG, San Jose, 2000*, pp. 1–11, 2000. [cited at p. 14, 169]
- [17] S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, and J. Yamada, “1v high-speed digital circuit technology with 0.5 μm multi-threshold cmos,” in *ASIC Conference and Exhibit, 1993. Proceedings., Sixth Annual IEEE International*, Sep 1993, pp. 186–189. [cited at p. 14]
- [18] K. Seta, H. Hara, T. Kuroda, M. Kakumu, and T. Sakurai, “50% active-power saving without speed degradation using standby power reduction (spr) circuit,” in *Solid-State Circuits Conference, 1995. Digest of Technical Papers. 41st ISSCC, 1995 IEEE International*, Feb 1995, pp. 318–319. [cited at p. 15]
- [19] T. da Rosa, V. Larrea, N. Calazans, and F. Moraes, “Power consumption reduction in mpsoCs through dfs,” in *Integrated Circuits and Systems Design (SBCCI), 2012 25th Symposium on*, Aug 2012, pp. 1–6. [cited at p. 17]
- [20] T. Burd, T. Pering, A. Stratakos, and R. Brodersen, “A dynamic voltage scaled microprocessor system,” in *Solid-State Circuits Conference, 2000. Digest of Technical Papers. ISSCC. 2000 IEEE International*, Feb 2000, pp. 294–295. [cited at p. 17, 39]
- [21] C. Albea, D. Puschini, S. Lesecq, and Y. Akgul, “Advanced coupled voltage-frequency control for power efficient dvfs management,” in *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*, Oct 2012, pp. 2168–2173. [cited at p. 17]

- [22] M. Altieri, W. Lombardi, D. Puschini, and S. Lesecq, "Coupled voltage and frequency control for dvfs management," in *Power and Timing Modeling, Optimization and Simulation (PATMOS), 2013 23rd International Workshop on*, Sept 2013, pp. 207–214. [cited at p. 17]
- [23] W. Lombardi, M. Altieri, Y. Akgul, D. Puschini, and S. Lesecq, "Multivariable voltage and frequency control for dvfs management," in *Multi-conference on Systems and Control (MSC), 2014 IEEE*, Oct 2014. [cited at p. 17]
- [24] G. Semeraro, G. Magklis, R. Balasubramonian, D. Albonesi, S. Dwarkadas, and M. Scott, "Energy-efficient processor design using multiple clock domains with dynamic voltage and frequency scaling," in *High-Performance Computer Architecture, 2002. Proceedings. Eighth International Symposium on*, Feb 2002, pp. 29–40. [cited at p. 18]
- [25] K. Choi, R. Soma, and M. Pedram, "Fine-grained dynamic voltage and frequency scaling for precise energy and performance tradeoff based on the ratio of off-chip access to on-chip computation times," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 24, no. 1, pp. 18 – 28, jan. 2005. [cited at p. 18]
- [26] S. Herbert and D. Marculescu, "Variation-aware dynamic voltage/frequency scaling," *High Performance Computer Architecture, 2009. HPCA 2009. IEEE 15th International Symposium on*, pp. 301 –312, feb. 2009. [cited at p. 18, 39]
- [27] V. Gutnik and A. Chandrakasan, "An efficient controller for variable supply-voltage low power processing," in *VLSI Circuits, 1996. Digest of Technical Papers., 1996 Symposium on*, June 1996, pp. 158–159. [cited at p. 19]
- [28] E. Beigne and P. Vivet, "An innovative local adaptive voltage scaling architecture for on-chip variability compensation," *New Circuits and Systems Conference (NEW-CAS), 2011 IEEE 9th International*, pp. 510–513, 2011. [cited at p. 19, 41, 73]
- [29] L. Vincent, E. Beigne, S. Lesecq, J. Mottin, D. Coriat, and P. Maurine, "Dynamic variability monitoring using statistical tests for energy efficient adaptive architectures," *Circuits and Systems I : Regular Papers, IEEE Transactions on*, vol. 61, no. 6, pp. 1741–1754, June 2014. [cited at p. 20]
- [30] S. Kumar, C. Kim, and S. Sapatnekar, "Body bias voltage computations for process and temperature compensation," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 16, no. 3, pp. 249 –262, march 2008. [cited at p. 20]
- [31] C. Kim and K. Roy, "Dynamic vth scaling scheme for active leakage power reduction," in *Design, Automation and Test in Europe Conference and Exhibition, 2002. Proceedings, 2002*, pp. 163 –167. [cited at p. 20, 31]
- [32] S. Kumar, C. Kim, and S. Sapatnekar, "Adaptive techniques for overcoming performance degradation due to aging in cmos circuits," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 19, no. 4, pp. 603 –614, april 2011. [cited at p. 20]
- [33] H. Mostafa, M. Anis, and M. Elmasry, "A novel low area overhead direct adaptive body bias (d-abb) circuit for die-to-die and within-die variations compensation,"

- Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. PP, no. 99, pp. 1–13, 2010. [cited at p. 20]
- [34] L. Yan, J. Luo, and N. Jha, “Joint dynamic voltage scaling and adaptive body biasing for heterogeneous distributed real-time embedded systems,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 24, no. 7, pp. 1030–1041, july 2005. [cited at p. 20, 21, 34]
- [35] T. Kuroda, T. Fujita, S. Mita, T. Nagamatsu, S. Yoshioka, K. Suzuki, F. Sano, M. Norishima, M. Murota, M. Kako, M. Kinugawa, M. Kakumu, and T. Sakurai, “A 0.9-v, 150-mhz, 10-mw, 4 mm², 2-d discrete cosine transform core processor with variable threshold-voltage (vt) scheme,” *Solid-State Circuits, IEEE Journal of*, vol. 31, no. 11, pp. 1770–1779, nov 1996. [cited at p. 20]
- [36] A. Sathanur, A. Pullini, L. Benini, G. De Micheli, and E. Macii, “Physically clustered forward body biasing for variability compensation in nanometer cmos design,” in *Design, Automation Test in Europe Conference Exhibition, 2009. DATE '09.*, april 2009, pp. 154–159. [cited at p. 21, 34]
- [37] S. Martin, K. Flautner, T. Mudge, and D. Blaauw, “Combined dynamic voltage scaling and adaptive body biasing for lower power microprocessors under dynamic workloads,” in *Computer Aided Design, 2002. ICCAD 2002. IEEE/ACM International Conference on*, Nov 2002, pp. 721–725. [cited at p. 21, 34]
- [38] A. Bonnoit, S. Herbert, D. Marculescu, and L. Pileggi, “Integrating dynamic voltage/frequency scaling and adaptive body biasing using test-time voltage selection,” in *ACM/IEEE ISPLED*, 2009. [Online]. Available : <http://doi.acm.org/10.1145/1594233.1594284> [cited at p. 21, 34, 71, 91]
- [39] Accessed October 03, 2014. [Online]. Available : http://www.cadence.com/rl/Resources/white_papers/low_power_impl_wp.pdf [cited at p. 22, 23, 175]
- [40] S. C. Gupta, “Phase-locked loops,” *Proceedings of the IEEE*, vol. 63, no. 2, pp. 291–306, Feb 1975. [cited at p. 22]
- [41] S.-A. Yu and P. Kinget, “A 0.042-mm² fully integrated analog pll with stacked capacitor-inductor in 45nm cmos,” in *Solid-State Circuits Conference, 2008. ESSCIRC 2008. 34th European*, Sept 2008, pp. 94–97. [cited at p. 24]
- [42] S. Lesecq, D. Puschini, E. Beigne, P. Vivet, and Y. Akgul, “Low-cost and robust control of a dfl for multi-processor system-on-chip,” in *Int. Federation of Automatic Control*, 2011. [cited at p. 25, 33, 73, 75]
- [43] P. Hazucha, T. Karnik, B. Bloechel, C. Parsons, D. Finan, and S. Borkar, “Area-efficient linear regulator with ultra-fast load regulation,” *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 4, pp. 933–940, April 2005. [cited at p. 25, 26, 169]
- [44] M. Mulligan, B. Broach, and T. Lee, “A 3mhz low-voltage buck converter with improved light load efficiency,” in *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*, Feb 2007, pp. 528–620. [cited at p. 26, 27, 169]
- [45] J. Wen, T. Jin, and K. Smedley, “A new interleaved isolated boost converter for high power applications,” in *Applied Power Electronics Conference and Exposition, 2006. APEC '06. Twenty-First Annual IEEE*, March 2006, pp. 6 pp.–. [cited at p. 26]

- [46] B. Sahu and G. Rincon-Mora, "A low voltage, dynamic, noninverting, synchronous buck-boost converter for portable applications," *Power Electronics, IEEE Transactions on*, vol. 19, no. 2, pp. 443–452, March 2004. [cited at p. 26]
- [47] G. Patounakis, Y. Li, and K. L. Shepard, "A fully integrated on-chip dc-dc conversion and power management system," *Solid-State Circuits, IEEE Journal of*, vol. 39, no. 3, pp. 443–451, March 2004. [cited at p. 26]
- [48] H.-P. Le, S. Sanders, and E. Alon, "Design techniques for fully integrated switched-capacitor dc-dc converters," *Solid-State Circuits, IEEE Journal of*, vol. 46, no. 9, pp. 2120–2131, Sept 2011. [cited at p. 27, 28, 169, 175]
- [49] T. V. Breussegem and M. Steyaert, "A 82% efficiency 0.5% ripple 16-phase fully integrated capacitive voltage doubler," in *VLSI Circuits, 2009 Symposium on*, June 2009, pp. 198–199. [cited at p. 28]
- [50] D. Somasekhar, B. Srinivasan, G. Pandya, F. Hamzaoglu, M. Khellah, T. Karnik, and K. Zhang, "Multi-phase 1 ghz voltage doubler charge pump in 32 nm logic process," *Solid-State Circuits, IEEE Journal of*, vol. 45, no. 4, pp. 751–758, April 2010. [cited at p. 28]
- [51] L. Chang, R. Montoye, B. Ji, A. Weger, K. Stawiasz, and R. Dennard, "A fully-integrated switched-capacitor 2 :1 voltage converter with regulation capability and 90% efficiency at 2.3a/mm²," in *VLSI Circuits (VLSIC), 2010 IEEE Symposium on*, June 2010, pp. 55–56. [cited at p. 28]
- [52] R. Jevtić, H.-P. Le, M. Blagojević, S. Bailey, K. Asanović, E. Alon, and B. Nikolić, "Per-core dvfs with switched-capacitor converters for energy efficiency in manycore processors," *IEEE Transactions on VLSI Systems*, 2014. [cited at p. 28, 29, 169]
- [53] S. Lee and T. Sakurai, "Run-time voltage hopping for low-power real-time systems," in *Design Automation Conference, 2000. Proceedings 2000*, 2000, pp. 806–809. [cited at p. 31, 34]
- [54] K. Nose, M. Hirabayashi, H. Kawaguchi, S. Lee, and T. Sakurai, "V_{TH}-hopping scheme for 82% power saving in low-voltage processors," in *Custom Integrated Circuits, 2001, IEEE Conference on.*, 2001, pp. 93–96. [cited at p. 31, 32, 33, 73, 75, 169]
- [55] J. Tschanz, J. Kao, S. Narendra, R. Nair, D. Antoniadis, A. Chandrakasan, and V. De, "Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage," in *Solid-State Circuits Conference, 2002. Digest of Technical Papers. ISSCC. 2002 IEEE International*, vol. 1, Feb 2002, pp. 422–478 vol.1. [cited at p. 31, 32, 33, 73]
- [56] D. Jacquet, F. Hasbani, P. Flatresse, R. Wilson, F. Arnaud, G. Cesana, T. Di Gilio, C. Lecocq, T. Roy, A. Chhabra, C. Grover, O. Minez, J. Uginet, G. Durieu, C. Adobati, D. Casalotto, F. Nyer, P. Menut, A. Cathelin, I. Vongsavady, and P. Magarshack, "A 3 ghz dual core processor arm cortex tm -a9 in 28 nm utbb fd-soi cmos with ultra-wide voltage range and energy efficiency optimization," *Solid-State Circuits, IEEE Journal of*, vol. 49, no. 4, pp. 812–826, April 2014. [cited at p. 32, 34, 71]

- [57] M. Javidan, E. Zianbetov, F. Anceau, D. Galayko, A. Korniienko, E. Colinet, G. Scorletti, J. Akre, and J. Juillard, “All-digital pll array provides reliable distributed clock for socs,” in *IEEE Int. Symp. on Circuits and Systems (ISCAS)*, 2011. [cited at p. 33, 73]
- [58] J. Tschanz, S. Narendra, A. Keshavarzi, and V. De, “Adaptive circuit techniques to minimize variation impacts on microprocessor performance and power,” in *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*, May 2005, pp. 9–12 Vol. 1. [cited at p. 34]
- [59] F. Firouzi, A. Yazdanbakhsh, H. Dorosti, and S. Fakhraie, “Dynamic soft error hardening via joint body biasing and dynamic voltage scaling,” in *Euromicro Conf. on Digital System Design*, 2011. [cited at p. 34, 71]
- [60] S. Herbert, S. Garg, and D. Marculescu, “Exploiting process variability in voltage/frequency control,” *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 20, no. 8, pp. 1392–1404, Aug 2012. [cited at p. 38]
- [61] S. Garg, D. Marculescu, R. Marculescu, and U. Ogras, “Technology-driven limits on dvfs controllability of multiple voltage-frequency island designs : A system-level perspective,” in *Design Automation Conference, 2009. DAC '09. 46th ACM/IEEE*, July 2009, pp. 818–821. [cited at p. 38]
- [62] K. Nose and T. Sakurai, “Optimization of V_{DD} and V_{TH} for low-power and high-speed applications,” in *Asia and South Pacific Design Automation Conf.*, 2000. [cited at p. 38]
- [63] K. Nose, M. Hirabayashi, H. Kawaguchi, S. Lee, and T. Sakurai, “ V_{TH} -hopping scheme to reduce subthreshold leakage for low-power processors,” *Solid-State Circuits, IEEE Journal of*, vol. 37, no. 3, pp. 413–419, mar 2002. [cited at p. 38]
- [64] M. Nomura, Y. Ikenaga, K. Takeda, Y. Nakazawa, Y. Aimoto, and Y. Hagihara, “Delay and power monitoring schemes for minimizing power consumption by means of supply and threshold voltage control in active and standby modes,” *Solid-State Circuits, IEEE Journal of*, vol. 41, no. 4, pp. 805–814, 2006. [cited at p. 38]
- [65] S. Martin, K. Flautner, T. Mudge, and D. Blaauw, “Combined dynamic voltage scaling and adaptive body biasing for lower power microprocessors under dynamic workloads,” in *Computer Aided Design, 2002. ICCAD 2002. IEEE/ACM International Conference on*, Nov 2002, pp. 721–725. [cited at p. 38]
- [66] Y. Akgul, D. Puschini, S. Lesecq, I. Miro-Panades, P. Benoit, L. Torres, and E. Beigne, “Power mode selection in embedded systems with performance constraints,” in *Faible Tension Faible Consommation (FTFC), 2012 IEEE*, 2012, pp. 1–4. [cited at p. 38, 76]
- [67] R. Rao and S. Vrudlhula, “Energy optimal speed control of devices with discrete speed sets,” in *Design Automation Conference*, 2005. [cited at p. 38, 39, 75, 76, 79, 80]
- [68] B. Rebaud, M. Belleville, E. Beigne, M. Robert, P. Maurine, and N. Azemard, “On-chip timing slack monitoring,” in *IFIP Int. Conf. on Very Large Scale Integration*, 2009. [cited at p. 41, 72, 86, 90, 94, 102, 116]

- [69] S. Das, C. Tokunaga, S. Pant, W.-H. Ma, S. Kalaiselvan, K. Lai, D. Bull, and D. Blaauw, “RazorII : In Situ Error Detection and Correction for PVT and SER Tolerance,” *Solid-State Circuits, IEEE Journal of*, vol. 44, pp. 32–48, 2009. [cited at p. 41]
- [70] L. Vincent, P. Maurine, S. Lesecq, and E. Beigne, “Embedding statistical tests for on-chip dynamic voltage and temperature monitoring,” in *Design Automation Conference*, 2012. [cited at p. 42, 74, 94]
- [71] A. Valentian and E. Beigne, “Automatic gate biasing of an sccmos power switch achieving maximum leakage reduction and lowering leakage current variability,” *Solid-State Circuits, IEEE Journal of*, vol. 43, no. 7, pp. 1688–1698, 2008. [cited at p. 42]
- [72] A. Nelson, A. Nejad, A. Molnos, M. Koedam, and K. Goossens, “Comik : A predictable and cycle-accurately composable real-time microkernel,” in *Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014*, March 2014, pp. 1–4. [cited at p. 42, 74]
- [73] I. Miro-Panades, E. Beigne, Y. Thonnart, L. Alacoque, P. Vivet, S. Lesecq, D. Puschini, A. Molnos, F. Thabet, B. Tain, K. Ben Chehida, S. Engels, R. Wilson, and D. Fuin, “A fine-grain variation-aware dynamic vdd-hopping avfs architecture on a 32 nm gals mpsoc,” *Solid-State Circuits, IEEE Journal of*, vol. 49, no. 7, pp. 1475–1486, July 2014. [cited at p. 44]
- [74] R. Wilson, E. Beigne, P. Flatresse, A. Valentian, F. Abouzeid, T. Benoist, C. Bernard, S. Bernard, O. Billoint, S. Clerc, B. Giraud, A. Grover, J. Le Coz, I. Panades, J.-P. Noel, B. Pelloux-Prayer, P. Roche, O. Thomas, Y. Thonnart, D. Turgis, F. Clermidy, and P. Magarshack, “27.1 a 460mhz at 397mv, 2.6ghz at 1.3v, 32b vliw dsp, embedding fmax tracking,” in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International*, Feb 2014, pp. 452–453. [cited at p. 61, 71, 116, 122, 134, 141, 144]
- [75] Y. Akgul, D. Puschini, S. Lesecq, P. Benoit, L. Torres, and E. Beigné, “Power optimization in fd-soi circuits with vdd-hopping and adaptive body bias,” *submitted in Design, Automation Test in Europe Conference Exhibition (DATE), 2015*, pp. 1–6, 2015. [cited at p. 69]
- [76] Y. Akgul, D. Puschini, S. Lesecq, E. Beigné, I. Miro-Panades, P. Benoit, and L. Torres, “Power management through dvfs and dynamic body biasing in fd-soi circuits,” in *Proceedings of the The 51st Annual Design Automation Conference on Design Automation Conference*, ser. DAC '14. New York, NY, USA : ACM, 2014, pp. 183 :1–183 :6. [Online]. Available : <http://doi.acm.org/10.1145/2593069.2593185> [cited at p. 70, 89]
- [77] Y. Akgul, D. Puschini, S. Lesecq, E. Beigne, P. Benoit, and L. Torres, “Methodology for power mode selection in fd-soi circuits with dvfs and dynamic body biasing,” in *Power and Timing Modeling, Optimization and Simulation*, 2013. [cited at p. 70, 80]
- [78] Y. Akgul, E. Beigné, S. Lesecq, D. Puschini, and P. Benoit, “Procédé de commande d’un circuit électronique,” FR Patent FR Patent 13/58 541, 2013. [cited at p. 70, 80, 89]

- [79] J.-L. Lions, *Petite encyclopédie des mathématiques*. K. Pagoulatos, 1980. [cited at p. 73]
- [80] Accessed June 06, 2014. [Online]. Available : http://magnoludoviciens.perso.sfr.fr/cours%20cours/14_%20fonctions%20par%20morceaux.pdf [cited at p. 73]
- [81] Accessed June 16, 2014. [Online]. Available : <http://spiral.univ-lyon1.fr/mathsv/cours/stats/chap3/c3p2/c3p2.html> [cited at p. 73]
- [82] J. Peddersen and S. Parameswaran, “Clipper : Counter-based low impact processor power estimation at run-time,” in *Design Automation Conference, 2007. ASP-DAC '07. Asia and South Pacific*, Jan 2007, pp. 890–895. [cited at p. 74]
- [83] K. Ganesan, L. John, V. Salapura, and J. Sexton, “A performance counter based workload characterization on blue gene/p,” in *Parallel Processing, 2008. ICPP '08. 37th International Conference on*, Sept 2008, pp. 330–337. [cited at p. 74]
- [84] U. Yüceer, “Discrete convexity : convexity for functions defined on discrete spaces,” *Discrete Applied Mathematics*, vol. 119, pp. 297–304, 2002. [cited at p. 75, 109]
- [85] B. Bereanu, “A property of convex piecewise linear functions with applications to mathematical programming,” *Unternehmensforschung*, vol. 9, no. 2, pp. 112–119, 1965. [Online]. Available : <http://dx.doi.org/10.1007/BF01919478> [cited at p. 85, 92]
- [86] R. Graham, “An efficient algorithm for determining the convex hull of a finite planar set,” in *Information Processing Letters 1*, January 1972, pp. 132–133. [cited at p. 92, 123]
- [87] R. A. Jarvis, “On the identification of the convex hull of a finite set of points in the plane,” *Information Processing Letters*, vol. 2, no. 1, pp. 18–21, 1973. [cited at p. 92, 123]
- [88] I. Mansouri, P. Benoit, L. Torres, and F. Clermidy, “Fine-grain dynamic energy tracking for system on chip,” *Circuits and Systems II : Express Briefs, IEEE Transactions on*, vol. 60, no. 6, pp. 356–360, 2013. [cited at p. 94]
- [89] S. Miermont, E. Beigne, and P. Vivet, “Electronic circuit power supply device and electronic circuit,” US Patent US008 018 093B2, 09 13, 2011. [cited at p. 101]
- [90] E. Beigne, F. Clermidy, S. Miermont, and P. Vivet, “Dynamic voltage and frequency scaling architecture for units integration within a gals noc,” in *Networks-on-Chip, 2008. NoCS 2008. Second ACM/IEEE International Symposium on*, April 2008, pp. 129–138. [cited at p. 141]
- [91] V. Gupta, P. Brett, D. Koufaty, D. Reddy, S. Hahn, K. Schwan, and G. Srinivasa, “Heteromates : Providing high dynamic power range on client devices using heterogeneous core groups,” in *Green Computing Conference (IGCC), 2012 International*, June 2012, pp. 1–10. [cited at p. 156]

Bibliographie Personnelle

Publications internationales

S. Lesecq, D. Puschini, E. Beigné, P. Vivet, and Y. Akgul, “Low-Cost and Robust Control of a DFLL for Multi-Processor System-On-Chip,” *18th International Federation of Automatic Control, IFAC 2011*, 2011. [cited at p. -]

C. Albea, D. Puschini, S. Lesecq, and Y. Akgul, “Advanced coupled voltage-frequency control for power efficient DVFS management,” *IECON 2012, 38th Annual Conference on IEEE Industrial Electronics Society*, pp. 2168–2173, 2012. [cited at p. -]

Y. Akgul, D. Puschini, S. Lesecq, I. Miro-Panades, P. Benoit, L. Torres, and E. Beigné, “Power mode selection in embedded systems with performance constraints,” *Faible Tension Faible Consommation (FTFC), 2012 IEEE*, pp. 1–4, 2012. [cited at p. -]

Y. Akgul, D. Puschini, S. Lesecq, E. Beigné, P. Benoit, and L. Torres, “Methodology for Power Mode selection in FD-SOI circuits with DVFS and Dynamic Body Biasing,” *Power and Timing Modeling, Optimization and Simulation (PATMOS), 2013 23rd International Workshop on*, pp. 199–206, 2013. [cited at p. -]

W. Lombardi, M. Altieri Scarpato, Y. Akgul, D. Puschini, and S. Lesecq, “Multivariable Voltage and Frequency Control for DVFS Management,” *Multi-conference on Systems and Control (MSC), 2014 IEEE*, pp. 1–6, 2014. [cited at p. -]

Y. Akgul, D. Puschini, S. Lesecq, P. Benoit, I. Miro-Panades, E. Beigné, and L. Torres, “Power Management through DVFS and Dynamic Body Biasing in FD-SOI circuits,” *Design Automation Conference (DAC), 2014 51st ACM/EDAC/IEEE*, pp. 1–6, 2014. [cited at p. -]

Brevet

Y. Akgul, E. Beigné, S. Lesecq, D. Puschini, and P. Benoit, “Procédé de commande d’un circuit électronique,” FR Patent 13/58541, 2013. [cited at p. -]

Table des figures

1.1	Tendance des futures évolutions du nombre de cœurs par SoC (source : [2]). . .	3
2.1	Différentes technologies de transistors (source : [7]).	8
2.2	Tension de polarisation du substrat en technologie BULK (source : [8]).	9
2.3	Nouvelle entrée permettant d'agir sur la tension de body bias en technologie FD-SOI (source : [8]).	11
2.4	Représentation du système par un modèle entrées-sorties.	12
2.5	Circuit de clock gating basé sur une bascule Latch (source : [16]).	14
2.6	Gestion de la consommation et de la performance avec les techniques au niveau circuit et système.	16
2.7	Séquence de modification de la fréquence F et de la tension d'alimentation V_{dd}	18
2.8	Profil consommation-performance DVFS et DVFS discret.	19
2.9	Structure de base d'une FLL et d'une PLL.	24
2.10	Schéma bloc du régulateur linéaire rapide (source : [43]).	26
2.11	Schéma bloc d'un convertisseur Buck (source : [44]).	27
2.12	Rendement η attendu pour un convertisseur capacitif en fonction de la tension en sortie de l'actionneur V_{out} disposant de différents ratio de conversion n , pour différentes capacités (source : [48]).	28
2.13	Schéma bloc d'un convertisseur capacitif (source : [52]).	29
2.14	Schéma bloc de l'alimentation d'une architecture divisée en plusieurs îlots avec un mécanisme de V_{dd} -Hopping.	30
2.15	Schéma bloc du V_{th} -Hopping (source : [54]).	32
3.1	Schéma d'un îlot avec 2 actionneurs. d est le temps imparti pour exécuter la tâche et N_c correspond au nombre de cycles d'horloge nécessaires pour exécuter la tâche.	38
3.2	Exemple de positionnement des PM dans le plan (F, P_{tot})	43
3.3	Hopping entre deux fréquences pour atteindre une performance équivalente à la fréquence F_{target} avec α le rapport cyclique et d le temps imparti connu pour la tâche considérée.	45
3.4	Gain en consommation G_{3PM} en ayant 3 PM disponibles par rapport à 2 PM disponibles, avec la fréquence de PM_{middle} choisie à 820 MHz.	47

3.5	Gain en consommation G_{3PM} en ayant 3 PM disponibles par rapport à 2 PM disponibles, avec la fréquence de PM_{middle} choisie à 652 MHz.	48
3.6	Gain en consommation maximal G_{max} obtenu par l'ajout d'un PM intermédiaire pour toutes les possibilités en fonction de la fréquence du PM intermédiaire F_{middle}	49
3.7	Contribution de l'ajout d'un PM intermédiaire sur le gain en consommation avec 3 PM par rapport à 2 PM disponibles, pour toutes les possibilités.	50
3.8	Contribution de l'ajout d'un PM intermédiaire sur le gain en consommation avec 3 PM par rapport à 2 PM disponibles, pour une gamme de fréquence.	51
3.9	Consommation en fonction de la fréquence, avec l'apport d'un PM intermédiaire ($F_{middle}, P_{tot}(F_{middle})$) sur toute la gamme de fréquence.	53
3.10	Consommation en fonction de la fréquence, avec l'apport d'un PM intermédiaire ($F_{middle}, P_{tot}(F_{middle})$) sur la gamme de fréquence $[F_a ; F_b]$	54
3.11	Consommation en fonction de la fréquence, avec l'apport d'un PM intermédiaire ($F_{middle}, P_{tot}(F_{middle})$) sur la gamme de fréquence $[F_c ; F_d]$	56
3.12	Réductions en consommation moyenne $\overline{PR_M}$ et $\overline{PR_{Tri}}$ obtenues par l'ajout d'un PM intermédiaire en fonction de la fréquence du PM intermédiaire F_{middle} , pour toutes les possibilités.	57
3.13	Profil de consommation en fonction de la fréquence $P_{tot}(F)$ et modèle cubique sur LoCoMoTIV.	60
3.14	Réduction en consommation moyenne $\overline{PR_{Tri}}$ obtenue par l'ajout d'un PM intermédiaire en fonction de la fréquence du PM intermédiaire F_{middle} , pour quatre V_{bb}	62
3.15	Réduction en consommation moyenne $\overline{PR_{Tri}}$ obtenue par l'ajout d'un PM intermédiaire en fonction de la fréquence du PM intermédiaire F_{middle} en modifiant (F, V_{dd}, V_{bb}).	63
3.16	Profil de consommation en fonction de la fréquence pour la position optimale F_{middle}^{opt} en modifiant (F, V_{dd}, V_{bb}).	63
3.17	Consommation en fonction de la fréquence, avec l'apport de N PM sur toute la gamme de fréquence.	65
3.18	Consommation en fonction de la fréquence, avec l'apport de N PM sur la gamme de fréquence $[F_g ; F_h]$	66
4.1	Schéma d'un îlot avec 3 actionneurs. d est le temps imparti et N_c correspond au nombre de cycles d'horloge.	70
4.2	Ensemble de PM dans le plan (F, P_{tot}).	74
4.3	Exemple d'un ensemble convexe croissant $P_{tot}(F)$	76
4.4	Exemple de 3 fréquences applicables (F_{n-1}, F_n, F_{n+k}) dans le plan (F, P_{tot}).	77
4.5	Exemple de 3 fréquences applicables (F_{n-k}, F_n, F_{n+1}) dans le plan (F, P_{tot}).	78
4.6	Phase d'identification du DCS permettant de déterminer les modes de fonctionnement possibles et les modes optimaux (étapes 1 à 4)	81
4.7	Phase de fonctionnement (exécution de la tâche) (étapes 5 à 7).	81
4.8	$P_{tot}(F)$ avec PM_n appartenant au DCS.	82
4.9	$P_{tot}(F)$ avec PM_n n'appartenant pas au DCS.	82
4.10	$P_{tot}(F)$ avec F_{target} correspondant à la fréquence d'un PM appartenant au DCS.	83

4.11	$P_{tot}(F)$ avec F_{target} correspondant à la fréquence d'un PM n'appartenant pas au DCS.	83
4.12	$P_{tot}(F)$ avec F_{target} ne correspondant à la fréquence d'aucun PM.	84
4.13	Consommation P_{tot} d'un oscillateur en anneau en fonction de $V_{bb} \in [V_{bb_{min}} ; V_{bb_{max}}]$ et de $V_{dd} \in \{V_{dd_{min}} ; V_{dd_{max}}\}$. La coloration correspond à la fréquence maximale (F_{max}) applicable pour chaque couple (V_{dd}, V_{bb})	85
4.14	Fréquence maximale (F_{max}) applicable en fonction de $V_{bb} \in [V_{bb_{min}} ; V_{bb_{max}}]$ pour 4 valeurs de V_{dd}	86
4.15	Exemple avec 4 valeurs de V_{dd} , V_{bb} continu et F_{target} atteignable par un seul couple (V_{dd}, V_{bb})	87
4.16	Exemple avec 4 valeurs de V_{dd} , V_{bb} continu et F_{target} atteignable par deux couples (V_{dd}, V_{bb})	87
4.17	Exemple avec 4 valeurs de V_{dd} , V_{bb} continu et F_{target} non atteignable par un couple (V_{dd}, V_{bb})	88
4.18	<i>Hopping</i> entre deux fréquences pour atteindre une performance équivalente à la fréquence F_{target} , avec α le rapport cyclique et d le temps imparti, connu pour la tâche considérée.	89
4.19	Puissance P_{tot} en fonction de la fréquence pour 4 V_{dd} et $V_{bb} \in [V_{bb_{min}} ; V_{bb_{max}}]$	89
4.20	Phase d'identification du PWCS permettant de déterminer les modes de fonctionnement possibles et les modes optimaux (étapes 1 à 5).	90
4.21	Puissance P_{tot} en fonction de la fréquence pour 4 V_{dd} différents et $V_{bb} \in [V_{bb_{min}} ; V_{bb_{max}}]$	91
4.22	Identification des morceaux de courbe $P_{tot}(F)$ appartenant au PWCS.	92
4.23	Détermination des limites basses et hautes pour T° et A données.	93
4.24	Modèles de consommation avec les PM appartenant au PWCS en vert pour le cas I.	97
4.25	Modèles de consommation avec les PM appartenant au PWCS tracés en vert pour le cas II.	98
4.26	Phase de fonctionnement (exécution de la tâche) (étapes 6 à 8).	99
4.27	Ensemble de PM et modes d'application possibles. Mode 1 ou M1 : on applique un unique PM. Mode 2 ou M2 : on applique 2 PM en mode <i>hopping</i>	99
4.28	Exemple d'application du Mode 1.	101
4.29	Exemple d'application du Mode 2.	102
4.30	Mode 2 : Exécution en mode <i>hopping</i> avec un <i>hopping</i> en F et en V_{dd} , avec une tension de body bias constante V_{bb_x}	104
5.1	Ensemble de fréquences disponibles F en fonction de la tension d'alimentation $V_{dd} : F(V_{dd})$	110
5.2	Ensemble de PM dans le plan (F, P_{tot}) avec P_{tot} la consommation totale du circuit et F la fréquence d'horloge.	110
5.3	$P_{tot}(F)$ avec PM_n appartenant ou non au DCS croissant.	111
5.4	Ensemble de PM dans le plan (F, P_{tot}) avec P_{tot} la consommation totale et F la fréquence avec les PM appartenant au DCS en vert.	112
5.5	Consommation P_{tot} en fonction de la fréquence F (en haut). Gain en consommation G avec la méthode proposée en fonction de la fréquence F (en bas).	114
5.6	Lien entre le gain en consommation et le plan (F, P_{tot})	115

5.7	Ensemble de fréquences disponibles F en fonction de la tension d'alimentation V_{dd} de Frisbee.	117
5.8	Ensemble de PM dans le plan (F, P_{tot}) avec P_{tot} la consommation totale du circuit et F la fréquence d'horloge de Frisbee.	118
5.9	Ensemble de PM dans le plan (F, P_{tot}) avec P_{tot} la consommation totale et F la fréquence avec les PM appartenant au DCS en vert de Frisbee.	119
5.10	Consommation P_{tot} en fonction de la fréquence F du circuit Frisbee (en haut). Gain en consommation G avec la méthode proposée en fonction de la fréquence F sur Frisbee (en bas).	120
5.11	Ensemble de PM dans le plan (F, P_{tot}) pour 2 niveaux de V_{dd} avec les PM appartenant au PWCS en vert.	124
5.12	Ensemble de PM dans le plan (F, P_{tot}) pour 3 niveaux de V_{dd} avec les PM appartenant au PWCS en vert.	125
5.13	Ensemble de PM dans le plan (F, P_{tot}) pour 4 niveaux de V_{dd} avec les PM appartenant au PWCS en vert.	126
5.14	Consommation P_{tot} en fonction de la fréquence F (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 2 niveaux de V_{dd} (en bas).	128
5.15	Consommation P_{tot} en fonction de la fréquence F (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 3 niveaux de V_{dd} (en bas).	129
5.16	Consommation P_{tot} en fonction de la fréquence F (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 4 niveaux de V_{dd} (en bas).	131
5.17	Lien entre le maximum de gain en consommation et le plan (F, P_{tot}) pour Frisbee.	132
5.18	Mode 2 : Exécution en mode <i>hopping</i> avec un <i>hopping</i> en V_{dd} et en F , avec une tension de body bias constante V_{bb_x}	134
5.19	Gain en consommation G_{so} en exécutant avec 2 PM (en ligne continue), 1 PM (en ligne mixte) et 0 PM (en ligne pointillée) appartenant au PWCS avec V_{bb_x} constant pour les 3 gains maximaux de la Figure 5.16.	136
5.20	Gain en consommation G_{so} en exécutant avec 2 PM (en ligne continue), 1 PM (en ligne mixte) et 0 PM (en ligne pointillée) appartenant au PWCS avec V_{bb_x} constant pour 3 F_{target} choisies arbitrairement.	139
5.21	Consommation P_{tot} en fonction de la fréquence F avec les sous-ensembles PWCS et PN (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 4 niveaux de V_{dd} pour un coût de commutation constant (en bas).	142
5.22	Consommation P_{tot} en fonction de la fréquence F avec les sous-ensembles PWCS et PN (en haut). Gain en consommation G avec la méthode proposée (en Mode 2) en fonction de la fréquence F pour 4 niveaux de V_{dd} pour un coût de commutation non constant (en bas).	143
5.23	Ensemble de PM dans le plan (F, P_{tot}) pour 4 niveaux de V_{dd} et 3 puces différentes avec les PM appartenant au PWCS en vert.	145
5.24	Ensemble de PM dans le plan (F, P_{tot}) pour 4 niveaux de V_{dd} pour 16 puces avec les PM appartenant au PWCS en vert.	146

5.25	Ensemble des limites basses et hautes du PWCS dans le plan (F, P_{tot}) pour 4 niveaux de V_{dd} et 16 puces.	147
5.26	Pic de gain et gain maximal G_{max} pour chaque puce lorsque 4 niveaux de V_{dd} sont disponibles.	148
6.1	Profil consommation en fonction de la performance pour des architectures hétérogènes.	156

Liste des tableaux

2.1	Synthèse des techniques de gestion du compromis consommation-performance (source : [39]).	23
2.2	Synthèse des avantages et des inconvénients sur la gestion du compromis consommation-performance en agissant sur les entrées F , V_{dd} , V_{bb}	24
2.3	Comparaison entre une PLL et une FLL	25
2.4	Comparaison de convertisseurs capacitifs continu-continu (<i>Switched-Capacitor DC-DC Converter</i>) (source : [48]).	28
2.5	Comparaison de convertisseurs de tension	29
2.6	Table de vérité du schéma bloc de la Figure 2.15.	32
3.1	Valeurs de la consommation P_{tot} , de la fréquence F et de la tension d'alimentation V_{dd} associées à chaque PM, avec la fréquence de PM _{middle} choisie à 820 MHz.	47
3.2	Valeurs de la consommation P_{tot} , de la fréquence F et de la tension d'alimentation V_{dd} associées à chaque PM, avec la fréquence de PM _{middle} choisie à 652 MHz.	48
3.3	La moyenne du gain en consommation $\overline{G_G}$ obtenue par l'ajout du PM intermédiaire défini avec la consommation $P_{tot}(F_{middle})$ associée à la fréquence F_{middle} et à $V_{dd_{middle}}$ sur l'intervalle [700 ; 900] MHz.	52
3.4	La réduction en consommation moyenne $\overline{PR_M}$ avec MATLAB et la réduction en consommation moyenne $\overline{PR_{Tri}}$ avec la méthode du triangle obtenues par l'ajout du PM intermédiaire défini avec la consommation $P_{tot}(F_{middle})$ associée à la fréquence F_{middle} et à $V_{dd_{middle}}$ sur toute la gamme de fréquence.	58
3.5	Tableau de signes de l'évolution de S_{Tri}	60
3.6	Variation de la réduction en consommation moyenne $\overline{PR_{Tri}}$ lorsque F_{middle} varie.	62
3.7	Meilleurs PM avec 3 actionneurs discrets.	62
4.1	Limites pour des gammes de T° et A données.	94
4.2	Limites pour des gammes de température T° et activité A données.	102
5.1	Valeurs de la tension d'alimentation V_{dd} , la tension de body bias V_{bb} , la fréquence normalisée F , la consommation normalisée P_{tot} pour chaque PM.	111
5.2	Valeur des pentes afin d'identifier les PM appartenant au DCS, avec $1 < n < N$	112

5.3	Valeurs de la tension d'alimentation V_{dd} , la tension de body bias V_{bb} , la fréquence F , la consommation P_{tot} pour chaque PM de Frisbee.	117
5.4	Valeur des pentes afin d'identifier les PM appartenant au DCS pour Frisbee, avec $1 < n < N$	119
5.5	Valeur du gain moyen $\overline{G_{CTr}}$ sur l'intervalle $[F_a ; F_b]$ considéré et valeur du gain en consommation maximal G_{max} pour la fréquence cible F_i du PM_i	121
5.6	Valeurs limites lorsque 2 niveaux de V_{dd} sont disponibles.	123
5.7	Valeurs limites lorsque 3 niveaux de V_{dd} sont disponibles.	124
5.8	Valeurs limites lorsque 4 niveaux de V_{dd} sont disponibles.	125
5.9	Gain en consommation G en exécutant avec le Mode 2 pour 2 V_{dd}	128
5.10	Gain en consommation G en exécutant avec le Mode 2 pour 3 V_{dd}	130
5.11	Gain en consommation G en exécutant avec le Mode 2 pour 4 V_{dd}	130
5.12	Nombre de PM appartenant au PWCS en exécutant avec les tensions V_{dd_i} et $V_{dd_{i+1}}$ pour la gamme V_{bb_x} et la gamme F_{target}	137
5.13	Gain G en exécutant avec le Mode 2 et gain minimal G_{somin} en exécutant avec 2 PM quelconques afin d'atteindre F_{target} en moyenne pour V_{bb_x} constant pour les 3 gains maximaux de la Figure 5.16.	137
5.14	Gain G en exécutant avec le Mode 2 et gain maximal G_{soma} en exécutant avec 2 PM quelconques afin d'atteindre F_{target} en moyenne pour V_{bb_x} constant pour les 3 gains maximaux de la Figure 5.16.	137
5.15	Gain G en exécutant avec le Mode 2 et gain minimal G_{somin} en exécutant avec 2 PM quelconques afin d'atteindre F_{target} en moyenne pour V_{bb_x} constant pour 3 F_{target} choisies arbitrairement.	140
5.16	Gain G en exécutant avec le Mode 2 et gain maximal G_{soma} en exécutant avec 2 PM quelconques afin d'atteindre F_{target} en moyenne pour V_{bb_x} constant pour 3 F_{target} choisies arbitrairement.	140

Liste des acronymes

ABB	Adaptive Body Biasing.
AVFS	Adaptive Voltage and Frequency Scaling.
BP	Back-Plane.
BTBT	Band-To-Band Tunneling.
BTI	Bias Temperature Instability.
CIN	Circuit Intégré Numérique.
CMOS	Complementary Metal Oxide Semiconductor.
continu	Une fonction est définie comme continue si la fonction est ininterrompue. En d'autres mots, cela revient à "ne pas lever son crayon" lors du tracé de cette fonction.
continu par morceaux	Une fonction est définie comme continue par morceaux sur l'intervalle $[a, b[$ si la fonction interrompue en quelques points seulement. En d'autres mots, cela revient à "lever son crayon" lors du tracé de cette fonction.
CPU	Central Processing Unit.
d	temps imparti.
DBB	Dynamic Body Biasing.
DCS	Discretely Convex Subset.
DFS	Dynamic Frequency Scaling.
discret	Une fonction est considérée comme discrète si elle ne prend qu'un nombre fini de valeurs discontinues.
DoF	Degree(s) of Freedom ou degré(s) de liberté en français.
DSP	Digital Signal Processing.
DVFS	Dynamic Voltage and Frequency Scaling.
DVS	Dynamic Voltage Scaling.
FBB	Forward Body Biasing.

FD-SOI	Fully Depleted SOI.
FinFET	Fin-shaped Field Effect Transistor.
FLL	Frequency Locked-Loop.
GP	Ground-Plane.
ITRS	International Technology Roadmap for Semiconductors.
LIALP	Laboratoire Infrastructure et Atelier Logiciel pour Puces.
LIRMM	Laboratoire d'Informatique, Robotique et Micro-électronique de Montpellier.
LISAN	Laboratoire Intégration Silicium et Architecture Numérique.
LoCoMoTIV	Local Compensation of Modern Technology Induced Variability.
MCD	Multiple Clock Domains.
MOSFET	Metal Oxide Semiconductor Field-Effect Transistor.
MPSoC	Multi-Processor System on Chip.
PD-SOI	Partially Depleted SOI.
PDF	densité de probabilité.
PE	Processing Element.
PGN	Power Gating Noise.
PLL	Phase Locked-Loop.
PM	Power Mode.
PVT	Process Voltage Temperature.
PWCS	Piece-Wise Convex Subset.
RBB	Reverse Body Biasing.
RO	Ring-Oscillator.
SoC	System on Chip.
SOI	Silicon On Insulator.
temps imparti	Il s'agit du temps que la tâche dispose pour être complètement exécutée.
TSM	Timing Slack Monitor.
UTBB	Ultra Thin Body and Box.

variations Die-to-Die	Il s'agit de la variation des paramètres intrinsèques aux transistors (concentration de porteurs (dopage), longueur, largeur, tension de seuil, <i>etc</i>) issus de différentes puces.
variations With-In-Die	Il s'agit de la variation des paramètres intrinsèques aux transistors (concentration de porteurs (dopage), longueur, largeur, tension de seuil, <i>etc</i>) issus d'une même puce.
VCO	Voltage-Controlled Oscillator.
VFI	Voltage/Frequency Island.

Titre : Gestion de la consommation basée sur l'adaptation dynamique de la tension, fréquence et body bias pour les systèmes sur puce en technologie FD-SOI

Résumé

Au-delà du nœud technologique CMOS BULK 28nm, certaines limites ont été atteintes dans l'amélioration des performances en raison notamment d'une consommation énergétique devenant trop importante. C'est une des raisons pour lesquelles de nouvelles technologies ont été développées, notamment celles basées sur Silicium sur Isolant (SOI). Par ailleurs, la généralisation des architectures complexes de type multi-cœurs, accentue le problème de gestion de la consommation à grain-fin. Les technologies CMOS FD-SOI offrent de nouvelles opportunités pour la gestion de la consommation en permettant d'ajuster, outre les paramètres usuels que sont la tension d'alimentation et la fréquence d'horloge, la tension de body bias. C'est dans ce contexte que ce travail étudie les nouvelles possibilités offertes et explore des solutions innovantes de gestion dynamique de la tension d'alimentation, fréquence d'horloge et tension de body bias afin d'optimiser la consommation énergétique des systèmes sur puce.

L'ensemble des paramètres tensions/fréquence permettent une multitude de points de fonctionnement, qui doivent satisfaire des contraintes de fonctionnalité et de performance. Ce travail s'intéresse donc dans un premier temps à une problématique de conception, en proposant une méthode d'optimisation du placement de ces points de fonctionnement. Une solution analytique permettant de maximiser le gain en consommation apporté par l'utilisation de plusieurs points de fonctionnement est proposée. La deuxième contribution importante de cette thèse concerne la gestion dynamique de la tension d'alimentation, de la fréquence et de la tension de body bias, permettant d'optimiser l'efficacité énergétique en se basant sur le concept de convexité. La validation expérimentale des méthodes proposées s'appuie sur des échantillons de circuits réels, et montre des gains en consommation moyens allant jusqu'à 35%.

Mots-Clés : DVFS, Adaptation Dynamique de Body Bias, FD-SOI, Gestion de la consommation, Optimisation de la consommation, Sélection de Power Mode

Title : Power Management based on Dynamic Voltage, Frequency and Body Bias Scaling on System On Chip in FD-SOI technology

Abstract

Beyond 28nm CMOS BULK technology node, some limits have been reached in terms of performance improvements. This is mainly due to the increasing power consumption. This is one of the reasons why new technologies have been developed, including those based on Silicon-On-Insulator (SOI). Moreover, the standardization of complex architectures such as multi-core architectures emphasizes the problem of power management at fine-grain. FD-SOI technologies offer new power management opportunities by adjusting, in addition to the usual parameters such as supply voltage and clock frequency, the body bias voltage. In this context, this work explores new opportunities and searches novel solutions for dynamically manage supply voltage, clock frequency and body bias voltage in order to optimize the power consumption of System on Chip.

Adjusting supply voltage, frequency and body bias parameters allows multiple operating points, which must satisfy the constraints of functionality and performance. This work focuses initially at design time, proposing a method to optimize the placement of these operating points. An analytical solution to maximize power savings achieved through the use of several operating points is provided. The second important contribution of this work is a method based on convexity concept to dynamically manage the supply voltage, the frequency and the body bias voltage so as to optimize the energy efficiency. The experimental results based on real circuits show average power savings reaching 35%.

Key-Words : DVFS, Adaptive Body Biasing, FD-SOI, Power Management, Power Optimization, Power Mode Selection.

Laboratoires : LIRMM - 161 rue Ada, 34392 Montpellier Cedex 5, France
CEA-Leti, Minatec Campus - 17 rue des Martyrs,
38054 Grenoble Cedex 9, France
