



Etude et intégration d'un circuit analogique, basse consommation et à faible surface d'empreinte, de neurone impulsif basé sur l'utilisation du BIMOS en technologie 28 nm FD-SOI

Thomas Bedecarrats

► To cite this version:

Thomas Bedecarrats. Etude et intégration d'un circuit analogique, basse consommation et à faible surface d'empreinte, de neurone impulsif basé sur l'utilisation du BIMOS en technologie 28 nm FD-SOI. Micro et nanotechnologies/Microélectronique. Université Grenoble Alpes, 2019. Français. NNT : 2019GREAT045 . tel-02520926v2

HAL Id: tel-02520926

<https://theses.hal.science/tel-02520926v2>

Submitted on 15 May 2020

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 DE LA COMMUNAUTE UNIVERSITE
GRENOBLE ALPES**

Spécialité : **Nano Electronique et Nano Technologies**

Arrêté ministériel : 25 mai 2016

Présentée par

Thomas BÉDÉCARRATS

Thèse dirigée par **Philippe GALY**
et codirigée par **Sorin CRISTOLOVEANU, Claire FENOUILLET-
BÉRANGER**

préparée au sein des **Laboratoires IMEP-LAHC et CEA-LETI**
dans l'**École Doctorale Electronique, Electrotechnique,
Automatique et Traitement du Signal (EEATS)**

**Etude et intégration d'un circuit
analogique basse consommation et à
faible surface d'empreinte de neurone
impulsionnel basé sur l'utilisation du
BIMOS en technologie 28 nm FD-SOI**

Thèse soutenue publiquement le « **4 septembre 2019** »,
devant le jury composé de :

Madame Lorena ANGHEL

Professeure, Université de Grenoble Alpes, Présidente du jury

Monsieur François DANNEVILLE

Professeur, Université de Lille, Rapporteur

Monsieur Damien QUERLIOZ

Chargé de recherches, HDR, Université de Paris-Saclay, Rapporteur

Monsieur Salvador MIR

Directeur de recherche, CNRS, Examineur

Monsieur Sylvain SAIGHI

Maître de conférences, HDR, Université Bordeaux, Examineur

Monsieur Philippe GALY

Directeur technique et professeur associé, HDR, STMicroelectronics et
Université de Sherbrooke, Directeur de thèse

Madame Claire FENOUILLET-BÉRANGER

Ingénieure et chercheuse, CEA-LETI, Co-encadrant de thèse, invité

Monsieur Sorin CRISTOLOVEANU

Directeur de recherche émérite CNRS, Co-encadrant de thèse, invité



Résumé

Avec la fin annoncée de la loi de Moore, les acteurs de la microélectronique cherchent de nouveaux paradigmes sur lesquels s'appuyer pour alimenter les développements futurs de notre société de l'information. En s'inspirant des systèmes nerveux biologiques, l'ingénierie neuromorphique offre des perspectives nouvelles qui révolutionnent d'ores et déjà l'intelligence artificielle. Pour que leurs performances permettent leur généralisation, les processeurs neuronaux se doivent d'intégrer des circuits de neurones les plus petits et les moins énergivores possible afin que les réseaux de neurones artificiels qu'ils implémentent atteignent une taille critique. Dans ce travail, nous montrons qu'il est possible de réduire le nombre de composants nécessaires à la conception d'un circuit analogique de neurone impulsif par la fonctionnalisation des courants de génération parasites dans un transistor BIMOS intégré en technologie 28 nm FD-SOI et dimensionné aux tailles minimales autorisées par la technologie. Après une caractérisation systématique de ces courants par des mesures quasi-statiques du FD-SOI BIMOS à température ambiante sous différentes polarisations, une modélisation compacte de ce composant adaptée à partir du modèle CEA-LETI UTSOI est proposée. Le circuit analogique de neurone impulsif à fuite, intégration et déclenchement basé sur le BIMOS (« BIMOS-based leaky, integrate-and-fire spiking neuron » : BB-LIF SN) est ensuite décrit. L'influence des différentes dimensions caractéristiques et polarisations de contrôle sur son fonctionnement observé lors des mesures sur des démonstrateurs fabriqués sur silicium est expliquée en détail. Un modèle analytique simple de ses limites de fonctionnement est proposé. La cohérence entre les résultats de mesures, ceux de simulations compactes et les prédictions du modèle analytique simple atteste la pertinence des analyses proposées. Dans sa version la plus aboutie, le BB-LIF SN occupe une surface de $15 \mu\text{m}^2$, consomme autour de 2 pJ/spike, fonctionne à des fréquences de déclenchement comprises entre 3 et 75 kHz pour des courant synaptique compris entre 600 pA et 25 nA et doit être alimenté par une tension de 3 V.

Abstract

As Moore's law reaches its limits, microelectronics actors are looking for new paradigms to ensure future developments of our information society. Inspired by biologic nervous systems, neuromorphic engineering is providing new perspectives which have already enabled breakthroughs in artificial intelligence. To mainstream neural processors, they will require better performance by integrating smaller and more energy efficient neuron circuits so that artificial neuron networks reach a critical size. My research shows that it is possible to reduce the number of components necessary to design an analogue spiking neuron circuit thanks to the functionalisation of parasitic generation currents in a BIMOS transistor integrated in 28 nm FD-SOI technology and using the minimum dimensions allowed by this technology. After a systematic characterization of the FD-SOI BIMOS currents under several biases through quasi-static measurements at room temperature, a compact model of this component, adapted from the CEA-LETI UTSOI is proposed. The BIMOS-based leaky, integrate-and-fire spiking neuron (BB-LIF SN) circuit is described. Influence of the different design and bias parameters on its behaviour is explained in detail based on experiments using a silicon-based demonstrator. A simple analytic model of its operating boundaries is proposed. The coherence between measurement and compact simulation results and predictions coming from the simple analytic model attests to the relevance of the proposed analysis. The optimal BB-LIF SN circuit tested measured $15 \mu\text{m}^2$, which consumed around 2 pJ/spike, triggered at a rate between 3 and 75 kHz for 600 pA to 25 nA synaptic currents and needed a 3 V power supply.

*À mes frères
Florent et Adrien,
À ma sœur Nadia,
À mon oncle Jean.*

Remerciements

Les travaux présentés dans ce manuscrit n'auraient pu être réalisés sans le soutien, l'aide et l'intervention de très nombreuses personnes que je souhaite, avant toutes choses, remercier dans ces quelques lignes.

Le premier de mes remerciements va tout naturellement à mon directeur de thèse Philippe Galy pour avoir proposé le sujet passionnant de cette thèse et pour avoir soutenu ma candidature lors du recrutement. Je le remercie pour son accompagnement bienveillant au fil des jours, pour les nombreuses discussions scientifiques qui ont jalonné les trois années passées à STMicroelectronics, pour le partage et la transmission de son état d'esprit et de sa créativité dans le travail et pour la confiance qu'il a su m'accorder.

Je souhaite ensuite remercier mes deux encadrants, Claire Fenouillet-Béranger et Sorin Cristoloveanu, pour le partage de leur expertise, pour les précieux conseils qu'ils m'ont donnés et pour leur bienveillance, leur soutien et leur disponibilité. Grâce à leur ouverture d'esprit, j'ai pu notamment profiter pleinement de l'environnement et des ressources humaines et matérielles de leurs deux laboratoires de rattachement respectifs.

Je remercie les membres du jury, la présidente Lorena Anghel, les rapporteurs François Danneville et Damien Querlioz et les examinateurs Salvador Mir et Sylvain Saïghi, pour avoir accepté d'y participer et d'apporter leur regards critiques et constructifs pour garantir l'exigence à laquelle mon travail doit répondre.

Je remercie également les membres de mon comité de suivi de thèse Edith Beigne, Elisa Vianello et Laurent Grenouillet, tous du CEA-LETI, pour leur conseils avisés et pertinents lors des réunions de suivi annuel.

Je remercie Martine Gri et Xavier Mescot de l'IMEP-LAHC et Charles-Alexandre Legrand et Ruddy Costanzi de STMicroelectronics pour leur aide indispensable lors des campagnes de mesures effectuées dans les laboratoires de leurs structures respectives.

Je remercie Thierry Poiroux et Olivier Roseaux du CEA-LETI sans qui les travaux de modélisation compacte n'auraient pas été possibles.

J'ai eu la chance d'effectuer ma thèse au sein d'une grande entreprise technologique qui a su s'entourer de collaborateurs à très haut niveau de compétence et d'expertise, qui, malgré leur intense charge de travail, savent se rendre disponibles pour les besoins des doctorants. Nombre d'entre eux m'ont apporté un support décisif dans l'avancement de mes travaux. Je remercie ainsi particulièrement Charlotte Beylier et Christian Gardin pour nos nombreux échanges sur les contraintes photolithographiques lors de la conception des masques. Je remercie Jean-Pierre Blanc, Philippe Cathelin, Sylvain Clerc et Thierry Di Gilio pour leurs conseils sur la conception de circuits. Je remercie Franck Arnaud, Olivier Weber, Philippe Boivin, Simon Jeannot, Benjamin Dumont et Bertand Le Gratiet pour nos échanges à propos du procédé de fabrication en technologie 28 FD-SOI. Je remercie Clément Chabruillet pour son support en gestion de l'environnement informatique, Frédéric Monsieur pour celui en simulations 3D TCAD, Raffaele Bianchini pour avoir garanti la conformité des échantillons envoyés en fabrication et Andrea Cathelin pour l'accompagnement qu'elle offre aux doctorants.

Je remercie les doctorants que j'ai côtoyés pendant ma thèse avec lesquels j'ai pu partager mes expériences : Sotirios Athanasious, Renan Lethiecq, Mohammed Tmimi, Valérian Cincon, Geoffrey Delahaye et Ioanna Kriekouki de l'équipe dirigée par Philippe Galy ; Joris Jourdon, Mathilde Lemang,

Raphaël Guillaume, Florian Voineau, Matteo Causo, Hanni Sherry, Hani Malloug, Reda Kasri, Hanae Zegmout, Hassan El-Dirhani, Florent Torres de STMicroelectronics ; Lina Kadura, Léo Bourdet, Giulia Usai et Daphnée Bosc du CEA-LETI ; Kyunghwa Lee, Hyungjin Park et Louison Legrand de l'IMEP-LAHC.

Un remerciement tout particulier va à ma collègue et amie Louise De Conti qui a commencé sa thèse en même temps que moi, et qui a été à mes côtés pendant ces trois ans. C'est une véritable chance de vivre l'aventure si particulière de préparer une thèse de doctorat accompagné d'une alter-ego avec qui partager ses joies et ses peines, ses doutes et ses succès. Merci Louise pour ton écoute, tes encouragements, ta spontanéité et ta sincérité au quotidien.

Enfin, je remercie ma famille et mes proches, qui m'ont soutenu et encouragé dans ma reprise d'études scientifiques, commencée en 2012 et qui se termine aujourd'hui par la défense de cette thèse.

Table des matières

Résumé	2
Abstract	2
Remerciements	4
Table des figures.....	11
Table des tableaux.....	19
Glossaire des abréviations, symboles et acronymes.....	20
Introduction.....	25
Chapitre 1 : Contexte et État de l’art	28
1.1 Le neuromorphisme	28
1.1.1 Système nerveux et neurone biologique	28
1.1.2 L’ingénierie neuromorphique.....	29
1.2 Les réseaux de neurones artificiels	30
1.2.1 Les modèles de neurone	31
1.2.1.1 Le neurone formel	31
1.2.1.2 Le neurone impulsionnel.....	31
1.2.2 Topologies de réseaux neuronaux	34
1.2.2.1 Les FNN	34
1.2.2.2 Les RNN.....	35
1.2.3 Les règles d’apprentissage	35
1.2.3.1 Apprentissage dans les réseaux de neurones formels	36
1.2.3.2 Apprentissage dans les réseaux de neurones impulsionnels	37
1.3 Implémentations matérielles des ANN	38
1.3.1 Les GPU pour les ANN formels	38
1.3.2 Les processeurs « event-based » pour les SNN.....	38
1.3.2.1 Les systèmes de communication AER	39
1.3.2.2 Les circuits de synapses.....	40
1.3.2.3 Les circuits de neurones	42
1.4 État de l’art et figure de mérite des circuits analogiques de neurone impulsionnel	44
1.4.1 État de l’art	44
1.4.2 Les figures de mérite	44
1.4.2.1 Surface de l’empreinte du circuit	45
1.4.2.2 L’efficacité énergétique.....	45
1.4.2.3 Autres caractéristiques.....	45
1.5 Les objectifs de la thèse	46

1.6	Conclusion	47
1.7	Références	49
Chapitre 2 : Notions de physique du semiconducteur, de physique des composants et technologie 28 nm UTBB FD-SOI		56
2.1	Notion de physique du semiconducteur	56
2.1.1	Théorie des bandes d'énergie	56
2.1.2	Transport électronique et mécanismes de génération-recombinaison.....	58
2.2	Notions de physique du composant micro-électronique.....	60
2.2.1	La jonction PN.....	60
2.2.2	Le transistor bipolaire (« bipolar junction transistor » : BJT)	64
2.2.3	La capacité métal-oxyde-semiconducteur (MOSCAP).....	67
2.2.3.1	Notion de potentiel de surface et de tension de bande plate	68
2.2.3.2	Régimes de la MOSCAP	68
2.2.3.3	Caractéristique C-V de la MOSCAP	70
2.2.4	Le transistor à effet de champ métal-oxyde-semiconducteur (MOSFET)	71
2.3	La technologie 28nm FD-SOI	73
2.3.1	« Downscaling » et effet de canaux courts	73
2.3.2	Les technologies SOI.....	74
2.3.2.1	Les transistors PD-SOI.....	76
2.3.2.2	Les transistors FD-SOI	77
2.3.2.3	Les technologies UTBB FD-SOI.....	77
2.3.2.4	La technologie 28nm UTBB FD-SOI.....	78
2.4	Références	80
Chapitre 3 : Le 28 nm FD-SOI BIMOS : Caractérisation et modélisation		81
3.1	Principe général du BIMOS.....	81
3.2	Topologie et dimensions du BIMOS de référence.....	82
3.3	Caractérisation du FD-SOI BIMOS en mode « gated-diode »	83
3.3.1	Courbe I_A - V_A avec V_{FG} comme paramètre.....	84
3.3.2	Courbes I_A - V_{FG} avec V_{A-K} comme paramètre	87
3.3.3	Courbes I_A - V_A avec $V_{FG} = V_A$	89
3.4	Caractérisation du FD-SOI BIMOS en mode MOSFET.....	90
3.4.1	Courbe I_D - V_{FG}	91
3.4.1.1	Courbe I_D - V_{FG} à $V_B = 0$ V	91
3.4.1.2	Courbe I_D - V_{FG} lorsque le body-contact est flottant	94
3.4.1.3	Influence de V_B sur les courbes I_D - V_{FG}	95
3.4.2	Courbes I_D - V_D	99

3.4.2.1	Courbes I_D - V_D à $V_B = 0$ V	99
3.4.2.2	Courbe I_D - V_D à $I_B = 0$ A.....	101
3.5	Caractérisation du BIMOS en mode BJT en émetteur commun	103
3.5.1	Régime de fonctionnement normal altéré.....	106
3.5.1.1	Courants de collecteur et de base.....	106
3.5.1.2	Gain en courant β_F	106
3.5.2	Régime de fonctionnement anormal	107
3.5.2.1	Courants de collecteur et de base.....	107
3.5.2.2	Gain en courant β_F	108
3.5.2.3	Caractérisation du BJT parasite de la structure MOSFET	108
3.5.3	Résumé	109
3.6	Implémentation dans le modèle UTSOI	109
3.6.1	Le modèle UTSOI	109
3.6.2	Ajout du body-contact au modèle UTSOI.....	110
3.6.2.1	Courants de génération par BTBT et II	110
3.6.2.2	Courant de gated-diode	112
3.6.2.3	Courant de BJT.....	112
3.6.2.4	Modèle UTSOI modifié	114
3.7	Conclusion	114
3.8	Références.....	115
Chapitre 4 : Le circuit de neurone impulsionnel basé sur le BIMOS : « BIMOS-based spiking neuron circuit »		117
4.1	Potentiel de membrane et potentiel d'action des neurones biologiques	117
4.2	Le modèle de neurone « leaky, Integrate and Fire »	118
4.3	Le circuit de déclenchement à fuite basé sur le BIMOS (BB-LTC)	120
4.3.1	Fonction de bascule de Schmitt du BB-LTC	122
4.3.2	Fonction de conductance du BB-LTC.....	124
4.3.3	Fonction d'interrupteur du BB-LTC	125
4.3.4	Principes de fonctionnement du BB-LTC.....	126
4.3.5	Etudes de l'influence des paramètres de polarisation et de design sur le comportement quasi-statique du BB-LTC.....	132
4.3.5.1	Influence de V_{FG}^N sur la caractéristique I_C - V_{MEM} du BB-LTC.....	132
4.3.5.2	Influence de la largeur W_1 de BIMOS sur la caractéristique I_C - V_{MEM} du BB-LTC.....	135
4.3.5.3	Influence de V_{BG}^B sur la caractéristique I_C - V_{MEM} du BB-LTC.....	137
4.3.5.4	Influence de l'épaisseur d'oxyde sur la caractéristique I_C - V_{MEM} du BB-LTC.....	138
4.3.6	Confrontation du modèle SPICE aux mesures quasi-statiques	140

4.3.7	Comportement dynamique du BB-LTC.....	142
4.3.7.1	Influence du temps de montée de V_{MEM} sur V_{TH}	142
4.3.7.2	Durée de déclenchement τ_{ON}	143
4.3.7.3	Résumé	144
4.4	Le neurone impulsionnel basé sur le BIMOS.....	145
4.4.1	Preuve de concept du fonctionnement du BB-LIF SN	145
4.4.1.1	Régime de comportement fonctionnel du BB-LIF SN.....	146
4.4.1.2	Régime de comportement non-fonctionnel du BB-LIF SN	148
4.4.2	Influence du potentiel V_{FG}^N sur le fonctionnement du BB-LIF SN	149
4.4.2.1	Régime de comportement fonctionnel du BB-LIF SN.....	150
4.4.2.2	Régime de comportement non-fonctionnel du BB-LIF SN	150
4.4.2.3	Résumé	151
4.4.3	Influence du potentiel V_{BG}^B sur le fonctionnement du BB-LIF SN	151
4.4.4	Confrontation du modèle SPICE aux mesures $V_{MEM} = f(t)$	153
4.4.4.1	Comportement global	153
4.4.4.2	Potentiel d'axone V_{AXO}	153
4.4.4.3	Bilan	155
4.4.5	Modélisation des limites de fonctionnement du BB-LTC.....	155
4.4.5.1	Limites de fonctionnement inférieures $I_{SYN,MIN}$ et f_{MIN}	155
4.4.5.2	Limites de fonctionnement supérieures $I_{SYN,MAX}$ et f_{MAX}	156
4.4.5.3	Application au circuit mesuré.....	156
4.4.5.4	Limites du modèle	157
4.5	Intégration du BB-LIF SN en 28nm FD-SOI	157
4.5.1	Dimensionnement du BB-LTC intégré	158
4.5.2	Dimensionnement du buffer intégré.....	158
4.5.3	Dimensionnement de la capacité C_{MEM}	159
4.5.4	Layout du BB-LIF SN.....	160
4.5.5	Evaluation des limites de fonctionnement en fonction de V_{FG}^N	160
4.5.6	Mesure de la réponse temporelle du BB-LIF SN intégré	162
4.5.6.1	Dispositif de mesure	162
4.5.6.2	Configuration des polarisations.....	163
4.5.6.3	Fréquence de déclenchement f en fonction du potentiel V_{IN}	163
4.5.6.4	Fréquence de déclenchement f en fonction du courant I_{SYN}	164
4.5.6.5	Consommation énergétique du BB-LIF SN	165
4.5.7	Bilan	166
4.6	Conclusion	166

4.7	Références	169
	Conclusion	170
	Liste des brevets et publications	174
	Brevets.....	174
	Papiers de journal.....	174
	Papiers de conférence.....	174

Table des figures

Figure 1-1 – Schéma descriptif (a) d'un neurone biologique et (b) d'une synapse biologique (dans (a), les parties grisées représentent les dendrites et terminaisons axonales d'autres neurones).	28
Figure 1-2 – (A) Comportement d'une population de canaux ioniques actifs dans une membrane de neurone. (B) Courant de saturation d'un transistor MOS en fonction de la tension de grille. D'après [3].	30
Figure 1-3 – Photographie du perceptron MARK I construit par Rosenblatt, provenant de [7].	30
Figure 1-4 – Schéma descriptif du modèle mathématique de neurone formel.	31
Figure 1-5 – Résumé des comportements identifiés dont sont capables les SN biologique. Depuis [13].	32
Figure 1-6 – Comparaison des comportements que sont capables de reproduire les modèles de neurone et du nombre d'opération à effectuer pour les simuler à chaque pas de temps. Depuis [13].	33
Figure 1-7 – Schéma illustrant la topologie du perceptron multi couches à trois couches et fonctions mathématiques correspondantes.	35
Figure 1-8 – Variation du poids synaptique Δw en fonction de l'intervalle de temps $\Delta \tau$ qui caractérise la STDP. $\Delta \tau < 0$ conduit à une dépréciation à long terme (« long term depression » : LTD), $\Delta \tau > 0$ conduit à une appréciation à long terme (« long term potentiation » : LTP)	37
Figure 1-9 – Schéma illustrant le système AER.	40
Figure 1-10 – (a) Circuit Axon-Hillock et (b) chronogramme correspondant, depuis [87] et d'après [1].	42
Figure 1-11 – Liste de modèles de SN dont on trouve une implémentation matérielle dans la littérature. La taille des boîtes correspond au nombre d'articles publiés et leur couleur à la famille de modèle à laquelle ils appartiennent. Depuis [71].	43
Figure 2-1 – Schéma d'une jonction PN à l'équilibre thermodynamique ($V_{A-K} = 0$). Profil de densité de charge, de champ électrique et de potentiel électrique correspondant.	61
Figure 2-2 – Illustration des mécanismes de BTBT et TAT dans une jonction PN polarisée en inverse	63
Figure 2-3 – Comparaison des courants typiques de diode idéale et de diode incluant le phénomène de génération SRH, depuis [5].	63
Figure 2-4 - Vue schématique de la topologie d'un BJT NPN et d'un BJT PNP	64
Figure 2-5 – Exemple typique du profil de dopants d'un BJT NPN ; d'après [5]	64
Figure 2-6 – (a) Vue schématique des flux de porteurs dans un BJT NPN en fonctionnement et (b) schéma de bande associé, d'après [4], [5].	65
Figure 2-7 – (a) Diagramme de Gummel et (b) gain en courant en émetteur commun β_F en fonction du courant de collecteur d'un BJT typique, d'après [4]. $V_{C-E} = 3 \text{ V}$	66
Figure 2-8 – Vue schématique de deux MOSCAP de type P et N	67
Figure 2-9 – Diagramme de bande d'une P-MOSCAP en régime (a) d'accumulation, (b) de bande-plate, (c) de déplétion, (d) d'inversion faible et (e) d'inversion forte.	68
Figure 2-10 – Tendence de la courbe C-V d'une capacité P-MOSCAP (d'après [4]). V_G correspond à V_{M-SC}	70

Figure 2-11 – Vue schématique d’un N-MOSFET et d’un P-MOSFET	71
Figure 2-12 – Caractéristique typique I_D-V_D d’un N-MOSFET en régime au-dessus du seuil [4]	72
Figure 2-13 – Caractéristique typique I_D-V_G d’un N-MOSFET, en échelle semi-logarithmique (axe de gauche) et linéaire (axe de droite) [4]	73
Figure 2-14 – Dépendance de la tension de seuil en fonction de la longueur de grille et de la polarisation du drain [12]	74
Figure 2-15 – A : Vue en coupe montrant le chemin de latch-up dans un inverseur CMOS bulk. B : Vue en coupe d’un inverseur CMOS SOI. Les capacités parasites de drain sont aussi représentées. D’après [14].	75
Figure 2-16 – Vue en coupe d’un inverseur CMOS en A : bulk, B : PD-SOI et C : FD-SOI. D’après [14].	75
Figure 2-17 – Diagramme de bande d’énergie en (A) bulk, (B) PD-SOI et (C) FD-SOI. Tous les composants sont représentés au seuil ($V_{FG} = V_{T,MOS}$). Les zones grisées représentent les zones déplétées. Les composants SOI sont représentés en situation d’inversion faible (et donc, sous le seuil) à l’interface arrière. D’après [14].	75
Figure 2-18 – Réseau de capacité équivalent dans un transistor FD-SOI en l’absence de couche d’inversion dans le canal avant ou arrière (transistor sous le seuil) et en considérant la charge de la zone de déplétion du body négligeable.	77
Figure 2-19 – Vue schématique en coupe d’un transistor en technologie 28 nm UTBB FD-SOI.....	79
Figure 3-1 – Schéma de principe du BIMOS.	81
Figure 3-2 – Topologie et dimensions du 28nm FD-SOI BIMOS de référence.	82
Figure 3-3 – Mise en évidence des structures “gated-diode” dans le schéma du FD-SOI BIMOS. (a) Vue de dessus et (b) vue en coupe.	83
Figure 3-4 – (a) Vue de dessus du BIMOS connecté en mode gated diode et (b) schéma correspondant. A : anode, BC : body contact, BG : grille-arrière, D : drain, FG : grille-avant, K : cathode, S : source.....	83
Figure 3-5 – Mesure, sur un BIMOS connecté en mode gated-diode, du courant d’anode I_A en fonction de la tension d’anode V_A avec la tension de grille-avant V_{FG} comme paramètre, en échelle (a) semi-logarithmique et (b) linéaire. $V_{BG} = V_K = 0$ V.	86
Figure 3-6 – (a) Tension de transition $V_{T,PN}$ en fonction du potentiel de grille-avant V_{FG} . (b) Facteur d’idéalité η en fonction de la tension V_{A-K} pour différentes valeur de potentiel de grille-avant V_{FG}	86
Figure 3-7 – Courant d’anode I_A en fonction de la tension grille-avant – anode V_{FG-A} pour différent valeurs de potentiel de grille V_{FG} . $V_{BG} = V_K = 0$ V.	87
Figure 3-8 – Simulation de la concentration verticale d’électrons et de trous dans le body d’une gated-diode sur SOI ($W = L = 3\mu m$, $V_{A-K} = 0.3$ V, densité de dopants du body : $N_A = 10^{17} cm^{-3}$) lorsque $n=p$ à l’interface body/BOX et pour une interface body/oxyde de grille déplétée (D), accumulée (A) ou inversée (I), d’après [18]. Ces différentes courbes sont obtenues en jouant sur les potentiels des grilles avant et arrière dont les valeurs ne sont pas explicitées.	88
Figure 3-9 – (a) Mesures à température ambiante du courant d’anode I_A en fonction du potentiel de grille-avant V_{FG} sous différentes tensions d’anode-cathode V_{A-K} dans le BIMOS de référence connecté en mode gated-diode et en régime de polarisation légèrement en direct ($0.35 < V_{A-K} < 0.7$ V). Le potentiel de grille-arrière est pris comme référence. L’anode et la cathode sont polarisées à $\pm V_{A-K}/2$ respectivement. (b) Schéma électrique correspondant.	88

- Figure 3-10 – (a) Mesure effectuée sur le BIMOS de référence et à température ambiante (ligne pleine) et simulation selon le modèle présenté à la figure (d) (ligne en pointillé) du courant d'anode I_A en fonction de la tension d'anode V_{A-K} sous la condition $V_A = V_{FG}$, en échelle semi-logarithmique (axe de gauche) et linéaire (axe de droite). $V_K = V_{BG} = 0$ V. (b) Schéma électrique correspondant. (c) Facteur d'idéalité η (axe de gauche), dérivée de la conductance $\partial^2 I_A / \partial V_{A-K}^2$ (axe de droite) et tension de transition $V_{T,PN}$ extraits de la mesure présenté à la figure (a). (d) Schéma électrique et paramètres du modèle utilisé à la figure (a). Les deux diodes sont décrites par l'équation (2.35)..... 90
- Figure 3-11 – (a) Vue de-dessus du BIMOS et du NMOS et longueurs évoquées au Tableau 3-2. (b) Schéma électrique correspondant..... 91
- Figure 3-12 – (a et b) Courants de drain I_D circulant dans le BIMOS de référence et dans un NMOS de type et de dimensions comparables en fonction du potentiel de la grille-avant V_{FG} lors d'un fonctionnement des composant en régime (a) linéaire ($V_D = 0.2$ V) et (b) saturé ($V_D = 1.8$ V). (c) Courant de source et (d) courant de body-contact du BIMOS de référence en fonction de la grille-avant V_{FG} en régime linéaire et saturé. Les données qui concernent le BIMOS sont issues de mesures à température ambiante et celles qui concernent le NMOS sont issues de simulations SPICE calibrées. Dans les deux expériences, $V_S = V_{BG} = V_B = 0$ V. 92
- Figure 3-13 – Schémas électrique qui permettent d'obtenir un body-contact flottant : (a) sans connecter le body contact (b) en contraignant le courant $I_B = 0$ à l'aide d'une source de courant. Cela permet de mesurer le potentiel. 94
- Figure 3-14 – (a) Courants de drain du BIMOS de référence en fonction du potentiel de grille-avant V_{FG} , en régime saturé ($V_D = 1.8$ V), sous les conditions $V_B = 0$, $I_B = 0$ et body-contact flottant, et courant de drain du NMOS de dimensions comparables et sous les mêmes polarisations. (b) Courbes de (a) en échelle linéaire pour $V_{FG} > 0.5$ V. (b) Agrandissement de (a) pour $V_{FG} < 0.1$ V. (d) Potentiel de body-contact du BIMOS de référence en fonction du potentiel de grille-avant V_{FG} , en régime saturé ($V_D = 1.8$ V) et sous la condition $I_B = 0$ 95
- Figure 3-15 – Mesures à température ambiante des courants de drain I_D , de source I_S et de body-contact I_B en fonction du potentiel de grille-avant V_{FG} du FD-SOI BIMOS décrit au Tableau 3-2, en régime linéaire ($V_D = 0.20$ V) et pour un potentiel de body-contact V_B entre 0 V et 1 V, par pas de 0.25 V ((a) à (e)). $V_S = 0$ V et $V_{BG} = 0$ V. 96
- Figure 3-16 – Mesures à température ambiante des courants de drain I_D , de source I_S et de body-contact I_B en fonction du potentiel de grille-avant V_{FG} du FD-SOI BIMOS décrit au Tableau 3-2, en régime saturé ($V_D = 1.80$ V) et pour un potentiel de body-contact V_B entre 0 V et 1 V, par pas de 0.25 V ((a) à (e)). $V_S = 0$ V et $V_{BG} = 0$ V. 98
- Figure 3-17 – (a) Courants de drain I_D du BIMOS de référence et du NMOS aux dimensions comparables (Tableau 3-2) en fonction du potentiel de drain V_D et à différents $V_{FG} < V_{T,MOS}$. (b) Courant de de body-contact $-I_B$ du BIMOS de référence en fonction du potentiel de drain V_D et à différents $V_{FG} < V_{T,MOS}$. $V_B = 0$ V, $V_S = 0$ V et $V_{BG} = 0$ V. Les données qui concernent le BIMOS sont obtenues par mesure sous pointe à température ambiante. Celles qui concernent le NMOS par simulation SPICE calibrée. 100
- Figure 3-18 – (a) Courants de drain I_D du BIMOS de référence et du NMOS aux dimensions comparables (Tableau 3-2) en fonction du potentiel de drain V_D et à différents $V_{FG} > V_{T,MOS}$. (b) Courant de de body-contact $-I_B$ du BIMOS de référence en fonction du potentiel de drain V_D et à différents $V_{FG} > V_{T,MOS}$. $V_B = 0$ V, $V_S = 0$ V et $V_{BG} = 0$ V. Les données qui concernent le BIMOS sont obtenues par mesure sous pointe à température ambiante. Celles qui concernent le NMOS par simulation SPICE calibrée. 100

Figure 3-19 – Mesures du courant de drain I_D (a) et du potentiel de body-contact V_B (b) en fonction du potentiel de drain V_D , à différentes valeurs de potentiel de grille-avant V_{FG} . effectuées simultanément sur le BIMOS de référence (Tableau 3-2). $V_S = V_{BG} = 0$ V et $I_B = 0$ A.	101
Figure 3-20 – Courbes paramétrés du courant de body $I_{B,V0}$ mesuré lorsque $V_B = 0$ V, en fonction du potentiel de body $V_{B,10}$ mesuré lorsque $I_B = 0$ A. L'abscisse et l'ordonnée de chaque point des courbes proviennent de deux expériences distinctes mais pour lesquelles le couple $[V_D ; V_{FG}]$ est identique.	102
Figure 3-21 – Schéma électrique qui modélise la courbe $I_{B,V0}$ - $V_{B,10}$ de la Figure 3-20. Les fonctions de transfert de la résistance R et de la source de tension V_S ainsi que l'équation de la diode idéale sont reportées également.	103
Figure 3-22 – (a) Vue de dessus et (b) vue schématique du montage de la caractérisation du BIMOS en mode BJT émetteur-commun	104
Figure 3-23 – Mesure des courants de collecteur I_B , d'émetteur I_E et de base I_B en fonction de la tension base – émetteur V_{B-E} avec V_{C-E} comme paramètre, en échelle (a) semi-logarithmique et (b) linéaire. La dérivée de la conductance est également tracée ((b) axe de droite)	104
Figure 3-24 – Mesure des courants de collecteur I_B , d'émetteur I_E et de base I_B en fonction de la tension base – émetteur V_{B-E} avec V_{FG-E} comme paramètre, en échelle (a) semi-logarithmique et (b) linéaire. La dérivée de la conductance est également tracée ((b) axe de droite)	104
Figure 3-25 – Courant de collecteur en base ouverte I_{CE0} , courant de collecteur en base mise à la masse $I_{C,VB0} = I_C(V_B = 0)$ (axe de gauche) et rapport entre ces deux courants $I_{CE0}/I_{C,VB0}$ (axe de droite), en fonction des tensions (a) collecteur – émetteur et (b) grille-avant – émetteur.	105
Figure 3-26 – Gain β_F en fonction de V_{B-E} , (a) avec V_{C-E} comme paramètre et $V_{FG-E} = 0$ V, et (b) V_{FG-E} comme paramètre et $V_{C-E} = 1.8$ V. Données extraites de la Figure 3-23 et de la Figure 3-24.	105
Figure 3-27 – Gain β_F en fonction de I_C (a) avec V_{C-E} comme paramètre et $V_{FG-E} = 0$ V, et (b) V_{FG-E} comme paramètre et $V_{C-E} = 1.8$ V. Données extraites de la Figure 3-23 et de la Figure 3-24.	106
Figure 3-28 – Schéma électrique des courants de sortie calculés dans le modèle UTSOI [36]. La source est notée S, le drain D, la grille-avant G et la grille-arrière B.	110
Figure 3-29 – Courants de body-contact mesurés sur le BIMOS de référence (lignes continues) et simulés par modélisation compacte selon les équations (3.9) et (3.10) et le jeu de paramètres $(a_1 ; a_2 ; a_3 ; a_4 ; K_{a2}) = (100 ; 18 ; 1 ; 0 ; 0)$ (lignes pointillées) en fonction du potentiel de grille-avant V_{FG} et avec le potentiel drain V_D pris comme paramètre. Les autres terminaux sont mis à la masse.	111
Figure 3-30 – Schéma électrique et paramètres utilisés pour modéliser les gated-diodes body-contact – drain et body-contact – source.	112
Figure 3-31 – Gain en courant en mode émetteur commun β_F en fonction de la tension base – émetteur V_{B-E} et pour une tension émetteur – collecteur V_{C-E} entre 0.2 V et 2.2 V, extrait à partir de mesures effectuées à température ambiante sur le BIMOS de référence, et modèle de $\beta_F = f(V_{B-E})$ selon l'équation (3.13)	113
Figure 3-32 – Schéma électrique des courants de sortie calculés dans le modèle UTSOI modifié [36]. La source est notée S, le drain D, la grille-avant G, la grille-arrière B et le body-contact BC.	113
Figure 4-1 – Schéma d'un neurone. Les éléments grisés sont les synapses et dendrites d'autres neurones. D'après [1].	117

Figure 4-2 – Vue schématique de la forme du potentiel de membrane au cours du temps sous l’effet d’activités synaptiques qui conduisent ou non au déclenchement du potentiel d’action.	118
Figure 4-3 – (a) Schéma électrique et équation conditionnelle correspondants au modèle de neurone LIF. (b) Chronogrammes d’un courant I_{SYN} constant et du potentiel V_{MEM} correspondant prédit par le modèle LIF. (c) Chronogrammes d’un courant I_{SYN} sinusoïdal et du potentiel V_{MEM} correspondant prédit par le modèle LIF.	119
Figure 4-4 – Vues schématiques du potentiel de membrane au cours du temps (a) d’un neurone biologique et (b) d’un neurone simulé par le modèle LIF, dans le cas d’une excitation synaptique constante.	119
Figure 4-5 – (a) Schéma du circuit qui implémente entièrement le modèle LIF. (b) Chronogrammes du courant I_{SYN} et des potentiels V_{MEM} et V_{AXO} , pour un courant I_{SYN} constant, obtenus selon le schéma exposé en (a).	120
Figure 4-6 – Schéma électrique d’un LTC utilisant une bascule de Schmitt typique implémentée en technologie CMOS.	120
Figure 4-7 – Schéma électrique d’un BB-LTC composé d’un FD-SOI BIMOS et d’un FD-SOI NMOS. ..	121
Figure 4-8 – Layout typique d’un BB-LTC avec les dimensions caractéristiques.....	121
Figure 4-9 – (a) Caractéristique I_C - V_{MEM} mesurée à température ambiante sur BB-LTC décrit au Tableau 4-1. $V_E = V_S = V_{\text{BG}} = V_{\text{FG}}^N = 0$ V. (b) Vue schématique de la caractéristique I_C - V_{MEM} où sont repérées les coordonnées des points de la fenêtre d’hystérésis.	122
Figure 4-10 – Mesures quasi-statiques à température ambiante, sur un BIMOS de référence, du courant de collecteur I_C en fonction du potentiel d’axone V_{AXO} , à différents potentiels de membrane V_{MEM} . $V_{\text{FG}} = V_B = V_{\text{AXO}}$, $V_C = V_{\text{MEM}}$ et $V_E = V_{\text{BG}} = 0$ V. (a) Plage de balayage entière. (b) Zoom sur la plage $V_{\text{MEM}} \in [0 ; 0.2]$ V.	123
Figure 4-11 – (a) Potentiel d’axone V_{AXO} en fonction du potentiel de membrane V_{MEM} dans le BB-LTC, extrait à partir des données I_C des Figure 4-9 et Figure 4-10. (b) Vue schématique de la caractéristique V_{AXO} - V_{MEM} , où sont repérées les coordonnées des points de la fenêtre d’hystérésis.	124
Figure 4-12 – Extraction de G_{LK} à partir des données de la Figure 4-9 et selon l’équation (4.3).	124
Figure 4-13 – Extraction de G_{ON} à partir des données de la Figure 4-9 et selon l’équation (4.4).	125
Figure 4-14 – (a) Décomposition de la courbe caractéristique I_C - V_{MEM} en 5 régions différentes. (b) Vue schématique du BB-LTC lorsque $V_{\text{MEM}} = 0$ V mettant en évidence la vue en coupe horizontale du film de silicium du FD-SOI BIMOS.....	126
Figure 4-15 – (a) Vue schématique de la courbe $I_C - V_{\text{MEM}}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 1 (ligne bleu). (b) Vue schématique des courants qui circulent dans le BB-LTC.....	127
Figure 4-16 – (a) Vue schématique de la courbe $I_C - V_{\text{MEM}}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 2 (ligne orange). (b) Vue schématique des courants qui circulent dans le BB-LTC.....	128
Figure 4-17 – (a) Vue schématique de la courbe $I_C - V_{\text{MEM}}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 3 (ligne grise). (b) Vue schématique des courants qui circulent dans le BB-LTC.....	128

- Figure 4-18 – (a) Vue schématique de la courbe $I_C - V_{MEM}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 4 (ligne jaune). (b) Vue schématique des courants qui circulent dans le BB-LTC. 129
- Figure 4-19 – (a) Vue schématique de la courbe $I_C - V_{MEM}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 5 (ligne bleue). (b) Vue schématique des courants qui circulent dans le BB-LTC. 130
- Figure 4-20 – (a) Caractéristiques $I_C - V_{MEM}$ pour différents potentiels $V_{FG}^N \geq -0.2$ V, mesurées à température ambiante sur un BB-LTC décrit au Tableau 4-1. $V_E = V_S = V_{BG} = 0$ V. (b) Extraction depuis (a) des grandeurs V_{TH} et V_{RST} en fonction de V_{FG}^N . (c) Extraction depuis (a) des grandeurs $I_{ON,TH}$, $I_{OFF,TH}$, $I_{ON,RST}$ et $I_{OFF,RST}$ en fonction de V_{FG}^N 133
- Figure 4-21 – (a) Caractéristiques $I_C - V_{MEM}$ pour différents potentiels $V_{FG}^N \leq -0.2$ V, mesurées à température ambiante sur un BB-LTC décrit au Tableau 4-1. $V_E = V_S = V_{BG} = 0$ V. (b) Extraction depuis (a) de V_{TH} en fonction de V_{FG}^N . (c) Extraction depuis (a) des grandeurs $I_{ON,TH}$, $I_{OFF,TH}$ en fonction de V_{FG}^N 134
- Figure 4-22 – (a) Caractéristique $I_C - V_{MEM}$ pour différentes largeurs W_1 , mesurée à température ambiante sur un BB-LTC à oxyde épais (EOT = 3.4 nm) de dimension $L_1 = L = 135$ nm, $W_2 = W = 350$ nm et $L_2 = 144$ nm. Le BIMOS et le NMOS sont de type RVT. $V_E = V_S = V_{BG} = V_{FG}^N = 0$ V. (b) Polarisation caractéristique en fonction de la largeur W_1 , extraites depuis (a). (c) Courants caractéristiques en fonction de la largeur W_1 , extraits depuis (a). (d) Polarisation caractéristique normalisée par leur valeur à $W_{1,MIN}$ en fonction de la largeur W_1 normalisée par $W_{1,MIN}$. (e) Courants caractéristiques normalisés par leur valeur à $W_{1,MIN}$ en fonction de la largeur W_1 normalisée par $W_{1,MIN}$ 136
- Figure 4-23 – (a) Courants de drain I_D et de body-contact I_B en fonction du potentiel de grille-avant du BIMOS V_{FG}^B , mesurés à température ambiante sur un BIMOS de type GO2 RVT aux dimensions $L_1 = 135$ nm, $W_1 = 900$ nm, $L_2 = 144$ nm, $W_2 = 350$ nm. $V_D = 1.8$ V et $V_S = V_B = 0$ V. (b) Tension de seuil $V_{T,MOS}$ du BIMOS en fonction du potentiel de grille-arrière V_{BG}^B , extrait depuis (a). 137
- Figure 4-24 – (a) Caractéristiques $I_C - V_{MEM}$ pour différents potentiels V_{BG}^B , mesurées à température ambiante sur un BB-LTC à oxyde épais (EOT = 3.4 nm) de dimension $L_1 = L = 135$ nm, $W_1 = 144$ nm, $W_2 = W = 350$ nm et $L_2 = 144$ nm. Le BIMOS est de type LVT et le NMOS de type RVT. $V_E = V_S = V_{BG} = 0$ V. (b) Extraction depuis (a) des grandeurs V_{TH} et V_{RST} en fonction de V_{FG}^N . (c) Extraction depuis (a) des grandeurs $I_{ON,TH}$, $I_{OFF,TH}$, $I_{ON,RST}$ et $I_{OFF,RST}$ en fonction de V_{FG}^N 138
- Figure 4-25 – (a) Caractéristiques quasi-statiques $I_C - V_{MEM}$ mesurées à température de deux BB-LTC aux épaisseurs d'oxyde de grille différents (EOT = 1.1 nm et 3.4 nm) et aux autres dimensions égales : $L_1 = L = 135$ nm, $W_1 = 144$ nm, $W_2 = W = 350$ nm et $L_2 = 144$ nm. Le BIMOS est de type LVT et le NMOS de type RVT. $V_E = V_S = V_{BG} = V_{FG}^N = 0$ V. (b) Tensions de seuils V_{TH} et V_{RST} en fonction de l'EOT extraites depuis (a). (c) Courants caractéristiques $I_{ON,TH}$, $I_{OFF,TH}$, $I_{ON,RST}$ et $I_{OFF,RST}$ en fonction de l'EOT extraits depuis (a). 139
- Figure 4-26 – Potentiel V_{AXO} en fonction du potentiel de membrane V_{MEM} et à différents EOT, extrait depuis les données de la Figure 4-25, et reconstruit à partir de mesures des caractéristique $I_D - V_{FG}$ à différents V_D et à $V_B = V_{FG}$ (même méthodologie que celle utilisée à la section 4.3.1). 140
- Figure 4-27 – (a) Courants de collecteur I_C en fonction du potentiel de membrane V_{MEM} obtenus par simulation et par la mesure. (b) Potentiels V_{AXO} en fonction du potentiel de membrane V_{MEM} obtenus par simulation et par la mesure. Les dimensions du BB-LTC sont répertoriées au Tableau 4-5. $V_E = V_S = V_{BG} = V_{FG}^N = 0$ V. 141

Figure 4-28 – Comparaisons des valeurs obtenues par la mesure et par la simulation des potentiels (a) V_{TH} et (b) V_{RST} et des courants (c) $I_{ON,TH}$, (d) $I_{ON,RST}$, (e) $I_{OFF,TH}$ et (f) $I_{OFF,RST}$, en fonction du potentiel V_{FG}^N , pour un BB-LTC aux dimensions répertoriées au Tableau 4-5. $V_{BG} = 0$ V. 141	141
Figure 4-29 – (a) Forme d'onde du potentiel de membrane pendant la caractérisation $I_C - V_{MEM}$ dynamique. (b) Courbes $I_C - V_{MEM}$ obtenues. (c) Potentiel de seuil de V_{TH} en fonction de la vitesse de montée dV_{MEM}/dt , extrait depuis (b). $V_{FG}^N = V_E = V_{BG} = 0$ V. Dimensions du BB-LTC : $L_1 = 135$ nm, $W_1 = 540$ nm, $L_2 = 144$ nm, $W_2 = 350$ nm, $L = 135$ nm, $W = 350$ nm. BIMOS et NMOS de type GO2 RVT. 143	143
Figure 4-30 – (a) Courbes $I_C = f(t)$ pour différents $V_{MEM} = V_{MEM,MAX}$ d'un BB-LTC dont les dimensions sont données au Tableau 4-6, mesurées à température ambiante. (b) forme de la courbe $V_{MEM} = f(t)$. (c) Durée de déclenchement τ_{ON} en fonction de $V_{MEM,MAX}$ extraite depuis (a). 144	144
Figure 4-31 – Circuit de neurone impulsionnel BB-LIF, implémenté à l'aide d'une capacité et d'un BB-LTC composé d'un FD-SOI BIMOS et d'un NMOS. 145	145
Figure 4-32 – Schéma du montage utilisé pour les mesures apportant la preuve de concept du fonctionnement du neurone impulsionnel BB-LIF. 146	146
Figure 4-33 – Chronogrammes $V_{MEM} = f(t)$ pour différents courants I_{SYN} entre 10 pA et 3 nA obtenus par mesure à température ambiante d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7. 147	147
Figure 4-34 – (a) Fréquence de déclenchement f en fonction du courant synaptique I_{SYN} . (b) Tension de seuil V_{TH} en fonction du courant synaptique I_{SYN} . (c) Période réfractaire τ_{RFR} en fonction du courant synaptique I_{SYN} . Valeurs extraites depuis les données de la Figure 4-33. 148	148
Figure 4-35 – Chronogrammes $V_{MEM} = f(t)$ pour des courants I_{SYN} de 5 nA et 10 nA obtenus par mesure à température ambiante d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7. 148	148
Figure 4-36 – Chronogrammes $V_{MEM} = f(t)$ à $I_{SYN} = 100$ pA pour différentes valeurs du potentiel de grille-avant V_{FG}^N obtenus par mesure à température ambiante d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7. 149	149
Figure 4-37 – (a) Fréquence de déclenchement f en fonction du potentiel de grille-avant V_{FG}^N . (b) Tension de seuil V_{TH} en fonction du potentiel de grille-avant V_{FG}^N . (c) Période réfractaire τ_{RFR} en fonction du potentiel de grille-avant V_{FG}^N . Valeurs extraites depuis les données de la Figure 4-36. 150	150
Figure 4-38 – Chronogrammes $V_{MEM} = f(t)$ à $I_{SYN} = 1$ nA pour différentes valeurs du potentiel de grille-arrière V_{BG}^B obtenus par mesure à température ambiante d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7. 152	152
Figure 4-39 – (a) Fréquence de déclenchement f en fonction du potentiel de grille-arrière V_{BG} . (b) Tension de seuil V_{TH} en fonction du potentiel de grille-arrière V_{BG} . (c) Période réfractaire τ_{RFR} en fonction du potentiel de grille-arrière V_{BG} . Valeurs extraites depuis les données de la Figure 4-38. 152	152
Figure 4-40 - Chronogrammes $V_{MEM} = f(t)$ et $V_{AXO} = f(t)$ pour un courant I_{SYN} de 1 nA et 5 nA, obtenus par simulation SPICE d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7. 154	154
Figure 4-41 – (a) Fréquence de déclenchement f , (b) potentiel de seuil V_{TH} , (c) période réfractaire τ_{RFR} et (d) potentiel d'axone maximal $V_{AXO,MAX}$, en fonction du courant synaptique I_{SYN} , extraits de données issues de simulations et de mesures. 154	154
Figure 4-42 – Schéma du design du BB-LIF SN intégré. 158	158

Figure 4-43 – (a) Schéma électrique du buffer intégré au BB-LIF SN. Potentiel de sortie V_{OUT} et courant circulant dans le buffer I_{BUF} en fonction du potentiel d'axone V_{AXO} , obtenue par simulation SPICE calibrée.....	159
Figure 4-44 – Capacités différentielle $C_{MEM,DIFF}$ et effective $C_{MEM,EFF}$ de la NMOSCAP de $7.7 \mu m^2$ utilisée comme capacité de membrane C_{MEM} du BB-LIF SN intégré, en fonction du potentiel de membrane V_{MEM} , obtenues par simulation SPICE et pour une implémentation en technologie 28nm FD-SOI.	160
Figure 4-45 – Layout du BB-LIF SN intégrant le BB-LTC (BIMOS et NMOS) la capacité de membrane (C_{MEM}) et le buffer de sortie asymétrique (N_1 , P_1 , N_2 , et P_2), occupant une surface totale d'environ $15 \mu m^2$	161
Figure 4-46 – Valeurs limites de la fréquence de déclenchement (a) et du courant synaptique (b) pour lesquels le BB-LIF SN est fonctionnel, en fonction du potentiel de grille-avant du NMOS V_{FG}^N , et d'après le modèle développé à la section 4.4.5.	161
Figure 4-47 – Schéma du montage utilisé pour la mesure de la réponse temporelle de BB-LIF SN intégré.	162
Figure 4-48 – Fréquence de déclenchement du BB-LIF SN intégré en fonction du potentiel de contrôle de la source de courant V_{IN} , et pour différentes valeurs de VDD_{SN} . Les données sont issues de mesures à température ambiante.	163
Figure 4-49 – Capture d'écran de l'oscilloscope lors d'une mesure du potentiel de sortie du buffer amplificateur sous les polarisations suivantes : $V_{FG}^N = 0.3 V$, $V_{IN} = 0.42 V$, $VDD_{CS} = 3 V$, $VDD_{SN} = 0.81 V$	164
Figure 4-50 – Fréquence de déclenchement f en fonction du courant synaptique I_{SYN} du BB-LIF SN intégré, obtenue par la mesure à $V_{FG}^N = 0.3 V$ et par simulation à $V_{FG}^N = 0.26 V$, et fenêtre de fonctionnement prédit par le modèle simplifié.....	164
Figure 4-51 – Consommation énergétique dynamique E_{DYN} du BB-LIF SN intégré en fonction de la fréquence de déclenchement f . Donnée obtenue par simulation SPICE avec $VDD_{SN} = 0.75 V$, $VDD_{CS} = 3 V$ et $V_{FG}^N = 0.26 V$	165

Table des tableaux

Tableau 1-1 – Comparaison des processeurs neuronaux « event-based », d’après [72].	39
Tableau 1-2 – Modèles de neurones IF et références où trouver leur description mathématique et leur implémentation matérielle.	43
Tableau 1-3 – Caractéristiques et figures de mérite des circuits analogiques de SN de l’état de l’art.	44
Tableau 2-1 – Tensions de seuil approximatives en mV des transistors MOSFET typiques de la technologie 28nm UTBB FD-SOI. $W = L = 1 \mu\text{m}$	79
Tableau 3-1 – Caractéristiques des BIMOS présents dans la littérature.	82
Tableau 3-2 – Caractéristiques physiques du BIMOS et du NMOS.	91
Tableau 4-1 - Dimensions du BB-LTC utilisé pour la mesure de la Figure 4-9.a	121
Tableau 4-2 – Paramètres de la fenêtre d’hystérésis extraits de la Figure 4-9.	122
Tableau 4-3 – Paramètres de la fenêtre d’hystérésis extraits de la Figure 4-11.	123
Tableau 4-4 – Caractéristiques principales de l’état du BB-LTC dans chacune des régions.	131
Tableau 4-5 - Dimensions du BB-LTC utilisé pour les mesures présentées à la section 4.3.5.3.	137
Tableau 4-6 – Dimensions et polarisation du BB-LTC utilisé pour les mesures de τ_{ON} effectuées à la Figure 4-30	144
Tableau 4-7 - Dimension du BB-LTC utilisé pour les mesures de $V_{\text{MEM}} = f(t)$	146
Tableau 4-8 – Grandeurs nécessaires aux calculs de $I_{\text{SYN,MIN}}$, $I_{\text{SYN,MAX}}$, f_{MIN} et f_{MAX} du BB-LIF SN mesuré à la section 4.4.1	157
Tableau 4-9 – Comparaison de grandeurs $I_{\text{SYN,MIN}}$, $I_{\text{SYN,MAX}}$, f_{MIN} et f_{MAX} obtenues par le calcul et par la mesure.	157
Tableau 4-10 - Dimension du BB-LTC utilisé pour l’intégration du BB-LIF SN.	158
Tableau 4-11 – Paramètres de design des MOSFET qui constitue le buffer intégré au BB-LIF SN.	159
Tableau 4-12 – Valeurs utilisées pour calculer les limites de fonctionnement du BB-LIF SN intégré selon le modèle développé à la section 4.4.5.	161
Tableau 4-13 – Performances du BB-LIF SN intégré.	166
Tableau 4-14 – Derniers travaux à l’état de l’art des circuits analogiques de SN auquel est ajouté le circuit développé dans cette thèse.	168

Glossaire des abréviations, symboles et acronymes

aEIF	: IF à adaptation exponentielle (« adaptive exponential IF »)	C _{OX}	: Capacité de l'oxyde
AER	: Représentation d'adresse d'évènement (« adress event representation »)	CS	: Source de courant (« current source »)
ANN	: Réseau de neurones artificiels (« artificial neural network »)	C _{SI}	: Capacité du film de silicium
BB-LIF	: Neurone impulsif à fuite, intégration et déclenchement, basé sur le BIMOS (« BIMOS-based spiking neurone »)	D _n	: Coefficient de diffusion de électrons
SN		D _p	: Coefficient de diffusion des trous
BB-LTC	: Circuit de déclenchement à fuite basé sur le BIMOS	E	: Energie
BG	: Grille arrière (« back gate »)	$\vec{E}$	: Vecteur de champ électrique
BJT	: Transistor à jonction bipolaire (« bipolar junction transistor »)	E _C	: Niveau d'énergie du minimum de la bande de conduction
BJT _{PAR}	: BJT parasite	E _{DYN}	: Consommation dynamique
BOX	: Oxyde enterré (« buried oxide »)	E _F	: Énergie du niveau de Fermi
BTBT	: Effet tunnel bande à bande (« band-to-band tunneling »)	E _{Fn}	: Quasi-niveau de Fermi des électrons
C _{AXO}	: Capacité d'axone	E _{Fp}	: Quasi-niveau de Fermi des trous
C _{AXO,EFF}	: Capacité d'axone effective	EG	: Grille étendue (« extended gate »)
CBM	: Modèle basé sur la conductance (« conductance-based model »)	E _G	: Énergie du gap
C _{BOX}	: Capacité de l'oxyde enterré	E _i	: Niveau de Fermi du semi-conducteur intrinsèque
CBRAM	: Mémoire vive à pont conducteur (« conductive bridging RAM »)	EOT	: Épaisseur d'oxyde équivalent (« equivalent oxide thickness »)
C _D	: Capacité de la zone de déplétion	ESD	: Décharge électrostatique (« electrostatic discharge »)
C _{DIFF}	: Capacité différentielle	ESN	: « echo state networks »
C _{EFF}	: Capacité effective	E _{STA}	: Consommation statique
C _{MEM}	: Capacité de membrane	E _t	: Niveau d'énergie des pièges
C _{MEM,DIFF}	: Capacité de membrane différentielle	E _V	: Niveau d'énergie du maximum de la bande de valence
C _{MEM,EFF}	: Capacité de membrane effective	F	: Farad
CMOS	: Métal oxyde semiconducteur complémentaire (« complementary metal oxide semiconductor »)	FD	: Entièrement déplété (« fully depleted »)
C _{MOSCAP}	: Capacité de la MOSCAP	FeRAM	: Mémoire vive ferroélectrique (« ferroelectric random-access memory »)
CNN	: Réseau de neurones convolutif (« convolutional neural network »)	FG	: Grille avant (« front gate »)
		FLOPS	: Opérations en virgule flottante par seconde (« floating point operations per second »)
		f _{MAX}	: Fréquence de déclenchement maximale
		f _{MIN}	: Fréquence de déclenchement minimale

FNN	: Réseau de neurones à propagation directe (« feedforward neural network »)	IF	: Intégration et déclenchement (« leaky integrate-and-fire »)
G	: Taux de génération de porteur	I_{GEN}	: Courant de génération
GIDL	: Fuite du drain induite par la grille (« gate-induced drain leakage »)	I_{GR}	: Courant de génération/recombinaison
GISL	: Fuite de la source induite par la grille (« gate-induced source leakage »)	II	: Ionisation par impact (« impact ionization »)
G_{LK}	: Conductance de fuite	I_{II}	: Courant de génération par II
G_n	: Taux de génération d'électron	I_{OFF}	: Courant à l'état OFF
G_{ON}	: Conductance à l'état ON	$I_{OFF,RST}$	: Courant à l'état OFF lorsque $V_{MEM} = V_{RST}$
G_p	: Taux de génération de trou	$I_{OFF,TH}$	: Courant à l'état OFF lorsque $V_{MEM} = V_{TH}$
GPU	: Unité de traitement graphique (« graphical processor unit »)	I_{ON}	: Courant à l'état ON
HH	: Hodgkin et Huxley	$I_{ON,RST}$	: Courant à l'état ON lorsque $V_{MEM} = V_{RST}$
HN	: Réseau de Hopfield (« Hofield network »)	$I_{ON,TH}$	: Courant à l'état ON lorsque $V_{MEM} = V_{TH}$
Hz	: Hertz	I_{PN}	: Courant de jonction PN
$I_{0,SRH}$	: Constante de proportionnalité de courant I_{SRH}	I_R	: Courant de recombinaison
I_A	: Courant d'anode	I_S	: Courant de source
I_B	: Courant de base ou de body-contact	I_{SAT}	: Courant de saturation
$I_{B,VBO}$	: Courant de base lorsque le potentiel de base est nul	I_{SRH}	: Courant de de génération/recombinaison par mécanisme SRH
$I_{B,PAR}$	: Courant I_B parasite	I_{SYN}	: Courant synaptique
I_{BTBT}	: Courant de génération par BTBT	$I_{SYN,LK}$	: Courant synaptique de fuite
I_{BUF}	: Courant du circuit tampon	$I_{SYN,MAX}$	: Courant synaptique maximal
I_C	: Courant de collecteur	$I_{SYN,MIN}$	: Courant synaptique minimal
$I_{C,MAX}$	: Courant de collecteur maximal	I_{TAT}	: Courant de génération par TAT
$I_{C,VBO}$	: Courant I_C lorsque le potentiel de base est nul	J	: Joule
I_{CEO}	: Courant I_C en base ouverte	$\vec{J}$	: Vecteur de densité de courant
$I_{CEO,PAR}$	: Courant I_C en base ouverte du BJT parasite	$\vec{J}_{DIFF,n}$	: Vecteur de densité de courant de diffusion d'électrons
I_{CH}	: Courant de canal	$\vec{J}_{DIFF,p}$	: Vecteur de densité de courant de diffusion de trous
$I_{C,PAR}$	: Courant I_C parasite	$\vec{J}_{DRIFT,p}$	: Vecteur de densité de courant de drift de trous
I_{CS}	: Courant de la source de courant	$\vec{J}_{DRIFT,n}$	: Vecteur de densité de courant de drift d'électrons
I_D	: Courant de drain	$\vec{J}_p$	: Vecteur de densité de courant de trous
$I_{D,SAT}$	: Courant de drain en saturation	$\vec{J}_n$	: Vecteur de densité de courant d'électrons
I_{DIFF}	: Courant de diffusion	J/spike	: Joule par impulsion (« joule per spike »)
I_E	: Courant d'émetteur	K	: Kelvin
$I_{e,BJT,PAR}$	: Courant d'électron du BJT parasite	k	: Constante de boltzmann
I_{EEO}	: Courant I_E en base ouverte	$\vec{k}$	: Vecteur d'onde

L	: Longueur	PCRAM	: Mémoire vive à changement de phase (« phase change random-access memory »)
L_G	: Longueur de grille	PD	: Partiellement déplété (« Partially depleted »)
LIF	: À fuite, intégration et déclenchement (« leaky integrate-and-fire »)	PMOS	: MOSFET de type P
LSM	: « Liquid state machines »	PMOSFE	: MOSFET de type P
LSTM	: « Long short term memory »	T	
LTC	: Circuit de déclenchement à fuite	q	: Charge élémentaire
LVT	: Tensions de seuil faibles (« low V_T »)	Q_{AXO}	: Charge stockée dans C_{AXO}
MIM	: Métal isolant métal	Q_D	: Charge de la zone de déplétion
MOM	: Métal oxyde métal	Q_{SC}	: Charge dans le semiconducteur
MOS	: Metal oxyde semiconcteur	R	: Taux de recombinaison des porteurs
MOSCAP	: Capacité MOS (« MOS capacitor »)	R_n	: Taux de recombinaison des électrons
MOSFET	: Transistor à effet de champ MOS (« MOS field effect transistor »)	RNN	: Réseau de neurones récurrent (« recurrent neural network »)
MRAM	: Mémoire vive magnétique (« magnetic random-access memory »)	R_{ON}	: Résistance à l'état ON
n	: Densité d'électrons présent dans le bande de conduction	R_p	: Taux de recombinaison des trous
n_{AXO}	: Nœud d'axone	R_S	: Résistance série d'une jonction PN
N_A	: Densité de dopants accepteurs	$R_{S,SRH}$	: Taux de recombinaison surfacique par mécanisme SRH
N_B	: Densité de dopants dans la base	RVT	: Tensions de seuil régulière (« regular V_T » :).
N_C	: Densité d'états du minimum de la bande de conduction ou densité de dopants dans le collecteur	SCE	: Effet de canaux courts (« short channel effect »)
N_D	: Densité de dopants donneurs	SCR	: Redresseur silicium commandé (« silicon controlled rectifier »)
N_E	: Densité de dopants dans l'émetteur	SG	: Grille standard (« standard gate »)
n_i	: Densité de porteurs intrinsèque	SMU	: Unité de source et de mesure (« source and measure unit »)
n_{MEM}	: Nœud de membrane	SN	: Neurone impulsionnel (« spiking neuron »)
NMOS	: MOSFET de type N	SNN	: Réseau de neurones impulsionnels (« spiking neural network »)
NMOSCA	: MOSCAP de type N	SOI	: Silicium sur isolant (« silicon on insulator »)
P		SPICE	: Simulation program with integrated circuit emphasis
NMOSFE	: MOSFET de type N	S_{PN}	: Section d'une jonction PN
T		SRAM	: Mémoire vive statique (« static random-access memory »)
$N_{s,t}$	: Densité de pièges surfacique	SRH	: Shockley-Read-Hall
N_t	: Densité de pièges	SS	: Pente sous le seuil
N_v	: Densité d'état du maximum de la bande de valence	SSD	: Disque dur à semiconducteur (« solide state drive »)
OxRAM	: Mémoire vive à base d'oxyde (« oxide random-access memory »)		
p	: Densité de trous présents dans la bande de conduction		

STDP	: Plasticité fonction du temps d'occurrence des impulsions (« spike-timing-dependent plasticity »)	VDD _{CS}	: Tension d'alimentation de la source de courant
STI	: Isolation par tranchée peu profonde (« Shallow Trench Isolation »)	VDD _{EG}	: Tension nominale des composants EG
t	: Temps	VDD _{SN}	: Tension d'alimentation du neurone impulsif
T	: Température	V _{D-S}	: Tension drain – source
TAT	: Effet tunnel assisté par piège (« trap-assisted tunneling »)	V _E	: Potentiel d'émetteur
T _{KD}	: Température du composant	V _{FB}	: Tension de bandes plates
T _{KR}	: Température de référence	V _{FG}	: Potentiel de grille-avant
t _{OX}	: Épaisseur de l'oxyde	V _{FG-E}	: Tension grille-avant – émetteur
U _T	: Potentiel thermique	V _{G-SUB}	: Tension grille – substrat
UTBB	: Corps et BOX ultra fin (« ultra thin body and box »)	V _{IN}	: Potentiel d'entrée
V _A	: Potentiel d'anode	VLSI	: Intégration à très grande échelle (« very large scale intégration »)
VA-K	: Tension anode – cathode	V _{MEM}	: Potentiel de membrane
V _{AXO}	: Potentiel d'axone	V _{MEM,MAX}	: Potentiel de membrane maximal
V _{AXO,INV}	: Potentiel d'axone lorsque I _D = I _{PN}	V _{M-SC}	: Tension métal – semi-conducteur
V _{AXO,MAX}	: potentiel d'axone maximal	V _{N^{BG}}	: Potentiel de grille-arrière du NMOSFET
V _{AXO,OFF,RS}	: V _{AXO} à l'état OFF lorsque V _{MEM} = V _{RST}	V _{N^{FG}}	: Potentiel de grille-avant du NMOSFET
V _{AXO,OFF,TH}	: V _{AXO} à l'état OFF lorsque V _{MEM} = V _{TH}	V _{OUT}	: Potentiel de sortie
V _{AXO,ON,RS}	: V _{AXO} à l'état ON lorsque V _{MEM} = V _{RST}	V _{RST}	: Potentiel de seuil de réinitialisation
V _{AXO,ON,TH}	: V _{AXO} à l'état ON lorsque V _{MEM} = V _{TH}	V _S	: Potentiel de source
V _B	: Potentiel de base ou de body-contact	V _{T,BUF}	: Potentiel de seuil du circuit tampon
V _{B^{BG}}	: Potentiel de grille-arrière BIMOS	V _{T,MOS}	: Tension de seuil du transistor MOSFET
V _{BC-S}	: Tension body-contact – source	V _{T,MOS,BG}	: Tension de seuil de la grille-arrière
V _{B-E}	: Tension base – émetteur	V _{T,MOS,FG}	: Tension de seuil de la grille-avant
V _{B^{FG}}	: Potentiel de grille-avant du BIMOS	V _{T,PN}	: Tension de transition de la diode
V _{BG}	: Potentiel de grille-arrière	V _{TH}	: Potentiel de seuil de déclenchement
V _{bi}	: Potentiel interne d'une jonction PN (« built-in potential »)	v _{TH}	: Vitesse thermique
V _C	: Potentiel de collecteur	W	: Watt ou largeur
V _{C-B}	: Tension collecteur – base	W _{MIN}	: Largeur minimale
V _{C-E}	: Tension collecteur – émetteur	WTA	: « winner-take-all »
V _D	: Potentiel de drain	W _{ZCE}	: Largeur de la ZCE
VDD	: Tension d'alimentation	x _D	: Profondeur de la zone de déplétion
		x _{D,MAX}	: Profondeur maximale de la zone de déplétion

ZCE	: Zone de charge d'espace
α_F	: Gain en base commune
α_T	: Facteur de transport dans la base
β_F	: Gain en émetteur commun
$\beta_{F,PAR}$	: Gain en émetteur commun parasite
γ_F	: Rendement d'émetteur
ΔV_{MEM}	: Variation du potentiel de membrane
$\Delta\psi$	: Variation du potentiel de surface
ϵ_{OX}	: Permittivité diélectrique de l'oxyde
ϵ_{SC}	: Permittivité diélectrique du semiconducteur
η	: Coefficient d'idéalité d'une jonction PN
μ_n	: Mobilité des électrons
μ_p	: Mobilité des trous
σ_n	: Section efficace de capture des électrons
σ_p	: Section efficace de capture des trous
τ_0	: Durée de vie des porteurs
τ_{AXO}	: Durée de décroissance de V_{AXO}
τ_{FALL}	: Temps de descente de V_{MEM}
$\tau_{MEM,MAX}$	: Durée de montée maximale de V_{MEM}
τ_n	: Durée de vie des électrons
τ_{ON}	: Durée pour atteindre l'état ON
τ_p	: Durée de vie des trous
τ_{RFR}	: Durée de la période réfractaire

τ_{RISE}	: Temps de montée de V_{MEM}
τ_{TRG}	: Durée de déclenchement
$\vec{\phi}$	: Vecteur de densité de flux
$\vec{\phi}_{DIFF,n}$	: Vecteur $\vec{\phi}_{DIFF}$ d'électrons
$\vec{\phi}_{DIFF,p}$	: Vecteur $\vec{\phi}_{DIFF}$ de trous
$\vec{\phi}_{DRIFT,n}$	: Vecteur $\vec{\phi}_{DRIFT}$ d'électrons
$\vec{\phi}_{DRIFT,p}$	: Vecteur $\vec{\phi}_{DRIFT}$ de trous
$\vec{\phi}_{DIFF}$	: Vecteur $\vec{\phi}$ de diffusion
$\vec{\phi}_{DRIFT}$	: Vecteur $\vec{\phi}$ de drift
$\vec{\phi}_n$	: Vecteur $\vec{\phi}$ d'électrons
$\vec{\phi}_p$	: Vecteur $\vec{\phi}$ de trous
ϕ_F	: Distance énergétique entre le niveau de Fermi du semi-conducteur à l'équilibre et le niveau de Fermi intrinsèque.
ϕ_M	: Travail de sortie du métal
ϕ_{M-SC}	: Différence des travaux de sortie du métal et du semi-conducteur
$\phi_{M-SC,B}$	: ϕ_{M-SC} du côté de la grille arrière
$\phi_{M-SC,F}$	: ϕ_{M-SC} du côté de la grille avant
$\phi_{n,BC}$	: Flux d'électrons qui traversent la jonction base – collecteur
$\phi_{n,BE}$	: Flux d'électrons qui traversent la jonction base – émetteur
$\phi_{p,BE}$	: Flux de trous qui traversent la jonction base – émetteur
ϕ_s	: Potentiel de surface
ϕ_{SC}	: Travail de sortie du semi-conducteur
ψ_{SD}	: Potentiel de surface côté drain
ψ_{SS}	: Potentiel de surface côté source
Ω	: ohm

Introduction

Incontestablement, nos sociétés contemporaines sont façonnées par le déploiement massif des technologies de l'information. Leur systématisation génère aujourd'hui des bouleversements profonds dans de très nombreux champs de nos vies individuelles et collectives. Le déploiement de l'internet haut débit, des réseaux de communications sans fils et la prolifération des terminaux multimédias changent significativement au quotidien nos modes de communication, de divertissement et de consommation, nos façons de nous informer, de nous sociabiliser ou de nous organiser, jusqu'à notre rapport à la connaissance. A l'échelle des corps sociaux, économiques ou politiques, ces technologies éclairent les politiques publiques, favorisent la circulation des connaissances scientifiques, soutiennent la mondialisation des échanges commerciaux, renforcent la financiarisation de l'économie, facilitent la multinationalisation des entreprises et transforment le marché du travail.

La généralisation des technologies de l'information est intrinsèquement liée au progrès constant de la microélectronique de par la nature des dispositifs sur lesquels elles reposent. Depuis son invention au milieu du siècle dernier, la microélectronique n'a cessé de fournir des composants toujours plus performants, intégrés dans des microprocesseurs toujours plus efficaces. Jusqu'à récemment, ces progrès étaient principalement gouvernés par la loi de Moore. Enoncée initialement en 1965, puis révisée en 1975, cette loi généralisait le constat que le nombre de composants intégrés par processeur doublait tous les 2 ans. En réalité, plus qu'une loi, elle était une feuille de route que se fixait l'industrie de la microélectronique pour diriger ses efforts en recherche et développement. De fait, pendant plus de 40 ans, cette loi performative a tenu ses promesses, allant jusqu'à permettre la construction des téléphones intelligents depuis la fin des années 2000, produits emblématiques des révolutions qui accompagnent l'époque actuelle.

Cependant, la loi de Moore atteint aujourd'hui ses limites [1]. La réduction de la taille des transistors, composants essentiels de la microélectronique, est devenue si importante qu'elle se heurte désormais aux limites physiques de la matière. La longueur de grille des transistors, dimension caractéristique de ces composants, devient maintenant comparable aux distances interatomiques des cristaux de silicium. À ces échelles, les équilibres électrostatiques à la base du fonctionnement des transistors changent de régime. Des phénomènes parasites tel que l'effet tunnel deviennent prépondérants et dégradent fortement les performances des composants. De plus, l'augmentation des fréquences de rayonnement laser nécessaire pour atteindre la résolution requise en photolithographie pour façonner les composants de taille toujours plus réduite est limitée par l'opacification de l'air à de telles longueurs d'onde. Pour pallier ces effets indésirables, la complexification des procédés de fabrication industrielle des puces microélectroniques s'est intensifiée avec pour conséquence l'augmentation significative des coûts des équipements. Ainsi, le coût unitaire des transistors tend aujourd'hui à se stabiliser [2] alors qu'il décroissait systématiquement d'une génération technologique à l'autre, pendant l'ère du « happy scaling » des décennies précédentes.

Face à la fin annoncée de la loi de Moore, les acteurs du champ de la microélectronique cherchent de nouveaux paradigmes sur lesquels s'appuyer pour poursuivre la marche du progrès. Ce mouvement, appelé « More-than-Moore », a d'abord été initié par la modification des architectures des systèmes avec l'émergence des processeurs multicœurs au milieu des années 2000. La parallélisation des tâches qu'autorise ce type d'architecture permet un gain de performances sans augmenter la fréquence d'horloge, contournant ainsi les problèmes d'auto-échauffement qui commençaient à apparaître. Cette réflexion sur l'organisation des architectures se poursuit aujourd'hui par la conception de systèmes basés sur la cointégration et l'interconnexion de puces spécialisées et dédiées à certaines fonctions, en particulier dans le cadre de l'intégration 3D [3] qui est en plein essor aujourd'hui.

Cependant, parmi les stratégies « More-than-Moore » actuellement en développement, deux d'entre elles sont particulièrement des stratégies de rupture : l'informatique quantique et le neuromorphisme. La première consiste à s'appuyer sur les propriétés quantiques de la matière pour effectuer certains types de calculs avec un gain de performance sans précédent. La seconde consiste à construire des systèmes qui s'inspirent du mode de traitement de l'information observé dans les systèmes nerveux biologiques et qui ont la particularité de devoir être entraînés à la tâche à laquelle ils sont dédiés. Ces deux approches n'ont pas la même maturité. L'informatique quantique est encore à ses balbutiements, bien qu'un premier ordinateur basé sur ce principe ait été récemment proposé par IBM en 2019. De son côté, le neuromorphisme est à la base de nombreux services déjà disponibles pour le grand public, comme les assistants vocaux ou la reconnaissance de contenus dans les photographies. Il est par ailleurs à l'origine des progrès récents en intelligence artificielle qui a vu exploser le nombre de ses débouchés applicatifs.

Les systèmes neuromorphiques reposent sur l'exécution d'algorithmes qui implémentent des modèles mathématiques de réseaux de neurones. Historiquement, ces algorithmes de réseaux de neurones artificiels (« artificial neural network » : ANN) étaient exécutés sur des processeurs standards à l'architecture de type Von Neumann. Aujourd'hui encore, les ANN les plus performants sont exécutés de la sorte. Cependant, ces dernières années voient l'émergence de processeurs dont l'architecture elle-même est calquée sur celle des systèmes nerveux. Ces processeurs sont composés d'une multitude de copies de circuits élémentaires interconnectés qui reproduisent chacun le comportement d'un neurone ou d'une synapse et qui communiquent entre eux au travers de l'émission de courtes impulsions électriques. Ce mode de communication impulsif et asynchrone réduit drastiquement la consommation électrique de ces processeurs neuronaux.

La taille de ANN, c'est-à-dire le nombre d'unités neuronales et synaptiques qu'ils interconnectent, est le premier facteur de leur performance. Lorsqu'ils sont exécutés sur des machines standards, cette taille est limitée par la capacité de stockage de l'élément de mémoire du système, qui peut être arbitrairement grande. En revanche, lorsqu'ils sont exécutés par des processeurs neuronaux, cette taille dépend du nombre de circuits élémentaires de neurones et de synapses présents sur la puce. Or, les surfaces des puces sont limitées par le champ d'insolation de photolithographie qui ne peut excéder quelques centimètres carrés. Dans ce contexte, la réduction de l'empreinte de ces circuits élémentaires est un enjeu majeur pour permettre aux processeurs neuronaux de rivaliser avec les ANN exécutés sur des processeurs standards.

Les travaux de recherche présentés dans cette thèse s'inscrivent dans le champ du neuromorphisme. Ils portent sur l'étude et le développement d'un circuit analogique de neurone à impulsion intégré en technologie 28 nm FD-SOI, avec la réduction de la surface de l'empreinte du circuit comme priorité. L'originalité du travail réside dans l'exploitation du phénomène de génération de paires électron-trou dans un transistor à l'architecture particulière appelé BIMOS, pour réaliser la fonction de transfert du neurone. En s'inspirant notamment de certains circuits de protection contre les décharges électrostatiques.

Quatre chapitres composent ce document :

- Le premier détaille le contexte dans lequel se place le travail, dresse l'état de l'art des circuits analogiques de neurone à impulsion et énonce les objectifs de la thèse.
- Le second introduit les notions de physique des composants semiconducteurs nécessaires à l'interprétation des caractérisations électriques présentées dans ce document. Il décrit également la technologie 28 nm FD-SOI de STMicroelectronics avec laquelle les dispositifs sont fabriqués.

- Le troisième chapitre est une étude détaillée du comportement électrique quasi-statique à température ambiante du FD-SOI BIMOS, selon que ses polarisations le font fonctionner à la manière d'une diode à grille, d'un transistor à effet de champ et d'un transistor bipolaire. L'analyse des mesures effectuées sur les dispositifs permet de construire et de paramétrer un modèle compact du BIMOS qui est une version modifiée du modèle UTSOI.
- Le dernier chapitre présente le circuit analogique de neurone impulsif à fuite, intégration et déclenchement et basé sur le BIMOS (« BIMOS-based leaky integrate-and-fire spiking neuron » : BB-LIF SN) issu des travaux de recherche. En s'appuyant sur la description du comportement électrique du BIMOS effectuée au chapitre 3, il décrit précisément les mécanismes de déclenchement et de production d'impulsions du circuit de neurone. L'influence des polarisations de contrôle sur le comportement du circuit est étudiée par la mesure à température ambiante d'un premier démonstrateur partiellement intégré. Un modèle analytique simple d'évaluation des limites de fonctionnement du circuit est proposé. Les mesures sont confrontées aux simulations et la cohérence des résultats est discutée. Les performances d'un second démonstrateur entièrement intégré sont évaluées à l'aune des figures de mérite évoquées au chapitre 1. La cohérence entre les mesures du circuit à température ambiante, les simulations compactes utilisant le modèle construit au chapitre 3 et le modèle d'évaluation des limites de fonctionnement est établie.

A la suite de ces chapitres, une conclusion est finalement proposée.

Les débuts des travaux de recherche ont été guidés par l'exécution de très nombreuses simulations en trois dimensions du comportement électrique des composants FD-SOI utilisant la méthode des éléments finis. Elles ont permis de dimensionner les premiers échantillons de BIMOS et de BB-LIF SN. Cependant, par souci de clarté vis-à-vis du lecteur, les résultats de ces investigations ne sont pas évoqués dans ce document qui se concentre exclusivement sur la discussion des résultats de mesures effectuées sur les échantillons.

Chapitre 1 : Contexte et État de l'art

Ce chapitre a pour but d'apporter au lecteur les éléments de contexte et une description des contributions à l'état de l'art qui lui permettent d'appréhender les enjeux des travaux présentés dans cette thèse. Son contenu est organisé ainsi. Après une introduction du concept de neuromorphisme, les réseaux de neurones artificiels (« artificial neural network » : ANN) sont présentés par le biais des trois éléments qui les constituent : les modèles de neurones, les topologies de réseaux et les règles d'apprentissage. Ensuite, les différents types d'implémentation matérielle de ces réseaux sont déclinés. Enfin, après avoir présenté les figures de mérite des circuits analogiques de neurone impulsif (« spiking neuron » : SN) et les travaux de la littérature à l'état de l'art, les objectifs de la thèse sont énoncés.

1.1 Le neuromorphisme

Cette section introduit le concept de neuromorphisme. Elle s'articule autour d'une description succincte du système nerveux et des neurones biologiques, et d'une présentation de l'ingénierie neuromorphique.

1.1.1 Système nerveux et neurone biologique

Le système nerveux est un système biologique animal dont le rôle est, d'une part, de réguler le fonctionnement des organes vitaux et, d'autre part, de coordonner les actions physiques à effectuer selon la configuration de l'environnement et les besoins de l'individu. Pour ce faire, il met en relation les organes sensoriels, moteurs et vitaux, desquels et auxquels il reçoit et envoie de l'information. D'un point de vue fonctionnel, il a pour rôle principal de transmettre et traiter de l'information.

Le système nerveux est constitué d'un réseau de cellules nerveuses, appelées neurones, interconnectées au travers de terminaisons, appelées synapses. Le neurone est composé d'un corps cellulaire, ou soma, qui contient le noyau et auxquels sont reliés des prolongements : les dendrites et l'axone. L'axone se subdivise en son extrémité en terminaisons axonales. Le neurone est séparé de son milieu extérieur par sa membrane. Les terminaisons axonales sont reliées aux dendrites d'autres neurones par les synapses (Figure 1-1.a).

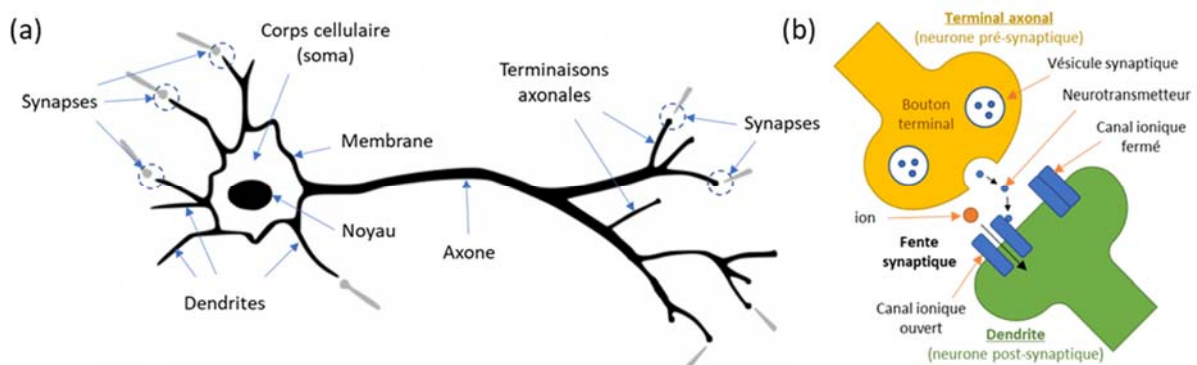


Figure 1-1 – Schéma descriptif (a) d'un neurone biologique et (b) d'une synapse biologique (dans (a), les parties grisées représentent les dendrites et terminaisons axonales d'autres neurones).

Lorsqu'un neurone s'active, il émet une impulsion électrique (le potentiel d'action) qui se propage le long de l'axone vers les terminaisons axonales. L'arrivée du potentiel d'action aux terminaisons axonales a pour effet de libérer des neurotransmetteurs, initialement stockés dans les vésicules

synaptiques présentes dans le bouton terminal, dans le milieu extracellulaire qui, dans la région des synapses, porte le nom de fente synaptique. Une fois libérés, ces neurotransmetteurs diffusent dans la fente synaptique jusqu'à atteindre la membrane de la dendrite du neurone post-synaptique. Au contact des neurotransmetteurs, des canaux ioniques situés sur cette membrane s'activent et s'ouvrent. Des flux ioniques locaux – principalement d'ions Na^+ , K^+ , Ca^{2+} et Cl^- – apparaissent entre les milieux extra et intra cellulaire du neurone post-synaptique. Ces flux ioniques font varier son potentiel de membrane qui est la différence de potentiel électrique entre le milieu interne et externe du neurone. Cette variation s'ajoute à celles provenant de l'ensemble des synapses afférentes au neurone post-synaptique. Lorsque son potentiel de membrane atteint un certain seuil, ce dernier s'active et émet un potentiel d'action à son tour.

1.1.2 L'ingénierie neuromorphique

L'ingénierie neuromorphique est un concept introduit par le professeur Carver Mead de l'institut Caltech (Californie, États-Unis) à la fin des années 1980 [1]. A l'origine, elle consiste en l'utilisation de circuits électroniques analogiques intégrés à très grande échelle (« very large scale integration » : VLSI) pour construire des systèmes de traitement de l'information inspirés du fonctionnement des réseaux de neurones biologiques. De nos jours, elle désigne tout système électronique analogique, digital ou mixte ou tout système informatique (logiciel) qui implémente des modèles de neurones et de synapses mis en réseau les uns avec les autres [2].

Pour développer ce concept, Carver Mead s'est appuyé notamment sur la grande similarité entre les courbes caractéristiques des courants ioniques qui circulent à travers la membrane des neurones en fonction de la tension de part en part de celle-ci, et celle du courant de saturation des transistors en fonction de leur tension de grille (Figure 1-2) [3]. La principale motivation pour l'étude et le développement des systèmes neuromorphiques réside dans l'efficacité énergétique du cerveau (estimée par C. Mead à 10^{-16} J/opération) comparée à celle des circuits intégrés classiques (10^{-9} J/opération) [3].

L'architecture des systèmes neuromorphiques se distingue des architectures de type Von-Neumann [4] sur lesquelles sont basés les systèmes de calculs classiques. En particulier, les unités de mémoire (synapses) sont distribuées parmi les unités de calcul (neurone et synapse) au sein d'un réseau, contrairement aux architectures Von-Neumann où les deux types d'unités sont regroupés en deux ensembles distincts.

Bien que l'invention du terme « neuromorphique » soit attribuée à Carver Mead dans les années 1980, l'intérêt des sciences de l'informatique et de l'électronique pour les réseaux de neurones est présent dès le début de leur histoire. Ainsi, dès 1943, McCulloch et Pitts proposent un formalisme mathématique pour représenter la façon dont les réseaux de neurones traitent l'information [5]. Ce formalisme a donné naissance à un modèle de neurone, dit « formel » (voir section 1.2), encore utilisé aujourd'hui dans les réseaux de neurones artificiels (« artificial neural network » : ANN) les plus performants. En 1958, Franck Rosenblatt introduit une première architecture de calculateur inspirée des réseaux de neurones biologiques appelée « perceptron » [6] en s'appuyant sur le modèle du neurone formel. La généralisation du perceptron donnera naissance au perceptron multi couches et, plus généralement, aux réseaux de neurones à propagation directe (« feedforward neural network » : FNN). Comme tous les systèmes neuromorphiques, le perceptron doit d'abord être entraîné pour exécuter une tâche : les paramètres internes (les poids synaptiques et les seuils d'activation des neurones) doivent être ajustés lors d'une phase d'apprentissage. Cet apprentissage peut être

supervisé ou non-supervisé, selon que l'ajustement des paramètres est guidé depuis l'extérieur ou effectué automatiquement en fonction de l'activité du réseau.

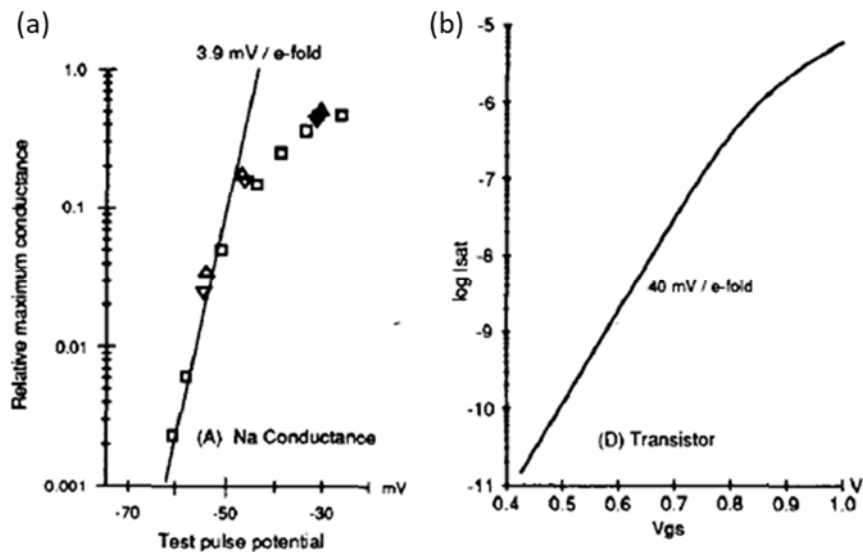


Figure 1-2 – (A) Comportement d'une population de canaux ioniques actifs dans une membrane de neurone. (B) Courant de saturation d'un transistor MOS en fonction de la tension de grille. D'après [3].

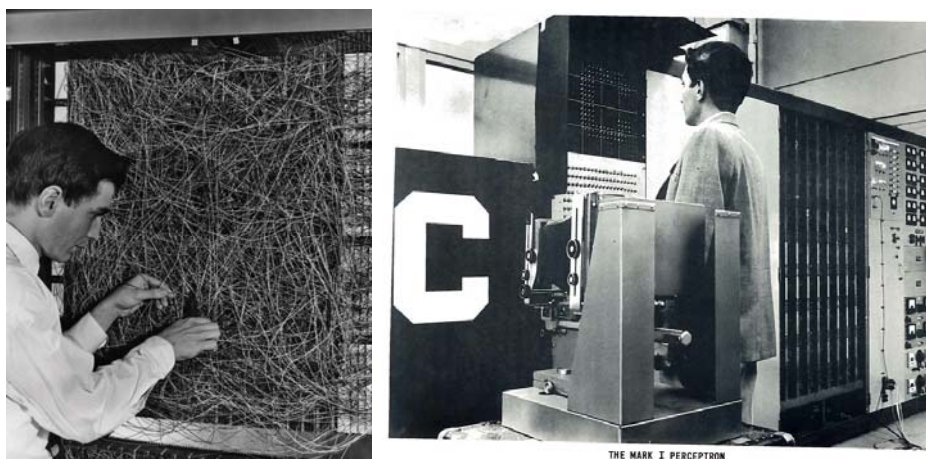


Figure 1-3 – Photographie du perceptron MARK I construit par Rosenblatt, provenant de [7].

Pour concevoir le perceptron, Rosenblatt s'est appuyé notamment sur les travaux du psychologue Donald O. Hebb qui énonça en 1949 un principe général sur la plasticité neuronale [8] (i.e. la façon dont les connections synaptiques évoluent selon l'activité des neurones au cours du temps) et qui prendra le nom de « règle de Hebb ». Elle peut être résumée par le fait que « des neurones qui se déclenchent ensemble se lient entre eux » [9]. Il est le premier exemple de règle d'apprentissage non-supervisé.

1.2 Les réseaux de neurones artificiels

Un réseau de neurones artificiels correspond à une configuration spécifique des trois éléments suivants : le modèle de neurone qui inclue également les synapses, la topologie du réseau et la règle d'apprentissage. Cette section en donne une vue d'ensemble.

1.2.1 Les modèles de neurone

Il existe deux familles de modèle de neurone, les modèles de neurone formel et les modèles de neurone impulsionnel. Bien que les travaux exposés dans cette thèse portent sur la seconde, elles sont toutes deux présentées ici. En effet, la description du neurone formel facilite la compréhension du fonctionnement des modèles de neurone en général.

1.2.1.1 Le neurone formel

Le neurone formel est une unité élémentaire qui dispose de plusieurs entrées, les synapses, et d'une seule sortie, l'axone. Les valeurs d'entrée présentées à ses synapses sont multipliées chacune par un poids synaptique dédié. Les valeurs ainsi pondérées sont sommées et utilisées comme argument pour le calcul de la fonction d'activation qui donne la valeur de sortie du neurone (Figure 1-4).

À l'origine, dans les modèles de neurone dits de première génération [10], cette fonction d'activation était une fonction de Heaviside à seuil ($y = H(x-\theta)$), restreignant la sortie du neurone à une sortie binaire (0 ou 1). Par la suite, cette fonction s'est diversifiée, constituant les modèles de seconde génération [10]. Il peut s'agir par exemple des fonctions sigmoïde ($y = [1 + \exp(-x)]^{-1}$), arc tangente ($y = \arctan(x)$) ou rectifieur linéaire ReLU ($y = \max(0; x)$). Comparées à la fonction Heaviside, ces fonctions ont l'avantage d'être continues et dérivables (à l'exception de ReLU en $x = 0$). Elles facilitent notamment l'application de la méthode rétropropagation du gradient [11], [12] présentée à la section 1.2.3.1 pour l'algorithme d'apprentissage.

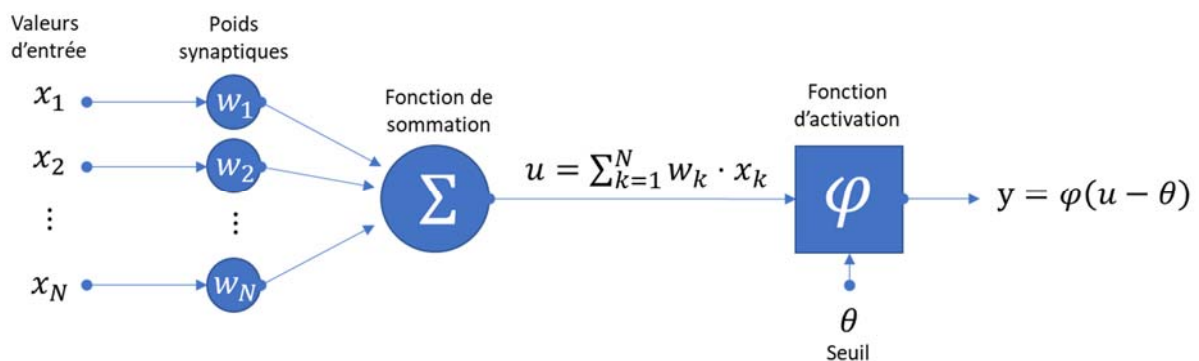


Figure 1-4 – Schéma descriptif du modèle mathématique de neurone formel.

1.2.1.2 Le neurone impulsionnel

Les modèles de neurone impulsionnel (« spiking neuron » : SN) sont des modèles de neurone qui se rapprochent davantage du comportement des neurones biologiques. Ils rendent compte de la dimension temporelle de l'activation qui se manifeste par l'émission d'une impulsion repérée dans le temps. Tous les neurones d'un même réseau émettent des impulsions de même amplitude et de même durée (donnée temporelle binaire). Leurs amplitudes sont modifiées par leur passage au travers des synapses, selon le poids synaptique de ces dernières. La grandeur que représentent les impulsions change également de nature : de potentiel d'action, elles deviennent un courant synaptique d'excitation. Les impulsions pondérées sont ensuite sommées au cours du temps (intégration temporelle) au niveau des neurones. Par analogie avec les neurones biologiques, le résultat de cette intégration est nommé potentiel de membrane. Les courants synaptiques peuvent être positifs ou négatifs. Les synapses d'où ils proviennent sont alors excitatrices ou inhibitrices respectivement. L'impulsion d'un neurone est émise lorsque son potentiel de membrane atteint un certain seuil.

L'émission de l'impulsion s'accompagne d'un retour du potentiel de membrane à son potentiel de repos.

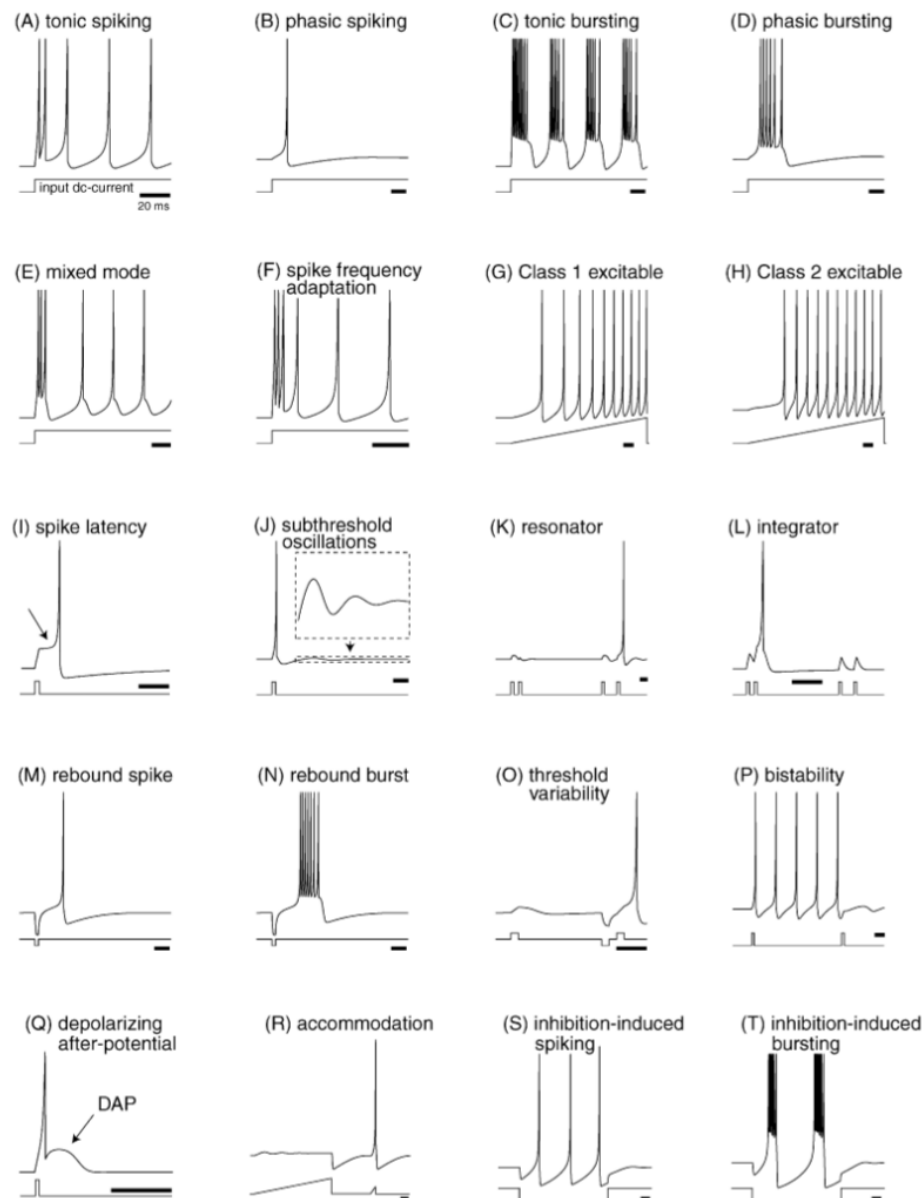


Figure 1-5 – Résumé des comportements identifiés dont sont capables les SN biologiques. Depuis [13].

Le modèle décrit ici – le modèle à intégration et déclenchement (« integrate and fire » : IF) – est très simplifié. En effet, les neurones biologiques sont capables de générer une grande diversité des trains d'impulsions [13] (Figure 1-5), selon leurs types et leurs fonctions, dont le modèle IF ne peut pas rendre compte. Il existe cependant d'autres modèles capables d'émuler davantage de comportements. Ils sont répertoriés dans [13] et classés selon le type de comportement qu'ils sont capables d'émuler et leur coût calculatoire (i.e. nombre d'opération à effectuer pour calculer la valeur du potentiel de membrane à chaque pas de temps) (Figure 1-6).

Le plus ancien de ces modèles est celui du neurone à intégration à fuite et déclenchement (« leaky, integrate and fire » : LIF), proposé par Lapicque en 1907 [14], [15]. En plus des caractéristiques du IF, le LIF dispose d'une fonction d'oubli implémentée par l'ajout d'une dynamique de retour du potentiel de membrane au potentiel de repos indépendante de l'excitation. Le LIF reste pauvre en capacité

d'émulation de la diversité des dynamiques neuronales. Néanmoins, il a servi de base à de nombreux autres modèles. D'après [16], sa généralisation s'est effectuée selon trois axes : 1 – une extension de l'équation différentielle en V_{MEM} à des formes quadratiques [17]–[19] ou exponentielles [20], 2 – l'ajout d'une seconde variable dynamique autre que V_{MEM} [21], [22] et 3 – une modélisation d'injection de conductance plutôt que d'injection de courant [23].

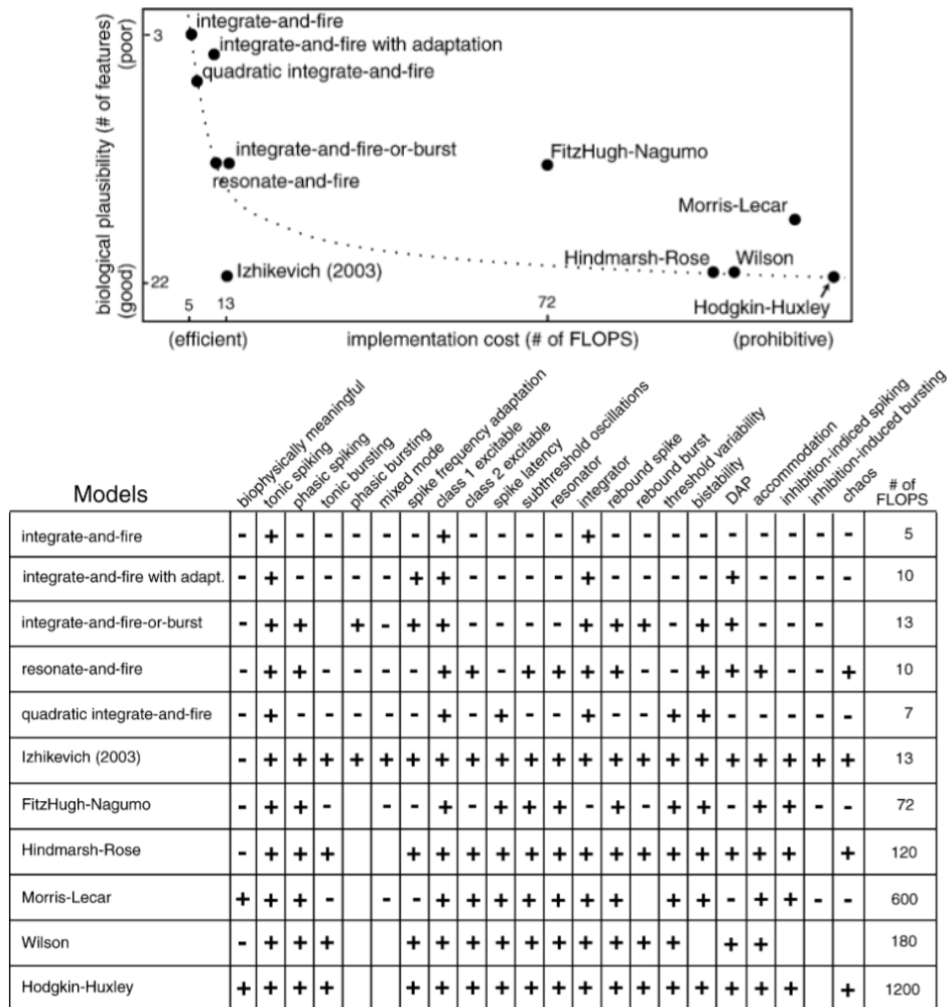


Figure 1-6 – Comparaison des comportements que sont capables de reproduire les modèles de neurone et du nombre d'opération à effectuer pour les simuler à chaque pas de temps. Depuis [13].

Une autre famille de modèles est constituée par des modèles basés sur la conductance (« conductance-based models » : CBM). Le premier d'entre eux est proposé par Hodgkin et Huxley en 1952 et porte leur nom [24]. Ce type de modèle s'attache à modéliser la dynamique des différents canaux ioniques (Na^+ et K^+ dans [24]) présents sur la membrane du neurone. Ainsi, grâce à un jeu d'équations différentielles couplées, ils peuvent émuler les mécanismes internes aux neurones responsables de la génération du potentiel d'action lorsque le potentiel de membrane atteint son seuil, là où les modèles de type LIF se contentent d'introduire une discontinuité sous la forme d'une équation

conditionnelle. Les CBM peuvent notamment rendre compte de la présence de la période réfractaire¹ qui est une caractéristique importante des neurones biologiques. Selon les besoins, des canaux ioniques peuvent être ajoutés pour reproduire des dynamiques globales spécifiques [25]. Certains modèles trouvent leur filiation à la fois dans les modèles LIF et Hodgkin et Huxley (HH) comme les modèles « adaptive exponential IF » (aEIF) [16], et « Generalized LIF » [26]–[28].

1.2.2 Topologies de réseaux neuronaux

Les ANN sont classées en deux catégories selon leur topologie. Ceux qui contiennent des cycles (i.e. succession de synapses et de neurones formant une boucle) sont appelés réseaux de neurones récurrents (« recurrent neural network » : RNN), les autres, réseaux de neurones à propagation directe (« feedforward neural network » : FNN).

1.2.2.1 Les FNN

Les FNN sont généralement utilisés pour traiter des jeux de données statiques. Deux types de FNN sont particulièrement populaires : le perceptron multi-couche et les réseaux de neurones convolutifs (« convolutional neural network » : CNN) [29], [30].

Le perceptron multi couche :

Le perceptron multi couche est une structure typique des FNN. Elle est constituée de plusieurs couches de neurones en série (Figure 1-7). Elle porte ce nom car elle généralise la topologie du perceptron à des configurations contenant plus de 2 couches. On peut y distinguer une couche d'entrée, une ou plusieurs couches cachées, et une couche de sortie. Le nombre de neurones par couche peut être différent d'une couche à l'autre. Chaque neurone d'une couche est connecté via des synapses à tous les neurones de la couche suivante (on parle de réseau entièrement connecté). Le $k^{\text{ème}}$ ensemble des poids synaptiques $\{w_{j,i}^{(k)}\}$ entre deux couches successives contenant i et j neurones respectivement est représenté par une matrice de poids $\widehat{W}_k$. Dans le cas d'un réseau constitué de neurones formels, les neurones sont définis par une fonction d'activation φ , a priori commune à tous les neurones, et des biais θ propres à chacun. Les données de sortie, représentées par un vecteur $\vec{Y}$, sont calculées en fonction des données d'entrée, représentées par un vecteur $\vec{X}$, et de l'ensemble des paramètres w et θ du réseau, associés aux synapses et aux neurones. De ce point de vue, le réseau de neurone est une fonction non-linéaire de R^n dans R^m où n et m sont les dimensions des vecteurs $\vec{X}$ et $\vec{Y}$ respectivement. A l'état initial, les paramètres w et θ sont fixés arbitrairement. De façon générale, la phase d'apprentissage consiste à faire converger le jeu de paramètres vers un ensemble spécifique qui définit la fonction de transfert souhaitée.

Les perceptrons multi couches implémentent un type d'ANN appelé classificateur. Après la phase d'apprentissage, ces ANN classent les vecteurs de données d'entrée $\vec{X}$ selon ce qu'ils représentent. Pour ce faire, chaque dimension du vecteur de sortie $\vec{Y}$ est associée à une classe. L'apprentissage consiste à trouver le jeu de paramètre du réseau $\{w, \theta\}$ qui maximise la coordonnée y_i et minimise les

¹ La période réfractaire est une période de quelques millisecondes pendant laquelle le neurone ne peut plus être excitée. Elle advient pendant les instants qui suivent la production du potentiel d'action, et permet d'éviter l'emballement du réseaux soumis à une sollicitation trop intense.

coordonnées $\{y_{j \neq i}\}$ du vecteur $\vec{Y}$ lorsque des vecteurs $\vec{X}$ différents appartenant à la $i^{\text{ème}}$ classe sont présentés en entrée du réseau.

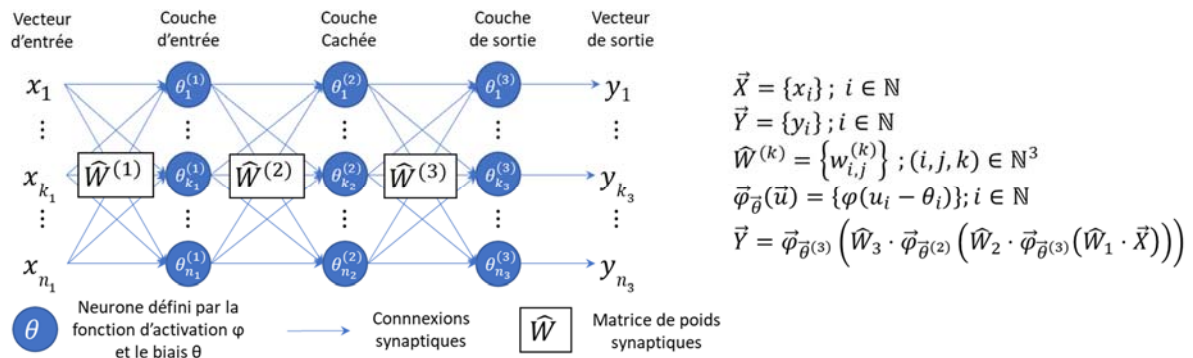


Figure 1-7 – Schéma illustrant la topologie du perceptron multi couches à trois couches et fonctions mathématiques correspondantes.

Les réseaux de neurones convolutifs :

Dans les CNN, certaines couches de neurones effectuent des extractions de caractéristiques sur leur couche précédente par l'application d'une convolution. Du point de vue des connexions synaptiques, cela se traduit par la restriction du champ de réception des neurones, et le partages de poids synaptiques. Les CNN sont très largement utilisés pour l'apprentissage profond (« deep learning ») [31]–[34] qui est à l'origine des avancées significatives en intelligence artificielle depuis les années 2010, dans le champ des techniques de reconnaissance de motif en général, et d'images en particulier.

1.2.2.2 Les RNN

De façon générale, les RNN sont des ANN dont la topologie contient des cycles. Il existe de nombreux types de RNN comme les réseaux de Hopfield (« Hofield network » : HN) [35], [36], les réseaux « long short term memory » (LSTM) [37]–[39], les « liquid state machines » (LSM) [40]–[42] ou les « echo state networks » (ESN) [43]–[45]. De par leur comportement dynamique (à l'exception des HN), ils sont particulièrement adaptés au traitement de flux de données. Ils trouvent leur application dans des champs comme la reconnaissance de discours [38], [46], la reconnaissance scripturale [47], [48], la robotique [49] ou la détection d'anomalies dans les séries temporelles [50], [51] ou leur prédiction à court terme [52]. D'autres types d'application sont listés dans [53].

Les LSM et les ESN sont composés d'une couche de neurones d'entrée, d'une couche dite de « réservoir » dont les neurones sont interconnectés de façon récursive et dont les poids synaptiques sont fixés aléatoirement, et d'une couche de neurones de sortie. La différence entre les LSM et le ESN réside dans le type de neurone utilisé : neurone à déclenchement dans le premier cas et neurone formel dans le second. Pendant l'apprentissage, seuls les poids synaptiques des couches de sortie sont entraînés. A cause de cette caractéristique particulière, ces systèmes sont les supports d'un type de traitement de données appelé « reservoir computing » [54].

1.2.3 Les règles d'apprentissage

Pour fonctionner, un ANN doit être entraîné pendant une phase dite d'apprentissage. Cet apprentissage peut être supervisé ou non-supervisé selon qu'il requiert l'intervention d'une entité extérieure ou selon qu'il apprend par sa simple exécution. Les stratégies d'apprentissages dépendent fortement du type de neurones, formels ou impulsionnels, qui constituent le réseau.

1.2.3.1 Apprentissage dans les réseaux de neurones formels

Les réseaux de neurones formels sont aujourd'hui les plus matures. Ils ont été les plus largement étudiés et développés, notamment parce qu'ils peuvent être émulés efficacement sur des calculateurs standards. De nombreuses stratégies d'apprentissage, supervisé ou non, ont été développées pour eux. Beaucoup d'entre elles reposent sur l'algorithme de rétropropagation du gradient (« back propagation algorithm ») introduit par Rumelhart en 1986 [12] et dont l'efficacité est à l'origine de sa large diffusion.

La rétropropagation du gradient :

La méthode de rétropropagation du gradient repose en premier lieu sur le calcul d'une fonction d'erreur² ε . Dans le cadre d'un apprentissage supervisé, elle est une fonction de la différence entre les normes des vecteurs de sorties cibles $\{\vec{y}_i^{(t)}\}$ et inférés $\{\vec{y}_i\}$. Sa forme la plus commune est celle issue de la méthode des moindres carrés $\varepsilon = \sum_i \|\vec{y}_i - \vec{y}_i^{(t)}\|^2$, mais d'autres formes sont possibles [55]. Dans le cadre d'apprentissages non-supervisés, les vecteurs $\{\vec{y}_i^{(t)}\}$ sont générés selon un certain jeu de contraintes. ε dépend des paramètres du réseau $\{w_j; \theta_j\}$ (poids synaptiques et biais neuronaux) au travers des $\{\vec{y}_i^{(t)}\}$. L'apprentissage consiste à minimiser ε par descente de gradient en mettant à jour les paramètres selon³ $\vec{w} \otimes \vec{\theta} \leftarrow \vec{w} \otimes \vec{\theta} - \lambda \cdot \vec{\nabla}_{\vec{w} \otimes \vec{\theta}}(\varepsilon)$, où λ est une constante appelée taux d'apprentissage, et $\vec{\nabla}_{\vec{w} \otimes \vec{\theta}}(\varepsilon)$ la divergence de ε . L'algorithme de rétropropagation est une méthode de calcul des $\partial \varepsilon / \partial w_j$ et $\partial \varepsilon / \partial \theta_j$ optimisée pour la structure des réseaux de neurones formels. Elle consiste à calculer d'abord les dérivées partielles qui concernent les paramètres de la dernière couche du réseau, puis de remonter de couche en couche vers le début du réseau, en retro-propageant l'erreur. La force de cette méthode réside dans sa généralisation : elle s'adapte à n'importe quelle structure de réseaux FNN [56], dont les CNN [57].

Cette méthode a été adaptée aux RNN [58], [59], notamment sous le nom de « backpropagation through time » [60], [61]. Cependant, l'application de cette méthode aux RNN donne lieu aux problèmes de disparition ou d'explosion du gradient (« gradient vanishing » ou « exploding gradient ») [62]–[64]. Le premier empêche le réseau de capturer les dépendances à long terme, le second rend le système très sensible au bruit. Les travaux pour résoudre ces problèmes ont donné naissance aux réseaux LSTM évoqués à la section 1.2.2.2 dont l'efficacité a permis l'émergence de nombreuses applications.

Apprentissages supervisés et non-supervisés :

L'apprentissage supervisé implique la construction de jeux de données labellisées (couple de données d'entrée $\{\vec{x}_i\}$ et de vecteur de sortie cible $\{\vec{y}_i^{(t)}\}$). Le calcul de la fonction d'erreur est alors évident et l'application de la méthode de rétropropagation du gradient est applicable de façon générale.

Dans le cadre de l'apprentissage non-supervisé, des implémentations de règles de Hebb – i.e. de règles d'apprentissage locales – ont été tentées. Pour pallier aux problèmes d'instabilité de la première formulation proposée par Hebb [8], des variantes ont été développées comme, par exemple, la règle de Oja [65] ou la règle de Sanger [66]. Cependant, peu de développements de ces stratégies sont présents dans la littérature actuelle, signe de leur efficacité limitée.

² La fonction d'erreur porte aussi le nom de fonction de perte (« loss function ») ou fonction de coût (« cost function »).

³ Le symbole $\otimes$ correspond au produit tensoriel.

Les stratégies les plus utilisées aujourd'hui consistent en la génération de vecteur cible $\{\vec{y}_i^{(t)}\}$ par un jeu de contraintes globales. De cette façon, elles permettent de tirer profit des avancées déjà effectuées dans le cadre de l'apprentissage supervisé tel que la rétropropagation du gradient et ses diverses optimisations. Souvent, les contraintes s'expriment sous la forme de structures de réseau spécifiques qui donnent lieu à des sous-types d'apprentissages comme notamment le « deep reinforcement learning », les « deep belief networks » ou encore l'« autoencoding ».

1.2.3.2 Apprentissage dans les réseaux de neurones impulsionnels

Contrairement aux neurones formels, les SN ont un codage temporel. Leur activité n'est pas codée dans une valeur de leur fonction d'activation, mais dans le train d'impulsions qu'ils génèrent. Chaque impulsion constitue un évènement repéré dans le temps. Dès lors, plusieurs types de codage peuvent être envisagés : fréquence moyenne d'émission d'impulsions, nombre ou ordre d'arrivée des impulsions émises [67]. Ces spécificités de codage nécessitent des mécanismes d'apprentissage adaptés.

Apprentissage non-supervisé : STDP et WTA :

Les codages temporels des réseaux de neurones impulsionnels (« spiking neural network » : SNN) rendent les algorithmes d'apprentissage de type rétro propagation assez inefficaces. En revanche, ils sont particulièrement adaptés aux règles d'apprentissages de type Hebbien pour la modification des poids synaptiques (plasticité synaptique) [68], [69]. La plus connue de ces règles est la plasticité fonction du temps d'occurrence des impulsions (« spike-timing-dependent plasticity » : STDP) qui est observée dans les réseaux de neurones biologiques [70]. Les poids synaptiques varient d'une quantité Δw selon l'écart temporel entre les impulsions pré et post synaptique $\Delta\tau$. La fonction $\Delta w = f(\Delta\tau)$ peut avoir différentes formes. Généralement, la synapse est renforcée (son poids synaptique augmente) lorsque le neurone pré-synaptique se déclenche juste avant le neurone post-synaptique et elle est dépréciée dans le cas contraire. L'amplitude du changement est d'autant plus forte que l'intervalle du temps est court (Figure 1-8).

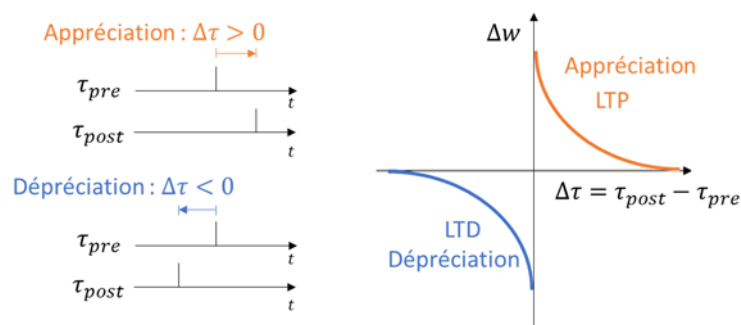


Figure 1-8 – Variation du poids synaptique Δw en fonction de l'intervalle de temps $\Delta\tau$ qui caractérise la STDP. $\Delta\tau < 0$ conduit à une dépréciation à long terme (« long term depression » : LTD), $\Delta\tau > 0$ conduit à une appréciation à long terme (« long term potentiation » : LTP)

La structure du réseau peut également jouer un rôle dans l'apprentissage non-supervisé par la présence de connexion inhibitrice entre neurones d'une même couche. En particulier, la structure « winner-take-all » (WTA) permet, à l'échelle d'une couche ou d'une assemblée de neurones, que seul un neurone se déclenche à la fois. De cette façon, chaque neurone de l'assemblée se spécialise dans la réaction à un type de train d'impulsion provenant d'autres assemblées.

Apprentissage supervisé :

La méthode d'apprentissage supervisé la plus répandue dans les SNN consiste à présenter au réseau les vecteurs d'entrée que l'on souhaite classer et à forcer le déclenchement des neurones que l'on souhaite associer à ces vecteurs, pendant une durée suffisamment longue pour que les mécanismes de plasticité synaptique agissent.

1.3 Implémentations matérielles des ANN

Certaines architectures de processeur sont particulièrement pertinentes pour implémenter des ANN. Elles diffèrent selon qu'il s'agit d'ANN formel ou de SNN.

1.3.1 Les GPU pour les ANN formels

Les calculateurs classiques (architectures de type von-Neumann) sont adaptés pour l'implémentation de réseaux de neurones formels et de leurs algorithmes d'apprentissage et d'inférences. Cependant, l'architecture des unités de traitement graphique (« graphical processor unit » : GPU), capable d'exécuter des opérations simples sur des nombreuses données en parallèle, est particulièrement efficace pour les calculs matriciels effectués dans les ANN formels [34]. Aujourd'hui, les datacenters dédiés au deeplearning intègrent préférentiellement ces processeurs. C'est d'ailleurs une des raisons majeures de la forte croissance de la compagnie NVIDIA ces dernières années, constructeur historique de cartes graphiques initialement dédiées aux jeux vidéos 3D, qui propose des processeurs entièrement dédiés au deep-learning.

1.3.2 Les processeurs « event-based » pour les SNN

Depuis le milieu des années 2000, les avancées dans le développement des circuits VLSI hybrides analogiques, digitales ou mixtes de SN permettent la construction de processeurs neuromorphiques au sens entendu par Carver Mead, où chaque neurone et chaque synapse du système est un sous-circuit physiquement localisé sur la puce. La capacité de calcul de ces systèmes est basée notamment sur la transmission d'évènements – les impulsions émises par les neurones – entre les neurones et synapses. De ce fait, ils portent le nom de processeur neuronal basés sur les évènements « event-base neural processor ». Ils sont des systèmes intrinsèquement asynchrones. Leur grande efficacité énergétique comparée aux SNN simulés sur des processeurs standards ou aux ANN formels est leur principal atout.

Le développement des SNN et leur étude s'articulent autour de deux principaux centres d'intérêt. Tout d'abord, un intérêt scientifique pour comprendre les propriétés de traitement de l'information des systèmes nerveux biologiques. En effet, l'émulation de modèles de réseaux de neurones biologiques à grande échelle par les SNN offre un outil sans précédent pour l'étude et la compréhension du fonctionnement des systèmes nerveux biologiques, et la découverte des mécanismes responsables de l'émergence de l'intelligence et de la cognition. D'autre part, un intérêt applicatif d'ingénierie. Si aujourd'hui, les applications de rupture des systèmes neuromorphiques n'ont pas encore été identifiées [71], les SNN suscitent une grande attente concernant leur application dans des tâches de reconnaissance et de classification en temps réel d'images, de discours ou de données en général, de contrôle d'automate, et de détection d'anomalies. On suppose également qu'ils trouveront de nombreuses applications dans les technologies de prochaine génération, comme les voitures autonomes, les drones ou les interfaces cerveau-machine [72].

Tableau 1-1 – Comparaison des processeurs neuronaux « event-based », d'après [72].

Chip name	Technology	Process (nm)	Neurons type	No. neurons	No. synapse	Area per neuron	Energy per event	Synaptic plasticity
MNIFAT	Mixed Signal	500	LIF/ M-N	6,120	–	1,495 μm^2	360 pJ	Programmable
DeepSouth	Digital	28	LIF	200 K	–	–	–	No Plasticity
Dynap-SEL	Mixed Signal	28	I&F	1,088	78,080	20 μm^2	2.8 pJ	STDP
BrainScaleS	Mixed Signal	180	AdEx IF	512	100 K	1,500 μm^2	100 pJ	Hebbian learning, STDP
2DIFWTA	Analog	350	I&F	2,048	28,672	–	–	No Plasticity
HiAER-IFAT board with 4 chips	Analog	90	I&F	256 K	256 M	140 μm^2	22 pJ	No Plasticity
Transistor-Channel	Analog	350	Floating Gate MOSFET	100	30,000	4 cm^2	10 pJ	STDP
Neurogrid	Mixed signal	180	Adaptive Quad IF	65 K	100 M	1,800 μm^2	31.2 pJ	No Plasticity
TrueNorth	Digital	28	Adaptive Exp IF	1 M	256 M	3,325 μm^2	45 pJ	STDP
Spinnaker	Digital	130	Programmable	16 K	16 M	–	43 nJ	STDP
Loihi	Digital	14	Adaptive LIF	130 K	130 M	0.4 mm^2	23.6 pJ	Epoch-based, STDP

Les principales caractéristiques des processeurs neuronaux les plus aboutis sont présentées au Tableau 1-1. La majorité de ces systèmes sont développés par des centres de recherche académique et deux d'entre eux sont proposés par de grandes compagnies industrielles (TrueNorth par IBM et Loihi par Intel). Ces puces sont plus ou moins récentes comme le montrent les différents nœuds technologiques avec lesquels elles ont été fabriquées (colonne « process » du Tableau 1-1). Leurs objectifs applicatifs sont également différents. Les processeurs neuronaux peuvent être classés en trois catégories principales selon la nature de leur technologie : digitale, analogique ou mixte.

Du point de vue fonctionnel, les processeurs neuronaux sont composés de trois types de sous-systèmes : les circuits de neurones, les circuits de synapses et le système de communication qui assure la transmission entre les neurones et les synapses.

1.3.2.1 Les systèmes de communication AER

Les processeurs neuronaux implémentent un très grand nombre de connexions synaptiques au travers desquels sont acheminés les événements de déclenchement des neurones. Pour pallier l'impossibilité de dédier, dans la puce électronique, une ligne de transmission par connexion, un système de multiplexage temporel, basé sur un protocole appelé représentation d'adresse d'évènement (« adress event representation » : AER) est généralement utilisé. Ce multiplexage est permis par la différence des fréquences de fonctionnement caractéristiques des circuits de neurones – typiquement le kHz – et des circuits digitaux des technologie actuelles – typiquement le GHz.

Les systèmes AER sont composés de quatre types d'élément fonctionnel : les blocs digitaux d'encodage, ceux de décodage, ceux de routage et les lignes de transmissions. Les neurones et leurs synapses afférentes sont organisés par groupes connectés chacun à un encodeur à leur sortie et un décodeur à leur entrée. Les impulsions de déclenchement collectées en sortie des neurones par le bloc encodeur sont transmises au bloc routeur le plus proche via les lignes de transmission sous la forme d'adresses qui correspondent aux neurones émetteurs. Toujours via des lignes de transmission, le bloc routeur envoie ensuite ces adresses aux autres blocs routeurs concernés pour la transmission selon une carte du réseau mémorisée. Les adresses arrivent finalement au bloc de décodeur de destination qui se charge d'activer les synapses adéquates des neurones cibles.

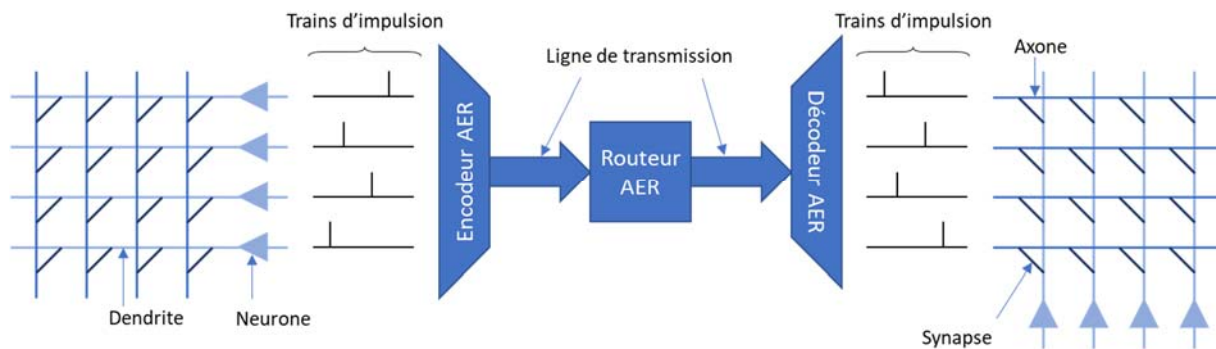


Figure 1-9 – Schéma illustrant le système AER.

La brève description ci-dessus laisse entrevoir les nombreuses problématiques que les systèmes AER doivent prendre en charge. Comment coder les adresses ? Comment coder la carte du réseau ? Comment gérer les conflits lorsque, par exemple, deux neurones d'un même paquet s'activent en même temps ? Différentes stratégies sont adoptées selon les réponses données à ces questions [73].

1.3.2.2 Les circuits de synapses

Les circuits de synapse pondèrent les impulsions qui sont sommées par les neurones. Pour implémenter les mécanismes de plasticité synaptique et permettre au système d'apprendre, ces pondérations doivent évoluer au cours du temps. Elles sont à la fois une mémoire et une fonction de transfert.

De nombreuses stratégies existent pour leur implémentation. Elles dépendent des contraintes fonctionnelles que les concepteurs leur imposent. Sont-elles cointégrées dans un système digital, analogique ou mixte ? Sont-elles le support de l'algorithme d'apprentissage, ou ce dernier est-il effectué hors de la puce et les poids synaptiques simplement chargés lors d'une phase de mise à jour ? Si l'algorithme d'apprentissage est exécuté sur la puce, quelle règle implémente-t-il ? Cet algorithme est-il reconfigurable ? Les poids synaptiques sont-ils conservés lorsque la puce est mise hors tension ? Combien de valeurs différentes ces poids prennent-ils ?

Selon les réponses données à ces questions, les circuits de synapses ont des degrés de complexité divers et des spécifications différentes, en termes de consommation énergétique et de surface de silicium notamment. Ce dernier point est critique, car, de par leur plus grand nombre, la surface de silicium dédiée aux circuits synaptiques est généralement la plus étendue.

Il existe plusieurs technologies pour l'implémentation du point mémoire. Certaines d'entre-elles sont compatibles avec les technologies CMOS déjà fortement matures de par leur utilisation massive par l'industrie, comme les mémoires vives statiques (« static random-access memory » : SRAM), ou les transistors à grille flottante. D'autres nécessitent l'intégration de nouveaux matériaux et donc de nouvelles étapes de procédé, comme les memristors.

- **SRAM** : Pour les systèmes entièrement digitaux, les SRAM sont le support naturel des poids synaptiques. Très largement utilisées dans tous les processeurs standards, elles bénéficient de plusieurs décennies d'optimisation par les compagnies industrielles. Un point mémoire SRAM est un circuit d'au moins 6 transistors capable de stocker un bit. Il peut fonctionner à basse tension – quelques centaines de millivolts – et dispose d'une bonne densité d'intégration (moins de 1 μm^2 par SRAM). Son principal défaut est sa volatilité : l'information est perdue lorsque le circuit est mis hors tension. Le stockage d'un poids synaptique pouvant prendre N valeurs nécessite $\log_2(N)$ points mémoires. Pour maximiser la densité d'intégration, le choix se

porte souvent sur des synapses binaires qui n'ont besoin que d'une seule SRAM par synapse [74].

- Transistor à grille flottante : Les transistors à grille flottante sont les composants à la base des mémoires FLASH non-volatiles digitales utilisées aujourd'hui dans les disques durs à semiconducteur (« solide state drive » : SSD). Bien qu'utilisées à l'origine pour stocker des bits (0 ou 1), il est possible de les faire fonctionner de manière analogique pour implémenter des synapses [75], [76]. Comme leur fonctionnement repose sur le franchissement par les électrons de barrières de potentiel élevées, elles nécessitent l'application de tensions de plusieurs volts. C'est un inconvénient pour la conception de circuits faible consommation. La valeur du point mémoire est donnée par le décalage de tension de seuil du transistor.
- Les FeRAM : Le support du point mémoire des FeRAM est la polarisation d'un matériau ferroélectrique. Généralement intégré dans l'empilement de grille, la valeur du point mémoire est donnée par la mesure de la tension de seuil du transistor, comme pour les transistors à grille flottante.
- Les memristors : Les memristors sont des résistances variables non-volatiles : l'information est stockée dans la valeur de leur résistance [77], [78]. Leur fonctionnement s'appuie sur des mécanismes singuliers de la physique des matériaux à l'échelle nanométrique. Cette échelle permet a priori de construire des points mémoires à très haute densité d'intégration. De plus, ils peuvent être analogiques (capables de stocker un continuum de valeurs). Ils suscitent un grand intérêt pour l'ingénierie neuromorphique. Tout d'abord, du point de vue conceptuel, l'exploitation de mécanismes physiques à très petite échelle correspond parfaitement à l'esprit du neuromorphisme tel qu'évoqué par Carver Mead, dont la philosophie est de s'appuyer sur les analogies entre les mécanismes biochimiques du vivant et les mécanismes électro-physiques de la microélectronique. Ensuite, leur forte densité d'intégration est un atout pour l'implémentation des synapses qui sont les éléments les plus nombreux dans une puce neuronale. Enfin, les mécanismes de leur programmation par l'application d'impulsion de tension ou de courant sont compatibles avec les SNN. Plusieurs types de memristors sont développés aujourd'hui. Ils diffèrent par la nature des matériaux qui les constituent et la nature des mécanismes physiques mis en jeu :
 - Les OxRAM : Dans une OxRAM (« oxyde RAM »), l'information est codée par la constitution réversible d'un filament conducteur obtenue par l'agrégation de lacunes d'oxygène, dans une couche isolante d'oxyde. La présence ou l'absence de ce filament modifie la résistance de la couche d'oxyde.
 - Les PCRAM : Dans une PCRAM (« phase change RAM »), l'information est codée par la phase, amorphe ou cristalline, d'un matériau généralement chalcogénide [79]. Une phase amorphe aura une plus grande résistivité qu'une phase cristalline. La nature de la phase est déterminée par la forme de l'impulsion de courant appliqué pour faire fondre localement le matériau. Selon cette forme (durée, amplitude, temps de montée, temps de descente) le matériau se trouvera en forme amorphe ou cristalline après refroidissement.
 - Les CBRAM : Le principe de fonctionnement des CBRAM (« conductive bridging RAM ») est similaire à celui des OxRAM : un filament conducteur est créé dans une couche isolante – d'oxyde métallique en l'occurrence. Ce filament est ici créé par la diffusion d'atomes de métal depuis l'anode vers la cathode à travers l'oxyde. Cette diffusion est autorisée par un mécanisme d'oxydoréduction.
 - Les MRAM : Les MRAM (« magnetic RAM ») exploitent la valeur de résistance entre deux couches magnétiques séparées par une jonction tunnel. Cette valeur dépend de l'alignement des spins de ces deux couches, dont l'une est figée et l'autre manipulable.

Une revue plus exhaustive des composants développés pour l'implémentation des synapses se trouve dans [71].

1.3.2.3 Les circuits de neurones

Depuis l'introduction du neuromorphisme par Carver Mead, de nombreux designs de circuit de neurone ont été proposés dans la littérature. De façon schématique, deux approches différentes sont exploitées pour leur conception, l'une bio-réaliste et l'autre simplificatrice.

L'approche bio-réaliste tend à reproduire la diversité et la complexité du comportement des neurones biologiques de la façon la plus fidèle possible, en particulier par la modélisation de la dynamique des canaux ioniques proposée par Hodgkin et Huxley (modèle H-H) [24]. Elle conduit à des circuits d'une certaine complexité qui nécessitent l'application de nombreuses tensions de contrôle [80]–[82] et qui souffrent d'une faible densité d'intégration. Elle est d'un grand intérêt pour l'exécution de simulation en neurobiologie.

L'approche simplificatrice cherche à réduire le circuit à une fonction primaire, généralement d'intégration et de déclenchement (« integrate and fire » : IF). Les circuits conçus de cette façon ne sont pas capables de reproduire la diversité des comportements des neurones biologiques mais présentent une certaine simplicité et peuvent être intégrés à plus grande échelle. Un exemple caractéristique de ce type de circuit est le « Axon-Hillock circuit », proposé à l'origine par Carver Mead (Figure 1-10), implémenté dans des circuits de rétine artificielle notamment [83], [84], ou des processeurs neuronaux [85] et optimisé récemment pour atteindre une très faible consommation [86].

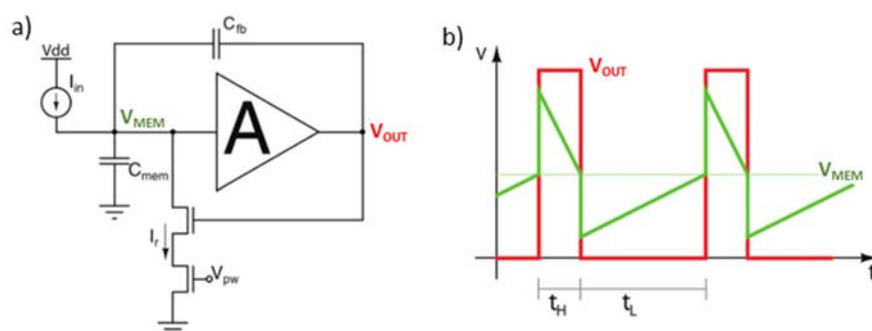


Figure 1-10 – (a) Circuit Axon-Hillock et (b) chronogramme correspondant, depuis [87] et d'après [1].

De nombreux circuits sont le fruit d'une synthèse entre ces deux approches. Par un jeu de compromis, ils tentent de tirer profit des avantages de chacune d'elles. Ils peuvent s'appuyer sur des modèles de neurones dont les équations différentielles sont plus simples et en moindre nombre que celles introduites par le modèle H-H de référence [88], comme les modèles de FitzHugh-Nagumo [89], Hindmarsh-Rose [90] ou Morris-Lecar [91]. Certains intègrent la discontinuité présente dans le modèle IF de base par la réinitialisation de variable comme le modèle d'Izhikevich [21] ou le modèle « adaptive exponential IF » [16]. Ils sont souvent regroupés sous l'appellation « generalized IF ».

Le Tableau 1-2 répertorie une sélection de modèle de neurones, ainsi que les références où trouver leur description mathématique et leur implémentation matérielle. Le lecteur trouvera dans [71] une revue de ces modèles davantage exhaustive (Figure 1-11).

Les circuits évoqués jusqu'à présent utilisent des méthodes de design classiques qui intègrent des composants standards (transistors PMOS et NMOS et capacités). Ces méthodes exploitent les correspondances entre les équations qui décrivent les comportements courant-tension des

composants et les équations des modèles de neurones à implémenter [92], [111], [115]–[117]. Une autre approche exploite les mécanismes physico-électroniques spécifiques de composants moins conventionnels. Elle est similaire à celle qui a conduit à l'utilisation de mémoires résistives pour l'implémentation des synapses. L'ambition de cette stratégie est de minimiser au maximum la taille du circuit de neurone. Dans cet état d'esprit, on trouve dès la fin des années 1980 des travaux qui exploitent la présence d'impulsion dans le courant direct d'une diode PIN refroidie à très basse température et couplée à un circuit RC [118], [119]. Plus récemment, l'effet électrostatique du stockage temporaire de trous générés par ionisation par impact dans des composants silicium a été exploité pour la fabrication de « composants neuronaux », dans une structure NPN sur SOI optimisé [120], dans un PD-SOI MOSFET [121], [122] et dans une structure NIPIN [123].

Tableau 1-2 – Modèles de neurones IF et références où trouver leur description mathématique et leur implémentation matérielle.

Nom	Modèle mathématique	Implémentation physique
Hodgkin-Huxley	[24]	Analogique : [80]–[82], [92]
Generalized LIF	[14]	Analogique : [93]–[102] Digitale : [74], [101]
aEIF	[16]	Analogique : [103], [104] Digitale : [105] Mixte : [106]
Izhikevitch	[21]	Analogique : [107]–[110]
Axon-Hillock	-	Analogique : [1], [85], [86]
Morris-Lecar	[91]	Analogique : [111]
Saddle-node	[112]	Analogique : [112]
Quartic Model	[113]	Digitale : [114]

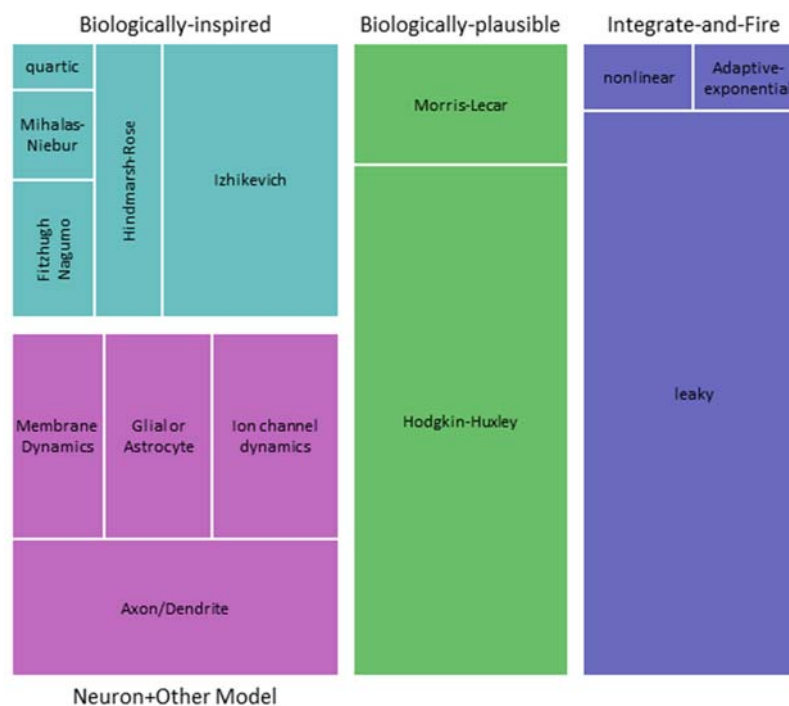


Figure 1-11 – Liste de modèles de SN dont on trouve une implémentation matérielle dans la littérature. La taille des boîtes correspond au nombre d'articles publiés et leur couleur à la famille de modèle à laquelle ils appartiennent. Depuis [71].

Certains composants à la base des nouvelles mémoires résistives et utilisés notamment pour l'implémentation de synapses, sont aussi étudiés pour la fabrication de neurones. Par exemple, un matériau à changement de phase, développé à l'origine pour les mémoires PCRAM est exploité comme élément intégrateur des impulsions synaptiques [124]. On trouve également de nombreux travaux

basés sur l'utilisation de composants magnétiques utilisés pour les MRAM [125]–[129]. Du point de vue de l'industrialisation, le point faible de ces approches est le surcoût qu'elles engendrent par la nécessité d'intégrer de nouveaux matériaux et donc, des étapes supplémentaires dans le procédé de fabrication.

1.4 État de l'art et figure de mérite des circuits analogiques de neurone impulsif

Dans cette section, l'état de l'art de circuit de neurone est d'abord présenté par un tableau qui reporte les différentes figures de mérite et caractéristiques des contributions sélectionnées. Ensuite, ces figures de mérite et caractéristiques sont discutées.

1.4.1 État de l'art

Depuis l'introduction du neuromorphisme par Carver Mead à la fin des années 1980, de nombreuses conceptions de circuit analogique de SN ont été proposées. Le Tableau 1-3 répertorie une sélection de ces contributions.

Tableau 1-3 – Caractéristiques et figures de mérite des circuits analogiques de SN de l'état de l'art.

Auteurs	Modèle de neurone	Technologie	Tension d'alimentation	Nombre de composant	Surface du circuit*	Valeur des capacités**	Surface des Capacités	Fréquence de fonctionnement	Consommation statique	Consommation dynamique	Efficacité énergétique*	Maturité
Danneville et al., 2019 [86]	Axon-Hillock	65 nm	0.2 V	6 : 5T1C	31 μm^2	CF = 4 fF	-	[0.29 ; 15.6] kHz	11 pW	30 pW à 15,7 kHz	2 fJ/spike	Circuit (silicium)
Sourikopoulos et al., 2017 [111]	Morris-Lecard based	65 nm	0.2 V	6 : 6T	35 μm^2	CM = 4 fF, CK = 8 fF	-	[- ; 25] kHz	25 pW*	100 pW à 25 kHz	4 fJ/spike à 25 kHz	Circuit (silicium)
Sourikopoulos et al., 2017 [111]	Morris-Lecard based	65 nm	0.2 V	10 : 8T2C	200 μm^2	CM = 50 fF, CK = 100 fF	-	[0.02 ; 1,2] kHz	18 pW*	94 pW à 1.2 kHz	78.3 fJ/spike à 1.2 kHz	Circuit (silicium)
Qiao et Indiveri, 2016, 2017 [130], [131]	Adaptive exponential I&F	28 nm FD-SOI	1 V	52T3C	50 μm^2	CM = 0.5 pF, CA = 0.2 pF, CR = 0.2 pF	50 μm^2	[0 ; 12] kHz	-	1.5 nW à 30 Hz	50 pJ à 30 Hz	Système
Cruz-Albrecht et al., 2012 [100]	LIF	90 nm	0.6 V	16T1C1R	442 μm^2	-	-	100 Hz	-	40 pW à 100 Hz	78.3 fJ à 100 Hz	Système
Joubert et al., 2011 [99]	LIF	65 nm	1.2 V	32T1C	538 μm^2	500 fF	100 μm^2	[1 kHz ; 1.9 MHz]	-	78 μW à 1.9 MHz	41 pJ à 1.9 MHz	Circuit
Basu and Hasler, 2010 [112]	« Saddle »	0.35 μm	-	8T2C	2740 μm^2 (optimisé 1887 μm^2)	CK = 300 fF	-	100 Hz	-	1.74 nW	17.4 pJ	Circuit
Wijekoon and Dudeck, 2009 [107]	Izhikevich	0.35 μm	3.3 V	25 : 23T2C	4900 μm^2	-	1200 μm^2 800 μm^2	[1 Hz ; 1 kHz]	-	-	-	Circuit (simulation)
Wijekoon and Dudeck, 2008 [132]	Izhikevich	0.35 μm	3.3 V	14T2C	2800 μm^2	CM : 100 fF, CK : 1 pF	-	1 MHz	-	8 – 40 μW	8,5 – 9 pJ	Circuit
Indiveri, 2006 [102]	Optimized LIF	AMS 0.8 μm	5.5 V	-	2573 μm^2	CM : 432 fF	244 μm^2	200 Hz	-	-	900 pJ	Système
Lee et al., 2004 [116]	Hindmarsh and Rose	0.25 μm	2 V	90T	-	-	-	10 kHz	-	163.4 μW	16.34 nJ/spike	Circuit (simulation)

* Figures de mérite principales

** Les capacités contribuent fortement à l'augmentation de la surface du circuit

1.4.2 Les figures de mérite

Les performances des circuits analogiques de SN sont généralement évaluées selon deux figures de mérite principales : la surface de leur empreinte sur le silicium et leur efficacité énergétique. En revanche, leur fréquence de fonctionnement n'est pas un critère significatif. En effet, la pertinence de sa valeur dépend du type d'application auquel le circuit est destiné. De faibles fréquences proches de celles des neurones biologiques (Hz – KHz), seront adaptées au traitement de données physiques en temps-réel, fournies par des capteurs. De hautes fréquences (kHz-MHz) conviendront davantage aux traitements de données par lots ou aux simulations accélérées de réseaux bio-réalistes [87].

1.4.2.1 Surface de l'empreinte du circuit

Généralement, le concepteur d'un circuit cherche à réduire la surface de son empreinte du circuit afin d'en intégrer un maximum sur une surface donnée. La taille de cette empreinte dépend, d'une part, du nombre de composants qui constituent le circuit et d'autre part, de leur dimensionnement.

Le nombre de composant dépend :

- De la diversité et de la complexité des comportements neuronaux que le circuit cherche à reproduire – autrement dit, du modèle de neurone à implémenter.
- De la stratégie d'implémentation. Par exemple, l'utilisation systématique d'amplificateurs opérationnels pour générer toutes les fonctions d'un modèle est une stratégie robuste et fiable mais qui maximise le nombre de composants [116].

Le dimensionnement des composants dépend :

- Du nœud technologique dans lequel ils sont fabriqués. A ce titre, à architecture identique, les circuits les plus anciens, et donc intégrés dans les technologies les plus lâches, seront plus étendus, donc moins performants du point de vue de la densité d'intégration.
- Des spécifications du circuit. Ainsi, une conception faiblement consommatrice d'énergie nécessitera une augmentation de la taille des transistors (des longueurs de canal en particulier) et donc une augmentation de la surface du circuit. De la même façon, une conception émettrice d'impulsions à basse fréquence s'appuiera sur l'augmentation des constantes de temps caractéristiques du circuit par l'augmentation de la taille des composants capacitifs et résistifs.

Notons que les capacités sont en général les composants qui occupent la plus grande surface. La diminution de leur valeur est un enjeu important dans la réduction de la surface totale du circuit.

1.4.2.2 L'efficacité énergétique

La mesure pertinente de l'efficacité énergétique d'un circuit de SN est le joule par impulsion (« joule per spike » : J/spike). Cette mesure est le pendant du FLOPS par watt pour la mesure de l'efficacité énergétique des processeurs aux architectures standards. Le J/spike permet de comparer des circuits de neurones qui ont différentes fréquences typiques d'émission d'impulsion.

Pour l'utilisation des SN pour des applications embarquées, la consommation énergétique est une problématique de premier plan. Les SNN à faible fréquence de fonctionnement sont dans ces cas particulièrement adaptés.

1.4.2.3 Autres caractéristiques

Outre les deux premières figures de mérite évoquées plus haut, d'autres caractéristiques intéressantes sont répertoriées au Tableau 1-3 :

- La technologie utilisée pour l'implémentation des circuits nous renseigne sur leurs contraintes de dimensionnement des circuits et permet de relativiser leur surface d'empreinte.
- De la même façon, la tension d'alimentation permet de mettre en perspective les consommations.
- Le nombre de composants nous renseigne sur la complexité du design.
- Les capacités sont de loin les composants qui occupent la plus grande surface. Il est donc pertinent de relever leurs valeurs et leur surface propre. Il est d'usage de noter C_M la capacité

qui émule la capacité de membrane du neurone. Pour certains modèles, d'autres capacités sont implémentées pour émuler les dynamiques requises : C_K pour la dynamique des canaux de potassium, C_R et C_A pour les dynamiques des mécanismes de période réfractaire et d'adaptation de fréquence, respectivement.

- Les consommations statiques (en l'absence de courant synaptique) et dynamiques (i.e. en fonctionnement). La consommation dynamique est égale au produit de la fréquence de fonctionnement par l'efficacité énergétique.
- Enfin, une notion de maturité est introduite. Elle distingue les travaux selon qu'ils concernent des circuits de neurones seuls, qu'ils aient été fabriqués sur silicium ou simplement simulés (maturité « circuit ») ou selon qu'ils sont intégrés dans des systèmes complets comprenant des neurones, des synapses et des circuits de communication AER (maturité « système»). Les travaux à maturité « circuit » sont à nouveau différenciés selon qu'il s'agit de circuits fabriqués sur silicium ou simplement simulés. La distinction de maturité système et circuit est importante en ce que les circuits intégrés dans des systèmes doivent intégrer des fonctions d'interfaçage avec le reste du système qui complexifient leur conception et donc la surface de leur empreinte.

Parmi les travaux présentés au Tableau 1-3, deux sous-groupes méritent une attention particulière. Tout d'abord, les contributions [102], [130], [131] issues du groupe *Neuromorphic Cognitive Systems* de l'institut de neuro-informatique de Zürich et dans lesquelles les circuits de neurone proposés sont intégrés dans des puces qui contiennent le système neuromorphique dans son entier. Ainsi, bien que leurs figures de mérite ne soient pas les plus performantes, ce sont les circuits de plus haute maturité, et de ce point de vue, ils dépassent toutes les autres contributions. Ensuite, les réalisations récentes [86], [111] du groupe *Spiking Neural Circuits and Systems* de l'université de Lille où les efficacités énergétiques des circuits dépassent celles des autres contributions d'au minimum un ordre de grandeur, en réduisant également la surface de silicium consommée. Ces performances sont atteintes par une stratégie simple et originale de réduction de la tension d'alimentation jusqu'à 200 mV.

1.5 Les objectifs de la thèse

Dans ce contexte, l'objectif de la thèse est de concevoir un circuit analogique de SN qui se positionne de façon ambitieuse face à ceux de l'état de l'art, notamment en ce qui concerne la réduction de la surface de silicium consommée par l'empreinte du circuit. La stratégie de développement s'articule autour de deux axes : d'une part, l'utilisation d'une technologie d'intégration avancée et, d'autre part, l'exploitation de mécanismes « de second ordre » de la physique du transistor semiconducteur (il s'agit en particulier de la génération de paires électron-trou, un des phénomènes parasite à l'origine de leur non-idéalité). Ces deux axes permettent en effet de minimiser la surface du circuit de neurone. Le premier par la réduction possible de la taille des composants grâce à la grande finesse de gravure. Le second par la réduction du nombre de composants nécessaires pour la réalisation de la fonction de transfert du neurone, comme cela est exposé au chapitre 4. L'exploitation de mécanismes spécifiques de la physique des composants inscrit ce travail dans la continuité des approches décrites dans [118]–[123], déjà évoquées à la section 1.3.2.3, ou de celle exposée dans [133] pour l'intégration de neurones formels.

Dans cet objectif, la technologie 28 nm FD-SOI proposée par STMicroelectronics est choisie pour l'intégration. Comme cela est expliqué au Chapitre 2, les technologies SOI minimisent la consommation en supprimant les courants de fuite par le substrat. De plus, la finesse de gravure pouvant atteindre 28 nm assure une première minimisation de l'empreinte du circuit. Pour l'utilisation de composant non-conventionnel, le point de départ du travail est la constatation que la fonction de déclenchement du

neurone IF est similaire à celle des circuits de protections contre les décharges électrostatiques (« electrostatic discharge » : ESD). Or, certains de ces circuits se basent sur l'implémentation d'un seul composant tel que le thyristor – nommé également redresseur silicium commandé (« silicon controlled rectifier » : SCR [134]) – ou le BIMOS [135]. Un circuit de neurone qui utilise un thyristor est notamment proposé dans [136]. Pour le travail présenté dans cette thèse, le choix s'est porté sur l'intégration du BIMOS pour la conception d'un nouveau circuit analogique de SN.

1.6 Conclusion

La longue évolution des organismes vivants a fait émerger un système de traitement de l'information singulier, basé sur la mise en réseaux de cellules neuronales par le biais de synapses, dont l'ensemble constitue le système nerveux. Dès les débuts de l'informatique au milieu du siècle dernier, et grâce aux découvertes alors récentes en neurobiologie, ce mode de traitement de l'information a guidé la fabrication de premiers calculateurs neuro-inspirés, dont le principe de fonctionnement diffère fondamentalement de celui des calculateurs standards, notamment par le fait que ces calculateurs doivent être entraînés pour fonctionner. C'est ainsi que les premiers ANN sont créés, basés sur l'intégration de modèles de neurones formels à fonction d'activation d'abords binaire (modèle de 1^{ère} génération), puis continue (modèle de 2^{ème} génération). À la fin des années 1980, Carver Mead approfondit cette approche par l'invention du concept de neuromorphisme. Il constate qu'il est possible de construire des circuits électroniques analogiques qui émulent le comportement des neurones et des synapses. Il s'appuie pour cela sur les similitudes entre les lois qui gouvernent les mécanismes biochimiques du vivant et les mécanismes électro physiques de la microélectronique. Il ouvre ainsi la voie aux modèles de SN, qui constituent les modèles de 3^{ème} génération. Aujourd'hui, les ANN formels sont massivement utilisés dans le domaine de l'intelligence artificielle qui a connu récemment et grâce à eux des progrès sans précédent. Les SNN, quant à eux, sont encore au stade de développement mais suscitent de grandes attentes.

Les ANN reposent sur trois éléments : le modèle de neurone, la topologie du réseau et la règle d'apprentissage :

- Le modèle de neurone peut être formel ou impulsionnel.
- Les topologies des ANN diffèrent selon les applications auxquels ils sont dédiés. Les topologies de type FNN et RNN sont plutôt adaptées au traitement de données statiques et dynamiques respectivement.
- La règle d'apprentissage peut être supervisée ou non-supervisée. Les méthodes de rétropropagation du gradient et de STDP conviennent particulièrement pour la mise à jour des poids synaptiques dans les ANN formels et impulsionnels respectivement.

L'évolution de l'état d'un ANN est calculée par un processeur. Il constitue l'implémentation matérielle du réseau. Bien que les processeurs standard, en tant que machines de Turing universelle, sont capables de simuler n'importe quel ANN, des architectures spécifiques sont davantage adaptées pour exécuter cette tâche. Ainsi, l'architecture des GPU est particulièrement performante pour l'implémentation des ANN formels. De leur côté, l'émergence des SNN a initié le développement de processeurs « event-based », généralement de type mixte analogique/digital asynchrone, qui se rapproche davantage du concept de neuromorphisme introduit par Carver Mead. A la différence des ANN formels implémentés sur GPU où l'état du réseau est stocké dans une mémoire mise à jour par les calculs du processeur, les neurones et synapses qui constituent les ANN impulsionnels sont des sous-éléments du circuit : chacun d'eux est localisé physiquement sur la puce. Dès lors, outre le

pilotage du fonctionnement de ces sous blocs élémentaires, le système a aussi pour fonction de gérer les communications entre eux au travers d'un protocole AER.

Les travaux présentés dans cette thèse portent sur le développement d'un nouveau circuit analogique de neurone pouvant servir de bloc élémentaire pour la construction d'un processeur neuronal « event-based ». La stratégie de conception repose sur l'exploitation non-conventionnelle de phénomènes de la physique des composants semi-conducteurs. Elle s'appuie sur l'intégration du BIMOS, composant utilisé notamment dans les circuits de protection ESD, en technologie 28 nm FD-SOI. Par ce moyen, la surface de l'empreinte du circuit sur le silicium peut être minimisée et dépasser ainsi de ce point de vue les circuits aujourd'hui à l'état de l'art.

1.7 Références

- [1] C. Mead, *Analog VLSI and Neural Systems*. Addison-Wesley, 1989.
- [2] Wikipedia contributors, *Neuromorphic engineering — Wikipedia, The Free Encyclopedia*. 2019.
- [3] C. Mead, “Neuromorphic electronic systems,” *Proc. IEEE*, vol. 78, no. 10, pp. 1629–1636, Oct. 1990.
- [4] J. von Neumann, “First draft of a report on the EDVAC,” *IEEE Ann. Hist. Comput.*, vol. 15, no. 4, pp. 27–75, 1993.
- [5] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, 1943.
- [6] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, 1958.
- [7] H. D. Block, “The Perceptron: A Model for Brain Functioning. I,” *Rev. Mod. Phys.*, vol. 34, no. 1, pp. 123–135, Jan. 1962.
- [8] D. O. Hebb, *The Organization of Behavior: A Neuropsychological Theory*. Wiley, 1949.
- [9] S. Lowel and W. Singer, “Selection of intrinsic horizontal connections in the visual cortex by correlated neuronal activity,” *Science*, vol. 255, no. 5041, pp. 209–212, 1992.
- [10] W. Maass, “Networks of spiking neurons: The third generation of neural network models,” *Neural Netw.*, vol. 10, no. 9, pp. 1659–1671, Dec. 1997.
- [11] Y. Le Cun, “Learning Process in an Asymmetric Threshold Network,” in *Disordered Systems and Biological Organization*, 1986, pp. 233–240.
- [12] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.
- [13] E. M. Izhikevich, “Which model to use for cortical spiking neurons?,” *IEEE Trans. Neural Netw.*, vol. 15, no. 5, pp. 1063–1070, Sep. 2004.
- [14] L. Lapique, “Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarization,” *J. Physiol. Pathololgy*, vol. 9, pp. 620–635, 1907.
- [15] L. F. Abbott, “Lapicque’s introduction of the integrate-and-fire model neuron (1907),” *Brain Res. Bull.*, vol. 50, no. 5, pp. 303–304, 1999.
- [16] R. Brette and W. Gerstner, “Adaptive Exponential Integrate-and-Fire Model as an Effective Description of Neuronal Activity,” *J. Neurophysiol.*, vol. 94, no. 5, pp. 3637–3642, Nov. 2005.
- [17] G. Ermentrout and N. Kopell, “Parabolic Bursting in an Excitable System Coupled with a Slow Oscillation,” *SIAM J. Appl. Math.*, vol. 46, no. 2, pp. 233–253, Apr. 1986.
- [18] A. Tonnelier, H. Belmabrouk, and D. Martinez, “Event-Driven Simulations of Nonlinear Integrate-and-Fire Neurons,” *Neural Comput.*, vol. 19, no. 12, pp. 3226–3238, 2007.
- [19] N. Brunel and P. E. Latham, “Firing Rate of the Noisy Quadratic Integrate-and-Fire Neuron,” *Neural Comput.*, vol. 15, no. 10, pp. 2281–2306, Oct. 2003.
- [20] N. Fourcaud-Trocmé, D. Hansel, C. van Vreeswijk, and N. Brunel, “How Spike Generation Mechanisms Determine the Neuronal Response to Fluctuating Inputs,” *J. Neurosci.*, vol. 23, no. 37, pp. 11628–11640, Dec. 2003.
- [21] E. M. Izhikevich, “Simple model of spiking neurons,” *IEEE Trans. Neural Netw.*, vol. 14, no. 6, pp. 1569–1572, Nov. 2003.
- [22] M. J. E. Richardson, N. Brunel, and V. Hakim, “From Subthreshold to Firing-Rate Resonance,” *J. Neurophysiol.*, vol. 89, no. 5, pp. 2538–2554, 2003.
- [23] A. Destexhe, M. Rudolph, and D. Paré, “The high-conductance state of neocortical neurons in vivo,” *Nat. Rev. Neurosci.*, vol. 4, p. 739, Sep. 2003.
- [24] A. L. Hodgkin and A. F. Huxley, “A quantitative description of membrane currents and its application to conduction and excitation in nerve,” *J. Physiol.*, pp. 500–544, 1952.
- [25] A. Erisir, D. Lau, B. Rudy, and C. S. Leonard, “Function of Specific K⁺ Channels in Sustained High-Frequency Firing of Fast-Spiking Neocortical Interneurons,” *J. Neurophysiol.*, vol. 82, no. 5, pp. 2476–2489, 1999.

- [26] Y. Liu and X. Wang, "Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron.," *J. Comput. Neurosci.*, vol. 10, no. 1, pp. 25–45, 2001.
- [27] Ş. Mihalas and E. Niebur, "A Generalized Linear Integrate-and-Fire Neural Model Produces Diverse Spiking Behaviors," *Neural Comput.*, vol. 21, no. 3, pp. 704–718, Oct. 2008.
- [28] R. Jolivet, T. J. Lewis, and W. Gerstner, "Generalized Integrate-and-Fire Models of Neuronal Activity Approximate Spike Trains of a Detailed Model to a High Degree of Accuracy," *J. Neurophysiol.*, vol. 92, no. 2, pp. 959–976, 2004.
- [29] Y. Lecun and Y. Bengio, "Convolutional networks for images, speech, and time-series," in *The handbook of brain theory and neural networks*, M. A. Arbib, Ed. MIT Press, 1995.
- [30] Y. Lecun, "Generalization and network design strategies," in *Connectionism in perspective*, R. Pfeifer, Z. Schreter, F. Fogelman, and L. Steels, Eds. Elsevier, 1989.
- [31] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.
- [32] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [33] M. Z. Alom *et al.*, "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches," *CoRR*, vol. abs/1803.01164, 2018.
- [34] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, 2015.
- [35] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci.*, vol. 79, no. 8, pp. 2554–2558, 1982.
- [36] W. A. Little, "The existence of persistent states in the brain," *Math. Biosci.*, vol. 19, no. 1, pp. 101–120, Feb. 1974.
- [37] S. Hochreiter and J. Schmidhuber, "Long Short-term Memory," *Neural Comput.*, vol. 9, pp. 1735–80, 1997.
- [38] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in *Fifteenth annual conference of the international speech communication association*, 2014.
- [39] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, "Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 4580–4584.
- [40] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Comput.*, vol. 14, no. 11, pp. 2531–2560, 2002.
- [41] H. Hazan and L. M. Manevitz, "Topological constraints and robustness in liquid state machines," *Expert Syst. Appl.*, vol. 39, no. 2, pp. 1597–1606, Feb. 2012.
- [42] W. Maass, R. Legenstein, and H. Markram, "A New Approach towards Vision Suggested by Biologically Realistic Neural Microcircuit Models," in *Biologically Motivated Computer Vision*, 2002, pp. 282–293.
- [43] H. Jaeger, "The 'echo state' approach to analysing and training recurrent neural networks-with an erratum note'," *Bonn Ger. Ger. Natl. Res. Cent. Inf. Technol. GMD Tech. Rep.*, vol. 148, 2001.
- [44] H. Jaeger, "Adaptive Nonlinear System Identification with Echo State Networks," in *Proceedings of the 15th International Conference on Neural Information Processing Systems*, Cambridge, MA, USA, 2002, pp. 609–616.
- [45] M. Lukoševičius, "A Practical Guide to Applying Echo State Networks," in *Neural Networks: Tricks of the Trade: Second Edition*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 659–686.
- [46] Q. Wang, Y. Li, and P. Li, "Liquid state machine based pattern recognition on FPGA with firing-activity dependent power gating and approximate computing," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2016, pp. 361–364.

- [47] P. Voigtlaender, P. Doetsch, and H. Ney, "Handwriting Recognition with Large Multidimensional Long Short-Term Memory Recurrent Neural Networks," in *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, 2016, pp. 228–233.
- [48] N. Schaetti, M. Salomon, and R. Couturier, "Echo State Networks-Based Reservoir Computing for MNIST Handwritten Digits Recognition," in *2016 IEEE Intl Conference on Computational Science and Engineering (CSE) and IEEE Intl Conference on Embedded and Ubiquitous Computing (EUC) and 15th Intl Symposium on Distributed Computing and Applications for Business Engineering (DCABES)*, 2016, pp. 484–491.
- [49] B. Bakker, V. Zhumatiy, G. Gruener, and J. Schmidhuber, "A robot that reinforcement-learns to identify and memorize important previous observations," in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)*, Las Vegas, Nevada, USA, 2003, vol. 1, pp. 430–435.
- [50] L. Bontemps, V. L. Cao, J. McDermott, and N.-A. Le-Khac, "Collective Anomaly Detection Based on Long Short-Term Memory Recurrent Neural Networks," in *Future Data and Security Engineering*, 2016, pp. 141–152.
- [51] S. Chauhan and L. Vig, "Anomaly detection in ECG time signals via deep long short-term memory networks," in *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2015, pp. 1–7.
- [52] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transp. Res. Part C Emerg. Technol.*, vol. 54, pp. 187–197, May 2015.
- [53] Wikipedia contributors, *Recurrent neural network — Wikipedia, The Free Encyclopedia*. 2019.
- [54] G. Tanaka *et al.*, "Recent advances in physical reservoir computing: A review," *Neural Netw.*, vol. 115, pp. 100–123, Jul. 2019.
- [55] K. Liano, "Robust error measure for supervised neural network learning with outliers," *IEEE Trans. Neural Netw.*, vol. 7, no. 1, pp. 246–250, Jan. 1996.
- [56] R. Rojas, *Neural Networks : A Systematic Introduction*, Springer-Verlag Berlin Heidelberg. 1996.
- [57] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [58] F. J. Pineda, "Generalization of back-propagation to recurrent neural networks," *Phys. Rev. Lett.*, vol. 59, no. 19, pp. 2229–2232, Nov. 1987.
- [59] Y. LeCun, "A theoretical framework for Back-Propagation," in *Proceedings of the 1988 Connectionist Models Summer School*, CMU, Pittsburgh, Pa, 1988, pp. 21–28.
- [60] P. J. Werbos, "Backpropagation through time: what it does and how to do it," *Proc. IEEE*, vol. 78, no. 10, pp. 1550–1560, Oct. 1990.
- [61] A. M. Ahmad, S. Ismail, and D. F. Samaon, "Recurrent neural network with backpropagation through time for speech recognition," in *IEEE International Symposium on Communications and Information Technology, 2004. ISCIT 2004.*, 2004, vol. 1, pp. 98–102 vol.1.
- [62] R. Pascanu, T. Mikolov, and Y. Bengio, "Understanding the exploding gradient problem," *CoRR Abs12115063*, vol. 2, 2012.
- [63] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.
- [64] S. Hochreiter, "The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions," *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, vol. 06, no. 02, pp. 107–116, 1998.
- [65] E. Oja, "Simplified neuron model as a principal component analyzer," *J. Math. Biol.*, vol. 15, no. 3, pp. 267–273, Nov. 1982.
- [66] T. Sanger, *Optimal unsupervised learning in feedforward neural networks*. AI Lab TR 1086. MIT, 1989.
- [67] R. V. Rullen and S. J. Thorpe, "Rate Coding Versus Temporal Order Coding: What the Retinal Ganglion Cells Tell the Visual Cortex," *Neural Comput.*, vol. 13, no. 6, pp. 1255–1283, Jun. 2001.

- [68] A. Gavrilo and K. O Panchenko, "Methods of Learning for Spiking Neural Networks. A Survey," 2016.
- [69] R. Kempter, W. Gerstner, and J. L. Van Hemmen, "Hebbian learning and spiking neurons," *Phys. Rev. E*, vol. 59, no. 4, p. 4498, 1999.
- [70] A. Kepecs, M. van Rossum, S. Song, and J. Tegner, "Spike-timing-dependent plasticity: Common themes and divergent vistas," *Biol. Cybern.*, vol. 87, pp. 446–58, 2003.
- [71] C. D. Schuman *et al.*, "A Survey of Neuromorphic Computing and Neural Networks in Hardware," *ArXiv E-Prints*, May 2017.
- [72] C. S. Thakur *et al.*, "Large-Scale Neuromorphic Spiking Array Processors: A Quest to Mimic the Brain," *Front. Neurosci.*, vol. 12, p. 891, 2018.
- [73] S. C. Liu, T. Delbruck, G. Indiveri, R. Douglas, and A. Whatley, *Event-Based Neuromorphic Systems*. Wiley, 2015.
- [74] J. Seo *et al.*, "A 45nm CMOS neuromorphic chip with a scalable architecture for learning in networks of spiking neurons," in *2011 IEEE Custom Integrated Circuits Conference (CICC)*, 2011, pp. 1–4.
- [75] Y. J. Park *et al.*, "3-D Stacked Synapse Array Based on Charge-Trap Flash Memory for Implementation of Deep Neural Networks," *IEEE Trans. Electron Devices*, vol. 66, no. 1, pp. 420–427, Jan. 2019.
- [76] P. E. Hasler, C. Diorio, B. A. Minch, and C. Mead, "Single transistor learning synapses," in *Advances in neural information processing systems*, 1995, pp. 817–824.
- [77] L. Chua, "Memristor-The missing circuit element," *IEEE Trans. Circuit Theory*, vol. 18, no. 5, pp. 507–519, Sep. 1971.
- [78] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 453, no. 7191, pp. 80–83, May 2008.
- [79] A. V. Kolobov, J. Tominaga, and P. Fons, "Phase-Change Memory Materials," in *Springer Handbook of Electronic and Photonic Materials*, S. Kasap and P. Capper, Eds. Cham: Springer International Publishing, 2017, pp. 1–1.
- [80] T. Yu and G. Cauwenberghs, "Analog VLSI Biophysical Neurons and Synapses With Programmable Membrane Channel Kinetics," *IEEE Trans. Biomed. Circuits Syst.*, vol. 4, no. 3, pp. 139–148, Jun. 2010.
- [81] S. Saighi, J. Tomas, Y. Bornat, and S. Renaud, "A Conductance-Based Silicon Neuron with Dynamically Tunable Model Parameters," in *Conference Proceedings. 2nd International IEEE EMBS Conference on Neural Engineering, 2005.*, 2005, pp. 285–288.
- [82] M. F. Simoni, G. S. Cymbalyuk, M. E. Sorensen, R. L. Calabrese, and S. P. DeWeerth, "A multiconductance silicon neuron with biologically matched dynamics," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 2, pp. 342–354, Feb. 2004.
- [83] M. Azadmehr, J. P. Abrahamsen, and P. Hafliger, "A foveated AER imager chip [address event representation]," in *2005 IEEE International Symposium on Circuits and Systems*, 2005, pp. 2751–2754 Vol. 3.
- [84] J. A. M. Olsson and P. Hafliger, "Mismatch reduction with relative reset in integrate-and-fire photo-pixel array," in *2008 IEEE Biomedical Circuits and Systems Conference*, 2008, pp. 277–280.
- [85] E. Chicca *et al.*, "A VLSI recurrent network of integrate-and-fire neurons connected by plastic synapses with long-term memory," *IEEE Trans. Neural Netw.*, vol. 14, no. 5, pp. 1297–1307, Sep. 2003.
- [86] F. Danneville, C. Loyez, K. Carpentier, I. Sourikopoulos, E. Mercier, and A. Cappy, "A Sub-35 pW Axon-Hillock artificial neuron circuit," *Solid-State Electron.*, vol. 153, pp. 88–92, Mar. 2019.
- [87] G. Indiveri *et al.*, "Neuromorphic Silicon Neuron Circuits," *Front. Neurosci.*, vol. 5, May 2011.
- [88] O. Erdener and S. Ozoguz, "A new neuron model suitable for low power VLSI implementation," in *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, 2015, pp. 15–19.

- [89] R. FitzHugh, "Impulses and Physiological States in Theoretical Models of Nerve Membrane," *Biophys. J.*, vol. 1, no. 6, pp. 445–466, Jul. 1961.
- [90] Rose R. M., Hindmarsh J. L., and Huxley Andrew Fielding, "The assembly of ionic currents in a thalamic neuron I. The three-dimensional model," *Proc. R. Soc. Lond. B Biol. Sci.*, vol. 237, no. 1288, pp. 267–288, Aug. 1989.
- [91] C. Morris and H. Lecar, "Voltage oscillations in the barnacle giant muscle fiber," *Biophys. J.*, vol. 35, no. 1, pp. 193–213, Jul. 1981.
- [92] C. Toumazou, J. Georgiou, and E. M. Drakakis, "Current-mode analogue circuit representation of Hodgkin and Huxley neuron equations," *Electron. Lett.*, vol. 34, no. 14, pp. 1376–1377, Jul. 1998.
- [93] S. A. Aamir, P. Müller, A. Hartel, J. Schemmel, and K. Meier, "A highly tunable 65-nm CMOS LIF neuron for a large scale neuromorphic system," in *2016 46th European Solid-State Device Research Conference (ESSDERC)*, 2016, pp. 63–66.
- [94] J. Shamsi, K. Mohammadi, and S. B. Shokouhi, "A low power circuit of a leaky integrate and fire neuron with global reset," in *2017 Iranian Conference on Electrical Engineering (ICEE)*, 2017, pp. 366–369.
- [95] C. Mayr *et al.*, "A Biological-Realtime Neuromorphic System in 28 nm CMOS Using Low-Leakage Switched Capacitor Circuits," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 1, pp. 243–254, Feb. 2016.
- [96] S. Srivastava and S. S. Rathod, "Silicon neuron-analog CMOS VLSI implementation and analysis at 180nm," in *2016 3rd International Conference on Devices, Circuits and Systems (ICDCS)*, 2016, pp. 28–32.
- [97] V. Kornijcuk *et al.*, "Leaky Integrate-and-Fire Neuron Circuit Based on Floating-Gate Integrator," *Front. Neurosci.*, vol. 10, p. 212, 2016.
- [98] X. Wu, V. Saxena, K. Zhu, and S. Balagopal, "A CMOS Spiking Neuron for Brain-Inspired Neural Networks With Resistive Synapses and In Situ Learning," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 62, no. 11, pp. 1088–1092, Nov. 2015.
- [99] A. Joubert, B. Belhadj, and R. Héliot, "A robust and compact 65 nm LIF analog neuron for computational purposes," in *New Circuits and Systems Conference (NEWCAS), 2011 IEEE 9th International*, 2011, pp. 9–12.
- [100] J. M. Cruz-Albrecht, M. W. Yung, and N. Srinivasa, "Energy-Efficient Neuron, Synapse and STDP Integrated Circuits," *IEEE Trans. Biomed. Circuits Syst.*, vol. 6, no. 3, pp. 246–256, Jun. 2012.
- [101] A. Joubert, B. Belhadj, O. Temam, and R. Héliot, "Hardware spiking neurons design: Analog or digital?," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*, 2012, pp. 1–5.
- [102] G. Indiveri, E. Chicca, and R. Douglas, "A VLSI array of low-power spiking neurons and bistable synapses with spike-timing dependent plasticity," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 211–221, Jan. 2006.
- [103] S. Millner, A. Grübl, K. Meier, J. Schemmel, and M. Schwartz, "A VLSI Implementation of the Adaptive Exponential Integrate-and-Fire Neuron Model," 2010.
- [104] P. Livi and G. Indiveri, "A current-mode conductance-based silicon neuron for address-event neuromorphic systems," in *2009 IEEE International Symposium on Circuits and Systems*, 2009, pp. 2898–2901.
- [105] A. Makhlooghpoor, H. Soleimani, A. Ahmadi, M. Zwolinski, and M. Saif, "High accuracy implementation of Adaptive Exponential integrated and fire neuron model," in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 192–197.
- [106] S. A. Aamir, "A Mixed-Signal Structured AdEx Neuron for Accelerated Neuromorphic Cores," 2018.
- [107] J. H. B. Wijekoon and P. Dudek, "A CMOS circuit implementation of a spiking neuron with bursting and adaptation on a biological timescale," in *2009 IEEE Biomedical Circuits and Systems Conference*, 2009, pp. 193–196.

- [108] A. van Schaik, C. Jin, A. McEwan, and T. J. Hamilton, "A log-domain implementation of the Izhikevich neuron model," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, 2010, pp. 4253–4256.
- [109] V. Rangan, A. Ghosh, V. Aparin, and G. Cauwenberghs, "A subthreshold aVLSI implementation of the Izhikevich simple neuron model," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2010, pp. 4164–4167.
- [110] A. S. Demirkol and S. Ozoguz, "A low power VLSI implementation of the Izhikevich neuron model," in *New Circuits and Systems Conference (NEWCAS), 2011 IEEE 9th International*, 2011, pp. 169–172.
- [111] I. Sourikopoulos *et al.*, "A 4-fJ/Spike Artificial Neuron in 65 nm CMOS Technology," *Front. Neurosci.*, vol. 11, p. 123, 2017.
- [112] A. Basu and P. E. Hasler, "Nullcline-Based Design of a Silicon Neuron," *IEEE Trans. Circuits Syst. Regul. Pap.*, vol. 57, no. 11, pp. 2938–2947, Nov. 2010.
- [113] J. Touboul, "Bifurcation Analysis of a General Class of Nonlinear Integrate-and-Fire Neurons," *SIAM J. Appl. Math.*, vol. 68, no. 4, pp. 1045–1079, Jan. 2008.
- [114] F. Grassia, T. Levi, T. Kohno, and S. Saighi, "Silicon neuron: digital hardware implementation of the quartic model," *Artif. Life Robot.*, vol. 19, no. 3, pp. 215–219, Nov. 2014.
- [115] J. H. B. Wijekoon and P. Dudek, "Spiking and Bursting Firing Patterns of a Compact VLSI Cortical Neuron Circuit," in *2007 International Joint Conference on Neural Networks*, 2007, pp. 1332–1337.
- [116] Y. J. Lee *et al.*, "Low power real time electronic neuron VLSI design using subthreshold technique," in *2004 IEEE International Symposium on Circuits and Systems (IEEE Cat. No. 04CH37512)*, 2004, vol. 4, pp. IV–744.
- [117] E. Chicca, F. Stefanini, C. Bartolozzi, and G. Indiveri, "Neuromorphic Electronic Circuits for Building Autonomous Cognitive Systems," *Proc. IEEE*, vol. 102, no. 9, pp. 1367–1388, Sep. 2014.
- [118] D. D. Coon and A. G. U. Perera, "Neuron-like transient phenomena in silicon p-i-n structures," *Solid-State Electron.*, vol. 31, no. 5, pp. 851–862, May 1988.
- [119] D. D. COON and A. G. U. PERERA, "Semiconductor electronic concepts for neural network emulation," *Int. J. Electron.*, vol. 63, no. 1, pp. 61–69, Jul. 1987.
- [120] V. Ostwal, R. Meshram, B. Rajendran, and U. Ganguly, "An ultra-compact and low power neuron based on SOI platform," in *2015 International Symposium on VLSI Technology, Systems and Applications*, 2015, pp. 1–2.
- [121] S. Dutta, V. Kumar, A. Shukla, N. R. Mohapatra, and U. Ganguly, "Leaky Integrate and Fire Neuron by Charge-Discharge Dynamics in Floating-Body MOSFET," *Sci. Rep.*, vol. 7, no. 1, p. 8257, Aug. 2017.
- [122] M.-W. Kwon, H. Kim, J. Park, and B.-G. Park, "Integrate-and-Fire Neuron Circuit and Synaptic Device using Floating Body MOSFET with Spike Timing- Dependent Plasticity," *J. Semicond. Technol. Sci.*, vol. 15, no. 6, pp. 658–663, Dec. 2015.
- [123] B. Das, J. Schulze, and U. Ganguly, "Ultra-Low Energy LIF Neuron Using Si NIPIN Diode for Spiking Neural Networks," *IEEE Electron Device Lett.*, vol. 39, no. 12, pp. 1832–1835, Dec. 2018.
- [124] T. Tuma, A. Pantazi, M. Le Gallo, A. Sebastian, and E. Eleftheriou, "Stochastic phase-change neurons," *Nat Nano*, vol. 11, no. 8, pp. 693–699, Aug. 2016.
- [125] N. Locatelli, V. Cros, and J. Grollier, "Spin-torque building blocks," *Nat. Mater.*, vol. 13, no. 1, pp. 11–20, Jan. 2014.
- [126] M. Sharad, C. Augustine, G. Panagopoulos, and K. Roy, "Spin-Based Neuron Model With Domain-Wall Magnets as Synapse," *IEEE Trans. Nanotechnol.*, vol. 11, no. 4, pp. 843–853, Jul. 2012.
- [127] A. Sengupta, S. H. Choday, Y. Kim, and K. Roy, "Spin orbit torque based electronic neuron," *Appl. Phys. Lett.*, vol. 106, no. 14, p. 143701, Apr. 2015.
- [128] A. Jaiswal, S. Roy, G. Srinivasan, and K. Roy, "Proposal for a Leaky-Integrate-Fire Spiking Neuron Based on Magnetoelectric Switching of Ferromagnets," *IEEE Trans. Electron Devices*, vol. 64, no. 4, pp. 1818–1824, Apr. 2017.

- [129] D. Fan, Y. Shim, A. Raghunathan, and K. Roy, "STT-SNN: A Spin-Transfer-Torque Based Soft-Limiting Non-Linear Neuron for Low-Power Artificial Neural Networks," *IEEE Trans. Nanotechnol.*, vol. 14, no. 6, pp. 1013–1023, Nov. 2015.
- [130] N. Qiao and G. Indiveri, "Analog circuits for mixed-signal neuromorphic computing architectures in 28 nm FD-SOI technology," in *2017 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, 2017, pp. 1–4.
- [131] N. Qiao and G. Indiveri, "Scaling mixed-signal neuromorphic processors to 28 nm FD-SOI technologies," in *2016 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2016, pp. 552–555.
- [132] J. H. B. Wijekoon and P. Dudek, "Compact silicon neuron circuit with spiking and bursting behaviour," *Neural Netw.*, vol. 21, no. 2–3, pp. 524–534, Mar. 2008.
- [133] P. Galy, P. Dehan, J. Jimenez, and B. Heitz, "Preliminary results on low power sigmoid neuron transistor response in 28 nm high-k metal gate Fully Depleted SOI technology," *Solid-State Electron.*, vol. 89, pp. 17–21, Nov. 2013.
- [134] J. Bourgeat, C. Entringer, P. Galy, M. Bafleur, and D. Marin-Cudraz, "Evaluation of the ESD performance of local protections based on SCR or bi-SCR with dynamic or static trigger circuit in 32nm," *Microelectron. Reliab.*, vol. 50, no. 9, pp. 1379–1382, Sep. 2010.
- [135] P. Galy *et al.*, "BIMOS transistor and its applications in ESD protection in advanced CMOS technology," in *2012 IEEE International Conference on IC Design Technology*, 2012, pp. 1–4.
- [136] R. Ranjan, M.-W. Kwon, J. Park, H. Kim, and B.-G. Park, "Neuron Circuit Using a Thyristor and Inter-neuron Connection with Synaptic Devices," *J. Semicond. Technol. Sci.*, vol. 15, no. 3, pp. 365–373, Jun. 2015.

Chapitre 2 : Notions de physique du semiconducteur, de physique des composants et technologie 28 nm UTBB FD-SOI

Ce chapitre expose les notions de physique du semiconducteur et de physique des composants nécessaires à la compréhension des mécanismes impliqués dans le fonctionnement du circuit de neurone analogique développé dans cette thèse. Il présente également la technologie 28nm UTBB FD-SOI utilisée pour la fabrication des échantillons et des circuits de test.

2.1 Notion de physique du semiconducteur

Cette section introduit brièvement les notions de physique du semiconducteur essentielles pour la description du comportement des composants microélectroniques.

2.1.1 Théorie des bandes d'énergie

Le comportement électrique des matériaux semiconducteurs est décrit par la théorie des bandes d'énergie, qui dérive de l'application des principes de la mécanique quantique à la description des matériaux solides cristallins [1], [2]. Cette théorie indique que chaque électron du matériau occupe un des états électroniques disponibles dans le cristal. Chaque état électronique est décrit par un unique couple constitué d'un vecteur d'onde $\vec{k}$ et d'un niveau d'énergie E (la dégénérescence de spin n'est ici pas prise en compte). Deux états peuvent être éloignés ou proches selon que leur différence de $\vec{k}$ et de E est grande ou petite. En raison du très grand nombre d'états disponibles, ces états forment des ensembles continus que l'on appelle « bande » : deux états d'une même bande sont reliés entre eux par une succession d'états extrêmement proches les uns des autres et qui forment ainsi un continuum. Cependant, chaque bande admet un état d'énergie maximale et un état d'énergie minimale. Cet intervalle est caractéristique de la bande. Il est possible de repérer ces bandes sur un axe d'énergie en y plaçant les intervalles qui leurs sont associées. Elles sont ainsi classées de la bande d'énergie la plus basse à la bande d'énergie la plus haute. Deux bandes successives peuvent s'interpénétrer (i.e. il existe des états de même énergie dans chaque bande) ou être séparées par une bande d'énergie interdite appelée « gap », dont la largeur est l'énergie du gap E_G . Le matériau ne contient aucun état électronique dont le niveau d'énergie se situe dans le gap. Ces gaps sont caractéristiques de chaque matériau.

En l'absence d'énergie thermique ($T = 0$ K), les électrons occupent les états de plus basses énergies. Le niveau d'énergie du dernier état occupé est appelé niveau de Fermi E_F . Lorsque E_F se trouve au milieu d'une bande, le matériau est conducteur et la bande où se trouve le niveau de Fermi est appelée bande de conduction. Lorsque le dernier état occupé est l'état d'énergie maximale d'une bande et que celle-ci se situe juste avant un « gap », le matériau est semiconducteur, ou isolant selon que E_G est faible ou grand. La dernière bande totalement remplie s'appelle la bande de valence, et la première bande totalement vide s'appelle la bande de conduction. Ces deux bandes se trouvent de part et d'autre du gap. Dans ce cas particulier, le niveau de Fermi se trouve approximativement au milieu du gap. Plus précisément, on montre que le niveau de Fermi vaut :

$$E_F = \frac{E_C + E_V}{2} + \frac{k \cdot T}{2} \cdot \ln\left(\frac{N_V}{N_C}\right) \quad (2.1)$$

[J]

où N_C et N_V sont les densités d'états du minimum de la bande de conduction et du maximum de la bande de valence respectivement, E_C et E_V leur niveau d'énergie, k la constante de Boltzmann, et T la température.

Il est possible de faire transiter un électron depuis la bande de valence vers la bande de conduction en lui fournissant une énergie au moins égale à E_G . Cette énergie peut être apportée par différents mécanismes. En particulier, la présence d'énergie thermique permet la transition d'un certain nombre d'électrons, nombre d'autant plus grand que la température est élevée. La relation entre la densité d'électrons présents dans la bande de conduction, notée n , et la température est donnée par la statistique de Fermi-Dirac :

$$n = \frac{N_C}{1 + \exp\left(\frac{E_C - E_F}{k \cdot T}\right)} \quad [\text{m}^{-3}] \quad (2.2)$$

Les transitions laissent des états électroniques inoccupés dans la bande de valence appelés « trous ». La densité de trous présents dans la bande de valence, notée p , est aussi donnée par la statistique de Fermi-Dirac :

$$p = \frac{N_V}{1 + \exp\left(\frac{E_F - E_V}{k \cdot T}\right)} \quad [\text{m}^{-3}] \quad (2.3)$$

Dans les cas où $E_C - E_F \gg kT$ et $E_F - E_V \gg kT$, c'est-à-dire lorsque le niveau de Fermi est suffisamment éloigné des bandes de valence et de conduction, n et p peuvent être approximées par la statistique de Maxwell-Boltzmann :

$$n = N_C \cdot \exp\left(-\frac{E_C - E_F}{k \cdot T}\right) \quad \text{et} \quad p = N_V \cdot \exp\left(-\frac{E_F - E_V}{k \cdot T}\right) \quad [\text{m}^{-3}] \quad (2.4)$$

En l'absence de dopants dans le semiconducteur (i.e., semiconducteur intrinsèque) et à l'équilibre thermodynamique, les concentrations n et p sont toutes deux égales à la densité de porteur intrinsèque n_i .

$$n = p = n_i \quad [\text{m}^{-3}] \quad (2.5)$$

Dans ce cas particulier, le niveau d'énergie du niveau de Fermi est noté E_i et vaut :

$$E_F = E_i = E_V + \frac{E_G}{2} + \frac{k \cdot T}{2} \cdot \ln\left(\frac{N_V}{N_C}\right) \approx E_V + \frac{E_G}{2} \quad [\text{J}] \quad (2.6)$$

Ce qui correspond à une densité de porteur intrinsèque :

$$n_i = \sqrt{N_V \cdot N_C} \cdot \exp\left(-\frac{E_G}{2 \cdot k \cdot T}\right) \quad [\text{m}^{-3}] \quad (2.7)$$

Il est possible de réécrire les densités n et p en fonction de n_i , de E_i et de E_F :

$$n = n_i \cdot \exp\left(\frac{E_F - E_i}{k \cdot T}\right) \quad \text{et} \quad p = n_i \cdot \exp\left(\frac{E_i - E_F}{k \cdot T}\right) \quad [\text{m}^{-3}] \quad (2.8)$$

On remarque que quelle que soit la hauteur du niveau de Fermi, le produit $n \cdot p$ est constant et vaut n_i^2 (loi d'action de masse). Il s'agit d'une constante associée au matériau semiconducteur :

$$n \cdot p = n_i^2 = N_V \cdot N_C \cdot \exp\left(-\frac{E_G}{k \cdot T}\right) \quad [\text{m}^{-6}] \quad (2.9)$$

En dehors de l'état d'équilibre thermodynamique, (2.9) n'est plus vrai, et le produit $n \cdot p$ est alors relié au potentiel électrique local V par :

$$n \cdot p = n_i^2 \cdot \exp\left(\frac{q \cdot V}{k \cdot T}\right) \quad [m^{-6}] \quad (2.10)$$

Hors équilibre, il est également possible de définir 2 quasi-niveaux de Fermi E_{Fn} et E_{Fp} différents tels que les relations suivantes soient respectées :

$$n = n_i \cdot \exp\left(\frac{E_{Fn} - E_i}{k \cdot T}\right) \quad \text{et} \quad p = n_i \cdot \exp\left(\frac{E_i - E_{Fp}}{k \cdot T}\right) \quad [m^{-3}] \quad (2.11)$$

d'où l'on obtient :

$$V = \frac{1}{q} \cdot (E_{Fn} - E_{Fp}) \quad [V] \quad (2.12)$$

A l'équilibre, il est possible de casser la relation $n = p$ en substituant des atomes du cristal semiconducteur par des atomes d'un autre élément chimique, appelé dopant. On dénombre deux types de dopants selon qu'ils fournissent des électrons à la bande de conduction (dopant de type donneur) ou qu'il capte des électrons de la bande valence et crée des trous (dopant de type accepteur). Les nouveaux états électroniques qu'apportent les dopants donneurs se trouvent dans le gap, proche de la bande de conduction. A $T = 0$ K, ils sont occupés par un électron. Les nouveaux états électroniques qu'apportent les dopants accepteurs sont dans le gap, proches de la bande valence. A $T = 0$ K, ils sont inoccupés. Pour que les dopants puissent fournir les électrons ou les trous, ils doivent être ionisés. Cela advient lorsque la température est suffisamment grande pour permettre aux électrons de la bande de valence de transiter vers les états libres des dopants accepteurs, ou pour permettre aux électrons des états des dopants donneurs de transiter vers la bande de conduction.

En présence de dopants donneurs de densité N_D et entièrement ionisés, les densités d'électrons et de trous deviennent :

$$n = N_D \quad \text{et} \quad p = \frac{n_i^2}{N_D} \quad [m^{-3}] \quad (2.13)$$

Comme généralement $N_D \gg n_i$, on a $n \gg p$. Les électrons sont alors appelés porteurs majoritaires et les trous porteurs minoritaires.

En présence de dopants accepteurs de densité N_A et entièrement ionisés, les densités d'électrons et de trous deviennent :

$$n = \frac{n_i^2}{N_A} \quad \text{et} \quad p = N_A \quad [m^{-3}] \quad (2.14)$$

Comme généralement $N_A \gg n_i$, on a $p \gg n$. Les trous sont alors appelés porteurs majoritaires et les électrons porteurs minoritaires.

2.1.2 Transport électronique et mécanismes de génération-recombinaison

Seuls les porteurs de charge électrique libres (électrons présents dans la bande de conduction et trous présents dans la bande de valence) participent au transport du courant électrique, de densité $\vec{J}$, par leur déplacement. Ce déplacement, quantifié sous forme de densité de flux $\vec{\Phi}$, peut advenir sous l'effet d'un champ électrique $\vec{E}$ (densité de flux de drift $\vec{\Phi}_{\text{DRIFT}}$) ou sous l'effet d'un gradient de

concentration (densité de flux de diffusion $\vec{\Phi}_{DIFF}$). Les expressions des densités de flux d'électrons et de trous sont :

$$\vec{\Phi}_n = \vec{\Phi}_{DRIFT,n} + \vec{\Phi}_{DIFF,n} = -n \cdot \mu_n \cdot \vec{E} - D_n \cdot \vec{\nabla} n \quad [m^{-2} \cdot s^{-1}] \quad (2.15)$$

$$\vec{\Phi}_p = \vec{\Phi}_{DRIFT,p} + \vec{\Phi}_{DIFF,p} = p \cdot \mu_p \cdot \vec{E} - D_p \cdot \vec{\nabla} p \quad [m^{-2} \cdot s^{-1}] \quad (2.16)$$

où μ_n et μ_p sont les mobilités des électrons et des trous et D_n et D_p leurs coefficients de diffusion. Les densités de courant associées s'obtiennent en multipliant ces densités de flux par les charges $-q$ et $+q$ des électrons et des trous :

$$\vec{J}_n = -q \cdot \vec{\Phi}_n = q \cdot n \cdot \mu_n \cdot \vec{E} + q \cdot D_n \cdot \vec{\nabla} n = \vec{J}_{DRIFT,n} + \vec{J}_{DIFF,n} \quad [A \cdot m^{-2}] \quad (2.17)$$

$$\vec{J}_p = q \cdot \vec{\Phi}_p = q \cdot p \cdot \mu_p \cdot \vec{E} - q \cdot D_p \cdot \vec{\nabla} p = \vec{J}_{DRIFT,p} + \vec{J}_{DIFF,p} \quad [A \cdot m^{-2}] \quad (2.18)$$

Le courant total s'obtient en ajoutant ces deux contributions :

$$\vec{J} = \vec{J}_n + \vec{J}_p \quad (2.19)$$

Il est aussi possible de décomposer le courant total en une contribution de drift et une contribution de diffusion :

$$\begin{cases} \vec{J} = \vec{J}_{DIFF} + \vec{J}_{DRIFT} & [A \cdot m^{-2}] \quad (2.20) \\ \vec{J}_{DIFF} = \vec{J}_{DIFF,n} + \vec{J}_{DIFF,p} & [A \cdot m^{-2}] \quad (2.21) \\ \vec{J}_{DRIFT} = \vec{J}_{DRIFT,n} + \vec{J}_{DRIFT,p} & [A \cdot m^{-2}] \quad (2.22) \end{cases}$$

L'équation de continuité, conséquence de la loi de conservation de la charge électrique, permet de relier les divergences des densités de courant $\vec{J}_n$ et $\vec{J}_p$ au taux de génération G_n et G_p et de recombinaison R_n et R_p des électrons et des trous :

$$\frac{1}{q} \cdot \vec{\nabla} \cdot \vec{J}_n + G_n - R_n = \frac{\partial n}{\partial t} \quad [m^{-3} \cdot s^{-1}] \quad (2.23)$$

$$-\frac{1}{q} \cdot \vec{\nabla} \cdot \vec{J}_p + G_p - R_p = \frac{\partial p}{\partial t} \quad [m^{-3} \cdot s^{-1}] \quad (2.24)$$

A l'état stationnaire, $\partial n / \partial t$ et $\partial p / \partial t$ sont nuls, et dans la majorité des cas, $G_n = G_p = G$ et $R_n = R_p = R$, ce qui correspond à des transitions locales d'électrons.

L'origine des phénomènes de génération-recombinaison qui gouverne G et R est diverse. Notamment, la présence d'états électroniques dans le gap induits par la présence d'impureté ou de défaut cristallin (appelés piège) déclenche un mécanisme de génération-recombinaison décrit par la statistique SRH [3]. Cette statistique relie le taux de génération/ recombinaison R_{SRH} aux densités de porteur n et p , à la densité de piège N_t et au niveau d'énergie de ces pièges E_t :

$$R_{SRH} = \frac{\sigma_n \cdot \sigma_p \cdot v_{th} \cdot N_t \cdot (n \cdot p - n_i^2)}{\sigma_n \cdot \left[n + n_i \cdot \exp\left(\frac{E_t - E_i}{k \cdot T}\right) \right] + \sigma_p \cdot \left[p + n_i \cdot \exp\left(\frac{E_i - E_t}{k \cdot T}\right) \right]} \quad [m^{-3} \cdot s^{-1}] \quad (2.25)$$

où σ_n et σ_p sont les sections efficaces de capture des électrons et des trous et v_{th} la vitesse thermique. Selon que $n \cdot p > n_i^2$ ou $n \cdot p < n_i^2$, la présence de piège conduit à un mécanisme de recombinaison ($R_{SRH} >$

0) ou de génération ($R_{SRH} < 0$). De plus, l'équation (2.25) montre que $|R_{SRH}|$ est maximale lorsque $E_t = E_i$, c'est-à-dire lorsque le niveau d'énergie des pièges se situe au centre de gap.

Une autre famille de phénomènes de génération-recombinaison implique le transfert d'énergie cinétique de porteurs. Ainsi, un électron peut transiter de la bande de valence vers la bande de conduction en ponctionnant une énergie E_G à un porteur ayant une énergie cinétique suffisante, à l'occasion d'un choc. Ce mécanisme s'appelle l'ionisation par impact (« impact ionization » : II). Le mécanisme inverse est appelé recombinaison Auger : un électron transite de la bande de conduction à la bande de valence en fournissant une quantité d'énergie E_G à l'énergie cinétique d'autres porteurs.

Une autre famille de phénomènes de génération recombinaison implique la présence d'un champ électrique suffisamment fort pour que les bandes de valence et de conduction en deux points de l'espace suffisamment proches se retrouvent à la même énergie. La transition de la bande de valence au premier point vers la bande de conduction au second point s'effectue par effet tunnel bande à bande (« band-to-band tunneling », BTBT). La présence de piège dans le gap peut assister ce mécanisme (« trap-assisted tunneling » : TAT).

2.2 Notions de physique du composant micro-électronique

Cette section décrit succinctement les comportements électriques et les mécanismes physiques sous-jacents de 4 composants fondamentaux en micro-électronique : la jonction PN, le transistor bipolaire (« bipolar junction transistor » : BJT), la capacité métal – oxyde – semiconducteur (« metal – oxide – semiconductor capacitor » : MOSCAP), et le transistor métal – oxyde – semiconducteur à effet de champ (« metal – oxide – semiconductor field effect transistor » : MOSFET). Pour chaque composant, seuls les concepts pertinents pour le déroulement des discussions dans les chapitres suivants sont abordés. Pour de plus amples détails, le lecteur pourra se tourner vers les nombreux livres sur la physique des semiconducteurs, tel que [4] ou [5], sur lesquels cette section s'appuie.

2.2.1 La jonction PN

Une jonction PN s'obtient par l'aboutement de deux zones d'un même semiconducteur (homo-jonction) dopées de type P et N. La zone dopée P est l'anode et la zone dopée N est la cathode. A l'équilibre thermodynamique, cet aboutement conduit à l'apparition d'une zone de charge d'espace (ZCE), ou zone de déplétion, de part et d'autre de l'interface entre les deux zones dopées différemment. Dans la ZCE, les densités de porteur sont négligeables face aux densités de dopants ($n, p \ll N_D, N_A$). Cette inégalité fait apparaître une charge électrique locale. En faisant l'hypothèse d'une ZCE entièrement déplétée, cette charge est égale à $q \cdot N_D$ dans la zone N et $-q \cdot N_A$ dans la zone P. La présence de ces charges induit un champ électrique interne auquel est associé une barrière de potentiel électrique interne V_{bi} (*bi* pour « built-in ») (Figure 2-1). Sa valeur est établie en considérant l'alignement des niveaux de Fermi de part et d'autre de la ZCE :

$$V_{bi} = U_T \cdot \ln \left(\frac{N_A \cdot N_D}{n_i^2} \right) \quad [V] \quad (2.26)$$

où $U_T = k \cdot T / q$ est le potentiel thermique. Dans le cas de densités de dopants constantes dans la ZCE et d'une jonction abrupte, la largeur de la ZCE W_{ZCE} vaut :

$$W_{ZCE} = \sqrt{\frac{2 \cdot \epsilon_S}{q} \cdot \frac{(N_A + N_D) \cdot V_{bi}}{N_A \cdot N_D}} \quad [m] \quad (2.27)$$

La largeur W_{ZCE} est modulée par l'application d'une tension électrique externe V_{A-K} aux bornes de la jonction PN :

$$W_{ZCE}(V_{A-K}) = \sqrt{\frac{2 \cdot \epsilon_S}{q} \cdot \frac{(N_A + N_D) \cdot (V_{bi} - V_{A-K})}{N_A \cdot N_D}} \quad [m] \quad (2.28)$$

Une tension V_{A-K} négative a pour effet d'augmenter W_{ZCE} tandis qu'une tension V_{A-K} positive la réduit jusqu'à devenir nulle lorsque $V_{A-K} = V_{bi}$.

L'application de V_{A-K} porte la jonction PN en dehors de l'équilibre thermodynamique : un courant I_{PN} apparaît dans la jonction. Dans le cas d'une diode idéale, ce courant est un courant de diffusion I_{DIFF} dont la valeur est donnée par :

$$I_{PN} = I_{DIFF} = I_{SAT} \left[\exp\left(\frac{V_{A-K}}{U_T}\right) - 1 \right] \quad [A] \quad (2.29)$$

avec :

$$I_{SAT} = S_{PN} \cdot q \cdot n_i^2 \cdot \left(\frac{1}{N_A} \cdot \sqrt{\frac{D_n}{\tau_n}} + \frac{1}{N_D} \cdot \sqrt{\frac{D_p}{\tau_p}} \right) \quad [A] \quad (2.30)$$

où τ_n et τ_p sont les durées de vie de électrons et des trous dans les zones P et N respectivement, et S_{PN} la section de la jonction. Le courant est indépendant de V_{A-K} pour $V_{A-K} \ll -U_T$ et croit comme $\exp(V_{A-K}/U_T)$ pour $V_{A-K} \gg U_T$. Cette croissance exponentielle n'est cependant pas infinie. En particulier, lorsque $V_{A-K} = V_{bi}$, la ZCE disparaît, et les hypothèses qui ont conduit à établir (2.29) ne sont plus vérifiées. Lorsque V_{A-K} se rapproche de V_{bi} la résistance de la ZCE devient extrêmement faible et ce sont les résistances R_s des zones neutres de part et d'autre de la ZCE qui limitent le courant. La diode est en régime de résistance série. Dans ce contexte, (2.29) devient :

$$I_{PN} = I_{DIFF} = I_{SAT} \left[\exp\left(\frac{V_{A-K} - R_s \cdot I_{PN}}{U_T}\right) - 1 \right] \quad [A] \quad (2.31)$$

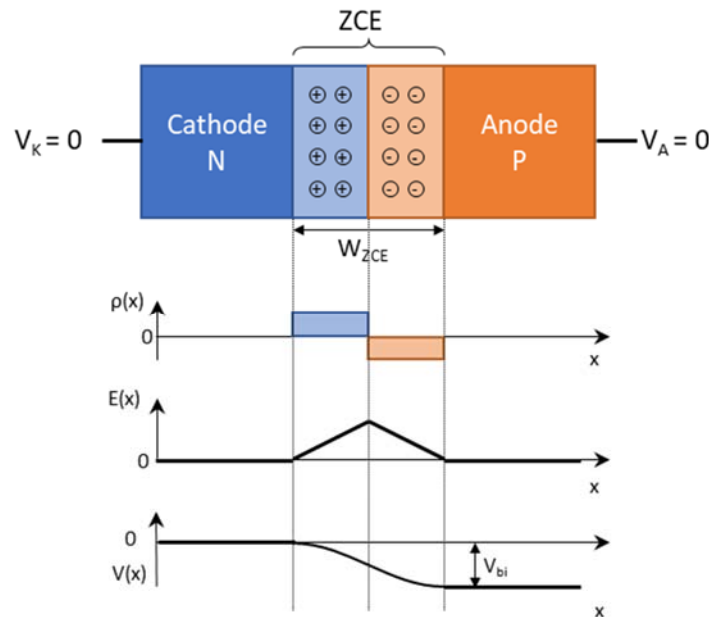


Figure 2-1 – Schéma d'une jonction PN à l'équilibre thermodynamique ($V_{A-K} = 0$). Profil de densité de charge, de champ électrique et de potentiel électrique correspondant.

Dans une diode non-idéale, des courants de génération/recombinaison I_{GR} s'ajoutent au courant de diffusion I_{DIFF} de la diode idéale :

$$I_{PN} = I_{DIFF} + I_{GR} \quad [A] \quad (2.32)$$

En polarisation directe, la contribution dominante à I_{GR} est le mécanisme décrit par la statistique SRH (équation (2.25)) [6]. Comme $n \cdot p > n_i^2$, le mécanisme SRH est un phénomène de recombinaison. En faisant l'hypothèse que les sections efficaces de capture des électrons et des trous sont égales, que l'énergie des centres de recombinaison se trouvent au milieu du gap, et que la recombinaison est maximale dans toute l'épaisseur de la ZCE, on montre à partir de (2.25) que [4], [7]:

$$I_{GR} = I_{SRH} = I_{0,SRH} \sqrt{(V_{bi} - V_{A-K})} \cdot \left[\exp\left(\frac{V_{A-K}}{2 \cdot U_T}\right) - 1 \right] \quad [A] \quad (2.33)$$

avec

$$I_{0,SRH} = S_{PN} \cdot \frac{q \cdot n_i}{2 \cdot \tau_0} \cdot \sqrt{\frac{2 \cdot \epsilon_{SC}}{q} \cdot \frac{N_A + N_D}{N_A \cdot N_D}} \quad [A] \quad (2.34)$$

où τ_0 est la durée de vie des porteurs n et p . I_{GR} varie comme $\exp[V_{A-K}/(2 \cdot U_T)]$ tandis que I_{DIFF} varie comme $\exp[V_{A-K}/U_T]$. Comme $I_{0,SRH}$ est généralement plus grand que I_{SAT} , cela conduit à un courant de diode polarisée en direct dominé par I_{SRH} à faible V_{A-K} et par I_{DIFF} à fort V_{A-K} . Un façon simplifiée de réécrire (2.32) consiste à introduire dans (2.31) un coefficient d'idéalité η compris entre 1 et 2, et d'écrire :

$$I_{PN} = I_{DIFF} + I_{SRH} = I_{0,PN} \cdot \left[\exp\left(\frac{V_{A-K} - R_S \cdot I_{PN}}{\eta \cdot U_T}\right) - 1 \right] \quad [A] \quad (2.35)$$

où $I_{0,PN}$ est un paramètre de l'équation qui correspond au courant en polarisation inverse.

En polarisation inverse, le mécanisme SRH est également actif selon l'équation (2.33) et sa contribution est généralement supérieure à I_{DIFF} . Notons que comme $n \cdot p < n_i^2$, le mécanisme SRH est un phénomène de génération. Deux autres mécanismes de génération peuvent aussi être mis en jeu : le BTBT et le TAT (Figure 2-2). Pour rappel, le BTBT intervient lorsque, sous l'effet d'un champ électrique intense, la courbure de bande est suffisamment importante pour autoriser les transitions d'électrons de la bande de valence d'un point du semiconducteur vers la bande de conduction d'un autre point du semiconducteur. Le TAT est un mécanisme similaire mais où les états présents dans le gap, i.e. les pièges, servent de relais.

On montre que, sous les hypothèses développées dans [8], le calcul du courant de BTBT prend la forme

$$I_{BTBT} = A_{BTBT} \cdot V_{A-K} \cdot (V_{bi} - V_{A-K}) \cdot \exp\left(-\frac{B_{BTBT}}{\sqrt{V_{bi} - V_{A-K}}}\right) \quad [A] \quad (2.36)$$

où A_{BTBT} et B_{BTBT} sont des constantes du modèle.

Comme le TAT est un mécanisme qui fait intervenir les pièges déjà évoqués dans le mécanisme SRH, ces deux phénomènes sont liés et on montre que le TAT a aussi une dépendance en V_{A-K} de la forme [8] :

$$I_{TAT} \propto \exp\left(\frac{V_{A-K}}{2 \cdot U_T}\right) \quad [A] \quad (2.37)$$

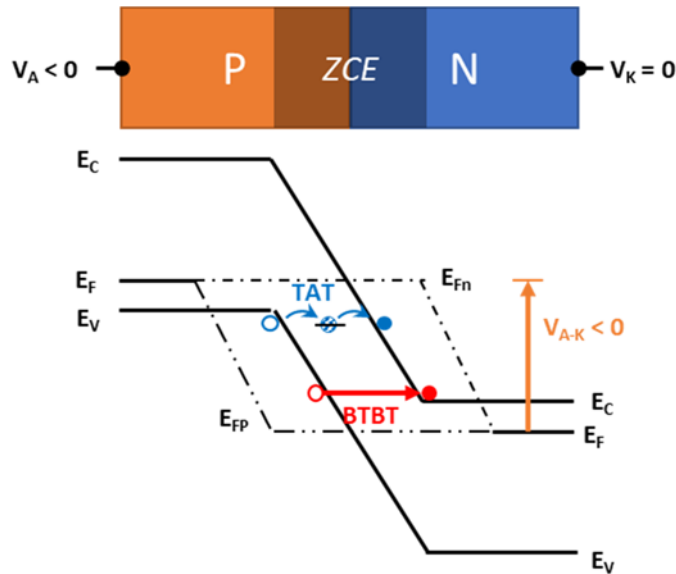


Figure 2-2 – Illustration des mécanismes de BTBT et TAT dans une jonction PN polarisée en inverse

Le courant total de la diode non-idéale est la somme de toutes ces contributions, dont certaines sont dominantes selon le régime de polarisation et les caractéristiques physiques de la diode :

$$I_{PN} = I_{DIFF} + I_{SRH} + I_{BTBT} + I_{TAT} \quad (2.38)$$

A titre d'exemple, une comparaison des courbes de courant de diode idéale et de diode incluant le phénomène de génération/recombinaison SRH est présentée à la Figure 2-3. Le courant de claquage de la diode (« junction breakdown ») observé à forte polarisation inverse n'est pas abordé dans ce document.

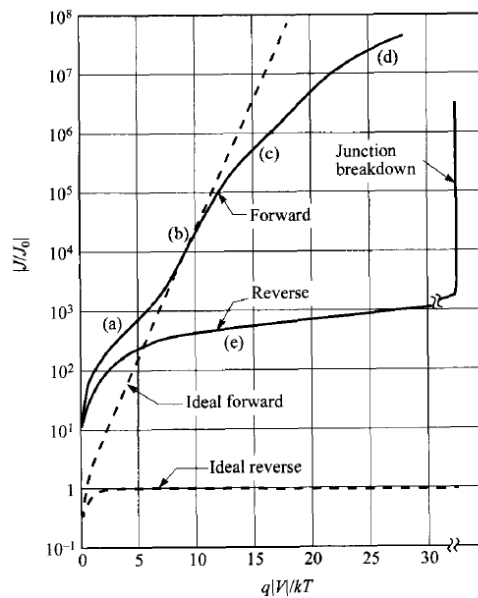


Figure 2-3 – Comparaison des courants typiques de diode idéale et de diode incluant le phénomène de génération SRH, depuis [5].

2.2.2 Le transistor bipolaire (« bipolar junction transistor » : BJT)

Le BJT est une structure obtenue par l'aboutement de 2 jonctions PN. Il peut-être de type NPN ou PNP. Les 2 régions extérieures et de même dopage sont appelées émetteur et collecteur. La région centrale, de dopage opposé, est appelée base (Figure 2-4). Sous certaines conditions de dimensionnement, de dopage et de polarisation, il est possible de contrôler le courant de collecteur I_C avec le courant de base I_B , avec un facteur d'amplification A tel que $I_C = A \cdot I_B$. C'est l'effet transistor.

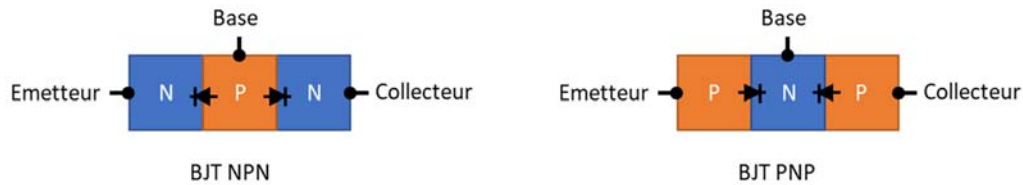


Figure 2-4 - Vue schématique de la topologie d'un BJT NPN et d'un BJT PNP

Pour fonctionner en régime actif, le BJT doit avoir sa jonction base-émetteur polarisée en direct et sa jonction base-collecteur polarisée en inverse. De plus, ces 2 jonctions doivent être suffisamment proches l'une de l'autre et répondre à certains critères de dimensionnement et de dopage pour qu'un couplage adéquat apparaisse entre elles :

- La longueur de la base doit être inférieure à la longueur de diffusion de ses porteurs minoritaires (n pour la base P du BJT NPN, et p pour la base n du BJT PNP).
- La densité de dopants de la base N_B doit être très supérieure à la densité de dopants du collecteur N_C : $N_B \gg N_C$. De cette façon, la ZCE de la jonction base-collecteur s'étend majoritairement du côté du collecteur, notamment lors de la polarisation inverse de la jonction.
- La densité de dopants de l'émetteur N_E doit être très supérieure à la densité de dopants de la base : $N_E \gg N_B$. De cette manière, le courant de la diode base-émetteur est majoritairement composé des porteurs majoritaires de l'émetteur (n pour l'émetteur N du BJT NPN, et p pour l'émetteur P du BJT PNP).
- Enfin, les densités de dopant, la longueur de la base et les tensions appliquées doivent être mises en cohérence de sorte que les deux ZCE ne se rejoignent jamais.

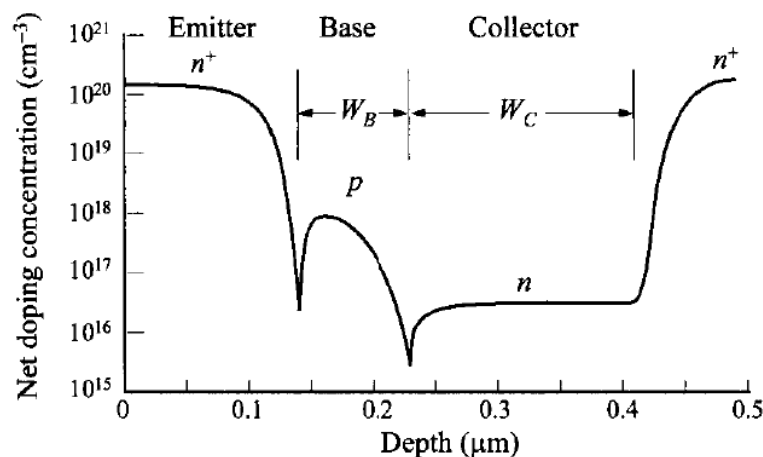


Figure 2-5 – Exemple typique du profil de dopants d'un BJT NPN ; d'après [5]

Un exemple typique de dimensionnement d'un BJT NPN est donné à la Figure 2-5. En plus des points évoqués, la densité de dopants de la base contient un gradient, et le collecteur est très fortement dopé loin de la jonction. Ces deux points améliorent encore le fonctionnement du BJT.

Pour décrire le fonctionnement du BJT, prenons l'exemple d'un transistor NPN.

Avec $V_{C-B} < 0$, la jonction base – collecteur est polarisée en inverse. Le courant qui traverse la jonction est égal au courant de saturation de la diode. Il est très faible et peut être considéré comme négligeable. De plus, à cause de rapport de dopage $N_B \gg N_C$, l'augmentation de l'extension de la ZCE du côté de la base est négligeable également.

Avec $V_{B-E} > 0$, la jonction base – émetteur est polarisée en direct. Un flux de trous ϕ_p circule dans la jonction, de la base vers l'émetteur, et un flux d'électrons ϕ_n circule dans le sens inverse. On montre que dans un BJT NPN correctement dimensionné ($N_B \ll N_E$), le flux d'électrons est très grand devant le flux de trous ($\phi_n \gg \phi_p$)[4].

Après avoir traversé la jonction base – émetteur, les électrons diffusent dans la base. Une première partie d'entre eux se recombine avec les trous présents en très grand nombre. Une seconde partie parvient à atteindre la jonction base – collecteur sans se recombiner. Ces derniers sont alors acheminés vers le collecteur sous l'effet du champ électrique qui règne dans la ZCE de la jonction base – collecteur (Figure 2-6).

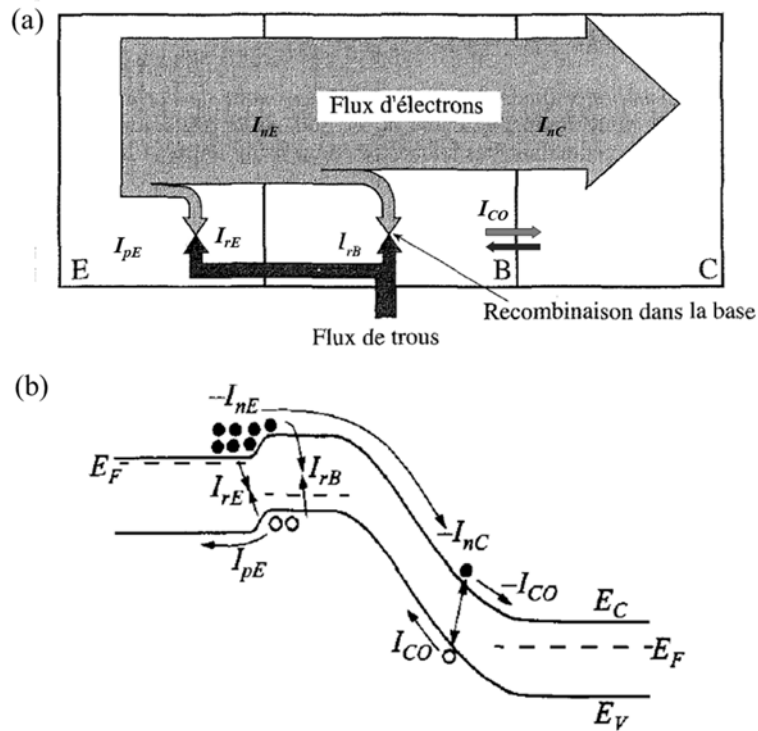


Figure 2-6 – (a) Vue schématique des flux de porteurs dans un BJT NPN en fonctionnement et (b) schéma de bande associé, d'après [4], [5]

La part de flux d'électrons dans le flux total de charge qui circule dans la jonction base – émetteur s'appelle le *rendement d'émetteur* γ_F :

$$\gamma_F = \frac{\phi_{n, BE}}{\phi_{n, BE} + \phi_{p, BE}} = \frac{1}{1 + \frac{\phi_{p, BE}}{\phi_{n, BE}}} \quad (2.39)$$

où $\phi_{n,BE}$ et $\phi_{p,BE}$ sont les flux d'électrons et de trous qui traversent la jonction base – émetteur. On montre que le rapport des deux flux est proportionnel au rapport des dopants des deux régions[4] : $\phi_{p,BE} / \phi_{n,BE} \propto N_B/N_E$. Dans le cas d'un BJT correctement dimensionné, $N_B \ll N_E$, donc $\phi_{p,BE} \ll \phi_{n,BE}$, et $\gamma_F \approx 1$.

La proportion d'électrons arrivée dans la base et qui parvient à rejoindre le collecteur sans se recombiner s'appelle le *facteur de transport dans la base* et se note α_T :

$$\alpha_T = \frac{\phi_{n,BC}}{\phi_{n,BE}} \quad (2.40)$$

Plus la largeur de la base est petite devant la longueur de diffusion des électrons, plus α_T est proche de 1. Dans ces conditions, le produit $\alpha_F = \gamma_F \cdot \alpha_T$, appelé *gain en base commune*, est également proche de 1. On constate qu'il correspond finalement à la part du courant d'émetteur qui atteint le collecteur :

$$\alpha_F = \gamma_F \cdot \alpha_T = \frac{\phi_{n,BC}}{\phi_{n,BE} + \phi_{p,BE}} = \frac{I_C}{-I_E} \quad (2.41)$$

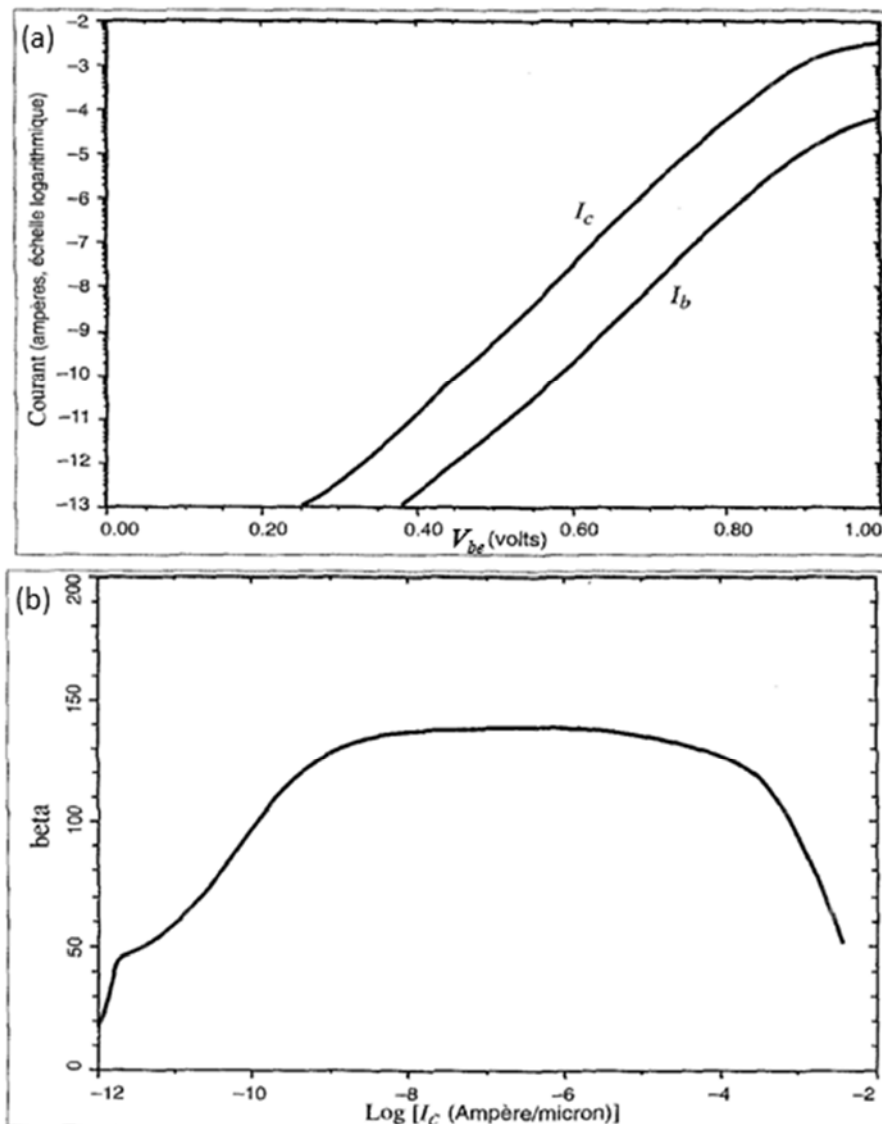


Figure 2-7 – (a) Diagramme de Gummel et (b) gain en courant en émetteur commun β_F en fonction du courant de collecteur d'un BJT typique, d'après [4]. $V_{C-E} = 3$ V.

La part de courant restant circule par la base. Le principe de conservation de la charge $I_B + I_E + I_C = 0$ (loi de Kirchhoff) permet d'établir que :

$$\frac{I_C}{I_B} = \frac{\alpha_F}{1-\alpha_F} \equiv \beta_F \quad (2.42)$$

où $\beta_F = \alpha_F / (1 - \alpha_F)$ est le *gain en émetteur commun*. Plus α_F est proche de 1, plus β_F est grand, typiquement de l'ordre de 100 à 300. Dans ce contexte, le courant de collecteur I_C est indépendant de la tension de base – collecteur, et vaut $\beta_F \cdot I_B$. Pour caractériser le gain, les courants de collecteur et de base sont mesurés en fonction de V_{B-E} et tracés sur un même graphe appelé diagramme de Gummel. β_F est ensuite extrait de ces courbes (Figure 2-7).

Dans l'équation (2.42), le courant de saturation de la jonction base – collecteur a été négligé. Une écriture plus exacte de β_F qui aura son importance pour la suite est [5]:

$$\beta_F = \frac{I_C - I_{CEO}}{I_B} \quad (2.43)$$

où I_{CEO} est le courant de collecteur lorsque la base est laissée flottante.

Pour résumer, afin d'obtenir un BJT capable d'une amplification dans le régime actif, 3 notions sont importantes :

- Avoir un dopage plus fort dans l'émetteur que dans la base, pour maximiser γ_F .
- Avoir une base courte pour maximiser α_T .
- Avoir un collector moins dopé que la base, pour que la ZCE ne s'étende pas du côté de la base.

2.2.3 La capacité métal-oxyde-semiconducteur (MOSCAP)

Une capacité MOS est une structure obtenue en superposant une couche de semiconducteur dopée (n ou p), une couche d'oxyde (matériau isolant) et une couche de métal (Figure 2-8). En appliquant une tension électrique V_{M-SC} entre le métal au potentiel V_M et le semiconducteur au potentiel V_{SC} ($V_{M-SC} = V_M - V_{SC}$), deux charges $+Q$ et $-Q$ apparaissent de part et d'autre de l'oxyde. Dans le métal, cette charge se répartit à l'interface métal – oxyde. Dans le semiconducteur, cette charge se répartit soit à l'interface oxyde-semiconducteur, soit dans le volume du semiconducteur, soit les deux à la fois, selon que la MOSCAP est en régime d'accumulation, de déplétion, d'inversion faible ou d'inversion forte.

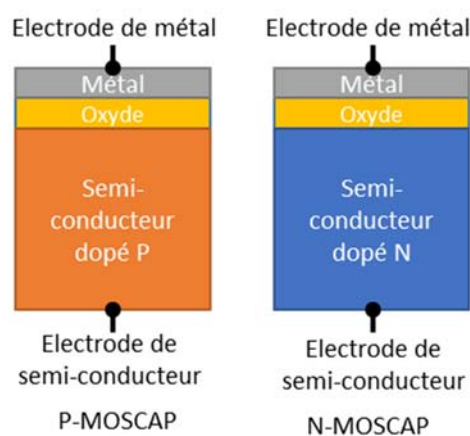


Figure 2-8 – Vue schématique de deux MOSCAP de type P et N

2.2.3.1 Notion de potentiel de surface et de tension de bande plate

- Notion de potentiel de surface

Les différents régimes de la MOSCAP sont distingués par la position du niveau de Fermi dans le gap à l'interface oxyde – semiconducteur. Cette position définit le potentiel de surface ϕ_s qui a pour origine la position du niveau de Fermi dans le gap du semiconducteur loin de l'interface. Lorsque $\phi_s = 0$, aucune charge n'est présente dans la structure.

Par ailleurs, la position du niveau de Fermi dans le gap du semiconducteur loin de l'interface est repérée par rapport au niveau intrinsèque par la distance énergétique ϕ_F , qui joue un rôle dans la définition des régimes de la MOSCAP.

- Notion de tension de bande plate

La différence ϕ_{M-SC} des travaux de sortie du métal ϕ_M et du semiconducteur dopé ϕ_{SC} ($\phi_{M-SC} = \phi_M - \phi_{SC}$) conduit à l'apparition de charge dans la structure MOSCAP en l'absence de tension appliquée ($V_{M-SC} = 0$). Dans cette situation, $\phi_s \neq 0$. Ces charges disparaissent lorsque que $V_{M-SC} = \phi_{M-SC}$. Dans cette condition la tension appliquée V_{M-SC} est égale, par définition, à la tension de bande plate V_{FB} . Les critères de définition du régime de la structure reposent sur la valeur de la différence $V_{M-SC} - V_{FB}$.

2.2.3.2 Régimes de la MOSCAP

La MOSCAP admet plusieurs régimes qui conduisent à différentes valeurs de capacité selon la position ϕ_s dans le gap et la valeur de $V_{M-SC} - V_{FB}$ associée (Figure 2-9).

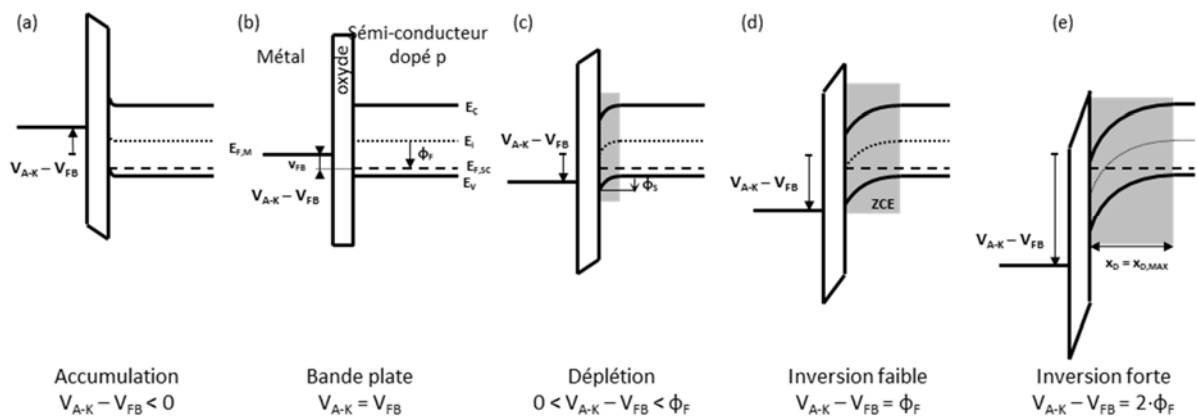


Figure 2-9 – Diagramme de bande d'une P-MOSCAP en régime (a) d'accumulation, (b) de bande-plate, (c) de déplétion, (d) d'inversion faible et (e) d'inversion forte.

- Régime d'accumulation

En régime d'accumulation (Figure 2-9.a), la polarisation de la MOSCAP favorise la présence des porteurs majoritaires dans le semiconducteur (ils sont « accumulés »). Cette accumulation est concentrée à l'interface entre l'oxyde et le semiconducteur. Du point de vue de diagramme de bande, cela se traduit par un rapprochement, au voisinage de l'oxyde, du niveau de Fermi vers la bande de valence ($\phi_s < 0$) ou de conduction ($\phi_s > 0$) selon qu'il s'agit d'une P-MOSCAP ou d'une N-MOSCAP. Dans ce régime, la variation de charge dans le semiconducteur δQ_{SC} induit par une variation de tension δV_{M-SC} est localisée dans la couche d'accumulation. Dans ce contexte, la capacité de la MOSCAP C_{MOSCAP} correspond à la capacité de l'oxyde C_{OX} définie par la permittivité électronique du matériau isolant ϵ_{OX} et son épaisseur t_{OX} :

$$C_{\text{MOSCAP}} = C_{\text{OX}} = \frac{\epsilon_{\text{OX}}}{t_{\text{OX}}} \quad [\text{F} \cdot \text{m}^{-2}] \quad (2.44)$$

- Régime de déplétion

En régime de déplétion (Figure 2-9.c), la polarisation de la MOSCAP repousse les porteurs majoritaires de l'interface oxyde – semiconducteur. Cela fait apparaître une zone de déplétion dans le volume du semiconducteur. Dans la zone de déplétion, les charges des dopants ionisés (de concentration $+qN_D$ ou $-qN_A$ selon le type de semiconducteur) ne sont plus compensées par les charges des porteurs majoritaires. Une densité de charge nette $\rho = \pm qN_{D,A} \neq 0$ apparaît. Dans cette zone, le potentiel électrique V varie selon l'équation de Poisson $\partial^2 V / dx^2 = \pm q \cdot N_{D,A} / \epsilon_{\text{SC}}$, où ϵ_{SC} est la permittivité électronique du semiconducteur. La profondeur de la zone de déplétion x_D augmente avec l'augmentation de la polarisation $V_{\text{M-SC}}$ (positive pour une P-MOSCAP et négative pour une N-MOSCAP) :

$$x_D = \sqrt{\frac{\epsilon_{\text{SC}}^2}{C_{\text{OX}}^2} + \frac{2 \cdot \epsilon_{\text{SC}} \cdot |V_{\text{M-SC}} - V_{\text{FB}}|}{q \cdot N_{D,A}}} - \frac{\epsilon_{\text{SC}}}{C_{\text{OX}}} \quad [\text{m}] \quad (2.45)$$

La charge Q_D de la ZCE vaut alors :

$$Q_D = \pm q \cdot N_{D,A} \cdot x_D \quad [\text{C} \cdot \text{m}^{-2}] \quad (2.46)$$

Dans le diagramme de bande, au voisinage de l'oxyde, le régime de déplétion se traduit par un éloignement du niveau de Fermi de la bande de valence dans le cas d'une P-MOSCAP ($\phi_s > 0$), ou de la bande de conduction dans le cas d'une N-MOSCAP ($\phi_s < 0$). Ce déplacement du niveau de Fermi dans le gap a aussi pour effet d'augmenter la concentration en porteur minoritaire, qui reste cependant inférieure à la concentration des porteurs majoritaires. Dans ce régime, la variation de charge δQ_{SC} induit par $\delta V_{\text{M-SC}}$ est localisée à la frontière entre le ZCE et le semiconducteur. Dans ce contexte, C_{MOSCAP} est égale à la capacité de l'oxyde et de la zone de déplétion en série :

$$C_{\text{MOSCAP}} = \frac{C_{\text{OX}} \cdot C_D}{C_{\text{OX}} + C_D} = \frac{\epsilon_{\text{OX}} \cdot \epsilon_{\text{SC}}}{t_{\text{OX}} \cdot \epsilon_{\text{SC}} + x_D \cdot \epsilon_{\text{OX}}} \quad [\text{F} \cdot \text{m}^{-2}] \quad (2.47)$$

- Régime d'inversion faible

En régime d'inversion faible (Figure 2-9.d), la profondeur de la zone de déplétion continue à s'étendre. Cependant, le déplacement du niveau de Fermi dans le gap est tel que, au voisinage de l'oxyde, la concentration de porteur minoritaire devient plus importante que celle des porteurs majoritaires. Bien que cette situation soit mathématiquement singulière, la contribution des porteurs au bilan de charge électrique est encore négligeable face à la contribution des dopants et le régime d'inversion faible est peu distinguable du régime de déplétion du point de vue phénoménologique.

- Régime d'inversion forte

En régime d'inversion forte (Figure 2-9.e), le déplacement du niveau de Fermi dans le gap au voisinage de l'oxyde est tel que la concentration de porteurs minoritaires devient plus importante que la concentration de dopants. Cette forte densité de porteurs minoritaires est localisée à l'interface oxyde – semiconducteur et forme une couche d'inversion. La contribution des porteurs minoritaires au bilan de charge devient prédominante. Dans cette situation, tout nouvel apport de charge suffisamment lent sera fourni par cette couche d'inversion, et la capacité de la MOSCAP est à nouveau donnée par (2.44). La profondeur de la zone de déplétion x_D ne varie plus et a atteint sa valeur maximale $x_{D,\text{MAX}}$:

$$x_{D,MAX} = \sqrt{\frac{4 \cdot \epsilon_{SC} \cdot \Phi_F}{q \cdot N_D}} \quad [m] \quad (2.48)$$

2.2.3.3 Caractéristique C-V de la MOSCAP

La section précédente a montré que les différentes origines des contributions de charges dans le semiconducteur (couche d'accumulation, zone de déplétion et couche d'inversion) conduisent à des valeurs de capacité C_{MOSCAP} différentes selon les régimes. La caractéristique C-V de la MOSCAP est non-linéaire (Figure 2-10).

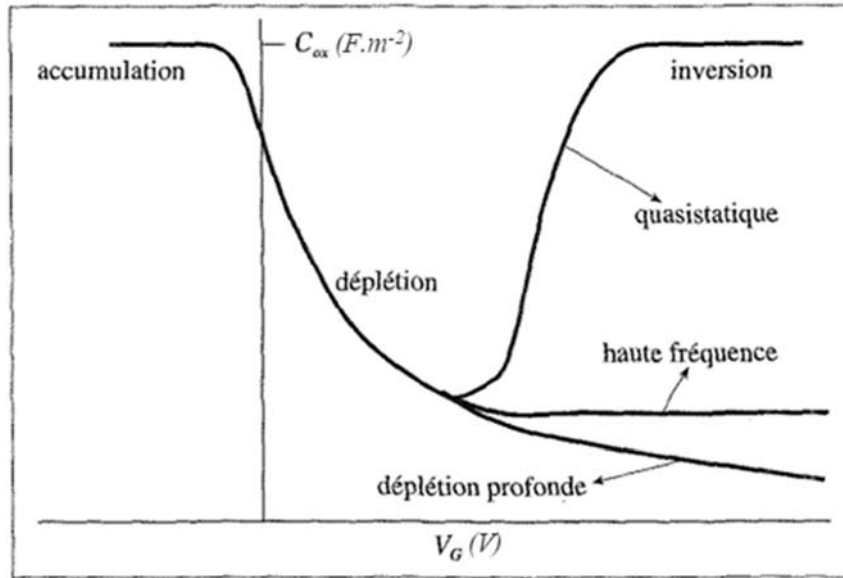


Figure 2-10 – Tendance de la courbe C-V d'une capacité P-MOSCAP (d'après [4]). V_G correspond à V_{M-SC} .

En régime d'inversion forte, plusieurs cas doivent être distingués selon la valeur du temps caractéristique associé à la variation de tension δV_{M-SC} qui induit la variation de charge dans le semiconducteur δQ_{SC} .

Aux temps longs, qui correspondent à l'équilibre thermodynamique, δQ_{SC} est fournie par la variation de concentration de porteur minoritaire dans la couche d'inversion, comme évoqué à la section précédente. La capacité correspondante est la même qu'en régime d'inversion (équation (2.44)). Ces porteurs minoritaires proviennent de la génération thermique dans la ZCE, qui met un temps non nul à fournir la charge δQ_{SC} [9], [10].

Aux temps courts, la génération thermique n'a pas le temps de fournir δQ_{SC} qui est alors compensée par l'extension de la zone de déplétion dans le semiconducteur. La capacité correspondante est la même qu'en régime de déplétion (équation (2.47)). Ces deux différentes natures de la contribution δQ_{SC} conduisent à des valeurs de C_{MOSCAP} qui dépendent de la fréquence de variation de δV_{M-SC} :

$$(\text{hautes fréquences}) \frac{C_{OX} \cdot C_D}{C_{OX} + C_D} < C_{MOSCAP} < C_{OX} \text{ (basses fréquences)} \quad [F \cdot m^{-2}] \quad (2.49)$$

Le cas de la déplétion profonde en régime de déséquilibre n'est pas abordé dans ce document.

2.2.4 Le transistor à effet de champ métal-oxyde-semiconducteur (MOSFET)

A l'instar du BJT, le MOSFET est aussi une structure obtenue par l'aboutement de deux jonctions PN (Figure 2-11), mais, contrairement au BJT, une sous-structure appelée grille est en plus disposée au-dessus de la zone centrale commune aux deux diodes qui, dans ce contexte, est appelée « body ». Cette grille est constituée d'un empilement d'une couche d'oxyde isolante et d'une couche conductrice en métal ou en poly-silicium fortement dopé. Par ailleurs, les contraintes de dopage, de dimensionnement et de polarisation sont différentes de celles du BJT. En particulier, les zones externes et de même dopages, appelées source et drain respectivement, ont des concentrations de dopage similaires ou même identiques (la structure est alors symétrique). En ce qui concerne les tensions appliquées lors du fonctionnement du MOSFET, la diode du côté source n'est généralement pas polarisée et la diode du côté du drain est polarisée en inverse. Il existe deux types de MOSFET selon que les source et drain sont de type N (N-MOSFET) ou P (P-MOSFET).

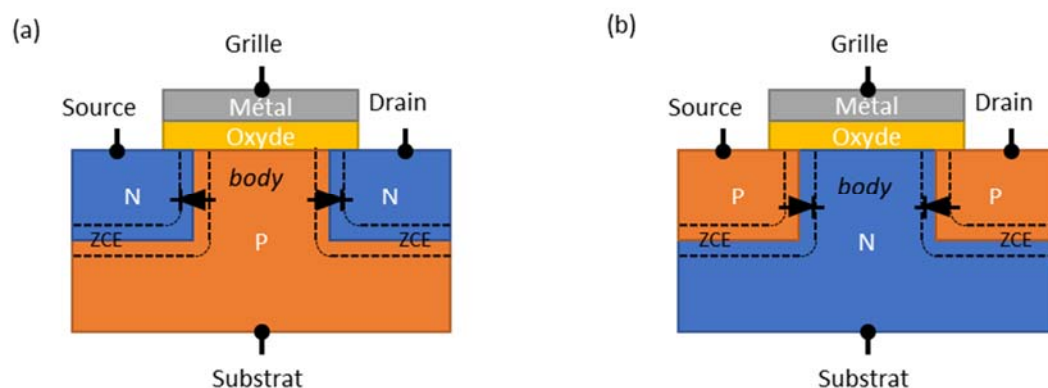


Figure 2-11 – Vue schématique d'un N-MOSFET et d'un P-MOSFET

Les courants qui circulent dans le MOSFET sont proportionnels au rapport W/L de la largeur W et la longueur L de la grille. Ces deux dimensions sont les deux paramètres principaux sur lesquels jouer pour construire ces circuits électroniques qui rassemblent plusieurs MOSFET. Les technologies qui utilisent les MOSFET comme briques élémentaires sont appelées « complementary metal oxide semiconductor » (CMOS).

La sous-structure constituée par le substrat et la grille est identique à celle de la MOSCAP. Les mêmes notions d'accumulation, de déplétion et d'inversion faible ou forte, développées à la section précédente, y sont applicables. Lorsqu'elle est présente, la ZCE de la MOSCAP fusionne avec celle des deux diodes. (MOSCAP en régime de déplétion ou d'inversion faible ou forte).

Le principe de fonctionnement du MOSFET est de contrôler la conductance entre la source et le drain avec la polarisation de la grille. En première approximation, la conductance est nulle lorsque la MOSCAP est en régime d'accumulation, de déplétion ou d'inversion faible, et elle est non-nulle lorsque la MOSCAP est en régime d'inversion forte. En effet, dans ce dernier cas, la couche d'inversion située à l'interface oxyde-substrat, est constituée de porteurs de même type que les porteurs majoritaires présents dans la source et dans le drain. Cette couche d'inversion constitue un plan conducteur, appelé canal, qui permet au courant de circuler entre la source et le drain. Notons par ailleurs que la source et le drain constituent des réservoirs de porteurs minoritaires qui fournissent les charges à la couche d'inversion. La variation de capacité en fonction de la fréquence des variations de tension de la grille n'est pas effective ici.

La tension grille – substrat V_{G-SUB} pour laquelle la MOSCAP atteint le régime d'inversion forte est appelée tension de seuil du transistor (« threshold voltage » : $V_{T,MOS}$). Dans le cas d'un N-MOSFET, cette tension vaut :

$$V_{T,MOS} = V_{FB} + 2 \cdot \phi_F - \frac{\sqrt{4 \cdot q \cdot N_A \cdot \epsilon_{SC} \cdot \phi_F}}{C_{OX}} \quad [V] \quad (2.50)$$

- Caractéristique courant-tension

En présence du canal d'inversion, le MOSFET est en régime au-dessus du seuil. Lorsque la tension drain-source V_{D-S} est faible, le MOSFET se comporte comme une résistance variable dont la valeur est proportionnelle à $(V_G - V_{T,MOS})$. Il est en régime *linéaire*. Le courant qui circule entre la source et le drain est asymptotiquement donné par :

$$I_D = \mu_n \cdot C_{OX} \cdot \frac{W}{L} \cdot (V_G - V_{T,MOS}) \cdot V_{D-S} \quad [A] \quad (2.51)$$

où μ_n est la mobilité des électrons, W et L la largeur et la longueur de la grille.

Lorsque V_D dépasse une valeur de saturation $V_{D,SAT} \approx V_G - V_{T,MOS}$, le MOSFET entre en régime de saturation. Une zone de pincement du canal (« pinch-off »), où la couche d'inversion disparaît, apparaît du côté du drain. Le courant $I_D = I_{D,SAT}$ est maximal et devient indépendant de V_{D-S} . Il vaut :

$$I_{D,SAT} = \mu_n \cdot C_{OX} \cdot \frac{W}{L} \cdot \frac{K}{2} \cdot (V_G - V_{T,MOS})^2 \quad [A] \quad (2.52)$$

où K est un facteur qui dépend de C_{OX} , q, N_A , ϵ_{SC} et ϕ_F et dont la valeur est proche de 1.

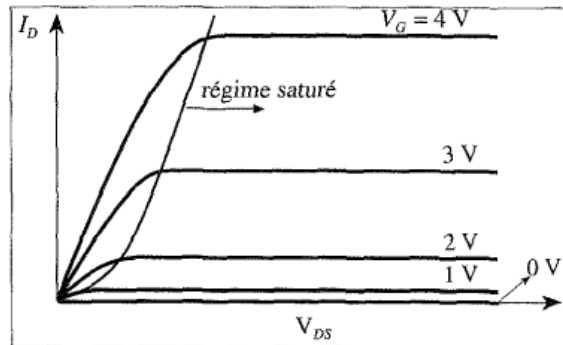


Figure 2-12 – Caractéristique typique I_D-V_D d'un N-MOSFET en régime au-dessus du seuil [4]

En dessous du seuil, le courant de drain I_D est indépendant de V_{D-S} dès que V_{D-S} est supérieur à quelques U_T . On montre que sa dépendance en V_{G-S} est de la forme [4] :

$$I_D = \frac{W}{L} \cdot I_0 \cdot \exp\left(\frac{V_{G-S}}{n \cdot U_T}\right) \quad [A] \quad (2.53)$$

Où I_0 est un courant de référence, qui dépend très légèrement de V_G , et n est un facteur qui implique le rapport des capacités de déplétion et d'oxyde :

$$n = 1 + \frac{C_D}{C_{OX}} \quad (2.54)$$

La pente sous le seuil SS, quantifie, en volt par décade de courant, la croissance exponentielle du courant I_D dans ce régime :

$$SS = \frac{\partial V_{G-S}}{\partial \log I_D} = \ln(10) \cdot n \cdot U_T = \ln(10) \cdot \left(1 + \frac{C_D}{C_{OX}}\right) \cdot U_T \quad [V \cdot \log(A)^{-1}] \quad (2.55)$$

Sa valeur est minimale pour $C_{OX} \gg C_D$, et vaut 60 mV/dec à 300 K. Elle est une limite fondamentale des transistors MOSFET. Une courbe typique de la caractéristique I_D - V_G est présentée à la Figure 2-13, en échelle semi-logarithmique et en échelle linéaire.

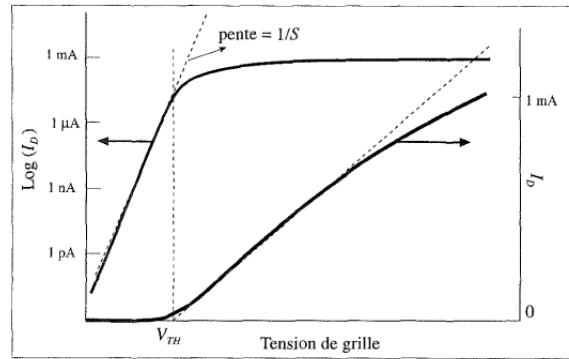


Figure 2-13 – Caractéristique typique I_D - V_G d'un N-MOSFET, en échelle semi-logarithmique (axe de gauche) et linéaire (axe de droite) [4]

Certains aspects de la physique des transistors MOSFET ne sont pas développés ici, notamment :

- La diminution de la mobilité μ à fort champ électrique conduit à une saturation de la vitesse des porteurs.
- La possibilité pour des charges d'être piégées dans l'oxyde de grille ou l'interface oxyde-semiconducteur modifie la tension de seuil du transistor, ajoute une capacité supplémentaire et augmente la variabilité entre les composants.

2.3 La technologie 28nm FD-SOI

Cette section présente la technologie utilisée pour la fabrication des échantillons présentés dans ce manuscrit de thèse. Les raisons du développement de cette technologie sont contextualisées par une évocation des défis technologiques liés à la miniaturisation des composants.

2.3.1 « Downscaling » et effet de canaux courts

La réduction de la taille des transistors MOSFET, appelée *downscaling* et théorisée de façon performative par Gordon Moore en 1965 (loi de Moore) [11], a été l'élément clef du développement de la microélectronique. Le *downscaling* permet à la fois d'améliorer la performance des circuits CMOS et de réduire leur coût de production unitaire.

Cependant, lorsque la longueur de grille devient trop petite, des effets indésirables qui nuisent au bon fonctionnement du MOSFET apparaissent. Ces effets portent le nom de « short channel effect » (SCE). Ils consistent principalement en une perte de contrôle électrostatique de la grille sur le canal. Les 3 principaux SCEs sont les suivants :

1. *Punch-through* : La polarisation du drain a pour effet d'étendre la ZCE de la jonction drain-substrat. Lorsque la longueur de grille est courte, cette extension peut rejoindre la ZCE de la jonction source-substrat et créer un chemin de conduction qui court-circuite le canal.

2. *Drain induced Barrier lowering* (DIBL) : La polarisation du drain a pour effet de réduire la hauteur de la barrière de potentiel dans la région du canal au voisinage de la source. Cela augmente le courant de fuite et réduit la tension de seuil du transistor.
3. V_T *roll-off* : Lorsque les profondeurs des ZCE des jonctions source-substrat et drain-substrat deviennent du même ordre de grandeur que la longueur de grille, la tension de seuil du transistor diminue avec la réduction de la longueur de grille (Figure 2-14).

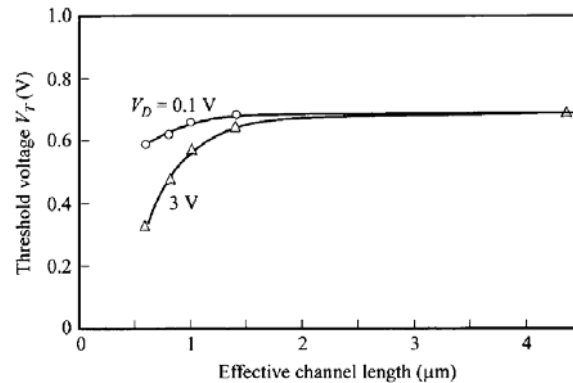


Figure 2-14 – Dépendance de la tension de seuil en fonction de la longueur de grille et de la polarisation du drain [12]

On montre que le V_T roll-off diminue lorsque la capacité de l'oxyde de grille augmente [13], notamment en réduisant l'épaisseur de grille t_{ox} . Il s'agit de redonner du pouvoir électrostatique à la grille sur le canal en augmentant C_{ox} . Cependant cette diminution amène aussi son lot d'effets indésirables :

1. Gate leakage current : En dessous de quelques nanomètres, les électrons peuvent circuler au travers de l'oxyde par effet tunnel, ou par injection de porteurs chauds. La grille n'est plus isolée des autres terminaux, et la consommation augmente.
2. Gate induced drain leakage (GIDL) : Lorsque la tension drain-grille est maximale (i.e. égale à la tension nominale utilisée dans le circuit), le champ électrique devient suffisamment fort à la jonction drain-canal pour générer des paires électrons-trous par « band-to-band-tunneling ». Ces porteurs sont séparés et évacués par le substrat et le drain. Ils forment ainsi un courant de fuite.

Le déploiement de stratégies d'ingénierie pour diminuer les SCEs complexifie les processus de fabrication et augmente leur coût. Une façon de réduire les SCEs, consiste à disposer la structure MOSFET sur un substrat de silicium-sur-isolant (« silicon-on-insulator » : SOI) de type particulier : les substrats SOI entièrement déplétés avec un body et un BOX ultra fins (« ultra thin body and box fully-depleted SOI » : UTBB FD-SOI). Le détail des caractéristiques de ce type de substrat et les raisons de son efficacité contre les SCEs sont l'objet des paragraphes suivants.

2.3.2 Les technologies SOI

Les substrats UTBB FD-SOI sont en premier lieu des substrats SOI. De telles structures contiennent une couche d'oxyde enterrée (« buried oxide » : BOX) disposée sous leur surface, qui permet d'isoler électriquement la partie supérieure du substrat, appelée « film », de la partie inférieure. Ainsi, dans les structures MOSFET sur substrat SOI, le drain, la source et la région du canal sont isolés du substrat par le BOX. La région du canal prend dès lors le nom de « body ». Dans cette configuration, les jonctions PN drain – substrat, source – substrat et well – substrat sont éliminées ainsi que les effets parasites

inhérents à ces jonctions : courant de fuite des jonctions en inverse, capacité de jonction et sensibilité au latch-up (Figure 2-15).

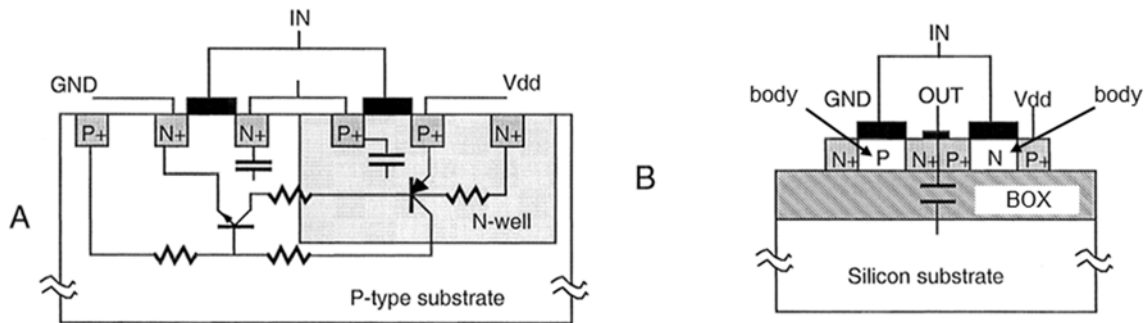


Figure 2-15 – A : Vue en coupe montrant le chemin de latch-up dans un inverseur CMOS bulk. B : Vue en coupe d'un inverseur CMOS SOI. Les capacités parasites de drain sont aussi représentées. D'après [14].

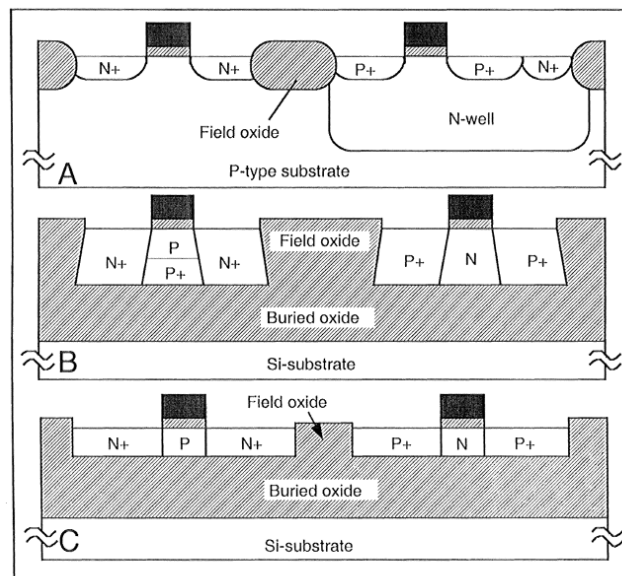


Figure 2-16 – Vue en coupe d'un inverseur CMOS en A : bulk, B : PD-SOI et C : FD-SOI. D'après [14]

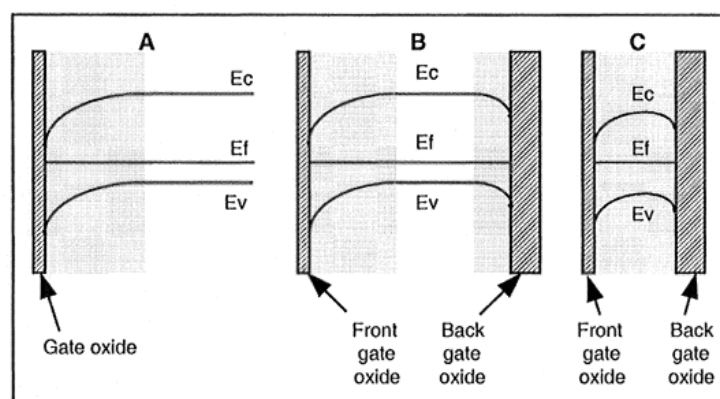


Figure 2-17 – Diagramme de bande d'énergie en (A) bulk, (B) PD-SOI et (C) FD-SOI. Tous les composants sont représentés au seuil ($V_{FG} = V_{T,MOS}$). Les zones grisées représentent les zones déplétées. Les composants SOI sont représentés en situation d'inversion faible (et donc, sous le seuil) à l'interface arrière. D'après [14].

La présence du BOX rajoute un couplage capacitif entre le substrat et les éléments du MOSFET. En particulier, la polarisation du substrat peut influencer le potentiel de surface à l'interface entre le body et le BOX jusqu'à créer un canal de conduction « arrière », par opposition au canal de conduction « avant » créé par la polarisation de la grille. Ainsi, un SOI MOSFET peut être vu comme un MOSFET comportant une grille-avant (« front gate » : FG) et une grille-arrière (« Back-gate » : BG). Il existe alors deux tensions de seuil associées à chacune de ces grilles, notées $V_{T,MOS,FG}$ et $V_{T,MOS,BG}$. La notion de grille-arrière n'est pertinente que si l'épaisseur d'oxyde est suffisamment fine pour que le couplage capacitif au travers du BOX entre le body et le substrat soit suffisamment important.

On distingue deux types de substrat SOI, selon que, après constitution d'une structure MOSFET, le body est partiellement déplété (« partially depleted SOI » : PD-SOI) ou entièrement déplété (« fully depleted » : FD-SOI) (Figure 2-16 et Figure 2-17). Autrement dit, le body des PD-SOI MOSFET contient une zone neutre et flottante alors que le body des FD-SOI MOSFET n'en contient pas. La qualité de la déplétion dépend de l'épaisseur du film de silicium et de la concentration de dopants du body. Un film épais et fortement dopé conduira à des structures PD-SOI tandis qu'un film mince et non dopé donnera lieu à des structures FD-SOI.

Dans la suite, les caractéristiques du PD-SOI sont d'abord exposées avant celle du FD-SOI. Cet ordre d'exposition permet, d'une part, d'introduire des concepts propres au PD-SOI mais néanmoins importants pour la compréhension des travaux exposés dans cette thèse, et, d'autre part, de mettre davantage en valeur l'intérêt du FD-SOI par la description des effets parasites liés au PD-SOI.

2.3.2.1 Les transistors PD-SOI

La présence d'une zone neutre dans le body des PD-SOI MOSFET a pour premier effet de découpler électriquement l'interface body – oxyde de l'interface body – BOX. En conséquence, les tensions de seuil des grilles avant et arrière sont indépendantes l'une de l'autre. En particulier, $V_{T,MOS,FG}$ ne dépend ni du potentiel du substrat, ni des charges piégées à l'interface body-BOX ou dans le BOX.

La zone neutre dans le body des PD-SOI MOSFET est à l'origine d'effets parasites comme le « kink effect » ou le « single-transistor latch-up ». En effet, le potentiel de la zone neutre n'est pas contrôlé par un terminal électrique. Elle est donc électriquement flottante. On parle de floating-body MOSFET. Le potentiel du body dépend des charges qui s'y accumulent et des couplages capacitifs avec le reste de la structure. Or, la variation de ce potentiel a pour effet de modifier la tension de seuil du transistor $V_{T,MOS,FG}$, comme dans les MOSFET bulk lorsque le substrat est polarisé. Elle peut également activer un effet bipolaire parasite par le couplage des jonctions source-body et drain-body. Ces deux phénomènes modifient chacun la conductance drain – source.

Outre les couplages capacitifs dont l'effet se manifeste à haute fréquence, la principale façon de modifier le potentiel de body se fait par l'accumulation de porteurs majoritaires (vis-à-vis du dopage du body). Ils sont issus des différents phénomènes de génération qui peuvent avoir lieu dans la structure : ionisation par impact, effet tunnel bande-à-bande, photo-génération. Une fois générés, ces porteurs sont acheminés vers la zone flottante par les champs électriques des ZCEs des jonctions drain – body et source – body et de la MOSCAP grille-avant – body. Le potentiel du body varie en conséquence. Cette variation s'arrête lorsque le potentiel atteint un seuil pour lequel il devient plus favorable aux porteurs majoritaires générés de rejoindre la source plutôt que le body.

Il est possible de supprimer ces effets parasites en empêchant l'accumulation de porteur majoritaire par leur évacuation dans un cinquième terminal électrique appelé « body contact ». Il est généralement disposé perpendiculairement à l'axe drain-source. On parle de body-contacted MOSFET

ou de body-tied MOSFET. Dans ce cas-là, le comportement du PD-SOI MOSFET est équivalent à celui d'un MOSFET sur substrat standard, avec le body contact en lieu et place du contact de substrat.

2.3.2.2 Les transistors FD-SOI

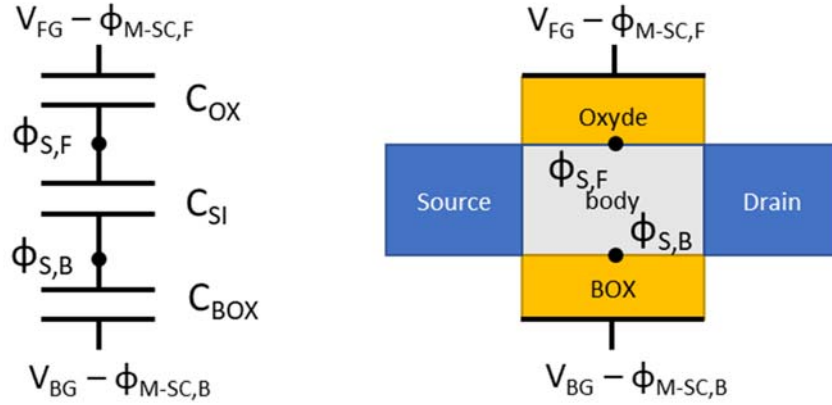


Figure 2-18 – Réseau de capacité équivalent dans un transistor FD-SOI en l'absence de couche d'inversion dans le canal avant ou arrière (transistor sous le seuil) et en considérant la charge de la zone de déplétion du body négligeable.

L'absence de zone neutre dans les FD-SOI MOSFET les affranchit a priori des effets parasites présents dans les PD-SOI MOSFET. Le champ électrique dû à la déplétion complète du body évacue systématiquement les porteurs majoritaires générés vers la source du composant sans qu'ils s'accumulent dans le body. L'intérêt de l'ajout d'un body-contact disparaît et les transistors FD-SOI en sont dépourvus. Par ailleurs, l'absence de zone neutre a pour effet de coupler fortement les interfaces avant et arrière : les tensions de seuils des grilles avant et arrière du transistor dépendent l'une de l'autre. La grandeur qui rend compte de ce couplage est notée α et correspond à la dérivée de la tension de seuil de la grille avant $V_{T,MOS,FG}$ par rapport au potentiel de la grille-arrière V_{BG} (équation (2.56)). Pour l'établir, il convient de considérer le réseau capacitif constitué par l'oxyde de grille, le film de silicium et le BOX (Figure 2-18).

$$\alpha = \frac{\delta V_{TH,FG}}{\delta V_{BG}} = \left. \frac{\delta V_{FG}}{\delta V_{BG}} \right|_{\phi_{FS}=cste} = - \frac{C_{SI}C_{BOX}}{C_{OX}(C_{SI} + C_{BOX})} \quad (2.56)$$

Le body entièrement déplété permet d'atteindre une pente sous le seuil très proche de sa limite théorique de 60 mV/dec à température ambiante.

2.3.2.3 Les technologies UTBB FD-SOI

Les paragraphes précédents ont montré que la présence du BOX dans les substrats SOI permet de supprimer les courants de fuite par le substrat et que l'utilisation d'un substrat FD-SOI permet en plus de s'affranchir des effets de body. Pour s'affranchir des SCEs il faut de plus que l'épaisseur du film et de l'oxyde de grille soient suffisamment fines. Ainsi on montre qu'en introduisant une « longueur naturel » λ définie par :

$$\lambda = \sqrt{\frac{\epsilon_{SI}}{\epsilon_{OX}} \cdot t_{OX} \cdot t_{SI}} \quad [m] \quad (2.57)$$

un transistor FD-SOI est dépourvu de SCE si sa longueur de grille effective est 5 à 10 fois plus grande que λ [15]. Les substrats FD-SOI avec un film très fin sont qualifiés de « ultra-thin body FD-SOI ». Ils

sont généralement pourvu d'un BOX très fin également. Cela permet d'avoir un contrôle efficace de la tension de seuil des transistors. On parle alors de substrat « ultra thin body and BOX » (UTBB FD-SOI).

2.3.2.4 La technologie 28nm UTBB FD-SOI.

L'utilisation de substrat UTBB FD-SOI a permis de développer des technologies dites « agressives » avec des longueurs de grille réduites et des performances augmentées. Les échantillons présentés dans cette thèse ont été fabriqués en technologie 28nm FD-SOI développée par STMicroelectronics, où les longueurs de grilles peuvent descendre jusqu'à 28nm. Les caractéristiques de cette technologie, pertinentes pour le travail présenté ici, sont les suivantes :

- L'épaisseur du BOX est de 25 nm,
- L'épaisseur du film de silicium est de 7-8 nm
- Les grilles sont de type « High-k metal gate ». Elles sont constituées d'un empilement d'oxyde de silicium, d'oxyde à haute permittivité, de métal, de silicium poly-cristallin dopé et de siliciure. L'utilisation d'oxyde high-k permet de réduire les courants de fuite de grille en augmentant l'épaisseur de l'oxyde sans réduire la valeur de la capacité C_{ox} .
- Deux types de grilles sont disponibles selon l'épaisseur d'oxyde de silicium équivalent (« equivalent oxide thickness » : EOT) :
 - Les grilles standards (« standard gate » : SG) avec un EOT de 1.1 nm
 - Les grilles étendues (« Extended gate » : EG) avec un EOT de 3.4 nmLes transistors utilisant ces grilles sont qualifiés respectivement de GO1 et de GO2
- Un dopage de type P ou N est disponible dans la région du substrat sous le BOX, pour former des puits (« well ») de type P ou de type N. Le profil de dopage de ces wells n'est pas constant : il décroît en fonction de la profondeur sous le BOX. Il est maximal au proche voisinage de l'interface BOX – substrat dans la région appelée plan arrière (« back plane »). La forte densité de dopage du back plane revient à changer le travail de sortie de la grille-arrière. De par le couplage entre les interfaces avant et arrière dans la technologie UTBB SOI, ce changement de travail de sortie change la tension de seuil de la grille-avant [16]–[19] :
 - Les transistor NMOSFET sur P-well et PMOSFET sur N-well ont des tensions de seuil régulière (« regular V_T » : RVT).
 - Les transistor NMOSFET sur N-well et PMOSFET sur P-well ont des tensions de seuil faibles (« low V_T » : LVT).

On parle de saveur RVT ou LVT. Les tensions de seuil approximatives des transistors de la technologie 28nm UTBB FD-SOI sont répertoriées au Tableau 2-1 selon leur saveur et leur épaisseur de grille.

Une vue schématique en coupe d'un transistor 28nm UTBB FD-SOI est présenté à la Figure 2-19. Outre l'architecture du transistor en lui-même, elle illustre la façon dont l'accès électrostatique à la grille arrière est implémentée grâce à une zone dite « hybride » où le BOX est retiré [20]. Ces zones hybrides permettent notamment la fabrication de composant en technologie bulk au côté des composants fabriqués en technologie 28nm UTBB FD-SOI. L'isolation entre cet accès et le transistor, ainsi qu'entre tous les transistors présents sur la puce, est assurée par la présence de tranchée d'isolation peu profonde (« Shallow Trench Isolation » : STI).

Tableau 2-1 – Tensions de seuil approximatives en mV des transistors MOSFET typiques de la technologie 28nm UTBB FD-SOI. $W = L = 1 \mu\text{m}$

Type	NMOSFET	PMOSFET
SG RVT	500	600
SG LVT	450	550
EG RVT	550	500
EG LVT	400	550

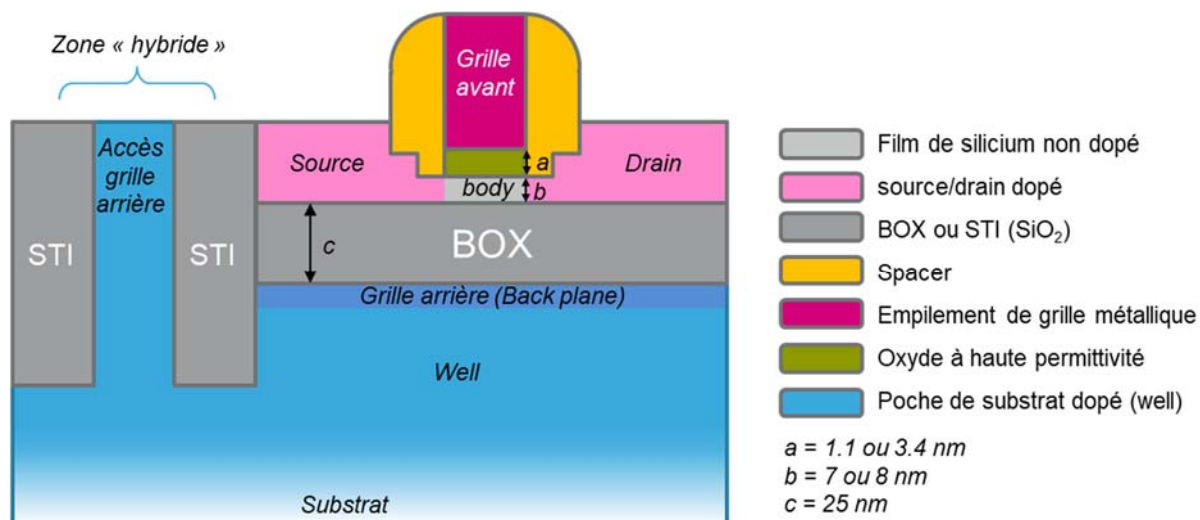


Figure 2-19 – Vue schématique en coupe d'un transistor en technologie 28 nm UTBB FD-SOI

2.4 Références

- [1] P. Galy and V. Galy, Description des monocristaux parfaits. Ellipses Marketing, 1999.
- [2] P. Galy and V. Galy, Description énergétique des constituants de la micro et nano-électronique. Ellipses Marketing, 2005.
- [3] W. Shockley and W. T. Read, "Statistics of the Recombinations of Holes and Electrons," *Phys Rev*, vol. 87, no. 5, pp. 835–842, Sep. 1952.
- [4] J.-P. Colinge and F. Van de Wiele, Physique des dispositifs semi-conducteurs. De Boeck, 1996.
- [5] S. M. Sze and K. K. Ng, Physics of Semiconductor Devices. John Wiley & Sons, 2006.
- [6] C. Sah, R. N. Noyce, and W. Shockley, "Carrier Generation and Recombination in P-N Junctions and P-N Junction Characteristics," *Proc. IRE*, vol. 45, no. 9, pp. 1228–1243, Sep. 1957.
- [7] A. S. Grove, Physics and technology of semiconductor devices. Wiley, 1967.
- [8] A. J. Scholten, G. D. J. Smit, M. Durand, R. V. Langevelde, and D. B. M. Klaassen, "The Physical Background of JUNCAP2," *IEEE Trans. Electron Devices*, vol. 53, no. 9, pp. 2098–2107, Sep. 2006.
- [9] A. S. Grove, B. E. Deal, E. H. Snow, and C. T. Sah, "Investigation of thermally oxidised silicon surfaces using metal-oxide-semiconductor structures," *Solid-State Electron.*, vol. 8, no. 2, pp. 145–163, 1965.
- [10] S. R. Hofstein and G. Warfield, "Physical limitations on the frequency response of a semiconductor surface inversion layer," *Solid-State Electron.*, vol. 8, no. 3, pp. 321–341, 1965.
- [11] G. E. Moore, "Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff.," *IEEE Solid-State Circuits Soc. Newsl.*, vol. 11, no. 3, pp. 33–35, Sep. 2006.
- [12] L. Yau, "A simple theory to predict the threshold voltage of short-channel IGFET's," *Solid-State Electron.*, vol. 17, no. 10, pp. 1059–1063, 1974.
- [13] Z.-Liu et al., "Threshold voltage model for deep-submicrometer MOSFETs," *IEEE Trans. Electron Devices*, vol. 40, no. 1, pp. 86–95, Jan. 1993.
- [14] J.-P. Colinge, Silicon-on-Insulator Technology: Materials to VLSI, 3rd edition. Boston: Springer, 2004.
- [15] R.-Yan, A. Ourmazd, and K. F. Lee, "Scaling the Si MOSFET: from bulk to SOI to bulk," *IEEE Trans. Electron Devices*, vol. 39, no. 7, pp. 1704–1710, Jul. 1992.
- [16] J.-Noel, O. Thomas, C. Fenouillet-Beranger, M.-Jaud, P. Scheiblin, and A. Amara, "A simple and efficient concept for setting up multi-VT devices in thin BOX fully-depleted SOI technology," in 2009 Proceedings of the European Solid State Device Research Conference, 2009, pp. 137–140.
- [17] J. Noel et al., "Multi-TUTBB FDSOI Device Architectures for Low-Power CMOS Circuit," *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2473–2482, Aug. 2011.
- [18] C. Fenouillet-Beranger et al., "Efficient multi-VTFDSOI technology with UTBOX for low power circuit design," in 2010 Symposium on VLSI Technology, 2010, pp. 65–66.
- [19] B. Pelloux-Prayer et al., "Fine grain multi-VT co-integration methodology in UTBB FD-SOI technology," in 2013 IFIP/IEEE 21st International Conference on Very Large Scale Integration (VLSI-SoC), 2013, pp. 168–173.
- [20] D. Golanski et al., "First demonstration of a full 28nm high-k/metal gate circuit transfer from Bulk to UTBB FDSOI technology through hybrid integration," in 2013 Symposium on VLSI Technology, 2013, pp. T124–T125.

Chapitre 3 : Le 28 nm FD-SOI BIMOS : Caractérisation et modélisation

Le circuit de neurone développé dans cette thèse s'appuie sur un composant particulier, appelé BIMOS, intégré en technologie 28nm FD-SOI. Dans un premier temps, ce chapitre décrit le principe général du BIMOS et la topologie ainsi que les dimensions du BIMOS utilisé comme référence. Ensuite, il présente les résultats des caractérisations effectuées sur ce dernier, et propose une modélisation compacte de son comportement avec pour objectif d'intégrer ce composant dans des simulateurs de type SPICE. Les caractérisations sont réparties en 3 groupes selon le mode de fonctionnement du BIMOS qu'elles exploitent (gated-diode, MOSFET ou BJT) et sont chacune l'objet d'analyses et de discussions autour de la nature des contributions aux différents courants avec pour objectif de préciser la stratégie de modélisation.

3.1 Principe général du BIMOS

Le transistors BJT NPN et le N-MOSFET comportent tous deux une zone P entourée de deux zones N. Dès lors, il est possible d'envisager un composant appelé BIMOS (contraction de *bipolar* et *MOSFET*), introduit notamment dans [1], qui comporte à la fois une grille et une base, et qui mêle les deux types de fonctionnement. Ce composant a 4 terminaux : la source ou émetteur, le drain ou collecteur, la base ou body-contact, et la grille (Figure 3-1).

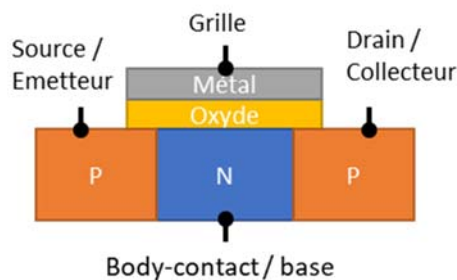


Figure 3-1 – Schéma de principe du BIMOS.

Le BIMOS a été étudié et intégré sur substrat BULK [2]–[5], sur substrat PD-SOI [6]–[11] et sur substrat FD-SOI [12], [13]. Selon les travaux, il est désigné de plusieurs façons (Tableau 3-1). On distingue 3 modes de fonctionnement du BIMOS : le mode MOSFET, BJT et hybride. Dans les deux premiers modes, les polarisations des terminaux sont telles que l'un des deux mécanismes MOSFET ou BJT est clairement dominant. Dans le mode hybride, obtenu en connectant la grille et la base ensemble, les deux mécanismes sont activés en même temps. Généralement cela conduit à de meilleures performances du gain en courant du BJT, une amélioration de la pente sous le seuil et de la transconductance du MOSFET[5]–[9]. Cependant, comme la base et la grille sont connectées ensemble, le courant de base augmente le courant de fuite lorsque le BIMOS est intégré dans des circuits CMOS. Il n'est à ce jour pas utilisé dans des circuits logiques. En revanche, lorsqu'il est couplé à une résistance, il est utilisé pour ses hautes performances en circuit de protection ESD [4], [12].

Le circuit de neurone analogique développé dans cette thèse s'inspire fortement de l'utilisation du BIMOS dans les protections ESD. Dans le but de construire à terme des circuits neuro-morphiques à haute densité d'intégration, le choix s'est porté sur la technologie 28nm FD-SOI qui offre des dimensions agressives. Le 28nm FD-SOI BIMOS a déjà été étudié comme protection ESD dans [12], avec un dimensionnement qui consomme une surface de silicium adaptée à cette application (grande surface pour pouvoir conduire une quantité importante de courant). L'étude des 28nm FD-SOI BIMOS

de taille minimale, proche des limites inférieures de dimensionnement en technologie 28nm FD-SOI n'a pas été faite à ce jour. Ainsi, la suite de ce chapitre s'attache à caractériser par la mesure un 28nm FD-SOI BIMOS de petite taille, optimisé pour l'intégration à grande échelle. L'objectif est d'identifier les mécanismes de physique du composant responsables de la forme des courbes courant-tension mesurées dans le but de proposer un modèle compact qui permette l'exécution de simulation SPICE. 3 types de caractérisation sont effectués, selon le point de vue sous lequel le BIMOS est considéré : gated-diode, MOSFET ou BJT.

Tableau 3-1 – Caractéristiques des BIMOS présents dans la littérature.

Référence	Nom donné au BIMOS	Type de substrat	Gain en courant β en mode hybrid
Galy [1]	BIMOS	Bulk	13.6
Olsson et al. [5]	Bipolar-MOS (BIPMOS)	Bulk	$2 \cdot 10^3$
Huang et al. [7]	Drive-in gate controlled hybrid transistor (DGCHT)	PD-SOI	10^4
Verdonckt-Vandebroek et al. [2]	Hybrid-mode device	Bulk	$2.5 \cdot 10^3$
Assaderaghi et al. [8]	Dynamic Threshold Voltage MOSFET (DTMOS)	PD-SOI	-
Colinge et al. [9]	Voltage-controlled bipolar-MOS (VCBM)	PD-SOI	$1.4 \cdot 10^1$
Athanasious et al. [13]	BIMOS	FD-SOI	$1.8 \cdot 10^7$

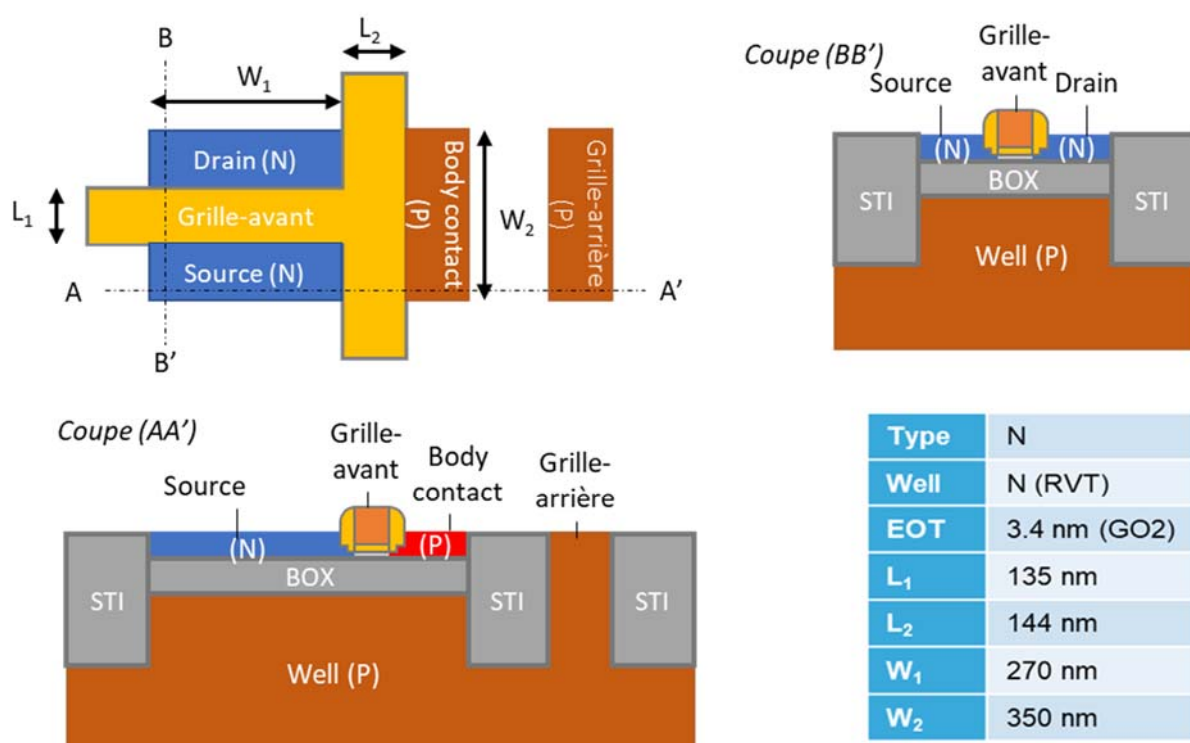


Figure 3-2 – Topologie et dimensions du 28nm FD-SOI BIMOS de référence.

3.2 Topologie et dimensions du BIMOS de référence

Les mesures de caractérisation présentées dans la suite (section 3.3 à 3.5) ont toutes été effectuées sur un 28nm FD-SOI BIMOS de référence (structure BIMOS dans le film mince), dont la topologie et les

dimensions sont présentées à la Figure 3-2. Le choix s'est porté sur un composant doté d'une grille à oxyde épais (EOT : 3.4 nm) afin de s'affranchir de courant de grille présent dans les composants à oxyde de grille fin. Les longueurs de grilles L_1 et L_2 sont prises au minimum permis par la technologie.

3.3 Caractérisation du FD-SOI BIMOS en mode « gated-diode »

Le BIMOS contient deux structures de type « gated-diode » (diode surplombée d'une grille) au niveau des ensembles body-contact/body/grille-avant/source et body-contact/body/grille-avant/drain (Figure 3-3).

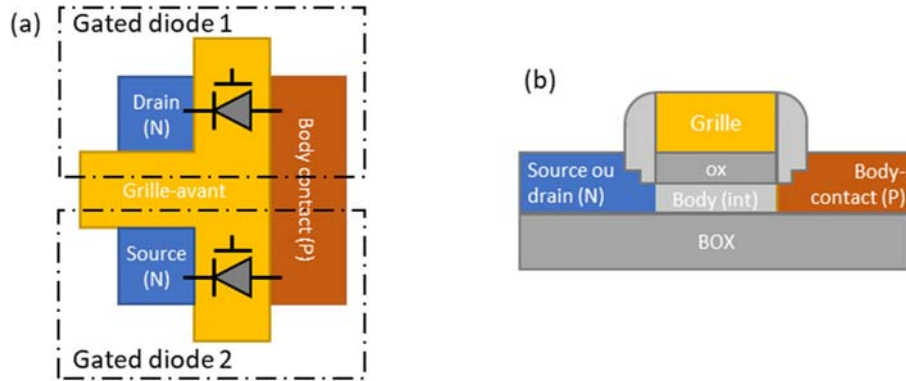


Figure 3-3 – Mise en évidence des structures « gated-diode » dans le schéma du FD-SOI BIMOS. (a) Vue de dessus et (b) vue en coupe.

Dans ce point de vue, les anodes (notées A) et les grilles des deux diodes sont communes. Leurs deux cathodes respectives (notées K) sont la source et le drain du BIMOS. Ces deux diodes sont équivalentes de par les symétries de la topologie et du procédé de fabrication (la source et le drain reçoivent les mêmes doses de dopants lors des étapes d'implantation).

En faisant l'hypothèse que leur contribution au courant est égale, elles peuvent être caractérisées en même temps, en connectant leurs cathodes (Figure 3-4). Ce type de connexion permet de désactiver les couplages qui pourraient apparaître entre les deux diodes lorsque leurs potentiels de cathode sont différents et de faciliter ainsi l'interprétation des résultats.

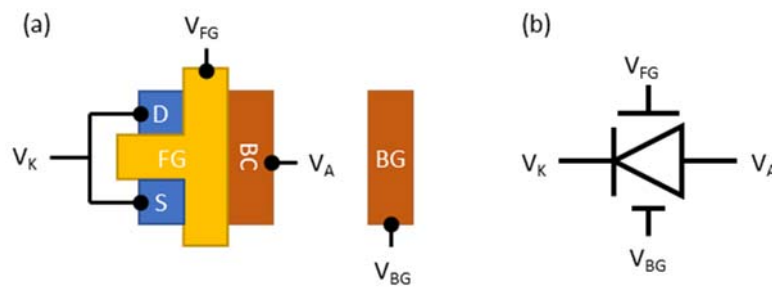


Figure 3-4 – (a) Vue de dessus du BIMOS connecté en mode gated diode et (b) schéma correspondant. A : anode, BC : body contact, BG : grille-arrière, D : drain, FG : grille-avant, K : cathode, S : source.

Dans ce qui suit, le courant d'anode I_A est d'abord caractérisé en fonction de la tension entre l'anode et la cathode V_{A-K} sous différentes polarisations de la grille-avant V_{FG} . Ensuite, I_A est caractérisé en fonction de V_{FG} à différentes tensions V_{A-K} . Enfin, parce que c'est la configuration dans laquelle se trouve le BIMOS lorsqu'il est utilisé dans le circuit de neurone développé dans cette thèse, I_A est caractérisé en fonction de V_{A-K} sous la condition $V_A = V_{FG}$. Pour toutes ces mesures, le potentiel de

cathode est pris comme référence, et la tension entre la grille-arrière et la cathode est nulle ($V_K = V_{BG} = 0$).

3.3.1 Courbe I_A - V_A avec V_{FG} comme paramètre

Dans cette section, le courant d'anode I_A en fonction de la tension entre l'anode et la cathode V_{A-K} est caractérisé par la mesure à température ambiante. L'influence de la tension de grille-avant V_{FG} sur le courant I_A est étudiée.

Pendant la mesure, la cathode et la grille-arrière sont branchées à la masse ($V_K = V_{BG} = 0$ V), le potentiel de la grille-avant est fixé à une valeur comprise entre -1 V et 1 V, et un balayage sur le potentiel d'anode est effectué, entre -2.5 V et 2.5 V. Le courant d'anode est tracé en échelle semi-logarithmique (Figure 3-5.a) et en échelle linéaire (Figure 3-5.b).

Comme attendu, la diode a deux comportements distincts selon qu'elle est polarisée en direct ($V_A > 0$) ou en inverse ($V_A < 0$) :

- Lorsque la diode est polarisée en direct ($V_A > 0$) :
 - Le courant d'anode I_A est positif et croît avec V_A . Sa croissance connaît deux régimes distincts : une croissance exponentielle pour $V_A < 1$ V et une croissance linéaire pour $V_A > 1.2$ V.
 - En régime linéaire, la tension de grille V_{FG} n'a pas d'effet significatif sur la caractéristique I_A - V_A (Figure 3-5.b). En particulier, la résistance différentielle $R_{PN} = \partial V_A / \partial I_A$ est indépendante de V_{FG} . Pour toutes les courbes, elle est minimale autour de $V_A = 1.6$ V pour une valeur de 650 Ω (soit une résistance de 1300 Ω pour chacune de deux diodes). La tension de transition $V_{T,PN}$ entre les deux régimes de fonctionnement de la diode est évaluée en utilisant la méthode du maximum de la dérivé de la conductance, inspirée de la méthode d'extraction de la tension de seuil $V_{T,MOS}$ par le calcul du maximum de la dérivée de la transconductance des MOSFET [14] :

$$V_{T,PN} = V_A \text{ tel que } \frac{\partial^2 I_A}{\partial V_A^2}(V_A) = \max \frac{\partial^2 I_A}{\partial V_A^2} \quad [V] \quad (3.1)$$

Pour s'affranchir du bruit de mesure lors de l'évaluation de $V_{T,PN}$, le maximum de $\partial^2 I_A / \partial V_A^2$ est obtenu par interpolation quadratique sur la plage $V_A \in [1 ; 1.3]$ V (Figure 3-6.a). $V_{T,PN}$ se situe autour de 1.165 V. Sa variation avec V_{FG} n'est pas monotone : $V_{T,PN}$ croît sur $V_{FG} \in [-1 ; 0.5]$ V et décroît ensuite. Son minimum est de 1.14 à $V_{FG} = -1$ V et son maximum est de 1.19 à $V_{FG} = 0.5$ V.

- En régime exponentiel, le courant d'anode I_A dépend fortement de V_{FG} pour $V_A < 0.9$ V (Figure 3-5.a). Cette dépendance s'accroît à mesure que la tension V_A se rapproche de 0 V. En particulier, le rapport $I_A(V_{FG}=1)/I_A(V_{FG}=-1)$ est maximal autour de $V_A = 0.4$ V pour une valeur de 10^3 .

Lorsqu'une diode est polarisée légèrement en direct ($0 < V_A \ll V_{T,PN}$) le courant de recombinaison I_{GR} domine le courant de diffusion I_{DIFF} de la diode : $I_A = I_R + I_{DIFF}$ avec $I_R \gg I_{DIFF}$. Du point de vue de la modélisation, cela se traduit par une augmentation du facteur d'idéalité η de la diode. Le facteur

d'idéalité η est lié à la pente de la courbe de la caractéristique I_A - V_A en échelle semi-logarithmique et peut-être extrait à partir de l'équation (2.35) :

$$\eta = \left[U_T \cdot \ln(10) \cdot \frac{\partial \log I_A}{\partial V_A} \right]^{-1} \quad (3.2)$$

où $U_T = kT/q$ est le potentiel thermique, avec k la constante de Boltzmann, T la température en Kelvin, et q la charge élémentaire.

Le phénomène de recombinaison connu pour contribuer majoritairement au courant I_R dans une diode standard polarisée en direct est la recombinaison SRH : $I_R = I_{SRH}$. Les modélisations compactes usuelles concluent à une dépendance de I_{SRH} en V_A de la forme $I_{SRH} \propto \exp[V_A/(2 \cdot U_T)]$, cela correspond à un facteur d'idéalité $\eta = 2$. Le courant de diffusion I_{DIFF} de la diode idéal est de la forme $I_{DIFF} \propto \exp[V_A/U_T]$, ce qui correspond à un facteur d'idéalité $\eta = 1$. Lors d'une mesure du courant $I_A = I_{DIFF} + I_R$, on s'attend à extraire un facteur d'idéalité compris entre ces deux valeurs : $1 < \eta < 2$. Or, les extractions faites sur les mesures (Figure 3-6.b) donnent $\eta > 2$, pour $V_{FG} \leq 0$ V et $V_A < 0.7$. Cela indique que d'autres mécanismes de génération/recombinaison que la recombinaison SRH sont en jeu.

Les deux autres mécanismes de génération/recombinaison connus sont l'effet tunnel assisté par piège (« trap-assisted tunneling » : TAT) et l'effet tunnel bande à bande (« band-to-band-tunneling » : BTBT). Cependant, ces mécanismes sont connus pour contribuer au courant de diode en polarisation inverse et non en polarisation directe. En effet, le TAT et le BTBT dépendent fortement de l'intensité du champ électrique E dans la ZCE de la diode. En particulier, l'expression du BTBT donnée à l'équation (2.36) est une évaluation en fonction de V_{A-K} de l'expression [15]:

$$I_{BTBT} = A \cdot V_{A-K} \cdot |\vec{E}_{MAX}|^2 \cdot \exp\left(-\frac{B}{|\vec{E}_{MAX}|}\right) \quad [A] \quad (3.3)$$

Où A et B sont deux paramètres du modèle et E_{MAX} est la valeur maximale du champ électrique dans la ZCE, qui se situe en son milieu. Dans une diode classique, $\vec{E}$ est exclusivement longitudinal ($\vec{E} = \vec{E}_{||}$) et il se réduit à mesure que la polarisation en direct augmente. De ce point de vue, ni le TAT, ni le BTBT ne semblent être en mesure d'expliquer $\eta > 2$. Or, dans une gated-diode, du fait de la présence de la grille, la contribution transverse $\vec{E}_{\perp}$ peut être conséquente et doit également être prise en compte ($|\vec{E}|^2 = |\vec{E}_{||}|^2 + |\vec{E}_{\perp}|^2$). La situation $|\vec{E}_{\perp}| \gg |\vec{E}_{||}|$ est d'ailleurs souvent réalisée. En outre, dans une structure FD-SOI, la valeur maximale de $|\vec{E}_{\perp}|$ ($|\vec{E}_{\perp}|_{MAX}$) dépend de la tension entre chacune des grilles et chacun des terminaux où circule le courant (tension grille-avant – anode V_{FG-A} , grille-avant – cathode V_{FG-K} , grille-arrière – anode V_{BG-A} , et grille-arrière – cathode V_{BG-K}). Ainsi, la contribution à I_A de I_{BTBT} généré par $\vec{E}_{\perp}$ pourrait expliquer la forte modulation de I_A par V_{FG} et le fait que $\eta > 2$. Cependant, une autre hypothèse, impliquant une modulation du mécanisme SRH sera développée à l'occasion de la section 3.3.2.

- Lorsque la diode est polarisée en inverse ($V_A < 0$) :
 - Le courant d'anode I_A est négatif et dépend fortement de V_{FG} .
 - Le courant d'anode I_A ne sature pas et croît avec $|V_A|$.

L'absence de saturation du courant témoigne de la présence de courants de génération I_G . Les phénomènes physiques connus pour leur contribution au courant I_G dans une diode polarisée en inverse sont la génération SRH, le TAT et le BTBT : $I_G = I_{SRH} + I_{TAT} + I_{BTBT}$ [16]. La forte dépendance de I_A en V_{FG} tend à montrer que, tout comme I_{RG} en polarisation directe, I_{RG} en polarisation inverse est majoritairement composé de I_{BTBT} . En effet, lorsque l'on trace le courant I_A en fonction de la tension

V_{FG-A} pour $V_A < 0$, on constate que les courbes obtenues à différents V_{FG} se regroupent (Figure 3-7). Cette observation suggère que, d'une part, le BTBT est le processus dominant dans le courant génération de la gated-diode polarisée en inverse : $I_A = I_{RG} = I_{BTBT}$ et que, d'autre part, le BTBT est localisé dans le film de silicium à la jonction body – anode (c'est en effet là que le champ électrique est maximal dans la configuration de polarisation étudiée). Cependant, le fait que toutes les courbes I_A - V_{FG-A} ne se chevauchent pas exactement témoigne du fait que le phénomène de génération impliqué ne dépend pas uniquement de la tension V_{FG-A} . Cela est cohérent avec le fait que, dans les structures FD-SOI, le profil de champ électrique dans le film déplété a également une dépendance importante vis-à-vis du potentiel de grille-arrière V_{BG} . Pour mettre en évidence cette dépendance, il serait intéressant d'effectuer de nouvelles caractérisations en prenant V_{BG} comme paramètre.

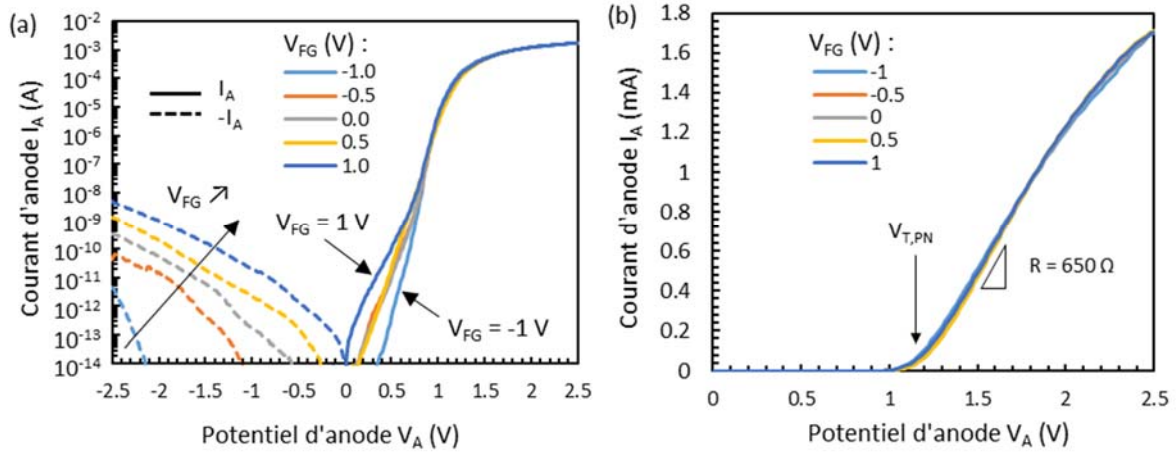


Figure 3-5 – Mesure, sur un BIMOS connecté en mode gated-diode, du courant d'anode I_A en fonction de la tension d'anode V_A avec la tension de grille-avant V_{FG} comme paramètre, en échelle (a) semi-logarithmique et (b) linéaire. $V_{BG} = V_K = 0$ V.

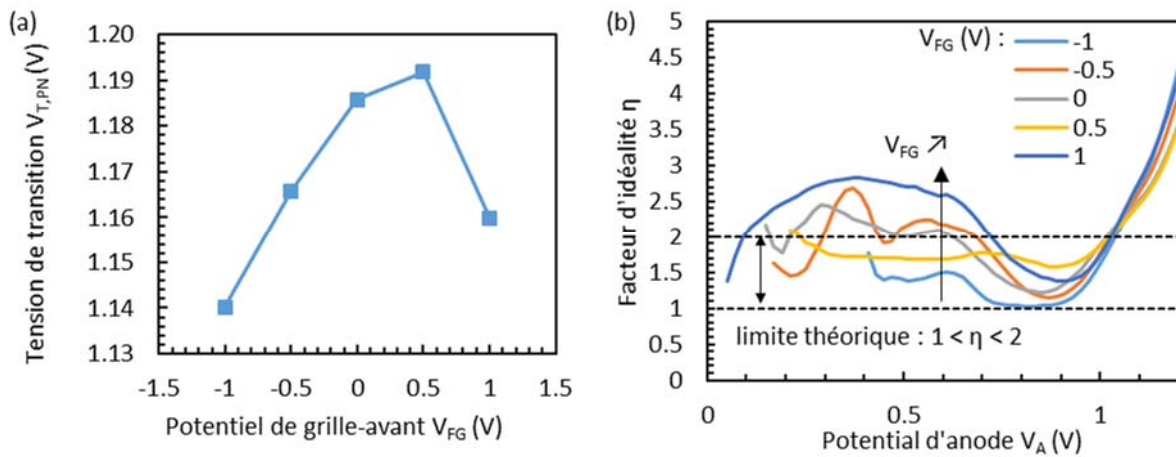


Figure 3-6 – (a) Tension de transition $V_{T,PN}$ en fonction du potentiel de grille-avant V_{FG} . (b) Facteur d'idéalité η en fonction de la tension V_{A-K} pour différentes valeurs de potentiel de grille-avant V_{FG} .

Pour résumer, le BIMOS connecté en gated-diode a bien les caractéristiques globales d'une diode. En polarisation directe, le courant I_A est positif. Il croît exponentiellement puis linéairement avec V_{A-K} . En croissance linéaire, la résistance différentielle R_{PN} vaut environ 650Ω et ne dépend pas de V_{FG} . La tension de transition $V_{T,PN}$ est autour de 1.165 V et est faiblement dépendante de V_{FG} . En régime exponentiel ($V_A < V_{T,PN}$), le courant I_A est fortement modulé par V_{FG} et conduit dans certains cas à un coefficient d'idéalité $\eta > 2$. Un courant de génération I_{BTBT} pourrait en être responsable. En polarisation inverse, I_A est négatif et dépend fortement de V_{FG} . Il ne sature pas lorsque $|V_A|$ augmente. Le courant

de génération I_{BTBT} généré à la frontière anode – body domine dans ce régime. Notons que toutes ces observations sont cohérentes avec les études effectuées dernièrement sur une autre structure de type FD-SOI gated-diode appelée diode virtuelle [17].

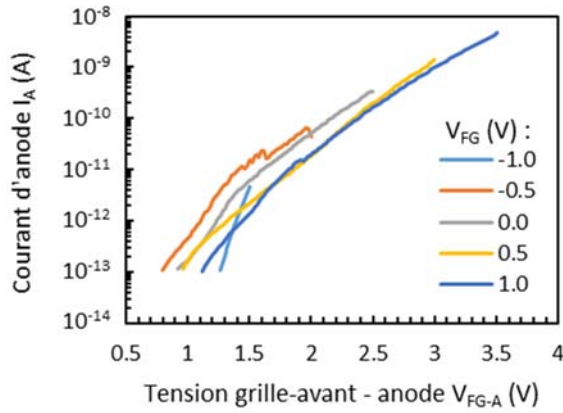


Figure 3-7 – Courant d’anode I_A en fonction de la tension grille-avant – anode V_{FG-A} pour différentes valeurs de potentiel de grille V_{FG} . $V_{BG} = V_K = 0$ V.

3.3.2 Courbes I_A - V_{FG} avec V_{A-K} comme paramètre

La section précédente a montré que le potentiel de grille V_{FG} module le courant d’anode I_A lorsque que la gated-diode est polarisée légèrement en direct. Ce phénomène a déjà été observé dans la littérature, mais sans que cela conduise à un facteur d’idéalité η supérieur à 2 [18]–[20]. Il est expliqué par la modulation par V_{FG} du taux de recombinaison SRH dans le volume du body et à ses interfaces avec le BOX et l’oxyde de grille.

De façon plus détaillée, V_{A-K} contrôle l’injection de porteur dans le canal, selon (2.9) :

$$n \cdot p = n_i^2 \cdot \exp\left(\frac{V_{A-K}}{U_T}\right) \quad [m^{-6}] \quad (3.4)$$

tandis que V_{FG} (ainsi que V_{BG} dans les structures FD-SOI), de par son influence sur le potentiel dans le body, contrôle la répartition des porteurs entre les électrons n et les trous p (Figure 3-8). Or, l’équation (2.25) indique que ce sont ces densités n et p qui interviennent dans la recombinaison SRH.

Il est possible de simplifier l’expression de R_{SRH} à l’équation (2.22) en admettant que les durées de vie τ sont les mêmes pour les trous et pour les électrons ($\tau_n = \tau_p = \tau$) et que l’énergie des centres de recombinaison sont au centre du gap ($E_t = E_i$) :

$$R_{SRH} = \frac{2}{\tau} \cdot \frac{n \cdot p \cdot n_i^2}{n + p + 2 \cdot n_i} \quad [m^{-3} \cdot s^{-1}] \quad (3.5)$$

Notons que les simplifications apportées pour établir (3.5) maximisent la valeur de R_{SRH} . Cette expression peut également être adaptée au cas d’une recombinaison surfacique $R_{S,SRH}$, en introduisant une vitesse de recombinaison aux interfaces σ , considérée égale pour les électrons et les trous [19]:

$$R_{S,SRH} = 2\sigma \cdot \frac{n \cdot p \cdot n_i^2}{n + p + 2 \cdot n_i} [m^{-2} \cdot s^{-1}] \quad (3.6)$$

Comme σ et $1/\tau$ sont respectivement proportionnels aux densités de centre de recombinaisons dans le volume N_t et sur les interfaces $N_{S,t}$, R_{SRH} et $R_{S,SRH}$ sont d’autant plus grands que les densités de pièges sont importantes. De plus, (3.5) et (3.6) montrent que R_{SRH} et $R_{S,SRH}$ sont maximaux pour $n = p$ (qui valent alors $n_i \cdot \exp[V_{A-K}/(2 \cdot U_T)]$).

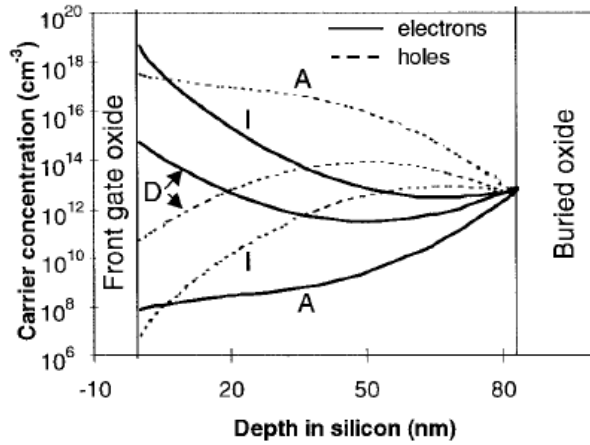


Figure 3-8 – Simulation de la concentration verticale d'électrons et de trous dans le body d'une gated-diode sur SOI ($W = L = 3\mu\text{m}$, $V_{A-K} = 0.3\text{ V}$, densité de dopants du body : $N_A = 10^{17}\text{ cm}^{-3}$) lorsque $n=p$ à l'interface body/BOX et pour une interface body/oxyde de grille déplétée (D), accumulée (A) ou inversée (I), d'après [18]. Ces différentes courbes sont obtenues en jouant sur les potentiels des grilles avant et arrière dont les valeurs ne sont pas explicitées.

Le calcul de I_{SRH} s'obtient en intégrant R_{SRH} dans le volume du body et $R_{S,SRH}$ sur les interfaces body/BOX et body/oxyde de grille.

La Figure 3-8 montre comment le point $n = p$ où la recombinaison est maximale se déplace dans le body en fonction des polarisations des grilles. Autrement dit, les pièges présents dans le body et aux interfaces sont successivement activés à mesure que le point $n = p$ se déplace. De cette façon, en évaluant par simulation les profils de densité de porteur, la mesure de I_{SRH} permet de construire une carte des densités N_t et $N_{S,t}$ dans le body.

A la manière des expériences menées dans [18]–[20], des mesures à température ambiante de I_A en fonction de V_{FG} sous différentes tensions V_{A-K} sont effectuées sur le BIMOS de référence connecté en gated-diode (Figure 3-9.a). Le potentiel de grille-arrière est pris comme référence. Pour conserver une symétrie du potentiel de body entre les différentes mesures, l'anode et la cathode sont polarisées à $V_{A-K}/2$ et $-V_{A-K}/2$ respectivement (Figure 3-9.b).

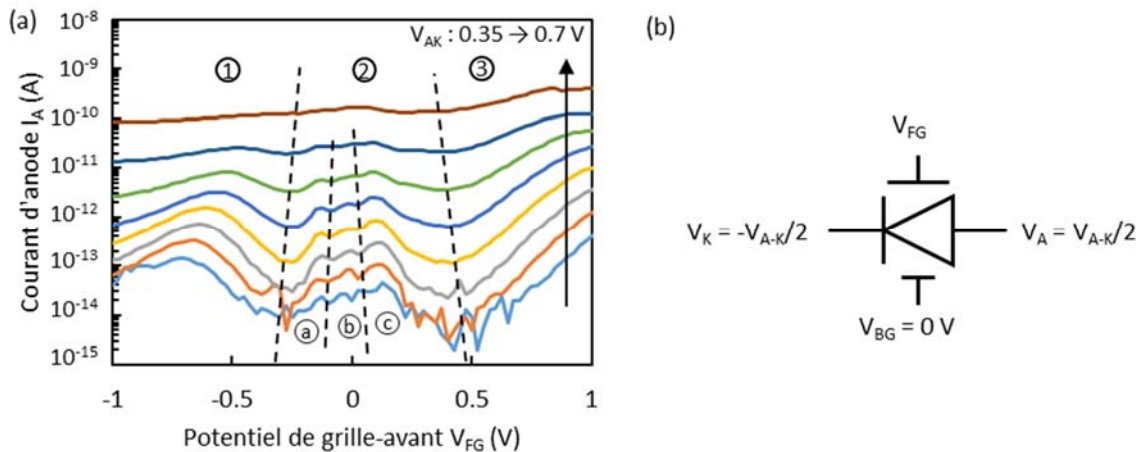


Figure 3-9 – (a) Mesures à température ambiante du courant d'anode I_A en fonction du potentiel de grille-avant V_{FG} sous différentes tensions d'anode-cathode V_{A-K} dans le BIMOS de référence connecté en mode gated-diode et en régime de polarisation légèrement en direct ($0.35 < V_{A-K} < 0.7\text{ V}$). Le potentiel de grille-arrière est pris comme référence. L'anode et la cathode sont polarisées à $\pm V_{A-K}/2$ respectivement. (b) Schéma électrique correspondant.

La modulation du courant I_A par V_{FG} apparaît clairement pour $V_{A-K} < 0.6\text{ V}$ et s'estompe lorsque V_{A-K} atteint 0.7 V . A cette tension, la forte injection de porteur dans la diode réduit la contribution de I_{SRH} au courant total au profit du courant de diffusion I_{DIFF} de la diode idéale. Pour $V_{A-K} < 0.6\text{ V}$, I_A connaît trois pics successifs sur les plages de polarisation de V_{FG} repérées par les zones 1, 2 et 3 de la Figure 3-9.a. Le pic central est lui-même composé de trois pics distincts repérés par les zones a, b, et c. Le

nombre de ces pics est important en comparaison aux nombre de pic reportés dans [19]. Cela témoigne de la complexité du mécanisme SRH dans la structure UTBB FD-SOI BIMOS étudié.

Une analyse plus profonde qui s'appuierait sur des simulations des profils de concentration de porteurs dans le body soumis aux tensions appliquées permettrait d'attribuer chacun de ces pics à des zones du body où le taux de recombinaison contribue majoritairement à I_{SRH} .

Quoi qu'il en soit, l'absence de description qualitative de la modulation de I_A par V_{FG} ne permet pas de proposer de modélisation analytique de ce phénomène. Ce point est donc ignoré pour la modélisation proposée à la section 3.6 au détriment de sa précision.

3.3.3 Courbes I_A - V_A avec $V_{FG} = V_A$

Dans la suite de ce document, le Chapitre 4 montre comment utiliser le BIMOS pour implémenter le circuit de déclenchement au cœur du circuit de neurone développé dans cette thèse. Dans le circuit de déclenchement, le body-contact et la grille-avant du BIMOS sont connectés : les deux gated-diodes ont leurs anodes connectées avec leurs grilles. Dans ce contexte, il est pertinent de caractériser ces diodes sous la condition $V_A = V_{FG}$ et de proposer un modèle compact correspondant qui permettra d'effectuer des simulations circuits.

Par ailleurs, le choix est fait de ne caractériser que la gated-diode body-contact – source, en laissant l'électrode de drain flottante pendant la mesure. Cette configuration correspond davantage à la situation rencontrée lorsque le BIMOS est utilisé dans le circuit de déclenchement.

Enfin, seule la plage de tension $V_{A-K} > 0$ est étudiée, car, comme nous le verrons à la section 3.6.2.1, le courant I_{BTBT} qui constitue le courant I_A sur la plage $V_{A-K} < 0$ est déjà pris en charge par la modélisation de type MOSFET.

Une mesure du courant d'anode I_A en fonction de la tension d'anode $V_{A-K} > 0$ sous les conditions $V_A = V_{FG}$ et $V_K = V_{BG} = 0$ V est effectuée sur le BIMOS de référence et à température ambiante (Figure 3-10.a et Figure 3-10.b). La dérivée de la conductance est calculée (Figure 3-10.c, axe de droite) et la tension de transition $V_{T,PN} = 1.108$ est extraite d'après l'équation (3.1). Le facteur d'idéalité η est calculé en fonction de V_A (Figure 3-10.c, axe de gauche) d'après l'équation (3.2). Pour s'affranchir du bruit de mesure, chaque étape de dérivation lors de calculs de η et $\partial^2 I_A / \partial V_{A-K}^2$ est évaluée en effectuant une interpolation linéaire sur les cinq points de mesure qui entourent l'abscisse V_A .

Notons ici que, comme pour les mesures effectuées à la section 3.3.1, η est supérieur à 2 sur une partie de la plage de croissance exponentielle $V_{A-K} < V_{T,PN}$. En particulier, ici, $\eta > 2$ pour $V_{A-K} < 0.5$ V.

Le courant I_A en fonction de V_{A-K} est modélisé par l'équation (2.35). Comme le η extrait de la mesure varie beaucoup, il n'est pas possible de trouver un jeu de paramètre [I_s ; η ; R_s] qui fasse correspondre la courbe mesurée et la courbe modélisée sur toute la plage de V_{A-K} . Une solution consiste à placer dans le modèle deux diodes en parallèle, l'une dotée d'un $\eta = 3.1$ proche de celui mesuré à faible V_A et d'un $R_s = 1$ M Ω pour réduire sa contribution à V_A fort, l'autre dotée d'un $\eta = 1.2$ proche de celui à V_A fort et d'un $R_s = 1200$ Ω égal à la résistance différentielle extraite du régime linéaire de la mesure (Figure 3-10.b). Notons qu'une telle modélisation « ad-hoc » qui implique deux diodes en parallèle ne repose pas sur une description physique du comportement du composant.

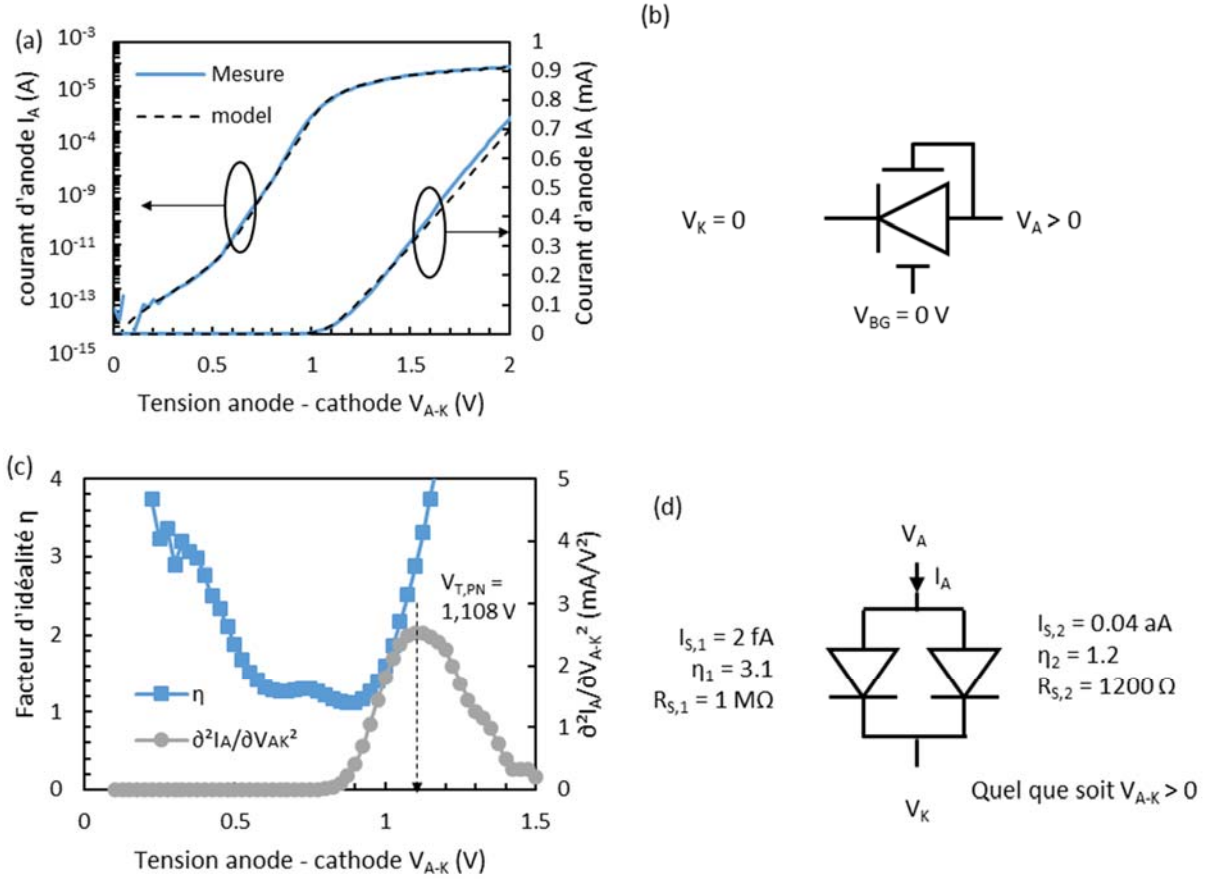


Figure 3-10 – (a) Mesure effectuée sur le BIMOS de référence et à température ambiante (ligne pleine) et simulation selon le modèle présenté à la figure (d) (ligne en pointillé) du courant d'anode I_A en fonction de la tension d'anode V_{A-K} sous la condition $V_A = V_{FG}$, en échelle semi-logarithmique (axe de gauche) et linéaire (axe de droite). $V_K = V_{BG} = 0$ V. (b) Schéma électrique correspondant. (c) Facteur d'idéalité η (axe de gauche), dérivée de la conductance $\partial^2 I_A / \partial V_{A-K}^2$ (axe de droite) et tension de transition $V_{T,PN}$ extraits de la mesure présentée à la figure (a). (d) Schéma électrique et paramètres du modèle utilisé à la figure (a). Les deux diodes sont décrites par l'équation (2.35).

3.4 Caractérisation du FD-SOI BIMOS en mode MOSFET

Dans les sections précédentes, le BIMOS a été étudié dans son mode gated-diode. Le but des sections qui suivent est de caractériser le FD-SOI BIMOS en fonctionnement de type MOSFET. En particulier, il s'agit d'évaluer l'impact de l'ajout et de la polarisation du body-contact sur les courbes caractéristiques courant-tension du MOSFET I_D - V_{FG} et I_D - V_D . Nous verrons que sur une plage de tension body-contact – source V_{B-S} suffisamment petite, la présence du body-contact ne perturbe pas le fonctionnement MOSFET du BIMOS.

Notons par ailleurs que les MOSFET pourvu de body-contact « (body-contacted MOSFET » : BC-MOSFET) ont été largement caractérisés et modélisés dans la littérature dans le cas d'intégration sur substrat PD-SOI. En effet, comme évoqué à la section 2.3.2.1, sur ce type de substrat notamment utilisé en radio fréquence, les effets de body sont importants et ne peuvent pas être négligés. En revanche, comme le FD-SOI n'est pas sujet à ces effets, peu d'effort ont été fourni par les communautés scientifique et industrielle pour étudier le BC-MOSFET sur ce type de substrat.

3.4.1 Courbe I_D-V_{FG}

Dans cette section, la courbe caractéristique I_D-V_{FG} du BIMOS de référence est mesurée et discutée selon 3 configurations distinctes : lorsque le potentiel de body-contact est nulle, lorsqu'il est flottant, et lorsqu'il varie positivement.

3.4.1.1 Courbe I_D-V_{FG} à $V_B = 0$ V

Dans un premier temps, nous comparons la caractéristique I_D-V_{FG} du BIMOS de référence et celle d'un NMOS standard équivalent : tous deux sont de type N, GO2 (oxyde de grille épais) et RVT (well de type P) et aux dimensions comparables (Tableau 3-2 et Figure 3-11). Dans l'expérience, la grille-arrière et la source des composants sont mises à la masse. Pour le BIMOS, la body-contact l'est également. Le drain est polarisé successivement à 0.2 V et 1.8 V, afin d'évaluer les réponses des composants en régime linéaire et saturé, respectivement. La tension de grille-avant est balayée entre -1 et 2 V (Figure 3-12).

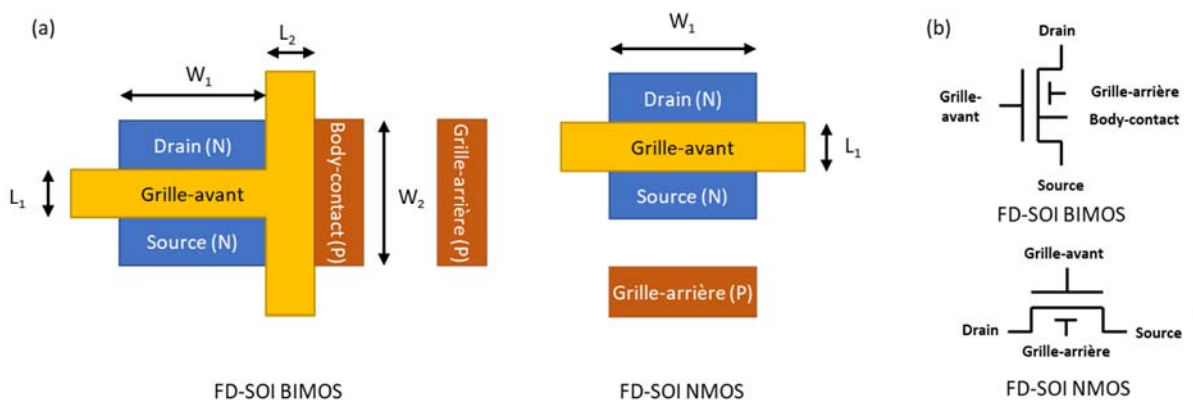


Figure 3-11 – (a) Vue de-dessus du BIMOS et du NMOS et longueurs évoquées au Tableau 3-2. (b) Schéma électrique correspondant.

Tableau 3-2 – Caractéristiques physiques du BIMOS et du NMOS.

Composant	Type	Saveur	EOT (nm)	L_1 (nm)	L_2 (nm)	W_1 (nm)	W_2 (nm)
BIMOS de référence	N	RVT	3.4	135	144	270	350
NMOS comparable	N	RVT	3.4	135	-	270	-

Pour le BIMOS, les données sont récoltées en effectuant une mesure sous pointe à température ambiante. Afin d'observer l'influence de la présence du body contact sur la circulation des courants, les courants de source (Figure 3-12.c) et de body (Figure 3-12.d) sont tracés en plus du courant de drain (Figure 3-12.a et Figure 3-12.b). Les courants de grille-avant et arrière, s'ils existent, sont trop faibles pour être mesurés et sont considérés comme nuls. Pour le NMOS, les données sont issues d'une simulation SPICE qui utilise le modèle compact UTSOI 2.1.1 [21]–[23] et les cartes modèles fournies par STMicroelectronics. Seul le courant de drain est tracé. En effet, comme les grilles à oxyde épais de type GO2 sont dépourvues de courant de fuite sur la plage de tension étudiée, le courant de drain du NMOS est égal en valeur absolue au courant de source. Il contient donc à lui seul toutes les informations sur les courants qui circulent dans le composant. Les courbes de courant de drain du NMOS sont tracées au côté des courbes du courant de drain du BIMOS (Figure 3-12.b et Figure 3-12.b).

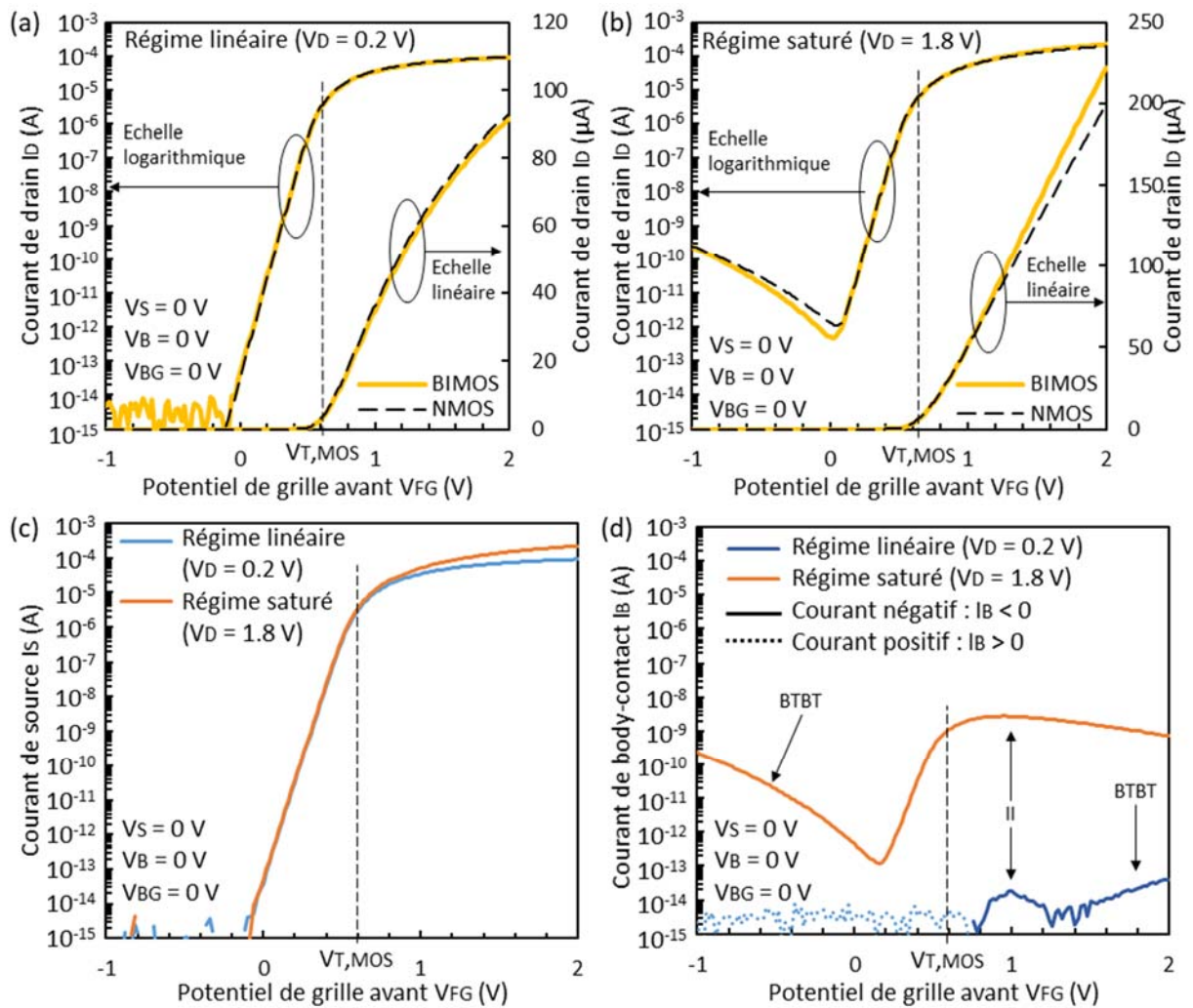


Figure 3-12 – (a et b) Courants de drain I_D circulant dans le BIMOS de référence et dans un NMOS de type et de dimensions comparables en fonction du potentiel de la grille-avant V_{FG} lors d'un fonctionnement des composant en régime (a) linéaire ($V_D = 0.2$ V) et (b) saturé ($V_D = 1.8$ V). (c) Courant de source et (d) courant de body-contact du BIMOS de référence en fonction de la grille-avant V_{FG} en régime linéaire et saturé. Les données qui concernent le BIMOS sont issues de mesures à température ambiante et celles qui concernent le NMOS sont issues de simulations SPICE calibrées. Dans les deux expériences, $V_S = V_{BG} = V_B = 0$ V.

En régime linéaire ($V_D = 0.2$ V, Figure 3-12.a et courbes bleues des Figure 3-12.b et Figure 3-12.c) :

- Le courant de drain du BIMOS est superposé au courant de drain du NMOS (Figure 3-12.a). La présence du body-contact mis à la masse dans le BIMOS n'affecte donc pas son comportement MOSFET.
- Le courant de drain et de source du BIMOS sont opposés et égaux en valeur absolue (Figure 3-12.a et courbe bleue de la Figure 3-12.c). Le courant de drain circule donc entièrement vers la source.
- Le courant de body est nul pour $V_{FG} < 0.8$ V, et très faible pour $V_{FG} > 0.8$ V ($|I_B| < 100$ fA) (courbe bleue de la Figure 3-12.d). Il circule vers l'extérieur du composant ($I_B < 0$). La forme de sa courbe se décompose en deux régions. Dans la première ($0.8 < V_{FG} < 1.26$ V), la courbe est en forme de cloche, et dans la seconde ($V_{FG} > 1.26$ V), elle suit une croissance exponentielle. Ces deux formes respectives sont caractéristiques de courants de trous générés dans le body par ionisation par impact (« impact ionization » : II) pour la première région et par BTBT pour la seconde. Notons que le courant d'II est très faible à cause de la faible valeur de V_D . Notons également que la génération par BTBT (pour $V_{FG} > 1.26$ V) doit avoir lieu à la jonction body-

contact – body où, étant donné l'état de polarisation, la configuration du diagramme de bande est la plus propice à la présence de ce phénomène.

En régime saturé ($V_D = 1.8 \text{ V}$, Figure 3-12.b et courbes oranges des Figure 3-12.b et Figure 3-12.c) :

- Pour $0.1 \text{ V} < V_{FG} < V_{T,MOS}$:
 - Les courants de source et de drain du BIMOS sont égaux en valeur absolue au courant de drain du NMOS. Ainsi, sur cette plage, la présence du body-contact n'affecte pas le comportement MOSFET du BIMOS.
 - Le courant sortant du body-contact du BIMOS ($I_B < 0$) est un courant de génération I_{GEN} . Il est constitué de trous issus de la génération de paires électron/trou par II : $I_{GEN} = I_{II}$. Les électrons générés rejoignent le drain, mais leur contribution au courant I_D total est difficilement détectable car le courant de canal I_{CH} qui constitue majoritairement I_D est trop important ($|I_B|/(|I_D| - |I_B|) < 1 \%$ pour $V_{FG} > 0.16 \text{ V}$). Les trous générés rejoignent le body-contact plutôt que la source. Cette préférence est due à la présence d'une barrière de potentiel interne en direction de la source (induite par les dopants de type N de la source), et d'un puit de potentiel interne en direction du body-contact (induit par les dopants de type P du body-contact). Notons que cette génération existe aussi dans le NMOS : dans ce cas, les électrons générés rejoignent le drain et les trous rejoignent la source. Comme le courant de génération I_{GEN} s'ajoute au courant de canal I_{CH} qui circule dans les deux terminaux, le phénomène est difficilement détectable pour la même raison que celle avancée dans le cas du courant d'électrons dans le drain du BIMOS.
- Pour $V_{FG} > V_{T,MOS}$:
 - Un décalage apparaît entre les courants de drain du BIMOS et du NMOS. Il conduit à la relation $I_{D,BIMOS} > I_{D,NMOS}$. Ce décalage sera discuté à la section 3.4.2, où les courbes caractéristiques I_D - V_D sont investiguées.
 - Le courant $I_{GEN} = I_{II}$ est toujours présent au niveau du body-contact du BIMOS. Il décroît avec V_{FG} . La forme en cloche de la courbe du courant de body-contact sur la plage $V_{FG} > 0.2 \text{ V}$ est caractéristique du courant d'II lors d'une mesure I_D - V_{FG} [24], [25]. Elle est discutée à la section 3.6 qui traite de la modélisation des courants dans le BIMOS.
- Pour $V_{FG} < 0.1 \text{ V}$
 - Les courants de drain du BIMOS et du NMOS sont également décalés. Néanmoins, la forme de leurs courbes permet de les identifier tous deux comme des courants de fuite induits par la grille (« gate induced drain leakage » : GIDL). Le GIDL est issu de la génération de paires électron-trou par BTBT au voisinage de la jonction drain-body, sous l'effet du champ électrique intense produit par la tension drain – grille-avant V_D - V_{FG} . [26]–[31]. De la même façon que l' I_{II} , le GIDL circule du drain vers la source dans le NMOS, et du drain vers le body-contact dans le BIMOS. Comme pour l' I_{II} , Cette différence est due, d'une part, à la nature des porteurs impliqués dans le courant et, d'autre part, à la différence de dopage des terminaux. Ainsi, pour $V_{FG} < 0.1 \text{ V}$, le courant de body-contact I_B est un courant de génération I_{GEN} engendré par BTBT I_{BTBT} : $I_B = I_{GEN} = I_{BTBT}$ (partie gauche de la courbe orange de la Figure 3-12.c).
 - Concernant le décalage entre les courbes I_D du NMOS et du BIMOS, deux hypothèses sont émises. La première attribue le décalage à l'incertitude sur les prédictions du modèle ou à la variabilité du procédé de fabrication. La seconde affirme que, dans le BIMOS, la dérivation du flux de trous depuis la source vers le body-contact empêche l'activation d'un effet bipolaire qui, dans le NMOS ajoute sa contribution au courant total.

Pour résumer, en comparaison à un FD-SOI NMOS de dimension équivalente, la présence du body-contact dans le FD-SOI BIMOS et sa mise à la masse n'influencent pas le courant de drain de la caractéristique I_D - V_{FG} . La particularité du FD-SOI BIMOS vis-à-vis de FD-SOI NMOS réside dans le fait que son body-contact, du fait de son dopage de type P, collecte les trous générés dans la structure par II ou BTBT, comme le fait le substrat dans les NMOS en technologie bulk, tandis que dans le FD-SOI NMOS, ces trous rejoignent la source. Cette analyse permet d'écrire, dans le BIMOS, $I_B = I_{GEN} = I_{II} + I_{BTBT}$

3.4.1.2 Courbe I_D - V_{FG} lorsque le body-contact est flottant

La section précédente a montré que, dans le BIMOS et lors d'une caractérisation de type I_D - V_{FG} , la présence du body-contact mis à la masse ne modifie quasiment pas les caractéristiques du courant de drain. Mais qu'en est-il lorsque le body-contact est laissé flottant ?

Pour répondre à cette question, l'effet d'un body-contact flottant sur la mesure I_D - V_{FG} du BIMOS en saturation ($V_D = 1.8$ V) est investiguée. Elle est comparée à la courbe I_D - V_{FG} du BIMOS lorsque la base est à la masse ($V_B = 0$) et à la courbe I_D - V_{FG} du NMOS de dimensions comparables soumis aux mêmes polarisations, obtenue par simulation SPICE.

Du point de vue de la mesure, deux méthodes sont appliquées pour obtenir un body-contact flottant. La première consiste à ne pas connecter de pointe sur le plot relié au body-contact (Figure 3-13.a). Dans la seconde, une pointe est connectée au relié au body contact et la condition $I_B = 0$ A lui est imposée (Figure 3-13.b). Cette dernière méthode permet de mesurer le potentiel de body-contact nécessaire pour maintenir son courant nul.

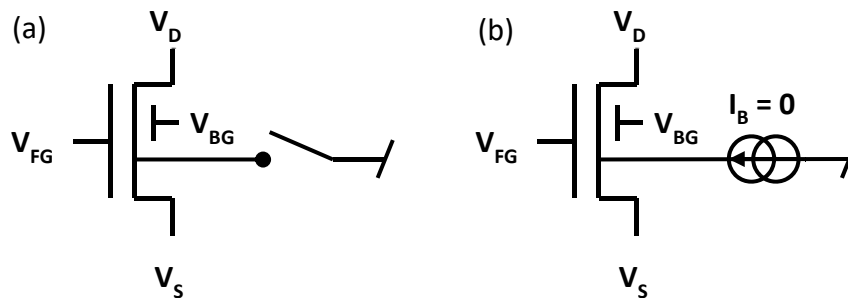


Figure 3-13 – Schémas électrique qui permettent d'obtenir un body-contact flottant : (a) sans connecter le body contact (b) en contraignant le courant $I_B = 0$ à l'aide d'une source de courant. Cela permet de mesurer le potentiel.

Lorsque l'on compare les quatre courbes de courant de drain (Figure 3-14.a à Figure 3-14.c) :

- En régime sous le seuil ($V_{FG} < V_{T,MOS}$) et $V_{FG} > 0.15$ V, toutes les courbes sont superposées (Figure 3-14.a).
- Au-dessus du seuil (Figure 3-14.a et Figure 3-14.b), seules les courbes issues des mesures sur le BIMOS sont superposées. La courbe issue des simulations sur le NMOS est décalée vers le bas ($I_{D,NMOS} < I_{D,BIMOS}$). Ce décalage sera discuté à la section 3.4.2. qui traite des mesures de type I_D - V_D .
- En régime très profondément sous le seuil ($V_{FG} < 0$), les courbes ont la même forme générale mais sont toutes décalées (Figure 3-14.a et Figure 3-14.b). Le décalage est maximal entre $I_D(V_B=0)$ et $I_D(I_B=0)$ à $V_{FG} = 0.05$ V avec un ratio $I_D(I_B=0)/I_D(V_B=0) = 3$. Le décalage entre $I_D(I_B=0)$ et $I_D(\text{floating body})$ pour V_{FG} autour de 0 V montre que les deux méthodes de mesure ne sont pas exactement équivalentes à très faible courant.

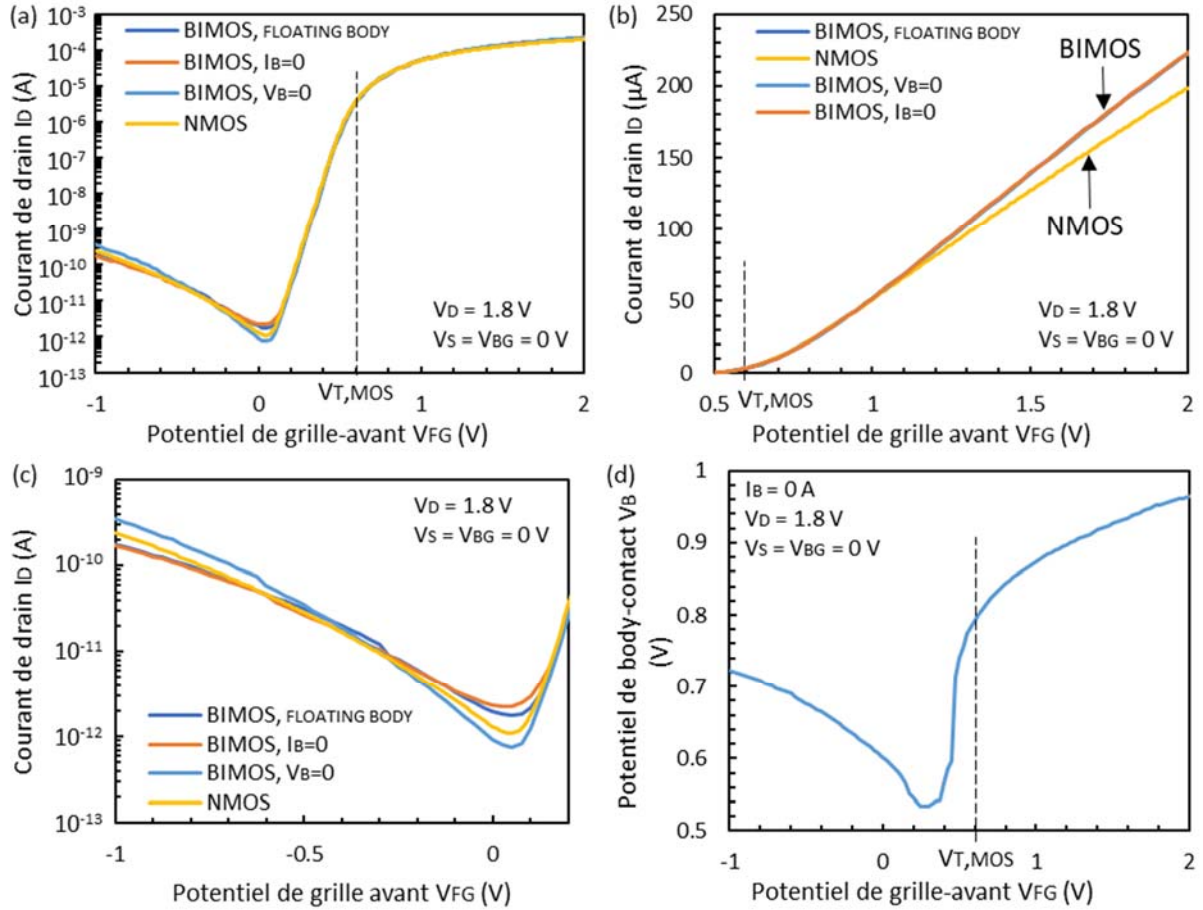


Figure 3-14 – (a) Courants de drain du BIMOS de référence en fonction du potentiel de grille-avant V_{FG} , en régime saturé ($V_D = 1.8$ V), sous les conditions $V_B = 0$, $I_B = 0$ et body-contact flottant, et courant de drain du NMOS de dimensions comparables et sous les mêmes polarisations. (b) Courbes de (a) en échelle linéaire pour $V_{FG} > 0.5$ V. (c) Agrandissement de (a) pour $V_{FG} < 0.1$ V. (d) Potentiel de body-contact du BIMOS de référence en fonction du potentiel de grille-avant V_{FG} , en régime saturé ($V_D = 1.8$ V) et sous la condition $I_B = 0$.

Lorsque la condition $I_B = 0$ A est imposée, le potentiel de body-contact varie entre $V_B = 0.5$ V et $V_B = 1$ V (Figure 3-14.d). Il décroît de 0.73 V à 0.54 V sur la plage $V_{FG} \in [-1 \text{ V} ; 0.25 \text{ V}]$, et croît de 0.54 V à 0.96 V sur la plage $V_{FG} \in [0.25 \text{ V} ; 1 \text{ V}]$. Une première façon d'interpréter ces variations consiste à considérer que V_B est le potentiel pour lequel il est plus favorable pour les trous générés dans le body de rejoindre la source plutôt que le body-contact. Une seconde façon d'interpréter ces variations consiste à introduire le courant I_{PN} de la « gated-diode » (diode surplombée d'une grille) constitué par le body-contact, le body, la grille et la source. Le potentiel de body-contact mesuré est alors celui pour lequel I_{PN} est égal et opposé au courant de génération I_{GEN} qui circule vers le body contact : $I_B = I_{PN} + I_{GEN} = 0$. Cette équation est utilisée pour l'établissement du modèle compact décrit à la section 3.6.

3.4.1.3 Influence de V_B sur les courbes I_D - V_{FG}

Les deux sections précédentes ont montré que la mise à la masse du body-contact ou son maintien en nœud électrique flottant ont un effet extrêmement limité sur la caractéristique I_D - V_{FG} du BIMOS. Pour compléter l'analyse, l'influence de la polarisation positive du body-contact sur la caractéristique I_D - V_{FG} est investiguée dans cette section. Nous verrons que sur la plage explorée $0 \text{ V} \leq V_B \leq 1 \text{ V}$, la polarisation de V_B ajoute à I_B le courant de gated-diode body-source et body-drain caractérisé à la section 3.3, sans que cela modifie la tension de seuil MOSFET du BIMOS.

Les courants de drain I_D , de source I_S et de body-contact I_B du BIMOS sont caractérisés par la mesure à température ambiante en fonction de la polarisation de la grille-avant V_{FG} et sous une polarisation du body-contact variant de 0 V à 1 V, par pas de 0.25 V, dans un régime de fonctionnement linéaire ($V_D = 0.2$ V, Figure 3-15) et saturé ($V_D = 1.8$ V, Figure 3-16).

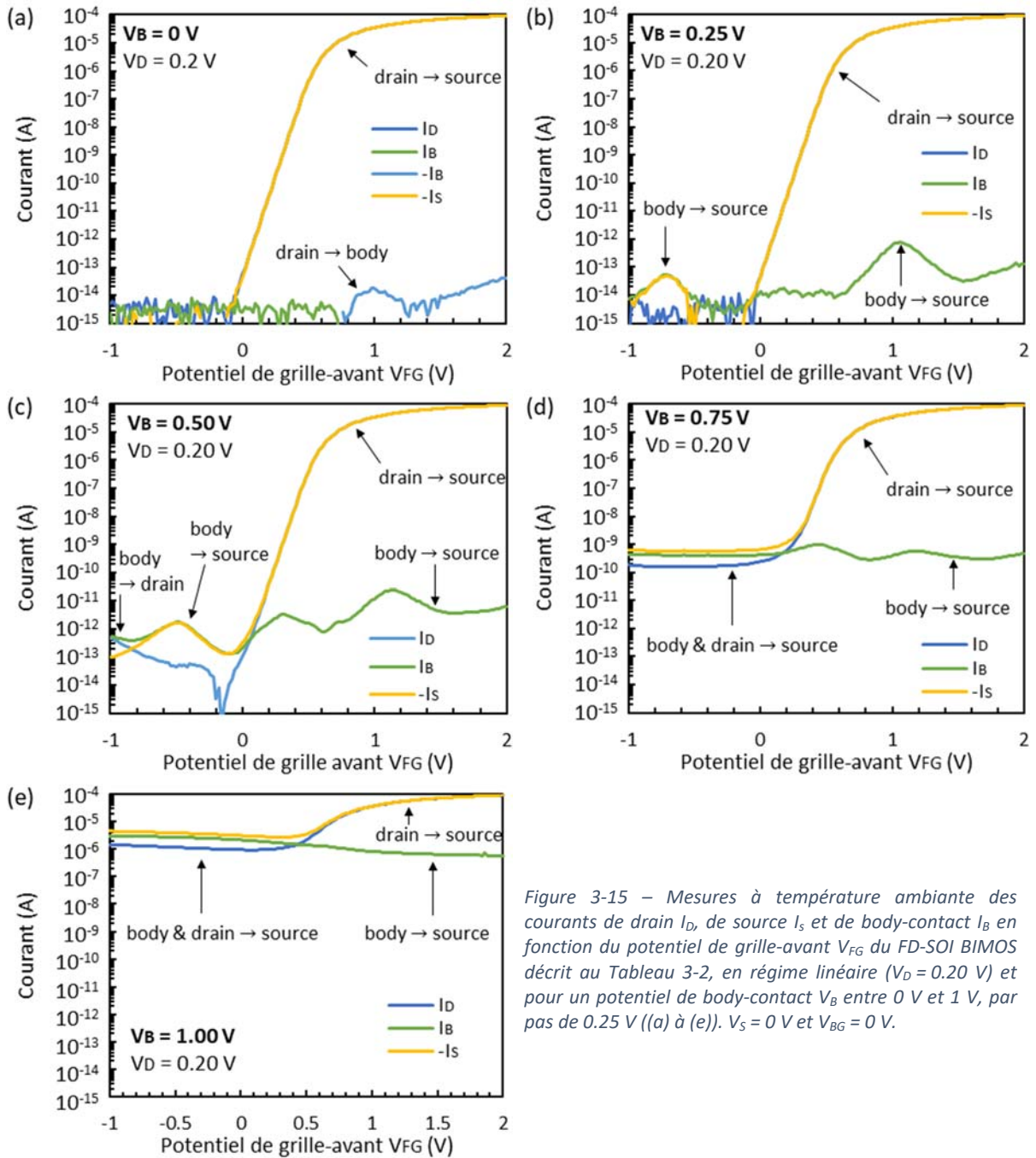


Figure 3-15 – Mesures à température ambiante des courants de drain I_D , de source I_S et de body-contact I_B en fonction du potentiel de grille-avant V_{FG} du FD-SOI BIMOS décrit au Tableau 3-2, en régime linéaire ($V_D = 0.20$ V) et pour un potentiel de body-contact V_B entre 0 V et 1 V, par pas de 0.25 V ((a) à (e)). $V_S = 0$ V et $V_{BG} = 0$ V.

En régime linéaire (Figure 3-15) :

- La polarisation du body-contact a pour effet premier d'injecter du courant dans le composant par ce terminal ($I_B > 0$ pour $V_B \geq 0$ V) de la même manière que lorsque le BIMOS fonctionne en régime de gated-diode. Ce courant circule préférentiellement vers la source, comme le montre la superposition des courbes I_B et $-I_S$ observable lorsque le courant de drain est suffisamment faible (courbes en forme de bosse aux Figure 3-15.b et Figure 3-15.c pour $V_{FG} < 0$ V). Une exception est constatée à la Figure 3-15.c pour $V_{FG} \approx -1$ V où le courant I_B circule

préférentiellement vers le drain. En négligeant dans un premier temps les fluctuations du courant I_B avec les variations du potentiel de grille, on constate que I_B a une dépendance exponentielle en V_B , comme attendu de la part du courant de gated-diode analysé à la section 3.3.1. L'oscillation du courant avec V_{FG} lorsque V_B est compris entre 0.25 V et 0.75 V (Figure 3-15.b à Figure 3-15.d) correspond à la modulation par la grille du courant d'anode de la gated-diode body-contact-source légèrement polarisée en direct déjà caractérisée à la section 3.3.2. Pour rappel, la tension de grille fait varier la répartition des porteurs dans le body ce qui modifie le courant de génération/recombinaison SRH qui domine lorsque la diode est légèrement polarisée en direct.

- Un second effet de la polarisation du body-contact est la modulation du courant drain-source. Cette modulation s'effectue lorsque $V_B > 0.5$ V (Figure 3-15.d et Figure 3-15.e). Elle n'est visible que si le courant de MOSFET est suffisamment faible pour permettre de l'observer ($V_{FG} < V_{T,MOS}$). Autrement dit, elle est masquée par le courant de MOSFET lorsque V_{FG} devient trop important. La modulation du courant drain-source par la polarisation du body-contact est un comportement typique de transistor bipolaire. L'étude du BIMOS dans ce régime de comportement est l'objet de la section 3.5. Notons cependant dès à présent que le courant de body-contact – qui est la base du point de vue BJT – est supérieur au courant de drain – qui est le collecteur du BJT. Cela correspond à une amplification en courant inférieure à 1. Notons de plus que le contrôle du courant de drain par V_B advient alors que la jonction drain-body-contact est polarisée en direct, ce qui est une caractéristique inattendue dans le cadre théorique des BJT standards.
- Enfin, il convient de constater que, sur la plage étudiée, la polarisation du body-contact n'a aucun effet sur la tension de seuil du transistor $V_{T,MOS}$, qui demeure inchangée autour de 0.60 V, quel que soit V_B .

En régime saturé (Figure 3-16), les analyses faites pour le régime linéaire restent valides :

- La polarisation du body-contact V_B :
 - i. Injecte un courant de gated-diode I_{PN} vers la source du BIMOS.
 - ii. Module le courant drain-source lorsque le courant de canal n'est pas dominant.
 - iii. Ne modifie pas la tension de seuil du BIMOS.

Cependant, la présence des courants de body provenant du GIDL et de l'II décale le seuil de polarisation V_B à partir duquel I_{PN} est observable sur le terminal de body-contact. En particulier, I_{PN} est observable sur le terminal de body contact dès $V_B = 0.25$ V en régime linéaire (Figure 3-15.b) et seulement à $V_B = 0.50$ V en régime saturé (Figure 3-16.b). De la même manière, I_{PN} constitue la totalité du courant de body-contact dès $V_B = 0.25$ V en régime linéaire (Figure 3-15.b) et seulement pour $V_B = 1$ V en régime saturé (Figure 3-16.e).

Lorsque l'on compare les courbes obtenues en régime linéaire et saturé, il est important de noter que des courants demeurent quasiment inchangés. Il en est ainsi :

- i. Des courants de source I_S et de drain I_D pour $0 \text{ V} < V_G < 0.6 \text{ V}$ et $V_B \leq 0.75 \text{ V}$ (Figure 3-15.a à Figure 3-15.d et Figure 3-16.a à Figure 3-16.d)
- ii. Du courant de source I_S pour $V_B < 0.5 \text{ V}$ et $V_G < 0 \text{ V}$ (Figure 3-15.c et Figure 3-16.c)
- iii. De la totalité des courants I_D , I_S et I_B pour $V_B \geq 0.75 \text{ V}$ et $V_G < 0 \text{ V}$ (Figure 3-15.d, Figure 3-15.e, Figure 3-16.d et Figure 3-16.c).

Il s'agit respectivement du (i) courant de MOSFET en régime sous le seuil, (ii) du courant de gated-diode et (iii) du courant de BJT. Ces invariances attestent d'une relative absence de couplage entre le MOSFET et la gated-diode d'une part et entre le MOSFET et le BJT d'autre part. Cette absence de couplage est un élément clef dans la perspective d'une modélisation compacte du FD-SOI BIMOS (section 3.6). Notons qu'évoquer un couplage entre la gated-diode et le BJT n'a pas de sens, en ce que le phénomène BJT est lui-même le résultat du couplage entre les gated-diodes body-contact – source et body-contact – drain.

Enfin, la chute du courant de body à $V_B = 1\text{ V}$ et $V_{FG} > 1\text{ V}$, lorsque V_D passe de 0.2 V à 1.8 V demeure inexpliquée et mériterait d'être davantage investiguée. Cette chute est significative, notamment à $V_{FG} = 2\text{ V}$ où I_B diminue d'un facteur 1000 (Figure 3-15.e. et Figure 3-16.e).

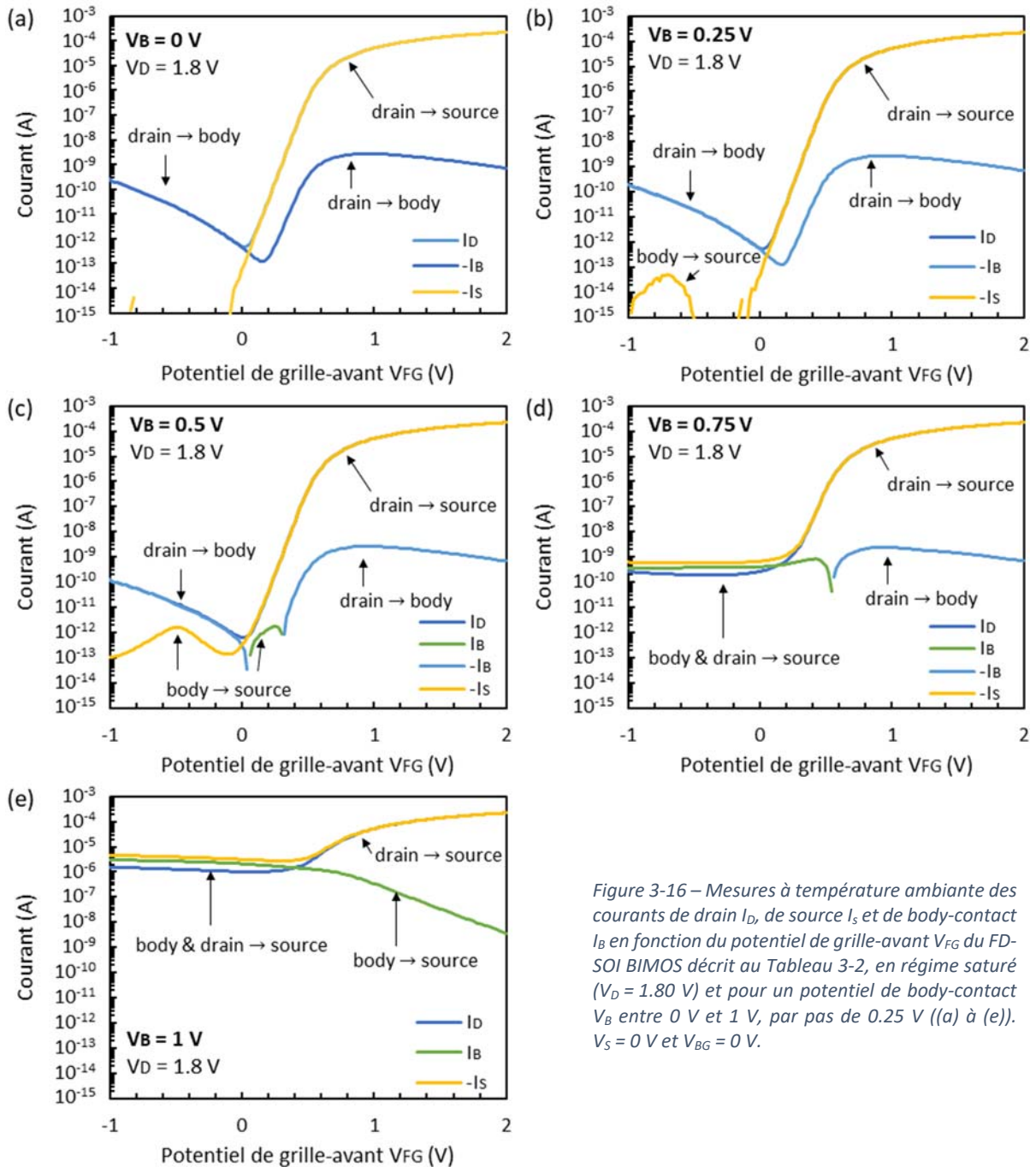


Figure 3-16 – Mesures à température ambiante des courants de drain I_D , de source I_S et de body-contact I_B en fonction du potentiel de grille-avant V_{FG} du FD-SOI BIMOS décrit au Tableau 3-2, en régime saturé ($V_D = 1.80\text{ V}$) et pour un potentiel de body-contact V_B entre 0 V et 1 V , par pas de 0.25 V ((a) à (e)). $V_S = 0\text{ V}$ et $V_{BG} = 0\text{ V}$.

Pour résumer, en régime linéaire comme en régime saturé, la polarisation du body-contact active la gated diode body-contact–source et le BJT de faible gain sans modifier les caractéristiques du MOSFET. En première approximation, les courants issus de ces trois composants se superposent, sans se coupler les uns aux autres. Cependant, deux principales étrangetés demeurent. D’une part, le BJT de faible gain est activé même lorsque sa jonction base-collecteur est polarisé en direct. D’autre part, le courant de body-contact chute fortement lorsque $V_{FG} > 0$ et $V_D = 1.8$ V.

3.4.2 Courbes I_D - V_D

Dans les sections précédentes, le BIMOS a été caractérisé en mode de fonctionnement MOSFET par des mesures de type I_D - V_{FG} . Dans cette section, cette caractérisation se poursuit mais en s’appuyant cette fois sur la mesure à température ambiante de la caractéristique courant tension I_D - V_D du BIMOS de référence, en imposant d’abord la condition $V_B = 0$ V, puis $I_B = 0$ A.

3.4.2.1 Courbes I_D - V_D à $V_B = 0$ V

Dans un premier temps, la caractéristique I_D - V_D du BIMOS de référence est effectuée à $V_B = 0$. Les terminaux de grille-arrière et de source sont branchés à la masse ($V_{BG} = V_S = 0$ V). Le potentiel de drain est balayé de 0 V à 2.4 V pour différents potentiels de grille-avant V_{FG} compris entre 0 V et 1.8 V. De la même façon qu’à la section 3.4.1.1, les courbes mesurées sont comparées avec celles obtenues par simulations SPICE calibrées d’un NMOSFET de dimensions comparables et soumis aux mêmes polarisations. L’ensemble des courbes est réparti en deux groupes selon que les électrodes de grille-avant sont polarisées en-dessous du seuil ($V_{FG} < V_{T,MOS}$) (Figure 3-17) ou au-dessus du seuil ($V_{FG} > V_{T,MOS}$) (Figure 3-18).

Sur la totalité de la plage de polarisation de grille-avant considérée, les courants de drain du BIMOS et du NMOS sont comparables :

- Les courbes I_D se superposent presque idéalement pour $V_{FG} \leq 1$ V.
- Pour $V_{FG} > 1$ V, un décalage apparaît lorsque que les composants entrent en régime de saturation (Figure 3-18). Il conduit à l’inégalité $I_{D,BIMOS} > I_{D,NMOS}$. Bien qu’aucune investigation particulière n’ait été menée sur ce décalage, on peut émettre l’hypothèse que le surplus de courant observé dans le BIMOS provient de la contribution de la partie de la grille en T qui longe le body-contact et qui est parallèle à la direction source-drain (voir Figure 3-11).
- Pour $V_{FG} = 0$ V et $V_D > 1.5$ V (Figure 3-17), l’augmentation du courant de drain est attribuée au courant de BTBT, activé par l’augmentation de la tension drain – grille-avant V_{D-FG} . Dans le BIMOS, ce courant de génération de paires électron-trou circule du drain vers le body-contact. en conformité avec les analyses faites sur la caractéristique I_D - V_{FG} (section 3.4.1.1).

Concernant le courant de body-contact I_B du BIMOS :

- I_B n’est pas mesurable pour de faibles valeurs de V_D , et augmente avec V_{FG} pour de fortes valeurs de V_D .
- Pour $V_D > 1$ V et $V_{FG} < V_{T,MOS}$ (Figure 3-17), I_B effectue un « aller-retour » avec l’augmentation de V_{FG} : il diminue puis augmente en valeur absolue ($|I_B(V_{FG}=0.2 \text{ V})| < |I_B(V_{FG}=0.0 \text{ V})| < |I_B(V_{FG}=0.4 \text{ V})|$).
- Pour $V_D > 1.5$ V et $V_{FG} > V_{T,MOS}$ (Figure 3-18), I_B effectue aussi un « aller-retour » à mesure que V_{FG} augmente.

La dépendance de I_B en V_D s'explique par les variations des courants de générations par BTBT et II avec V_D : le BTBT augmentent avec la tension $V_{D-FG} = V_D - V_{FG}$ augmente [32], et l'II augmente avec la tension $V_{D-S} = V_D - V_S$ [33]. A fort V_D et en régime au-dessous du seuil ($V_{FG} < V_{T,MOS}$), l'aller-retour de I_B avec V_{FG} témoigne de la transition depuis le BTBT vers l'II dans le mécanisme de génération de paires électron-trou. A fort V_D et en régime au-dessus du seuil, ($V_{FG} > V_{T,MOS}$), l'aller-retour de I_B avec V_{FG} est cette fois dû à la réduction de l'II lorsque V_{FG} augmente [33]. Ces deux aller-retours correspondent aux extrema locaux de la caractéristique I_D - V_{FG} à $V_D = 1.8$ V (section 3.4.1.1, Figure 3-12).

Autrement dit, en présence de phénomènes de génération (V_D fort), et lorsque le courant de canal est suffisamment faible ($V_{FG} \ll V_{T,MOS}$), le courant de drain est un courant d'électrons généré par BTBT dont la dépendance en V_{D-FG} est exponentielle. Ce courant circule sous la forme d'un courant de trous vers la source dans le NMOS et vers le body-contact dans le BIMOS.

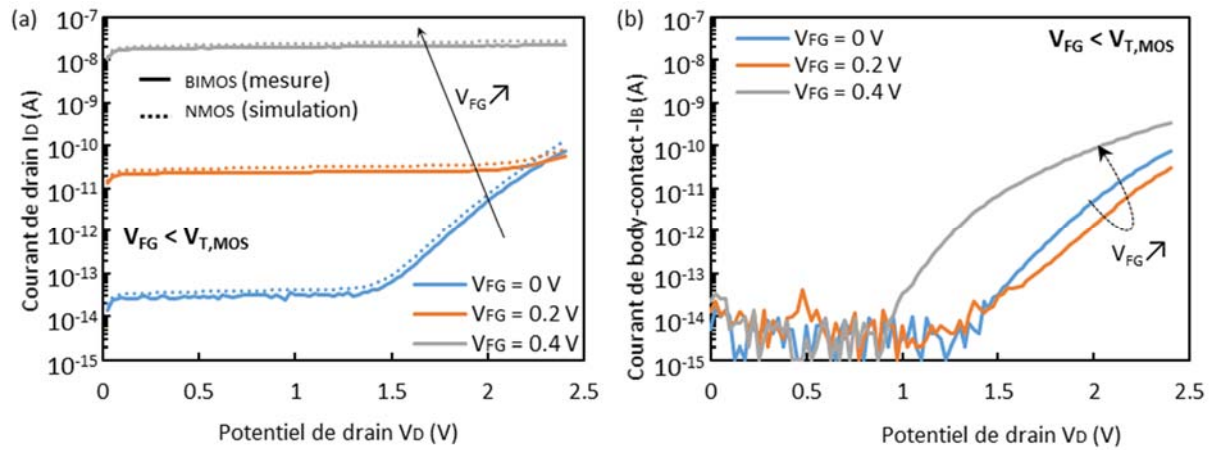


Figure 3-17 – (a) Courants de drain I_D du BIMOS de référence et du NMOS aux dimensions comparables (Tableau 3-2) en fonction du potentiel de drain V_D et à différents $V_{FG} < V_{T,MOS}$. (b) Courant de de body-contact $-I_B$ du BIMOS de référence en fonction du potentiel de drain V_D et à différents $V_{FG} < V_{T,MOS}$. $V_B = 0$ V, $V_S = 0$ V et $V_{BG} = 0$ V. Les données qui concernent le BIMOS sont obtenues par mesure sous pointe à température ambiante. Celles qui concernent le NMOS par simulation SPICE calibrée.

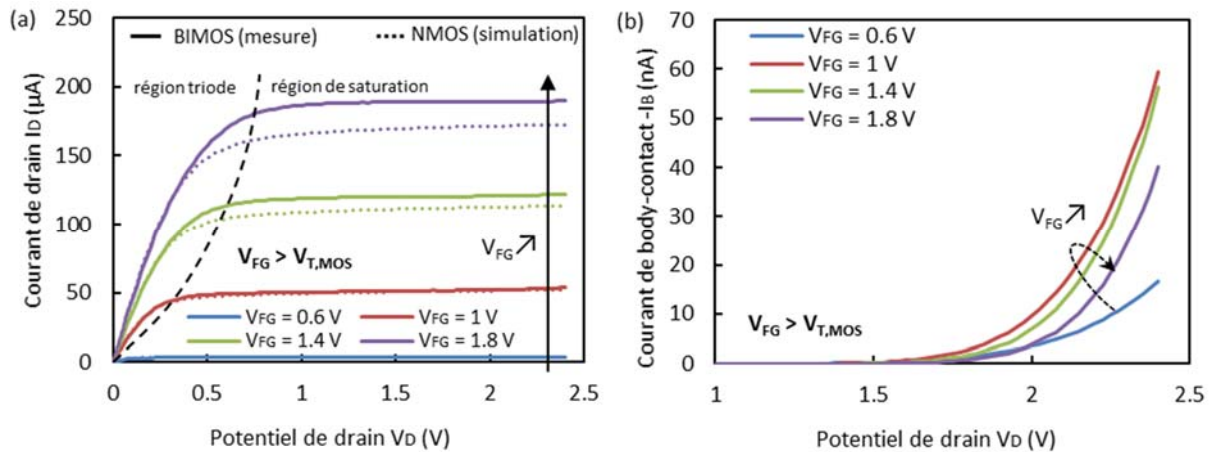


Figure 3-18 – (a) Courants de drain I_D du BIMOS de référence et du NMOS aux dimensions comparables (Tableau 3-2) en fonction du potentiel de drain V_D et à différents $V_{FG} > V_{T,MOS}$. (b) Courant de de body-contact $-I_B$ du BIMOS de référence en fonction du potentiel de drain V_D et à différents $V_{FG} > V_{T,MOS}$. $V_B = 0$ V, $V_S = 0$ V et $V_{BG} = 0$ V. Les données qui concernent le BIMOS sont obtenues par mesure sous pointe à température ambiante. Celles qui concernent le NMOS par simulation SPICE calibrée.

Par ailleurs, en présence de phénomène de génération (V_D fort) et lorsque le courant MOSFET est fort ($V_{FG} > V_{T,MOS}$), le courant de body-contact du BIMOS est un courant de trous généré par II qui circule

depuis le drain sous forme d'un courant d'électrons. Ce courant d'électrons est masqué par la haute intensité du courant de canal. La génération par II intervient a priori aussi dans le MOSFET, mais il circule en parallèle du courant de canal depuis le drain vers la source et se trouve donc totalement masqué.

Pour résumer, dans la caractéristique I_D - V_D , les courants de drain du BIMOS et du NMOS sont comparables et conformes au fonctionnement MOSFET typique. Le courant de body-contact du BIMOS est un courant de trous observable uniquement lorsque V_D est suffisamment fort pour déployer un champ électrique suffisamment intense à la jonction drain-body pour activer les mécanismes de génération. Selon la configuration des polarisations, le mécanisme de génération à l'origine de I_B est soit le BTBT ($V_{FG} < 0.2$ V), soit l'II ($V_{FG} > 0.2$ V).

3.4.2.2 Courbe I_D - V_D à $I_B = 0$ A

Dans cette section, le même type de mesure I_D - V_D qu'à la section précédente est effectué, mais en imposant cette fois $I_B = 0$ A. De cette manière, il est possible de mesurer le potentiel du body-contact en fonction de V_D et V_{FG} . Cela revient également à maintenir virtuellement le body contact en nœud flottant.

Ainsi des mesures du courant de drain du BIMOS (Figure 3-19.a) et du potentiel de body-contact (Figure 3-19.b) sont effectuées, à température ambiante, en fonction de la tension de drain V_D et à plusieurs polarisations de grille-avant $V_{FG} > 0$, en imposant $I_B = 0$. Les courbes du courant de drain obtenues sont comparées à celles d'un NMOS de dimensions comparables et soumis aux mêmes polarisations, obtenues par simulation SPICE (Figure 3-19.a).

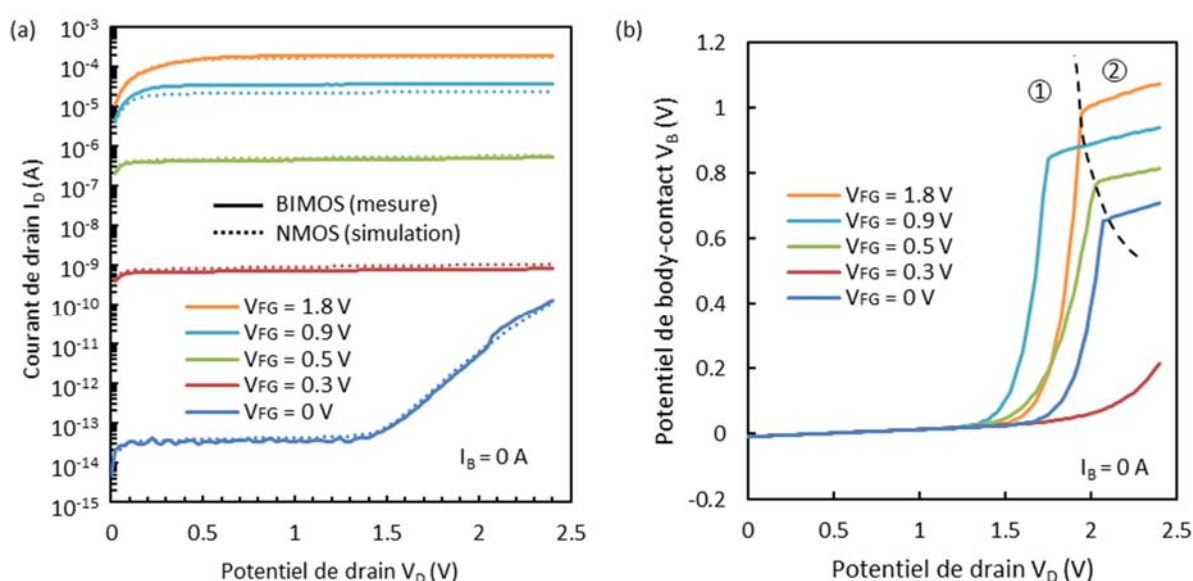


Figure 3-19 – Mesures du courant de drain I_D (a) et du potentiel de body-contact V_B (b) en fonction du potentiel de drain V_D , à différentes valeurs de potentiel de grille-avant V_{FG} , effectuées simultanément sur le BIMOS de référence (Tableau 3-2). $V_S = V_{BG} = 0$ V et $I_B = 0$ A.

Les courbes de courant de drain sont identiques à celles obtenues lorsque $V_B = 0$ V était imposé. Cela témoigne une fois de plus de la faible influence du potentiel de body-contact sur le fonctionnement MOSFET du BIMOS.

Le potentiel de body-contact V_B reste au voisinage de 0 V, tant que $V_D < 1.5$ V. Au-delà, il croît avec V_D , selon deux régimes distincts et successifs (notés ① et ② à la Figure 3-19.b), observables sur toutes

les courbes à l'exception de celle qui correspond à $V_{FG} = 0.3$ V. L'accroissement dV_B/dV_D est plus important dans le premier régime que dans le second. Par ailleurs, la dépendance de V_B en V_{FG} n'est pas monotone : pour $V_D > 1.9$ V, V_B décroît lorsque $0 \leq V_{FG} \leq 0.3$ V, puis croît lorsque $V_{FG} > 0.3$ V.

Une façon d'interpréter les courbes V_B - V_D consiste à considérer que V_B est le potentiel nécessaire pour s'opposer à la sortie du courant I_B mesuré à $V_B = 0$. Dans cette perspective, il est intéressant de tracer les courbes du courant de body $I_{B,V0}$, mesuré lorsque $V_B = 0$ V (Figure 3-17 et Figure 3-18), en fonction du potentiel de body $V_{B,I0}$ mesuré lorsque $I_B = 0$ A, à différent V_{FG} (Figure 3-19). Il s'agit de courbes paramétrées par le potentiel V_D .

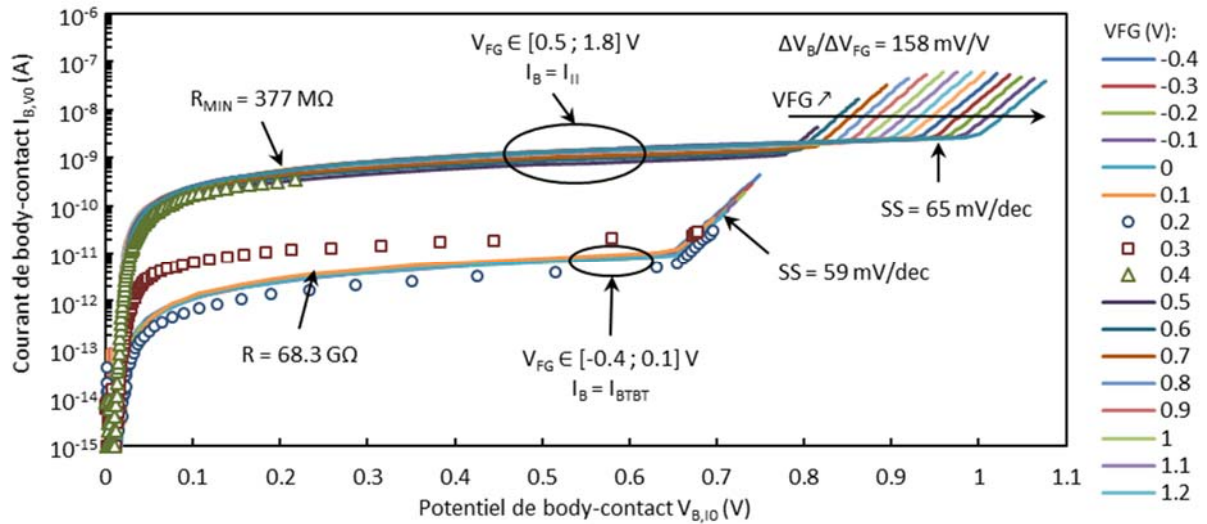


Figure 3-20 – Courbes paramétrées du courant de body $I_{B,V0}$ mesuré lorsque $V_B = 0$ V, en fonction du potentiel de body $V_{B,I0}$ mesuré lorsque $I_B = 0$ A. L'abscisse et l'ordonnée de chaque point des courbes proviennent de deux expériences distinctes mais pour lesquelles le couple $[V_D; V_{FG}]$ est identique.

Présentées de cette manière (Figure 3-20), les courbes se regroupent en 2 groupes distincts selon le potentiel V_{FG} : l'un pour $V_{FG} \leq 0.1$ V, localisée à $I_{B,V0}$ faible, et l'autre pour $V_{FG} \geq 0.4$ V localisée à $I_{B,V0}$ fort, (les courbes obtenues pour $V_{FG} \in [0.2; 0.3]$ V sont les courbes de transitions). Ces regroupements sont liés à la nature du mécanisme de génération à l'origine du courant de body-contact : BTBT dans le premier groupe et II dans le second.

Ainsi ce graphe montre que, pour une même intensité $I_{B,V0}$, V_B doit être davantage augmenté pour bloquer un courant généré par BTBT que pour bloquer un courant provenant de l'II. Cette façon de voir les choses est simpliste, car elle suppose l'absence de couplage entre le potentiel V_B et le mécanisme de génération de I_B . En effet, on peut également supposer que la polarisation de V_B influence les mécanismes des générations, et que cette influence est différente dans le cas du BTBT et de l'II.

Quoi qu'il en soit, les courbes des deux groupes ont des formes générales similaires, composées d'une région de croissance linéaire, dont on peut extraire une résistance R , suivie, lorsque V_B atteint une valeur frontière $V_{B,F}$, d'une région de croissance exponentielle quantifiée par le paramètre SS , en mV/dec.

Les courbes de la famille générée par BTBT sont toutes superposées ($R = 68.3$ GΩ, $SS = 59$ mV/dec, $V_{B,F} = 0.66$ V). Les courbes de la famille générée par II sont superposées dans la région linéaire pour $V_{FG} \geq 1$ V ($R = 377$ MΩ). Pour V_{FG} plus petit, elles sont décalées vers le bas ($R = 839$ MΩ à $V_{FG} = 0.4$ V). Dans la région exponentielle, toutes les courbes ont un accroissement égal ($SS = 65$ mV/dec). Cependant la tension de frontière $V_{B,F}$ se décale avec V_{FG} de façon régulière ($\Delta V_{B,F}/\Delta V_{FG} = 158$ mV/V).

La région de croissance exponentielle évoque le comportement de gated-diode du BIMOS. Notamment, les portions exponentielles des courbes $V_{FG} \in [-0.4 ; 0.5]$ V sont toutes alignées, et correspondent à la courbe d'une diode idéale définie par $\eta = 1$ et $I_{SAT} = 0.07$ aA. Pour $V_{FG} \geq 0.6$ V, η reste très proche de 1 également. Ces valeurs de I_{SAT} et η sont cependant différentes de celles extraites des mesures du mode gated-diode du BIMOS (section 3.3).

Il est possible de modéliser les courbes de la Figure 3-20 par un circuit comprenant une résistance variable commandée par V_{FG} en parallèle à une diode idéale en série avec une source de tension également commandée par V_{FG} (Figure 3-21). La source de tension et la résistance ont deux comportements bien distincts selon que le BIMOS est en dessous ou au-dessus du seuil.

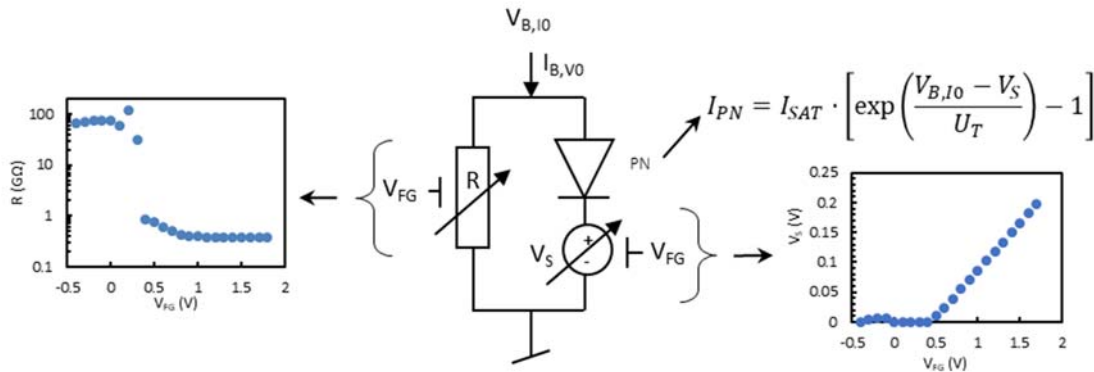


Figure 3-21 – Schéma électrique qui modélise la courbe $I_{B,V0}$ - $V_{B,I0}$ de la Figure 3-20. Les fonctions de transfert de la résistance R et de la source de tension V_S ainsi que l'équation de la diode idéale sont reportées également.

Les raisons physiques d'une telle modélisation n'ont pas été investiguées dans cette thèse. Cependant cette approche permet d'ouvrir des pistes de réflexion pour identifier les mécanismes sous-jacents d'un tel comportement. Cette modélisation n'est pas implémentée dans le modèle présenté à la section 3.6.

3.5 Caractérisation du BIMOS en mode BJT en émetteur commun

Dans cette section, le FD-SOI BIMOS est étudié dans son comportement en mode BJT émetteur commun. Dans ce qui suit, les électrodes du BIMOS en contact avec le film de silicium prennent la nomenclature relative au BJT : la source devient l'émetteur, le body-contact devient la base et le drain devient le collecteur. Nous verrons que deux régimes de fonctionnement sont identifiables : un régime « normal altéré », où le courant de collecteur est contrôlé par la tension base – émetteur V_{B-E} mais avec un gain en courant β_F inférieur à 1 ($I_C < I_B$), et un régime « anormal » où les courants sont indépendants de V_{B-E} , le courant de base est négatif et β_F perd son sens. Cependant, nous verrons qu'en régime anormal, le calcul de β_F correspond au gain du BJT parasite $\beta_{F,PAR}$ qui est présent dans le FD-SOI NMOSFET dépourvu de body contact.

Le mode de fonctionnement BJT du BIMOS a déjà été étudié en branchement base-commune [13]. Cependant, ce mode de branchement ne permet pas de distinguer le courant MOSFET du courant de BJT au niveau du collecteur. De plus, le contrôle du courant de collecteur par le courant de base n'a pas été démontré. Dans ce qui suit, le BIMOS est caractérisé en mode émetteur-commun ($V_E = 0$ V) avec une polarisation de la grille-avant tel que le BIMOS est maintenu en régime MOSFET profondément sous le seuil ($V_{FG-S} \leq 0.2 \ll V_{T,MOS}$). Sous cette condition, la contribution du courant de canal I_{CH} dans le courant de collecteur est réduite au maximum. Cela permet d'observer uniquement l'effet d'amplification bipolaire.

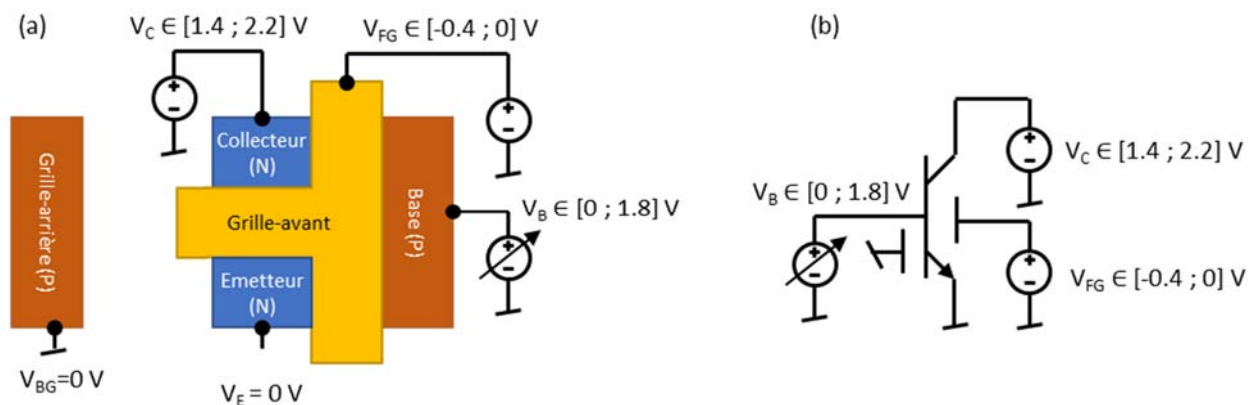


Figure 3-22 – (a) Vue de dessus et (b) vue schématique du montage de la caractérisation du BIMOS en mode BJT émetteur-commun

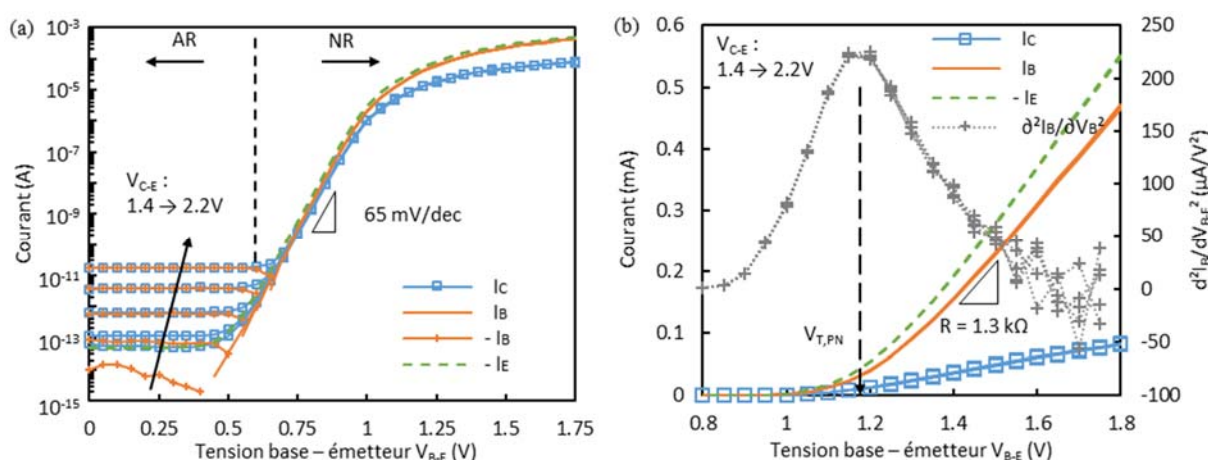


Figure 3-23 – Mesure des courants de collecteur I_C , d'émetteur I_E et de base I_B en fonction de la tension base – émetteur V_{B-E} avec V_{C-E} comme paramètre, en échelle (a) semi-logarithmique et (b) linéaire. La dérivée de la conductance est également tracée ((b) axe de droite)

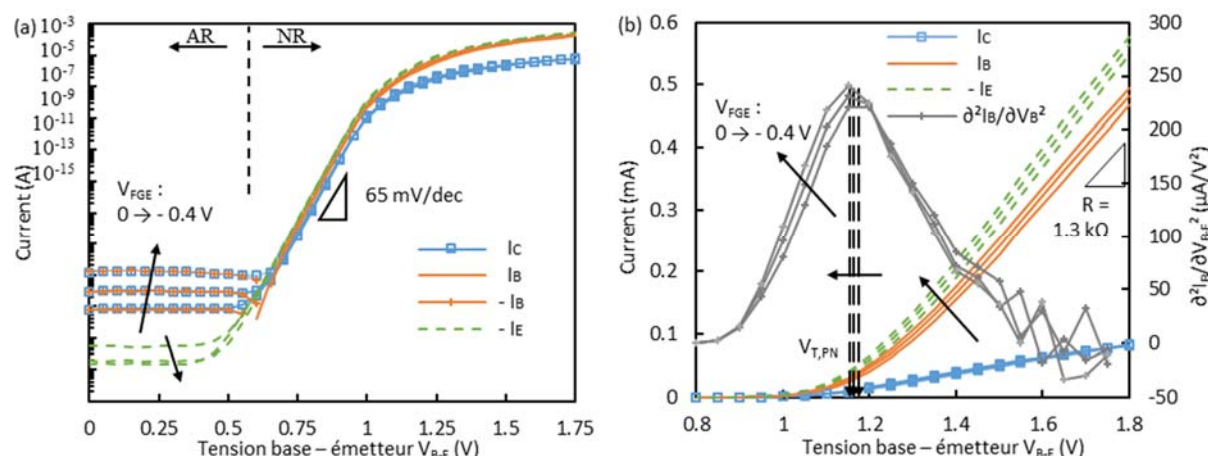


Figure 3-24 – Mesure des courants de collecteur I_C , d'émetteur I_E et de base I_B en fonction de la tension base – émetteur V_{B-E} avec V_{FGE} comme paramètre, en échelle (a) semi-logarithmique et (b) linéaire. La dérivée de la conductance est également tracée ((b) axe de droite)

Afin d'observer la circulation des courants entre tous les terminaux en contact avec le film de silicium, les courants d'émetteur I_E , de base I_B et de collecteur I_C sont mesurés. Une tension collecteur – émetteur $V_{C-E} > 0$ V est appliquée et un balayage sur la tension base – émetteur V_{B-E} entre 0 V et 1.8 V

est effectué (Figure 3-22). Les tensions V_{C-E} et V_{FG-E} sont successivement prises comme paramètres lors de deux expériences distinctes. Dans la première, la tension V_{C-E} varie entre 1.4 et 2.2 V, avec $V_{FG-E} = 0$ V (Figure 3-23), et dans la seconde, la tension V_{FG-E} varie entre -0.4 V et 0 V, avec $V_{C-E} = 1.8$ V (Figure 3-24).

Les courants I_C et I_B sont utilisés ensuite pour calculer le gain en courant en émetteur commun, selon (2.43) : $\beta_F = (I_C - I_{CEO}) / I_B$. Pour rappel, I_{CEO} est le courant de collecteur lorsque la base est laissée flottante. I_{CEO} est mesuré en fonction de V_{C-E} pour $V_{FG-E} = 0$ V (Figure 3-25.a) et en fonction de V_{FG-E} pour $V_{C-E} = 1.8$ V (Figure 3-25.b). S'il l'on considère le BIMOS en tant que body-contacted MOSFET, il s'agit des mêmes mesures que celles des caractéristiques I_D-V_D et I_D-V_{FG} respectivement, lorsque le body-contact est flottant (section 3.3). Pour les deux expériences, β_F est tracé en fonction de V_{B-E} (Figure 3-26), puis en fonction de I_B (Figure 3-27).

Les deux expériences mènent à des courbes très similaires. Dans chacune d'elles on peut définir un régime de fonctionnement normal altéré du BJT et un régime de fonctionnement anormal, en fonction de la tension V_{B-E} .

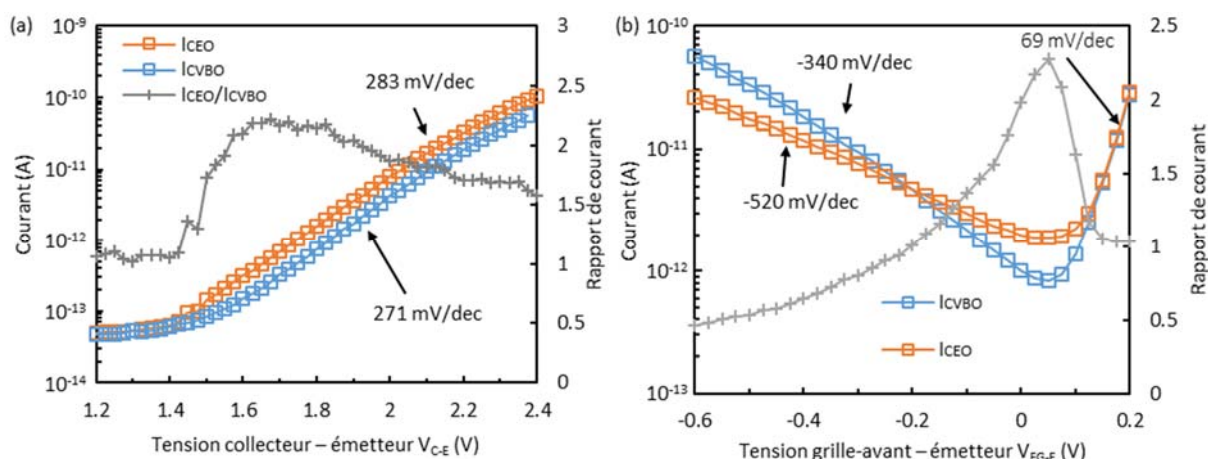


Figure 3-25 – Courant de collecteur en base ouverte I_{CEO} , courant de collecteur en base mise à la masse $I_{CVBO} = I_C(V_B = 0)$ (axe de gauche) et rapport entre ces deux courants I_{CEO}/I_{CVBO} (axe de droite), en fonction des tensions (a) collecteur – émetteur et (b) grille-avant – émetteur.

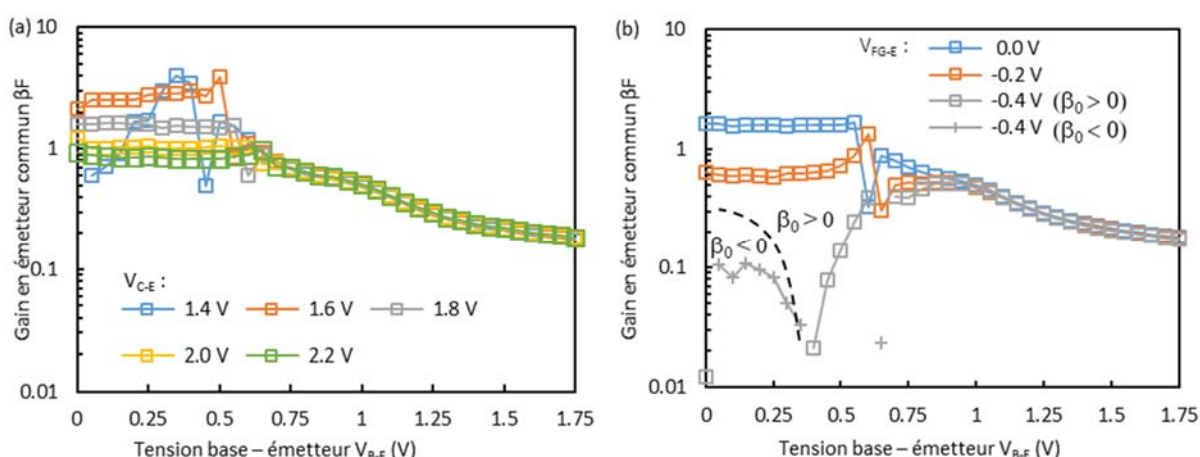


Figure 3-26 – Gain β_F en fonction de V_{B-E} , (a) avec V_{C-E} comme paramètre et $V_{FG-E} = 0$ V, et (b) V_{FG-E} comme paramètre et $V_{C-E} = 1.8$ V. Données extraites de la Figure 3-23 et de la Figure 3-24.

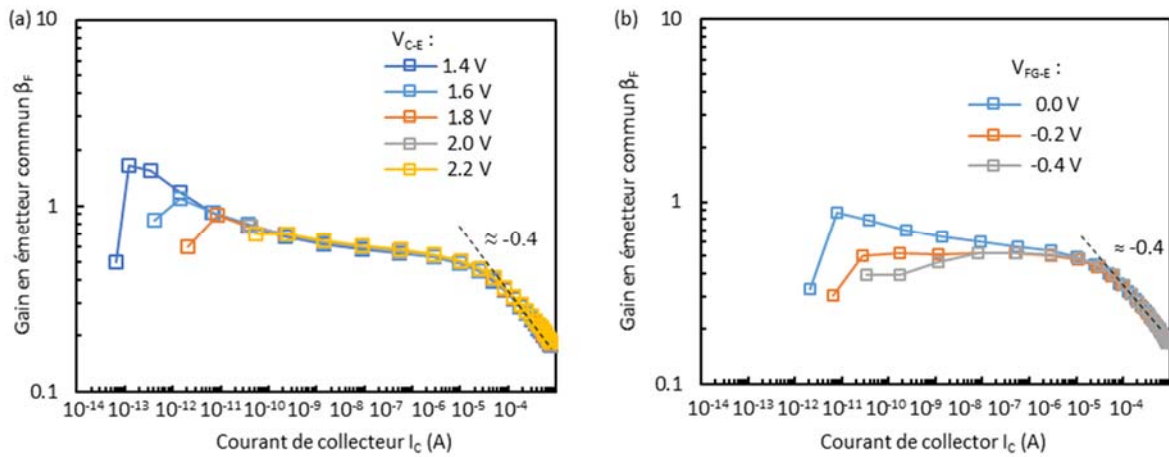


Figure 3-27 – Gain β_F en fonction de I_C (a) avec V_{CE} comme paramètre et $V_{FG-E} = 0$ V, et (b) V_{FG-E} comme paramètre et $V_{CE} = 1.8$ V. Données extraites de la Figure 3-23 et de la Figure 3-24.

3.5.1 Régime de fonctionnement normal altéré

3.5.1.1 Courants de collecteur et de base

Lorsque V_{B-E} est suffisamment grand ($V_{B-E} > 0.7$ V) le BIMOS a un régime de fonctionnement normal altéré. Le comportement BJT est observé (Figure 3-23 et Figure 3-24) :

- Le courant de base I_B est un courant de diode dépendant de V_{B-E} qui circule de la base vers l'émetteur. Ce courant a d'abord une croissance exponentielle sur la plage $0.7 \text{ V} < V_{B-E} < 1 \text{ V}$, avec un facteur d'idéalité η proche de 1 ($SS = 65 \text{ mV/dec}$). La diode fonctionne en régime de diffusion où le courant de génération/recombinaison devient négligeable ($I_{RG} \ll I_{DIFF}$ et $I_{PN} = I_{DIFF}$). Ensuite, la croissance du courant devient linéaire sur la plage $V_{B-E} > 1.2 \text{ V}$ avec un inverse de pente de $1.3 \text{ k}\Omega$. La diode fonctionne en régime de très haute injection où les résistances séries des zones neutres dominent. Notons que cette résistance correspond à celle observée lors de la caractérisation du BIMOS en mode gated-diode à la section 3.3.
- Le courant de collecteur I_C est contrôlé par V_{B-E} , avec le même comportement que I_B (régime exponentiel pour $0.7 \text{ V} < V_{B-E} < 1 \text{ V}$ et linéaire pour $V_{B-E} > 1.2 \text{ V}$). On remarque qu'il est inférieur à I_B , ce qui témoigne d'un gain en courant β_F inférieur à 1.

La tension de transition $V_{T,PN}$ entre les deux régimes de fonctionnement de la diode est évaluée en utilisant l'équation (3.1) de la section 3.3.1. Le maximum de $\partial^2 I_B / \partial V_B^2$ est déterminé en faisant une interpolation quadratique sur les 3 points de la courbe les plus hauts. Sur les plages de polarisation explorées, $V_{T,PN}$ est indépendant de V_{CE} (Figure 3-23.b) et dépendant de V_{FG-E} (Figure 3-24.b) avec la relation $\partial V_{T,PN} / \partial V_{FG-E} > 0$. ($V_{T,PN} = 1.174 \text{ V}$ à $V_{FG-E} = 0 \text{ V}$ et $V_{T,PN} = 1.154 \text{ V}$ à $V_{FG-E} = -0.4 \text{ V}$).

Le décalage de $V_{T,PN}$ correspond à un décalage de la tension pour laquelle la diode atteint le régime de résistance série. Si la diode atteint ce régime plus tôt, cela veut dire que la résistivité de la ZCE de la diode chute plus rapidement au voisinage de la tension de transition.

3.5.1.2 Gain en courant β_F

Bien qu'il présente un régime de fonctionnement normal BJT, les caractéristiques du gain en courant (Figure 3-26 et Figure 3-27) suggèrent que ce régime est altéré, et ce, pour deux raisons principales :

- β_F est inférieur à 1 sur toute la plage du régime de fonctionnement normal. Il s'agit davantage d'une atténuation plutôt que d'une amplification.
- β_F diminue à mesure que V_{B-E} augmente ($\partial\beta_F/\partial V_{B-E} < 0$), comme décrit par l'effet Webster dans la théorie du BJT [34]. Cependant, l'effet Webster prédit une décroissance $\partial\log(\beta_F)/\partial\log(I_C)$ égale à -1 alors que dans le BIMOS cette décroissance est égale à -0.45. Cette différence s'explique par le fait que la prédiction par l'effet Webster concerne le régime de haute injection dans lequel I_B et I_C sont théoriquement proportionnels à $\exp(qV_{B-E}/kT)$ et $\exp(qV_{B-E}/2kT)$ respectivement, alors que la décroissance observée dans le BIMOS s'effectue en régime de très haute injection où $\partial I_B/\partial V_{B-E}$ et $\partial I_C/\partial V_{B-E}$ sont tous deux constants. Notons qu'en régime d'injection forte, $|\partial\log(\beta_F)/\partial\log(I_C)| < 0.4$ est encore plus éloigné de la valeur -1 théorique.

Les BJT que l'on trouve dans le commerce ont généralement un β_F de l'ordre de 100. En comparaison, le β_F du BIMOS, de l'ordre de 1, est donc très faible. Il correspond à un gain en base commune α_F de l'ordre de 1/2 (équation (2.41) section 2.2.2). Pour rappel, α_F est le produit du rendement d'émetteur γ_F et du facteur de transport dans la base α_T , tous deux possiblement compris entre 0 et 1. Avec $\alpha_F = 1/2$, les cas limites sont $(\gamma_F; \alpha_T) = (1/2; 1)$ et $(\gamma_F; \alpha_T) = (1; 1/2)$. Dans le premier cas limite, autant d'électrons que de trous circulent dans la jonction émetteur – body ($\gamma = 1/2$) et tous les électrons injectés dans le body rejoignent la jonction body – collecteur ($\alpha_T = 1$). Dans le second cas limite, le flux de trous est négligeable devant le flux d'électrons au niveau la jonction émetteur – body ($\gamma_F = 1$), seule la moitié du flux d'électrons rejoint la jonction body – collecteur ($\alpha_T = 0.5$) et l'autre moitié se recombine avec les trous présents dans le body. Ce dernier cas est peu probable : les densités de dopants dans les terminaux de base et d'émetteur sont équivalentes, ce qui suggère des flux d'électrons et de trous équivalents, dans les jonctions base – body et émetteur – body (c.f. équation (2.39) section 2.2.2).

3.5.2 Régime de fonctionnement anormal

3.5.2.1 Courants de collecteur et de base

Lorsque V_{B-E} est suffisamment petit ($V_{B-E} < 0.7$ V) le BIMOS a un régime de fonctionnement anormal (Figure 3-23 et Figure 3-24):

- Les courants de base I_B et de collecteur I_C sont indépendants de la tension V_{B-E} .
- Le courant de base I_B est négatif : il circule du BIMOS vers le terminal de base. Le fait que ce courant rejoigne la base plutôt que l'émetteur lorsque $V_{B-E} = 0$ indique qu'il s'agit d'un courant de trous qui est généré dans le body.
- Le courant de base I_B est égal et opposé au courant de collecteur I_C . Il s'agit donc du même courant de génération qui circule du collecteur vers la base : les électrons générés dans le body rejoignent le collecteur.

I_B et I_C ont par ailleurs des dépendances exponentielles et opposées en V_{C-E} en V_{FG-E} (Figure 3-23.a et Figure 3-24.a respectivement). Ces dépendances correspondent aux pentes de courbes $I_{C,VB0}$ des Figure 3-25.a et Figure 3-25.b : $\partial V_{C-E}/\partial\log(I_{C,VB0}) = 271$ mV/dec et $\partial V_{FG-E}/\partial\log(I_{C,VB0}) = 340$ mV/dec. Au travers de ces deux dépendances, c'est en réalité la tension grille-avant – collecteur V_{C-FG} qui est impliquée. Cette dépendance en V_{C-FG} et le fait que, dans ces expériences $V_{FG} \leq 0$, montrent que le phénomène de génération impliqué est le BTBT (voir les caractérisations du BIMOS en mode MOSFET en régime de saturation à la section 3.4.1.3). Le fait que les dépendances en V_{C-E} et V_{FG-E} ne soient pas exactement égales indique que le BTBT a des dépendances vis-à-vis des autres tensions, comme par exemple la tension grille-arrière – collecteur V_{BG-C} , qui n'est pas la même dans les deux expériences.

3.5.2.2 Gain en courant β_F

Dans le régime anormal, comme I_B est négatif, le gain en courant β_F (Figure 3-26 et Figure 3-27) souffre d'une perte de signification. Cela est particulièrement évident lorsque β_F prend des valeurs négatives pour $V_{FG-E} = -0.4$ V et $V_{B-E} < 0.4$ V (Figure 3-26.b). Lorsque β_F est positif avec $I_B < 0$, cela tient au fait que $I_C - I_{CEO}$ est aussi négatif, c'est-à-dire que $I_C < I_{CEO}$. Autrement dit, le courant de collecteur est plus intense lorsque la base est flottante que lorsqu'elle est polarisée faiblement. Pour $V_B = 0$, la relation $I_C < I_{CEO}$ est mise en évidence par la position des courbes I_{CEO} situées au-dessus des courbes $I_{C,VB0}$ à la Figure 3-25.

3.5.2.3 Caractérisation du BJT parasite de la structure MOSFET

Le fait que $I_C < I_{CEO}$ peut néanmoins s'expliquer par l'activation dans la structure MOSFET d'un BJT parasite BJT_{PAR} [35] lorsque la base est laissée flottante. En effet, dans cette configuration, les trous générés dans le body s'évacuent par l'émetteur qui, en retour, injecte des électrons dans le body par effet bipolaire. Ceux-ci rejoignent le collecteur et augmentent le courant I_C .

Il est possible d'évaluer le gain en courant en émetteur commun $\beta_{F,PAR}$ associé au BJT_{PAR} . Pour ce faire, les courants qui circulent dans le BIMOS sont décrits de la manière suivante (pour faciliter la lecture, tous les courants sont considérés comme positifs, qu'ils sortent ou qu'ils entrent dans le BIMOS) :

- Lorsque la base est à la masse :
 - Le courant d'émetteur $I_{E,VB0}$ est constitué du courant de canal MOSFET : $I_{E,VB0} = I_{CH}$.
 - Le courant de base $I_{B,VB0}$ est constitué du courant de génération par BTBT : $I_{B,VB0} = I_{BTBT}$.
 - Le courant de collecteur contient à la fois le courant de canal MOSFET et le courant de génération par BTBT : $I_{C,VB0} = I_{CH} + I_{BTBT}$.

Dans cette configuration, le BJT_{PAR} est inactif. Le courant de collecteur peut être considéré comme étant le courant de collecteur en base ouverte $I_{CEO,PAR}$: $I_{CEO,PAR} = I_{C,VB0} = I_{CH} + I_{BTBT}$.

- Lorsque la base est laissée flottante :
 - Le courant de base est nul.
 - Le courant d'émetteur I_{EEO} est égal au courant de collecteur I_{CEO} . Ces courants sont constitués du courant de canal MOSFET I_{CH} , du courant de génération I_{BTBT} et du courant d'électrons $I_{e,BJT,PAR}$ injecté par l'activation du BJT_{PAR} : $I_{EEO} = I_{CEO} = I_{CH} + I_{BTBT} + I_{e,BJT,PAR}$.

Dans cette configuration, le BJT_{PAR} est actif. Son courant de base $I_{B,PAR}$ est le courant de trous généré dans le body par BTBT : $I_{B,PAR} = I_{BTBT}$. Son courant de collecteur $I_{C,PAR}$ est le courant de collecteur du BIMOS en base laissée flottante : $I_{C,PAR} = I_{CEO} = I_{CH} + I_{BTBT} + I_{e,BJT,PAR}$.

Dans ce contexte, l'équation (2.43) qui définit le gain en courant devient :

$$\beta_{F,PAR} = \frac{I_{C,PAR} - I_{CEO,PAR}}{I_{B,PAR}} = \frac{I_{e,BJT,PAR}}{I_{BTBT}} \quad (3.7)$$

Or, en remarquant que $I_{C,PAR} = I_{CEO}$, $I_{CEO,PAR} = I_{C,VB0}$ et $I_{B,PAR} = -I_{B,VB0}$, il vient :

$$\beta_{F,PAR} = \frac{I_{CEO} - I_{C,VB0}}{-I_{B,VB0}} = \frac{-(I_{C,VB0} - I_{CEO})}{-I_{B,VB0}} = \beta_F(V_{BE} = 0) \quad (3.8)$$

Ces calculs montrent que, en régime anormal, bien que le courant de base soit négatif et que la notion de gain en courant soit compromise, le calcul de β_F permet d'évaluer le gain en courant $\beta_{F,PAR}$ du BJT

parasite dans le canal du MOSFET. Ce gain a une valeur maximale de 4 sur la plage de tension étudiée, ce qui reste une valeur particulièrement faible. Elle est néanmoins en accord avec les valeurs mesurées dans [35] sur un FD-SOI MOSFET d'épaisseur de film de 7nm. Cette faible valeur explique pourquoi le BJT parasite est négligé dans les modèles compacts des transistors FD-SOI.

L'analyse faite ici a par ailleurs ses limites. Par exemple, aucun courant de recombinaison n'est pris en compte, et l'on considère que tous les trous générés participent au courant $I_{B,PAR}$. De plus, cette approche devient incohérente lorsque $I_C - I_{CEO}$ change de signe pour $V_{FG-E} < 0.5$ V (Figure 3-25.b).

3.5.3 Résumé

Pour résumer, lorsque le FD-SOI BIMOS fonctionne en mode BJT, deux régimes de comportement sont identifiés :

- Un régime normal altéré pour V_{B-E} faible, où la tension base – émetteur contrôle les courants de base et de collecteur, mais avec un gain en courant β_F inférieur à 1.
- Un régime anormal où I_B est négatif et où ni le courant de collecteur, ni le courant de base ne dépendent de V_{B-E} . I_B est un courant de génération par BTBT qui circule depuis le collecteur ($I_B = -I_C$) et qui dépend principalement de la tension grille-avant – collecteur. Bien que β_F soit dépourvue de signification dans ce régime, son calcul permet d'évaluer le gain en courant du BJT parasite de la structure MOSFET.

3.6 Implémentation dans le modèle UTSOI

Les caractérisations du BIMOS en mode gated-diode, MOSFET et BJT permettent d'envisager un modèle analytique, ou modèle compact, du comportement BIMOS.

Les modèles compacts de transistor sont utilisés dans des simulateurs électriques pour prévoir le comportement de circuits comprenant un nombre important de composants, en des temps raisonnables. Ces simulateurs sont systématiquement utilisés dans l'industrie pour dimensionner les composants d'un circuit dans le but d'implémenter des fonctions de transfert particulières.

Depuis les débuts de la micro-électronique dans les années 1950 des nombreux modèles de composants ont été développés et améliorés à mesure que les technologies de fabrication se diversifiaient et évoluaient. A titre d'exemple, un simulateur comme ELDO, propose plus de 40 modèles de MOSFET différents, dont la pertinence d'utilisation dépend de la technologie utilisée (substrat bulk, PD-SOI, FD-SOI, poly-silicon thin-film transistor, high-voltage transistor, etc...).

Le modèle adopté par STMicroelectronics pour modéliser les MOSFET de sa technologie 28nm UTBB FD-SOI est le modèle UTSOI v2 développé par le CEA-LETI. Il a donc été choisi comme base pour concevoir le modèle du FD-SOI BIMOS.

3.6.1 Le modèle UTSOI

Le modèle UTSOI est un modèle basé sur une description physique des transistors FD-SOI, notamment en ce qui concerne le couplage entre les interfaces avant et arrière du film de silicium (interface entre le film – oxyde de grille et film – BOX). Etant donné une polarisation des 4 terminaux de MOSFET (source, drain, grille-avant et grille-arrière), UTSOI calcule les courants qui circulent entre les terminaux, ainsi que les bruits, notés S, qui leur sont associés. Les courants sont répartis en deux catégories selon qu'il s'agit des courants dus au transport de charge dans la structure (notés I) ou aux variations de charge stockée dans chacun des terminaux (notés dQ/dt) (Figure 3-28).

$$I_{II} = a_1 \cdot (V_{DS} - a_3 \cdot \Delta\psi) \cdot \exp\left(-\frac{a_2^*}{V_{DS} - a_3 \cdot \Delta\psi}\right) \cdot I_{DS} \quad [A] \quad (3.9)$$

avec :

$$a_2^* = a_2 \cdot (T_{KD}/T_{KR})^{K_{a2}} \cdot [1 + a_4(\sqrt{2 \cdot \phi_B - V_{BC-S}} - \sqrt{2 \cdot \phi_B})] \quad [V] \quad (3.10)$$

où a_1 , a_2 , a_3 et a_4 sont des paramètres du modèle, $\Delta\psi = \psi_{SS} - \psi_{SD}$, ψ_{SS} et ψ_{SD} sont les potentiels de surface aux extrémités du canal côté source et côté drain respectivement, T_{KD} et T_{KR} sont respectivement la température du composant et celle de référence, V_{BC-S} est la tension body-contact – source, $\phi_B = \phi_T \cdot \ln(N_A/n_i)$ est le potentiel de Fermi, N_A la densité de dopants du body et n_i la densité de porteurs intrinsèques.

Les courbes des courants de body-contact mesurés et simulés avec le jeu de paramètres ($a_1 ; a_2 ; a_3 ; a_4 ; K_{a2}$) = (100 ; 18 ; 1 ; 0 ; 0) sont comparées à la Figure 3-29. Pour rappel, les courants de body-contact dans la région $V_{FG} < 0$ sont issus du BTBT, ceux dans la région $V_{FG} > 0$ sont issus de l'II. Les mises à 0 des paramètres a_4 et K_{a2} ont pour effet de désactiver les dépendances de l'II en température et vis-à-vis de la tension body-contact – source V_{BC-S} , qui n'ont pas été investiguées dans ce travail. Les valeurs des paramètres a_1 à a_3 sont le résultat d'une optimisation manuelle.

Le modèle d'II implémenté rend bien compte du comportement global du courant I_B mesuré sur la plage $V_{FG} > 0$ (courbe en forme de cloche). Cependant, il souffre d'un manque de précision, qui ne peut pas être amélioré par le simple ajustement des paramètres. Cela montre que les hypothèses sur lesquelles sont basées les équations (3.9) et (3.10) ne sont pas toutes valides dans le cas d'un FD-SOI BIMOS.

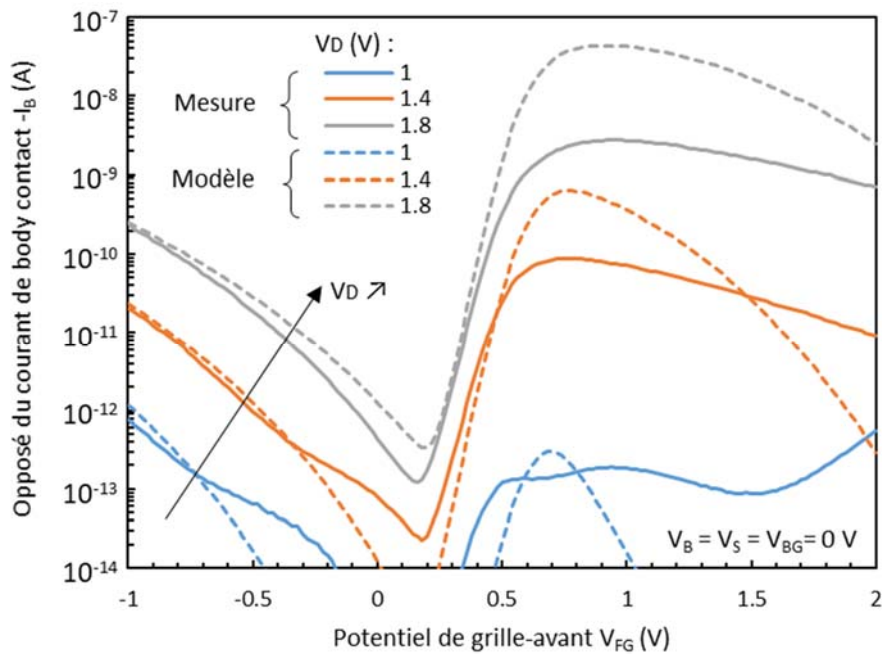


Figure 3-29 – Courants de body-contact mesurés sur le BIMOS de référence (lignes continues) et simulés par modélisation compacte selon les équations (3.9) et (3.10) et le jeu de paramètres ($a_1 ; a_2 ; a_3 ; a_4 ; K_{a2}$) = (100 ; 18 ; 1 ; 0 ; 0) (lignes pointillées) en fonction du potentiel de grille-avant V_{FG} et avec le potentiel drain V_D pris comme paramètre. Les autres terminaux sont mis à la masse.

3.6.2.2 Courant de gated-diode

En s'appuyant sur les mesures et la modélisation exposées à la Figure 3-10 (section 3.3.3), les courants des gated-diodes body-contact – source et body-contact – drain sont modélisés selon le schéma et les paramètres de la Figure 3-30, où chaque courant de diode est calculé selon l'équation (2.35).

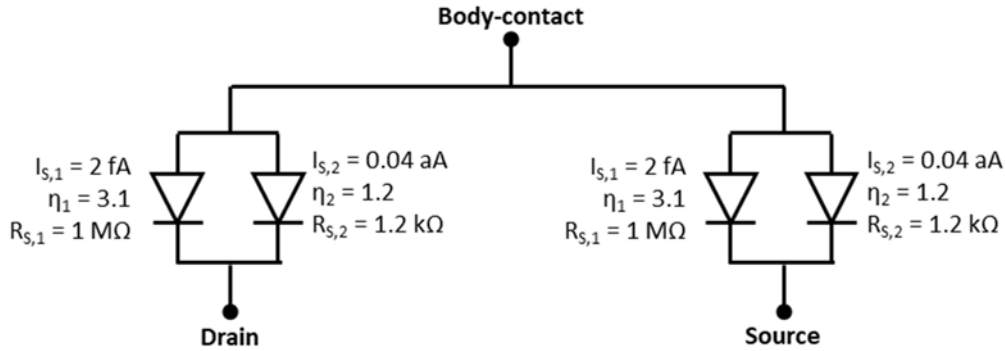


Figure 3-30 – Schéma électrique et paramètres utilisés pour modéliser les gated-diodes body-contact – drain et body-contact – source.

3.6.2.3 Courant de BJT

L'injection de courant depuis la body-contact vers la source induit un courant qui circule du drain vers la source, par mécanisme d'amplification bipolaire. Du fait de la topologie particulière du BIMOS (présence d'une grille, film de silicium entièrement déplété), les modèles classiques de description des BJT ne sont pas pertinents pour rendre compte de son comportement dans ce régime. Ainsi, le choix est fait d'évaluer la relation entre β_F et V_{B-E} à l'aide d'une fonction mathématique ajustée sur les données.

De nouvelles extractions de β_F en de fonction V_{B-E} sont effectuées, en suivant la méthodologie adoptée à la section 3.5, mais avec une plage de V_{C-E} explorée étendue à $V_{C-E} \in [0.2 ; 2.2]$ V et en retirant les points aberrants où $I_B < 0$ ou $(I_C - I_{CE0}) < 0$ (Figure 3-31). La fonction $\beta_F = f(V_{B-E})$ est approximée par :

$$\beta_{F,1} = a_{1,1} \cdot \exp(a_{1,2} \cdot V_{B-E}) \quad (3.11)$$

sur la plage $V_{B-E} \in [0.35 ; 0.6]$ V, avec $a_{1,1} = 3.12 \cdot 10^{-4}$ et $a_{1,2} = 12.5 \text{ V}^{-1}$, et par :

$$\beta_{F,2} = a_{2,1} \cdot \exp(a_{2,2} \cdot V_{B-E}) \quad (3.12)$$

sur la plage $V_{B-E} \in [0.7 ; 1.8]$ V, avec $a_{2,1} = 1.91$ et $a_{2,2} = -1.43 \text{ V}^{-1}$. La continuité entre ces deux fonctions est assurée par la fonction de raccordement MIN_FUNC utilisée dans UTSOI :

$$\beta_F = \text{MIN_FUNC}(\beta_{F,1}; \beta_{F,2}; a_3) = \frac{1}{2} \left(\beta_{F,1} + \beta_{F,2} - \sqrt{(\beta_{F,1} - \beta_{F,2})^2 + a_3} \right) \quad (3.13)$$

Où $a_3 = 0.1$ est un paramètre de la fonction. La courbe obtenue est tracée à la Figure 3-31.

Bien que ce modèle rende compte approximativement de la relation $\beta_F = f(V_{B-E})$ mesurée, les décalages observés entre le modèle et les mesures illustrent les limites de la démarche adoptée pour la modélisation. Pour aller plus loin, un travail supplémentaire doit être fourni pour construire un modèle physique de BJT adapté à la topologie du BIMOS.

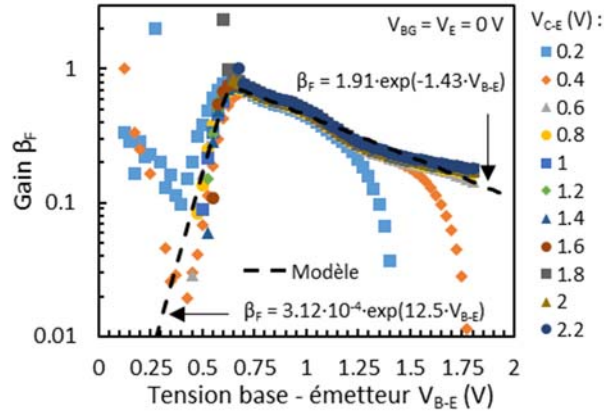


Figure 3-31 – Gain en courant en mode émetteur commun β_F en fonction de la tension base – émetteur V_{B-E} et pour une tension émetteur – collecteur V_{C-E} entre 0.2 V et 2.2 V, extrait à partir de mesures effectuées à température ambiante sur le BIMOS de référence, et modèle de $\beta_F = f(V_{B-E})$ selon l'équation (3.13)

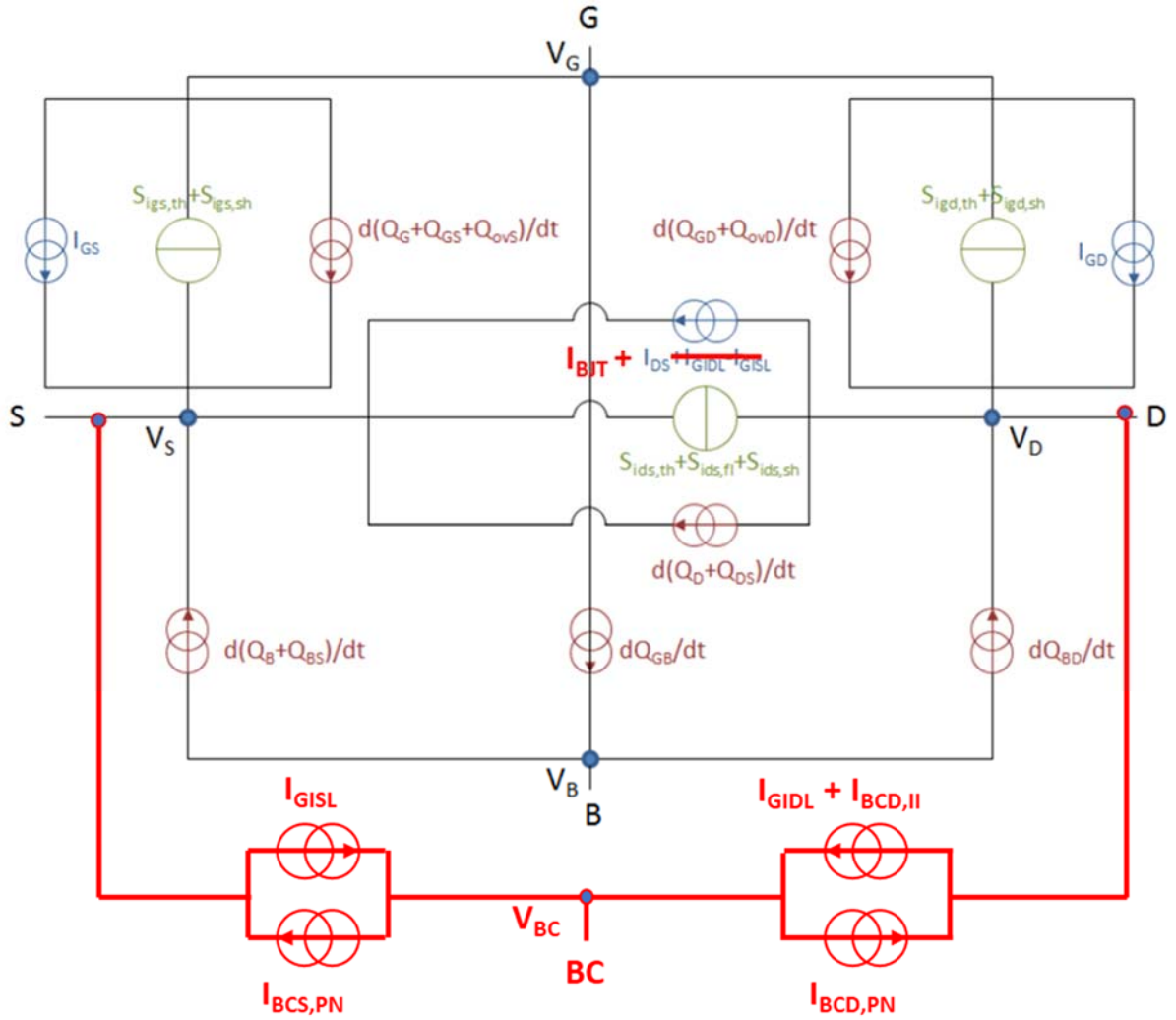


Figure 3-32 – Schéma électrique des courants de sortie calculés dans le modèle UTISOI modifié [36]. La source est notée S, le drain D, la grille-avant G, la grille-arrière B et le body-contact BC.

3.6.2.4 Modèle UTSOI modifié

Le modèle UTSOI est modifié selon le schéma électrique de la Figure 3-32. Par rapport au modèle UTSOI initial, un terminal de body-contact est ajouté. Les courants de GIDL et GISL qui circulaient initialement entre la source et le drain sont dérivés vers ce nouveau terminal et un courant d'ionisation par impact est ajouté entre le drain et le body-contact, comme décrit à la section 3.6.2.1. Les courants de diode circulant entre le body-contact et le drain d'une part et entre le body-contact et la source d'autre part sont ajoutés, comme décrit à la section 3.6.2.2. Un courant de BJT est ajouté entre la source et le drain, comme décrit à la section 3.6.2.3.

3.7 Conclusion

Dans ce chapitre, le composant BIMOS intégré en technologie 28nm UTBB FD-SOI et dimensionné aux tailles minimales permises par la technologie a été présenté, caractérisé et modélisé de façon compacte. D'un point de vue topologique, le BIMOS peut être considéré comme une diode PIN pourvue d'une grille (« gated-diode »), un NMOSFET pourvu d'un contact de body (« body-contacted NMOSFET ») ou un BJT pourvu d'une grille (« gated BJT »). Les caractérisations électriques à température ambiante du BIMOS selon ces trois modes de fonctionnement ont mis en évidence le rôle significatif des courants de génération dans les mesures de courants faibles.

En mode gated-diode, les mécanismes de génération identifiés comme dominants sont le BTBT en polarisation inverse et la génération/recombinaison SRH en polarisation légèrement directe. Ces deux mécanismes sont fortement modulés par le potentiel de grille.

En mode body-contacted NMOSFET, les mécanismes de génération dominants sont le BTBT et l'II. Comme attendu de la part d'une structure MOSFET, ces courants apparaissent pour de fortes tensions V_{D-S} , à V_{FG-S} faible ou négatif pour le BTBT et à V_{GS} faible ou positif pour l'II. Par ailleurs, ces deux mécanismes sont fortement modulés par les potentiels de grilles. Tant que la valeur du potentiel de body-contact est faible, ces courants circulent depuis le drain vers le body-contact et leurs intensités ne dépendent pas du potentiel de ce dernier. Par ailleurs, quelle que soit sa valeur sur la plage explorée, le potentiel de body-contact ne modifie pas la tension de seuil $V_{T,MOS}$ du MOSFET.

En mode gated-BJT branché en émetteur commun, le mécanisme de génération dominant est le BTBT, activé pour une tension collecteur – grille-avant suffisamment forte. Tant que le potentiel de base n'est pas suffisamment haut, le courant généré circule du collecteur vers la base, et conduit à un fonctionnement anormal du BJT où le courant de base I_B est négatif. En fonctionnement normal (potentiel V_B haut et courant I_B positif), le gain en émetteur commun est inférieur à 1.

A partir de l'analyse de ces mesures, un jeu d'équations a été ajouté au modèle UTSOI pour rendre compte de l'ensemble des comportements observés. En particulier, un modèle d'ionisation par impact adapté aux modèles basés sur le calcul du potentiel de surface est implémenté. Ce dernier, bien que reproduisant globalement la forme de la courbe, souffre d'un manque de précision qui ne peut pas être comblé par le simple ajustement des paramètres du modèle. Néanmoins, le modèle UTSOI ainsi modifié s'avère être un outil pertinent pour la simulation de circuit utilisant le BIMOS, comme cela est démontré au chapitre 4.

L'accent mis sur l'analyse des courants de générations prendra tout son sens lors de la description au chapitre suivant du fonctionnement du circuit de neurone développé dans cette thèse. En effet, ces courants qui circulent préférentiellement vers le terminal de body-contact, y jouent un rôle crucial.

3.8 Références

- [1] P. Galy, "Etude théorique, modélisation et analyse par simulateur tridimensionnel d'une cellule BICMOS optimisée : conception d'un transistor vertical BICMOS original," PhD Thesis, 1994.
- [2] S. Verdonckt-Vandebroek, S. S. Wong, J. C. S. Woo, and P. K. Ko, "High-gain lateral bipolar action in a MOSFET structure," *IEEE Trans. Electron Devices*, vol. 38, no. 11, pp. 2487–2496, Nov. 1991.
- [3] P. Galy and V. Berland, "The ideal NPN vertical BICMOS transistor analytical model simulation and experimental results of the collector current," *Int. J. Electron.*, vol. 81, no. 5, pp. 501–516, 1996.
- [4] P. Galy, J. Bourgeat, T. Lim, C. Fenouillet-Beranger, and D. Golanski, "ESD protection with BICMOS transistor for bulk amp; FDSOI advanced CMOS technology," in *CAS 2013 (International Semiconductor Conference)*, 2013, vol. 2, pp. 171–174.
- [5] J. Olsson, B. Edholm, A. Soderbarg, and K. Bohlin, "High current gain hybrid lateral bipolar operation of DMOS transistors," *IEEE Trans. Electron Devices*, vol. 42, no. 9, pp. 1628–1635, Sep. 1995.
- [6] S. A. Parke, C. Hu, and P. K. Ko, "Bipolar-FET hybrid-mode operation of quarter-micrometer SOI MOSFETs (MESFETs read MOSFETs)," *IEEE Electron Device Lett.*, vol. 14, no. 5, pp. 234–236, May 1993.
- [7] R. Huang, X. Zhang, and Y. Wang, "A high-performance SOI drive-in gate controlled hybrid transistor (DGCHT)," *IEEE Trans. Electron Devices*, vol. 45, no. 9, pp. 2079–2081, Sep. 1998.
- [8] F. Assaderaghi, S. Parke, D. Sinitsky, J. Bokor, P. K. Ko, and C. Hu, "A dynamic threshold voltage MOSFET (DTMOS) for very low voltage operation," *IEEE Electron Device Lett.*, vol. 15, no. 12, pp. 510–512, Dec. 1994.
- [9] J.- Colinge, "An SOI voltage-controlled bipolar-MOS device," *IEEE Trans. Electron Devices*, vol. 34, no. 4, pp. 845–849, Apr. 1987.
- [10] K. Yallup, S. Edwards, and O. Creighton, "A Novel Bipolar Device on SOI Wafers for Analog BICMOS Applications," 1994.
- [11] C. Entringer, P. Flatresse, P. Galy, F. Azais, and P. Nouet, "Partially Depleted SOI body-contacted MOSFET-triggered silicon controlled rectifier for ESD protection," in *2006 Electrical Overstress/Electrostatic Discharge Symposium*, 2006, pp. 166–171.
- [12] P. Galy, S. Athanasiou, and S. Cristoloveanu, "BICMOS transistor in thin silicon film and new solutions for ESD protection in FDSOI UTBB CMOS technology," in *EUROSOI-ULIS 2015: 2015 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon*, 2015, pp. 29–32.
- [13] S. Athanasiou, C. A. Legrand, S. Cristoloveanu, and P. Galy, "Novel Ultrathin FD-SOI BiMOS Device With Reconfigurable Operation," *IEEE Trans. Electron Devices*, vol. 64, no. 3, pp. 916–922, Mar. 2017.
- [14] H.-S. Wong, M. H. White, T. J. Krutsick, and R. V. Booth, "Modeling of transconductance degradation and extraction of threshold voltage in thin oxide MOSFET's," *Solid-State Electron.*, vol. 30, no. 9, pp. 953–968, 1987.
- [15] A. J. Scholten, G. D. J. Smit, M. Durand, R. V. Langevelde, and D. B. M. Klaassen, "The Physical Background of JUNCAP2," *IEEE Trans. Electron Devices*, vol. 53, no. 9, pp. 2098–2107, Sep. 2006.
- [16] G. A. M. Hurkx, H. C. de Graaff, W. J. Kloosterman, and M. P. G. Knuvers, "A new analytical diode model including tunneling and avalanche breakdown," *IEEE Trans. Electron Devices*, vol. 39, no. 9, pp. 2090–2098, Sep. 1992.
- [17] S. Cristoloveanu, K. H. Lee, and M. Bawedin, "A reconfigurable silicon-on-insulator diode with tunable electrostatic doping," *J. Appl. Phys.*, vol. 122, no. 8, p. 084502, Aug. 2017.
- [18] T. Ernst, S. Cristoloveanu, A. Vandooren, T. Rudenko, and J. P. Colinge, "Recombination current modeling and carrier lifetime extraction in dual-gate fully-depleted SOI devices," *IEEE Trans. Electron Devices*, vol. 46, no. 7, pp. 1503–1509, Jul. 1999.

- [19] T. Rudenko *et al.*, "Determination of film and surface recombination in thin-film SOI devices using gated-diode technique," *Solid-State Electron.*, vol. 48, no. 3, pp. 389–399, 2004.
- [20] X. Zhao and D. E. Ioannou, "'Gated-diode' in SOI MOSFETs: a sensitive tool for characterizing the buried Si/SiO/sub 2/ interface," *IEEE Trans. Electron Devices*, vol. 48, no. 4, pp. 685–687, Apr. 2001.
- [21] T. Poiroux *et al.*, "UTSOI2: A complete physical compact model for UTBB and independent double gate MOSFETs," in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 12.4.1–12.4.4.
- [22] T. Poiroux *et al.*, "Leti-UTSOI2.1: A Compact Model for UTBB-FDSOI Technologies - Part I: Interface Potentials Analytical Model," *IEEE Trans. Electron Devices*, vol. 62, no. 9, pp. 2751–2759, Sep. 2015.
- [23] T. Poiroux *et al.*, "Leti-UTSOI2.1: A Compact Model for UTBB-FDSOI Technologies - Part II: DC and AC Model Description," *IEEE Trans. Electron Devices*, vol. 62, no. 9, pp. 2760–2768, Sep. 2015.
- [24] N. D. Arora and M. S. Sharma, "MOSFET substrate current model for circuit simulation," *IEEE Trans. Electron Devices*, vol. 38, no. 6, pp. 1392–1398, Jun. 1991.
- [25] T. Y. Chan, P. K. Ko, and C. Hu, "A simple method to characterize substrate current in MOSFET's," *IEEE Electron Device Lett.*, vol. 5, no. 12, pp. 505–507, Dec. 1984.
- [26] T. Y. Chan, J. Chen, P. K. Ko, and C. Hu, "The impact of gate-induced drain leakage current on MOSFET scaling," in *1987 International Electron Devices Meeting*, 1987, pp. 718–721.
- [27] K. Kurimoto, Y. Odake, and S. Odanaka, "Drain leakage current characteristics due to the band-to-band tunneling in LDD MOS devices," in *International Technical Digest on Electron Devices Meeting*, 1989, pp. 621–624.
- [28] V. Nathan and N. C. Das, "Gate-induced drain leakage current in MOS devices," *IEEE Trans. Electron Devices*, vol. 40, no. 10, pp. 1888–1890, Oct. 1993.
- [29] L. Huang, P. T. Lai, J. P. Xu, and Y. C. Cheng, "Mechanism analysis of gate-induced drain leakage in off-state n-MOSFET," in *1997 IEEE Hong Kong Proceedings Electron Devices Meeting*, 1997, pp. 94–97.
- [30] Y.-K. Choi, D. Ha, T.-J. King, and J. Bokor, "Investigation of gate-induced drain leakage (GIDL) current in thin body devices: single-gate ultra-thin body, symmetrical double-gate, and asymmetrical double-gate MOSFETs," *Jpn. J. Appl. Phys.*, vol. 42, no. 4S, p. 2073, 2003.
- [31] Y. K. Lin *et al.*, "Modeling of Back-Gate Effects on Gate-Induced Drain Leakage and Gate Currents in UTB SOI MOSFETs," *IEEE Trans. Electron Devices*, vol. 64, no. 10, pp. 3986–3990, Oct. 2017.
- [32] G. Gildenblat *et al.*, "PSP: An advanced surface-potential-based MOSFET model for circuit simulation," *IEEE Trans. Electron Devices*, vol. 53, no. 9, pp. 1979–1993, 2006.
- [33] X. Gu, H. Wang, T. Chen, and G. Gildenblat, "Substrate current in surface-potential-based compact MOSFET models," in *Tech. Proc. Nanotechnol. Conf*, 2003, pp. 310–313.
- [34] W. M. Webster, "On the Variation of Junction-Transistor Current-Amplification Factor with Emitter Current," *Proc. IRE*, vol. 42, no. 6, pp. 914–920, Jun. 1954.
- [35] F. Liu, I. Ionica, M. Bawedin, and S. Cristoloveanu, "Extraction of the Parasitic Bipolar Gain Using the Back-Gate in Ultrathin FD SOI MOSFETs," *IEEE Electron Device Lett.*, vol. 36, no. 2, pp. 96–98, Feb. 2015.
- [36] T. Poiroux, O. Rozeau, S. Martinie, and M.-A. Jaud, "Leti-UTSOI 2.2.0 User's Manual." CEA-LETI.
- [37] H. J. Park, M. Bawedin, H. G. Choi, and S. Cristoloveanu, "Kink effect in ultrathin FDSOI MOSFETs," *Solid-State Electron.*, vol. 143, pp. 33–40, May 2018.
- [38] G. Gildenblat, *Compact Modeling: Principles, Techniques and Applications*. Springer Netherlands, 2010.

Chapitre 4 : Le circuit de neurone impulsif basé sur le BIMOS : « BIMOS-based spiking neuron circuit »

Ce chapitre présente le circuit de neurone à intégration à fuite et déclenchement basé sur le BIMOS (« BIMOS-based leaky integrate-and-fire spiking neuron » : BB-LIF SN), développé à l'occasion des travaux de recherche présentés dans cette thèse. Après une brève description du comportement dynamique du potentiel de membrane des neurones biologiques (section 4.1), le modèle de neurone à intégration à fuite et déclenchement (« leaky integrate-and-fire » : LIF) est présenté, ainsi qu'une façon standard de l'implémenter dans un circuit analogique (section 4.2). L'implémentation alternative proposée est ensuite décrite, mesurée et discutée, en s'attardant d'abord sur le sous circuit de déclenchement à fuite basé sur le BIMOS (« BIMOS-based leaky trigger circuit » : BB-LTC) (section 4.3) composé d'un BIMOS et d'un NMOS et dont le principe de fonctionnement est exposé en détail, avant d'aborder le circuit du BB-LIF SN complet, composé du BB-LTC et d'une capacité (section 4.4). Enfin, l'intégration complète du BB-LIF SN en technologie 28 FD-SOI est présentée, où tous les composants du circuit sont portés sur silicium et auquel un circuit tampon est ajouté pour permettre le découplage avec les synapses qui seront connectées en sortie du circuit (section 4.5). Certaines des mesures analysées au long du chapitre sont confrontées aux résultats de simulations qui utilisent le modèle compact de BIMOS développé au chapitre 3.

4.1 Potentiel de membrane et potentiel d'action des neurones biologiques

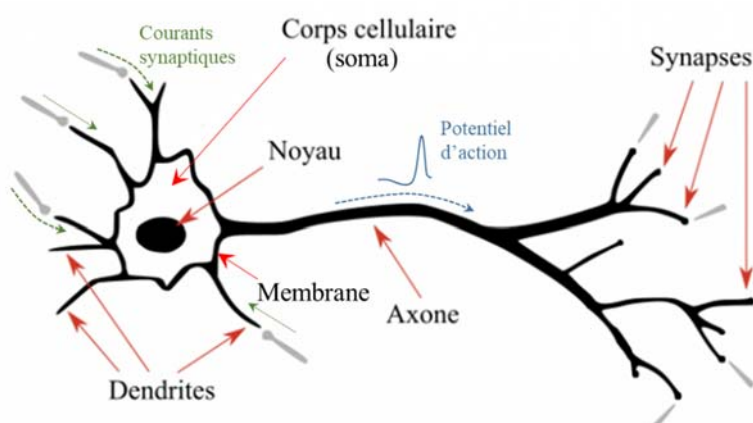


Figure 4-1 – Schéma d'un neurone. Les éléments grisés sont les synapses et dendrites d'autres neurones. D'après [1].

Le potentiel de membrane d'un neurone biologique varie selon l'activité des synapses auxquelles son soma est relié, via ses dendrites (Figure 4-1). Ce potentiel peut croître sous l'action de synapses excitatrices ou décroître sous l'action de synapses inhibitrices. Lorsque le potentiel atteint un certain seuil, des mécanismes de rétroactions internes aux neurones sont activés. Ces mécanismes font à leur tour varier le potentiel de membrane pour que celui-ci forme une impulsion appelée *potentiel d'action* et dont la forme est indépendante de l'activité des synapses. Cette variation de potentiel se propage à la manière d'une onde le long de l'axone vers les terminaux où se situent les synapses de sortie dont le rôle est de transmettre l'information de l'activité à d'autres neurones. La forme d'onde du potentiel d'action se décompose en trois périodes : une période de dépolarisation, une période de repolarisation et une période d'hyperpolarisation (Figure 4-2).

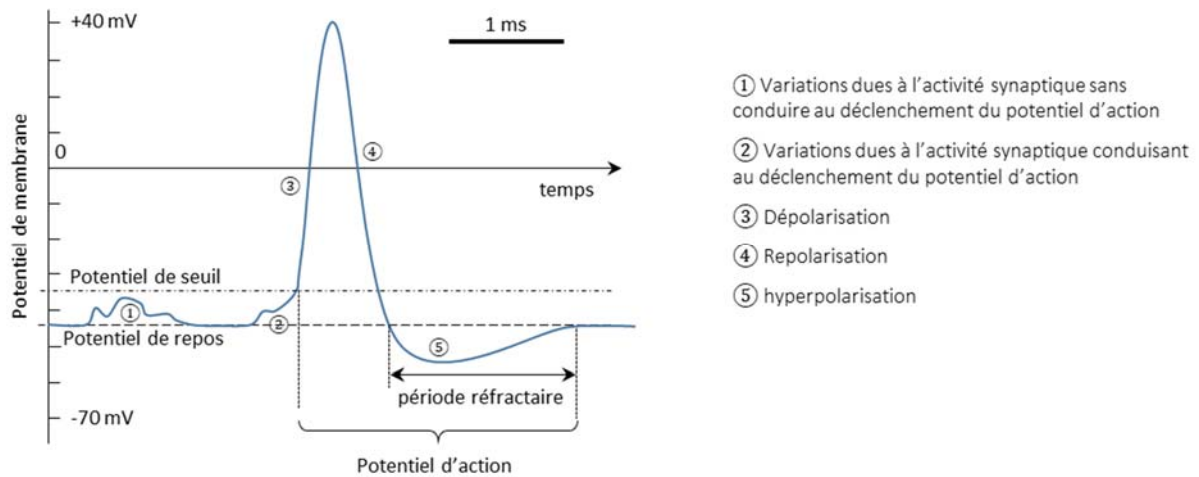


Figure 4-2 – Vue schématique de la forme du potentiel de membrane au cours du temps sous l'effet d'activités synaptiques qui conduisent ou non au déclenchement du potentiel d'action.

4.2 Le modèle de neurone « leaky, Integrate and Fire »

Le modèle le plus simple de neurone impulsionnel est le modèle d'*intégration à fuite et déclenchement* (« leaky integrate-and-fire » : LIF) [2], [3]. Ce modèle décrit la dynamique temporelle du potentiel de membrane V_{MEM} en fonction du courant synaptique I_{SYN} sans détailler les mécanismes de rétroaction responsables de la génération du potentiel d'action. Il se contente d'introduire une discontinuité au moment où V_{MEM} atteint le potentiel de seuil V_{TH} , qui est alors réinitialisé au potentiel de repos V_{RST} :

$$\text{si } V_{MEM} > V_{TH}, \text{ alors } V_{MEM} \leftarrow V_{RST} \quad (4.1)$$

Au cours du temps, V_{MEM} varie en présence d'un courant synaptique I_{SYN} et décroît en son absence. Le courant I_{SYN} est la somme des courants provenant de chacune des synapses d'entrée connectées au neurone. Les dynamiques associées aux variations de V_{MEM} sont gouvernées par la *capacité de membrane* C_{MEM} et la *conductance de fuite* G_{LK} selon l'équation différentielle :

$$C_{MEM} \cdot \frac{dV_{MEM}}{dt} = I_{SYN} - G_{LK} \cdot (V_{MEM} - V_{RST}) \quad [A] \quad (4.2)$$

Cette équation correspond au schéma électrique de la Figure 4-3.a. A titre d'exemple, les courbes typiques $V_{MEM} = f(t)$ établies selon (4.1) et (4.2) sont tracées pour un courant I_{SYN} constant à la Figure 4-3.c, et pour un courant I_{SYN} sinusoïdal à la Figure 4-3.e (la valeur minimale du courant sinusoïdal est non-nulle).

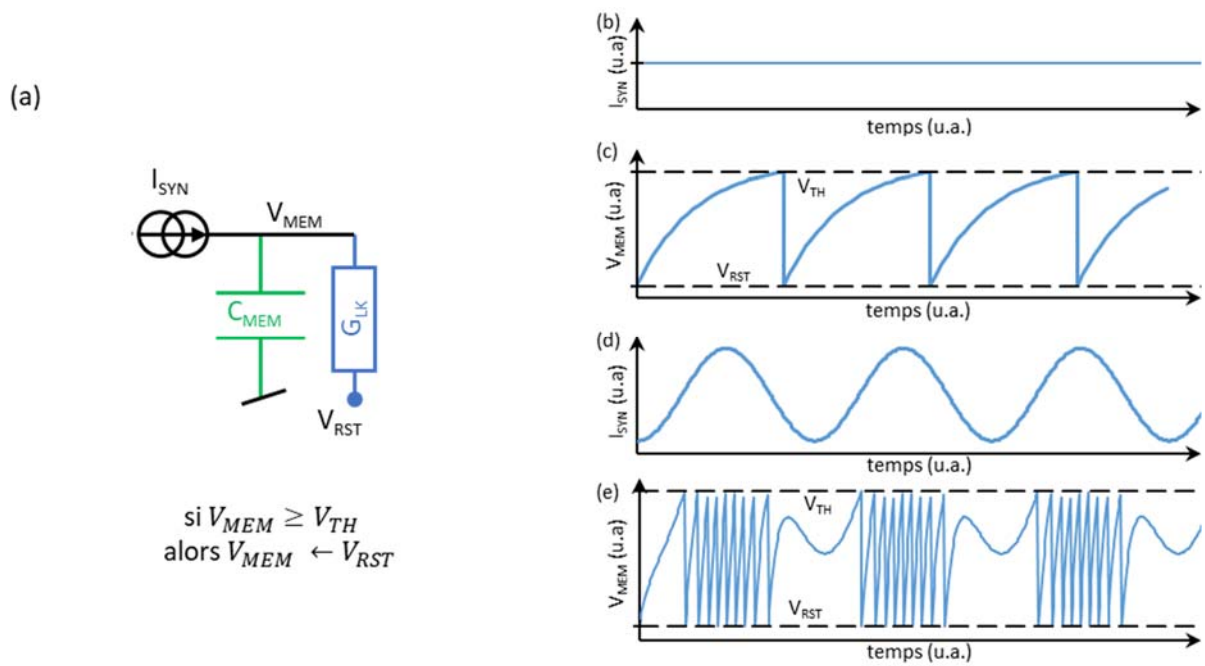


Figure 4-3 – (a) Schéma électrique et équation conditionnelle correspondants au modèle de neurone LIF. (b) Chronogrammes d'un courant I_{SYN} constant et du potentiel V_{MEM} correspondant prédit par le modèle LIF. (c) Chronogrammes d'un courant I_{SYN} sinusoïdal et du potentiel V_{MEM} correspondant prédit par le modèle LIF.

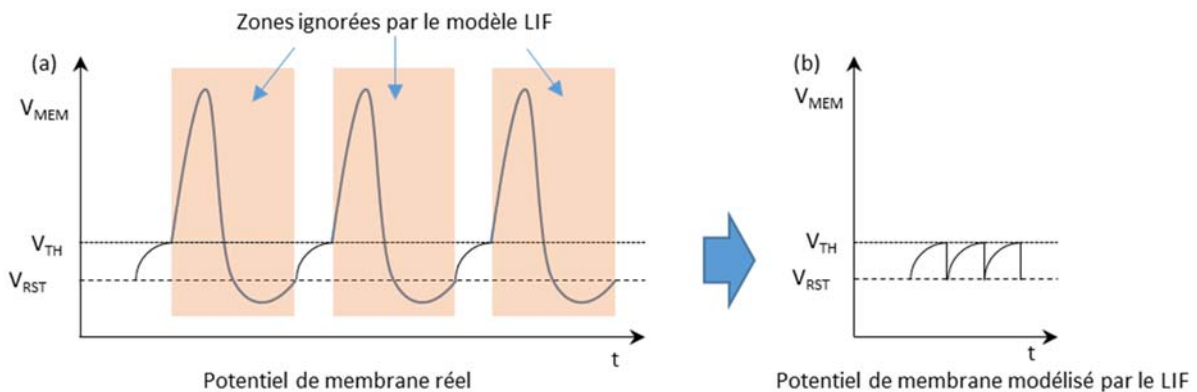


Figure 4-4 – Vues schématiques du potentiel de membrane au cours du temps (a) d'un neurone biologique et (b) d'un neurone simulé par le modèle LIF, dans le cas d'une excitation synaptique constante.

D'un certain point du vue, le modèle LIF rend compte uniquement des variations du potentiel de membrane avant le déclenchement d'un potentiel d'action (zones ① et ② de la Figure 4-2) (Figure 4-4). Par ailleurs, le schéma électrique de la Figure 4-3.a n'implémente pas l'évènement de réinitialisation. Pour pallier cette lacune, le schéma peut être complété en lui ajoutant une bascule à seuil (bascule de Schmitt) et un interrupteur commandé en tension, connectés tel que décrit à la Figure 4-5. Le seuil haut de la bascule de Schmitt est V_{TH} et son seuil bas est V_{RST} . En connectant ces deux éléments au reste du circuit, la fonction de réinitialisation est dès lors réalisée par des fonctions électroniques. De plus, la tension de sortie de la bascule est maintenant l'équivalent du potentiel d'action qui est transmis le long de l'axone jusqu'aux synapses de sortie connectées au neurone. Dans ce contexte le nœud électrique de sortie de la bascule devient le nœud de sortie du neurone. Il prend le nom de nœud d'axone, noté n_{AXO} , et sa tension est notée V_{AXO} .

Le bloc constitué par la conductance G_{LK} , la bascule de Schmitt et l'interrupteur commandé en tension est nommé *circuit de déclenchement à fuite* (« leaky trigger circuit » : LTC). Ainsi, le circuit électronique du neurone impulsif LIF est obtenu par l'assemblage d'une capacité et d'un LTC.

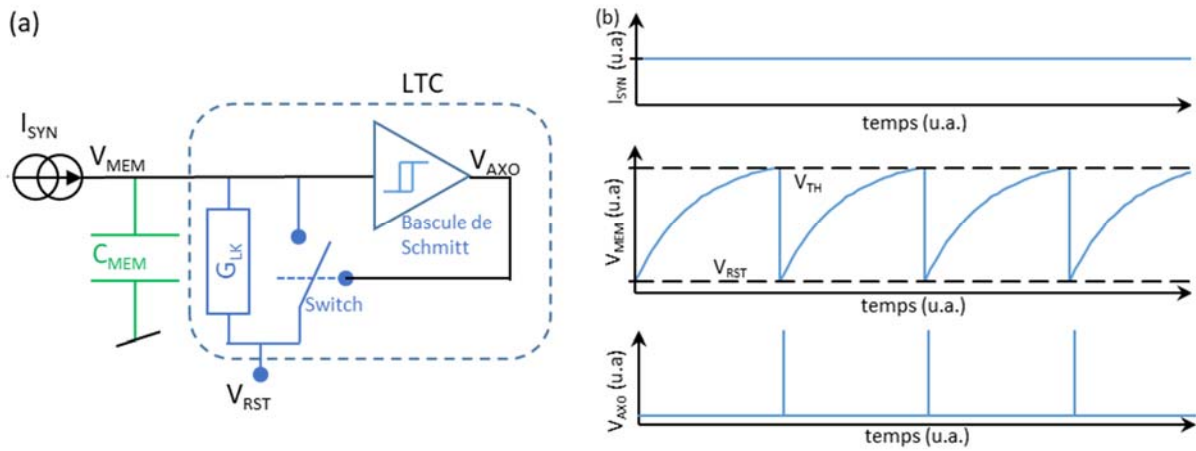


Figure 4-5 – (a) Schéma du circuit qui implémente entièrement le modèle LIF. (b) Chronogrammes du courant I_{SYN} et des potentiels V_{MEM} et V_{AXO} , pour un courant I_{SYN} constant, obtenus selon le schéma exposé en (a).

Dans le LTC, la conductance de valeur G_{LK} peut être simplement implémentée par une résistance de valeur $R_{LK} = 1/G_{LK}$. L'interrupteur commandé en tension peut être implémenté par un MOSFET de type N. La bascule de Schmitt peut être obtenue selon différents montages, basés sur des amplificateurs opérationnels [4], ou des circuits bipolaires ou CMOS plus simples [5]–[14], comprenant généralement 8 composants (Figure 4-6).

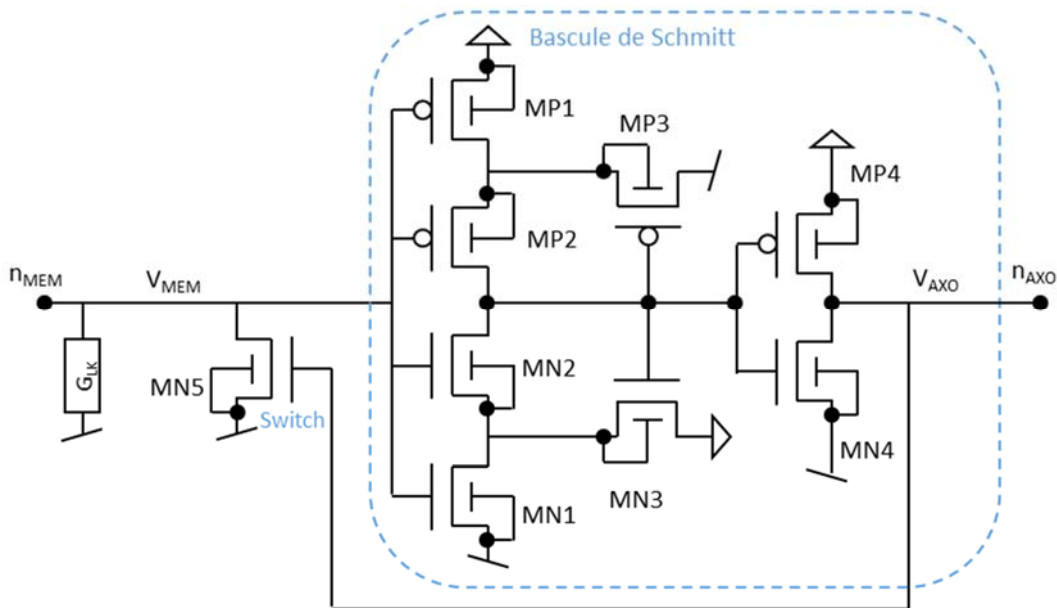


Figure 4-6 – Schéma électrique d'un LTC utilisant une bascule de Schmitt typique implémentée en technologie CMOS.

4.3 Le circuit de déclenchement à fuite basé sur le BIMOS (BB-LTC)

Une des contributions majeures du travail présenté dans cette thèse est de montrer qu'il est possible d'implémenter la fonction électronique du LTC en utilisant un montage comprenant uniquement un FD-SOI BIMOS et un FD-SOI NMOS : le LTC basé sur le BIMOS (« BIMOS-based LTC » : BB-LTC). Dans un

tel montage, le collecteur du BIMOS est connecté au nœud de membrane n_{MEM} , la source du NMOS, la base du BIMOS et la grille-avant du BIMOS sont connectées au nœud d'axone n_{AXO} . Le potentiel de grille-avant du NMOS est fixé à un potentiel V_{FG}^N dont la valeur contrôle les potentiels de seuil haut et bas de la fonction de bascule de Schmitt (voir plus loin, section 4.3.5.1). Les autres terminaux sont connectés à la masse. Le schéma électrique correspondant et un exemple typique de layout sont présentés à la Figure 4-7 et à la Figure 4-8 respectivement.

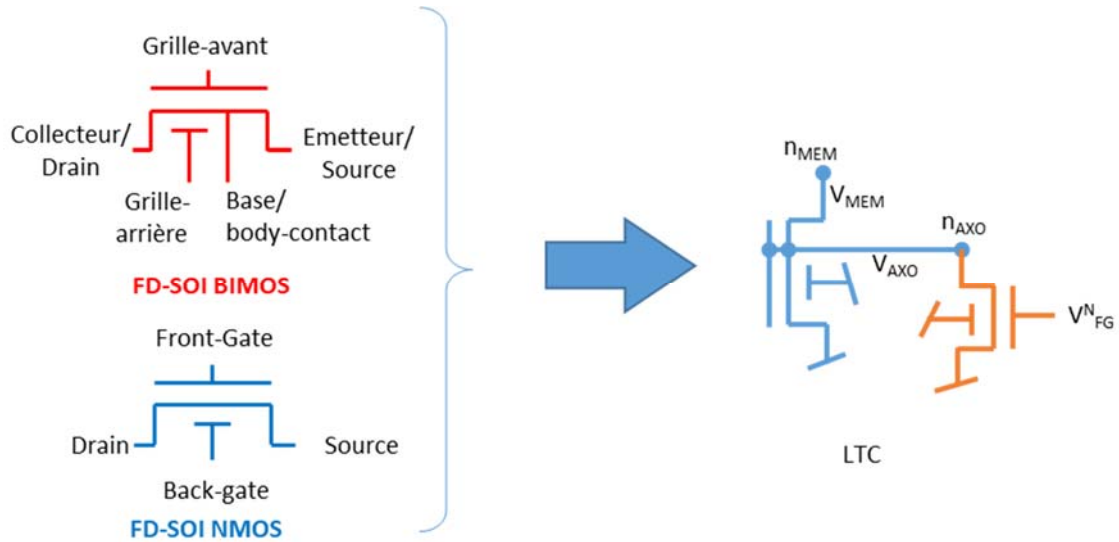


Figure 4-7 – Schéma électrique d'un BB-LTC composé d'un FD-SOI BIMOS et d'un FD-SOI NMOS.

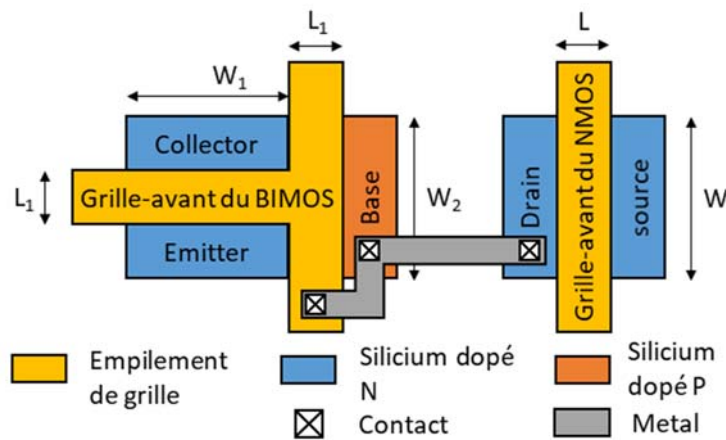


Figure 4-8 – Layout typique d'un BB-LTC avec les dimensions caractéristiques.

Dans le BB-LTC, la relation entre le potentiel de membrane V_{MEM} et le courant de collecteur I_C qui traverse le BIMOS a une courbe caractéristique tracée à la Figure 4-9.a pour un BB-LTC aux dimensions décrites au Tableau 4-1. Cette caractéristique présente un cycle d'hystérésis, que l'on peut décrire par les coordonnées des quatre points $(I_{OFF,RST} ; V_{RST})$, $(I_{OFF,TH} ; V_{TH})$, $(I_{ON,TH} ; V_{TH})$ et $(I_{ON,RST} ; V_{RST})$ qui définissent la fenêtre d'hystérésis (Figure 4-9.b). Ces coordonnées constituent les grandeurs caractéristiques du BB-LTC (Tableau 4-2).

Tableau 4-1 - Dimensions du BB-LTC utilisé pour la mesure de la Figure 4-9.a.

L_1 (nm)	W_1 (nm)	L_2 (nm)	W_2 (nm)	L (nm)	W (nm)	EOT (nm)	BIMOS	NMOS
135	270	270	350	135	350	3.4	RVT	RVT

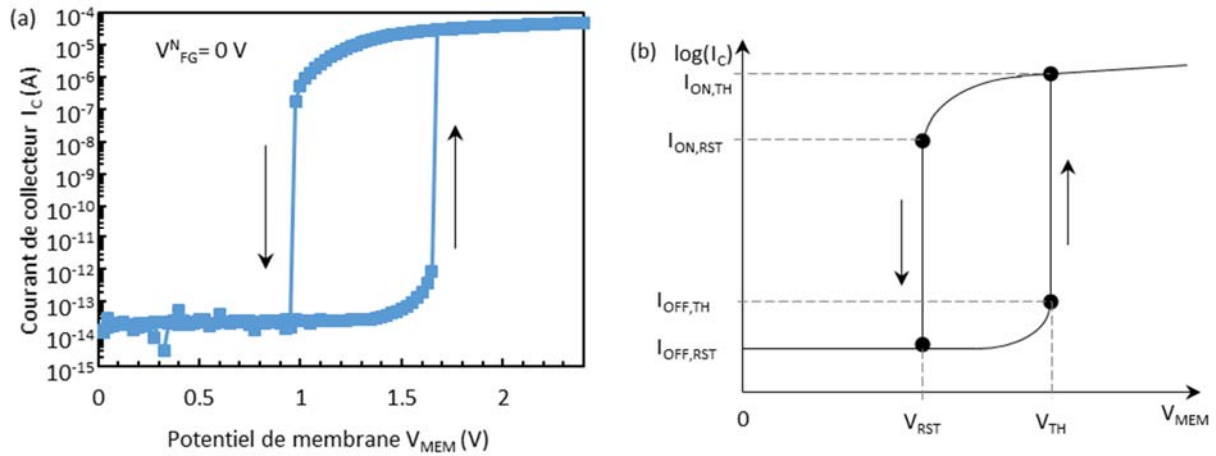


Figure 4-9 – (a) Caractéristique I_C - V_{MEM} mesurée à température ambiante sur BB-LTC décrit au Tableau 4-1. $V_E = V_S = V_{BG} = V_{FG}^N = 0$ V. (b) Vue schématique de la caractéristique I_C - V_{MEM} où sont repérées les coordonnées des points de la fenêtre d’hystérésis.

Tableau 4-2 – Paramètres de la fenêtre d’hystérésis extraits de la Figure 4-9.

V_{RST}	$I_{ON,RST}$	$I_{OFF,RST}$	V_{TH}	$I_{ON,TH}$	$I_{OFF,TH}$
962 ± 13 mV	175 nA	21 fA	1663 ± 13 mV	30.6 μ A	887 fA

Le comportement du BB-LTC peut être décrit ainsi :

- Lorsque V_{MEM} croît depuis une tension inférieure à V_{RST} , le courant I_C qui traverse le BIMOS est faible et indépendant de V_{MEM} . Il augmente ensuite exponentiellement avec V_{MEM} au voisinage de V_{TH} . Lorsque V_{MEM} atteint V_{TH} , le courant I_C augmente brusquement de plusieurs décades pour passer de la valeur $I_{OFF,TH}$ à $I_{ON,TH}$. Ensuite, I_C augmente linéairement avec V_{MEM} .
- Lorsque V_{MEM} décroît depuis une tension supérieure à V_{TH} , le courant I_C est fort et décroît linéairement avec V_{MEM} . Il diminue ensuite exponentiellement au voisinage de V_{RST} . Lorsque V_{MEM} atteint V_{RST} , le courant I_C chute brusquement de plusieurs décades pour passer de $I_{ON,RST}$ à $I_{OFF,RST}$.

En s’appuyant sur la caractéristique I_C - V_{MEM} , les 3 sections suivantes exposent en quoi le BB-LTC implémente les 3 fonctions qui définissent le LTC : la fonction de bascule de Schmitt (section 4.3.1), celle de conductance G_{LK} (section 4.3.2) et celle d’interrupteur (section 4.3.3).

4.3.1 Fonction de bascule de Schmitt du BB-LTC

Les caractérisations effectuées sur le BIMOS au chapitre 3 ont montré que le courant de collecteur (ou de drain, selon la terminologie utilisée) est majoritairement constitué du courant de canal I_{CH} de sa structure MOSFET intrinsèque. Ce courant I_{CH} dépend principalement de la tension grille-avant – émetteur, en particulier lorsque le BIMOS est en régime sous le seuil, ou en régime de saturation. Les variations brutales de la conductance du BIMOS observées dans le montage BB-LTC à la Figure 4-9 s’expliquent donc par une variation brutale de son potentiel de grille-avant pendant le balayage effectué sur son potentiel de collecteur. Pour rappel, dans le BB-LTC, le potentiel de grille-avant du BIMOS est égal au potentiel d’axone V_{AXO} tel que l’illustre le schéma à la Figure 4-7.

Il est possible d’extraire les valeurs du potentiel d’axone V_{AXO} pendant le balayage de V_{MEM} en s’appuyant sur la mesure quasi-statique et à température ambiante du courant de collecteur I_C du BIMOS lors d’un balayage sur les potentiels de grille-avant V_{FG} et de base V_B égaux entre eux (configuration constatée dans le BB-LTC), et pour différentes valeurs du potentiel de collecteur V_C

(Figure 4-10). Dans une telle configuration, $V_{AXO} = V_{FG} = V_B$ et $V_{MEM} = V_C$. L'extraction de $V_{AXO} = f(V_{MEM})$ s'effectue en faisant correspondre à chaque point $I_C = f(V_{MEM})$ de la courbe de la Figure 4-9, un point $I_C = f(V_{AXO})$ de la courbe $V_C = V_{MEM}$ de la Figure 4-10, tel que les deux courants I_C sont égaux.

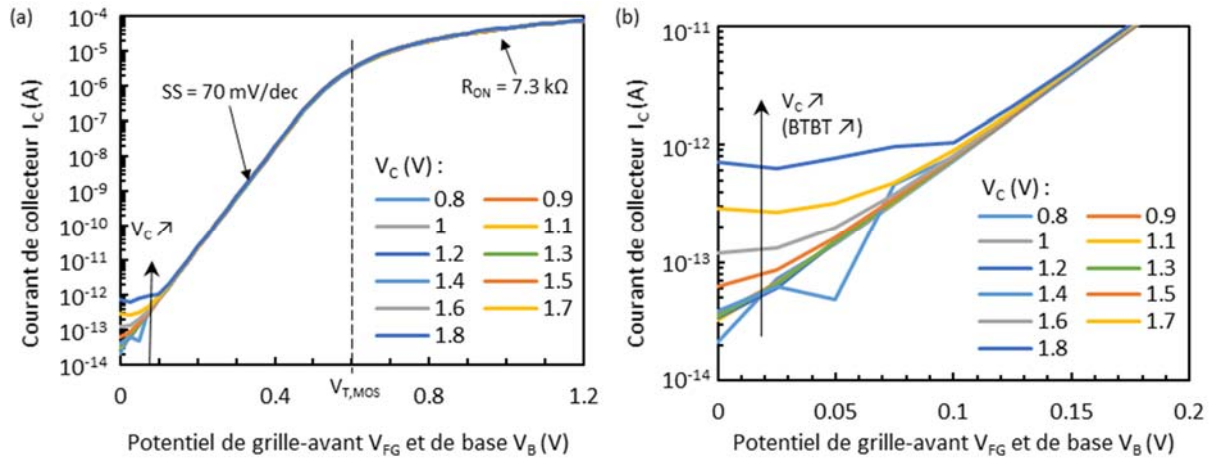


Figure 4-10 – Mesures quasi-statiques à température ambiante, sur un BIMOS de référence, du courant de collecteur I_C en fonction du potentiel d'axone V_{AXO} , à différents potentiels de membrane V_{MEM} . $V_{FG} = V_B = V_{AXO}$, $V_C = V_{MEM}$ et $V_E = V_{BG} = 0$ V. (a) Plage de balayage entière. (b) Zoom sur la plage $V_{MEM} \in [0 ; 0.2]$ V.

Remarquons ici que les courbes de la Figure 4-10 ne sont autres que les caractéristiques I_D - V_{FG} à différents V_D du BIMOS connecté en mode hybride ($V_{FG} = V_B$). On retrouve les croissances de I_D exponentielle ($SS = 70$ mV/dec) et linéaire ($R_{ON} = 7.3$ kΩ) selon que V_{FG} est inférieure ou supérieure à $V_{T,MOS} = 600$ mV (calculé par l'extraction du maximum de la dérivée de la transconductance [15]). L'augmentation de I_D à V_D faible ($V_D < 0.2$ V) est due à l'accroissement du BTBT à mesure que la tension drain – grille-avant V_{D-FG} augmente (phénomène de GIDL).

La courbe $V_{AXO} = f(V_{MEM})$ obtenue est tracée à la Figure 4-11, pour $V_{MEM} \in [0.8 ; 1.8]$ V (On suppose par extrapolation que $V_{AXO} = 0$ pour $V_{MEM} < 0.8$ V). La présence des barres d'erreur verticales de ± 25 mV résulte du pas de mesure utilisé à la Figure 4-10. On retrouve un cycle d'hystérésis similaire à celui observé sur la courbe $I_C(V_{MEM})$:

- Lorsque V_{MEM} croît depuis une valeur inférieure à V_{RST} , le potentiel d'axone V_{AXO} est nul. Il augmente linéairement avec V_{MEM} au voisinage de V_{TH} et à mesure qu'il s'en rapproche. Lorsque V_{MEM} atteint V_{TH} , V_{AXO} augmente brutalement de 0.1 à 0.875 V. Il augmente ensuite linéairement avec V_{MEM} .
- Lorsque V_{MEM} décroît depuis une valeur supérieure à V_{TH} , le potentiel d'axone est haut et diminue linéairement avec V_{MEM} . Lorsque V_{MEM} atteint V_{RST} , V_{AXO} chute brutalement de 0.5 à 0 V. V_{AXO} est ensuite nul quel que soit V_{MEM} .

La caractéristique $V_{AXO} = f(V_{MEM})$ est une fonction de transfert typique d'une bascule de Schmitt, à l'exception des variations du potentiel de sortie V_{AXO} en fonction du potentiel d'entrée V_{MEM} observées à l'état ON et au voisinage de V_{TH} à l'état OFF. Les paramètres caractéristiques de la fenêtre d'hystérésis extrait de la Figure 4-11 sont rassemblés au Tableau 4-3.

Tableau 4-3 – Paramètres de la fenêtre d'hystérésis extraits de la Figure 4-11.

V_{RST}	$V_{AXO,ON,RST}$	$V_{AXO,OFF,RST}$	V_{TH}	$V_{AXO,ON,TH}$	$V_{AXO,OFF,TH}$
975 ± 25 mV	500 ± 50 mV	0 ± 50 mV	1675 ± 25 mV	875 ± 50 mV	100 ± 50 mV

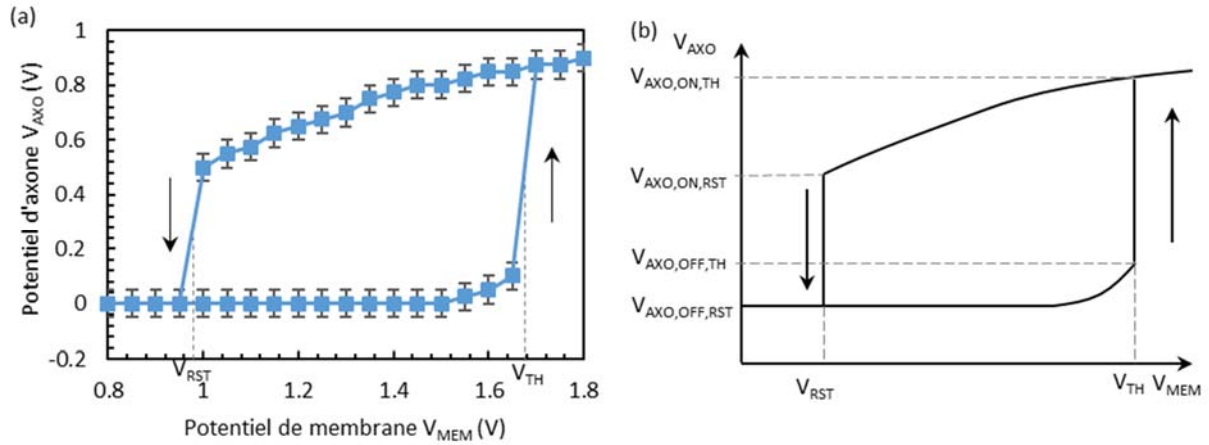


Figure 4-11 – (a) Potentiel d'axone V_{AXO} en fonction du potentiel de membrane V_{MEM} dans le BB-LTC, extrait à partir des données I_C des Figure 4-9 et Figure 4-10. (b) Vue schématique de la caractéristique V_{AXO} - V_{MEM} , où sont repérées les coordonnées des points de la fenêtre d'hystérésis.

4.3.2 Fonction de conductance du BB-LTC

A l'état OFF, un faible courant I_{OFF} traverse le BB-LTC. Comme $V_{FG} = V_{AXO} = 0$ V, il s'agit du courant de fuite de la structure MOSFET intrinsèque au BIMOS. La conductance G_{LK} associée à ce courant est donnée par l'expression :

$$G_{LK}(V_{MEM}) = \frac{I_{OFF}(V_{MEM})}{V_{MEM}} \quad [S] \quad (4.3)$$

Comme I_{OFF} ne croît pas linéairement avec V_{MEM} , G_{LK} n'est pas constant. En effet, I_{OFF} connaît deux régimes de variation avec le potentiel V_{MEM} : I_{OFF} est constant pour $V_{MEM} \ll V_{TH}$ et croît fortement lorsque V_{MEM} se rapproche de V_{TH} (Figure 4-9). La variation de G_{LK} est ainsi décroissante sur $V_{MEM} \in [0 ; 1.35]$ V et croissante sur $V_{MEM} \in [1.35 ; V_{TH} = 1.65]$ V (Figure 4-12).

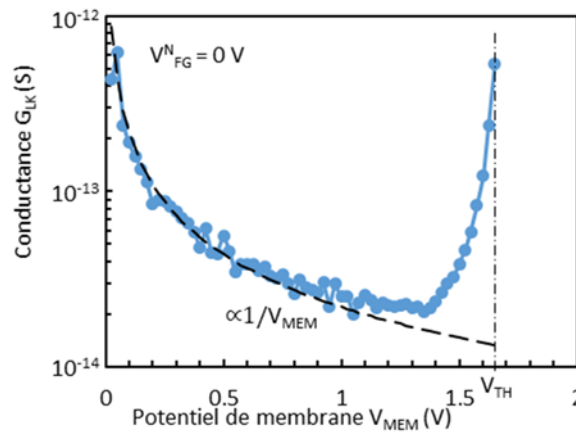


Figure 4-12 – Extraction de G_{LK} à partir des données de la Figure 4-9 et selon l'équation (4.3).

Les caractérisations du BIMOS en mode MOSFET effectuées au chapitre 3 section 3.4 permettent d'identifier la nature et le trajet du courant I_{OFF} dans ces deux régimes :

- Lorsque $V_{MEM} \ll V_{TH}$, I_{OFF} est constitué du courant de canal I_{CH} de la structure MOSFET intrinsèque au BIMOS en régime profondément sous le seuil ($V_{FG} = V_{AXO} = 0$ V). Ce courant circule du drain vers la source du BIMOS (i.e. du collecteur vers l'émetteur). Comme attendu pour ce type de courant dans ce régime de fonctionnement, I_{CH} est indépendant du potentiel

de drain V_{MEM} dès que la tension drain – source est supérieure à quelques kT/q , (i.e. $V_{MEM} \gtrsim 100$ mV). Dans ces conditions, la conductance décroît comme $1/V_{MEM}$ (ligne en pointillé de la Figure 4-12).

- Lorsque V_{MEM} se rapproche de V_{TH} , I_{OFF} est principalement constitué d'un courant de génération par BTBT I_{BTBT} qui circule du drain vers le body-contact (i.e. du collecteur vers la base). Ce type de courant est responsable du GIDL dans les structures MOSFET standard. Le courant de canal I_{CH} est toujours présent, mais il est négligeable face à I_{BTBT} . Pour rappel, la génération de paires électron/trou s'opère dans la région du film de silicium à la jonction drain – body et résulte de la présence d'un champ électrique intense sous l'effet de la haute tension drain – grille-avant $V_{D-FG} = V_{MEM} - V_{AXO}$. Les électrons générés rejoignent le drain (i.e. le collecteur) tandis que les trous rejoignent le body-contact (i.e. la base).

Notons dès à présents que le courant I_{BTBT} , activé lorsque la tension $V_{MEM} - V_{AXO}$ devient suffisamment importante, joue un rôle clef dans le mécanisme de bascule du BB-LTC. Ce rôle sera décrit à la section 4.3.4.

Grâce à la technologie FD-SOI, la conductance G_{LK} peut être modulée par la polarisation de la grille arrière du BIMOS V_{BG}^B . Mais cette polarisation a aussi pour effet de modifier d'autres grandeurs caractéristiques du BB-LTC, comme V_{TH} ou V_{RST} . Une étude de l'influence de V_{BG}^B sur le fonctionnement du BB-LTC est présenté plus loin à la section 4.3.5.3.

4.3.3 Fonction d'interrupteur du BB-LTC

Lorsque le BB-LTC est à l'état ON, un fort courant I_{ON} traverse le BIMOS, supérieur à I_{OFF} de plusieurs décades. Cela témoigne de la haute conductivité du BIMOS dans cet état. I_{ON} est alors constitué du courant de canal I_{CH} de la structure MOSFET intrinsèque au BIMOS, qui est polarisé en régime au-dessus du seuil ($V_{FG} = V_{AXO} \geq V_{T,MOS}$, Figure 4-11). Le BIMOS implémente ainsi une fonction d'interrupteur : une fois que le BB-LTC a basculé à l'état ON, le nœud électrique où est connecté le collecteur du BIMOS (ici, n_{MEM}) est en contact électrique avec le nœud où est connecté l'émetteur (ici, la masse). Ce contact est caractérisé par sa conductance G_{ON} qui est évaluée par une expression similaire à (4.3) :

$$G_{ON}(V_{MEM}) = \frac{I_{ON}(V_{MEM})}{V_{MEM}} \quad [S] \quad (4.4)$$

G_{ON} est extraite des données de la Figure 4-9 et tracée à la Figure 4-13. Elle est supérieure à G_{LK} d'environ 7 décades.

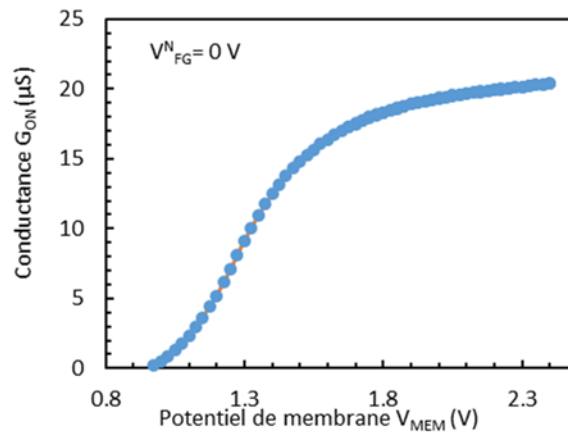


Figure 4-13 – Extraction de G_{ON} à partir des données de la Figure 4-9 et selon l'équation (4.4).

4.3.4 Principes de fonctionnement du BB-LTC

Les sections précédentes ont montré par la mesure en quoi un FD-SOI BIMOS connecté à un FD-SOI NMOS selon le schéma de la Figure 4-7 implémente un circuit de déclenchement à fuite (LTC). Cette section décrit les mécanismes physiques mis en jeu dans le circuit qui sont responsables d'un tel fonctionnement.

Pour ce faire, la courbe caractéristique I_C - V_{MEM} de la Figure 4-9 est décomposée en 5 régions distinctes (Figure 4-14.a). Pour chacune de ces régions, le BB-LTC est représenté de façon schématique (Figure 4-14.b). L'émetteur, le collecteur, la base et le body du BIMOS sont représentés par une vue en coupe horizontale de son film de silicium. La grille-avant du BIMOS est représentée par le symbole électronique d'une armature de capacité, située sur la droite de la vue en coupe du film de silicium, tournée vers le bord de la région de body et connectée à la base. Le NMOS est représenté par son symbole électronique. Aucune des grille-arrières n'est représentée. Cette manière de schématiser le BB-LTC permet d'y faire figurer les différentes contributions aux courants qui circulent dans le film de silicium du BIMOS.

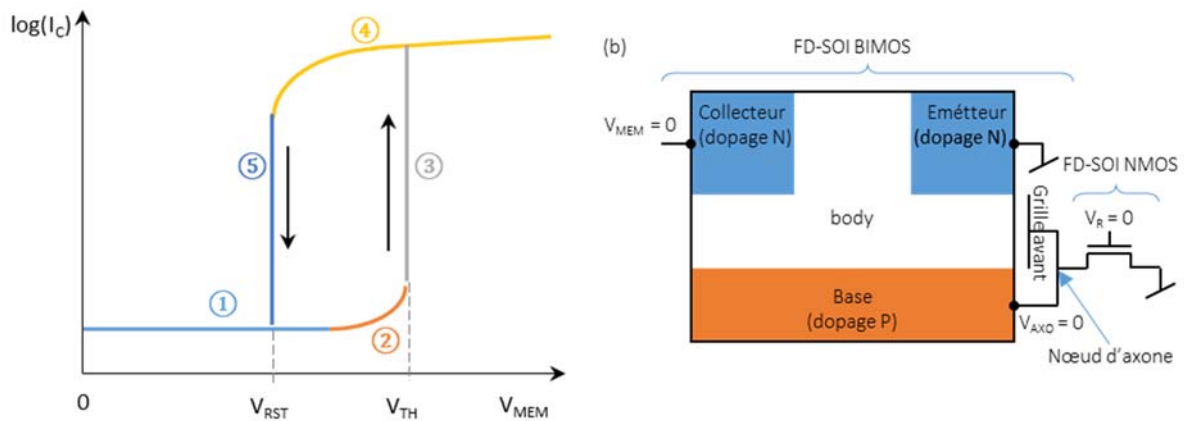


Figure 4-14 – (a) Décomposition de la courbe caractéristique I_C - V_{MEM} en 5 régions différentes. (b) Vue schématique du BB-LTC lorsque $V_{MEM} = 0$ V mettant en évidence la vue en coupe horizontale du film de silicium du FD-SOI BIMOS.

Lorsque $V_{MEM} = 0$, et que l'équilibre thermodynamique est vérifié, aucune charge ne circule dans le BB-LTC et le potentiel du nœud d'axone est nul ($V_{AXO} = 0$ V).

- **Région 1**

L'augmentation de V_{MEM} polarise, d'une part, la structure de gated-diode constituée par l'ensemble collecteur – body – base – grille-avant et, d'autre part, la structure MOSFET intrinsèque au BIMOS (la faiblesse du gain en émetteur commun du BJT intrinsèque au BIMOS, démontrée au chapitre 3, permet d'ignorer les effets liés à cette structure dans l'analyse).

La polarisation de la gated-diode est faite en inverse : un flux d'électrons circule depuis la base vers le collecteur et un flux de trous circule en sens inverse. Comme ces porteurs sont très fortement minoritaires dans leurs régions d'origine (les régions sont très fortement dopées), le courant associé à leurs flux est très faible et peut être considéré comme nul.

La polarisation de la structure MOSFET fait circuler un flux d'électrons depuis l'émetteur vers le collecteur. Le courant associé, qui circule en sens inverse, est le courant de canal. Comme $V_{AXO} = 0$ V, le BIMOS est en régime sous le seuil : le courant de canal est un courant de diffusion dont l'intensité est gouvernée par la tension entre l'émetteur et la grille-avant (cf . équation (2.53)). En l'absence

d'apport ou de flux de charge dans le nœud d'axone, V_{AXO} demeure constant sur toute la région 1 de la courbe $I_C - V_{MEM}$. Plus précisément, $V_{AXO} = V_E = 0$ V. Par conséquent, le courant de canal est constant et, comme il est la seule contribution au courant de collecteur, ce dernier l'est aussi.

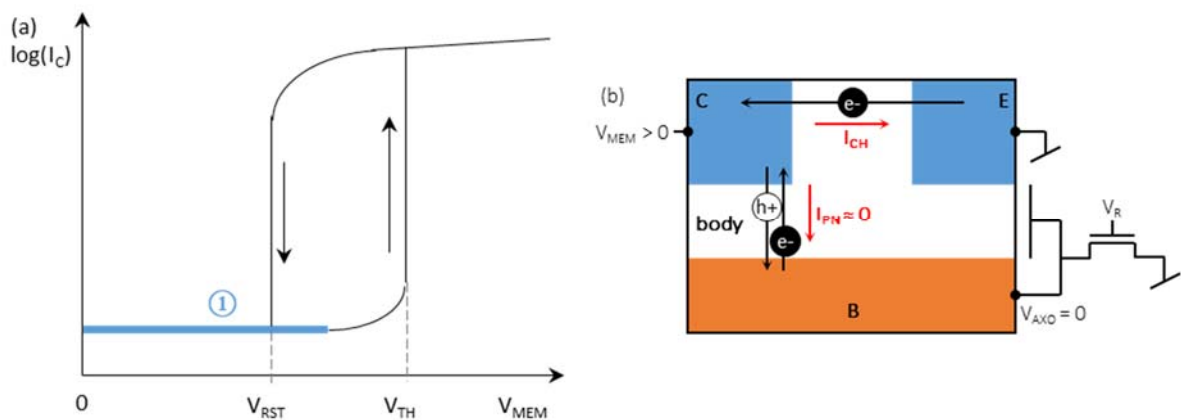


Figure 4-15 – (a) Vue schématique de la courbe $I_C - V_{MEM}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 1 (ligne bleu). (b) Vue schématique des courants qui circulent dans le BB-LTC.

• Région 2

Lorsque V_{MEM} atteint une valeur suffisamment importante, le champ électrique créé par la tension collecteur – grille-avant à la frontière entre le collecteur et le body devient suffisamment intense pour générer des paires électron-trou par BTBT. Les électrons rejoignent le collecteur qui est au potentiel le plus haut et les trous rejoignent la base. Bien que l'émetteur et la base sont tous deux à des potentiels bas, les trous circulent vers la base plutôt que vers l'émetteur à cause de la différence de type de dopants présents dans chacune des régions. Ces flux de porteurs constituent le courant I_{BTBT} .

Lorsqu'ils atteignent le contact métallique du nœud d'axone n_{AXO} , les trous arrivés jusqu'à la base se recombinent avec des électrons fournis par le contact. Une charge positive apparaît dans le nœud d'axone et le potentiel V_{AXO} augmentent en conséquence. L'augmentation du potentiel d'axone V_{AXO} a pour effet de polariser le drain du NMOS, la base du BIMOS et la grille-avant du BIMOS.

La polarisation du drain du NMOS fait circuler un courant de drain I_D à travers lui. Ce courant est un courant de canal de MOSFET sous le seuil ($V_{FG}^N = 0$ V) régi par l'équation (2-53) : c'est un courant de diffusion qui ne dépend que de la tension grille-avant – source dès que la tension drain-source dépasse quelques kT/q . Ce courant évacue les charges accumulées dans n_{AXO} , et fait décroître V_{AXO} .

La polarisation de la grille-avant de BIMOS augmente le courant de canal par effet MOSFET et fait décroître la tension collecteur – grille-avant du BIMOS. Cela diminue l'intensité du champ électrique dans le film du silicium à la frontière entre le collecteur et le body. Le taux de génération de paires électron-trou par BTBT diminue, ainsi que l'intensité de I_{BTBT} qui en résulte.

L'augmentation du potentiel de la base polarise en direct la structure de gated-diode constituée par la base, la grille-avant du BIMOS et l'émetteur. Un courant I_{PN} apparaît, constitué de trous circulant de la base vers l'émetteur et d'électrons circulant en sens inverse. Les mesures du chapitre 3 section 3.3 ont montré que ce courant est très faible tant que V_{AXO} n'est pas suffisamment fort. Dans la région 2 décrite ici, ce courant est négligeable face à I_D et I_{BTBT} .

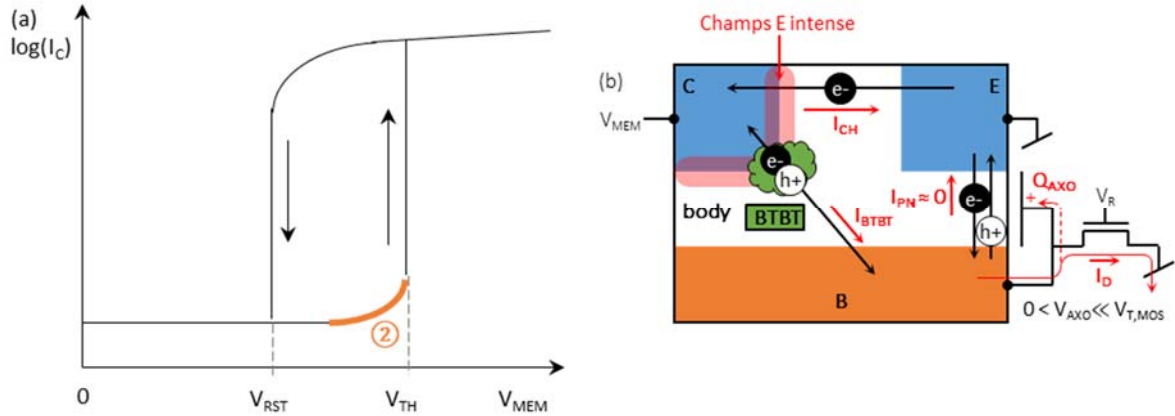


Figure 4-16 – (a) Vue schématique de la courbe $I_C - V_{MEM}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 2 (ligne orange). (b) Vue schématique des courants qui circulent dans le BB-LTC.

Finalement, un état stationnaire est atteint où $V_{AXO} > 0$, $I_{BTBT} = I_D$ et une charge Q_{AXO} est stockée dans le nœud d'axone. Autrement dit, dans la région 2, la relation $V_{AXO} = f(V_{MEM})$ tracée à la Figure 4-11 est gouvernée par l'égalité $I_{BTBT} = I_D$.

Les grandeurs Q_{AXO} et V_{AXO} sont par ailleurs reliées par la capacité de nœud d'axone C_{AXO} selon la relation $Q_{AXO} = C_{AXO} \cdot V_{AXO}$. La valeur de C_{AXO} n'est pas constante en fonction de V_{AXO} . En effet, C_{AXO} est la somme des capacités parasites entre le nœud d'axone et les autres nœuds du circuit, toutes mises en parallèle (capacités drain – source, drain – grille-avant et drain – grille-arrière du NMOS, capacités grille-avant – collecteur, grille-avant – émetteur, grille-avant – grille-arrière, base – collecteur, base – émetteur, base – grille-arrière du BIMOS et la capacité du métal qui relie les terminaux du BIMOS et du NMOS entre eux). Or, les capacités des structures MOS ne sont pas constantes en fonction des tensions qui leurs sont appliquées (voir chapitre 2 section 2.2.3).

Enfin, insistons ici sur le fait que c'est la relation $\partial I_{BTBT} / \partial V_{AXO} < 0$ (qui n'est autre que la relation $\partial I_{BTBT} / \partial V_{FG} < 0$ observée lors des caractérisations du BIMOS en mode MOSFET au chapitre 3 section 3.4.1) qui permet au système d'atteindre un état stationnaire.

• Région 3

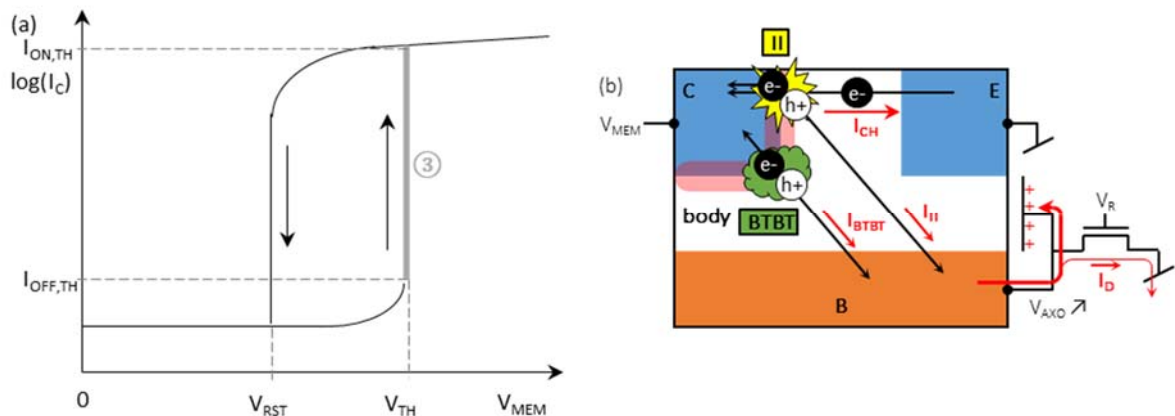


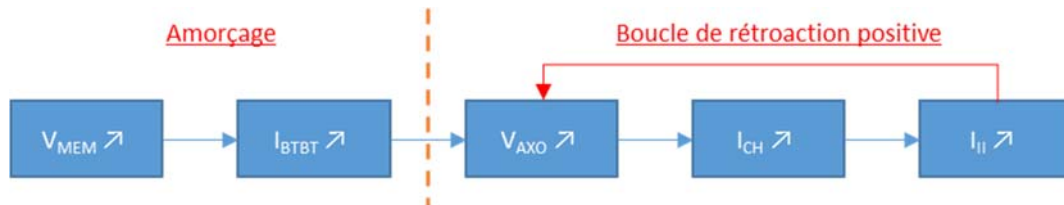
Figure 4-17– (a) Vue schématique de la courbe $I_C - V_{MEM}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 3 (ligne grise). (b) Vue schématique des courants qui circulent dans le BB-LTC.

La région 3 correspond à la transition entre l'état OFF et l'état ON du BB-LTC, atteinte pour $V_{MEM} = V_{TH}$. La configuration des courants dans cette région diffère de celle de la région 2 par l'apparition du courant de génération par ionisation par impact I_{II} . (En réalité, le courant I_{II} est déjà présent dans la région 2, mais sa contribution est négligeable face à I_{BTBT} : $I_{II} \ll I_{BTBT}$. Dans la région 3, l'inégalité s'inverse $I_{II} \gg I_{BTBT}$). Tout comme le courant I_{BTBT} , le courant I_{II} circule du collecteur vers la base (les trous générés rejoignent la base et les électrons rejoignent le collecteur). Ce nouveau courant I_{II} augmente l'apport de charge dans le nœud d'axone n_{AXO} et le potentiel d'axone V_{AXO} augmente en conséquence. Or, contrairement au BTBT et tant que le potentiel de grille-avant $V_{FG} = V_{AXO}$ du BIMOS n'est pas trop haut, l' I_{II} augmente avec V_{FG} : $\partial I_{II} / \partial V_{FG} > 0$ (c.f. chapitre 3 section 3.4.1). Il s'en suit un phénomène d'emballement qui fait croître V_{AXO} indépendamment du potentiel V_{MEM} . Notons ici que le courant de drain I_D du NMOS qui décharge le nœud d'axone n_{AXO} est limité par la tension grille-avant – source qui est fixée. I_D ne peut donc plus évacuer les nouvelles charges apportées par I_{II} .

Comme la structure MOSFET intrinsèque au BIMOS est en régime sous le seuil, la croissance de V_{AXO} augmente le courant de canal I_{CH} de façon exponentielle et conduit à la très forte augmentation du courant I_C observée dans la caractéristique I_C - V_{MEM} du BB-LTC.

Pour rappel, l'augmentation de I_{II} est directement liée à l'augmentation de I_{CH} . En effet, ce sont les électrons dont le flux constitue I_{CH} qui ionisent par impact des atomes du cristal de silicium, avec un rendement α : $I_{II} = \alpha \cdot I_{CH}$. Or, ce rendement est relativement constant lorsque le MOSFET est sous le seuil, comme c'est le cas du BIMOS dans le BB-LTC avant qu'il soit passé à l'état ON.

La boucle de rétroaction positive responsable de l'emballement peut être décrite ainsi :



• Région 4

Dans la région 4, la tension grille-avant – émetteur du BIMOS est supérieure ou égale à la tension de seuil MOSFET du BIMOS : $V_{AXO} \geq V_{T,MOS}$. Le courant I_C mesuré est majoritairement constitué du courant de canal I_{CH} de forte intensité.

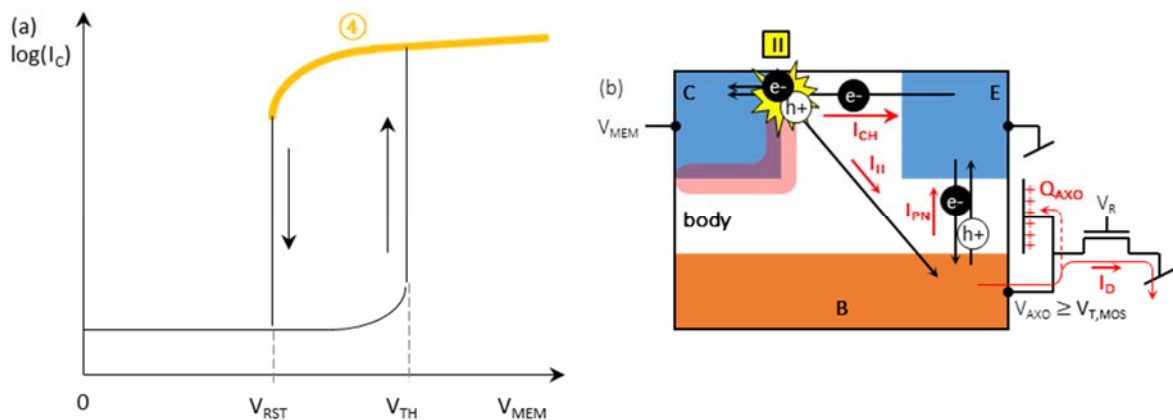


Figure 4-18 – (a) Vue schématique de la courbe $I_C - V_{MEM}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 4 (ligne jaune). (b) Vue schématique des courants qui circulent dans le BB-LTC.

La tension collecteur – grille-avant du BIMOS est désormais insuffisante pour générer des paires électron-trou par BTBT : $I_{BTBT} = 0$. Seul le courant I_{II} continue à fournir des charges au nœud n_{AXO} . La polarisation de la gated-diode base-émetteur est maintenant suffisamment importante pour faire circuler un courant I_{PN} significatif de la base vers l'émetteur, et face auquel le courant I_D qui traverse le NMOS est négligeable ($I_{PN} \gg I_D$). Ainsi, pour chaque valeur de V_{MEM} , un état stationnaire est atteint pour lequel $I_{II} = I_{PN}$ et auquel correspond un unique potentiel $V_{AXO} = f(V_{MEM})$. Autrement dit, les variations de V_{AXO} avec V_{MEM} sont maintenant déterminées par la modulation de I_{II} par V_{MEM} et par l'équilibrage $I_{II} = I_{PN}$. Rappelons par ailleurs que $\partial I_{II} / \partial V_{MEM} > 0$. (c.f chapitre 3 section 3.4.2).

• Région 5

La région 5 correspond à la transition entre l'état ON et l'état OFF du BB-LTC, atteinte pour $V_{MEM} = V_{RST}$. Dans cette configuration, la réduction du courant I_{II} , et donc de V_{AXO} , conduit à une situation où I_{PN} et I_{II} deviennent tous deux négligeables face à I_D . Dans ce contexte, la charge Q_{AXO} stockée dans n_{AXO} s'évacue par le drain du NMOS vers la masse et le potentiel V_{AXO} retombe à 0 V. Le courant de canal I_{CH} chute exponentiellement, depuis une valeur définie comme $I_{ON,RST}$ jusqu'à atteindre le niveau qu'il avait dans la région 1, qui définit la valeur $I_{OFF,RST}$. De ce point de vue, $I_{OFF,RST}$ est égale au courant de saturation du MOSFET intrinsèque au BIMOS lorsque son potentiel de grille est à 0 V.

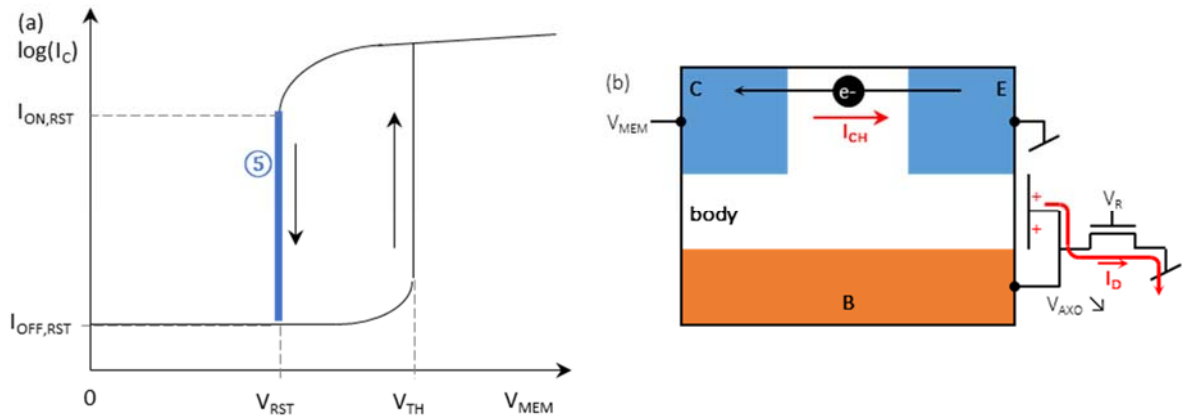


Figure 4-19 – (a) Vue schématique de la courbe $I_C - V_{MEM}$ du BB-LTC (ligne noire) et portion de la courbe correspondant à la région 5 (ligne bleue). (b) Vue schématique des courants qui circulent dans le BB-LTC.

• Synthèse

Pour synthétiser, l'état du BB-LTC est gouverné par le potentiel V_{AXO} du nœud d'axone n_{AXO} . Ce potentiel est lié à la charge Q_{AXO} stockée dans la capacité C_{AXO} de n_{AXO} selon la relation :

$$Q_{AXO} = C_{AXO} \cdot V_{AXO} \quad [C] \quad (4.5)$$

Les variations de Q_{AXO} sont liées aux courants qui circulent dans n_{AXO} par la loi de Kirchhoff (loi des nœuds) :

$$I_B - I_D - \frac{dQ_{AXO}}{dt} = 0 \quad [A] \quad (4.6)$$

où I_B est le courant de body-contact du BIMOS, compté positif lorsqu'il circule vers n_{AXO} , et I_D le courant de drain du NMOS, compté positif lorsqu'il circule vers le NMOS. En reprenant l'équation (4.5), l'équation (4.6) peut se réécrire comme :

$$I_B - I_D - C_{AXO} \cdot \frac{dV_{AXO}}{dt} = 0 \quad [A] \quad (4.7)$$

Le courant I_B se décompose en 3 courants distincts, qui dépendent de V_{MEM} et de V_{AXO} :

$$I_B(V_{AXO}, V_{MEM}) = I_{BTBT}(V_{AXO}, V_{MEM}) + I_{II}(V_{AXO}, V_{MEM}) - I_{PN}(V_{AXO}) \quad [A] \quad (4.8)$$

Le courant I_D est aussi une fonction de V_{AXO} :

$$I_D = I_D(V_{AXO}) \quad [A] \quad (4.9)$$

Dans les différentes régions de la courbe I_C - V_{MEM} , les courants prédominants diffèrent, selon l'état des polarisations. Selon les cas, ils conduisent à un état stable où V_{AXO} est constant (i.e. $dV_{AXO}/dt = 0$) (régions 1, 2, et 4) ou à un état instable où V_{AXO} varie au cours du temps (régions 3 et 5). Les caractéristiques principales de l'état du BB-LTC dans chacune des régions sont résumées au Tableau 4-4.

Tableau 4-4 – Caractéristiques principales de l'état du BB-LTC dans chacune des régions.

Région	Relation entre les courants prédominants dans n_{AXO}	$\frac{dQ_{AXO}}{dt}$	Potentiel d'axone V_{AXO}	Courant majoritairement constitutif de I_C
1	-	= 0	$V_{AXO} = 0$	I_{CH}
2	$I_{BTBT} = I_D$	= 0	$0 < V_{AXO} \ll V_{T,PN}$	I_{BTBT}
3	$I_{II} > I_D$	> 0	$0 < V_{AXO} < V_{T,PN}$	I_{CH}
4	$I_{II} = I_{PN}$	= 0	$V_{AXO} \geq V_{T,PN}$	I_{CH}
5	$I_{II} < I_D$	< 0	$0 < V_{AXO} < V_{T,PN}$	I_{CH}

En appliquant la théorie de la stabilité des systèmes dynamiques au potentiel V_{AXO} du BB-LTC, on montre que les états des régions 1, 2 et 4 sont des points fixes stables du système, définis, d'une part, par la condition d'équilibre :

$$\frac{dV_{AXO}}{dt} = 0 \leftrightarrow I_B - I_D = 0 \quad (4.10)$$

et, d'autres part, par la condition de stabilité :

$$\frac{\partial}{\partial V_{AXO}} (I_B - I_D) < 0 \quad [A \cdot V^{-1}] \quad (4.11)$$

Au moment des déclenchements, le système passe d'un état d'équilibre stable, à un état de déséquilibre. Pour ce faire, il transite par un état d'équilibre instable, où l'équation (4.10) est toujours vérifiée, mais plus l'équation (4.11). Dans ce contexte, la condition de déclenchement du BB-LTC devient :

$$\frac{\partial}{\partial V_{AXO}} (I_B - I_D) > 0 \quad [A \cdot V^{-1}] \quad (4.12)$$

En particulier, lorsque le NMOS a atteint son régime de saturation sous le seuil (V_{AXO} supérieure à quelques $k \cdot T/q$), $\partial I_D / \partial V_{AXO} = 0$ est vérifié, et la condition (4.12) se simplifie en :

$$\frac{\partial I_B}{\partial V_{AXO}} > 0 \quad [A \cdot V^{-1}] \quad (4.13)$$

Les caractérisations du I_D - V_{FG} à fort V_D du BIMOS en mode MOSFET montrent que cette condition est atteinte pour (section 3.4.1) :

$$I_{II} > I_{BTBT} \quad [A] \quad (4.14)$$

4.3.5 Etudes de l'influence des paramètres de polarisation et de design sur le comportement quasi-statique du BB-LTC

Dans cette section, les influences des polarisations V_{FG}^N et V_{BG} , de la largeur W_1 du BIMOS et de l'EOT des deux composants sur le comportement quasi-statique du BB-LTC sont étudiées.

4.3.5.1 Influence de V_{FG}^N sur la caractéristique I_C - V_{MEM} du BB-LTC

Le potentiel V_{FG}^N qui polarise la grille-avant du NMOS est un paramètre contrôlable du BB-LTC. Cette section explore l'influence de ce potentiel sur la caractéristique du BB-LTC décrit au Tableau 4-1.

Le courant I_D joue un rôle décisif dans les basculements du BB-LTC entre les états ON et OFF de par son rôle dans l'équation d'évolution du potentiel V_{AXO} (4.7) et dans les conditions de stabilité (4.10) et (4.11). En tant que courant de drain d'une structure MOSFET, le courant I_D est limité par la tension grille-avant – source V_{FG-S} du NMOS. En particulier, polarisé sous le seuil ($V_{FG-S} < V_{T,MOS}$) tel que c'est le cas pour le NMOS du BB-LTC, I_D atteint sa valeur de saturation $I_{D,SAT}$ dès que la tension drain-source dépasse quelques kT/q (voir équation (2.53) chapitre 2).

Dans le BB-LTC, la tension V_{FG-S} du NMOS est déterminée par le potentiel V_{FG}^N . Ainsi, une modification de V_{FG}^N modifie la valeur de I_D du NMOS, et toutes les grandeurs caractéristiques du BB-LTC (V_{TH} , V_{MEM} , $I_{ON,TH}$, $I_{OFF,TH}$, $I_{ON,RST}$ et $I_{OFF,RST}$) s'en trouvent modifiées.

- **Polarisation de V_{FG}^N fonctionnelle**

Lorsque $V_{FG}^N \geq -0.2$ V, le BB-LTC est fonctionnel : dans la caractéristique I_C - V_{MEM} , les deux transitions ON/OFF et OFF/ON sont observées et la fonction de bascule de Schmitt est assurée (Figure 4-20). Le mécanisme par lequel V_{FG}^N modifie I_C - V_{MEM} peut-être décrit ainsi :

- Dans le cas où V_{FG}^N augmente, I_D et $I_{D,SAT}$ augmentent en conséquence. Pour que le BB-LTC transite de l'état OFF à l'état ON, la condition $\partial I_B / \partial V_{AXO} > 0$ doit être vérifiée. Elle correspond à l'état de polarisation où $I_{II} > I_{BTBT}$, qui est atteint pour une valeur de $V_{AXO} > 0$. Or, l'augmentation de V_{AXO} avant le déclenchement résulte de l'équilibrage de I_{BTBT} et I_D . Ainsi, pour que la transition OFF/ON ait lieu, l'augmentation de I_D impose que I_{BTBT} soit plus intense, et donc que V_{MEM} soit plus important : V_{TH} augmente.
- En ce qui concerne la transition ON/OFF, sa dépendance en I_D s'appuie cette fois sur la condition d'équilibre (4.10) : le BB-LTC bascule à l'état OFF lorsque $I_B < I_D$, ce qui advient pour de plus fortes valeurs de I_B , et donc de V_{MEM} , lorsque I_D augmente : V_{RST} augmente aussi (Figure 4-20.b). Les courants de collecteur relatifs à ces deux transitions sont modifiés en conséquence (Figure 4-20.c).

Des mesures ont été effectuées jusqu'à $V_{FG}^N = 0.1$ V. Cependant, le BB-LTC reste fonctionnel pour des valeurs de V_{FG}^N supérieures. Elles n'ont pas été investiguées car l'augmentation de V_{TH} pourrait alors dégrader le circuit à cause des trop hautes tensions mises en jeu dans la structure.

Pour des besoins de modélisation qui apparaitront plus loin dans ce document (section 4.5.5), les relations $I_{OFF,TH}/\mu m = f(V_{FG}^N)$ et $V_{TH} = f(V_{FG}^N)$ sont modélisées par une régression exponentielle et linéaire respectivement :

$$I_{OFF,TH} = 7.3 \cdot \exp\left(\frac{0.4 \cdot V_{FG}^N}{U_T}\right) \quad [\text{pA}/\mu\text{m}^2] \quad (4.15)$$

$$V_{TH} = 1.74 + 2.96 \cdot V_{FG}^N \quad [\text{V}] \quad (4.16)$$

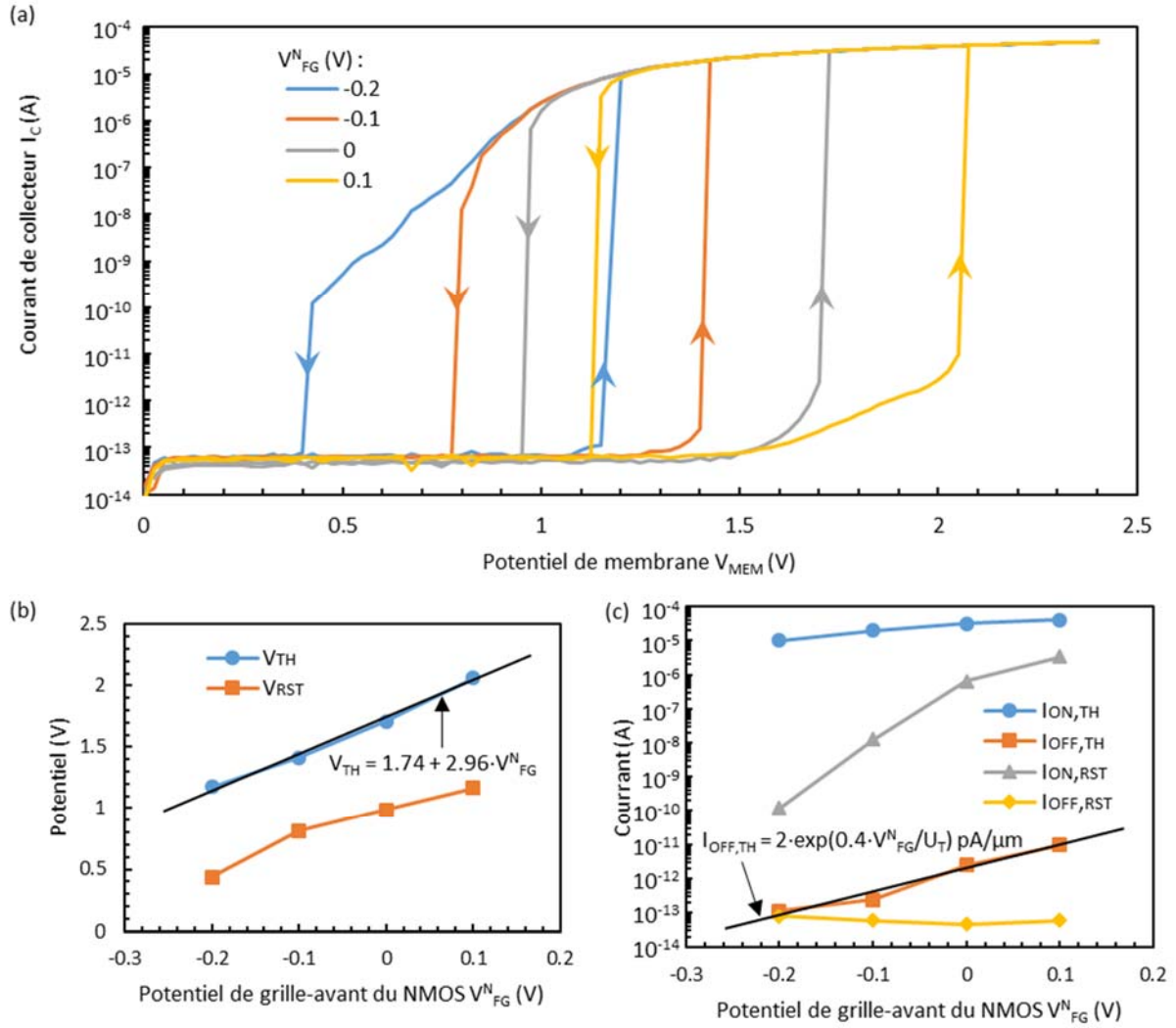


Figure 4-20 – (a) Caractéristiques I_C - V_{MEM} pour différents potentiels $V_{FG}^N \geq -0.2$ V, mesurées à température ambiante sur un BB-LTC décrit au Tableau 4-1. $V_E = V_S = V_{BG} = 0$ V. (b) Extraction depuis (a) des grandeurs V_{TH} et V_{RST} en fonction de V_{FG}^N . (c) Extraction depuis (a) des grandeurs $I_{ON,TH}$, $I_{OFF,TH}$, $I_{ON,RST}$ et $I_{OFF,RST}$ en fonction de V_{FG}^N .

• Limite de fonctionnement du BB-LTC avec V_{FG}^N

Lorsque la valeur de V_{FG}^N devient suffisamment négative ($V_{FG}^N < -0.2$ V), la forme globale de la caractéristique I_C - V_{MEM} change significativement (Figure 4-21.a) : la transition brutale de l'état ON à l'état OFF disparaît au profit d'une décroissance continue de I_C avec V_{MEM} . La fonction de bascule de Schmitt du BB-LTC n'est plus assurée. De plus la monotonie de V_{TH} en fonction de V_{FG}^N n'est plus vérifiée : V_{TH} passe par un minimum à $V_{FG}^N = -0.3$ V (Figure 4-21.b).

Dans cette configuration, le courant I_{D-SAT} devient très faible : la charge Q_{AXO} est difficilement évacuée par le NMOS. Elle rejoint la masse en circulant préférentiellement au travers de la gated-diode base – émetteur. L'absence de transition brutale de l'état ON à l'état OFF montre que, après que la structure est passée à l'état ON, ce nouveau chemin d'évacuation des charges Q_{AXO} lui permet de trouver une situation d'équilibre pour chaque valeur V_{MEM} décroissante.

Concernant l'état OFF du BB-LTC, ces observations peuvent-être faites :

- La lente croissance exponentielle de I_C sur $V_{MEM} \in [0 ; 0.8]$ V, particulièrement visible à $V_{FG}^N = -0.4$ V, montre que le potentiel V_{AXO} croît linéairement avec V_{MEM} , et ce, avant que le courant I_{BTBT} n'intervienne pour fournir des charges à C_{AXO} .
- L'augmentation de V_{TH} avec la réduction de V_{FG}^N peut s'expliquer par l'augmentation de la conductance du NMOS lorsque le potentiel de sa grille avant devient fortement négatif. Des mécanismes de conduction autres que la diffusion des porteurs dans le canal d'un transistor sous le seuil peuvent rendre compte d'un tel comportement, comme le courant de fuite par la grille, ou le courant de BTBT due à la haute tension drain – grille-avant.

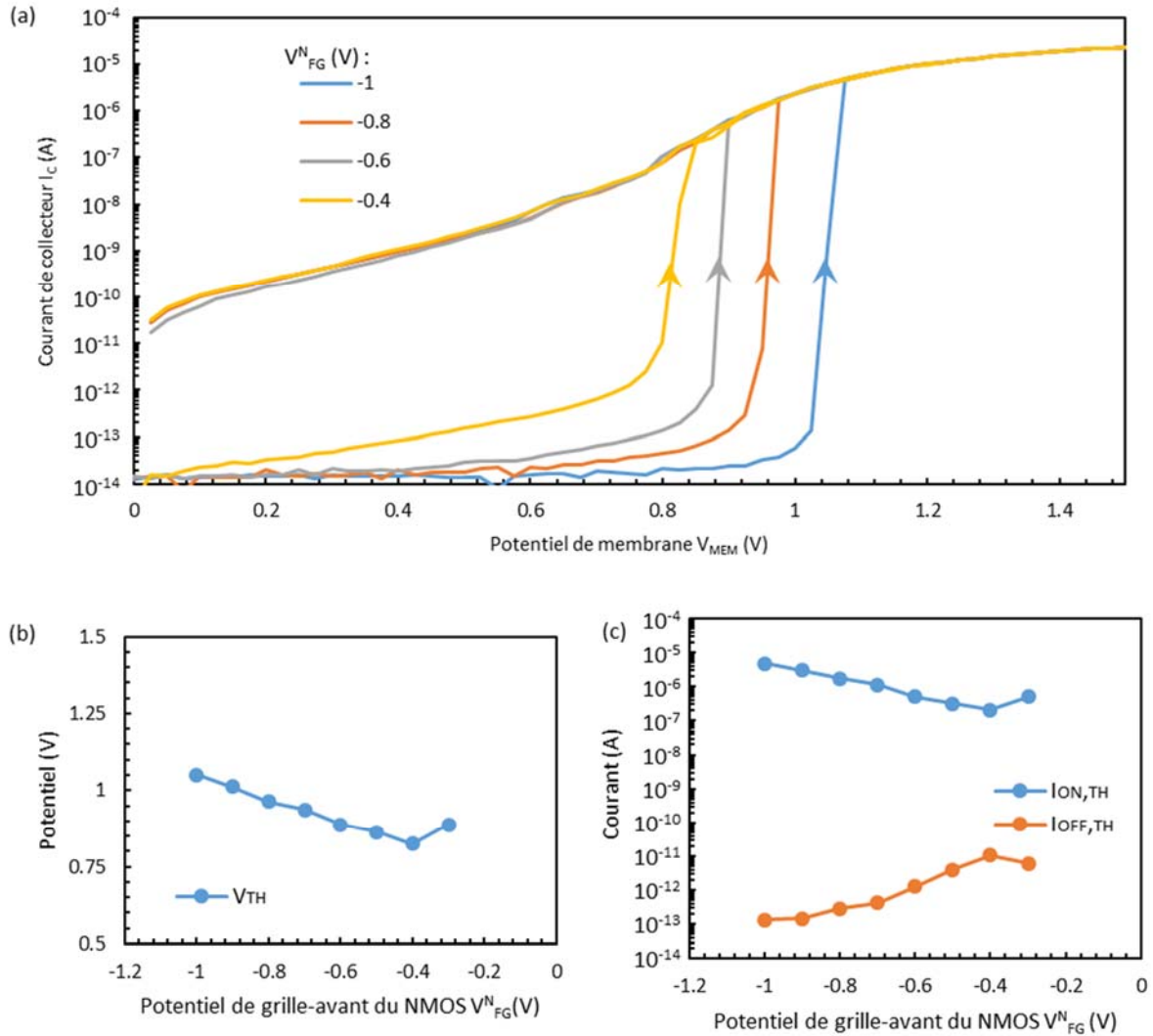


Figure 4-21 – (a) Caractéristiques I_C - V_{MEM} pour différents potentiels $V_{FG}^N \leq -0.2$ V, mesurées à température ambiante sur un BB-LTC décrit au Tableau 4-1. $V_E = V_S = V_{BG} = 0$ V. (b) Extraction depuis (a) de V_{TH} en fonction de V_{FG}^N . (c) Extraction depuis (a) des grandeurs $I_{ON,TH}$, $I_{OFF,TH}$ en fonction de V_{FG}^N .

Les hypothèses faites à la section 4.3.4 sur la nature et les caractéristiques des courants responsables du chargement de n_{AXO} ne permettent pas de décrire de façon pertinente les courbes I_C - V_{MEM} observées pour $V_{FG}^N < -0.1$ V. En particulier, l'hypothèse d'un courant de gated-diode base – collecteur négligeable lorsqu'elle est polarisée en inverse ne permet pas d'envisager les variations de V_{AXO} avant que le courant I_{BTBT} n'intervienne. De la même manière, l'hypothèse d'un courant de drain du NMOS exclusivement constitué du courant de canal ne permet pas d'envisager l'augmentation de V_{TH} lorsque la polarisation de sa grille-avant diminue (pour l'expliquer, il faut en effet que la conductance de NMOS s'accroisse avec la réduction de V_{FG}^N , comme c'est le cas en présence du phénomène de GIDL).

4.3.5.2 Influence de la largeur W_1 de BIMOS sur la caractéristique I_C - V_{MEM} du BB-LTC

Du point de vue du design, cette section explore l'influence de la largeur W_1 de BIMOS sur la caractéristique I_C - V_{MEM} du BB-LTC (Figure 4-22.a). Les autres dimensions caractéristiques du BB-LTC répertoriées au Tableau 4-1 sont laissées constantes. Les grandeurs caractéristiques sont extraites et tracées en fonction de W_1 (Figure 4-22.b et c). Afin de mieux observer les éventuels effets d'échelle, ces grandeurs sont également tracées en étant normalisées (Figure 4-22.d et e).

L'augmentation de W_1 a les effets suivants :

- V_{TH} et V_{RST} diminuent, dès que $W_1 \geq 270$ nm (Figure 4-22.b). Ils augmentent légèrement entre $W_1 = 144$ nm et $W_1 = 270$ nm.
- Les courants $I_{ON,TH}$ et $I_{OFF,RST}$ augmentent proportionnellement à W_1 (Figure 4-22.c et Figure 4-22.e).
- L'extraction des courants $I_{OFF,TH}$ et $I_{ON,RST}$ ne permet pas de définir une tendance dans leur dépendance en W_1 (Figure 4-22.c et Figure 4-22.e).

La proportionnalité des courants I_{CH} , I_{BTBT} et I_{II} avec W_1 , due à un simple effet d'échelle, permettent d'expliquer les variations observées sur V_{TH} , V_{RST} , $I_{ON,TH}$ et $I_{OFF,RST}$.

En effet, l'augmentation de $I_{OFF,RST}$ avec W_1 est due à la proportionnalité du courant I_{CH} avec W_1 . Rappelons que $I_{OFF,RST}$ correspond au courant de saturation de MOSFET intrinsèque au BIMOS quand son potentiel de grille est nul (section 2). Ainsi, l'augmentation de $I_{OFF,RST}$ correspond à l'augmentation du facteur W/L dans l'expression du courant de saturation du MOSFET en régime sous le seuil (voir équation (2.53)).

De même, la réduction de V_{TH} est due à la linéarité de I_{BTBT} avec W_1 . En effet, la génération par BTBT est proportionnelle à la longueur de la frontière collecteur – body où elle a lieu, qui vaut $W_1 + (W_2 - L_{G1})/2$. Comme I_D demeure constant entre les différents BB-LTC mesurés, la condition de stabilité $I_{BTBT} = I_D$ est atteinte pour de plus faibles valeurs de tension $V_{MEM} - V_{AXO}$. Autrement dit, à V_{MEM} égal et une fois que le BTBT est activé, le potentiel d'axone V_{AXO} est plus haut lorsque W_1 augmente. De la même manière, comme I_{II} est proportionnel à I_{CH} , qui est lui-même proportionnel à W_1 , la condition de déclenchement $\partial(I_B - I_D)/\partial V_{AXO} > 0$ (voir section 2) est atteinte pour une plus faible valeur de V_{MEM} , et V_{TH} décroît.

Concernant $I_{OFF,TH}$, le fait que, d'une part, juste avant la transition OFF/ON, le courant I_C est principalement constitué du courant I_{BTBT} et que, d'autre part, le courant I_D ne dépend pas de W_1 , devrait conduire à une indépendance de $I_{OFF,TH}$ en W_1 . Il est possible que cela soit vérifié ici, mais que les variabilités entre les dispositifs empêchent de l'observer clairement. Une étude de variabilité permettrait de lever ce doute.

L'augmentation de $I_{ON,TH}$ peut être le résultat de deux contributions qui se renforcent. D'une part, pour des valeurs de V_{AXO} et V_{MEM} données, I_{CH} augmente avec W_1 par effet d'échelle. D'autre part, comme I_{II} augmente aussi par effet d'échelle avec W_1 , mais pas I_{PN} , la condition de stabilité à l'état ON $I_{II} = I_{PN}$ est vérifiée pour un potentiel V_{AXO} plus haut, et donc pour un courant I_{CH} plus intense. Cependant, à la Figure 4-22.c, la proximité de la courbe normalisée $I_{ON,TH} = f(W_1)$ avec la courbe $y = x$ montre que la première contribution, directement liée au facteur d'échelle, est la plus dominante.

Enfin, la réduction de V_{RST} s'explique aussi par l'augmentation de I_{II} par facteur d'échelle qui conduit à la vérification de l'inégalité $I_{II} < I_D$ (responsable du basculement du BB-LTC de l'état ON à l'état OFF) pour des valeurs de V_{MEM} moins importantes lorsque W_1 est plus grand. De plus, la primauté du rôle

de l'inégalité $I_{II} < I_D$ dans la basculement ON/OFF est consistante avec l'invariance de $I_{ON,RST}$ en fonction de W_1 .

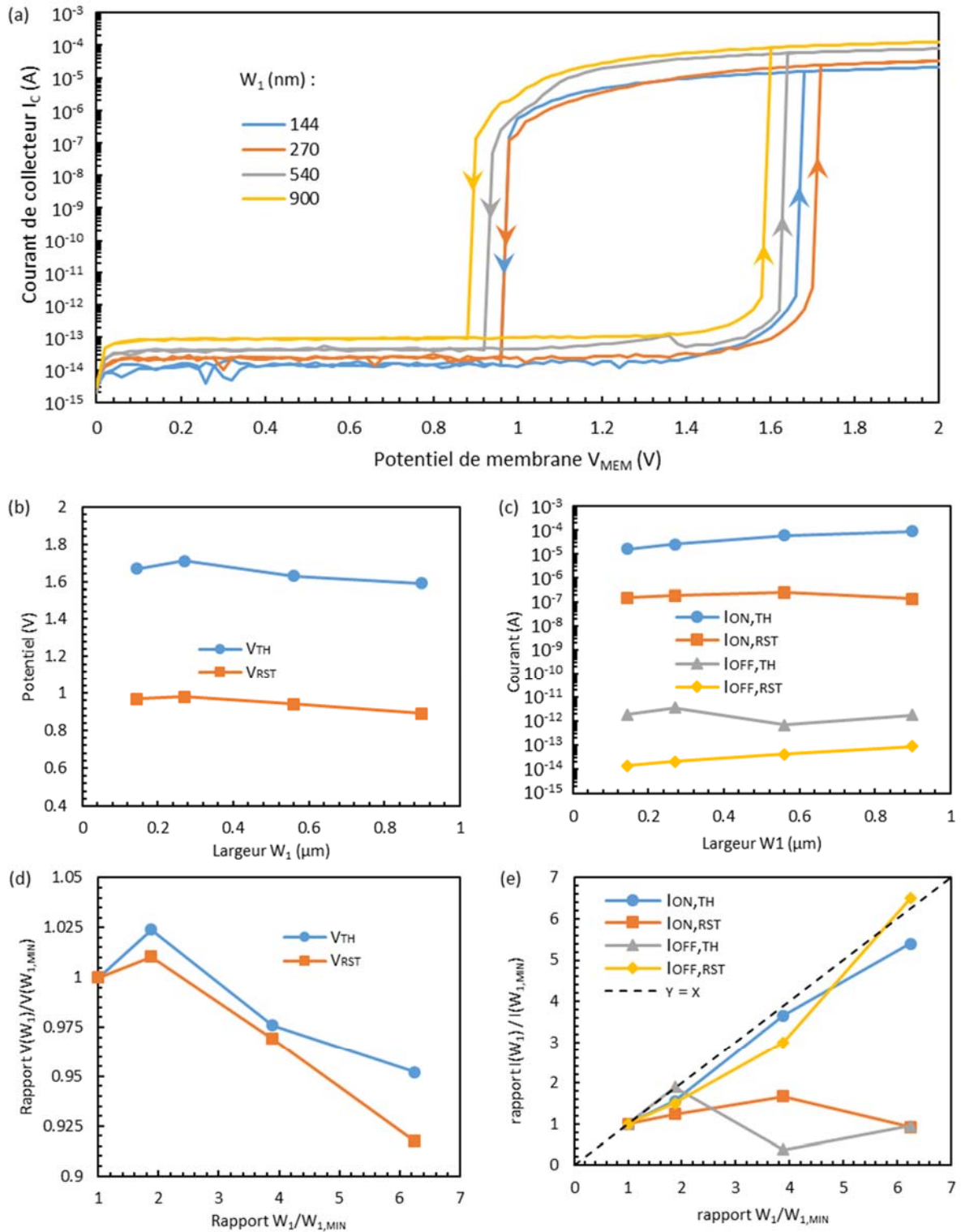


Figure 4-22 – (a) Caractéristique I_C - V_{MEM} pour différentes largeurs W_1 , mesurée à température ambiante sur un BB-LTC à oxyde épais ($EOT = 3.4$ nm) de dimension $L_1 = L = 135$ nm, $W_2 = W = 350$ nm et $L_2 = 144$ nm. Le BIMOS et le NMOS sont de type RVT. $V_E = V_S = V_{BG} = V_{FG}^N = 0$ V. (b) Polarisation caractéristique en fonction de la largeur W_1 , extraites depuis (a). (c) Courants caractéristiques en fonction de la largeur W_1 , extraits depuis (a). (d) Polarisation caractéristique normalisées par leur valeur à $W_{1,MIN}$ en fonction de la largeur W_1 normalisée par $W_{1,MIN}$. (e) Courants caractéristiques normalisés par leur valeur à $W_{1,MIN}$ en fonction de la largeur W_1 normalisée par $W_{1,MIN}$.

4.3.5.3 Influence de V_{BG}^B sur la caractéristique I_C - V_{MEM} du BB-LTC

Dans cette section l'influence du potentiel de grille-arrière du BIMOS, noté V_{BG}^B , sur la caractéristique I_C - V_{MEM} du BB-LTC est étudiée par la mesure. Le potentiel de grille arrière du NMOS, noté V_{BG}^N , est laissé inchangé à $V_{BG}^N = 0$ V. Les caractéristiques du BB-LTC utilisé sont données au Tableau 4-5. Elles diffèrent de celles du BB-LTC mesuré jusqu'à présent par le changement de saveur (et donc de tension seuil) du BIMOS, qui passe de RVT à LVT.

Tableau 4-5 - Dimensions du BB-LTC utilisé pour les mesures présentées à la section 4.3.5.3.

L_1 (nm)	W_1 (nm)	L_2 (nm)	W_2 (nm)	L (nm)	W (nm)	EOT (nm)	BIMOS	NMOS
135	270	270	350	135	350	3.4	LVT	RVT

Tout d'abord, rappelons qu'en technologie 28nm UTBB FD-SOI, la polarisation de la grille-arrière d'un MOSFET décale sa caractéristique I_D - V_{FG} selon l'axe des abscisses : sa tension de seuil, ses courants à l'état OFF et à l'état ON sont décalés en conséquence. En tant que structure MOSFET, ce décalage est observé sur la caractéristique I_D - V_{FG} du BIMOS (Figure 4-23.a). En particulier, pour le BIMOS mesuré de type GO2 en régime de saturation ($V_D = 1.8$ V > $V_{D,SAT}$ sur toute la plage de mesure $V_{FG} \in [-1 ; 2.5]$ V), le décalage de $V_{T,MOS}$ en fonction de V_{BG}^B est de -135 mV/V (Figure 4-23.b). Cela est proche de la valeur théorique -125 mV/V calculée d'après l'équation (2.56).

Concernant le courant de body-contact, la courbe $I_B = f(V_{FG})$ se décale également, à l'exception de la partie décroissante observée pour les hautes valeurs de V_{FG} (courbes en pointillés de la Figure 4-23.a, pour $V_{FG} > 0.9$ V). Pour rappel, la diminution du courant de body-contact dans cette région est due au rétrécissement de la zone de pincement du canal où le mécanisme d'ionisation par impact a lieu. Ces courbes montrent que la polarisation de la grille arrière V_{BG}^B n'a pas d'influence sur cette diminution.

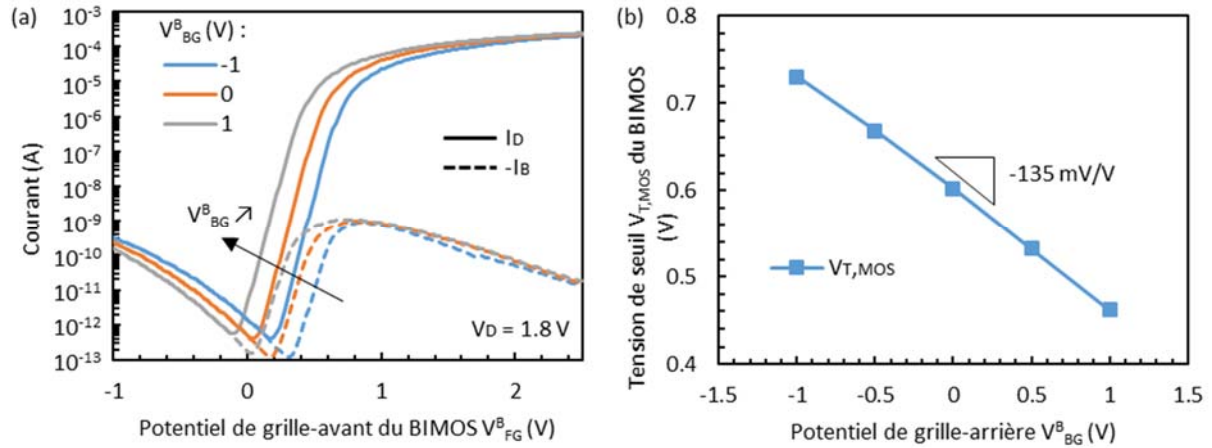


Figure 4-23 – (a) Courants de drain I_D et de body-contact I_B en fonction du potentiel de grille-avant du BIMOS V_{FG}^B , mesurés à température ambiante sur un BIMOS de type GO2 RVT aux dimensions $L_1 = 135$ nm, $W_1 = 900$ nm, $L_2 = 144$ nm, $W_2 = 350$ nm. $V_D = 1.8$ V et $V_S = V_B = 0$ V. (b) Tension de seuil $V_{T,MOS}$ du BIMOS en fonction du potentiel de grille-arrière V_{BG}^B , extrait depuis (a).

Les courbes I_D - V_{FG} de la Figure 4-23 permettent d'envisager l'influence de V_{BG}^B sur certaines grandeurs caractéristiques du BB-LTC observée lors des mesure (Figure 4-24.a) :

- Comme $I_{OFF,RST}$ est égal au courant I_{CH} à l'état OFF et que, dans les structures FD-SOI NMOSFET, I_{CH} augmente avec V_{BG}^B , $I_{OFF,RST}$ augmente avec V_{BG}^B (Figure 4-24.c).
- Par ailleurs, l'augmentation de I_{CH} avec V_{BG}^B conduit à des situations où $I_{CH} \gg I_{BTBT}$ au moment du déclenchement. Par conséquent, $I_{OFF,TH}$ augmente avec V_{BG}^B (Figure 4-24.c).

- Concernant V_{TH} , rappelons d'abord que le BB-LTC bascule de l'état OFF à l'état ON lorsque la condition $\partial(I_B - I_D)/\partial V_{AXO} > 0$ est vérifiée (section 4.3.5.2). Or, la Figure 4-23.a montre que $\partial I_B/\partial V_{AXO}$ augmente avec V_{BG}^B (pour rappel, dans le BB-LTC, $V_{AXO} = V_{BG}^B$). La condition de déclenchement est atteinte pour de plus faibles valeurs de V_{AXO} et donc pour de plus faibles valeurs de V_{MEM} : V_{TH} diminue avec V_{BG}^B . (Figure 4-24).

Le comportement des trois autres grandeurs caractéristiques V_{RST} , $I_{ON,TH}$ et $I_{ON,RST}$ avec V_{BG}^B n'est pas expliqué. On peut néanmoins constater que V_{RST} et V_{TH} demeurent relativement invariants avec V_{BG}^B et $I_{ON,RST}$ augmente avec V_{BG}^B .

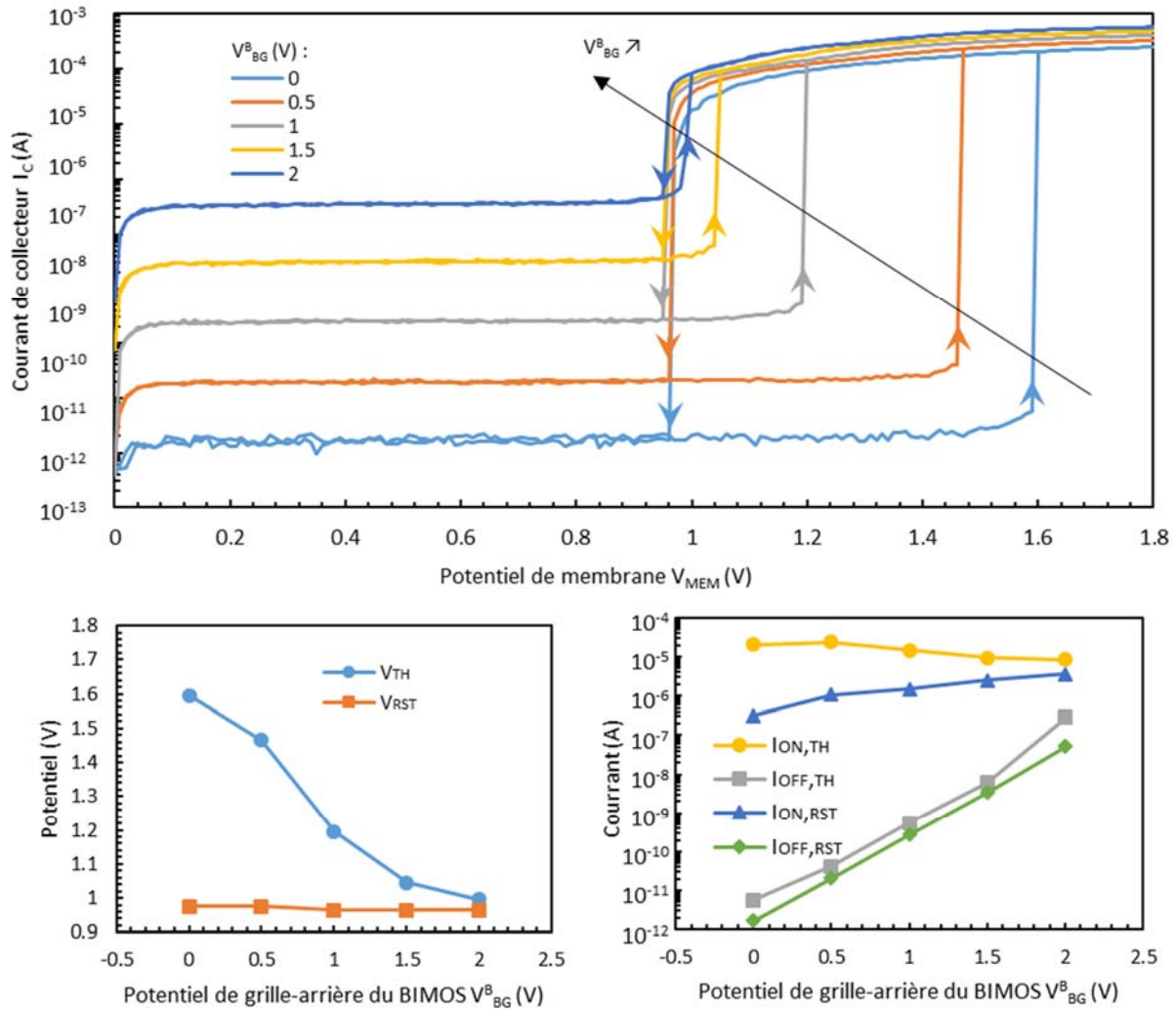


Figure 4-24 – (a) Caractéristiques I_C - V_{MEM} pour différents potentiels V_{BG}^B , mesurées à température ambiante sur un BB-LTC à oxyde épais ($EOT = 3.4$ nm) de dimension $L_1 = L = 135$ nm, $W_1 = 144$ nm, $W_2 = W = 350$ nm et $L_2 = 144$ nm. Le BIMOS est de type LVT et le NMOS de type RVT. $V_E = V_S = V_{BG} = 0$ V. (b) Extraction depuis (a) des grandeurs V_{TH} et V_{RST} en fonction de V_{BG}^B . (c) Extraction depuis (a) des grandeurs $I_{ON,TH}$, $I_{OFF,TH}$, $I_{ON,RST}$ et $I_{OFF,RST}$ en fonction de V_{BG}^B .

4.3.5.4 Influence de l'épaisseur d'oxyde sur la caractéristique I_C - V_{MEM} du BB-LTC

En technologie 28 nm FD-SOI, deux épaisseurs d'oxyde de grille différentes sont disponibles pour la fabrication des MOSFET : un oxyde fin qui implémente les grilles dites standards (SG), avec un EOT de 1.1 nm, et un oxyde épais qui implémente les grilles dites étendues (EG) avec un EOT de 3.4 nm. Jusqu'à présent, les BIMOS et les NMOS des différents BB-LTC caractérisés étaient de type EG. Dans cette section, nous comparons les caractéristique I_C - V_{MEM} de deux BB-LTC aux dimensions égales mais

implémentés l'un avec des composants SG et l'autre avec des composants EG (Figure 4-25).

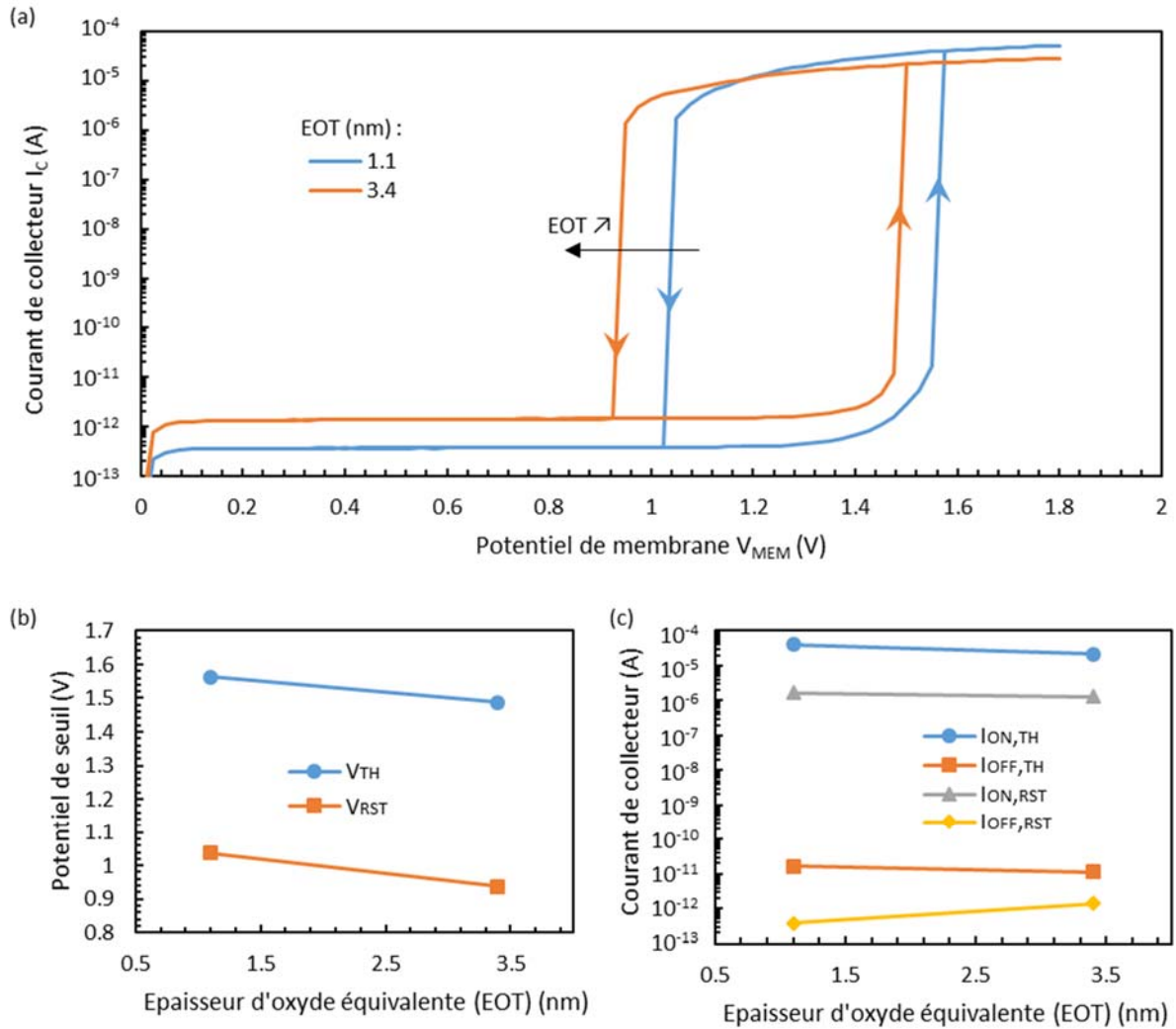


Figure 4-25 – (a) Caractéristiques quasi-statiques I_C - V_{MEM} mesurées à température de deux BB-LTC aux épaisseurs d'oxyde de grille différents ($EOT = 1.1$ nm et 3.4 nm) et aux autres dimensions égales : $L_1 = L = 135$ nm, $W_1 = 144$ nm, $W_2 = W = 350$ nm et $L_2 = 144$ nm. Le BIMOS est de type LVT et le NMOS de type RVT. $V_E = V_S = V_{BG} = V_{FG}^n = 0$ V. (b) Tensions de seuils V_{TH} et V_{RST} en fonction de l'EOT extraites depuis (a). (c) Courants caractéristiques $I_{ON,TH}$, $I_{OFF,TH}$, $I_{ON,RST}$ et $I_{OFF,RST}$ en fonction de l'EOT extraits depuis (a).

Globalement, la forme de la courbe I_C - V_{MEM} est identique dans les deux configurations.

Le courant à l'état OFF, dont la valeur est $I_{OFF,RST}$, est plus faible dans le BB-LTC à oxyde fin. Autrement dit, dans le BIMOS, pour un potentiel de grille-avant $V_{AXO} = 0$ V, le potentiel de surface à l'interface body – oxyde de grille-avant $\phi_{S,F}$ est plus bas lorsque l'épaisseur d'oxyde est plus faible. Cela se comprend en considérant 3 points :

- Tout d'abord, la capacité d'oxyde-avant C_{OX} diminue lorsque l'EOT augmente : $dC_{OX}/dEOT < 0$.
- Ensuite, le BIMOS mesuré est de type LVT. Par conséquent, le travail de sortie de la grille-avant $\phi_{M-SC,F}$ est supérieur au travail de sortie de la grille arrière $\phi_{M-SC,B}$: $\phi_{M-SC,F} > \phi_{M-SC,B}$.
- Enfin, en reprenant la modélisation de la structure FDSOI comme réseaux de capacités (voir section 2.3.2.2) et en considérant $V_{FG} = V_{BG} = 0$ V, la relation entre le potentiel de surface $\phi_{S,F}$ et la différence des travaux de sortie des grilles arrière et avant $\phi_{M-SC,B} - \phi_{M-SC,F}$ s'écrit :

$$\phi_{S,F} = \frac{1}{1 + \frac{C_{BOX} \cdot C_{SI}}{C_{OX} \cdot (C_{BOX} + C_{SI})}} \cdot (\phi_{M-SC,B} - \phi_{M-SC,F}) \quad [\text{eV}] \quad (4.17)$$

Ainsi, en considérant que $dC_{OX}/dEOT < 0$ et que $\phi_{M-SC,B} - \phi_{M-SC,F} < 0$, on montre facilement que :

$$\frac{\partial \phi_{S,F}}{\partial EOT} > 0 \quad [\text{eV} \cdot \text{nm}^{-1}] \quad (4.18)$$

Pour discuter des différences des autres grandeurs caractéristiques, il est nécessaire d'extraire la valeur de V_{AXO} lors du balayage de V_{MEM} , en suivant la même méthodologie qu'à la section 4.3.1 (Figure 4-26).

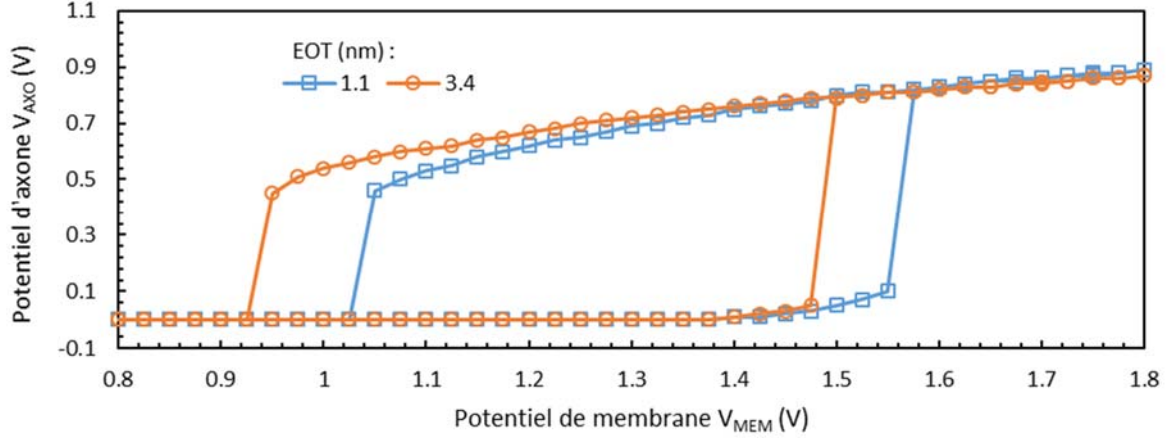


Figure 4-26 – Potentiel V_{AXO} en fonction du potentiel de membrane V_{MEM} et à différents EOT, extrait depuis les données de la Figure 4-25, et reconstruit à partir de mesures des caractéristique $ID-V_{FG}$ à différents V_D et à $V_B = V_{FG}$ (même méthodologie que celle utilisée à la section 4.3.1).

Bien que les tensions de seuils soient différentes, les deux courbes se chevauchent presque intégralement, sauf lorsque les BB-LTC sont à l'état ON et que $V_{MEM} < V_{TH}$. On a alors $V_{AXO}(EOT = 1.1 \text{ nm}) > V_{AXO}(EOT = 3.4 \text{ nm})$.

Le potentiel d'axone avant le déclenchement $V_{AXO,OFF}$ du BB-LTC à oxyde fin est supérieur à celui du BB-LTC à oxyde épais. Cela montre que le point de bascule (V_{MEM}, V_{AXO}) où $\partial(I_B - I_D)/\partial V_{AXO} > 0$ (qui marque le départ de l'emballage) est différent dans les deux structures.

Les potentiels V_{AXO} sont quasiment égaux après la transition OFF/ON. Leurs valeurs ne peuvent donc pas expliquer la différence de $I_{ON,TH}$ observée entre les structures aux deux épaisseurs d'oxyde différentes. Ainsi, la différence de $I_{ON,TH}$ s'explique simplement par le rôle de la capacité d'oxyde dans le courant de drain des MOSFET au-dessus du seuil (voir équation 2.51) : $I_D \propto C_{OX}$. Comme $C_{OX} \propto 1/EOT$, le courant $I_{ON,TH}$ est plus important dans le BB-LTC à l'oxyde le plus fin.

Lorsque le BB-LTC est à l'état ON, la baisse de courant I_C avec la diminution de V_{MEM} est plus importante dans la structure à oxyde fin. Cela correspond à une baisse de V_{AXO} plus importante. Dans les deux cas, la transition ON/OFF intervient pour la même valeur de V_{AXO} . Cependant, à cause de la chute de V_{AXO} plus rapide dans la structure à oxyde fin, V_{RST} est plus faible. Les courants $I_{ON,RST}$ sont équivalents dans les deux BB-LTC.

4.3.6 Confrontation du modèle SPICE aux mesures quasi-statiques

Dans cette section, des simulations SPICE quasi-statiques de la caractéristique I_C-V_{MEM} du BB-LTC qui utilisent le modèle du BIMOS développé à la section 3.6 sont comparées aux résultats obtenus par la

mesure. Les dimensions et les caractéristiques des BB-LTC mesurées et simulées sont celles répertoriées au Tableau 4-5.

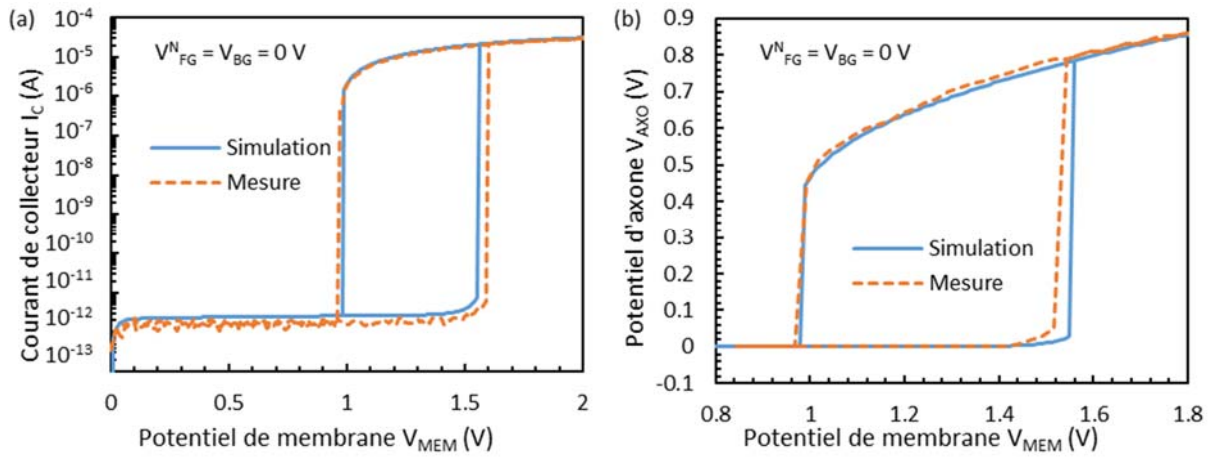


Figure 4-27 – (a) Courants de collecteur I_C en fonction du potentiel de membrane V_{MEM} obtenus par simulation et par la mesure. (b) Potentiels V_{AXO} en fonction du potentiel de membrane V_{MEM} obtenus par simulation et par la mesure. Les dimensions du BB-LTC sont répertoriées au Tableau 4-5. $V_E = V_S = V_{BG} = V_{FG}^N = 0$ V.

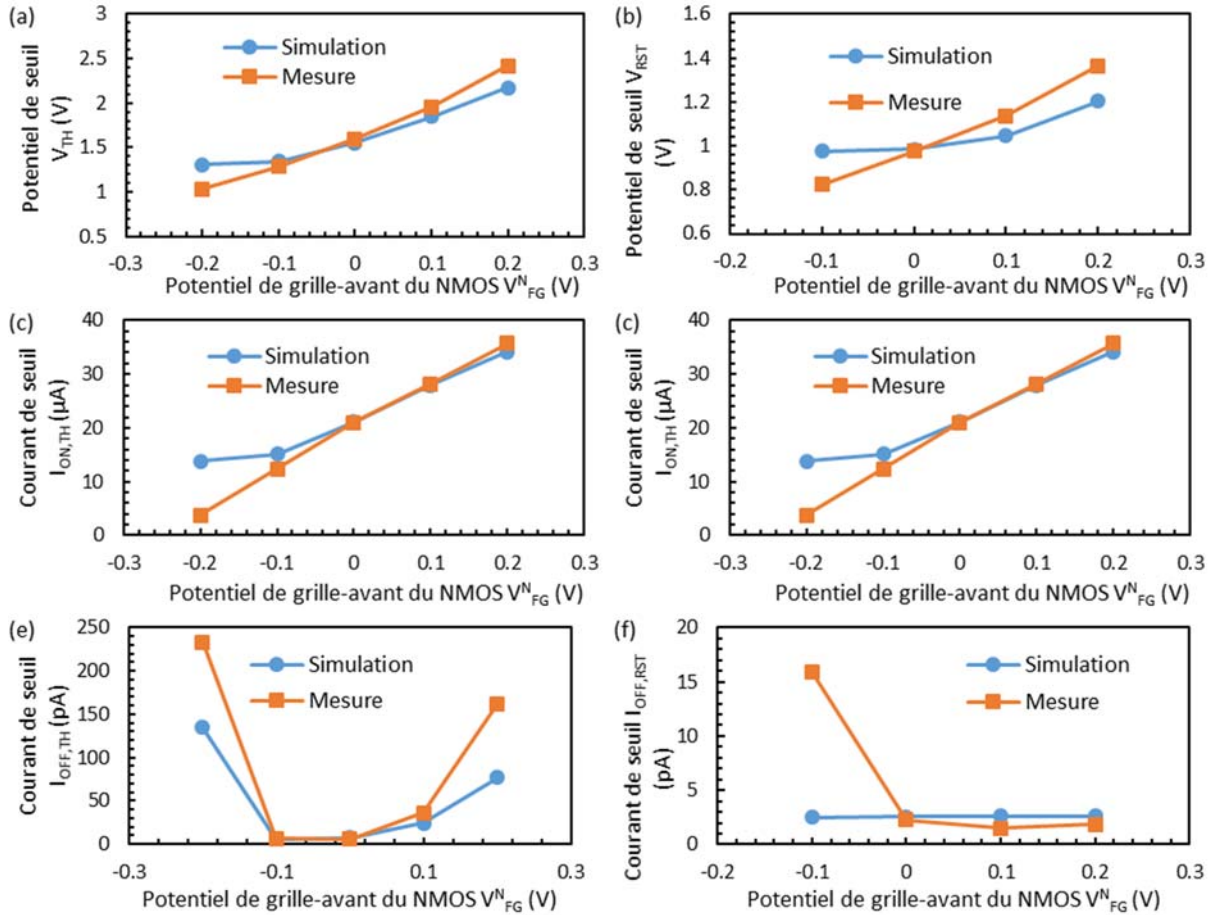


Figure 4-28 – Comparaisons des valeurs obtenues par la mesure et par la simulation des potentiels (a) V_{TH} et (b) V_{RST} et des courants (c) $I_{ON,TH}$, (d) $I_{ON,RST}$, (e) $I_{OFF,TH}$ et (f) $I_{OFF,RST}$, en fonction du potentiel V_{FG}^N , pour un BB-LTC aux dimensions répertoriées au Tableau 4-5. $V_{BG} = 0$ V.

Les courbes I_C - V_{MEM} et V_{AXO} - V_{MEM} obtenues en simulation et en mesure pour $V_{FG}^N = V_{BG} = 0$ V sont tracées à la Figure 4-27. Par ailleurs, les valeurs des six grandeurs caractéristiques V_{TH} , V_{RST} , $I_{OFF,TH}$, $I_{ON,TH}$, $I_{OFF,RST}$

et $I_{ON,RST}$ en fonction du potentiel V_{FG}^N obtenues en mesure et en simulation sont tracées à la Figure 4-28.

Bien que des décalages existent, les simulations sont globalement en accord avec les mesures sur la plage de polarisation explorée. Notamment, les tendances des variations des grandeurs caractéristiques avec V_{FG}^N sont reproduites par la simulation. Outre les variabilités inhérentes à la fabrication des échantillons, les décalages observés s'expliquent par l'imprécision du modèle du courant d'ionisation par impact déjà constatée à la section 3.6. Ces résultats montrent que, sans être un outil prédictif quantitatif, le modèle compact développé au chapitre 3 est pertinent pour effectuer des explorations qualitatives de circuit intégrant un BB-LTC. Notons que l'évaluation des incertitudes de mesure et de la variabilité des BB-LTC permettrait de statuer davantage sur la validité du modèle SPICE utilisé.

4.3.7 Comportement dynamique du BB-LTC

Jusqu'à présent, la caractéristique I_C - V_{MEM} du BB-LTC a été étudiée au travers de mesures quasi-statiques : chaque point de la caractéristique est mesuré en régime établi. Dans ce contexte, les courbes I_C - V_{MEM} présentent des discontinuités claires lors des transitions entre les états ON et OFF. Or, ces transitions se font sur des durées non nulles qui peuvent rendre les grandeurs caractéristiques du BB-LTC dépendantes de dV_{MEM}/dt lorsque dV_{MEM}/dt est suffisamment grand.

4.3.7.1 Influence du temps de montée de V_{MEM} sur V_{TH}

Pour mettre en évidence cette dépendance temporelle, des caractérisations I_C - V_{MEM} sont effectuées en variant le temps de montée τ_{RISE} et le temps de descente τ_{FALL} du potentiel de membrane V_{MEM} , à l'aide d'un générateur de forme d'onde Keysight B1530 (Figure 4-29). Par souci de simplicité, ces deux valeurs sont prises égales entre elles.

Comme le pas de temps avec lequel les mesures sont échantillonnées est très court, le calibre minimal est limité : il ne permet pas de mesurer des courants inférieurs au nanoampère. Dans ce contexte, seules les variations de V_{TH} en fonction de $\tau_{RISE} = \tau_{FALL}$ sont observables, mais pas celles de V_{RST} .

Deux observations principales peuvent être faites sur les courbes de la Figure 4-29.b :

- Plus le temps de montée est court, plus le potentiel de seuil V_{TH} est haut (Figure 4-29.c).
- Plus le temps de descente est court, moins la décroissance de I_{ON} avec V_{MEM} est importante.

Ces observations sont interprétées de la manière suivante.

En ce qui concerne le décalage de V_{TH} , rappelons que le BB-LTC transite de l'état OFF à l'état ON lorsque la condition $\partial(I_B - I_D)/\partial V_{AXO} > 0$ est vérifiée (voir 4.3.4). Dès lors, une boucle de rétroaction positive est enclenchée, qui augmente le potentiel V_{AXO} selon l'équation (4.7). Sous l'hypothèse que le courant I_{II} domine, elle se réécrit sous la forme :

$$\frac{dV_{AXO}}{dt} = \frac{1}{C_{AXO}} \cdot I_{II} \quad [V \cdot s^{-1}] \quad (4.19)$$

C_{AXO} et I_{II} sont des grandeurs qui varient avec V_{MEM} et V_{AXO} , ce qui rend le calcul analytique de dV_{AXO}/dt fastidieux. Néanmoins, (4.19) montre que l'augmentation de V_{AXO} au cours du temps est proportionnelle à l'intensité du courant I_{II} , qui est minimal au début du déclenchement et qui ne fera qu'augmenter à mesure que la transition OFF/ON s'effectue.

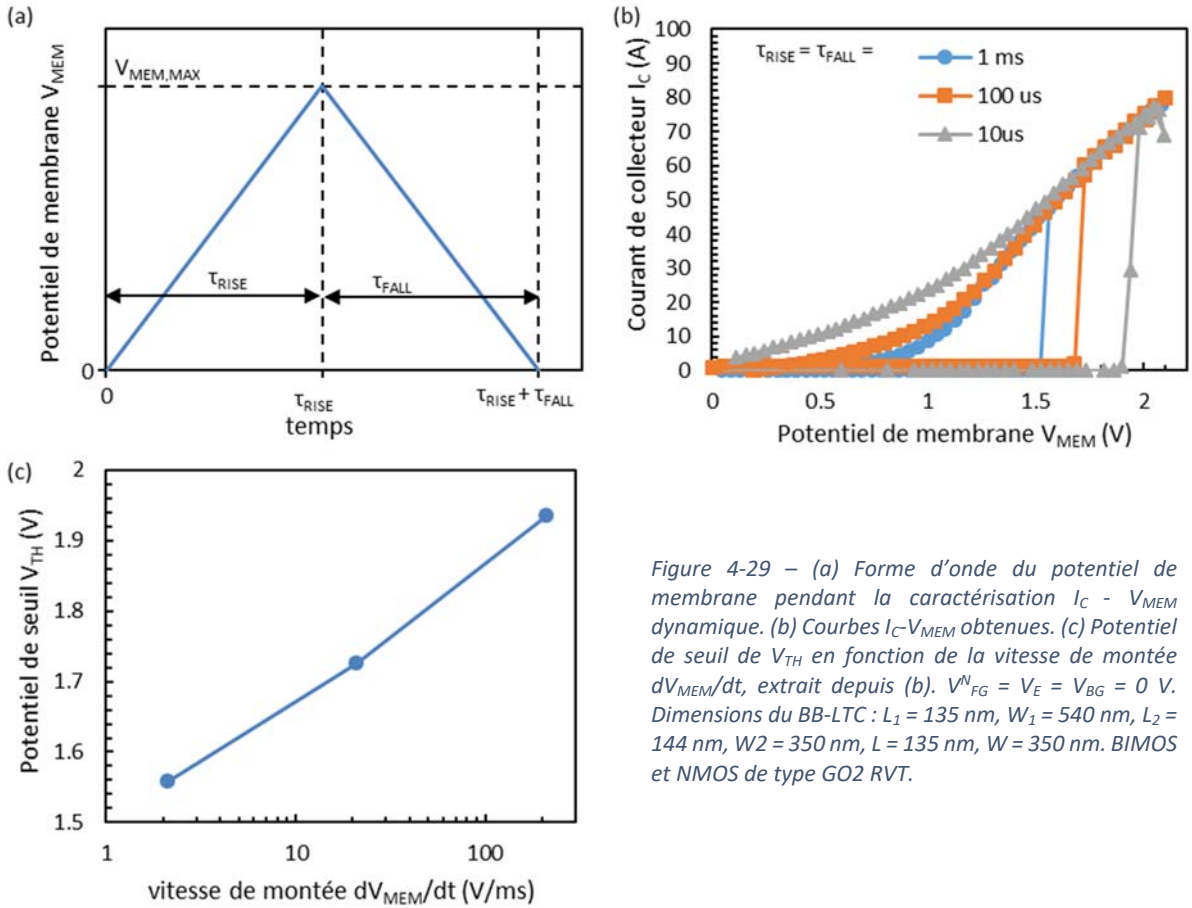


Figure 4-29 – (a) Forme d’onde du potentiel de membrane pendant la caractérisation $I_C - V_{MEM}$ dynamique. (b) Courbes $I_C - V_{MEM}$ obtenues. (c) Potentiel de seuil de V_{TH} en fonction de la vitesse de montée dV_{MEM}/dt , extrait depuis (b). $V_{FG}^N = V_E = V_{BG} = 0$ V. Dimensions du BB-LTC : $L_1 = 135$ nm, $W_1 = 540$ nm, $L_2 = 144$ nm, $W_2 = 350$ nm, $L = 135$ nm, $W = 350$ nm. BIMOS et NMOS de type GO2 RVT.

Si l’on considère que le BB-LTC est à l’état ON pour $V_{AXO} \geq V_{T,MOS}$, on peut définir une durée de déclenchement τ_{TRG} qui correspond à la durée que met V_{AXO} pour atteindre $V_{T,MOS}$ une fois que la transition vers l’état ON est enclenchée. Cette durée dépend de l’intensité de I_{II} : τ_{TRG} est d’autant plus court que I_{II} est intense. Pendant la durée de déclenchement τ_{TRG} , V_{MEM} varie de $\Delta V_{MEM} = \langle dV_{MEM}/dt \rangle \cdot \tau_{TRG}$, où $\langle dV_{MEM}/dt \rangle$ est la valeur moyenne de dV_{MEM}/dt pendant la durée du déclenchement. Ainsi, tant que le produit $\langle dV_{MEM}/dt \rangle \cdot \tau_{TRG}$ est petit, le potentiel de seuil V_{TH} semble indépendant du temps ($\Delta V_{MEM} \ll 1$ V). En revanche, si ce produit a une valeur significative, ΔV_{MEM} n’est pas négligeable, et le potentiel de membrane est plus haut au moment où le déclenchement s’achève : V_{TH} augmente.

En ce qui concerne la réduction de dI_C/dV_{MEM} lorsque τ_{FALL} diminue, elle est une conséquence de l’évacuation non-instantanée de la charge Q_{AXO} par les courants I_D et I_{PN} . Cette limitation de l’intensité de décharge conduit aux situations où, pour 2 valeurs de V_{MEM} égales, la charge Q_{AXO} présente dans n_{AXO} , et donc le potentiel V_{AXO} , est différente selon que V_{MEM} décroît lentement ou rapidement.

4.3.7.2 Durée de déclenchement τ_{ON}

Pour étudier davantage le comportement dynamique du BB-LTC, définissons la durée τ_{ON} comme la durée que met le BB-LTC pour passer de l’état OFF à l’état ON lorsque le potentiel V_{MEM} est monté instantanément (en 10 ns) à un potentiel $V_{MEM,MAX}$, puis maintenu à cette valeur. Notons que cette durée est plus longue que τ_{TRG} car elle prend également en compte la durée que met le courant I_{BTBT} pour précharger le nœud n_{AXO} et amorcer le déclenchement.

Pour évaluer $\tau_{ON} = f(V_{MEM,MAX})$, le courant de collecteur I_C d’un BB-LTC aux dimensions définies au Tableau 4-6 est mesuré au cours du temps (Figure 4-30.a), lorsque la fonction $V_{MEM} = f(t)$ est une fonction échelon avec un temps de montée de 10 ns (Figure 4-30.b). Cette durée est le pas de temps

minimal autorisé par le générateur d'impulsion Keysight B1530 utilisé pour la mesure. Le BB-LTC est considéré à l'état ON lorsque I_C a atteint 90 % de sa valeur maximale $I_{C,MAX}$. La durée τ_{ON} est extraite des courbes $I_C = f(t)$ selon cette définition (Figure 4-30.c).

Tableau 4-6 – Dimensions et polarisation du BB-LTC utilisé pour les mesures de τ_{ON} effectuées à la Figure 4-30

L_1 (nm)	W_1 (nm)	L_2 (nm)	W_2 (nm)	L (nm)	W (nm)	EOT (nm)	V_{FG}^N (V)
28	180	100	225	28	225	1.1	-0.2

Ces mesures montrent que la durée τ_{ON} est d'autant plus courte que le potentiel $V_{MEM,MAX}$ est haut. Cette relation est due aux dépendances de I_{BTBT} et I_{II} en V_{MEM} . En effet, ces deux courants sont plus intenses lorsque V_{MEM} est plus élevé. Ainsi, ils chargent la capacité C_{AXO} plus rapidement, l'augmentation du potentiel V_{AXO} est accéléré et le courant I_C atteint sa valeur maximale plus tôt.

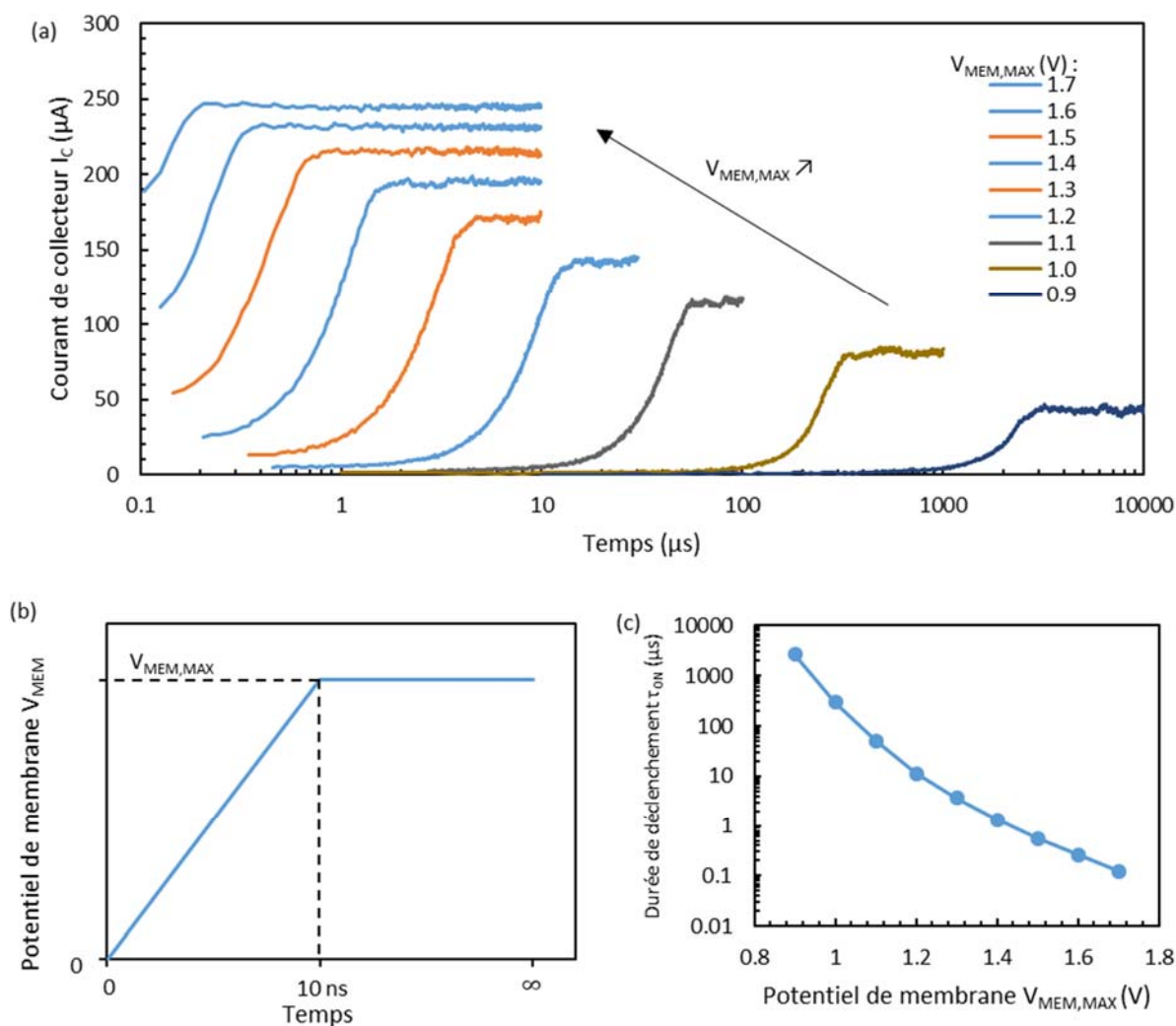


Figure 4-30 – (a) Courbes $I_C = f(t)$ pour différents $V_{MEM} = V_{MEM,MAX}$ d'un BB-LTC dont les dimensions sont données au Tableau 4-6, mesurées à température ambiante. (b) forme de la courbe $V_{MEM} = f(t)$. (c) Durée de déclenchement τ_{ON} en fonction de $V_{MEM,MAX}$ extraite depuis (a).

4.3.7.3 Résumé

Les transitions du BB-LTC entre les états ON et OFF se font sur des durées non nulles. Ces durées résultent de la dynamique de chargement de la capacité C_{AXO} du nœud n_{AXO} . Le chargement de C_{AXO} est

assuré par les courants I_{BTBT} et I_{II} et son déchargement par I_D et I_{PN} . Plus les courants I_{BTBT} et I_{II} sont intenses, plus la transition OFF/ON est rapide. Plus les courants I_{PN} et I_D sont intenses, plus la transition ON/OFF est brève.

Les courants I_{BTBT} et I_{II} dépendent de V_{MEM} et V_{AXO} . En particulier, ils sont tous deux plus intenses si V_{MEM} est plus haut (tandis que leurs dépendances en V_{AXO} sont opposées : une augmentation de V_{AXO} accroît I_{II} et décroît I_{BTBT}). Ainsi, la transition OFF/ON peut être accélérée en augmentant le potentiel V_{MEM} . Si V_{MEM} croît au cours du temps suffisamment rapidement, la durée non nulle de la transition OFF/ON induit une augmentation du potentiel de seuil V_{TH} : V_{TH} dépend de la vitesse de variation de V_{MEM} .

La dynamique de déchargement de C_{AXO} est déterminée par I_{PN} et I_D . L'intensité du courant I_{PN} dépend uniquement de V_{AXO} qui est un potentiel interne au BB-LTC. La constante de temps associée à la contribution de I_{PN} au déchargement de C_{AXO} est une constante liée au dimensionnement du circuit. En revanche, l'intensité du courant I_D dépend de V_{FG}^N qui est un paramètre du BB-LTC. La décharge de C_{AXO} par I_D peut être accélérée en augmentant V_{FG}^N . Cependant cela conduit à une augmentation de V_{TH} , car la condition $\partial(I_B - I_D)/\partial V_{AXO}$ nécessaire au déclenchement est vérifiée pour un $I_B = I_{II} + I_{BTBT}$ plus intense, et donc, un V_{MEM} plus haut.

4.4 Le neurone impulsif basé sur le BIMOS

Les sections précédentes ont montré que le schéma électrique du modèle de neurone LIF décrit à la Figure 4-5 (section 4.2) peut être implémenté en utilisant un BB-LTC pour intégrer les fonctions de bascule de Schmitt, de conductance et d'interrupteur. Le circuit de neurone correspondant contient dès lors 3 composants : une capacité, un FD-SOI BIMOS et un FD-SOI NMOS (Figure 4-31). Il est nommé neurone impulsif basé sur le BIMOS (« bimos-based LIF spiking neuron » : BB-LIF SN).

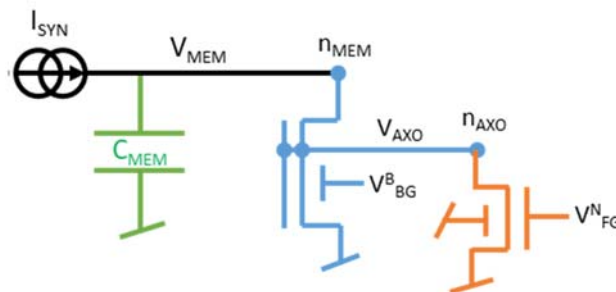


Figure 4-31 – Circuit de neurone impulsif BB-LIF, implémenté à l'aide d'une capacité et d'un BB-LTC composé d'un FD-SOI BIMOS et d'un NMOS.

4.4.1 Preuve de concept du fonctionnement du BB-LIF SN

Une façon d'apporter la preuve de concept du fonctionnement du BB-LIF SN consiste à imposer un courant constant I_{SYN} en entrée d'un BB-LTC et de mesurer la tension V_{MEM} au cours du temps. Dans une telle expérience, la capacité C_{MEM} du BB-LIF SN est implémentée par la capacité parasite du câble de mesure, qui vaut, dans le cas des mesures qui suivent, 6.5 pF (Figure 4-32). Ces mesures sont effectuées à température ambiante sur un BB-LTC aux dimensions et polarisations décrites au Tableau 4-7, à l'aide d'un analyseur de paramètres de semi-conducteur HP 4156C. Deux régimes de comportement, fonctionnel et non-fonctionnel, sont observés selon l'intensité du courant synaptique I_{SYN} .

Tableau 4-7 - Dimension du BB-LTC utilisé pour les mesures de $V_{MEM} = f(t)$.

L_1 (nm)	W_1 (nm)	L_2 (nm)	W_2 (nm)	L (nm)	W (nm)	EOT (nm)	BIMOS	NMOS	V_{FG}^N	V_{BG}
135	144	144	350	135	350	3.4	LVT	RVT	0 V	0 V

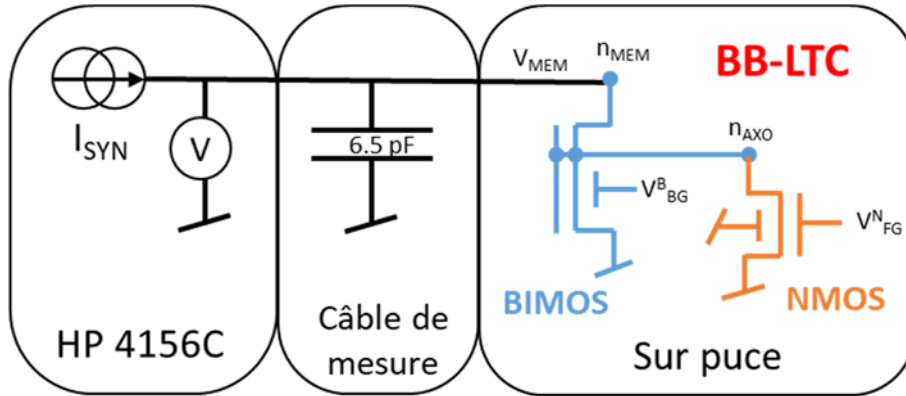


Figure 4-32 – Schéma du montage utilisé pour les mesures apportant la preuve de concept du fonctionnement du neurone impulsif BB-LIF.

4.4.1.1 Régime de comportement fonctionnel du BB-LIF SN

Un premier régime de comportement fonctionnel est identifié pour des intensités de courant I_{SYN} jusqu'à 3 nA (Figure 4-33).

Dans ce régime, le comportement LIF est observé :

- Lorsque le courant I_{SYN} est suffisamment fort (ici, $I_{SYN} \geq 10$ pA), le potentiel V_{MEM} croît jusqu'à atteindre V_{TH} puis chute brusquement lorsque le BB-LTC passe à l'état ON. La chute de V_{MEM} est due au déchargement de la capacité C_{MEM} au travers du BIMOS. Une fois la capacité déchargée, le BB-LTC repasse à l'état OFF et la tension V_{MEM} augmente à nouveau.
- Lorsque le courant I_{SYN} est trop faible ($I_{SYN} < 6$ pA), le potentiel V_{MEM} croît jusqu'à atteindre une valeur plafond. A cette valeur de V_{MEM} , la conductance collecteur – émetteur du BIMOS du BB-LTC à l'état OFF est suffisante pour faire circuler le courant I_{SYN} .

La fréquence de déclenchement f en fonction du courant synaptique I_{SYN} est extraite (Figure 4-34.a). La relation est linéaire (f est proportionnelle à I_{SYN}), et s'étend sur 2 décades. De ce point de vue, le BB-LIF SN implémente un neurone dont l'excitabilité est de classe 1, d'après la classification exposée dans [16].

Le potentiel de seuil V_{TH} en fonction du courant synaptique I_{SYN} est extrait (Figure 4-34.b). Il n'est pas constant. On retrouve ici un aspect du comportement dynamique du BB-LTC déjà évoqué à la section 4.3.7.1. En effet, à chaque courant synaptique I_{SYN} correspond une vitesse de montée du potentiel V_{MEM} dV_{MEM}/dt qui vaut :

$$\frac{dV_{MEM}}{dt} = \frac{1}{C_{MEM}} \cdot I_{SYN} \quad [V \cdot s^{-1}] \quad (4.20)$$

Or, les mesures I_C - V_{MEM} exposées à la section 4.3.7.1 ont montré que le potentiel de seuil V_{TH} dépend de dV_{MEM}/dt .

Par ailleurs, les chronogrammes $V_{MEM} = f(t)$ mesurés montrent que, pour $I_{SYN} \geq 2$ nA, le potentiel V_{TH} du premier déclenchement est supérieur au V_{TH} des déclenchements suivants. Cette chute de V_{TH} après le premier déclenchement montre que, lorsque le courant synaptique I_{SYN} est suffisamment intense, le

potentiel V_{AXO} ne redescend pas à 0 V avant que le déclenchement suivant advienne. En effet, la transition complète du BB-LTC vers l'état OFF se fait sur une durée non nulle. Cette durée correspond au temps que met la charge Q_{AXO} pour être évacuée de la capacité interne C_{AXO} par le courant I_{PN} de la gated-diode base – émetteur intrinsèque au BIMOS et par le courant de drain I_D du NMOS. Or, après un premier déclenchement et lorsque, sous l'effet d'un courant I_{SYN} intense, dV_{MEM}/dt est trop rapide, les courants I_{BTBT} et I_{II} sont réactivés avant que toute la charge Q_{AXO} soit évacuée. Ainsi, la capacité C_{AXO} est préchargée et un courant I_B plus faible suffit à ajouter les charges nécessaires au déclenchement, qui advient donc pour une tension V_{MEM} inférieure.

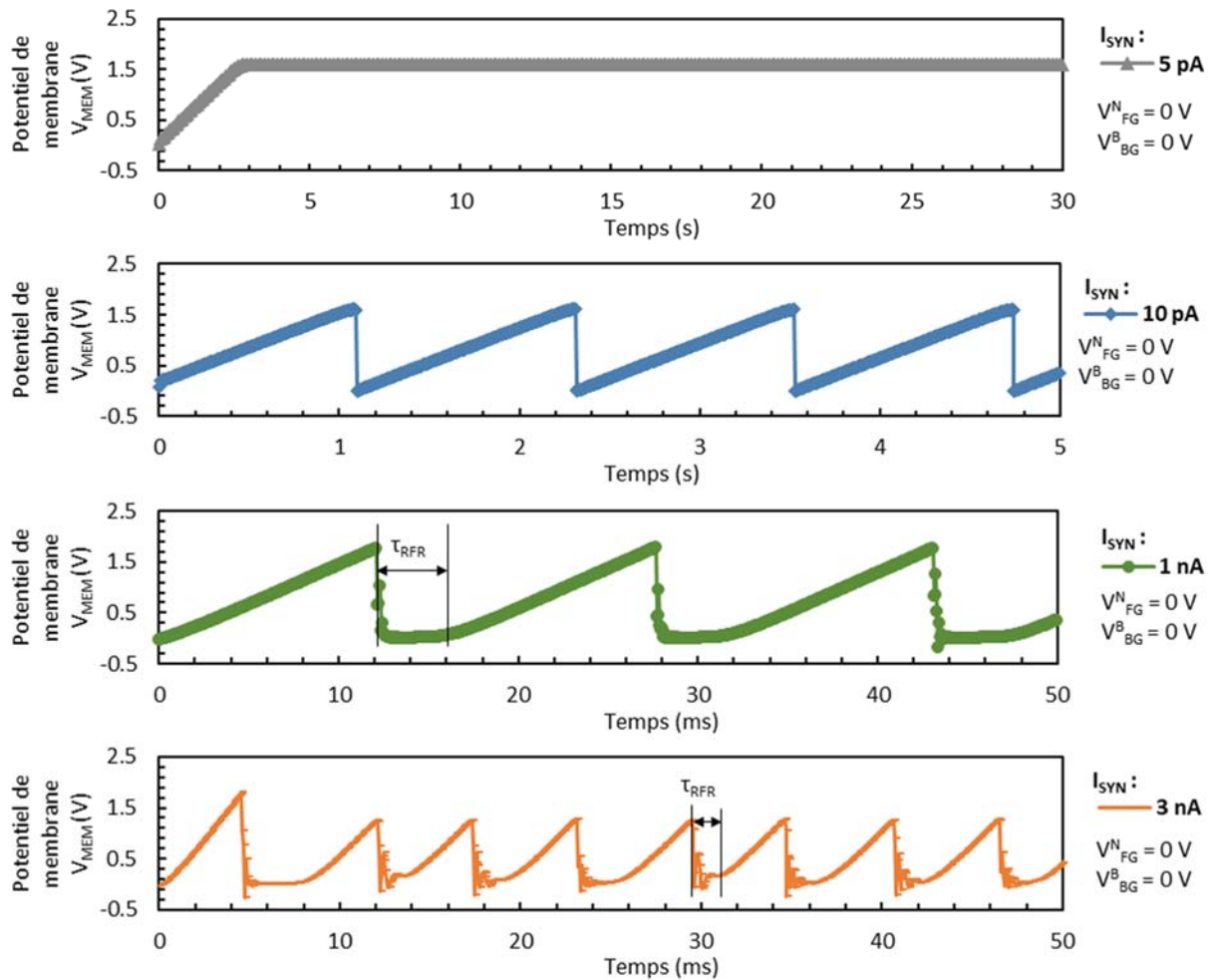
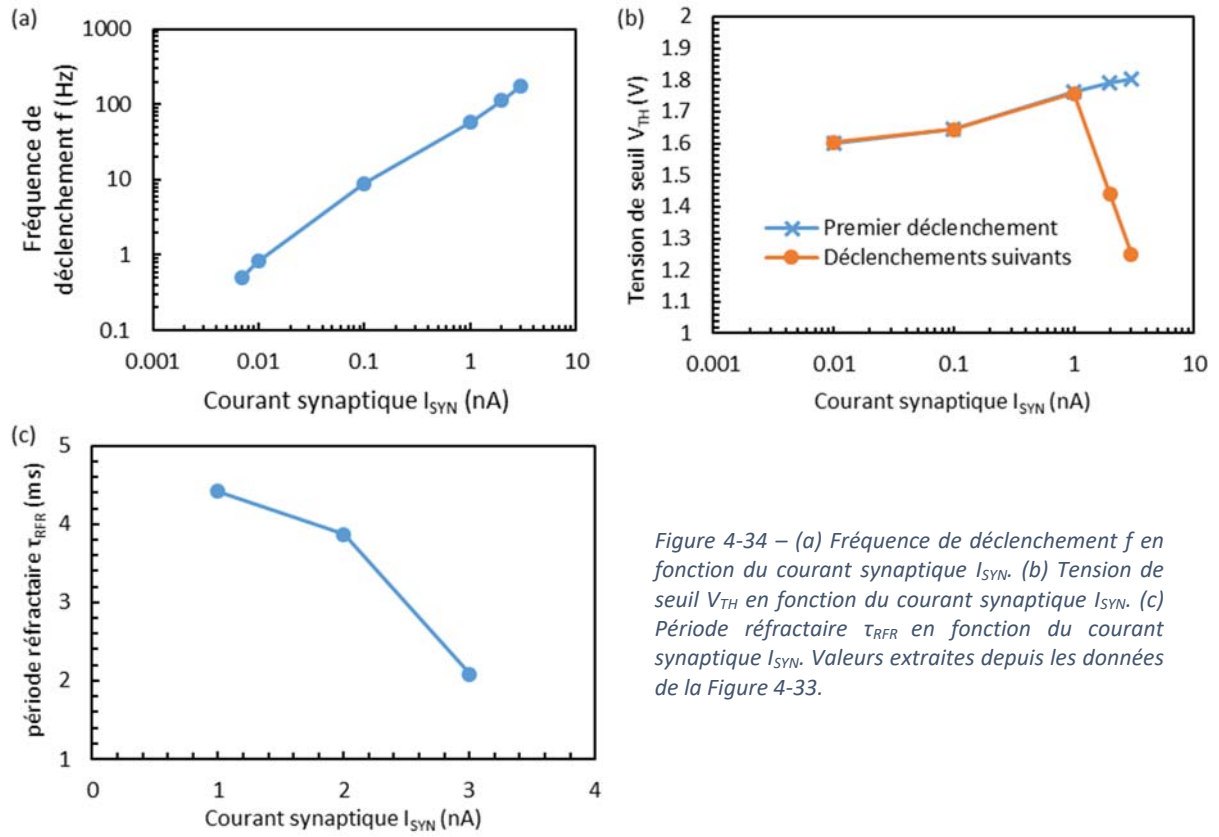


Figure 4-33 – Chronogrammes $V_{MEM} = f(t)$ pour différents courants I_{SYN} entre 10 pA et 3 nA obtenus par mesure à température ambiante d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7.

Notons ensuite que pour $I_{SYN} \geq 1$ nA, le potentiel de membrane V_{MEM} demeure proche de 0 V pendant une durée significative après avoir chuté lors du déclenchement. Ce comportement est caractéristique d'une période réfractaire pendant laquelle le neurone ne peut pas être excité. Comme pour la chute du potentiel V_{TH} , ce phénomène s'explique par la durée non nulle que met le circuit pour passer de l'état ON à l'état OFF. En effet, au moment de la chute du potentiel V_{MEM} , le potentiel V_{AXO} est au plus et la conductance collecteur – émetteur du BIMOS est maximale. Dans les instants suivants et malgré que V_{AXO} commence à décroître, un faible potentiel V_{MEM} est suffisant pour faire circuler le courant I_{SYN} au travers du BIMOS. Le potentiel V_{MEM} croît à nouveau lorsque le potentiel V_{AXO} a suffisamment diminué pour que la conductance collecteur – émetteur chute et que les charges fournies par I_{SYN} chargent C_{MEM} au lieu d'être évacuées directement par le BIMOS. La durée de la période réfractaire τ_{RFR}

est extraite des courbes de la Figure 4-33 (Figure 4-34.c). Elle est de l'ordre de quelques ms et décroît lorsque I_{SYN} augmente.



4.4.1.2 Régime de comportement non-fonctionnel du BB-LIF SN

Lorsque l'intensité du courant synaptique I_{SYN} est trop forte, le comportement LIF n'est plus observé (Figure 4-35). Après un premier déclenchement et une première période réfractaire, le potentiel V_{MEM} croît puis oscille sans retomber à $V_{MEM} = 0$ V. Le BB-LTC est bloqué dans un état intermédiaire entre ses états ON et OFF. Encore une fois, l'analyse et la comparaison des dynamiques de charge et de décharge de la capacité C_{AXO} expliquent ce phénomène.

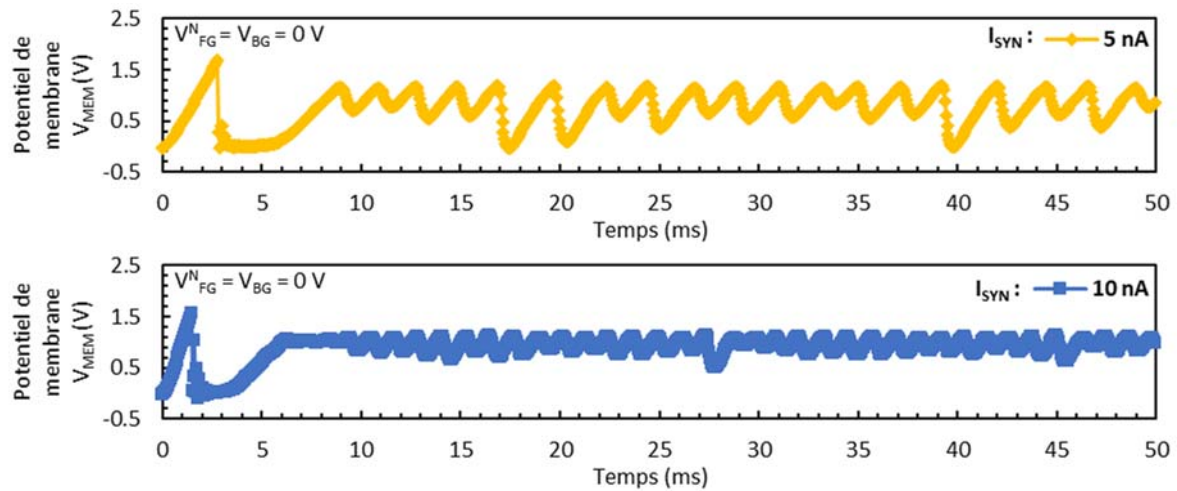


Figure 4-35 – Chronogrammes $V_{MEM} = f(t)$ pour des courants I_{SYN} de 5 nA et 10 nA obtenus par mesure à température ambiante d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7.

4.4.2 Influence du potentiel V_{FG}^N sur le fonctionnement du BB-LIF SN

Le courant I_D qui circule dans le NMOS joue un rôle de premier plan dans le fonctionnement du BB-LTC et, par extension, du BB-LIF SN. En effet, d'un point de vue statique, I_D fixe V_{TH} qui est le seuil de polarisation de V_{MEM} pour lequel la boucle de rétroaction positive qui fait basculer le BB-LTC de l'état ON à l'état OFF est activée. D'un point de vue dynamique, I_D pilote la durée de décharge de C_{AXO} qui a été identifiée comme un élément déterminant dans la limitation de la plage de fonctionnement de BB-LTC SN. I_D est fixé par le potentiel de la grille-avant du NMOS V_{FG}^N , selon l'équation (2.53). Sa valeur a donc un impact significatif sur le comportement du BB-LIF SN.

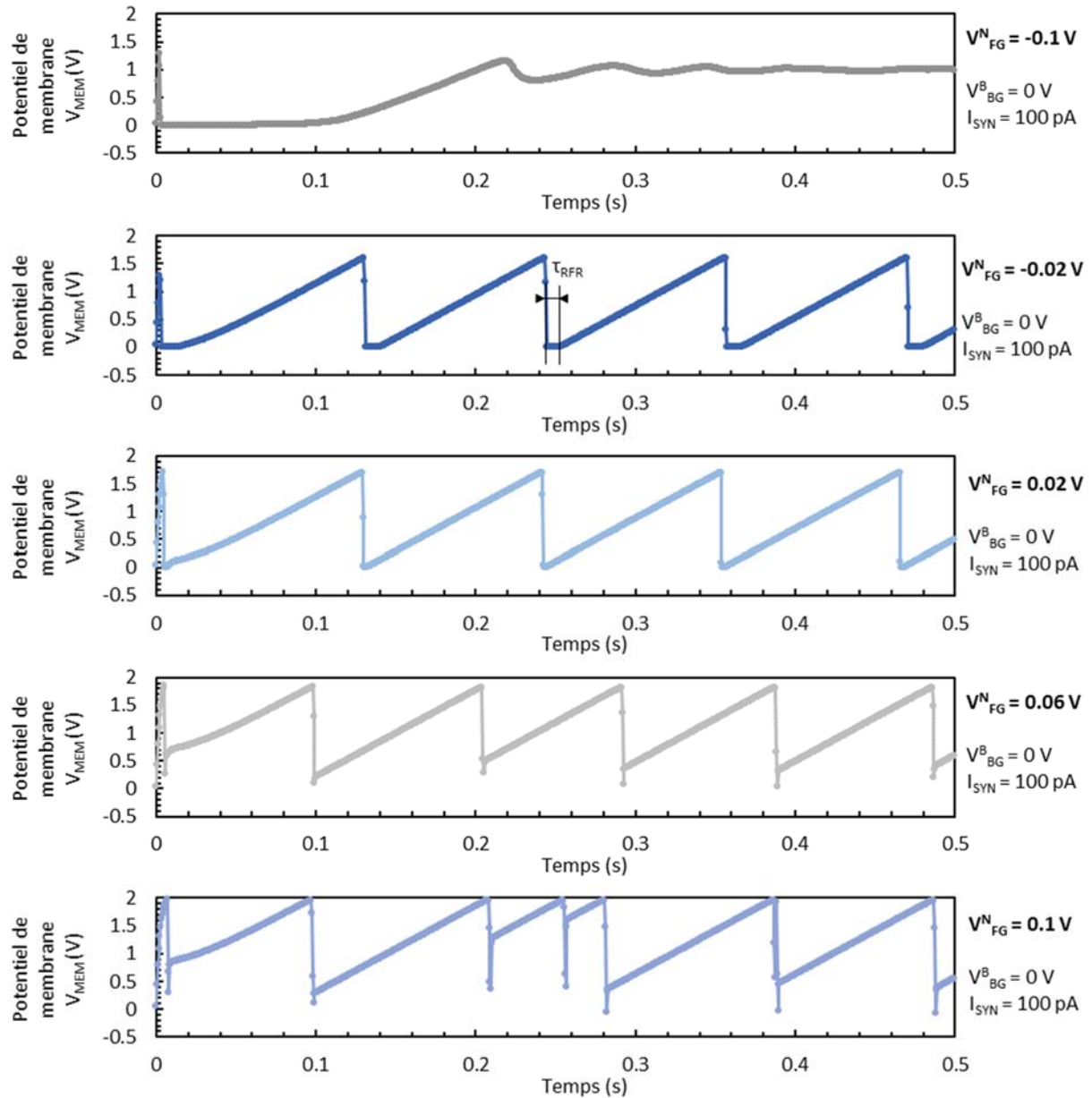


Figure 4-36 – Chronogrammes $V_{MEM} = f(t)$ à $I_{SYN} = 100$ pA pour différentes valeurs du potentiel de grille-avant V_{FG}^N obtenus par mesure à température ambiante d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7

Dans cette section, l'influence du potentiel V_{FG}^N sur le fonctionnement du BB-LIF SN est étudié. Le potentiel de membrane V_{MEM} du BB-LTC décrit au Tableau 4-7 est mesuré au cours du temps et à température ambiante pour un courant synaptique I_{SYN} de 100 pA et pour différentes valeurs de V_{FG}^N comprises entre -0.1 V et 0.1 V. Les chronogrammes obtenus sont présentés à la Figure 4-36. La

fréquence de déclenchement f , la tension de seuil V_{TH} et la durée de la période réfractaire τ_{RFR} sont extraites (Figure 4-37). Deux régimes de comportement sont identifiés, selon que le BB-LIF SN est fonctionnel ou non.

4.4.2.1 Régime de comportement fonctionnel du BB-LIF SN

Pour V_{FG}^N entre -0.06 V et 0.02 V, le BB-LIF SN est fonctionnel. L'augmentation du potentiel V_{FG}^N augmente la tension de seuil V_{TH} (Figure 4-37.b) en accord avec les mesures statiques I_C - V_{MEM} de la section 4.3.5.1. Par ailleurs, elle réduit la durée de la période réfractaire τ_{RFR} avec une relation de type exponentielle (Figure 4-37.c). Cette relation s'explique par le fait que, après le déclenchement et après qu'une partie de Q_{AXO} ait été évacuée par I_{PN} , le reste de la charge Q_{AXO} stockée dans C_{AXO} est évacuée par le courant I_D du NMOS. Or, dans le régime de polarisation considéré, le NMOS est en régime sous le seuil et I_D a une dépendance exponentielle en V_{FG}^N (NMOS est en régime sous le seuil).

En ce qui concerne la fréquence de déclenchement f , l'augmentation de V_{TH} a pour effet de la réduire, tandis que la réduction de τ_{RFR} a pour effet de l'augmenter. La combinaison de ces deux effets opposés semble les compenser pour conduire à une valeur de f relativement stable pour les valeurs de V_{FG}^N fonctionnel (Figure 4-37.a).

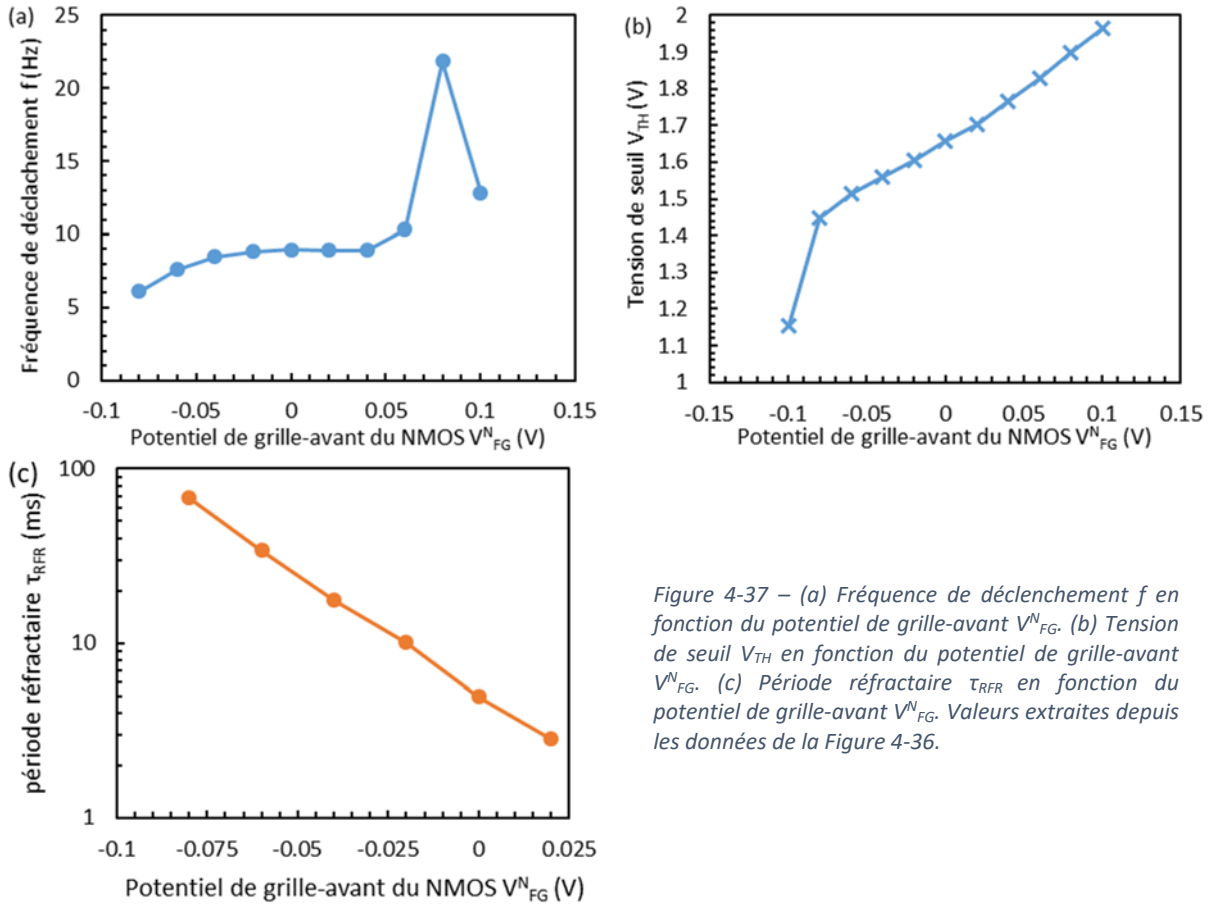


Figure 4-37 – (a) Fréquence de déclenchement f en fonction du potentiel de grille-avant V_{FG}^N . (b) Tension de seuil V_{TH} en fonction du potentiel de grille-avant V_{FG}^N . (c) Période réfractaire τ_{RFR} en fonction du potentiel de grille-avant V_{FG}^N . Valeurs extraites depuis les données de la Figure 4-36.

4.4.2.2 Régime de comportement non-fonctionnel du BB-LIF SN

Pour $V_{FG}^N \leq -0.1$ V, le BB-LIF SN n'est pas fonctionnel. Après un premier déclenchement, le potentiel de membrane V_{MEM} croît et se stabilise à une valeur autour de 1 V. La ressemblance de la courbe avec celles obtenues pour des courants forts à $V_{FG}^N = 0$ V (Figure 4-35) est significative. Elle suggère que,

dans les deux cas, le mode de défaillance du BB-LTC SN est issu du même mécanisme : une stabilisation du potentiel V_{AXO} à une valeur intermédiaire entre 0 V et $V_{T,MOS}$.

Pour $V_{FG}^N \geq 0.6$ V, le potentiel de membrane ne redescend pas jusqu'à 0 V entre les déclenchements mais reste à des valeurs positives qui peuvent être relativement élevées notamment lorsque $V_{FG}^N = 0.1$ V. Lors de ces événements, la capacité C_{MEM} ne se décharge pas entièrement lors du bref passage à l'état ON du BB-LTC. Un tel comportement constitue un autre mode de défaillance du BB-LIF SN. Il correspond à une évacuation trop rapide de la charge Q_{AXO} au travers du NMOS après le déclenchement : le BIMOS se rouvre avant que toute la charge stockée dans C_{MEM} ne soit évacuée.

4.4.2.3 Résumé

Le potentiel de grille-avant V_{FG}^N du NMOS modifie la tension de seuil V_{TH} ($\partial V_{TH}/\partial V_{FG}^N > 0$) par son influence sur le courant I_D impliqué dans le critère de déclenchement du BB-LTC $\partial(I_B - I_D)/\partial V_{AXO} > 0$. Par ailleurs, V_{FG}^N modifie la durée de la période réfractaire τ_{RFR} ($\partial \tau_{RFR}/\partial V_{FG}^N < 0$) par l'implication du courant I_D dans la décharge de C_{AXO} après le déclenchement. L'augmentation de V_{TH} et la réduction de τ_{RFR} avec V_{FG}^N ont des effets opposés sur la variation de la fréquence de déclenchement f , qui, en conséquence, varie peu avec V_{FG}^N . Le BB-LIF SN devient non fonctionnel si V_{FG}^N est trop faible ou trop fort. Pour V_{FG}^N trop faible, la charge Q_{AXO} est évacuée trop lentement par I_D , le potentiel V_{AXO} se stabilise à une valeur pour laquelle $I_{SYN} = I_{CH}$ et $I_B = I_D$. Pour V_{FG}^N trop fort, la charge Q_{AXO} est évacuée trop rapidement par I_D , si bien que la charge stockée dans C_{MEM} n'est pas entièrement évacuée par le BIMOS lorsque celui-ci se rouvre.

4.4.3 Influence du potentiel V_{BG}^B sur le fonctionnement du BB-LIF SN

Comparée aux technologies bulk ou PD-SOI, la technologie 28nm UTBB FD-SOI offre un levier supplémentaire de design grâce au fort couplage capacitif au travers du BOX entre le film de silicium et le back plane (partie du well à l'interface avec le BOX). En particulier, la polarisation du well permet de décaler la tension de seuil $V_{T,MOS}$ des MOSFET. Dans ce contexte, le well peut être considéré comme une grille-arrière dont le potentiel est V_{BG} . A la section 4.3.5.3, les mesures quasi-statiques de la courbe I_C - V_{MEM} du BB-LTC avec V_{BG}^B comme paramètre ont d'ores et déjà montré l'influence de ce potentiel sur les grandeurs caractéristiques du BB-LTC.

Dans cette section, l'influence du potentiel de grille-arrière du BIMOS V_{BG}^B sur le fonctionnement du BB-LIF SN est étudiée. Le potentiel de membrane V_{MEM} du BB-LTC décrit au Tableau 4-7 est mesuré au cours du temps et à température ambiante pour un courant synaptique I_{SYN} de 1 nA et pour différentes valeurs de V_{BG}^B comprises entre 0 V et 1 V. Comme la réduction du potentiel V_{BG}^B augmente la tension de seuil V_{TH} (c.f. Figure 4-24) que l'on souhaite a priori réduire, les valeurs de V_{BG}^B négatives ne sont pas explorées. Les chronogrammes obtenus sont présentés à la Figure 4-38. La fréquence de déclenchement f , la tension de seuil V_{TH} et la durée de la période réfractaire τ_{RFR} sont extraites (Figure 4-39). A nouveau, deux régimes, l'un fonctionnel et l'autre non, sont observés.

Le BB-LIF SN est fonctionnel pour $V_{BG}^B \in [0 ; 0.8]$ V. Comme attendu d'après la section 4.3.5.3, le potentiel de seuil V_{TH} décroît avec l'augmentation de V_{BG}^B (Figure 4-39.b). A l'inverse, la période réfractaire τ_{RFR} augmente avec V_{BG}^B (Figure 4-39.c). Comme la polarisation des grilles du NMOS demeure inchangée, aucune variation du courant I_D ne peut expliquer cette variation de τ_{RFR} . En revanche, un surplus de charge Q_{AXO} stocké dans C_{AXO} après les déclenchements, dû à l'augmentation du courant I_{II} avec V_{BG}^B (c.f. Figure 4-23), pourrait rendre compte de la dépendance de τ_{RFR} en V_{BG}^B . Quoiqu'il en soit, les variations opposées de V_{TH} et τ_{RFR} , conduisent à une relative indépendance de la

fréquence de déclenchement en V_{BG}^B (Figure 4-39.a). A $V_{BG} = 1$ V, le BB-LIF SN n'est plus fonctionnel. La forme de la courbe suggère à nouveau que le potentiel V_{AXO} est en train de se stabiliser.

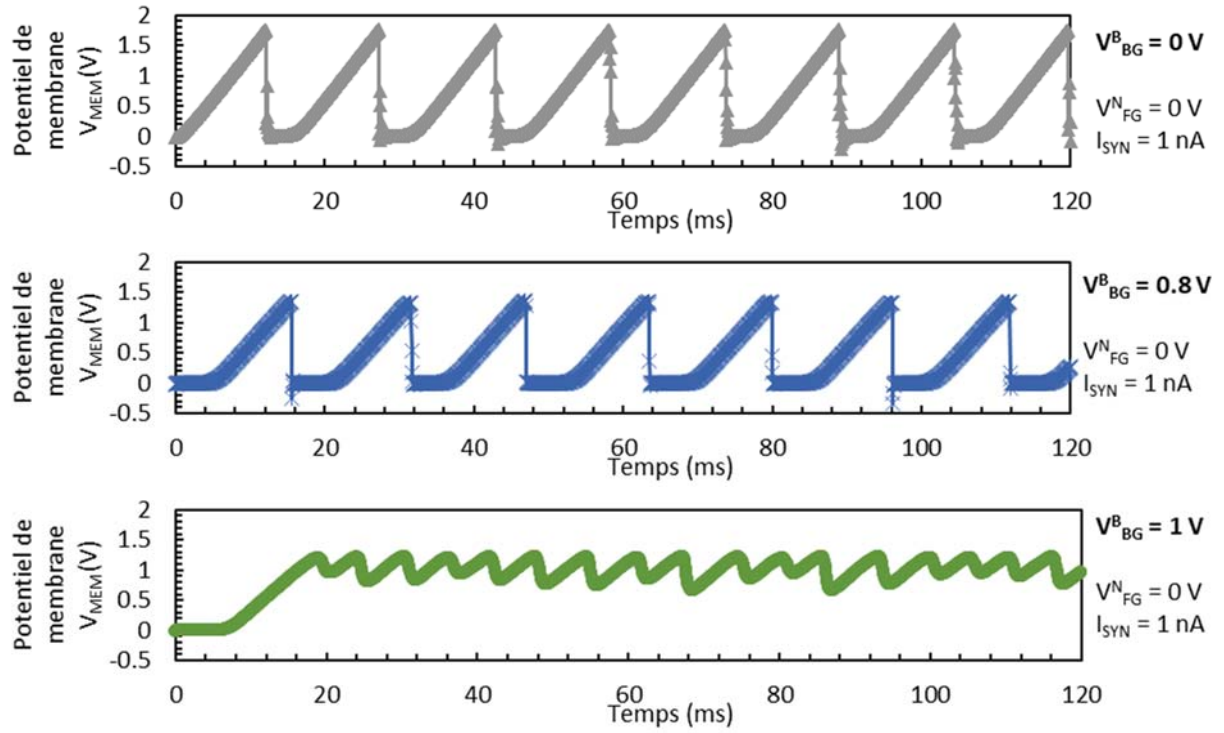


Figure 4-38 – Chronogrammes $V_{MEM} = f(t)$ à $I_{SYN} = 1$ nA pour différentes valeurs du potentiel de grille-arrière V_{BG}^B obtenus par mesure à température ambiante d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7

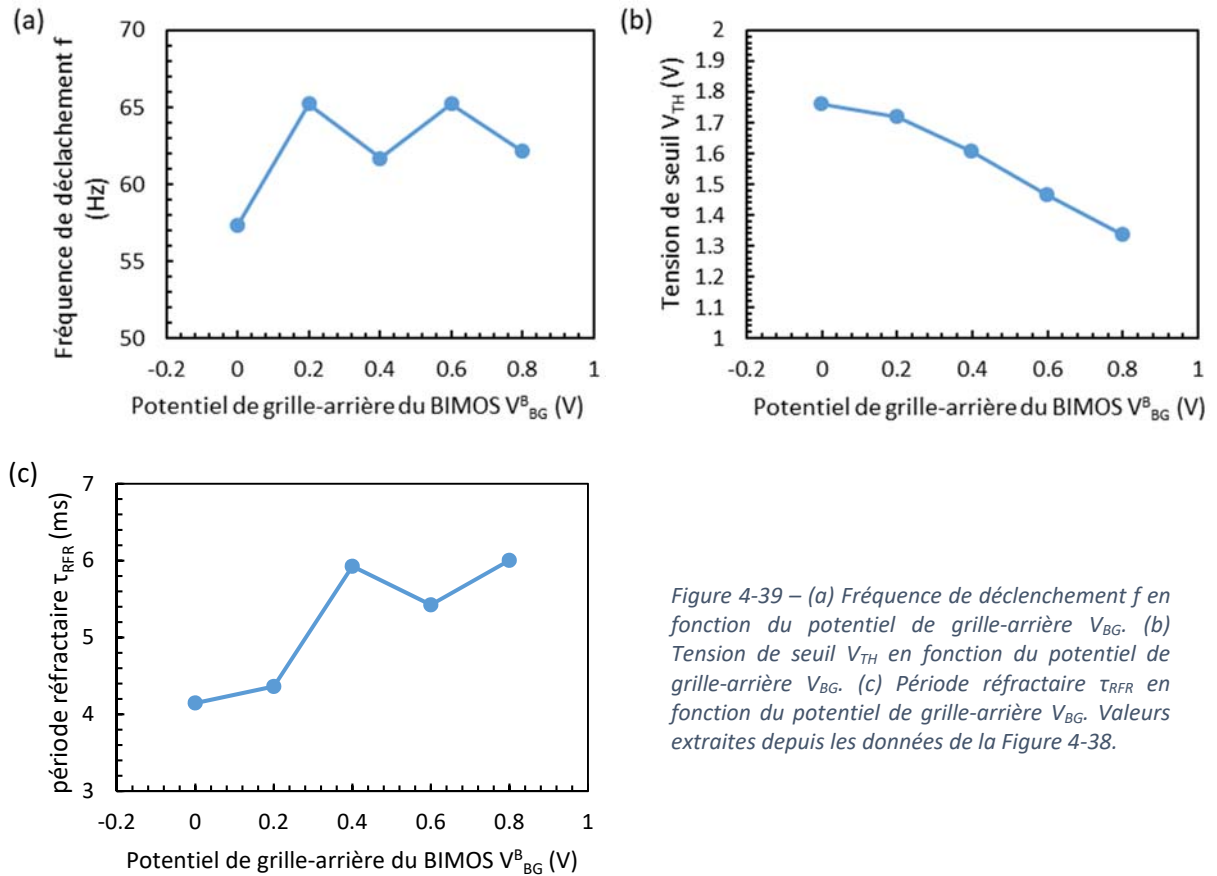


Figure 4-39 – (a) Fréquence de déclenchement f en fonction du potentiel de grille-arrière V_{BG} . (b) Tension de seuil V_{TH} en fonction du potentiel de grille-arrière V_{BG} . (c) Période réfractaire τ_{RFR} en fonction du potentiel de grille-arrière V_{BG} . Valeurs extraites depuis les données de la Figure 4-38.

4.4.4 Confrontation du modèle SPICE aux mesures $V_{MEM} = f(t)$

Dans cette section, la réponse temporelle du BB-LIF SN sous courant I_{SYN} constant est simulée sous ELDO en utilisant le modèle SPICE du BIMOS développé au chapitre 3. Elle est ensuite comparée aux mesures $V_{MEM} = f(t)$ effectuées à la section 4.4.1.

Le schéma électrique simulé est celui de la Figure 4-31. La capacité C_{MEM} est considérée comme idéale et vaut 6.5 pF. Les paramètres de design du BB-LTC sont ceux du Tableau 4-7. Les potentiels V_{BG}^B et V_{FG}^N sont maintenus à 0 V pour permettre la comparaison avec les mesures de la section 4.4.1. Les valeurs d'intensité du courant I_{SYN} explorées sont comprises entre 1 nA et 5 nA. Elles correspondent à la limite de fonctionnement du BB-LIF-SN observée en mesure. Les chronogrammes $V_{MEM} = f(t)$ obtenus pour $I_{SYN} = 1$ nA et 5 nA sont présentés à la Figure 4-40. Les courbes $V_{AXO} = f(t)$, disponibles en simulation mais pas en mesure, sont également tracées. La fréquence de déclenchement f , le potentiel de seuil V_{TH} et la durée de la période réfractaire τ_{RFR} sont extraits et comparés aux valeurs obtenues en mesure en fonction du courant synaptique I_{SYN} (Figure 4-41.a-c). Le potentiel d'axone maximal $V_{AXO,MAX}$ obtenu en simulation est également tracé (Figure 4-41.d).

4.4.4.1 Comportement global

Le comportement global simulé $V_{MEM} = f(t)$ est conforme à celui observé en mesure, que ce soit pour un régime de comportement fonctionnel ($I_{SYN} = 1$ nA, Figure 4-40.a) ou pour un régime de comportement non-fonctionnel ($I_{SYN} = 5$ nA, Figure 4-40.b). Cependant, des désaccords entre la simulation et la mesure sur les valeurs de f , V_{TH} et τ_{RFR} sont observés. Ils proviennent de l'imprécision du modèle SPICE du BIMOS déjà constatée au chapitre 3.

En régime fonctionnel ($I_{SYN} \leq 3$ nA), le potentiel V_{MEM} augmente linéairement au cours du temps sous l'effet du chargement de la capacité C_{MEM} par le courant synaptique I_{SYN} . Lorsque V_{MEM} atteint le potentiel de seuil V_{TH} , la capacité se décharge brutalement et V_{MEM} redescend à 0 V. Le chargement de C_{MEM} est ensuite empêché pendant une durée τ_{RFR} , puis le cycle recommence. Comme en mesure, la fréquence de déclenchement f est proportionnelle au courant synaptique I_{SYN} . Le décalage de f observé entre la simulation et la mesure (Figure 4-41.a) est une conséquence des décalages de V_{TH} et de τ_{RFR} (Figure 4-41.b et c). La réduction de V_{TH} après le premier déclenchement lorsque que l'on se rapproche de la limite de fonctionnement est aussi observée en simulation (Figure 4-40.b).

En régime non-fonctionnel, le potentiel V_{TH} se réduit après le premier déclenchement et la décharge de C_{MEM} n'est plus complète (Figure 4-40.b), conformément à ce qui est observé en mesure (Figure 4-35). Il en résulte une augmentation du coefficient de proportionnalité entre le courant I_{SYN} et la fréquence f (Figure 4-41.a).

4.4.4.2 Potentiel d'axone V_{AXO}

La simulation permet de relever le potentiel d'axone V_{AXO} au cours du temps (courbes oranges de la Figure 4-40). Comme attendu, les courbes $V_{AXO} = f(t)$ révèlent la présence d'un pic de potentiel lors des déclenchements de la structure, dont la valeur maximale $V_{AXO,MAX}$ est de 0.74 V pour $I_{SYN} = 1$ nA. Ce pic est composé d'un sommet étroit dont la durée caractéristique est de l'ordre de la microseconde, et d'une base large et asymétrique dont la durée caractéristique est de l'ordre de la milliseconde. Autrement dit, après que V_{AXO} a atteint rapidement sa valeur maximale, sa décroissance est d'abord rapide puis lente. Pour rappel, la croissance rapide de V_{AXO} est due au chargement de C_{AXO} par le courant I_{II} généré lors de la décharge brutale de C_{MEM} au travers du BIMOS. Concernant la décroissance de V_{AXO} , sa partie rapide entre $V_{AXO} = V_{AXO,MAX}$ et $V_{AXO} \approx 0.5$ V est due au déchargement de C_{AXO} par le

courant de gated-diode base – émetteur I_{PN} , tandis que sa partie lente est due au déchargement de C_{AXO} par le courant de drain I_D du NMOS. En effet, le courant I_D est constant (tant que V_{AXO} est supérieur à quelques kT/q), tandis que le courant I_{PN} décroît exponentiellement avec V_{AXO} . Ainsi, les courants I_{PN} et I_D vérifient $I_{PN} \gg I_D$ pour les hautes valeurs de V_{AXO} , et $I_{PN} \ll I_D$ pour les faibles valeurs de V_{AXO} .

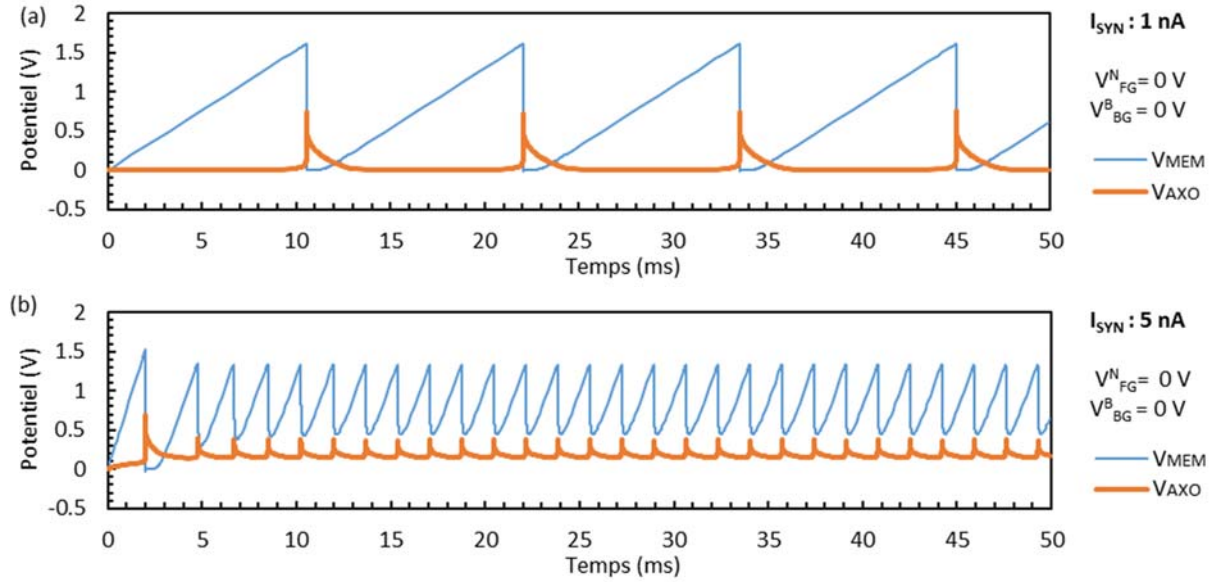


Figure 4-40 - Chronogrammes $V_{MEM} = f(t)$ et $V_{AXO} = f(t)$ pour un courant I_{SYN} de 1 nA et 5 nA, obtenus par simulation SPICE d'un BB-LTC dont les caractéristiques sont présentées au Tableau 4-7

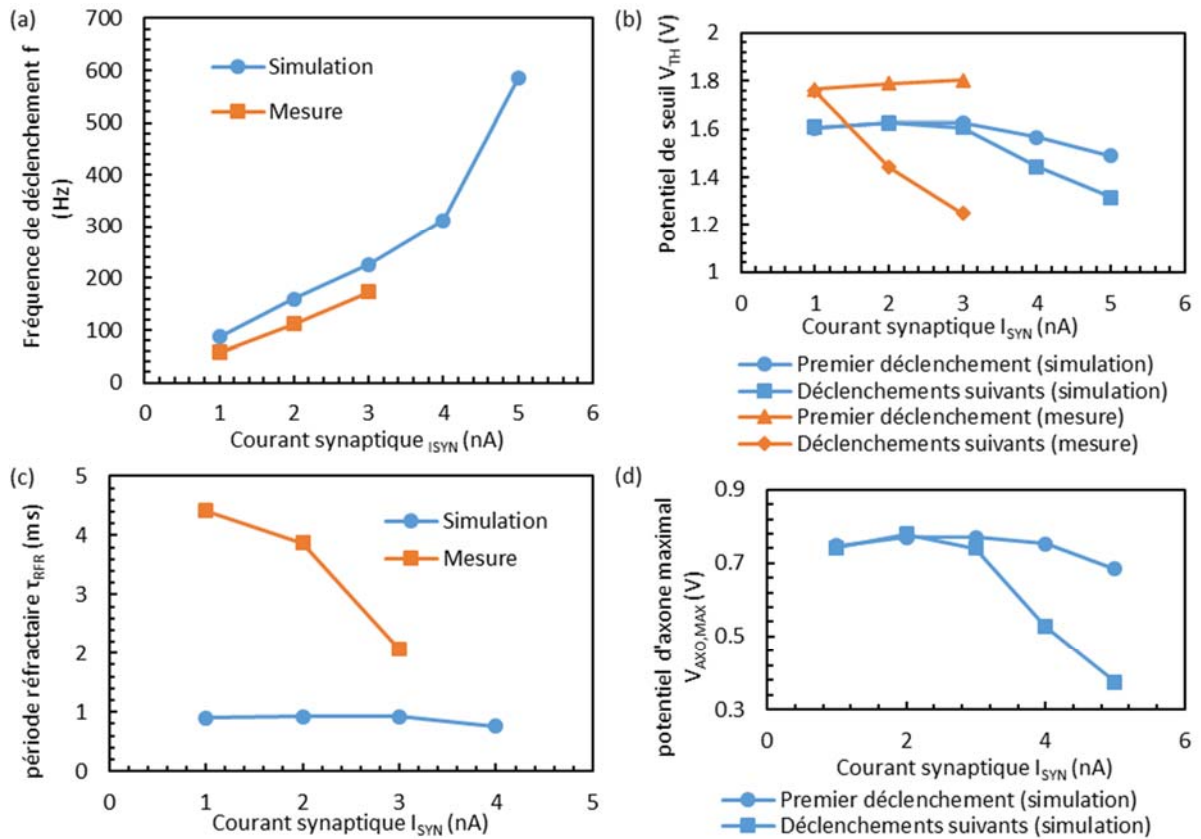


Figure 4-41 – (a) Fréquence de déclenchement f , (b) potentiel de seuil V_{TH} , (c) période réfractaire τ_{RFR} et (d) potentiel d'axone maximal $V_{AXO,MAX}$, en fonction du courant synaptique I_{SYN} , extraits de données issues de simulations et de mesures.

Les courbes $V_{AXO} = f(t)$ confirment l'explication donnée à la section 4.4.1.1 sur la raison de la présence de la période réfractaire : elle est due à la durée de la décroissance du potentiel d'axone V_{AXO} qui est relativement longue une fois que celle-ci est pilotée par le courant I_D .

En régime non-fonctionnel, on constate que l'amplitude des pics de V_{AXO} se réduit : la valeur $V_{AXO,MAX}$ diminue et la valeur minimale de V_{AXO} , qui était proche de 0 V en régime fonctionnel, augmente (elle vaut, par exemple, 150 mV lorsque $I_{SYN} = 5nA$).

4.4.4.3 Bilan

Les simulations SPICE de la réponse temporelle sous différents courants I_{SYN} du BB LIF-SN utilisant le modèle de BIMOS développé au chapitre 3 reproduisent qualitativement les données obtenues par les mesures, notamment en ce qui concerne les sens de variation des grandeurs caractéristiques f , V_{TH} et τ_{REFR} avec I_{SYN} . Des désaccords quantitatifs demeurent néanmoins sur ces grandeurs. Ils sont issus de l'imprécision du modèle SPICE du BIMOS tel que discuté au chapitre 3. D'autres simulations, non présentées ici, montrent que la correspondance qualitative des simulations avec les mesures est également assurée pour V_{BG}^B et V_{FG}^N comme paramètre.

La simulation permet par ailleurs de sonder le potentiel d'axone V_{AXO} et de suivre son évolution au cours du temps. Sa courbe révèle la présence de pics asymétriques aux moments des déclenchements, dont la base relativement large est responsable de la présence de la période réfractaire.

4.4.5 Modélisation des limites de fonctionnement du BB-LTC

Les mesures et les simulations présentées dans les sections précédentes ont montré qu'il existe un courant synaptique minimal $I_{SYN,MIN}$ sous lequel le BB-LIF SN ne se déclenche pas, et un courant synaptique maximal $I_{SYN,MAX}$ au-dessus duquel le BB-LIF SN ne se déclenche plus. A ces valeurs de courant correspondent des valeurs de fréquence de déclenchement extrémales f_{MIN} et f_{MAX} . Ces grandeurs dépendent des paramètres de design (W et L du BIMOS et du NMOS, et valeur de C_{MEM}) et de polarisation (V_{FG}^N et V_{BG}^B). Cette section présente un jeu d'équations simple qui modélise de façon approximative les relations entre ces paramètres et les plages de fonctionnement $[I_{SYN,MIN} ; I_{SYN,MAX}]$ et $[f_{MIN} ; f_{MAX}]$.

4.4.5.1 Limites de fonctionnement inférieures $I_{SYN,MIN}$ et f_{MIN}

Le courant synaptique minimal $I_{SYN,MIN}$ capable de conduire au déclenchement du BB-LTC SN est le courant $I_{OFF,TH}$:

$$I_{SYN,MIN} = I_{OFF,TH} \quad [A] \quad (4.21)$$

En effet, en dessous de ce courant limite, les charges fournies par le courant synaptique I_{SYN} au nœud de membrane n_{MEM} sont évacuées par le collecteur du BIMOS vers la masse sans que la boucle de rétroaction soit activée et que le BB-LTC passe à l'état ON. La durée de montée maximale de V_{MEM} , notée $\tau_{MEM,MAX}$, est alors donnée par :

$$\tau_{MEM,MAX} = \frac{C_{MEM} \cdot V_{TH}}{I_{SYN,MIN}} = \frac{C_{MEM} \cdot V_{TH}}{I_{OFF,TH}} \quad [s] \quad (4.22)$$

La fréquence de déclenchement minimale associée f_{MIN} est :

$$f_{MIN} = \frac{1}{\tau_{MEM,MAX} + \tau_{RFR}} \quad [Hz] \quad (4.23)$$

où τ_{REFR} est la durée de la période réfractaire. Dans le cas où $\tau_{\text{MEM},\text{MAX}} \gg \tau_{\text{REFR}}$, f_{MIN} devient :

$$f_{\text{MIN}} = \frac{I_{\text{OFF,TH}}}{C_{\text{MEM}} \cdot V_{\text{TH}}} \quad [\text{Hz}] \quad (4.24)$$

4.4.5.2 Limites de fonctionnement supérieures $I_{\text{SYN},\text{MAX}}$ et f_{MAX}

Lorsque le courant synaptique I_{SYN} est trop intense, le BB-LIF SN n'est plus fonctionnel. Il se déclenche une première fois, puis il se bloque dans un état intermédiaire entre ses états ON et OFF, où le potentiel d'axone V_{AXO} est maintenu à une valeur non nulle. De manière approximative, ce blocage est activé si le potentiel de membrane V_{MEM} croît plus vite que le potentiel d'axone V_{AXO} décroît au cours du temps : la fréquence de déclenchement maximale est limitée par la durée de la décroissance de V_{AXO} , notée τ_{AXO} .

Comme discuté à la section 4.4.4.2, cette décroissance est d'abord rapide, car conduite par le courant I_{PN} , puis lente, car gouvernée par le courant I_{D} . La bascule entre ces deux régimes de décroissance s'effectue lorsque V_{AXO} atteint une valeur $V_{\text{AXO,INV}}$ où l'ordre de l'intensité des courants I_{PN} et I_{D} s'inverse. En considérant la durée de décroissance conduite par I_{PN} négligeable devant celle conduite par I_{D} qui est alors égale au courant de saturation $I_{\text{D,SAT}}$, la durée τ_{AXO} s'écrit :

$$\tau_{\text{AXO}} = \frac{C_{\text{AXO}} \cdot V_{\text{AXO,INV}}}{I_{\text{D,SAT}}} \quad [\text{s}] \quad (4.25)$$

où C_{AXO} est la capacité de nœud d'axone n_{AXO} . Notons que l'établissement de (4.25) s'appuie sur les hypothèses où C_{AXO} et I_{D} sont indépendants de V_{AXO} . Cela est vérifié pour I_{D} tant que V_{AXO} est supérieur à quelques $k \cdot T/q$, mais pas pour C_{AXO} . En effet la capacité C_{AXO} est constitué de la mise en parallèle des capacités grille – collecteur, grille – émetteur et grille – substrat du BIMOS et des capacités drain – substrat, drain – grille et drain – source du NMOS, qui sont toutes des capacités non linéaires. Néanmoins, il est possible d'évaluer sa valeur par une valeur moyenne. La fréquence de déclenchement maximale f_{MAX} est alors donnée par :

$$f_{\text{MAX}} = \frac{1}{\tau_{\text{AXO}}} \quad [\text{Hz}] \quad (4.26)$$

Elle correspond à une valeur de courant synaptique $I_{\text{SYN},\text{MAX}}$ de :

$$I_{\text{SYN},\text{MAX}} = C_{\text{MEM}} \cdot V_{\text{TH}} \cdot f_{\text{MAX}} \quad [\text{A}] \quad (4.27)$$

4.4.5.3 Application au circuit mesuré

Dans le but de valider le modèle d'évaluation des limites de fonctionnement du BB-LIF SN, les équations (4.21) à (4.27) sont appliquées au circuit mesuré à la section 4.4.1. Les valeurs des grandeurs nécessaires aux calculs de $I_{\text{SYN},\text{MIN}}$, $I_{\text{SYN},\text{MAX}}$, f_{MIN} et f_{MAX} du BB-LIF SN sont répertoriées au Tableau 4-8. Les valeurs de C_{MEM} , V_{TH} et $I_{\text{OFF,TH}}$ proviennent des données mesurées, tandis que les valeurs de $C_{\text{AXO,INV}}$, $V_{\text{AXO,INV}}$ et I_{D} sont extraites de simulations SPICE.

Tableau 4-8 – Grandeurs nécessaires aux calculs de $I_{\text{SYN},\text{MIN}}$, $I_{\text{SYN},\text{MAX}}$, f_{MIN} et f_{MAX} du BB-LIF SN mesuré à la section 4.4.1

C_{MEM}	$C_{\text{AXO,INV}}$	V_{TH}	$V_{\text{AXO,INV}}$	$I_{\text{OFF,TH}}$	I_{D}
6.5 pF	360 aF	1.7 V	0.25 V	5.8 pA	5.23 fA

Tableau 4-9 – Comparaison de grandeurs $I_{\text{SYN},\text{MIN}}$, $I_{\text{SYN},\text{MAX}}$, f_{MIN} et f_{MAX} obtenues par le calcul et par la mesure.

Type de donnée	$I_{\text{SYN},\text{MIN}}$	f_{MIN}	$I_{\text{SYN},\text{MAX}}$	f_{MAX}
Mesurée	6 pA	0.50 Hz	3 nA	173 Hz
Calculée	5.8 pA	0.52 Hz	3.84 nA	581 Hz

Les valeurs de $I_{\text{SYN},\text{MIN}}$, $I_{\text{SYN},\text{MAX}}$, f_{MIN} et f_{MAX} obtenues sont comparées à leurs valeurs mesurées au Tableau 4-9. La modélisation est plutôt en accord avec les mesures, à l'exception de l'évaluation de f_{MAX} , dont la valeur calculée est plus de trois fois supérieure à la valeur mesurée. Néanmoins, les ordres de grandeur sont comparables et l'on peut admettre que la valeur mesurée est majorée par la valeur calculée. Ces résultats montrent que le modèle est un outil efficace pour mettre en cohérence les paramètres de design lors de la conception du BB-LIF SN.

4.4.5.4 Limites du modèle

Il est important de rappeler ici que, outre les dépendances de V_{TH} en V_{FG}^{N} et V_{BG}^{B} observées lors des mesures quasi statiques (voir section 4.3.5), V_{TH} dépend également de la vitesse de montée dV_{MEM}/dt lorsque qu'elle est supérieure à 1 V/ms environ (voir section 4.3.7.1). Comme d'autres grandeurs caractéristiques dépendent de V_{TH} , comme $I_{\text{OFF,TH}}$ par exemple, le vitesse dV_{MEM}/dt a aussi un impact sur ces grandeurs. Or, on montre facilement que $dV_{\text{MEM}}/dt = I_{\text{SYN}}/C_{\text{MEM}}$ (dans les cas où I_{SYN} est très supérieur au courant de fuite par le BIMOS). Ainsi, dès que $I_{\text{SYN}}/C_{\text{MEM}}$ est supérieur à 1 V/ms, il n'est plus pertinent d'utiliser les grandeurs V_{TH} et $I_{\text{OFF,TH}}$ obtenues par la mesures quasi statiques. Ainsi, à mesure que C_{MEM} diminue, le modèle perd en validité.

4.5 Intégration du BB-LIF SN en 28nm FD-SOI

L'objectif de la conception de circuit de neurone analogique est de pouvoir en intégrer un très grand nombre sur puce, afin d'atteindre des tailles de réseau importantes, et d'augmenter la complexité des tâches qu'ils doivent exécuter. Dans ce contexte, l'ensemble des composants qui constituent le BB-LIF SN doit être intégré sur silicium pour atteindre une densité d'intégration maximale.

Le démonstrateur du BB-LIF SN présenté à la section 4.4 utilise une capacité externe (implémenté par la capacité parasite du câble de mesure) couplée à un BB-LTC implémenté sur silicium. Pour satisfaire la contrainte de densité d'intégration, la capacité doit donc être également portée sur silicium, constituant ainsi un BB-LIF SN intégré.

Par ailleurs, un circuit tampon, ou « buffer », dont l'entrée est connectée au nœud d'axone n_{AXO} , doit être ajouté au circuit, pour transmettre l'information de l'activité du neurone aux synapses qui lui sont connectées en aval du réseau. En effet, lors du déclenchement du BB-LIF SN, l'information de l'activation est portée par le pic de potentiel V_{AXO} généré sur le nœud d'axone n_{AXO} . Or, ce pic de potentiel est conduit par les courants de génération (I_{BTBT} et I_{II}) qui proviennent du BIMOS et qui sont relativement faibles. La connexion directe des circuits synaptiques sur le nœud n_{AXO} aurait pour effet d'ajouter une charge trop importante (en ajoutant de la capacité à C_{AXO} et, éventuellement, des

courants de fuite) et conduirait à une dérive des grandeurs caractéristiques du BB-LTC SN intégré (décalage de V_{TH} vers des valeurs supérieures par exemple).

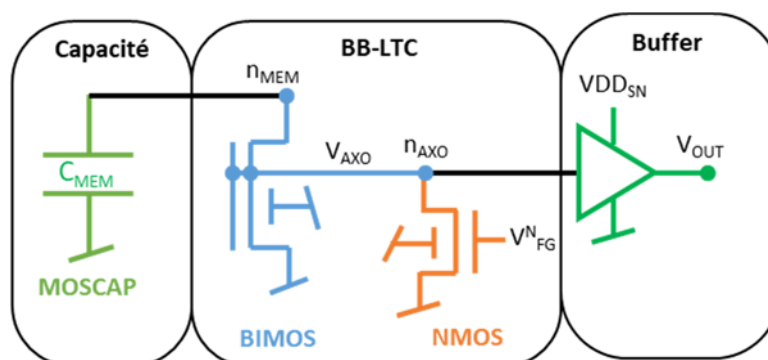


Figure 4-42 – Schéma du design du BB-LIF SN intégré.

4.5.1 Dimensionnement du BB-LTC intégré

Les tensions de seuils V_{TH} mesurées jusqu'à présent sont toutes supérieures à 1 V. Le BIMOS et le NMOS doivent donc avoir une résistance aux hautes tensions la plus grande possible. Pour satisfaire cette contrainte, ils sont donc pris de type EG (grande épaisseur d'oxyde de grille).

Le modèle des limites de fonctionnement du BB-LIF SN développé à la section 4.4.5 montre que la fréquence de fonctionnement minimale f_{MIN} est proportionnelle au courant $I_{OFF,TH}$ du BB-LTC. Pour obtenir une plage de fonctionnement la plus large possible, il est nécessaire de réduire ce courant au maximum. Pour ce faire, le BIMOS est pris de type RVT pour maximiser sa tension de seuil $V_{T,MOS}$. Par ailleurs, l'extraction de $I_{OFF,TH}$ en fonction de la largeur W_1 du BIMOS indique que ce courant est minimal et indépendant de W_1 pour $W_1 \leq 540$ nm (section 4.3.5.2 Figure 4-22). Dans le but de réduire au maximum la surface du circuit, W_1 est finalement pris au minimum autorisé par la technologie : $W_1 = 144$ nm. Pour la même raison, les dimensions L_1 , L_2 et W_2 sont également réduites au maximum. En ce qui concerne le NMOS, comme sa conductance peut être réglée par le potentiel V_{FG}^N appliqué, aucune contrainte particulière autre que la minimisation de la surface totale du BB LTC n'est adoptée. Ainsi, il est pris de type RVT, sa longueur de grille L est prise au minimum, et sa largeur W égale au W_2 du BIMOS. Notons ici qu'il aurait été pertinent de réduire W au minimum (144 nm) pour minimiser C_{AXO} . Le dimensionnement obtenu est récapitulé au Tableau 4-10.

Tableau 4-10 - Dimension du BB-LTC utilisé pour l'intégration du BB-LIF SN.

L_1 (nm)	W_1 (nm)	L_2 (nm)	W_2 (nm)	L (nm)	W (nm)	EOT (nm)	BIMOS	NMOS
135	144	144	350	135	350	3.4	RVT	RVT

Notons que la minimisation des dimensions des composants CMOS a généralement pour effet la maximisation de la variabilité de leurs caractéristiques. Cependant, les problématiques liées à ces variabilités ne sont pas considérées dans les travaux exposés ici.

4.5.2 Dimensionnement du buffer intégré

L'information de l'activation du neurone est une donnée binaire. L'architecture des buffers utilisés dans les circuits logiques est donc pertinente pour l'intégration dans le BB-LIF SN. Les buffers logiques sont constitués de deux circuits inverseurs en série, chacun composé d'un NMOSFET et d'un PMOSFET (Figure 4-43). Ces MOSFETs sont généralement dimensionnés pour que, d'une part, la bascule entre leurs états 1 et 0 se fasse lorsque la tension d'entrée vaut la moitié de la tension d'alimentation $V_{DD}/2$,

et, d'autre part, pour que les durées des transitions de l'état 1 vers l'état 0, et de l'état 0 vers 1 soient égales.

Tableau 4-11 – Paramètres de design des MOSFET qui constitue le buffer intégré au BB-LIF SN.

MOSFET	W (μm)	L (nm)	Saveur	V_{BG} (V)
N1	1.6	150	LVT	0
P1	0.16	150	RVT	0
N2	0.16	150	RVT	0
P2	1.6	150	LVT	0

Pour le dimensionnement du buffer intégré au BB-LIF SN, l'hypothèse d'une tension d'alimentation $V_{DD} = 1\text{V}$ est adoptée. Elle correspond à l'une des tensions nominales de la technologie 28 FD-SOI. Les simulations SPICE du BB-LIF SN effectuées à la section 4.4.4.2 montrent que, lors des déclenchements, le pic de potentiel d'axone V_{AXO} varie entre 0 V et quelques centaines de mV, et que, pour la majeure partie de sa durée, V_{AXO} se trouve dans la zone inférieure de cette plage. Ainsi, le choix est fait de dimensionner le premier inverseur de sorte que la tension de bascule se trouve plus proche de la masse que de V_{DD} , et le second pour que sa tension de bascule soit plus proche de V_{DD} que de la masse. Le potentiel de seuil du buffer se situe ainsi à 0.35 V. Comme les fréquences de déclenchement du BB-LIF SN sont bien plus faibles que les fréquences de coupures des inverseurs, l'impact sur la symétrie des durées transitions est considéré comme négligeable. Les grandeurs caractéristiques adoptées pour le design du buffer sont répertoriées au Tableau 4-11.

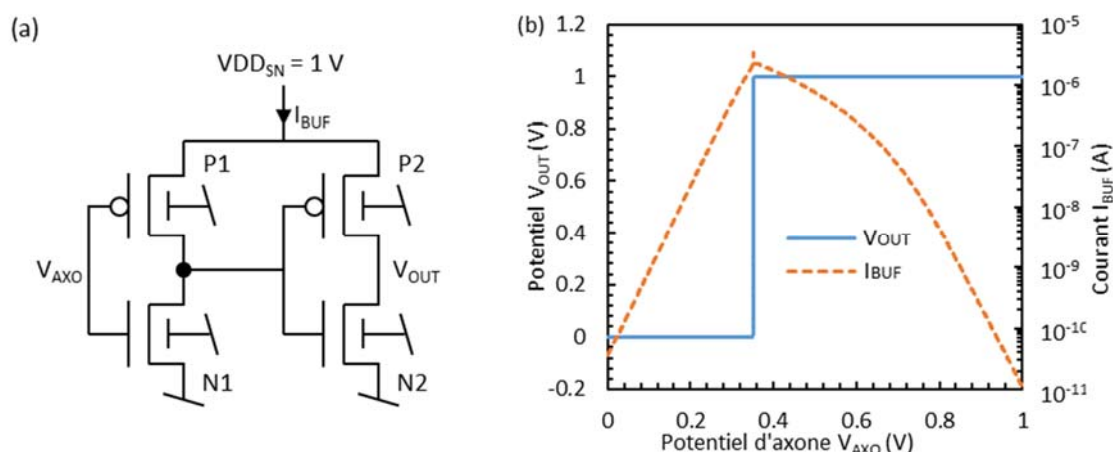


Figure 4-43 – (a) Schéma électrique du buffer intégré au BB-LIF SN. Potentiel de sortie V_{OUT} et courant circulant dans le buffer I_{BUF} en fonction du potentiel d'axone V_{AXO} , obtenue par simulation SPICE calibrée.

4.5.3 Dimensionnement de la capacité C_{MEM}

Il existe 3 types de capacité intégrée en technologie 28nm FD-SOI : la capacité métal-oxyde-métal (MOM), la capacité métal-isolant-métal (MIM) et la capacité MOS (MOSCAP) de type N ou P selon le sens de polarisation lors de leur utilisation. Les capacités MIM ont une capacité surfacique importante ($\sim 15\text{ fF}/\mu\text{m}^2$) mais nécessitent l'ajout de masques supplémentaires lors de leur fabrication qui induisent un surcoût. Les capacités MOM ont une densité plus faible ($\sim 1.2\text{ fF}/\mu\text{m}^2$) et rendent indisponibles les niveaux métalliques fins sur l'entièreté de leur surface. Ce dernier point est un désavantage significatif dans la perspective de la constitution d'un réseau à forte densité de connexion. Le choix est donc fait d'utiliser une capacité NMOSCAP pour l'intégration de la capacité C_{MEM} . Elle offre une capacité surfacique intermédiaire ($\sim 8\text{ fF}/\mu\text{m}^2$) et laisse libres les niveaux de métaux fins. Leur

faible facteur de qualité, leur comportement non linéaire et leur plus grande variabilité sont, dans un premier temps, considérés comme acceptables pour le circuit du BB-LIF SN.

Le dimensionnement de la capacité C_{MEM} est choisi de sorte que la surface totale du BB-LIF SN soit bien positionnée vis-à-vis de l'état de l'art. Ainsi, la contrainte adoptée est une surface totale de $15 \mu m^2$. Le respect des règles de dessin de la technologie 28nm FD-SOI conduit à une surface de capacité de $7.7 \mu m^2$. Les modèles SPICE mis à disposition par STMicroelectronics donnent la valeur de la capacité différentielle C_{DIFF} de leur MOSCAP en fonction de la tension appliquée V , définie comme :

$$C_{DIFF}(V) = \frac{dQ}{dV} \quad [F \cdot m^{-2}] \quad (4.28)$$

La capacité effective correspondante est alors donnée par :

$$C_{EFF}(V) = \frac{Q(V)}{V} = \frac{1}{V} \cdot \int_0^V C_{DIFF}(V) \cdot dV \quad [F \cdot m^{-2}] \quad (4.29)$$

Les deux fonctions $C_{MEM,DIFF}(V)$ et $C_{MEM,EFF}(V)$ obtenues par simulation sont tracées à la Figure 4-44. Comme attendu de la part d'une capacité MOS, la relation $C_{MEM,DIFF}(V)$ est non linéaire et, par extension, la relation $C_{MEM,EFF}(V)$ aussi. La valeur moyenne de C_{EFF} est de 59 fF.

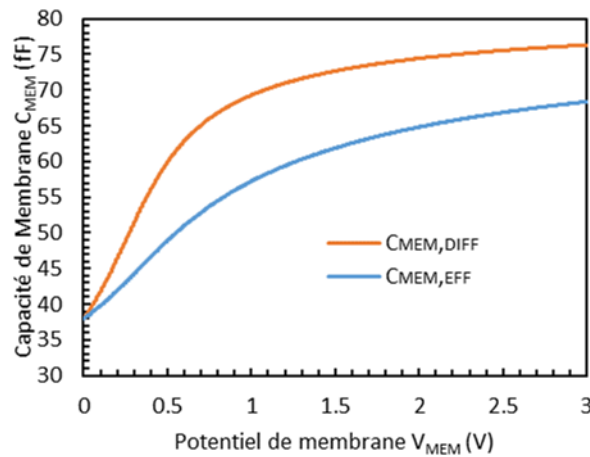


Figure 4-44 – Capacités différentielle $C_{MEM,DIFF}$ et effective $C_{MEM,EFF}$ de la NMOSCAP de $7.7 \mu m^2$ utilisée comme capacité de membrane C_{MEM} du BB-LIF SN intégré, en fonction du potentiel de membrane V_{MEM} , obtenues par simulation SPICE et pour une implémentation en technologie 28nm FD-SOI.

4.5.4 Layout du BB-LIF SN

Le layout du BB-LIF SN obtenu est présenté à la Figure 4-45. Il occupe une surface de $14.835 \mu m^2$.

4.5.5 Evaluation des limites de fonctionnement en fonction de V_{FG}^N

Dans cette section, les limites de fonctionnement du BB-LIF SN intégré $I_{SYN,MIN}$, $I_{SYN,MAX}$, f_{MIN} et f_{MAX} en fonction de V_{FG}^N sont évaluées à l'aide du modèle simple développé à la section 4.4.5. Pour ce faire, les relations $I_{OFF,TH}(V_{FG}^N)$, $V_{TH}(V_{FG}^N)$, $I_D(V_{FG}^N)$, $V_{AXO,INV}(V_{FG}^N)$, $C_{MEM}(V_{TH})$ et $C_{AXO}(V_{AXO,INV})$ doivent être établies.

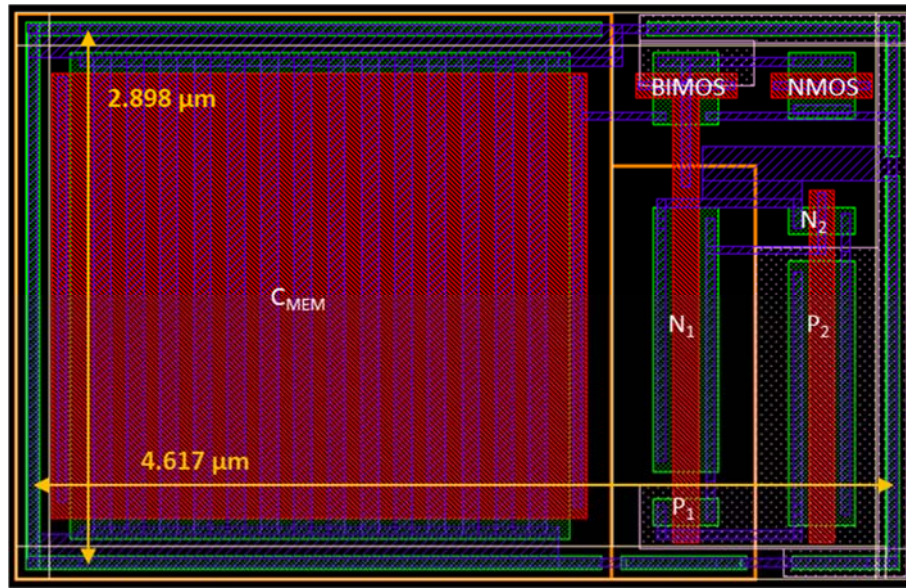


Figure 4-45 – Layout du BB-LIF SN intégrant le BB-LTC (BIMOS et NMOS) la capacité de membrane (C_{MEM}) et le buffer de sortie asymétrique (N_1 , P_1 , N_2 , et P_2), occupant une surface totale d'environ $15\mu m^2$.

Tableau 4-12 – Valeurs utilisées pour calculer les limites de fonctionnement du BB-LIF SN intégré selon le modèle développé à la section 4.4.5.

V_{FG}^N (V)	I_D (pA)	$V_{AXO,INV}$ (mV)	V_{TH} (V)	$I_{OFF,TH}$ (pA)	C_{MEM} (fF)	C_{AXO} (fF)
0	0.0491	261	1.74	7.30	63.4	1.75
0.1	1.44	504	1.95	34.0	64.5	2.12
0.2	413	644	2.15	158	65.5	2.38
0.3	1170	752	2.36	737	66.4	2.54

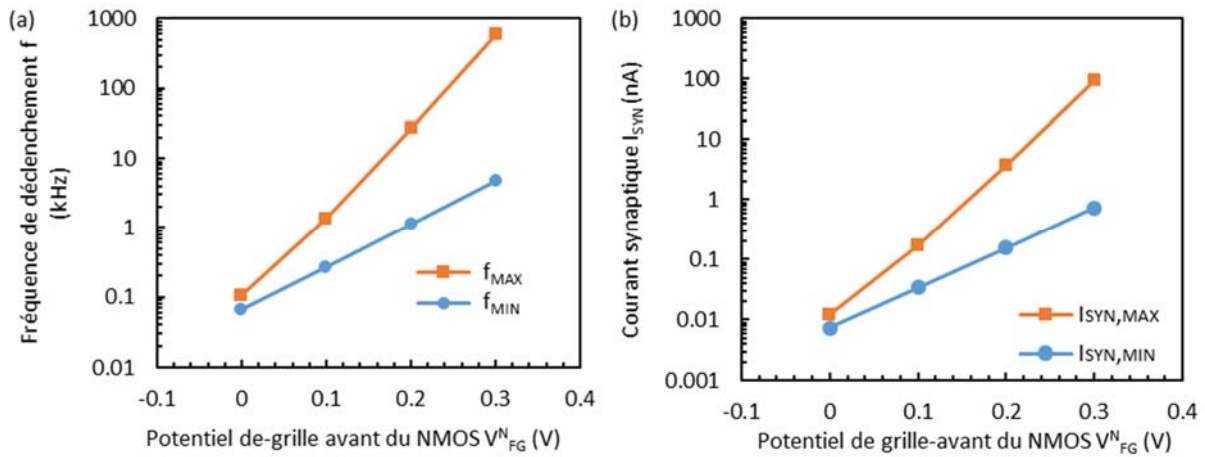


Figure 4-46 – Valeurs limites de la fréquence de déclenchement (a) et du courant synaptique (b) pour lesquels le BB-LIF SN est fonctionnel, en fonction du potentiel de grille-avant du NMOS V_{FG}^N , et d'après le modèle développé à la section 4.4.5.

$I_{OFF,TH}(V_{FG}^N)$ et $V_{TH} = f(V_{FG}^N)$ sont évalués en utilisant les équations (4.15) et (4.16), respectivement. Ces équations sont extraites de données concernant un BB-LTC dont la longueur $W_1 = 270$ nm est plus importante que celle utilisée ici $W_1 = 144$ nm. En s'appuyant sur les courbes de la Figure 4-22 (section 4.3.5.2), l'hypothèse est faite que le courant $I_{OFF,TH}$ et V_{TH} dépendent peu de W_1 lorsque W_1 est faible. $I_D(V_{FG}^N)$ et $V_{AXO,INV}(V_{FG}^N)$ sont obtenus par simulation SPICE. Pour rappel, $V_{AXO,INV}$ est la valeur de V_{AXO} pour laquelle $I_D = I_{PN}$. $C_{MEM}(V_{TH})$ est obtenu à partir de la courbe $C_{EFF}(V_{MEM})$ de la Figure 4-44 en prenant

$V_{MEM} = V_{TH} \cdot C_{AXO}(V_{AXO,INV})$ est obtenue à partir d'une courbe $C_{AXO,EFF}(V_{AXO})$ établie selon la même méthodologie que celle utilisée pour le calcul de $C_{MEM,EFF}$, et en prenant $V_{AXO} = V_{AXO,INV}$. Notons que, dans le cas du BB-LIF SN intégré, les capacités de grille des deux MOSFET d'entrée du buffer doivent être prises en compte pour le calcul de la valeur de C_{AXO} , qui passe ainsi de quelques centaines d'atto farads à quelques femto farads. L'ensemble des valeurs obtenues est récapitulé au Tableau 4-12.

Le calcul des valeurs de limite de fonctionnement (Figure 4-46) montre que toutes les grandeurs augmentent avec V_{FG}^N , et que la plage de fonctionnement s'élargit à mesure que V_{FG}^N augmente. La plage de fonctionnement en fréquence est quasiment nulle à $V_{FG}^N = 0$ V, et s'étend sur environ 2 décades à $V_{FG}^N = 0.3$ V.

4.5.6 Mesure de la réponse temporelle du BB-LIF SN intégré

Dans le but de mesurer la réponse temporelle du BB-LIF SN intégré, une source de courant et un circuit tampon amplificateur sont fabriqués sur silicium aux côtés du BB-LIF SN et sont connectés à celui-ci (Figure 4-47). La source de courant (« current source » : CS) est constituée d'un NMOS et d'un miroir de courant. Le courant I_{SYN} est contrôlé par le potentiel V_{IN} . La source du NMOS est connectée à la masse, et les sources des 2 PMOS du miroir de courant sont portées au potentiel VDD_{CS} . Le circuit tampon amplificateur est composé de 4 inverseurs en série, dont les tailles augmentent à chaque étage, de sorte que le dernier puisse conduire l'impédance de charge amenée par le câble de mesure et l'oscilloscope. Les sources des NMOS du circuit tampon amplificateur sont connectées à la masse, les sources des PMOS sont portées au potentiel VDD_{SN} . Dans le BB-LIF SN, la grille-avant du NMOS du BB-LTC est portée au potentiel V_{FG}^N , et les sources des PMOS de buffer sont portées au potentiel VDD_{SN} . Le nœud n_{MEM} est connecté à la sortie de la CS et le nœud de sortie du buffer au nœud d'entrée du circuit tampon amplificateur. Tous les autres terminaux sont connectés à la masse.

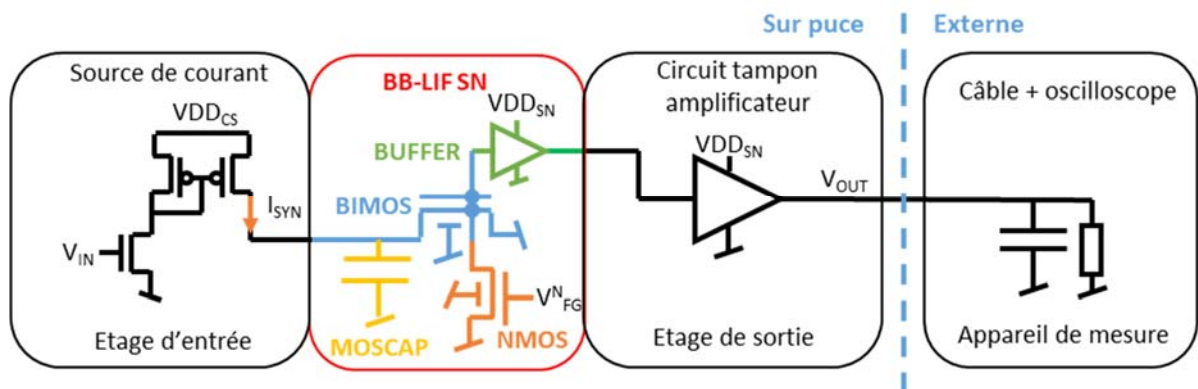


Figure 4-47 – Schéma du montage utilisé pour la mesure de la réponse temporelle de BB-LIF SN intégré.

4.5.6.1 Dispositif de mesure

Les tensions VDD_{CS} et V_{FG}^N sont appliquées par deux unités de source et mesure (« source and measure unit » : SMU) Keysight B1510A pilotées par un analyseur de paramètres de composants semi-conducteurs Keysight B1500. Les tensions VDD_{SN} et V_{IN} sont appliquées en utilisant 2 sources de tensions externes. La tension de sortie du buffer amplificateur V_{OUT} est sondée par un oscilloscope. Le courant I_{CS} qui traverse la source de courant est mesuré à l'aide du SMU. Pour évaluer le courant I_{SYN} , il faut déterminer quelle part du courant I_{CS} circule vers le BB-LIF SN. Elle est théoriquement de 50 %, mais les variations de V_{MEM} pendant le fonctionnement du BB-LIF SN conduisent à des appels de

courant qui modifie ce ratio. Ainsi des simulations SPICE montrent que, en fonctionnement, le rapport vaut plutôt :

$$\frac{I_{\text{SYN}}}{I_{\text{CS}}} \approx 0.63 \quad (4.30)$$

4.5.6.2 Configuration des polarisations

Pour obtenir une large plage de fonctionnement pendant les mesures, en se référant à la Figure 4-46, V_{FG}^{N} est prise égale à 0.3 V. Comme une telle valeur de V_{FG}^{N} augmente le potentiel de seuil V_{TH} , le potentiel V_{DDCS} est monté à 3 V. Notons que cette tension est au-delà de la tension nominale $V_{\text{DEG}} = 1.8$ V des transistors EG. La fiabilité à long terme des composants n'est donc plus garantie à cette tension. Afin de limiter la consommation du BB-LIF SN, le potentiel V_{DDSN} est fixé, dans un premier temps, à 0.75 V. Le potentiel V_{IN} est monté progressivement depuis la masse ($V_{\text{IN}} = 0$ V), par pas de 10 mV.

4.5.6.3 Fréquence de déclenchement f en fonction du potentiel V_{IN}

Tant que $V_{\text{IN}} < 0.37$ V, V_{OUT} demeure constant à 0 V. Pour $V_{\text{IN}} \in [0.37 ; 0.50]$ V, un signal périodique en forme de crête est observé sur V_{OUT} (Figure 4-49). La fréquence de variation de V_{OUT} augmente exponentiellement avec V_{IN} (Figure 4-48). Pour $V_{\text{IN}} > 0.5$ V, le signal crête disparaît, et V_{OUT} reste alors constant à 0.75 V. Le signal réapparaît lorsque que V_{DDSN} est augmenté, à 1.5 V par exemple. Le même phénomène survient pour $V_{\text{IN}} > 0.53$ V, où V_{DDSN} doit être augmenté à nouveau à 1.8 V pour que le signal réapparaisse. Au-delà de $V_{\text{IN}} = 0.55$ V, plus aucun signal n'est observé sur V_{OUT} , quel que soit le potentiel V_{DDSN} .

La sortie du buffer intégré au BB-LIF SN bascule entre ses états haut et bas lorsque le potentiel V_{AXO} franchit son potentiel de seuil $V_{\text{T,BUF}}$. Si V_{OUT} est constant et égal à 0 V, cela signifie qu'à tout instant $V_{\text{AXO}} < V_{\text{T,BUF}}$. De la même façon, si V_{OUT} est constant et égal à V_{DDSN} , cela signifie qu'à tout instant $V_{\text{AXO}} > V_{\text{T,BUF}}$. Ainsi, la disparition du signal sur V_{OUT} et sa réapparition lorsque V_{DDSN} est augmenté témoignent de la réduction de l'amplitude des variations de V_{AXO} à mesure que la fréquence de déclenchement augmente. Ce phénomène a déjà été mis en évidence par les résultats de simulation présentés à la Figure 4-40.b de la section 4.4.4.

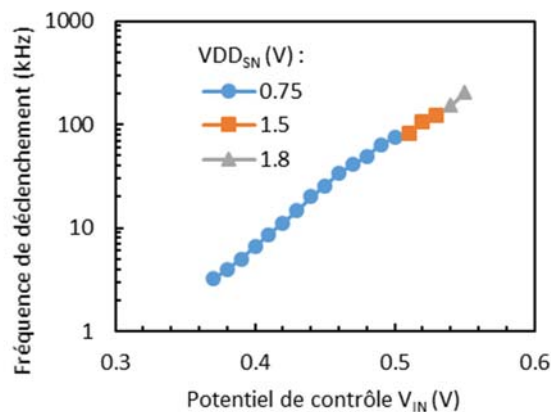


Figure 4-48 – Fréquence de déclenchement du BB-LIF SN intégré en fonction du potentiel de contrôle de la source de courant V_{IN} , et pour différentes valeurs de V_{DDSN} . Les données sont issues de mesures à température ambiante.

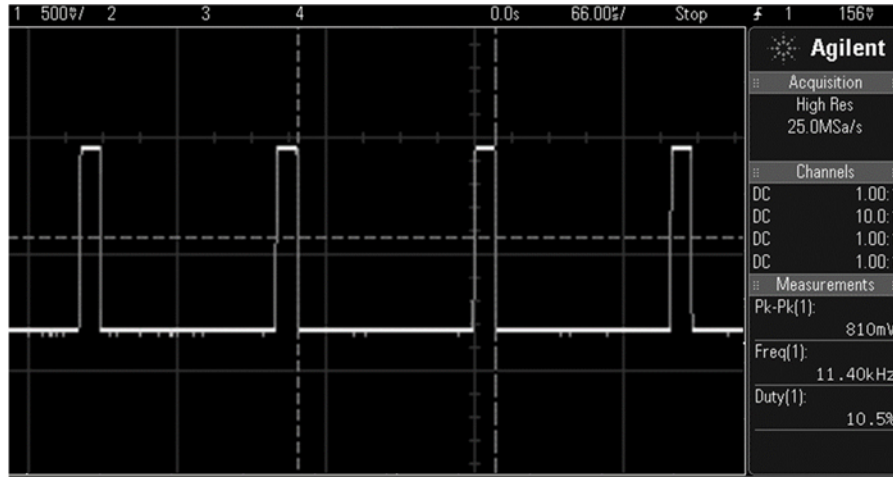


Figure 4-49 – Capture d'écran de l'oscilloscope lors d'une mesure du potentiel de sortie du buffer amplificateur sous les polarisations suivantes : $V_{FG}^N = 0.3$ V, $V_{IN} = 0.42$ V, $V_{DD_{CS}} = 3$ V, $V_{DD_{SN}} = 0.81$ V.

4.5.6.4 Fréquence de déclenchement f en fonction du courant I_{SYN}

Le courant I_{SYN} est extrait de la mesure de $I_{CS} = f(V_{IN})$ selon l'équation (4.30). La fréquence de déclenchement f est tracée en fonction du courant I_{SYN} à la Figure 4-50. Comme, en fonctionnement, les tensions d'alimentations doivent être constantes, seuls les points obtenus pour $V_{DD_{SN}} = 0.75$ V sont disposés sur le graphe. La plage de fonctionnement est ainsi réduite à [3.2 ; 74.6] kHz. Sur cette figure, la fenêtre de fonctionnement calculée à la section 4.5.5 est également tracée ainsi que la fréquence de déclenchement f obtenue par simulations, après qu'elles aient été calibrées en V_{FG}^N .

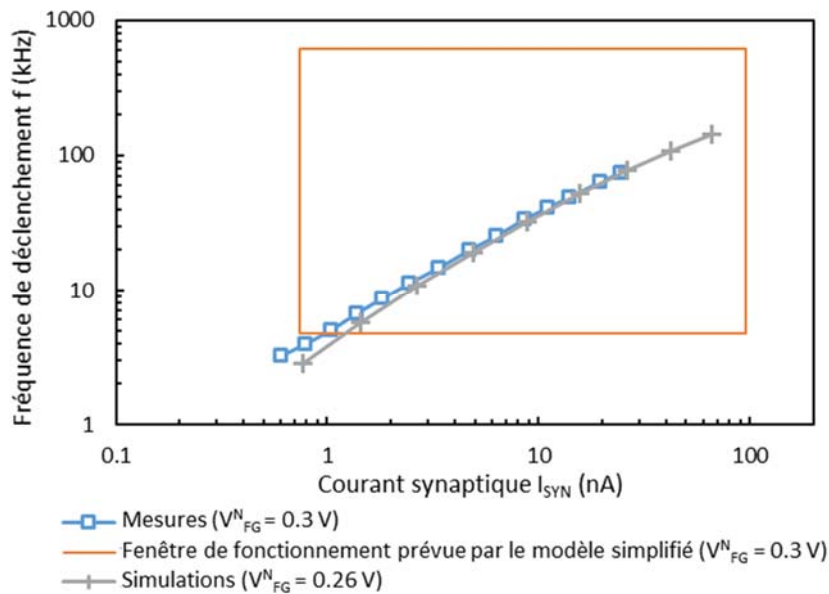


Figure 4-50 – Fréquence de déclenchement f en fonction du courant synaptique I_{SYN} du BB-LIF SN intégré, obtenue par la mesure à $V_{FG}^N = 0.3$ V et par simulation à $V_{FG}^N = 0.26$ V, et fenêtre de fonctionnement prédit par le modèle simplifié.

La calibration des simulations en V_{FG}^N est nécessaire car, dans la même configuration de polarisation que celle utilisée pendant les mesures, les simulations donnent un BB-LIF SN non-fonctionnel. En effet, en simulation et à $V_{FG}^N = 0.3$ V, la tension de seuil V_{TH} est supérieure à $V_{DD_{CS}} = 3$ V, qui est le potentiel maximal que peut atteindre V_{MEM} en mesure. Ce désaccord est une conséquence de l'imprécision du

modèle SPICE lors du calcul du courant de body-contact du BIMOS, déjà mise en évidence lors des caractérisations électriques du chapitre 3 et lors des comparaisons entre les mesures et les simulations du comportement du BB-LTC et du BB-LIF SN aux sections 4.3.6 et 4.4.4. Afin d'obtenir des simulations comparables à la mesure, le potentiel V_{FG}^N des simulations est ajusté pour que les fréquences de déclenchement à $I_{SYN} = 10$ nA soient égales entre les mesures et les simulations. Cette égalité est obtenue pour $V_{FG}^N = 0.26$ V.

Les courbes de la Figure 4-50 montrent un bon accord entre les simulations SPICE et la mesure. Au-delà de la quasi-superposition des courbes $f(I_{SYN})$ qui est le fruit de l'ajustement du potentiel V_{FG}^N utilisé pour calibrer les simulations, leurs bornages $[f_{MIN}; f_{MAX}]$ sont comparables ($[3.2; 74.6]$ kHz et $[2.8; 143]$ kHz) avec tout de même une surestimation de f_{MAX} en simulation. Le fenètre de fonctionnement issue du modèle simplifié donne également une bonne approximation de ce qui est mesuré.

4.5.6.5 Consommation énergétique du BB-LIF SN

La consommation du BB-LIF SN est de deux types : une consommation statique E_{STA} qui correspond à la consommation du circuit lorsqu'aucun courant synaptique ne lui est fourni, et une consommation dynamique E_{DYN} qui est à la consommation du circuit lorsqu'il se déclenche.

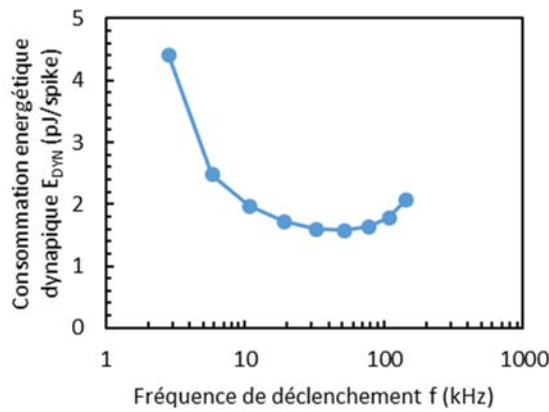


Figure 4-51 – Consommation énergétique dynamique E_{DYN} du BB-LIF SN intégré en fonction de la fréquence de déclenchement f . Donnée obtenue par simulation SPICE avec $VDD_{SN} = 0.75$ V, $VDD_{CS} = 3$ V et $V_{FG}^N = 0.26$ V.

Lorsqu'aucun courant synaptique ne lui est explicitement fourni, un courant de fuite $I_{SYN,LK}$ circule toutefois dans le BB-LIF SN depuis la source de courant. Dans le cas présent, le dimensionnement de cette source conduit à une valeur de $I_{SYN,LK}$ de 74 pA. Ce courant provient de la source de tension $VDD_{CS} = 3$ V. Le buffer intégré quant à lui consomme un courant I_{BUF} de 42 pA qui provient de la source de tension $VDD_{SN} = 0.75$ V (ces valeurs de courant sont obtenues par simulation). La consommation statique E_{STA} est donnée par :

$$E_{STA} = VDD_{CS} \cdot I_{SYN,LK} + VDD_{SN} \cdot I_{BUF} \quad (4.31)$$

et vaut 254 pW.

La consommation dynamique, exprimée en joule par déclenchement (J/spike) est calculée par :

$$E_{DYN} = \int_T (VDD_{CS} \cdot I_{SYN} + VDD_{SN} \cdot I_{BUF}) dt \quad (4.32)$$

où T est la période déclenchement. E_{DYN} est calculé en fonction de la fréquence de déclenchement (Figure 4-51). Calculée à l'aide de simulation SPICE, elle est de l'ordre de quelques pJ/spike, avec une valeur moyenne $E_{DYN} = 2$ pJ/spike.

4.5.7 Bilan

Les performances du BB-LIF SN intégré sont récapitulées au Tableau 4-13.

Tableau 4-13 – Performances du BB-LIF SN intégré.

Surface	15 μm^2
Capacité de membrane C_{MEM}	59 fF
Surface des Capacités	7,7 μm^2
Technologie	28nm UTBB FD-SOI
Tensions d'alimentation	3 V et 0.75 V
Fréquence de fonctionnement	[3.2 ; 74.6] kHz
Consommation statique	254 pW
Consommation dynamique	20 nW @ 10 kHz
Efficacité énergétique	2 pJ/spike @ 10 kHz
Modèle de neurone	LIF
Nombre de composant	7 : 6T1C
Maturité	Circuit

4.6 Conclusion

Le modèle LIF est le modèle de SN le plus simple. Avec seulement 4 paramètres, il décrit la dynamique du potentiel de membrane V_{MEM} en fonction du courant synaptique I_{SYN} selon une équation différentielle du premier ordre et une équation conditionnelle. Ce modèle peut être implémenté par un circuit électronique constitué d'une capacité et d'un LTC, lui-même composé d'une résistance couplée à une bascule de Schmitt et un interrupteur.

La principale contribution des travaux présentés ici consiste à montrer qu'il est possible, en s'inspirant des circuits de protection ESD, d'implémenter le LTC avec un circuit appelé BB-LTC, constitué uniquement de deux composants couplés : un FD-SOI BIMOS et un FD-SOI NMOS. L'ajout d'une capacité à ce circuit permet alors de construire un circuit de SN compact nommé BB-LIF SN.

Dans le BB-LTC, le couplage adéquat s'obtient par la connexion de la grille-avant et de la base du BIMOS au drain du NMOS. Dans le contexte des circuits de neurones, cette connexion constitue le nœud électrique d'axone n_{AXO} . Le collecteur du BIMOS est le nœud d'entrée du BB-LTC, destiné à être connecté à l'anode de la capacité de membrane C_{MEM} , constituant ainsi le nœud électrique de membrane n_{MEM} . L'émetteur du BIMOS et la source du NMOS sont connectés à la masse, tandis que le potentiel de grille avant du NMOS et le potentiel de grille arrière du BIMOS sont des potentiels de contrôle du circuit.

Selon la tension qui est appliquée sur le n_{MEM} , la résistance d'entrée du BB-LTC bascule entre deux valeurs différentes de plusieurs ordres de grandeurs. De plus, la caractéristique courant-tension du nœud d'entrée présente une hystérésis. Le déclenchement du BB-LTC – i.e. la bascule de l'état OFF (haute résistance) à l'état ON (résistance faible) – et sa réinitialisation – i.e. la bascule de l'état ON à l'état OFF – s'effectuent à deux valeurs de tension d'entrée V_{MEM} différentes, notées V_{TH} et V_{RST} respectivement et qui vérifient $V_{TH} > V_{RST}$. À ces deux valeurs de V_{MEM} particulières, le courant d'entrée bascule de $I_{OFF,TH}$ à $I_{ON,TH}$ et de $I_{ON,RST}$ à $I_{OFF,RST}$ respectivement. Avec V_{TH} et V_{RST} , ces quatre valeurs de courant constituent les grandeurs caractéristiques du BB-LTC.

Les analyses systématiques des multiples courbes courant-tensions du FD-SOI BIMOS effectuées au chapitre 3 permettent de décrire les mécanismes mis en jeu dans le déclenchement et la réinitialisation du BB-LTC. La génération de paires électron-trou dans le film de silicium à l'interface collecteur – body du BIMOS, sous l'effet du BTBT et de l'II, ainsi que la diffusion des trous générés vers le terminal de base y jouent un rôle majeur. La théorie de la stabilité des systèmes dynamiques indiquent que les transitions ON-OFF et OFF-ON interviennent lorsque $\partial(I_B - I_D)/\partial V_{AXO} > 0$. Si l'on considère qu'au moment des transitions, le NMOS est en saturation ($\partial I_D/\partial V_{AXO} = 0$), cette inégalité devient $\partial I_{II}/\partial V_{AXO} > -\partial I_{BTBT}/\partial V_{AXO}$ dans le cas de la transition OFF-ON et $\partial I_{II}/\partial V_{AXO} > \partial I_{PN}/\partial V_{AXO}$ dans le cas de la transition ON-OFF.

Des mesures expérimentales quasi-statiques effectuées à température ambiante sur des échantillons fabriqués en technologie 28 nm FD-SOI montrent comment les potentiels de contrôle V_{FG}^N , et V_{BG}^B , la largeur W_1 du BIMOS et l'EOT des composants du circuit influent sur les valeurs des six grandeurs caractéristiques. Par ailleurs, des mesures transitoires révèlent les dépendances de ces grandeurs vis-à-vis des temps de montée et de descente du potentiel V_{MEM} . En particulier, V_{TH} augmente avec dV_{MEM}/dt . Ces dépendances s'expliquent par les dynamiques de charge et de décharge propres au nœud d'axone. La cohérence des comparaisons entre les mesures et les simulations SPICE du BB-LTC, effectuées en s'appuyant sur le modèle de FD-SOI BIMOS développé au chapitre 3, atteste la validité du modèle descriptif établi précédemment.

La preuve de concept du fonctionnement du BB-LIF SN est apportée par la mesure au cours du temps et à température ambiante du potentiel de collecteur du BB-LTC – i.e. le potentiel de membrane V_{MEM} – sous l'application d'un courant d'entrée continu – i.e. le courant synaptique I_{SYN} . Dans une telle configuration, la capacité parasite du câble de mesure implémente la capacité de membrane C_{MEM} du BB-LIF SN. Le potentiel V_{MEM} croît linéairement au cours du temps et chute périodiquement et soudainement à sa valeur de repos $V_{MEM} = 0$ V lorsqu'il atteint la valeur seuil V_{TH} . Cette chute correspond au passage furtif du BB-LTC à l'état ON. Les mesures révèlent par ailleurs la présence d'une courte période réfractaire de durée τ_{RFR} pendant laquelle V_{MEM} demeure au potentiel de repos bien qu'un courant I_{SYN} non nul soit toujours présent. Comme attendu de la part d'un neurone LIF, le BB-LIF SN ne se déclenche pas lorsque $I_{SYN} < I_{OFF,TH}$. Pour $I_{SYN} > I_{ON,RST}$, V_{MEM} se stabilise après un premier déclenchement autour d'une valeur comprise entre V_{TH} et V_{RST} . Le BB-LIF SN cesse alors de fonctionner. La proportionnalité entre la fréquence de déclenchement et l'intensité du courant I_{SYN} fait du BB-LIF SN un neurone à l'excitabilité de type 1.

Les influences des potentiels V_{FG}^N et V_{BG}^B sur le comportement du BB-LIF SN sont par ailleurs mises en évidence. Elles permettent d'envisager un contrôle électrostatique du circuit qui peut être exploité dans le cadre de l'implémentation d'une règle d'apprentissage à l'échelle du réseau de neurones.

Tout comme pour le BB-LTC, les confrontations des mesures aux simulations SPICE attestent de la validité du modèle descriptif élaboré dans ce chapitre. De plus, comme attendu, les simulations révèlent la présence de pics dans le chronogramme du potentiel V_{AXO} au moment des déclenchements. Ainsi, le BB-LIF SN produit intrinsèquement l'impulsion qui porte l'information de l'activation du neurone.

Une méthode d'évaluation des limites de fonctionnement du BB-LIF SN en fonction du dimensionnement du circuit est proposé en s'appuyant sur une modélisation simple des dynamiques de chargement et déchargement des capacité C_{MEM} et C_{AXO} .

Enfin, une conception du BB-LIF SN complètement intégré en technologie 28 nm FD-SOI est proposée. Entièrement constituée par des composants EG à oxyde épais, elle comporte, en plus du BB-LTC, une capacité MOSCAP et un buffer digital adapté. L'ensemble du circuit dispose d'une surface d'empreinte

sur silicium de moins de $15 \mu\text{m}^2$, qui le place de ce point de vue au premier rang des circuits à l'état de l'art. Le bon fonctionnement du circuit est attesté par la mesure à température ambiante, avec une fréquence de déclenchement comprise entre 3.2 et 74.6 kHz pour un courant synaptique compris entre 605 pA et 24.6 nA. Cependant, pour que le circuit fonctionne, la tension d'alimentation de la source de courant cointégré au circuit et qui fournit le courant I_{SYN} doit être portée à 3 V. Cette valeur supérieure à la tension nominative des composants EG (1.8 V) ne permet plus de garantir la fiabilité des composants. La nécessité de cette surtension est par ailleurs correctement prédite par le modèle d'évaluation des limites de fonctionnement proposé précédemment.

Les mesures sont reproduites en simulation SPICE, avec cependant un désaccord quant au potentiel V_{FG}^{N} à appliquer pour obtenir des résultats concordants. Ce désaccord provient de celui observé au chapitre 3 entre la mesure et la simulation du courant de body-contact du BIMOS. Il est lié à l'incomplétude du modèle d'ionisation par impact implémenté dans le modèle UTSOI.

L'efficacité énergétique du BB-LIF SN est évaluée à 2 pJ/spike. La contribution provenant du sous-bloc constitué du BB-LTC et de la MOSCAP est obtenue par la mesure, tandis que celle provenant du buffer alimenté sous 0.75 V par simulation. Cette efficacité positionne le BB-LIF SN entre les meilleures contributions provenant des équipes de Zürich et de Lille évoquées au Chapitre 1. Une revue complète des performances et caractéristiques du BB-LIF SN aux côtés de celles des dernières contributions à l'état de l'art est présentée au Tableau 4-14.

Tableau 4-14 – Derniers travaux à l'état de l'art des circuits analogiques de SN auquel est ajouté le circuit développé dans cette thèse.

Auteurs	Ce travail	Danneville et al., 2019 [86]	Sourikopoulos et al., 2017 [111]	Qiao et Indiveri, 2016, 2017 [130], [131]
Modèle de neurone	LIF	Axon-Hillock	Morris-Lecard based	Adaptive exponential I&F
Technologie	28 nm FD-SOI	65 nm	65 nm	28 nm FD-SOI
Tension d'alimentation	3 V et 0.75 V	0.2 V	0.2 V	1 V
Nombre de composant	7 : 6T1C	6 : 5T1C	6 : 6T	52T3C
Surface du circuit*	15 μm^2	31 μm^2	35 μm^2	50 μm^2
Valeur des capacités	$C_M = 59 \text{ fF}$	$C_F = 4 \text{ fF}$	$C_M = 4 \text{ fF}$, $C_K = 8 \text{ fF}$	$C_R = 0.2 \text{ pF}$ $C_M = 0.5 \text{ pF}$ $C_A = 0.2 \text{ pF}$
Surface des Capacités	7,7 μm^2	-	-	50 μm^2
Fréquence de fonctionnement	[3.2 ; 74.6] kHz	[0.29 ; 15.6] kHz	[- ; 25] kHz	[0 ; 12] kHz
Consommation statique	254 pW	11 pW	25 pW*	-
Consommation dynamique	20 nW à 10 kHz	30 pW à 15,7 kHz	100 pW à 25 kHz	1.5 nW à 30 Hz
Efficacité énergétique*	2 pJ/spike à 10 kHz	2 fJ/spike	4 fJ/spike à 25 kHz	50 pJ/spike à 30 Hz
Maturité	Circuit (silicium)	Circuit (silicium)	Circuit (silicium)	Système

*Figures de mérite principales

4.7 Références

- [1] J. Chevallier, P. Reynaud-Bouret, and M. Monticelli, “Nos neurones se synchronisent-ils ?,” *Images des Mathématiques, CNRS*, 2017. .
- [2] L. Lapique, “Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarization.,” *J. Physiol. Pathology*, vol. 9, pp. 620–635, 1907.
- [3] L. F. Abbott, “Lapicque’s introduction of the integrate-and-fire model neuron (1907),” *Brain Res. Bull.*, vol. 50, no. 5, pp. 303–304, 1999.
- [4] J. K. Roberge, *Operational amplifiers: theory and practice*, vol. 197. Wiley New York, 1975.
- [5] K. Nagaraj and M. Satyam, “Novel CMOS Schmitt trigger,” *Electron. Lett.*, vol. 17, no. 19, pp. 693–694, Sep. 1981.
- [6] B. L. Dokic, “CMOS schmitt triggers,” *IEE Proc. G - Electron. Circuits Syst.*, vol. 131, no. 5, pp. 197–202, Oct. 1984.
- [7] B. L. DOKIĆ and Z. V. BUNDALO, “Regenerative logic circuits with CMOS transistors,” *Int. J. Electron.*, vol. 58, no. 6, pp. 907–920, 1985.
- [8] B. L. Dokić, “CMOS NAND and NOR Schmitt circuits,” *Microelectron. J.*, vol. 27, no. 8, pp. 757–765, 1996.
- [9] B. L. Dokic, “CMOS and BiCMOS Regenerative Logic Circuits,” in *Cutting Edge Research in New Technologies*, InTech, 2012.
- [10] A. Pfister, “Novel CMOS Schmitt trigger with controllable hysteresis,” *Electron. Lett.*, vol. 28, no. 7, pp. 639–641, Mar. 1992.
- [11] I. M. Filanovsky and H. Baltes, “CMOS Schmitt trigger design,” *IEEE Trans. Circuits Syst. Fundam. Theory Appl.*, vol. 41, no. 1, pp. 46–49, Jan. 1994.
- [12] S. F. Al-Sarawi, “Low power Schmitt trigger circuit,” *Electron. Lett.*, vol. 38, no. 18, pp. 1009–1010, Aug. 2002.
- [13] V. Katyal, R. L. Geiger, and D. J. Chen, “Adjustable hysteresis CMOS Schmitt triggers,” in *2008 IEEE International Symposium on Circuits and Systems*, 2008, pp. 1938–1941.
- [14] A. Marzaki, V. Bidal, R. Laffont, W. Rahajandraibe, J. Portal, and R. Bouchakour, “A new adustable Schmitt Trigger based on Dual Control Gate-Floating Gate Transistor (DCG-FGT),” in *2012 IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2012, pp. 643–645.
- [15] H.-S. Wong, M. H. White, T. J. Krutsick, and R. V. Booth, “Modeling of transconductance degradation and extraction of threshold voltage in thin oxide MOSFET’s,” *Solid-State Electron.*, vol. 30, no. 9, pp. 953–968, 1987.
- [16] E. M. Izhikevich, “Which model to use for cortical spiking neurons?,” *IEEE Trans. Neural Netw.*, vol. 15, no. 5, pp. 1063–1070, Sep. 2004.

Conclusion

Le neuromorphisme est un des nouveaux paradigmes « More-than-Moore » sur lequel les acteurs de la microélectronique peuvent s'appuyer pour poursuivre les progrès technologiques alors que la loi de Moore atteint aujourd'hui ses limites. Les processeurs neuronaux actuellement en développement en sont le support privilégié de par l'efficacité énergétique qu'ils offrent en mimant jusqu'au mode de communication impulsionnel des systèmes nerveux biologiques. Ils reposent sur l'intégration massive de circuits élémentaires de neurone et de synapses interconnectés. La performance calculatoire de ces processeurs est fondamentalement liée à la taille des réseaux qu'ils peuvent émuler. Dans ce contexte, la réduction de la surface de l'empreinte des circuits élémentaires sur le silicium est un élément clef pour permettre à ces processeurs de concurrencer les réseaux de neurones artificiels formels exécutés sur des machines standards et très majoritairement utilisés aujourd'hui dans les applications d'intelligence artificielle.

Pour proposer un circuit analogique de neurone impulsionnel de taille réduite, les travaux de recherche présentés dans cette thèse se sont appuyés sur l'exploitation du comportement électrique du BIMOS, transistor à l'architecture particulière utilisé notamment pour la constitution de circuit de protection ESD et intégré à l'occasion de ces travaux en technologie 28 nm FD-SOI.

Un premier pan du travail, présenté au chapitre 3, a consisté à caractériser électriquement le FD-SOI BIMOS lorsque celui-ci est dimensionné aux tailles minimales autorisées par cette technologie. Cela a permis d'identifier les contributions des différents phénomènes de transport électrique aux courants mesurés. Grâce à ces analyses, une modélisation compacte du comportement électrique du FD-SOI BIMOS à température ambiante a été proposée, adaptée à partir du modèle UTSOI développé par le CEA-LETI et des bibliothèques de composants 28 nm FD-SOI fournies par STMicroelectronics. De par sa topologie et selon le jeu de ses polarisations, le FD-SOI BIMOS peut fonctionner à la manière d'une diode munie d'une grille (« gated-diode »), d'un transistor à effet de champs muni d'un contact de corps (« body-contacted MOSFET ») ou d'un transistor bipolaire muni d'une grille (« gated-BJT »).

- Les caractérisations en mode « gated-diode » ont mis en évidence l'influence de la polarisation de la grille avant sur l'intensité des mécanismes de génération-recombinaison responsables de la non-idéalité de la diode (génération par effet tunnel bande-à-bande en polarisation inverse et recombinaisons SRH en polarisation directe). La modélisation compacte est restreinte au calcul du courant direct lorsque l'anode et la grille avant sont connectées ensemble, configuration rencontrée dans le circuit de neurone proposé.
- Les caractérisations en mode « body-contacted MOSFET » ont montré que le terminal de body-contact collecte les trous générés dans la structure par effet tunnel bande-à-bande ou par ionisation par impact. Un aspect remarquable est que le courant qui en résulte et qui sort du composant est indépendant du potentiel du terminal de body-contact tant que celui-ci n'est pas trop élevé.
- Les caractérisations en mode « gated-BJT » ont démontré la faiblesse du gain bipolaire ($\beta_0 < 1$) du FD-SOI BIMOS en mode émetteur commun, fabriqué en technologie CMOS. Cette valeur est cohérente avec la négligence assumée du phénomène d'amplification bipolaire parasite dans les modèles typiques de transistors MOSFET sur FD-SOI. Dans la perspective de la construction du modèle compact, un modèle empirique du gain β_0 est proposé par régression sur les données mesurées. Un mode de fonctionnement qualifié d'anormal est également mis en évidence, où le courant du terminal de base est négatif et totalement indépendant de son potentiel. Ce courant est issu du mécanisme de génération par effet tunnel bande-à-bande déjà observé dans les caractérisations de type MOSFET.

Pour rendre compte de l'ensemble de ces comportements en simulation compacte, le modèle UTSOI est modifié selon les quatre points suivants :

1. Un terminal électrique de body contact est ajouté. Le modèle de diode précédemment établi est implémenté entre les terminaux de body-contact et de source et entre les terminaux de body-contact et de drain.
2. Les courants de GIDL et GISL déjà calculés par le modèle UTSOI sont dérivés vers le terminal de body.
3. Un modèle de courant de génération par ionisation par impact est ajouté entre les terminaux de drain et de body-contact.
4. Enfin, un courant d'amplification bipolaire circulant du drain vers la source est ajouté, égal au produit du courant de gated-diode body-contact – source par le gain β_0 modélisé empiriquement.

Bien que la forme générale de la courbe du courant de body-contact obtenue par la mesure soit respectée par les simulations compactes, des désaccords quantitatifs sont observés entre elles malgré les efforts de calibration entrepris. Cela suggère de compléter lors de travaux futurs le modèle de courant d'ionisation par impact utilisé pour décrire ce mécanisme dans le FD-SOI BIMOS.

La seconde partie du travail a consisté à développer le circuit analogique de neurone à impulsion BB-LIF SN proprement dit. Les détails de ce développement, présentés au chapitre 4, montrent comment il est possible d'implémenter le modèle de neurone LIF avec seulement trois composants : une capacité et un BB-LTC comprenant un FD-SOI N-MOSFET et un FD-SOI BIMOS. Le circuit final comporte quatre composants de plus afin d'intégrer un circuit tampon capable de transmettre l'impulsion électrique présente sur le nœud d'axone au moment des déclenchements tout en découplant le BB-LTC du terminal de sortie du neurone.

Les mécanismes impliqués dans les différentes régions de la caractéristique courant-tension d'entrée du BB-LTC sont entièrement décrits et expliqués grâce aux analyses du comportement électrique du BIMOS exposées au chapitre 3. La génération de paires électron-trou par BTBT et II ainsi que la connexion au nœud d'axone des terminaux de base du BIMOS, de grille-avant du BIMOS et de drain du NMOS jouent un rôle décisif dans les deux variations abruptes de la conductance d'entrée de BB-LTC lorsque la tension V_{MEM} atteint V_{TH} et V_{RST} . La théorie de la stabilité des systèmes dynamiques appliquée au potentiel V_{AXO} décrit complètement ces transitions et permet d'énoncer le critère de stabilité du système. L'influence des différents paramètres de polarisation et de dimensionnement sur la courbe caractéristique du BB-LTC confirme la pertinence d'une telle analyse. Par ailleurs, la correspondance entre les simulations compactes du BB-LTC utilisant le modèle de FD-SOI BIMOS développé au chapitre 3 et les mesures valide la démarche adoptée pour la construction du modèle. L'impact de la vitesse de variation de V_{MEM} sur la courbe courant-tension d'entrée du BB-LTC est également mis en évidence et expliqué. En particulier, l'augmentation de dV_{MEM}/dt a pour conséquence l'augmentation du potentiel de déclenchement V_{TH} .

Des chronogrammes du potentiel V_{MEM} d'un premier démonstrateur de BB-LIF SN pourvu d'une capacité de membrane non-intégrée sur le silicium et soumis à des courants synaptiques constants ont été mesurés à température ambiante. Les résultats montrent que le BB-LIF SN a une excitabilité de type 1. Il implémente par ailleurs une période réfractaire due à la constante de temps associée à la dynamique de décharge du nœud d'axone après les déclenchements. Les influences des potentiels de contrôle V_{FG}^N et V_{BG}^B sur ces chronogrammes ont pu être mises en évidence et expliquées. Le modèle analytique simple proposé permet d'évaluer la fenêtre de fonctionnement du BB-LIF SN en fonction de ses paramètres de dimensionnement. Finalement, le dimensionnement d'un second

démonstrateur de BB-LIF SN dont tous les composants sont intégrés sur le silicium a été décrit et argumenté et l'échantillon correspondant mesuré. Ses fréquences de déclenchement s'étendent de 3 à 75 kHz en réponse à de courants synaptiques compris entre 600 pA et 25 nA. Ces mesures sont cohérentes avec la fenêtre de fonctionnement prévue par le modèle analytique simple et les simulations basées sur le modèle compact de BIMOS développé au chapitre 3. Les performances du BB-LIF SN sont positionnées de façon honorable vis-à-vis de l'état de l'art, avec une surface de $15\mu\text{m}^2$ et une efficacité énergétique de 2 pJ/impulsion à la fréquence de déclenchement de 10 kHz.

La force du travail présenté dans cette thèse se décline selon trois aspects :

- Du point de vue conceptuel, le circuit de neurone proposé est fondé sur la fonctionnalisation de mécanismes physiques considérés comme parasites dans les conceptions standards de circuits analogiques. Cette démarche permet la réduction du nombre de composants nécessaires pour réaliser la fonction de transfert du neurone.
- Du point de vue scientifique, les rôles de chacun des mécanismes physiques mis en jeu dans le fonctionnement du neurone sont clairement identifiés et quantifiés. Cela a permis la construction d'un modèle compact du FD-SOI BIMOS, dont la validité est attestée par la cohérence entre les résultats de simulation et de mesure. De plus, l'énonciation du critère de stabilité du système et la construction d'un modèle simple des limites de fonctionnement du neurone offrent des outils efficaces pour guider le dimensionnement du circuit.
- Du point de vue technologique, la livraison d'un démonstrateur fonctionnel entièrement intégré sur le silicium apporte la preuve de la validité globale de la démarche.

Par ailleurs, l'ensemble des résultats permet d'appréhender les limites de l'approche suivie pour la conception du BB-LIF SN. Ainsi, comme son fonctionnement repose sur la génération de paires électron-trou dans le corps du BIMOS, il est dès lors difficile de réduire le potentiel de déclenchement V_{TH} au-dessous de 1.5 V. De plus, de par la limitation de la dynamique de chargement de nœud d'axone par les intensités intrinsèquement faibles des mécanismes de génération, ce potentiel V_{TH} augmente avec dV_{MEM}/dt , ce qui advient notamment lorsque la surface de l'empreinte du circuit diminue (la réduction de la valeur de la capacité de membrane est directement impliquée). De ce fait, dans le cadre du démonstrateur entièrement intégré proposé, V_{TH} devient supérieur à la tension nominale des composants EG et leur fiabilité n'est plus garantie. Enfin, la dynamique de déchargement du nœud d'axone limitée par le courant de saturation du NMOS du BB-LTC majore la fréquence maximale de fonctionnement. L'accélération possible de cette dynamique de décharge par l'augmentation de la conductance du NMOS a pour effet collatéral d'augmenter le potentiel de seuil V_{TH} .

Plusieurs pistes de travail pourraient être envisagées pour poursuivre le développement du BB-LIF SN :

- Dans la perspective de son intégration à grande échelle dans un processeur neuronal, une quantification de la variabilité des grandeurs caractéristiques du BB-LTC serait indispensable.
- Dans le cadre de la modélisation compacte du FD-SOI BIMOS, l'équation qui décrit le courant de génération par ionisation par impact pourrait être modifiée afin que les simulations reproduisent plus fidèlement les mesures. Pour ce faire, les approximations des intégrales qui impliquent la norme du champ électrique seraient à redéfinir. Par ailleurs, le modèle compact du FD-SOI BIMOS pourrait être complété par l'évaluation et l'ajout de toutes les grandeurs manquantes qui impliquent le terminal de body-contact (capacités parasites, sources de bruit, courants de fuite par la grille).
- Pour l'amélioration des performances du BB-LIF SN, des étapes du procédé de fabrication pourraient être modifiées pour tenter d'augmenter le rendement des mécanismes de génération de paires électron-trou et réduire ainsi le potentiel de seuil.

- Toujours dans le cadre d'une amélioration des performances, le circuit du BB-LIF SN pourrait être modifié par l'ajout de bloc de rétroaction dynamique qui ajusterait les polarisations de contrôle.

Il serait par ailleurs intéressant de porter le circuit du BB-LIF SN sur d'autres technologies SOI afin d'évaluer sa possible généralisation. Les résultats issus des travaux de recherche présentés dans ce manuscrit ont permis la proposition d'un nouveau sujet de thèse qui porte, d'une part, sur l'extension de l'utilisation du BB-LTC dans des circuits de neurone impulsionnel plus complexes, et d'autre part, sur l'étude de SNN qui implémentent le BB-LIF SN.

Liste des brevets et publications

Brevets

T. Bédécarrats, P. Galy, « Circuit intégré avec double isolation du type tranchées profondes et peu profondes », 2019.

P. Galy et **T. Bédécarrats**, « Protection d'un circuit intégré contre un percement d'un contact source et/ou drain », 2019.

T. Bédécarrats, L. De Conti and Ph. Galy, « Non homogeneous silicide for thin-film carrier transport control for example to build an ESD high or low voltage protection in UTBB FD-SOI technology », 2018.

P. Galy et **T. Bédécarrats**, « Dispositif intégré de neurone artificiel », FR3064383A1, 2017.

P. Galy et **T. Bédécarrats**, « Bloc réfractaire pour dispositif intégré de neurone artificiel », FR3064384A1, 2017.

P. Galy et **T. Bédécarrats**, « Dispositif de génération d'un signal aléatoire », FR3072481A1, 2017.

H. El Dirani, **T. Bédécarrats**, et P. Galy, « Matrice Mémoire à point mémoire de type Z²-FET », FR3074352A1, 2017.

Papiers de journal

T. Bédécarrats, C. Fenouillet-Béranger, S. Cristoloveanu and P. Galy, « A BIMOS-based 2T1C analogue spiking neuron circuit integrated in 28nm FD-SOI technology for neuromorphic application », *Solid-State Electronics*, to be published.

T. Bédécarrats, P. Galy, C. Fenouillet-Béranger, and S. Cristoloveanu, « Investigation of built-in bipolar junction transistor in FD-SOI BIMOS », *Solid-State Electronics*, March 2019.

L. De Conti, **T. Bédécarrats**, S. Cristoloveanu, M. Vinet, and P. Galy, « GDNMOS and GDBIMOS devices for high voltage ESD protection in thin film advanced FD-SOI technology », *Solid-State Electronics*, March 2019.

Papiers de conférence

P. Galy, **T. Bédécarrats**, C. Fenouillet-Béranger and S. Cristoloveanu « Thin film BIMOS transistor for low power spiking neuron cell in 28nm FD-SOI CMOS technology », in *2019 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, 2019, pp. 1–4.

T. Bédécarrats, C. Fenouillet-Béranger, S. Cristoloveanu, and P. Galy, « A new sharp switch device integrated in 28 nm FD-SOI technology », in *2018 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, 2018, pp. 1–3.

T. Bédécarrats, P. Galy, C. Fenouillet-Béranger, and S. Cristoloveanu, « Investigation on built-in BJT in FD-SOI BIMOS », in *2018 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, 2018, pp. 1–4.

L. D. Conti, **T. Bédécarrats**, M. Vinet, S. Cristoloveanu, and P. Galy, « Toward Gated-Diode-BIMOS for thin silicon ESD protection in advanced FD-SOI CMOS technologies », in *2017 IEEE International Conference on IC Design and Technology (ICICDT)*, 2017, pp. 1–4.

L. Kadura, L. Grenouillet, **T. Bédécarrats**, O. Rozeau, N. Rambal, P. Scheiblin, C. Tabone, D. Blachier, O. Faynot, A. Chelnokov and M. Vinet, « Extending the functionality of FDSOI N- and P-FETs to light sensing », in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 32.6.1-32.6.4.