



HAL
open science

Hopping et Tunnelling assisté par pièges dans les mémoires non volatile à base d'oxyde métallique

Nazim Ait Abdelkader

► **To cite this version:**

Nazim Ait Abdelkader. Hopping et Tunnelling assisté par pièges dans les mémoires non volatile à base d'oxyde métallique. Micro et nanotechnologies/Microélectronique. Université Grenoble Alpes [2020-..], 2024. Français. NNT : 2024GRALT042 . tel-04708864

HAL Id: tel-04708864

<https://theses.hal.science/tel-04708864v1>

Submitted on 25 Sep 2024

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : EEATS - Electronique, Electrotechnique, Automatique, Traitement du Signal (EEATS)

Spécialité : Nano électronique et Nano technologies

Unité de recherche : Centre de Radiofréquences, Optique et Micro-nanoélectronique des Alpes

Hopping et Tunnelling assisté par pièges dans les mémoires non volatile à base d'oxyde métallique

Hopping and tunnelling trap assisted current in metallic oxide based non volatile memories

Présentée par :

Nazim AIT ABDELKADER

Direction de thèse :

Quentin RAFHAY

MAITRE DE CONFERENCES, Université Grenoble Alpes

Directeur de thèse

Rapporteurs :

Marc BOCQUET

PROFESSEUR DES UNIVERSITES, AIX-MARSEILLE UNIVERSITE

Damien QUERLIOZ

DIRECTEUR DE RECHERCHE, CNRS DELEGATION ILE-DE-FRANCE SUD

Thèse soutenue publiquement le **21 juin 2024**, devant le jury composé de :

Jean-Christophe TOUSSAINT,

PROFESSEUR DES UNIVERSITES, Grenoble INP

Président

Quentin RAFHAY,

MAITRE DE CONFERENCE HDR, Grenoble INP

Directeur de thèse

Marc BOCQUET,

PROFESSEUR DES UNIVERSITES, AIX-MARSEILLE UNIVERSITE

Rapporteur

Damien QUERLIOZ,

DIRECTEUR DE RECHERCHE, CNRS DELEGATION ILE-DE-FRANCE SUD

Rapporteur

Damien DELERUYELLE,

PROFESSEUR DES UNIVERSITES, INSA Lyon

Examineur

Abdelkader SOUFI,

PROFESSEUR DES UNIVERSITES, INSA Lyon

Examineur



Résumé

Les ReRAM, en particulier l'OxRAM, sont reconnues comme des technologies essentielles pour les applications émergentes en raison de leur haute densité, de leur accès rapide, de leur endurance et de leur intégration dans le back-end-of-line. Toutefois, les modèles régissant le mécanisme de conduction dans les OxRAM ont suscité de nombreux débats au cours des quinze dernières années. Cette thèse présente un nouveau formalisme qui implique la résolution des niveaux de Fermi des pièges pour assurer un courant constant entre eux. Ce formalisme a permis d'explorer deux modèles de courant sans recourir à des approches d'identification du chemin le plus conducteur. Un modèle généralisé de Tsu-Esaki, basé sur une conduction entre un continuum d'états, a été exposé mais n'a cependant pas été en mesure de reproduire de manière précise le comportement électrique dans les deux états de résistance de l'OxRAM. En revanche, un deuxième modèle basé sur l'équation de Miller et Abrahams, impliquant une conduction entre états liés, a démontré sa capacité à fournir un excellent ajustement aux données expérimentales, tout en utilisant un nombre de pièges beaucoup plus raisonnable par rapport aux approches existantes. Des énergies d'activation thermique ont également été calculées pour étudier la possibilité d'identifier ces mécanismes de manière expérimentale. Enfin, une partie expérimentale a été intégrée à cette thèse, consistant en une tentative de caractérisation électrique en fonction de la température afin d'identifier le ou les mécanismes de conduction dans une OxRAM.

Abstract

Resistive Random Access Memory, and more specifically Oxide RAM (OxRAM), are considered as key enabling technologies for emerging applications due to their high density, high speed, good endurance and integration in the back-end-of-line. However, the mechanisms that rule their conduction have been heavily debated in the last fifteen years. A new method, involving the solution of the traps Fermi levels is presented in this manuscript, to ensure constant current between traps, which allows to include two models for the conduction and avoid the identification of the most conducting path. A generalized Tsu-Esaki model is detailed, which is based on a conduction between continuum of states. It is found that this model is not able to provide a suitable fit of the current-voltages (I-V) curves in both resistance states of an OxRAM. A second model, based on Miller and Abrahams equation and hence on the conduction model between bound states has proved its ability to provide an excellent agreement with experimental data, employing a much more reasonable number of traps compared to already existing approaches. Thermal activation energy are finally computed to investigate the possibility to experimentally identify these mechanisms. Ultimately, an experimental component has been incorporated into this thesis. This involved an endeavor towards electrical characterization as a function of temperature, aiming to identify conduction mechanisms involved in OxRAMs.

Remerciements

La réalisation de cette thèse a été rendue possible grâce au soutien et à l'accompagnement de nombreuses personnes auxquelles je tiens à exprimer toute ma gratitude.

Je tiens tout d'abord à remercier chaleureusement Quentin Rafhay, mon directeur de thèse, pour son encadrement, ses précieux conseils et sa disponibilité tout au long de ce parcours. Son expertise et sa confiance m'ont permis de mener à bien ce projet avec sérénité.

Mes remerciements s'adressent également à Anna Kaminski, directrice de laboratoire, pour son soutien et sa vision scientifique, qui ont grandement contribué à l'évolution de mes recherches.

Je tiens à exprimer toute ma reconnaissance à mes collègues Léo, Chloé, Mathilde et Greg pour leurs encouragements constants, les discussions enrichissantes et le soutien inestimable qu'ils m'ont apporté au quotidien.

Je souhaite également remercier l'équipe technique, et tout particulièrement Xavier, pour son aide précieuse dans la réalisation de nombreux aspects pratiques de cette thèse.

Mes remerciements vont aussi à l'équipe administrative, en particulier Daliha, pour son efficacité et sa disponibilité, qui ont été des soutiens indispensables durant tout ce parcours.

Je tiens à exprimer toute ma reconnaissance à mes parents. Vous le savez, sans votre soutien indéfectible, je n'aurais jamais pu accomplir une telle étape. Les mots me manqueront toujours pour décrire l'ampleur des sacrifices que vous avez consentis pour mon éducation et pour que je ne manque de rien.

Je souhaite également remercier ma grand-mère, qui m'a élevé avec autant d'amour et de soin que si j'avais été son propre fils. Les valeurs qu'elle m'a transmises resteront à jamais gravées en moi. Enfin, je remercie mes tantes pour leur amour inconditionnel et leur soutien constant, qui ont été des sources de force tout au long de mon parcours.

À toutes ces personnes, ainsi qu'à celles que je n'ai peut-être pas mentionnées mais qui ont joué un rôle dans l'accomplissement de ce projet, je vous adresse mes plus sincères remerciements.

Table des matières

1	Introduction	2
2	Etat de l'art	4
2.1	Contexte	4
2.1.1	Croissance exponentielle de l'utilisation des données	4
2.1.2	Utilisation traditionnelle des mémoires	5
2.1.3	L'approche "More than Moore"	7
2.2	Mémoires non-volatiles émergentes	8
2.2.1	Mémoires ferroélectriques FeRAM	9
2.2.2	Mémoires Magnétiques MRAM	11
2.2.3	Mémoires à changement de phase PCM	11
2.3	Mémoires résistives ReRAM	13
2.3.1	Principe de fonctionnement d'une ReRAM	13
2.3.2	Classification des ReRAM	14
2.3.3	Mémoire CbRAM	15
2.3.4	Mémoire OxRAM	17
2.3.5	État de l'art en matière de performances	18
2.3.6	Comparaison des principales mémoires NVM émergentes	21
2.4	Applications des mémoires non volatiles émergentes	23
2.4.1	Mémoire de Stockage	23
2.4.2	Mémoire Principale	24
2.4.3	Cache de Dernière Niveau	24
2.4.4	In Memory Computation	24
2.5	État de l'art de la modélisation de la conduction	25
2.5.1	Distinction HRS/LRS	26
2.5.2	TAT Larcher et al.	27
2.5.3	TAT Ielmini et al.	31
2.5.4	QPC	34
2.5.5	Poole-Frenkel	35
2.6	Objectifs et problématique de la thèse	38

3	Modélisation de la conduction	40
3.1	Introduction	40
3.2	Méthode numérique pour résolution des niveaux de Fermi	41
3.2.1	Cadre de simulation	41
3.2.2	Nécessité de résoudre les niveaux de Fermi	42
3.2.3	La méthode de Newton-Raphson	43
3.2.4	Approche Nearest Neighbour	45
3.2.5	Approche Variable Range	46
3.3	Conduction assistée par piège	48
3.3.1	Formulation du modèle de TE (Tsu-Esaki) en Boltzmann	49
3.3.2	Formulation du modèle de TE en dégénéré	56
3.3.3	Comparaison en fonction du nombre de piège des modèles TE Boltzmann	67
3.3.4	Formulation du modèle TE dans le cadre générale	67
3.4	Conduction par saut : Hopping	74
3.4.1	Modélisation	76
3.4.2	Validation modèle	88
3.5	Prise en compte des simulations 3D	92
3.6	Exploitation des résultats	94
3.6.1	Discrimination des modèles de courants	94
3.6.2	Bilan sur les modèles de conduction	98
3.6.3	Comparaison avec des données expérimentales	99
4	Etude expérimentale	103
4.1	Introduction	103
4.2	Description du protocole expérimentale	104
4.2.1	Présentation de l'Elite 300	104
4.2.2	Présentation de l'Agilent B1500	104
4.2.3	Dispositif testé	105
4.2.4	Plan de mesure	106
4.3	Exploitation des données	107
5	Conclusion	112
6	Perspectives	115
6.1	Prise en compte du bruit RTN	115
6.2	Simulation Réaliste	115
6.3	Nouveau plan de mesure	116

Table des figures

2.1	Quantité de donnée générée chaque année [1].	5
2.2	Schéma de l'architecture de Von Neumann.	6
2.3	Schéma de la hiérarchie mémoire.	6
2.4	Graphe montrant : la tendance à la miniaturisation des transistor, les performances à un seul thread, le nombre de coeur logique et la puissance consommé par transistor [2].	8
2.5	Principe de fonctionnement d'une FeRAM [3]	10
2.6	Structure FeFET[4]	10
2.7	(a) Schéma d'une cellule MRAM (b) Caractéristique Résistance en fonction du courant d'une MRAM	12
2.8	(a) Schéma d'une cellule PCM (b) Caractéristique Courant en fonction de la Tension d'une PCM	13
2.9	Schéma des différents états de résistance d'une ReRAM	14
2.10	Courbe caractéristique courant-tension d'un dispositif ReRAM : (a) Unipolaire (b) Bipolaire. Source : [5]	16
2.11	Courbe I-V caractéristique d'une CbRAM.	17
2.12	Schéma représentant le mécanisme de commutation d'une OxRAM. Source : [6]	19
2.13	Ecart type expérimental et simulé de la résistance en fonction de la résistance moyenne. Source : [7]	20
2.14	Distributions N_{cmax} expérimentales et simulées. Source : [8]	20
2.15	Schéma représentatif du mécanisme de : (a) diffusion d'ions d'oxygène (b) diffusion de lacunes d'oxygène. Source : [9]	21
2.16	Schéma représentatif (a) D'une architecture IMC (b) Exemple de programme à exécuter avec un système IMC. Source : [10]	26
2.17	Schéma représentatif de la densité de piège importante en LRS dans une OxRAM. Source : [11]	27
2.18	Courbes I-V comparatives entre donnée expérimentales et simulation du modèle ohmique en LRS	27
2.19	Schéma récapitulatif du processus TAT	28
2.20	La capture depuis cathode et l'émission vers l'anode ne sont limitantes	29

2.21	Schéma représentant l'hypothèse du calcul de la transparence	29
2.22	Diagramme fonctionnelle de l'algorithme de calcul des chemins de percolation	30
2.23	Courbe I-V en fonction de la température, issu du modèle TAT larcher et <i>al.</i> Source : [12]	31
2.24	Schéma récapitulatif du modèle de Ielmini à 1 piège. Source :[13]	32
2.25	Schéma récapitulatif du modèle de Ielmini à deux pièges. Source : [14]	33
2.26	Diagramme de bande représentant la zone de constriction. Source : [15]	34
2.27	Schéma récapitulatif du modèle de QPC : (a) HRS (b) LRS. Source : [15]	36
2.28	Courbe courant en fonction de la tension en : (a) HRS (b) LRS. Source : [15]	36
2.29	Schéma du mécanisme Poole-Frenkel. Source : [16]	37
2.30	Courbe courant en fonction de la tension issu de données de simulation en HRS et LRS. Source : [17]	37
3.1	Schéma du diagramme de bande de l'OxRAM. x_i représente la position du piège le long de l'oxyde. E_{f_i} représente le niveau de Fermi du piège. E_{t_i} représente l'énergie du piège. V_i représente la tension appliquée sur le piège.	42
3.2	Schéma de l'approche Nearest Neighbour.	45
3.3	Schéma de l'approche Variable Range sur une chaîne de pièges : Tout les pièges avant la section i_S contribuent au courant.	47
3.4	Cas où $E_{t_i} < E_{t_{i-1}}$: Diagramme de bande simplifié sur trois pièges expliquant le modèle TE simplifié. Φ_i représente la hauteur de barrière, E_{f_i} (traits en pointillés)est le niveau de Fermi au niveau du piège i, E_{t_i} (traits plein) est le niveau d'énergie au niveau du piège i, V_i représente la tension au niveau du piège i.	50
3.5	Cas où $E_{t_i} > E_{t_{i-1}}$: Diagramme de bande simplifié sur trois pièges expliquant le modèle TE simplifié. Φ_i représente la hauteur de barrière, E_{f_i} (traits en pointillés)est le niveau de Fermi au niveau du piège i, E_{t_i} (traits plein) est le niveau d'énergie au niveau du piège i, V_i représente la tension au niveau du piège i.	51
3.6	Courant en en fonction de la position des pièges en utilisant le modèle TE simplifié Boltzmann. Les paramètres utilisés : $N_t=5$, $E_t=0.2$ eV et $V=0.5$ V.	53
3.7	Niveau de Fermi résolu en fonction de la position des pièges. Les paramètres utilisés : $N_t = 5$, $E_t = 0.2$ eV et $V=0.5$ V	53
3.8	(a) Courant en fonction de la tension avec 6 piège (HRS) (b) Courant en fonction de la tension avec 25 pièges (LRS). Les paramètres utilisée : $E_t = 0.2$ eV	54

3.9	Courbe courant en fonction de la tension pour le modèle TE simplifié Boltzmann avec l'approche NN et VR en LRS. Les paramètres utilisés : $N_t=20$, $E_t=0.2$ eV	55
3.10	Courbe courant en fonction de la tension pour le modèle TE simplifié Boltzmann avec l'approche NN et VR en HRS. Les paramètres utilisés : $N_t=6$, $E_t=0.2$ eV	55
3.11	Courbe courant en fonction de la tension en LRS sur un dispositif 1R mesurée issu de [18]	56
3.12	Courbe courant en fonction de la tension en LRS pour différents t_{ox} . Les paramètres utilisés : $E_t=0.2$ eV, $N_t=25$, $\Phi_l=3.1$ eV.	57
3.13	Courbe courant en fonction de la tension en LRS pour différents E_t . Les paramètres utilisés : $\Phi_l=3.1$ eV, $N_t=25$, $t_{ox}=5$ nm	57
3.14	Courbe courant en fonction de la tension en LRS pour différents Φ_l . Les paramètres utilisés : $E_t=0.2$ eV et $N_t=25$, $t_{ox}=5$ nm	58
3.15	Diagramme de bande simplifié sur trois pièges expliquant le modèle TE dégénéré. Φ_l représente la hauteur de barrière, E_{f_i} (traits en pointillés) est le niveau de Fermi au niveau du piège i, E_{t_i} (traits pleins) est le niveau d'énergie au niveau du piège i, V_{t_i} représente la tension au niveau du piège i.	58
3.16	Courbe niveaux de Fermi en fonction de la position en TE dégénéré en utilisant l'approche NN. Les paramètres utilisées : $N_t=5$, $E_t=0.2$ eV, $V=0.5$ V	62
3.17	Courbe courant en fonction de la position en TE dégénéré en utilisant l'approche NN. Les paramètres utilisés : $N_t=5$, $E_t=0.2$ eV, $V=0.5$ V	62
3.18	Courbe courant en fonction de la tension pour le modèle dégénéré en (a) HRS et en (b) LRS. Les paramètres utilisés : (a) $N_t=6$, $E_t=0.2$ eV et (b) $N_t=25$, $E_t=0.2$ eV	63
3.19	Courbe courant en fonction de la tension pour le modèle dégénéré et non dégénéré en HRS. Les paramètres utilisées : $N_t=6$, $E_t=0.2$ eV	64
3.20	Courbe courant en fonction de la tension pour le modèle dégénéré et non dégénéré en LRS. Les paramètres utilisées : $N_t=25$, $E_t=0.2$ eV	64
3.21	Courbe courant en fonction de la tension pour le modèle dégénéré en HRS en utilisant l'approche VR. Les paramètres utilisés : $N_t=6$, $E_t=0.2$ eV	65
3.22	Courbe courant en fonction de la tension pour le modèle dégénéré en LRS en utilisant l'approche NN et l'approche VR. Les paramètres utilisés : $N_t=25$, $E_t=0.2$ eV	65
3.23	Courbe courant en fonction de la tension pour le modèle dégénéré et non dégénéré en HRS en fonction de la température en utilisant l'approche NN. Les paramètres utilisés : $N_t=6$, $E_t=0.2$ eV	66

3.24	Courbe courant en fonction de la tension pour le modèle dégénéré et non dégénéré en LRS en fonction de la température en utilisant l'approche NN. Les paramètres utilisés : $N_t=25$, $E_t=0.2$ eV	66
3.25	Courbe courant en fonction du nombre de pièges en dégénéré et non dégénéré. Les paramètres utilisées : $E_t=0.2$ eV, $V = 1V$	68
3.26	Courant en fonction de la position à $V = 1$ V pour le modèle de Tsu-Esaki en NN et VR. Les paramètres utilisés : $N_t=12$, $E_t=0.2$ eV	71
3.27	(a) Niveau de Fermi en fonction de la position à $V = 0,5$ V pour le modèle de Tsu-Esaki en NN et VR. (b) Niveau de Fermi d'un piège au milieu de l'oxyde en fonction de la tension pour le modèle de Tsu-Esaki en NN et VR. Les paramètres utilisés : $N_t=12$, $E_t = 0.2$ eV . .	72
3.28	(a) Courbe I-V avec le modèle Tsu-Esaki en LRS (25 pièges). (b) Courbe I-V avec le modèle Tsu-Esaki en HRS (6 pièges). Les paramètres utilisés : $E_t = 0.2$ eV	73
3.29	Courbe I-V en fonction de la température pour le modèle Tsu-Esaki NN en HRS. Les paramètres utilisés : $E_t = 0.2$ eV et $N_t = 6$	73
3.30	Courbe I-V en fonction de la température pour le modèle Tsu-Esaki NN en LRS. Les paramètres utilisés : $E_t = 0.2$ eV et $N_t = 25$	74
3.31	Courbe I-V en fonction de la température pour le modèle Tsu-Esaki VR en HRS. Les paramètres utilisés : $E_t = 0.2$ eV et $N_t = 6$	75
3.32	Courbe I-V en fonction de la température pour le modèle Tsu-Esaki VR en LRS. Les paramètres utilisés : $E_t = 0.2$ eV et $N_t = 6$	75
3.33	Schéma explicatif du Hopping	76
3.34	Schéma explicatif de la transition entre deux états	79
3.35	Courant en fonction de la position à $V = 1$ V pour le modèle de Miller et Abrahams en utilisant l'approche Nearest Neighbour et Variable range. Les paramètres utilisés : $E_t = 0.2$ eV, $N_t = 25$, $\alpha = 10^{-10}$ m, $w_0 = 10^{17} s^{-1}$	89
3.36	(a) Niveau de Fermi en fonction de la position à $V = 0.5$ V pour le modèle de Miller et Abrahams en NN et VR. (b) Niveau de Fermi d'un piège en fonction de la tension appliqué pour le modèle de Miller et Abrahams en NN et VR. Les paramètres utilisés : $E_t = 0.2$ eV, $N_t = 25$, $\alpha = 10^{-10}$ m, $w_0 = 10^{17} s^{-1}$	90
3.37	(a) Courbe I-V pour le modèle de Miller et Abrahams en LRS (avec 25 pièges). (b) Courbe I-V pour le modèle de Miller et Abrahams en HRS (avec 6 pièges). Les paramètres utilisés : $E_t = 0.2$ eV, $\alpha = 10^{-10}$ m, $w_0 = 10^{17} s^{-1}$	91
3.38	(a) Courbe I-V à différentes températures pour le modèle de Miller et Abrahams en LRS (avec 25 pièges). Les paramètres utilisés : $E_t = 0.2$ eV, $\alpha = 10^{-10}$ m, $w_0 = 10^{17} s^{-1}$	91

3.39	Courbe I-V à différentes températures pour le modèle de Miller et Abrahams en HRS (avec 6 pièges). Les paramètres utilisés : $E_t = 0.2\text{eV}$, $\alpha = 10^{-10}\text{m}$, $w_0 = 10^{17}\text{s}^{-1}$	92
3.40	Schéma représentant la situation où plusieurs pièges sont générés pour un $x = 2$ nm dans un oxyde de 5nm d'épaisseur.	93
3.41	Courbe de l'énergie d'activation en fonction de la tension appliquée pour le modèle de (a) Miller et Abrahams en HRS (6 pièges) et (b) en LRS (25 pièges).	96
3.42	Courbe de l'énergie d'activation en fonction de la tension appliquée pour le modèle de Tsu-Esaki en HRS (6 pièges) et $E_t=0.2$ eV et (b) en LRS (25 pièges) et $E_t=0.2$ eV.	96
3.43	(a) Gamma en fonction de la tension appliquée en LRS pour (a) le modèle de Miller et Abrahams and (b) modèle de Tsu-Esaki. 25 pièges ont été utilisés pour les deux simulations et $E_t=0.2$ eV.	97
3.44	Bilan sur les deux modèles de conduction	98
3.45	Temps d'exécution des des codes utilisant l'approche NN	99
3.46	Temps d'exécution des des codes utilisant l'approche VR	100
3.47	Comparaison entre les courbes I-V issues du modèle de Miller et Abrahams model et les données expérimentales issues de : (a) Source : [18] et (b) Source : [19], en HRS. Les paramètres utilisées :(a) 6 pièges, $w_0 = 3.5 \times 10^{18} \text{ s}^{-1}$, $\alpha = 0.15 \times 10^{-10} \text{ m}$, $E_t = 0.13 \text{ eV}$ et (b) 6 pièges, $w_0 = 10^{17} \text{ s}^{-1}$, $\alpha = 10^{-10} \text{ m}$, $E_t = 0.2 \text{ eV}$	101
3.48	Comparaison entre les courbes I-V issues du modèle de Miller et Abrahams et les données expérimentales issues de : (a) [18] et (b) [19], en LRS. Les paramètres utilisées : (a) 33 pièges, $w_0 = 3.5 \times 10^{18} \text{ s}^{-1}$, $\alpha = 0.15 \times 10^{-10} \text{ m}$, $E_t = 0.13 \text{ eV}$ et (b) 16 pièges, $w_0 = 10^{17} \text{ s}^{-1}$, $\alpha = 10^{-10} \text{ m}$, $E_t = 0.2 \text{ eV}$	102
4.1	(a) Schéma de la structure de test 1T1R. (b) Schéma équivalent de la structure de test	105
4.2	Schéma du plan d'expérience	107
4.3	Courbe I-V en LRS issue des données expérimentale brute.	108
4.4	Courbe I-V en HRS issue des données expérimentale brute	108
4.5	Courbe I-V en LRS issue des données expérimentale ajustées.	109
4.6	Courbe I-V en HRS issue des données expérimentale ajustées.	109
4.7	Courbe E_a -V en LRS issue des données expérimentale : (a) Brute (b) Ajustées	110
4.8	Courbe E_a -V en HRS issue des données expérimentale : (a) Brute (b) Ajustées	110

6.1	Schéma du processus de suppression de piège pour modélisation du bruit RTN. Source : [20].	116
6.2	Nouveau plan de mesure pour caractérisation I-V en température. .	117

Chapitre 1

Introduction

Avec l'augmentation de la numérisation des tâches, il est devenu essentiel de conserver et de traiter de vastes volumes de données. Afin de satisfaire cette demande, il est indispensable de concevoir des technologies de mémoire électronique qui peuvent stocker et traiter ces données de manière efficace. Grâce aux progrès technologiques dans la production des composants semi-conducteurs, il est possible de concevoir des mémoires électroniques plus rapides, plus performantes et plus compactes. Les différentes étapes de miniaturisation des composants ont permis de réaliser ces progrès, ce qui a permis de diminuer la taille des transistors et d'accroître le nombre de composants sur une même puce. Cependant, même avec ces progrès, des difficultés demeurent en ce qui concerne le temps d'accès, en particulier entre les mémoires volatiles et les mémoires non-volatiles. Les performances globales du système de stockage peuvent être impactées par cette période de latence entre ces différents niveaux de mémoire. En outre, les mémoires traditionnelles à semi-conducteurs connaissent leurs limites en ce qui concerne la densité d'intégration. Il est donc essentiel de mettre au point de nouvelles technologies de mémoire afin de satisfaire les besoins actuels et futurs en matière de stockage de données.

Afin de relever les défis actuels liés au stockage et au traitement des données, l'étude se focalise sur des technologies émergentes qui peuvent présenter des bénéfices en termes de performances et de coût. L'OxRAM, également connu sous le nom de mémoire à base d'oxyde métallique, fait partie de ces technologies prometteuses. Selon [21], elle exploite la propriété électrique de certains oxydes métalliques afin d'obtenir deux états de résistance distincts et commutables, ce qui permet de stocker des données numériques. Cette technologie présente des avantages significatifs par rapport aux mémoires traditionnelles. Tout d'abord, elle offre une meilleure rétention des données, ce qui signifie que les données stockées sont moins susceptibles d'être perdues. De plus, elle a une endurance supérieure, ce qui signifie qu'elle peut subir un grand nombre de cycles d'écriture et de lecture sans se détériorer [22]. En outre, l'OxRAM a une vitesse de lecture et d'écriture plus rapide que les mémoires

traditionnelles, ce qui permet un accès plus rapide aux données stockées. Enfin, cette technologie peut résoudre la problématique de temps de latence entre les différents niveaux de mémoire, ce qui peut améliorer les performances globales du système de stockage [23].

Cependant, l'OxRAM rencontre un obstacle majeur : une variabilité significative. En effet, les caractéristiques courant-tension peuvent varier considérablement d'un cycle à l'autre ou d'un dispositif à l'autre, ce qui rend cette technologie difficilement industrialisable à grande échelle. Pour mieux comprendre cette variabilité, il est important de comprendre la physique de trois phénomènes observables dans une OxRAM : le forming, la commutation et la conduction.

Cette thèse a pour objectif d'amener à une compréhension approfondie de la physique sous-jacente au phénomène de conduction dans les OxRAMs.

Dans le chapitre suivant, nous commencerons par établir le contexte de cette thèse. Nous examinerons également les principales technologies de mémoires non volatiles émergentes, en évaluant leurs performances et en discutant des applications potentielles pour ce type de mémoire. Enfin, nous présenterons une revue de la littérature sur les différents modèles de conduction.

Chapitre 2

Etat de l'art

2.1 Contexte

Dans ce chapitre introductif, nous commencerons par décrire en détail le fonctionnement des systèmes informatiques, en mettant l'accent sur la manière dont ils utilisent la mémoire pour le stockage et le traitement des données. Nous analyserons ensuite les limitations des mémoires traditionnelles, notamment en ce qui concerne les temps d'accès et les transferts de données, qui peuvent freiner les performances globales des systèmes informatiques. Enfin, nous expliquerons comment les mémoires non volatiles (NVM) émergentes peuvent représenter une solution à ces problèmes. Nous montrerons comment ces nouvelles technologies offrent des temps d'accès réduits et permettent de développer des architectures innovantes, telles que l'In-Memory Computing, où les calculs sont réalisés directement dans la mémoire, réduisant ainsi les besoins en transferts de données.

2.1.1 Croissance exponentielle de l'utilisation des données

La croissance exponentielle de l'utilisation des données, amplifiée par la prolifération des dispositifs connectés, des applications en ligne et des systèmes de gestion des données, représente un défi majeur pour les infrastructures technologiques actuelles. Chaque jour, des quantités astronomiques de données sont générées, comme le montre le graphe 2.1, nécessitant des capacités de stockage et de traitement toujours plus grandes. Cette explosion des données, souvent qualifiée de "Big Data", requiert non seulement des espaces de stockage plus vastes, mais également des technologies de mémoire plus rapides et plus efficaces pour assurer une gestion optimale. Les technologies traditionnelles de mémoire, atteignent leurs limites en termes de vitesse, de capacité et d'efficacité énergétique. Ainsi, le développement de nouvelles technologies de mémoire, devient crucial pour soutenir cette croissance. Ces innovations permettront de répondre aux besoins croissants en matière de stockage et de

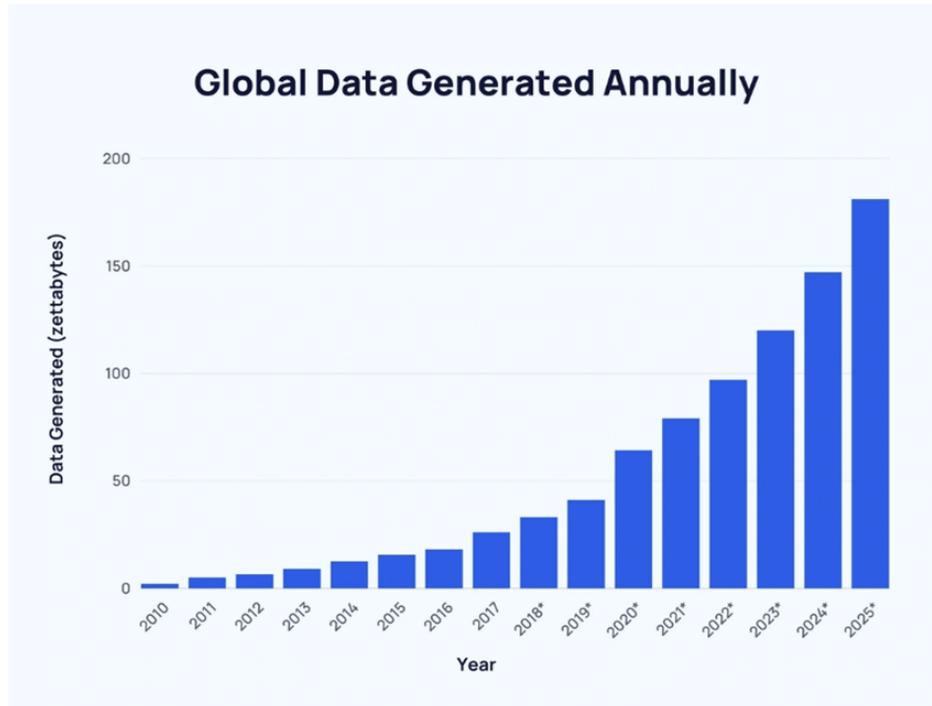


FIGURE 2.1 – Quantité de donnée générée chaque année [1].

traitement des données, tout en améliorant la performance et la fiabilité des systèmes informatiques modernes.

2.1.2 Utilisation traditionnelle des mémoires

Actuellement, la plupart des systèmes informatiques possèdent une architecture de Von Neumann qui sépare le calcul et le stockage. Le calcul est effectué dans un processeur, et les données nécessaires à ce calcul sont stockées en mémoire. Dans cette section, nous introduirons d'abord l'architecture de Von Neumann, puis la hiérarchie mémoire.

Architecture De Von Neumann

La Figure 2.2 montre l'architecture de Von Neumann. Le bloc central est l'unité de calcul. À l'intérieur, le processeur central récupère des données de la mémoire, c'est-à-dire à la fois des instructions et des données sur lesquelles appliquer ces instructions. De plus, il prend des entrées de l'environnement extérieur via des périphériques d'entrée, et il produit des sorties vers l'environnement extérieur via des périphériques de sortie. L'avantage de séparer le traitement et la mémoire est que le processeur central est indépendant du contenu de la mémoire. En d'autres termes, sa structure est toujours la même et n'a pas besoin de changer en fonction des différentes données. De plus, l'architecture de Von Neumann permet de développer ces

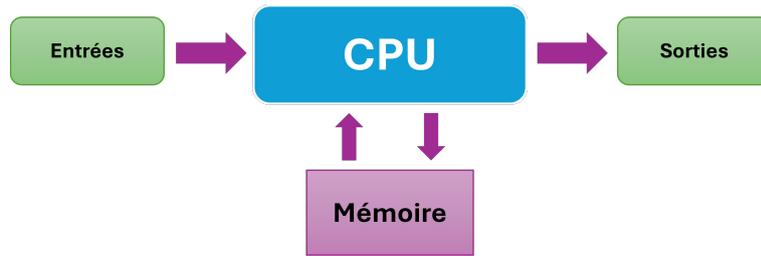


FIGURE 2.2 – Schéma de l'architecture de Von Neumann.

composants séparément.

Hiérarchie Mémoire

La vitesse de la mémoire n'augmente pas aussi rapidement que celle du processeur central. Cela limite la vitesse globale de calcul, car le processeur doit attendre que la mémoire charge et stocke les données. Ce goulot d'étranglement est connu sous le nom de goulot d'étranglement de Von Neumann [24]. Pour tenter de résoudre ce problème, la mémoire est divisée en plusieurs étages qui forment une hiérarchie de mémoire. L'objectif de cette hiérarchie est d'avoir suffisamment de mémoire rapide et coûteuse [25] près du processeur central, tout en ayant une mémoire plus lente et moins chère plus éloignée du processeur. Cette hiérarchie de mémoire est représentée dans la Figure 2.3 et se compose de cinq étages. Les étages supérieurs sont les plus rapides mais aussi les plus chers par bit, tandis que les étages inférieurs sont plus lents mais moins chers par bit. Toutefois l'approche qui a consisté à augmenter la

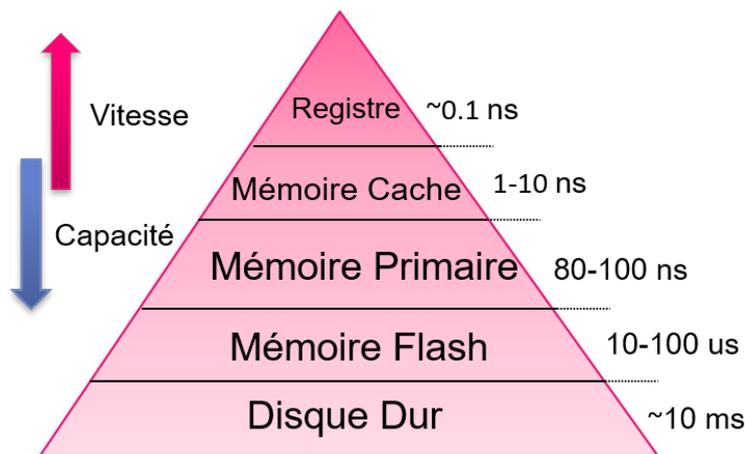


FIGURE 2.3 – Schéma de la hiérarchie mémoire.

taille de la mémoire cache, en augmentant ainsi l'empreinte de la SRAM sur les puces de CPU a finit par diminuer le débit d'information net. Bien que la DRAM offre une densité supérieure à celle de la SRAM, elle offre des temps d'accès beaucoup

plus lent que ceux de la SRAM [26]. Historiquement, pour avoir une mémoire non-volatile, on se reposait sur des supports de stockage externes tels que les disques durs magnétiques (HDD), qui ont des temps d'accès beaucoup plus lents que les mémoires volatiles. Les SSD à base de NAND Flash offrent une non-volatilité avec des latences inférieures à celles des HDD, mais ont un mécanisme d'effacement par page nécessitant une procédure de réécriture, et ont une endurance limitée [27]. Le développement d'une technologie de mémoire non volatile, combinant la vitesse de la DRAM, et l'endurance et la capacité de miniaturisation de la NAND Flash, pourrait transformer l'architecture des ordinateurs. En tirant parti, d'un processus de fabrication compatible aux technologies CMOS, une telle technologie de mémoire pourrait ouvrir la voie à des avancées dans les applications de mémoire autonomes et intégrées. Cela aurait été le cas idéal, sauf que le concept de mémoire universelle a été abandonné il y a une dizaine d'année [28]. Il est donc nécessaire de trouver un autre paradigme pour répondre aux besoins actuels. C'est ce qui sera exposé dans la section suivante.

2.1.3 L'approche "More than Moore"

Avant d'être le co-fondateur, Gordon Moore a imaginé en 1965 une loi, qui s'apparente à une prédiction économique, selon laquelle le nombre de transistors devrait doubler tout les deux ans. Jusqu'à il y a une vingtaine d'années, cette loi fut suivie à la lettre par l'industrie des semi-conducteurs. En début d'années 2024, Intel a déployé dans ces usines de production un nouveau noeud technologique de 3nm [29], tandis que International Roadmap for Devices and Systems de 2021 prévoit d'atteindre des noeuds technologique de 1 nanomètre [2]. Cependant, franchir la barrière du nanomètre pose des défis physiques considérables pour le processus de miniaturisation [30, 31]. Comme le montre la figure 2.4, depuis 2005, il est devenu de plus en plus difficile de maintenir la croissance exponentielle de l'intégration, principalement en raison de problèmes tels que la dissipation thermique excessive, la consommation d'énergie élevée et du processus de lithographie [30]. C'est pourquoi de nombreux acteurs du secteur privilégient désormais une approche dite "More than Moore". Cette stratégie repose sur l'identification des besoins spécifiques des applications avant de développer des solutions techniques adaptées, plutôt que de se focaliser uniquement sur la miniaturisation. Elle favorise la modification des architectures existantes et l'utilisation de nouveaux matériaux pour répondre aux exigences du marché.

Dans ce contexte, la recherche de nouvelles solutions de mémoire, outre les solutions traditionnelles proposé par la miniaturisation, devient cruciale pour accompagner cet évolution de paradigme. Les mémoires non volatiles émergentes, qui seront abordées dans la section suivante, représentent une solution prometteuse grâce à leur

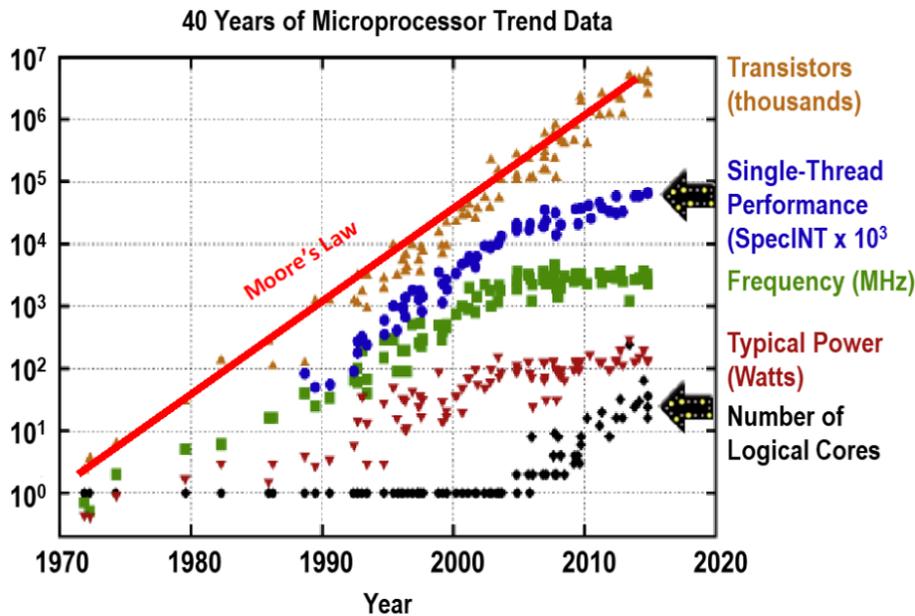


FIGURE 2.4 – Graphe montrant : la tendance à la miniaturisation des transistor, les performances à un seul thread, le nombre de coeur logique et la puissance consommé par transistor [2].

faible temps d'accès, leur faible consommation d'énergie et leur grande endurance.

Ces technologies contribueront également à réduire l'écart de performance entre la mémoire et le processeur, tout en favorisant le développement d'une nouvelle architecture mémoire, appelée "architecture Non-Von Neumann", qui intègre les unités de traitement et de stockage dans un même espace physique. Plus loin dans ce chapitre, nous expliquerons en détail en quoi consiste cette nouvelle architecture et quelles seront ses conséquences sur les performances des systèmes informatiques.

2.2 Mémoires non-volatiles émergentes

Au cours des deux dernières décennies, de nouvelles technologies de mémoire non volatile ont été développées pour surmonter les limitations des mémoires traditionnelles en termes de temps d'accès, d'endurance et de consommation d'énergie. Pour rappel, les mémoires volatiles perdent leurs données lorsque l'alimentation électrique est coupée, tandis que les mémoires non volatiles conservent les informations même en l'absence de courant. Les principales technologies de mémoires non volatiles émergentes sont :

- Mémoire ferroélectrique RAM (Random Access Memory).
- Mémoire à magnétique RAM
- Mémoire à changement de phase RAM
- Mémoire résistives RAM dont on peut distingués :

- Mémoire CbRAM : Conductive Bridge RAM
- Mémoire OxRAM : Oxide based RAM

Présentant des temps d'accès plus rapides, une capacité de stockage plus importante, une meilleure endurance et une meilleure efficacité énergétique, les technologies NVM émergentes présentent une solution viable pour surmonter les inconvénients des technologies de mémoire conventionnelles.

Dans la suite de cette section, un aperçu sera donné sur ces différents types de mémoires.

2.2.1 Mémoires ferroélectriques FeRAM

La FeRAM (Ferroelectric Random Access Memory) est un type de mémoire non volatile utilisant la polarisation réversible d'un matériau ferroélectrique. Comme le montre la figure 2.5, un matériau ferroélectrique est placé entre deux électrodes métalliques, créant ainsi une capacité. En appliquant un champ électrique coercitif ($+E_c$ et $-E_c$ sur la figure 2.5, la polarisation du matériau ferroélectrique peut être inversé car les différents dipôles électriques du matériau vont s'aligner avec le champ électrique externe. Ce processus est appelé commutation des domaines et cela se traduit par la courbe en forme d'hystérésis P-E, illustré sur la figure 2.5.

Après la commutation des domaines, l'amplitude de la polarisation à champ nulle est appelée polarisation rémanente (Pr), conférant à la mémoire son caractère non volatile. Le "0" et le "1" logique sont donc représentés par les deux états de polarisation rémanente $+Pr$ et $-Pr$. Pour résumer, l'écriture de donnée se fait en appliquant un champ $+E_c$ ou $-E_c$ à travers la couche ferroélectrique pour forcer les atomes à s'aligner dans une orientation "+" ou "-" [32]. Après avoir stocké l'information, il est crucial de pouvoir la récupérer. Dans les mémoires FeRAM, la lecture des données nécessite l'application d'un champ électrique positif, quelle que soit la polarisation initiale. Si la polarisation de la cellule FeRAM est initialement positive, la polarisation n'est pas perturbée, et ne change pas de signe, mais produit tout de même un courant j_0 . Si la polarisation de la cellule FeRAM est initialement négative, alors celle-ci va s'inverser, générant ainsi un courant j_1 supérieur à j_0 . Un comparateur analyse ces courants pour distinguer les deux états de polarisation et lire l'information [33].

Après la lecture, l'état de la cellule FeRAM doit être restauré à son état d'origine si la polarisation a été inversée pendant la lecture. Cette lecture est destructrice, car elle entraîne la perte de l'information stockée et nécessite une réécriture. Cette caractéristique constitue un défi majeur pour la fiabilité des mémoires FeRAM, car elle augmente le nombre de cycles d'écriture/lecture. On peut aussi citer une autre variante de mémoire ferroélectrique : le FeFET. Il a la même structure que celle d'une mémoire Flash sauf qu'à la place de l'oxyde de grille il y a un matériau fer-

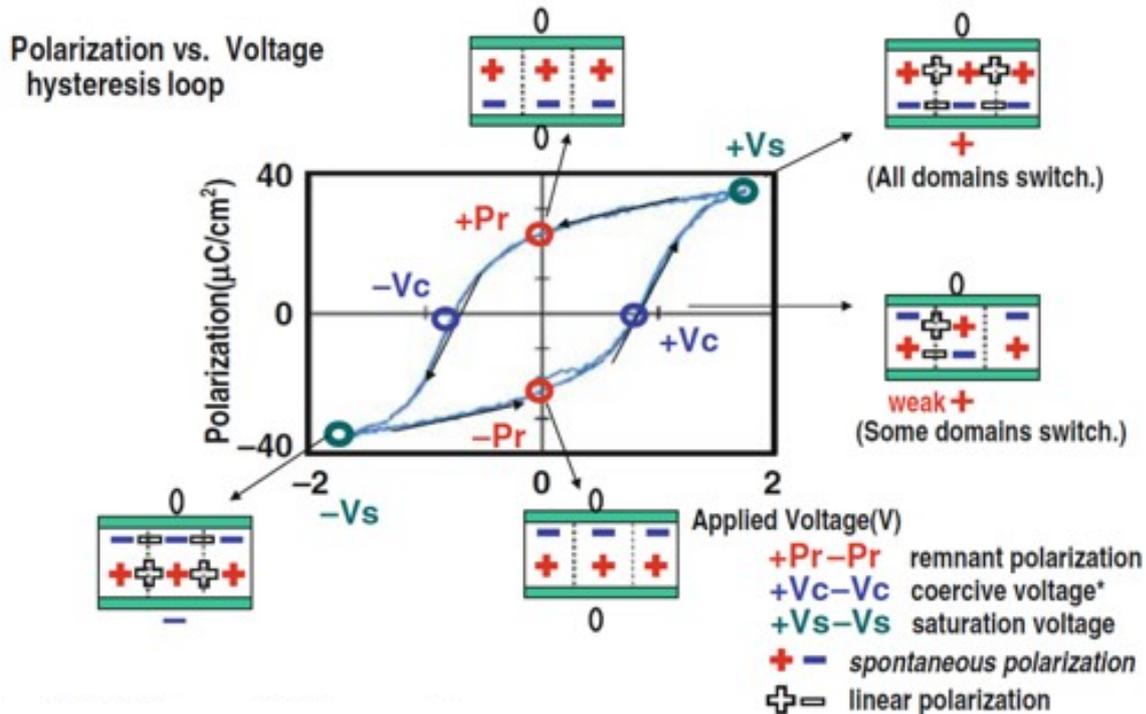


FIGURE 2.5 – Principe de fonctionnement d’une FeRAM [3]

roélectrique comme le montre la figure 2.6. La couche ferroélectrique présente une polarisation spontanée, qui peut être inversée en appliquant une tension. Cette polarisation reste même après que la tension est retirée, conférant ainsi un caractère non-volatile à ce composant. De nombreux matériaux ferroélectriques ont été étu-

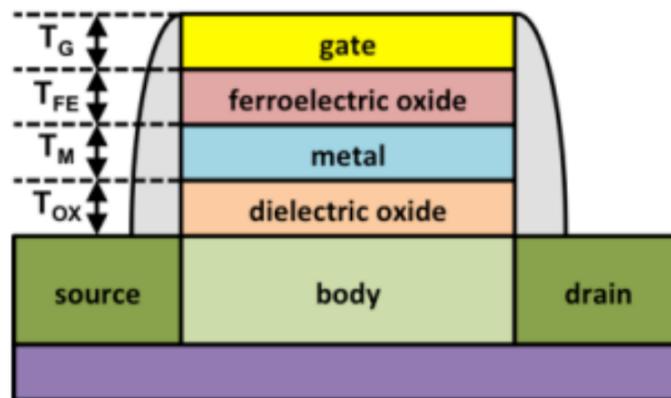


FIGURE 2.6 – Structure FeFET[4]

diés et, à l’heure actuelle, les trois matériaux suivants sont connus pour être les plus importants pour la fabrication des FeRAM : PZT, SBT et $(Bi, La)_4Ti_3O_{12}$ (BLT), mais aussi le HfO_2 [34]. Concernant les électrodes des FeRAM, on utilise généralement des métaux nobles tels que Pt et Ir ou des oxydes conducteurs tels que IrO_2 et $SrRuO_3$.

2.2.2 Mémoires Magnétiques MRAM

La STT-MRAM (Spin-Transfer Torque Magnetic Random Access Memory) est un type de mémoire non volatile exploitant le moment magnétique pour le stockage des données. Cette technologie repose sur des jonctions tunnel magnétiques (MTJ), composées de deux couches ferromagnétiques, la couche de référence (Reference Layer, RL) et la couche libre (Free Layer, FL), séparées par une couche diélectrique, souvent du MgO, comme le montre la figure 2.7 (a).

La couche libre (FL) peut voir son moment magnétique modifié pour stocker des données, tandis que la couche de référence (RL) conserve un moment magnétique stable, servant de référence pour l'orientation magnétique [35]. Quand un courant électrique traverse perpendiculairement la MTJ, les électrons qui passent exercent un couple de transfert de spin (spin-transfer torque) sur l'aimantation de la FL, suffisamment fort pour inverser son orientation magnétique. La direction de l'inversion du moment magnétique de la FL dépend du sens du courant appliqué. Un courant positif, allant de la RL vers la FL, permet aux électrons ayant un spin identique à ceux de la RL de traverser la jonction, transférant ainsi leur spin à la FL et alignant les moments magnétiques de manière parallèle. À l'inverse, un courant négatif fait passer les électrons de la FL vers la RL. À l'interface isolant-RL, les électrons avec un spin différent ont une faible probabilité de passage, ce qui transfère leur spin à la FL et conduit à un alignement antiparallèle [36]-[37].

Ces deux états d'alignement (parallèle et antiparallèle) se traduisent par deux états de résistances distincts : LRS (noté R_A dans la figure 2.7 (b)) pour l'alignement parallèle et HRS (noté R_{AP} dans la figure 2.7 (b)) pour l'alignement antiparallèle. Cette variation de résistance permet de lire l'état des données stockées dans la STT-MRAM.

Les matériaux les plus utilisés pour la STT-MRAM sont :

- Film de MgO : Utilisé comme barrière tunnel pour la jonction tunnel magnétique (MTJ).
- Des alliages comme Cobalt-Fer (CoFe) et le Cobalt-Nickel-Fer (CoNiFe) sont couramment employés comme matériaux ferromagnétiques dans les MRAM en raison de leurs excellentes propriétés magnétiques, notamment une polarisation de spin élevée et un faible amortissement.

2.2.3 Mémoires à changement de phase PCM

Les mémoires à changement de phase (PCM), qui sont des types de mémoires non-volatiles, se servent des propriétés électriques de matériaux présentant une différence de résistivité importante entre leur phase cristalline et leur phase amorphe [38]. En effet, la phase amorphe des matériaux à changement de phase est caractérisée par une forte résistivité électrique, tandis que la phase cristalline présente une faible

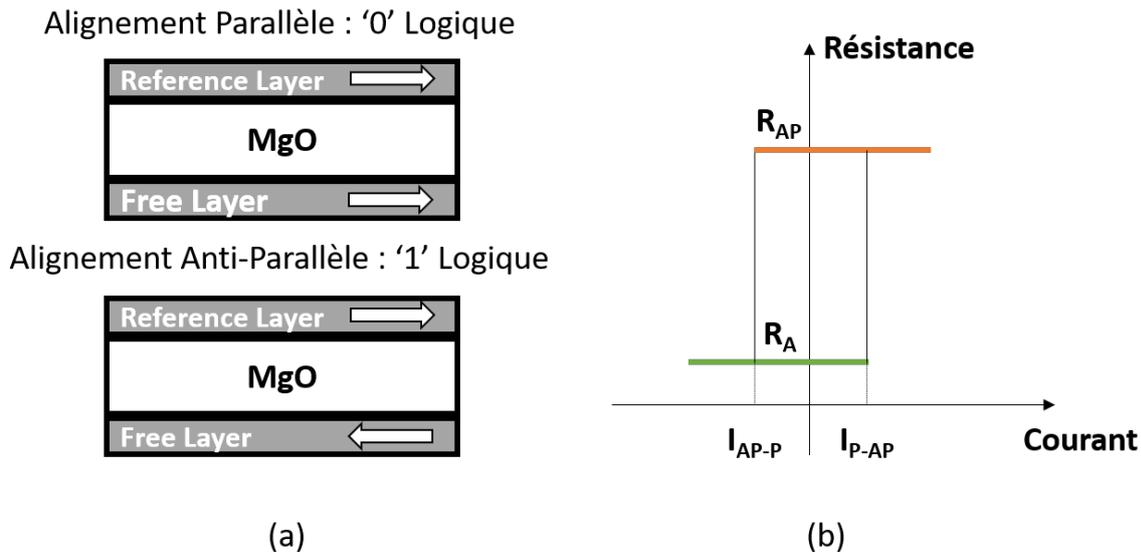


FIGURE 2.7 – (a) Schéma d'une cellule MRAM (b) Caractéristique Résistance en fonction du courant d'une MRAM

résistivité électrique. Cette différence de résistance peut être utilisée pour concevoir des dispositifs de mémoire [38].

Les matériaux couramment utilisés dans les PCM sont les chalcogénures, tels que le GST, $Ge_2Sb_2Te_5$ [39]. Ce sont généralement des alliages de matériaux de la colonne VI du tableau périodique on pourra citer à titre d'exemple : $AgSbSe_2$, $GeSeTe_2$ ou $GeTe$ [40].

La Figure 2.8 (a) illustre la structure en coupe d'une cellule de mémoire à changement de phase (PCM) en forme de T. Cette structure est principalement composée d'un matériau à changement de phase (généralement du GST) situé entre deux électrodes, à savoir une électrode supérieure et une électrode inférieure disposées verticalement. L'électrode inférieure sert d'élément chauffant qui fournit l'énergie thermique nécessaire pour le changement de phase. La zone active du matériau se situe au-dessus de l'élément chauffant, c'est là que le changement de phase se produit réellement et où la température est la plus élevée en raison de la concentration de courant dans cette région. [38].

La Figure 2.8 (b) présente la caractéristique courant-tension typique d'une cellule de mémoire à changement de phase (PCM). La transition de la phase amorphe à la phase cristalline est appelée opération de SET, et elle est réalisée en chauffant la région active au-dessus de sa température de cristallisation à l'aide de pulses en tension de courant moyen et de longue durée. En revanche, l'opération de RESET, qui correspond au passage de la phase cristalline à la phase amorphe, est obtenue en faisant fondre le matériau et en le refroidissant rapidement pour le ramener à la phase amorphe et désordonnée, à l'aide de pulses en tension de fort courant. Pour la

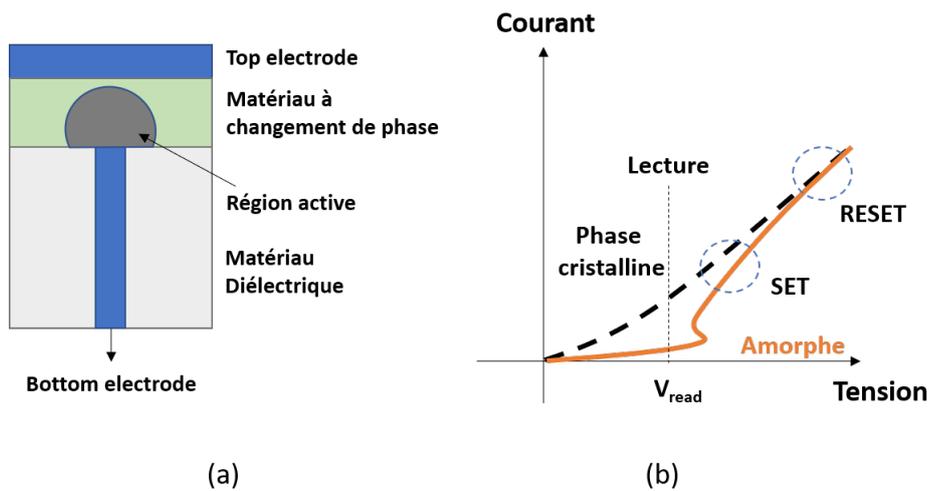


FIGURE 2.8 – (a) Schéma d'une cellule PCM (b) Caractéristique Courant en fonction de la Tension d'une PCM

lecture de l'information stockée, des pulses en tension de faible courant et de courte durée sont utilisées afin de minimiser la perte d'information [39].

2.3 Mémoires résistives ReRAM

Dans cette section, nous allons examiner le principe de fonctionnement des différentes technologies de mémoires résistives (ReRAM), ainsi que certaines de leurs caractéristiques telles que l'endurance et la variabilité.

2.3.1 Principe de fonctionnement d'une ReRAM

Les dispositifs ReRAM reposent sur le basculement réversible et reproductible, entre deux états de résistance d'un empilement MIM (Métal inférieur/Isolant/Métal supérieur), comme le montre la figure 2.9 (a). Ce phénomène de commutation de résistance a été observé pour la première fois dans les années 1960 sur des matériaux tels que le NiO , ZrO_2 et Ta_2O_5 . Cependant, ces recherches sont restées principalement académiques à l'époque en raison de performances défavorables, telles qu'une faible endurance et une tension de fonctionnement élevée [41, 42, 43, 44].

Entre 2005 et 2015, la recherche et le développement des ReRAM ont connu une période de forte croissance. Des avancées significatives ont été réalisées, notamment : la première intégration de ReRAM à pont conducteur (CBRAM) [45] a été faite, le premier dispositif HfO_2/Ti utilisant des matériaux de fabrication conventionnels [46], la première puce ReRAM vertical intégrée en 3D [47], la première puce intégrée de ReRAM CBRAM de 16 Gb [48], ainsi que le premier produit ReRAM intégré en technologie 40 nm. [49]. Ces avancées ont été rendues possibles grâce aux progrès en

science des matériaux, qui ont permis d'améliorer les performances des ReRAM et ainsi d'explorer le potentiel de ces mémoires dans différents domaines notamment l'IA et l'IoT.

Les dispositifs ReRAM utilisent un processus électrochimique pour permettre le changement de résistance. Ce processus implique la migration de cations métalliques ou d'anions d'oxygène, ce qui conduit à la création ou à la destruction d'un filament conducteur [50]. Ainsi, en appliquant un pulse électrique, la résistivité de la couche d'oxyde dite "active", qui correspond à l'isolant dans la structure MIM, peut être contrôlée. Cette couche "isolante" est capable de commuter entre deux états de résistance à de nombreuses reprises sans subir un claquage diélectrique irréversible. Chaque état de résistance est conservé même lorsque le dispositif n'est pas alimenté, ce qui confère à la mémoire un caractère non-volatile.

Trois états de résistance distincts sont considérés :

- Un état de résistance initiale très élevée (R_{init}), qui correspond à la résistance de l'empilement mémoire après sa fabrication. Il s'agit de l'état vierge ou "pré-électroforming", où le filament conducteur n'est pas formé (montré figure 2.9 (a)).
- Un état de résistance faible (R_{LRS} "Low Resistance State") qui intervient après une opération d'initialisation appelée "électroforming" ou après une opération d'écriture (ou SET) du dispositif, où le filament conducteur est complètement formé (montré figure 2.9 (b)).
- Un état de résistance élevée (R_{HRS} "High Resistance State") atteint après une opération d'effacement (ou RESET) du dispositif, où le filament conducteur est partiellement rompue (montré figure 2.9 (c)).

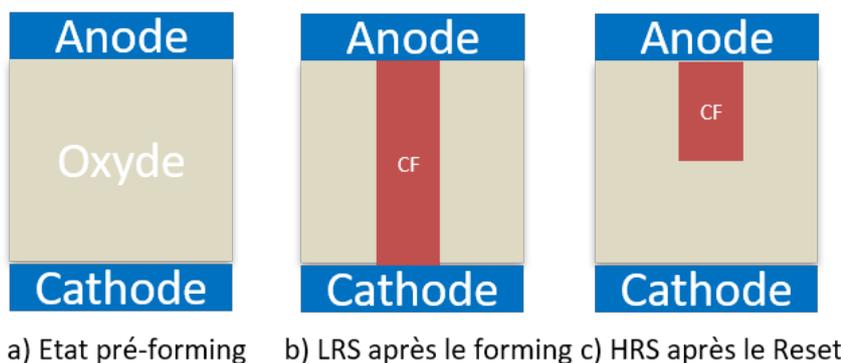


FIGURE 2.9 – Schéma des différents états de résistance d'une ReRAM

2.3.2 Classification des ReRAM

Les ReRAM (Resistive Random-Access Memory) peuvent être classifiées en fonction de différents critères. Voici deux critères de classifications couramment utilisés :

1. **La caractéristique Courant-Tension** : Afin de pouvoir distinguer les ReRAM en fonction de leur caractéristique I-V, les ReRAM peuvent être identifiées comme bipolaires ou unipolaires. Dans un dispositif unipolaire, la commutation ne peut se produire que dans un seul sens du courant. Cela est illustré dans la figure 2.10 (a), où les tensions des opérations SET et RESET ont le même signe, mais des amplitudes différentes. En effet, la tension de SET est plus élevée que celle de RESET. En revanche, dans un dispositif bipolaire, la tension des opérations SET et RESET doivent être effectués dans des polarisations spécifiques de signes différents, comme indiqué dans la figure 2.10 (b). L'opération SET s'effectue avec une tension positive, tandis que l'opération RESET s'effectue avec une tension négative.

Il convient de noter que le choix du matériau utilisé comme diélectrique n'est pas la source du caractère bipolaire ou unipolaire d'un dispositif. En effet, certains matériaux utilisés pour les dispositifs unipolaires peuvent également être utilisés dans des dispositifs bipolaires. Par exemple, le HfO_2 peut être utilisé pour fabriquer à la fois des dispositifs bipolaires et unipolaires [51]. La source de ce phénomène réside dans les électrodes utilisées. Si un métal noble tel que le platine (Pt) est utilisé comme électrode supérieure et inférieure, la caractéristique I-V de ce dispositif sera unipolaire. En revanche, si des métaux oxydables tels que le titane (Ti) ou le nitrure de titane (TiN) sont utilisés comme électrodes, la caractéristique I-V de ce dispositif sera bipolaire [50].

Dans la continuité de ce manuscrit, tous les dispositifs étudiés seront des ReRAM bipolaires. En effet, les dispositifs unipolaires présentent des tensions de RESET trop élevées et une variabilité trop importante. C'est pourquoi la quasi-totalité des études se concentrent sur les dispositifs bipolaires [50].

2. **Le mécanisme de commutation** : Les ReRAM peuvent opérer selon divers mécanismes de commutation, tels que la migration d'ions métalliques ou la migration d'anions d'oxygène. Ce critère permet de distinguer deux types de ReRAM : les CbRAM (Conductive Bridge RAM) et les OxRAM (Oxide based RAM). Dans la suite de ce document, le fonctionnement de ces deux types de mémoires sera expliqué en détail.

2.3.3 Mémoire CbRAM

La CbRAM (conductive bridge RAM) est un type de ReRAM. La figure 2.9 (a) illustre la structure d'une CbRAM, qui est un dispositif MIM (Métal inférieur/Isolant/Métal supérieur). L'électrode supérieure est conçue pour être électrochimiquement active sous l'application d'une tension, tandis que l'électrode inférieure est conçue pour être inerte. L'isolant entre ces deux électrodes peut être soit

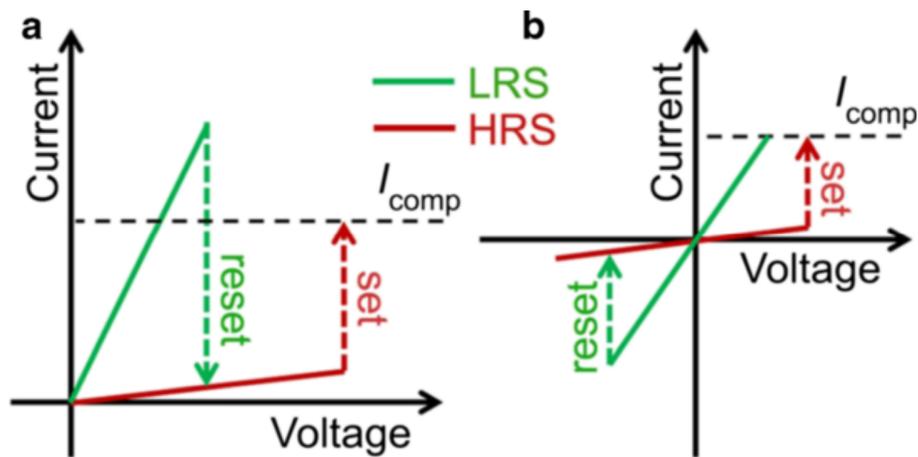


FIGURE 2.10 – Courbe caractéristique courant-tension d'un dispositif ReRAM : (a) Unipolaire (b) Bipolaire. Source : [5]

un oxyde métallique [52], soit un électrolyte solide [53]. Le principe de fonctionnement de ce dispositif repose sur l'électromigration d'ions métalliques de l'électrode supérieure électrochimiquement active à travers l'oxyde métallique, ce qui entraîne la formation d'un filament conducteur (CF) métallique. Ce processus de formation du CF (SET) peut être divisé en trois étapes successives : l'oxydation, la migration et la réduction [54, 55].

Au-delà d'une certaine tension, (V_{SET} sur la figure 2.11), des ions libres sont créés à l'interface de l'électrode électrochimiquement active, ce qui correspond à l'étape d'oxydation. Une fois que ces ions métalliques sont créés, sous l'effet du champ électrique, ils migrent à travers l'électrolyte solide en direction de l'électrode inférieure. Ce processus est rendu possible grâce au choix du matériau approprié pour l'électrode et la couche isolante.

Une fois que ces ions migrants en provenance de l'électrode supérieure atteignent l'électrode inférieure, ils sont réduits. Ainsi, le flux d'ions générés et migrant à travers l'électrolyte solide va permettre la formation d'un filament conducteur reliant les deux électrodes, ce qui entraîne une diminution significative de la résistance du dispositif. Cet état est appelé l'état de basse résistance (LRS, pour "Low Resistive State", comme le montre la courbe 2.11 I-V caractéristique de la CbRAM, dans la partie SET).

Pour le processus de RESET, un courant électrique de sens inverse à celui du SET est appliqué (V_{reset} sur la figure 2.11), ce qui entraîne la remigration partielle des ions métalliques du filament conducteur vers l'électrode supérieure. Cela conduit le dispositif à entrer dans un second état de résistance, appelé l'état de haute résistance (High Resistive State, HRS) [54], comme montré sur la figure 2.11, partie RESET.

Il est important de noter que les dispositifs CbRAM sont bipolaires.

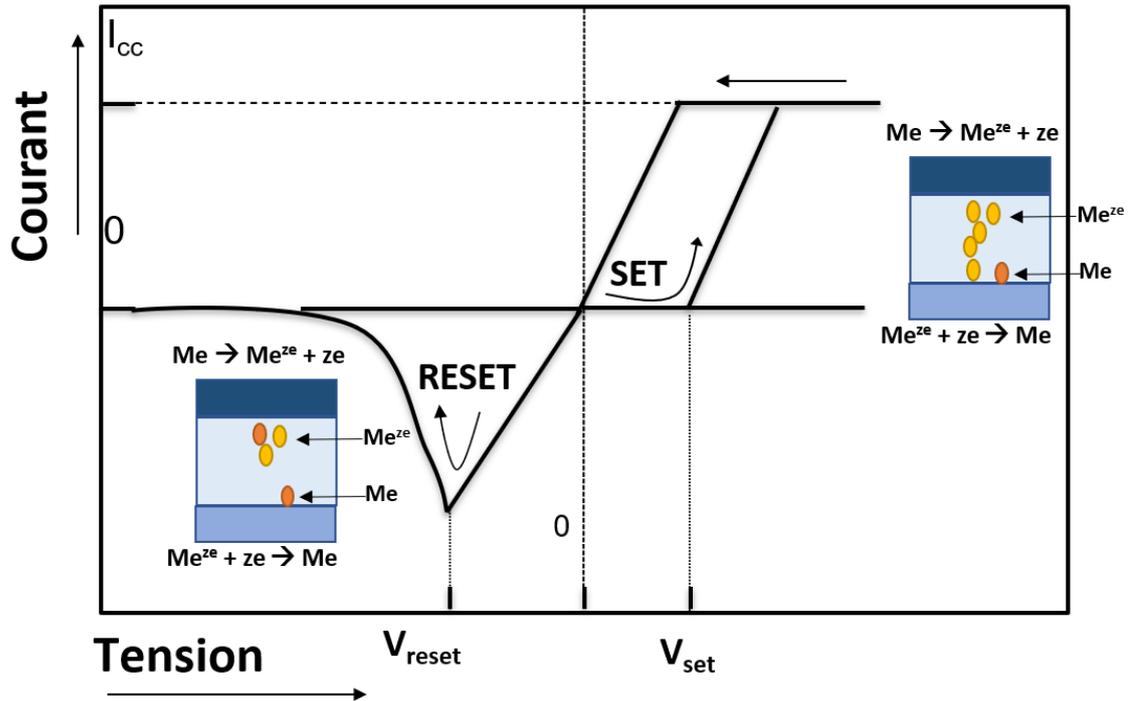


FIGURE 2.11 – Courbe I-V caractéristique d'une CbRAM.

2.3.4 Mémoire OxRAM

Principe de fonctionnement

L'OxRAM, ou "oxide based RAM", est un autre type de ReRAM, similaire en structure à la CbRAM. Son principe de fonctionnement est semblable, mais avec une différence clé : le filament conducteur est constitué de lacunes d'oxygène [56]. La Figure 2.12 illustre le mécanisme de commutation d'une OxRAM. On peut voir la structure d'un dispositif OxRAM, qui est composé d'une structure MIM (Métal-Isolant-Métal), avec une électrode inférieure en un matériau électrochimiquement inerte comme Pt ou Ru, et une électrode supérieure en un matériau électrochimiquement actif comme Ta, Ti, Cu ou Hf [57, 58]. La couche active quant à elle est constituée d'un oxyde de métal de transition comme le HfOx, TaOx ou AlOx [21, 59].

Dans son état vierge, l'OxRAM est hautement résistive, avec une valeur de résistance supérieure à celle qu'elle aura après des cycles de SET et RESET. Pour créer des lacunes d'oxygène et former un filament conducteur dans un dispositif à l'état vierge, une opération appelée "FORMATION" est nécessaire [6]. Pour cela, il faut appliquer une tension électrique, ce qui induit la création d'un champ électrique à l'intérieur de l'oxyde, ionisant ainsi certains atomes d'oxygène. Des ions d'oxygène avec une charge négative sont produits par oxydation. Les ions oxygène dérivent vers l'interface anode/oxyde par l'effet du champ électrique élevé, laissant donc des lacunes d'oxygène dans l'oxyde. Ces ions réagissent avec l'interface de l'anode et

forment une couche d'oxyde interfaciale. L'interface électrode/oxyde se comportera donc comme un réservoir d'oxygène. Le résultat de cette opération est que l'OxRAM se retrouve dans un état de basse résistance (LRS) [18]. Ce résultat est aussi reporté sur la courbe I-V caractéristique d'une OxRAM qui est représenté par la figure 2.10 (b). Pour une cellule mémoire en LRS, les électrons traversent le filament conducteur vers l'anode [60, 61]. Afin de prévenir les risques de défaillance majeure lors des opérations de FORMATION et de SET, il est nécessaire de limiter le courant circulant à travers le dispositif. Cette limitation peut être réalisée soit en utilisant un courant de compliance (noté I_{comp} dans la figure 2.10 (b)) fourni par l'analyseur lors des mesures sur les dispositifs 1R (1 résistance), soit en plaçant un transistor en série avec l'OxRAM pour restreindre le courant dans le dispositif. Cette approche est connue sous le nom de structure 1T1R (1 transistor - 1 résistance). Elle est utilisée afin d'éviter tout risque de détérioration de l'OxRAM. En effet, les pics de courants survenant lors du Forming sont principalement liées aux charges (et leur déchargent) "stocké" dans les capacités des câbles [62]. Il a été montré dans [62] que la résistance de post-forming ne peut être ajustée par la compliance sur le courant disponible sur les analyseurs de paramètres, puisque ces derniers sont inefficaces pour limiter le courant lors des premières microsecondes du pic de courant.

Pour commuter le dispositif et le ramener à son état de haute résistance, un RESET est effectué [6]. Une polarisation inverse, notée V_{reset} , est appliquée, ce qui provoque la ré-ionisation des atomes d'oxygène et leur recombinaison avec les lacunes d'oxygène du filament conducteur, le rompant ainsi partiellement. Cela ramène ainsi l'OxRAM à son état de haute résistance (HRS).

Toute opération de programmation ultérieure suit les mêmes étapes, bien qu'une tension plus basse, $V_{set} < V_{form}$, soit nécessaire pour amener le dispositif à son état de basse résistance (LRS).

Il est à noter que le fonctionnement décrit ci-haut est celui d'un dispositif OxRAM bipolaire.

2.3.5 État de l'art en matière de performances

Variabilité

Une mémoire OxRAM est sensé faire plusieurs cycle HRS/LRS sans pour cela détériorer la fenêtre résistible (écart de résistance entre l'état LRS et l'état HRS) entre ces deux états. Bien que les OxRAM démontrent des performances très intéressantes sur plusieurs figures de mérites (temps de rétention, endurance, faible de temps de latence), il a été démontré qu'elles présentent une forte variabilité en courant. Cette variabilité peut se manifester de deux manière soit d'un cycle à un autre, on parlera alors de variabilité C2C ou d'un dispositif à un autre, on parlera

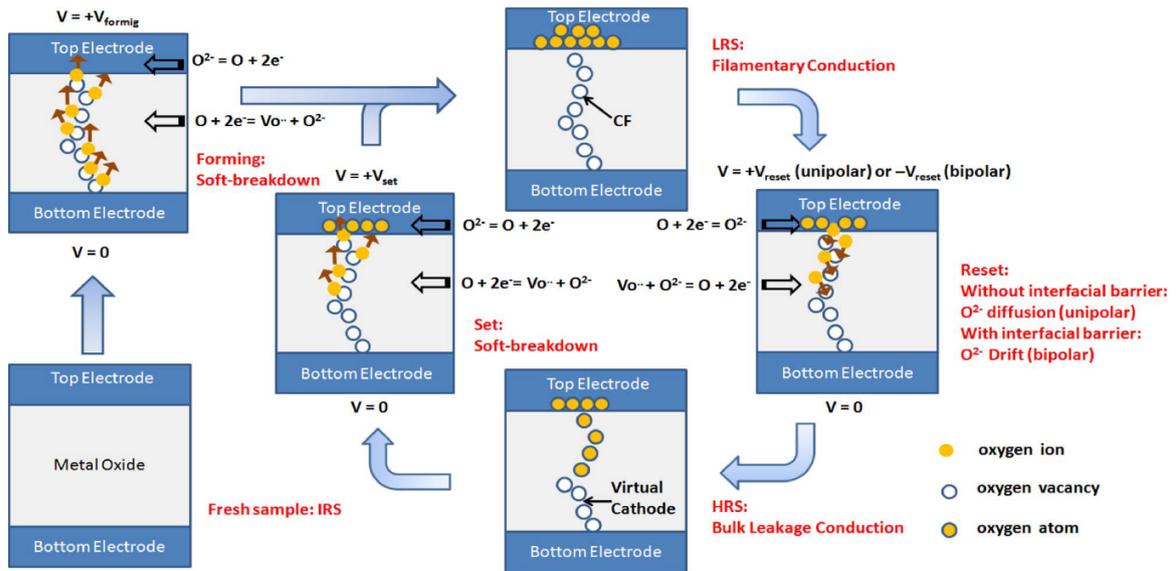


FIGURE 2.12 – Schéma représentant le mécanisme de commutation d'une OxRAM. Source : [6]

alors de variabilité D2D [63]. Ce phénomène est l'inconvénient majeur des OxRAM qui empêche l'industrialisation en masse de ce composant.

La variabilité en HRS est plus importante que celle en LRS. Divers études ont cherché à modéliser cette variabilité en HRS en introduisant notamment une variation de l'épaisseur de la barrière de tunnel entre les pièges [7]. Pour le LRS, il a été identifié que la variabilité géométrique de la forme du filament conducteur (CF), comme le rayon du CF, peut contribuer à cette variabilité. [6, 64].

Un modèle unifié permettant de modéliser la variabilité en HRS et en LRS a été proposé par D.Garbin et al.[7]. Il s'agit d'un modèle reposant sur un réseau de résistances. Il a été implémenté pour simuler l'effet de la variabilité sur les deux états de résistance. Sur la figure 2.13 l'écart type de la résistance en fonction de la résistance moyenne est représenté. Il est aisément remarquable que la courbe expérimentale est assez bien reproduite par les simulations.

Endurance

L'endurance se définit comme le nombre de cycles maximal d'écriture/lecture qu'un composant mémoire peut effectuer avant que sa fenêtre résistive ne se dégrade. Pour tenter de comprendre à quoi est due ce phénomène de détérioration de la fenêtre résistive, un modèle a été proposé par Robayo et *al.* [8]. Ce modèle suggère que le nombre de cycle d'endurance peut être expliqué par la formation d'état non commutable dans l'oxyde, avec une très faible probabilité en raison de l'énergie nécessaire à leur génération. De tels défauts peuvent être des lacunes d'oxygène l'OxRAM qui ne peuvent plus être éliminées une fois qu'elles sont créées. Cela implique

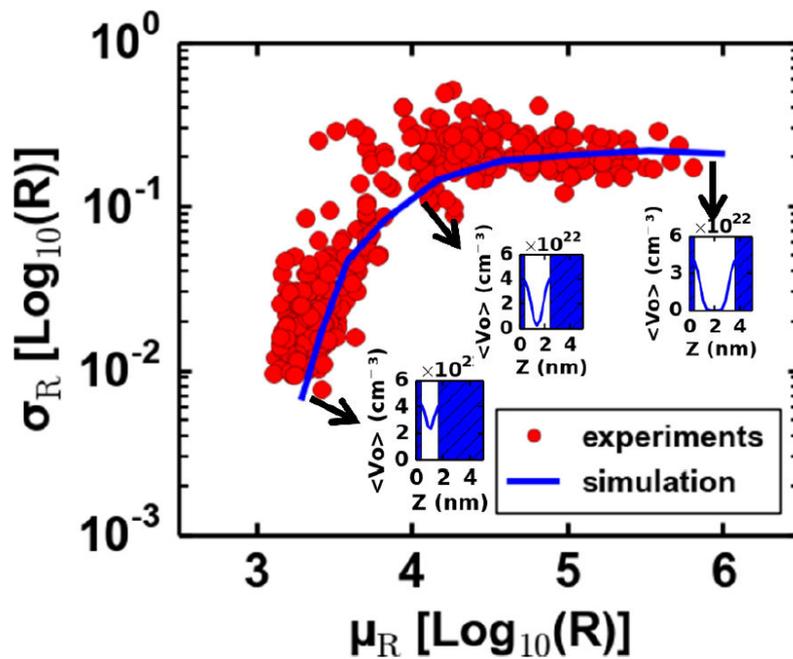


FIGURE 2.13 – Ecart type expérimental et simulé de la résistance en fonction de la résistance moyenne. Source : [7]

que deux filaments pourraient exister à l'intérieur de l'oxyde : l'un constitué d'états commutables et l'autre d'états non commutables ou que le filament est constitué de deux type de pièges.

La figure 2.14 représente l'écart type en résistance en fonction du nombre de cycle maximum. Un excellent accord entre mesures et simulations a été trouvé, permettant d'observer d'un point de vue phénoménologique comment se comportent les distributions de N_{cmax} (Nombre de cycle maximum).

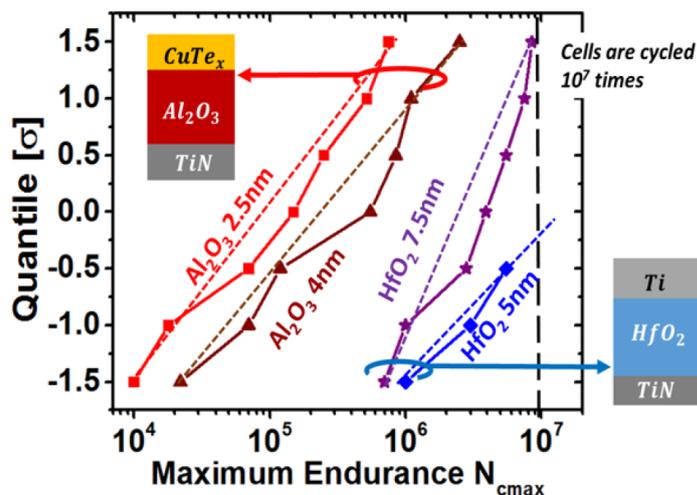


FIGURE 2.14 – Distributions N_{cmax} expérimentales et simulées. Source : [8]

Rétention de donnée

Les mémoires non volatiles sont connues pour leur capacité à conserver les données sur une longue période, et l'OxRAM en particulier a démontré une durée de rétention de dix ans à des températures allant jusqu'à 85°C [21]. Cependant, pour améliorer la rétention des données dans l'OxRAM, il est essentiel de comprendre les mécanismes qui sous-tendent la perte de données. Plusieurs hypothèses ont été formulées à cet égard, les deux plus importantes étant la diffusion des ions d'oxygène présents à l'interface du "scavenging layer" et leur recombinaison avec les lacunes d'oxygène présentes dans le filament conducteur, ce qui conduit également à la dissolution de ce dernier et à la disparition de l'état LRS, comme illustré dans la figure 2.15 (a). Le deuxième mécanisme proposé est la diffusion de lacunes d'oxygène hors du filament conducteur, ce qui dissout l'état LRS et le confond avec un état HRS, entraînant ainsi la perte de données, comme illustré dans la figure 2.15 (b). Pour remédier à cela, une solution pertinente proposée dans [9] consiste à effectuer un recuit après la fin du processus de fabrication, qui permet de réduire la diffusion des ions d'oxygène vers le filament conducteur. Des calculs *ab initio* ont également montré que l'incorporation d'une couche d'Al augmenterait significativement la rétention [65].

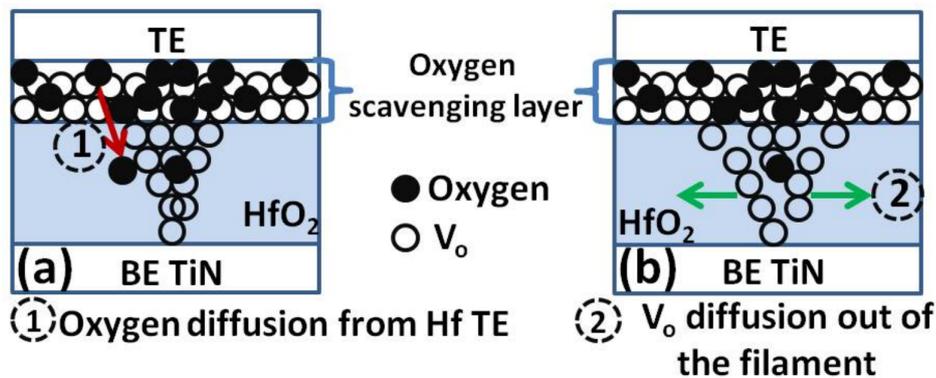


FIGURE 2.15 – Schéma représentatif du mécanisme de : (a) diffusion d'ions d'oxygène (b) diffusion de lacunes d'oxygène. Source : [9]

2.3.6 Comparaison des principales mémoires NVM émergentes

Les technologies de mémoires non volatiles émergentes offrent de nouvelles perspectives pour les dispositifs de stockage de données. Dans le tableau 2.1, est présenté une comparaison entre quelques-unes des principales technologies, à savoir la FeRAM, STT-MRAM, la PCM, la CBRAM et l'OxRAM ainsi que la NAND Flash.

La STT-MRAM a une endurance très grande et une bonne vitesse de lecture/écriture, ainsi qu'une faible consommation d'énergie et une rétention des don-

nées comparable aux autres NVM émergentes [66]. D'ailleurs plusieurs produits STT-MRAM ont été démontrés par Intel [67], GlobalFoundries [68], Samsung [69] et TSMC [70]. Cependant, sa capacité de miniaturisation est limitée en raison de la faible fenêtre résistive atteignable dans la MTJ. De plus, elle peut être sensible aux champs magnétiques externes, ce qui peut poser des défis pour certaines applications [71]. Le coût de fabrication du STT-MRAM peut également être potentiellement élevé [72].

La PCM quant à elle offre une grande capacité de miniaturisation, une endurance assez importante, une vitesse de lecture/écriture rapide et une faible consommation d'énergie [38]. Cependant, elle peut présenter une variabilité des performances en raison des changements de phase du matériau. Elle reste néanmoins la NVM émergente la plus mature avec des performances éprouvées [73]. Plusieurs produits sont déjà sur le marché, on pourra citer le P18 de STMicroelectronics [74] et l'Optane de Intel [75].

Le CBRAM et l'OxRAM ont un avantage majeur par rapport aux autres technologies : il s'agit de la simplicité de la structure de dispositif. En effet, c'est une structure MIM, utilisant des matériaux très utilisés dans l'industrie des semi-conducteurs et compatible avec les procédés CMOS [49]. Le CbRAM peut par contre être sensible aux conditions environnementales et présenter une instabilité des performances en raison de la formation de ponts conducteurs [53]. L'OxRAM, contrairement à la CbRAM a une grande endurance, une vitesse de lecture/écriture rapide, une faible consommation d'énergie [21]. En revanche, sa variabilité d'un composant à un autre (D2D) et d'un cycle à un autre (C2C) est un frein majeur à sa large commercialisation [21]. Toutefois, plusieurs produits ont été récemment mis sur le marché par Fujitsu [76] et Infineon [77]. Enfin la FeRAM présente la plus forte endurance de toutes les technologies de NVM émergentes. Elle a une faible consommation d'énergie et un temps de rétention comparable aux autres technologies [32]. Toutefois, elle peut avoir des limitations en termes de capacité de miniaturisation, car la taille des cellules de mémoire peut être relativement grande par rapport à d'autres technologies émergentes. Toutefois, plusieurs produits sont déjà commercialisés par : Fujitsu [78] ou encore Infineon [79].

Il est important de noter que chaque technologie a ses propres avantages et inconvénients, et que le choix de la technologie de mémoire non volatile émergente dépendra des exigences spécifiques de l'application, du coût de fabrication, de la densité de dispositif, de l'endurance, de la vitesse de lecture/écriture et d'autres facteurs pertinents. Il n'y a pas de solution unique qui conviendra à toutes les applications [28].

	FeRAM	PCM	STT-MRAM	NAND Flash	OxRAM	CbRAM
Taille dispositif (F)	15F ²	4F ²	6F ²	4F ²	6F ²	6F ²
Courant de programmation (A)	10 ⁻⁶	10 ⁻⁴	10 ⁻⁵	10 ⁻⁷	10 ⁻⁴	10 ⁻⁴
Temps d'écriture (ns)	<10 ns	100 ns	<10 ns	100 μ s-1 ms	<50 ns	<50 ns
Temps de lecture (ns)	<5 ns	<5 ns	<10 ns	10 μ s	<5 ns	<10 ns
Temps de rétention (années)	>10	>10	>10	>10	>10	>10
Endurance (nb cycles)	10 ¹¹	10 ⁷	10 ¹⁰	10 ⁵	10 ⁶	10 ⁶

TABLE 2.1 – Tableau récapitulatif des performances des différents types de mémoires non volatiles [80, 81, 82]

2.4 Applications des mémoires non volatiles émergentes

Les NVM émergentes sont susceptibles de réformer la hiérarchie de la mémoire, principalement en raison de sa vitesse élevée, de sa capacité de stockage non volatiles et de sa densité de stockage améliorée. Les diverses nouvelles applications de ces mémoires sont abordées dans cette section.

Cette partie est inspiré de l'article de review de G.Molas et *al.* [81].

2.4.1 Mémoire de Stockage

Les technologies émergentes comme les RRAMs pourraient potentiellement remplacer la technologie Flash. Cependant, leur adoption dépendra de leur compétitivité en termes de coût. Pour l'instant, la technologie Flash reste bien établie sur le marché et continue de bénéficier d'une densité accrue grâce à l'intégration 3D. En attendant que la RRAM devienne plus compétitive, ces nouvelles mémoires pourraient jouer un rôle de mémoire intermédiaire entre la mémoire principale et les solutions de stockage comme la Flash ou les disques durs. Par exemple, les SSD ont un temps de latence d'environ 100 μ s, tandis que la DRAM, utilisée comme mémoire principale, a un temps de latence d'environ 50 ns. La RRAM et la PCRAM, avec une latence

d'environ 1 μ s et une densité élevée, sont donc des candidats prometteurs pour ce rôle d'intermédiaire dit de Storage Class Memory.

2.4.2 Mémoire Principale

Les mémoires émergentes pourraient être utilisées en complément de la DRAM sur le même bus mémoire, voire la remplacer entièrement.

Elles offriraient un avantage significatif en améliorant la gestion de l'intégrité des données, en permettant une sécurisation rapide de celles-ci mot à mot. Cela simplifierait les processus actuels de journalisation et de pointage de contrôle, souvent complexes et fastidieux, utilisés pour protéger les données en cas de perte d'alimentation.

Ces nouvelles mémoires seraient également très utiles pour le Big Data. Si elles deviennent plus denses et moins coûteuses que la DRAM, elles offriraient une capacité de stockage accrue, indispensable pour les applications Big Data. Cela permettrait de garder de grands tableaux multidimensionnels en mémoire principale, apportant ainsi un avantage significatif en termes de performance par rapport aux mémoires traditionnelles.

Enfin, les mémoires comme la RRAM et la PCRAM, qui sont à la fois rapides (environ 50 nanosecondes) et très endurantes (jusqu'à un million de cycles), pourraient simplifier les opérations atomiques dans les bases de données transactionnelles. Elles offriraient une alternative plus efficace aux systèmes actuels qui combinent DRAM, Flash, et supercondensateurs pour assurer la fiabilité des données.

2.4.3 Cache de Dernière Niveau

Les technologies émergentes pourraient remplacer ou compléter les caches de dernier niveau, bien qu'elles ne soient probablement pas assez rapides pour les caches de premier niveau. Les exigences incluraient une vitesse inférieure à 30 ns et une endurance supérieure à 10^{16} cycles. La STT-MRAM est particulièrement adaptée en raison de sa haute endurance pour les caches.

Ainsi, les technologies de mémoire émergentes offrent des améliorations potentielles pour le stockage, la mémoire principale et les caches, chacune ayant ses propres avantages selon les exigences de performance et de coût.

2.4.4 In Memory Computation

L'architecture de Von Neumann a longtemps servi de base aux systèmes informatiques classiques, avec sa distinction claire entre les unités de mémoire et de traitement. Cependant, les mémoires non volatiles émergentes (NVM) encouragent le développement d'architectures alternatives, non-Von Neumann, qui sont capables

de surmonter ces limites. Cela est rendu possible grâce à une nouvelle approche de calcul connue sous le nom de "In Memory Computation" (IMC) [10]. Explorons ces applications en détail et voyons pourquoi les NVM sont particulièrement adaptées à ce type d'architecture.

L'architecture IMC (In-Memory Computing) est conçue pour intégrer les unités de traitement et le stockage au sein du même espace mémoire. Comme le montre la figure 2.16 (a), ce système repose sur un cœur IMC qui est constitué d'un réseau très dense de mémoires non volatiles (NVM) émergentes, associé à des circuits périphériques CMOS.

Cette puce IMC est efficace pour gérer les tâches de calcul intensives en mémoire qui sont généralement effectuées par le processeur. En procédant ainsi, elle accélère significativement le traitement tout en réduisant la consommation d'énergie, car elle limite les transferts de données massifs. La figure 2.16 (b) montre un exemple de programme qui peut être exécuté efficacement grâce à cette architecture. Dans cet exemple, plusieurs boucles de calcul peuvent être traitées directement dans le cœur IMC, tandis que le reste du programme continue de s'exécuter sur un cœur conventionnel.

À chaque fois qu'une boucle commence, le processeur envoie une instruction au cœur IMC. Celui-ci décode l'instruction, l'exécute localement, puis renvoie les résultats finaux, comme expliqué dans [10].

Les puces IMC (In-Memory Computing) bénéficient grandement de l'utilisation des mémoires non volatiles (NVM) émergentes pour plusieurs raisons. Tout d'abord, ces mémoires permettent de réduire la latence et la consommation d'énergie. Leur faible temps de latence et leur nature non volatile évitent les transferts de données fréquents entre le processeur et la mémoire, ce qui réduit le mouvement des données et diminue ainsi la consommation énergétique.

De plus, les NVM émergentes prolongent la longévité des systèmes IMC grâce à leur endurance élevée. Les systèmes IMC nécessitent de nombreuses mises à jour des données, car les calculs s'effectuent directement dans la mémoire, entraînant de fréquents cycles de lecture et d'écriture. Grâce à la robustesse des NVM, ces systèmes peuvent ainsi supporter ces opérations répétitives sur une plus longue période, améliorant ainsi leur endurance.

Par la suite, nous aborderons uniquement les OxRAM ; aucune autre technologie de NVM émergentes ne sera discutée dans ce manuscrit.

2.5 État de l'art de la modélisation de la conduction

Pour approfondir notre compréhension du fonctionnement de l'OxRAM, en particulier en ce qui concerne sa variabilité, il est essentiel d'analyser de près les trois mécanismes clés : forming, commutation et conduction. Dans cette thèse, nous nous

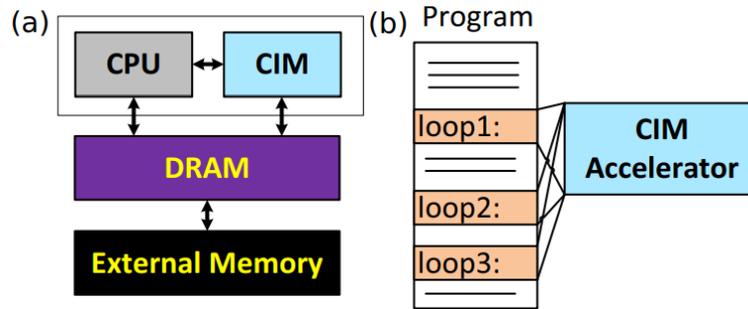


FIGURE 2.16 – Schéma représentatif (a) D'une architecture IMC (b) Exemple de programme à exécuter avec un système IMC. Source : [10]

concentrerons uniquement sur le mécanisme de conduction, et cette section en particulier examinera les modèles les plus pertinents de la littérature scientifique qui portent sur ce mécanisme dans les dispositifs OxRAM. Tout d'abord, nous parlerons de la distinction qui est faite entre le HRS et le LRS, en terme de modèle de conduction. Ensuite, nous examinerons le modèle le plus couramment employé pour les OxRam, à savoir le Trap Assisted Tunneling (TAT), dans ses deux variantes : celle de Larcher et *al.* ainsi que celle de Ielmini et *al.* Par la suite, nous aborderons deux autres modèles répertoriés dans la littérature, à savoir le modèle de Quantum Point Contact (QPC) et le modèle Poole-Frenkel.

2.5.1 Distinction HRS/LRS

Dans la littérature sur la conduction dans les dispositifs OxRAM, il est fréquent de constater que les états de HRS et LRS ne sont pas traités de la même manière à l'aide des mêmes modèles. En effet, le LRS est souvent considéré comme un état métallique en raison de la densité élevée de pièges présents dans cet état de résistance, comme le démontre la figure 2.17. En effet, il est rapporté un nombre de piège de plus 100 pièges sur une épaisseur d'oxyde de 5 nm [11].

Ainsi, le LRS est généralement modélisé par un modèle ohmique. La figure 2.18 illustre la caractéristique courant-tension en comparant les données expérimentales avec la simulation réalisée à l'aide du modèle ohmique. On constate un bon accord entre les données expérimentales et les simulations du modèle ohmique, comme indiqué dans l'étude de Vandelli et *al.* [11]. Cette approche est d'ailleurs partagé par de nombreuses équipes de recherche comme rapporté dans [83, 84, 6]. Le HRS est quant à lui traité par une multitudes de modèles dont certains vont être exposé dans la suite de ce chapitre. Un autre type de modèle est aussi répertorié, il s'agit de modèles qui ne font pas de distinction entre le HRS et le LRS, et dans ces cas-là, un seul modèle est utilisé pour représenter les deux états de résistance. C'est notamment le cas du modèle QPC proposé par Andrade et *al.* [15], qui sera exposé dans la suite

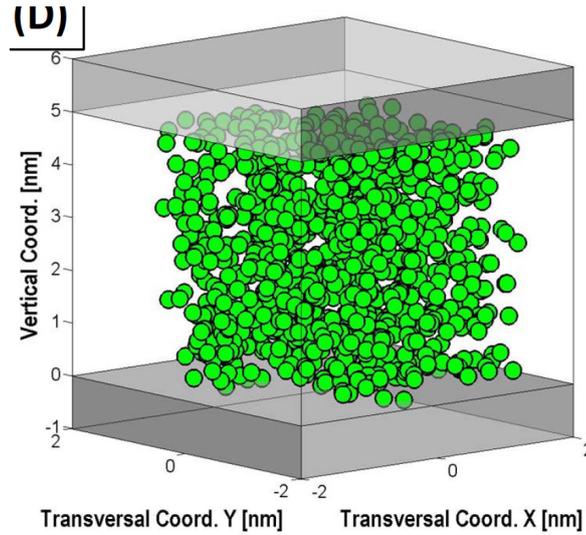


FIGURE 2.17 – Schéma représentatif de la densité de piège importante en LRS dans une OxRAM. Source : [11]

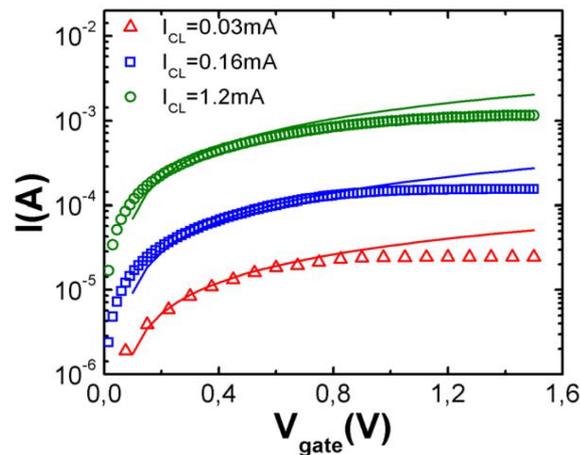


FIGURE 2.18 – Courbes I-V comparatives entre donnée expérimentales et simulation du modèle ohmique en LRS

Source : [11]

de ce chapitre.

2.5.2 TAT Larcher et al.

Les oxydes étant rarement des oxydes parfaits, ils présentent en général des défauts dans leur volume. Ces défauts peuvent donner lieu à une conduction dite TAT pour Trap Assisted Tunneling. Un exemple de ce mécanisme est schématisé sur la figure 2.19 [85]. Un premier processus tunnel consiste en la capture de l'électron par le piège et le second consiste en l'émission de cet électron vers la bande de conduction du métal. Ainsi, il s'agit du passage d'un électron en provenance d'une bande

de conduction, donc un continuum d'état vers un état lié (le piège) puis enfin vers un continuum d'états (bande de conduction). Ce mécanisme peut se faire sur un ou plusieurs pièges. Le mécanisme impliquant plusieurs pièges porte le nom MultiTAT. Ce modèle est utilisé par l'université de Modena pour simuler le transport électronique dans l'oxyde [83, 12]. En effet, dans ce mécanisme, la probabilité de tunnel des porteurs dépend, au premier ordre, de la barrière énergétique qu'ils doivent traverser. La présence de pièges à l'intérieur de l'oxyde divise donc la barrière d'énergie en deux parties, permettant ainsi le tunnel consécutif à travers des barrières d'énergie plus minces et augmentant de manière exponentielle la probabilité du processus de tunnel total. Pour calculer le courant, les formules de temps caractéristique de

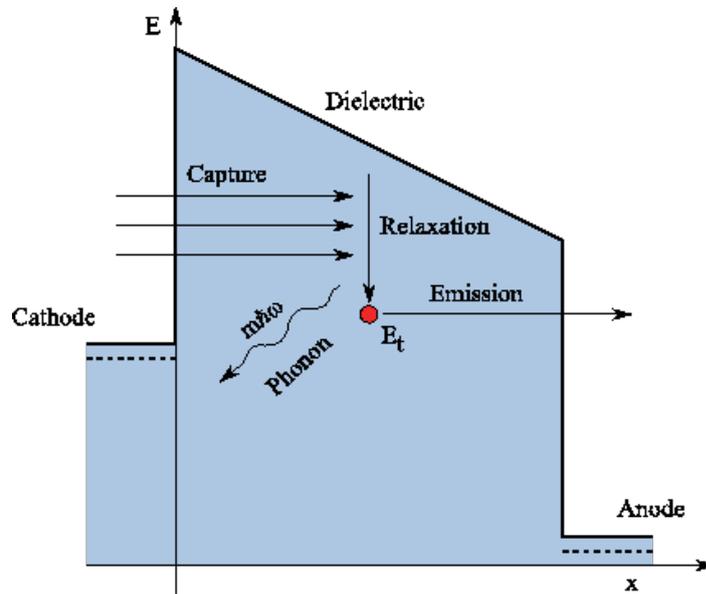


FIGURE 2.19 – Schéma récapitulatif du processus TAT
Source : [11]

capture et d'émission donné dans [85] sont utilisés.

$$\tau_{c,j,m} = N_{j-1}(E_{j,m}) f_{j-1}(E_{j,m}) C_{j,m} P_T \quad (2.1)$$

$$\tau_{e,j,n} = N_{j+1}(E_{j,n}) (1 - f_{j+1}(E_{j,n})) E_m m_{j,n} P_T \quad (2.2)$$

Dans les équations 2.1 et 2.2, N représente la densité d'état pour chaque piège, P_T représente la transparence d'une barrière trapézoïdale utilisé dans [85], les coefficients C et E_m sont respectivement les coefficients de capture et d'émission qui représentent le couplage entre les phonons et l'électron, f représente la fonction de Fermi-Dirac. Dans le cas, de la capture (équation 2.1), la fonction de Fermi-Dirac est utilisée pour évaluer si un électron est présent dans le piège $j-1$, et donc si le processus tunnel est possible. Pour l'émission on utilise la fonction de Fermi-Dirac pour estimer si le piège $j+1$ est libre, pour la même raison. Ainsi pour implémenter ce modèle les hypothèses suivantes ont été faite :

- 1^{ère} hypothèse : on considère que le métal est un réservoir à électron à l'équilibre thermodynamique, et donc que l'émission du piège vers l'anode ne peut être limitée par le métal et la capture du piège depuis la cathode ne peut être limitée par le métal. Ceci veut dire que le temps de capture et d'émission à partir du métal est proche de zéro, ceci est montré en figure 2.20.
- 2^{ème} hypothèse : dans le cas du tunnel on considère que si le piège i a une énergie supérieure à celle du piège j, le tunnel se produit à l'énergie du piège i. Ceci veut dire que l'électron tunnel tout le temps à l'énergie du piège la plus haute comme le montre la figure 2.21
- 3^{ème} hypothèse : on considère que les phonons qui interviennent dans le processus tunnel sont les phonons d'ordre n avec une énergie $n\hbar\omega_0$. Ainsi, dans le calcul des temps caractéristique (équation 2.1 et 2.2), on doit faire la somme sur l'ensemble des phonons impliqué dans ce processus.
- 4^{ème} hypothèse : la densité d'état des pièges est calculée en faisant l'approximation que celle-ci est la densité d'état d'un semi-conducteur 3D.

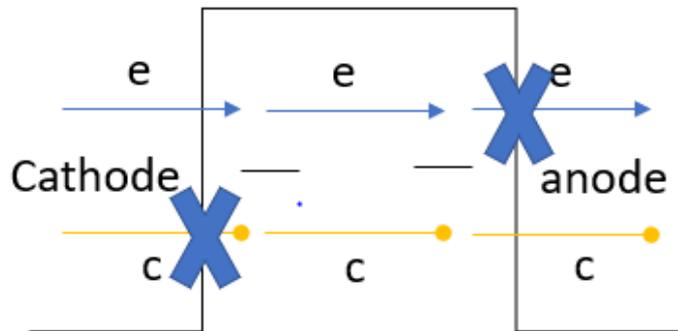


FIGURE 2.20 – La capture depuis cathode et l'émission vers l'anode ne sont limitantes

Source ; [12]

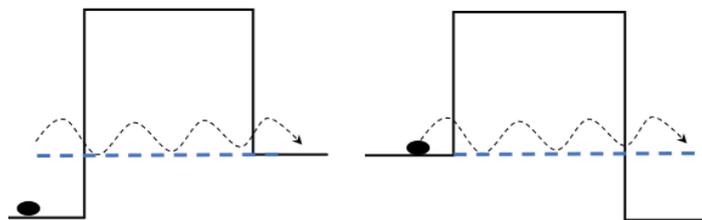


FIGURE 2.21 – Schéma représentant l'hypothèse du calcul de la transparence

Une fois les deux temps caractéristiques calculés, on retient le plus long des deux, c'est-à-dire celui qui limite le transport, en prenant le maximum des temps caractéristiques. À la fin de cette étape, il ne reste donc qu'une seule valeur de temps

caractéristique. Ensuite, il est nécessaire de déterminer les chemins de percolation qui permettront de calculer le courant. A cet effet, une procédure a été développée dans [12] pour calculer automatiquement le courant de fuite entraîné par chaque piège. Les contributions dues aux chemins de percolation uniques sont additionnées pour calculer le courant total. L'algorithme est décrit puis schématisé ci-dessous (figure 2.22) :

1. Le temps caractéristique $\tau_{j,anode}$ pour chaque piège est calculé en supposant que les électrons arrivent directement à l'anode sans passer par d'autres pièges ;
2. $\tau_{j,anode}$ est ensuite comparé aux temps caractéristiques $\tau_{j,k}$ calculés entre le piège j et les autres pièges k calculés en supposant que les électrons se déplacent d'un piège à un autre piège ;
3. s'il existe un $\tau_{j,k}$ inférieur à $\tau_{j,anode}$, un chemin de percolation comprenant le piège j et le piège k est formé, et la procédure revient au point 1) en considérant le piège k
4. On considère que le chemin de percolation est complet quand le dernier minimum est $\tau_{j,anode}$ (l'électron est arrivé à l'anode).

De cette façon, les contributions dues à chaque chemin de percolation de piège sont automatiquement calculées. Une fois les chemins de percolations construits, le courant peut facilement être calculer en prenant le temps maximal de chaque chemin de percolation et en utilisant la formule 2.3 donné dans [12] :

$$I = \frac{q}{\max(\tau)} \quad (2.3)$$

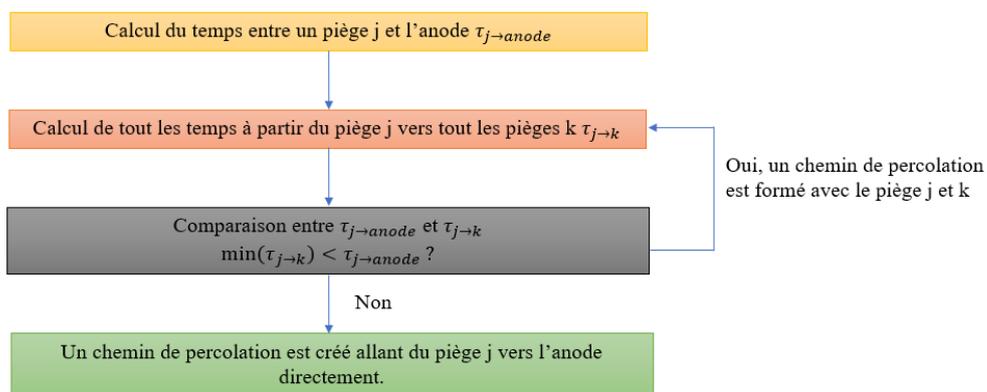


FIGURE 2.22 – Diagramme fonctionnelle de l'algorithme de calcul des chemins de percolation

Grâce à ce modèle, L. Larcher et *al.* parviennent à reproduire les courbes courant en fonction de la tension et de la température, comme illustré dans la figure 2.23.

On observe une concordance satisfaisante entre les courbes expérimentales et celles résultant des simulations. Il est à noter que cette courbe I-V, n'est pas une courbe en HRS issu d'une OxRAM.

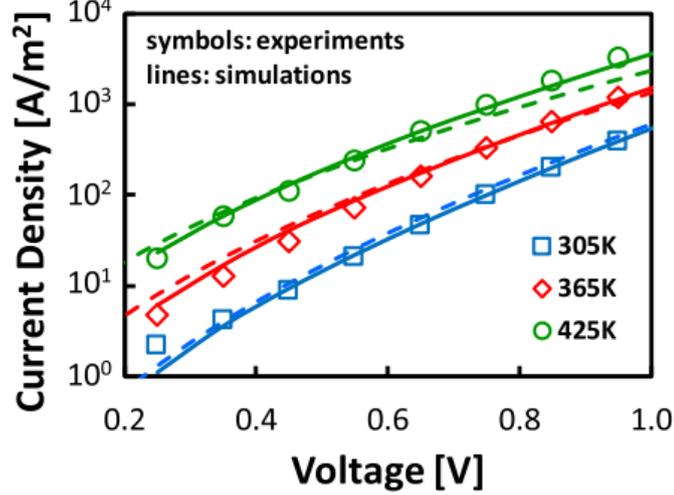


FIGURE 2.23 – Courbe I-V en fonction de la température, issu du modèle TAT larcher et *al.* Source : [12]

Il convient de souligner que ce qui a été décrit précédemment s'applique spécifiquement au HRS. L.Larcher et *al.* proposent une distinction dans le mécanisme de conduction en deux parties : il suggère que le HRS est régi par un transfert d'électrons tunnel (TAT), tandis que le LRS, en raison de la densité importante de pièges présents dans l'oxyde, est considéré comme un métal et est donc régi par un modèle ohmique. Cette distinction a été démontrée dans [86].

2.5.3 TAT Ielmini et al.

Le modèle de Ielmini [13] se base sur une approche de Tsu-Esaki [87]. Il s'agit d'un passage de la bande de conduction de la cathode (continuum d'états) (coté n dans la figure 2.24) vers le piège qui est ici considéré comme un continuum et enfin vers la bande de conduction du métal (continuum d'états) (coté n+ dans la figure 2.24). Pour calculer le courant, Ielmini et *al.* utilisent la formule de Tsu-Esaki donné dans [87].

$$J = \alpha \int \frac{1}{2\pi^3 \hbar} \frac{\partial E(k_x)}{\partial k_x} (f_C - f_A) T(k) d^3 k_x \quad (2.4)$$

Où $\alpha = \frac{4\pi q m_{eff}}{\hbar^3}$, T est la transparence, f_C la fonction d'occupation à droite de la barrière (coté n dans la figure 2.24) et f_A la fonction d'occupation à gauche de la barrière (coté n+ dans la figure 2.24). En notant J_{eC} (figure 2.24) le courant allant de la cathode au piège, J_{eA} (figure 2.24) le courant allant du piège à l'anode et après

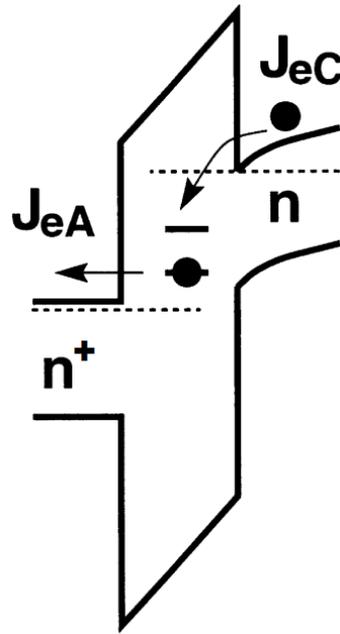


FIGURE 2.24 – Schéma récapitulatif du modèle de Ielmini à 1 piège. Source :[13]

l'intégration des fonctions d'occupations f sur les énergies transversales (énergies sur les axes y et z), on obtient les deux équations suivantes [13] :

$$J_{eC} = qN_T c_C \left(1 - \frac{f_T}{f_C(E_T)} \right) \quad (2.5)$$

$$J_{eA} = qN_T c_A \left(\frac{f_T}{f_A(E_T)} - 1 \right) \quad (2.6)$$

Avec :

$$\begin{aligned} c_C &= \frac{4\pi m_e k_B T \sigma_e}{\hbar^3} \int \ln \left(1 + \exp \left(-\frac{E - E_{FC}}{k_B T} \right) \right) T_{eC}(E) dE \\ c_A &= \frac{4\pi m_e k_B T \sigma_e}{\hbar^3} \int \ln \left(1 + \exp \left(-\frac{E - E_{FA}}{k_B T} \right) \right) T_{eA}(E) dE \end{aligned} \quad (2.7)$$

La particularité de ce modèle est de considérer que la fonction d'occupation du piège f_T est inconnue. La résolution de la continuité du courant entre les deux transitions permettra de déterminer la fonction d'occupation du piège f_T . En effet, en étant en état stationnaire, la condition de continuité stipule que $J_{eC} = J_{eA}$, conduisant ainsi à une équation du premier ordre à une inconnue, f_T . Résoudre cette équation permettra de déterminer la fonction d'occupation du piège, ouvrant ainsi la voie au calcul du courant traversant l'oxyde. Dans [14], Ielmini et *al.* ont proposé une extension de ce modèle à deux pièges. Il s'agit là de considérer la conduction non sur un seul mais sur deux pièges. Le formalisme avec lequel est traité le courant de la cathode vers le piège (noté J_1 sur figure 2.25) et du piège vers l'anode (noté J_2 sur figure 2.25) est le même que celui à un piège. La seule différence est la manière

avec laquelle est traité le courant inter-piège (noté J_{12} sur figure 2.25). En effet, pour écrire le courant du piège 1 vers le piège 2, Ielmini et *al.* introduisent un coefficient de capture inter piège c_{12} et c_{21} qui lui permet d'écrire le courant J_{12} en fonction des fonctions d'occupations du piège 1 f_1 et du piège 2 f_2 de la manière suivante :

$$J_{12} = qc_{12}f_1(1 - f_2) - qc_{21}f_2(1 - f_1) \quad (2.8)$$

Avec :

$$c_{12} = q \frac{D_{12}\Omega}{4\pi\tau_e} \exp\left(\frac{E_1 - E_2}{k_B T} \Theta(E_1 - E_2)\right) \quad (2.9)$$

où D_{12} est le coefficient de transmission tunnel incluant les effets d'abaissement de barrière des centres coulombiens (voir figure 2.25), τ_e est la durée de vie d'émission de piège, Ω est l'angle solide correspondant la section transversale du piège à la distance inter-piège R_{12} , et Θ est la fonction Heaviside. En intégrant ce coefficient de transmission tunnel D_{12} , Ielmini et *al.* supposent implicitement ici que les pièges possèdent une densité d'états en trois dimensions. Ils postulent que le courant tunnel entre deux pièges correspond à un tunnel entre deux continums d'états.

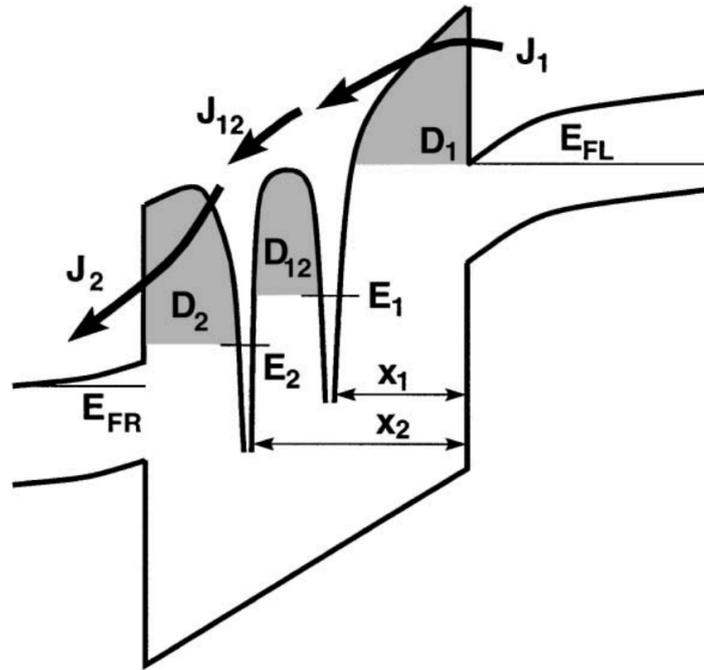


FIGURE 2.25 – Schéma récapitulatif du modèle de Ielmini à deux pièges. Source : [14]

De manière identique au modèle à 1 piège, les fonctions d'occupations étant inconnues, on utilise le fait que l'on soit en état stationnaire pour écrire $J_{12} = J_1 = J_2$. Un système d'équation de second ordre permettra ainsi de résoudre les fonctions d'occupations et donc de trouver le courant traversant l'oxyde.

Étant donné les résultats prometteurs du modèle de Ielmini et sa capacité à être généralisé à N pièges, nous avons pris la décision de poursuivre le développement de

ce modèle. Ainsi, nous pourrions obtenir un modèle basé sur l'équation de Tsu-Esaki, permettant de modéliser la conduction à travers N pièges. Ceci permettra d'avoir une nouvelle approche de la conduction dans une OxRAM, différente de celle de Larcher et *al.* Ce modèle sera exposé en détail dans le chapitre dédié à la conduction au sein des OxRAMs.

2.5.4 QPC

J. Andrade a utilisé le modèle Quantum Point Contact (QPC) [15] pour simuler la conduction dans les ReRAM à base de Ta_2O_5 . Ce modèle repose sur le formalisme de Landauer-Büttiker [88, 89] qui décrit le flux d'électrons à travers un conducteur quantique. Selon ce modèle, le CF contient une zone de passage étroite dont le diamètre est inférieur au libre parcours moyen des électrons, comme le montre la figure 2.26. Cette restriction conduit à une quantification transversale de l'énergie des électrons. L'énergie potentiel le long de l'axe du filament pour chaque niveau quantifié peut être considéré comme une sous-bande, agissant comme des barrières de potentiel.

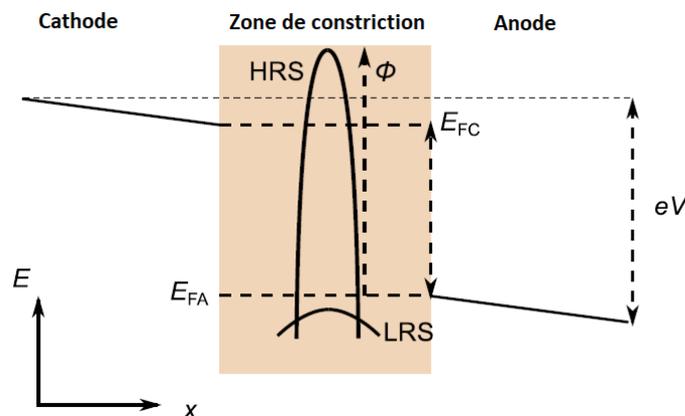


FIGURE 2.26 – Diagramme de bande représentant la zone de constriction. Source : [15]

En utilisant le formalisme de Landauer-Büttiker [88, 89], le courant à travers la zone de passage étroite peut être exprimé comme :

$$I = \frac{2e^2}{h} \int T(E)[f_C(E) - f_A(E)]dE \quad (2.10)$$

où e est la charge de l'électron, h est la constante de Planck, E est l'énergie, f_C et f_A sont la fonction de distribution de Fermi-Dirac pour la cathode et l'anode respectivement, et l'intégration est effectuée sur toutes les énergies. T est la transparence

de la barrière et peut s'exprimer comme suit :

$$T(E) = \Sigma_n \left(1 + \exp \left(-2\pi \frac{E - \hbar\omega_y(n+1) + eV_0}{\hbar\omega_x} \right) \right)^{-1} \quad (2.11)$$

où eV_0 est l'énergie potentielle au niveau de la zone de constriction, et ω_x et ω_y sont des fréquences liées à la courbure du potentiel, et donc de la forme de la constriction, le long de l'axe du filament et dans la direction transversale, respectivement. Des études antérieures [90, 91] ont montré que la fonction de transparence décroît rapidement au fur et à mesure que l'on s'éloigne de l'état fondamental. Par conséquent, seule la première sous-bande contribue significativement à la conduction. Cela simplifie considérablement la somme dans 2.11, qui se réduit à un seul élément. Ainsi, le courant peut être exprimé de la manière suivante :

$$I = \frac{2e^2}{h} N \left[eV + \frac{1}{\alpha} \ln \left(\frac{1 + \exp(\alpha(\phi - \beta eV))}{1 + \exp(\alpha(\phi + (1 + \beta)eV))} \right) \right] \quad (2.12)$$

où V est la tension appliquée, $\alpha = \frac{2\pi}{\hbar\omega_x}$. $\phi = \hbar\omega_y + eV_0$ considéré comme la hauteur de barrière. β détermine l'emplacement de la constriction. Le niveau de Fermi de la cathode est β au-dessus du niveau de Fermi au niveau de la zone de constriction. Le niveau de Fermi de l'anode est $(1 - \beta)$ en dessous de ce point. N est le nombre de canaux de conduction parallèle.

Pour modéliser à la fois le HRS et le LRS avec une seule approche, on met en place une modulation énergétique de la zone de constriction. La figure 2.27 (a) considère que, dans le cas du HRS, la zone de constriction est minimisée du point de vue géométrique. De plus, la figure 2.26 (a) indique que la hauteur de la barrière ϕ est également élevée dans le HRS. En revanche, pour le LRS, la figure 2.27 (b) et la figure 2.26 montrent une expansion de la zone de constriction et une réduction de la hauteur de la barrière par rapport au HRS.

J. Andrade et *al.* parviennent à reproduire les courbes de courant en fonction de la tension à l'aide de ce modèle, comme en témoigne la figure 2.28. On observe que les courbes obtenues par simulation sont en bons accords avec les courbes expérimentales.

2.5.5 Poole-Frenkel

Le modèle de Poole-Frenkel [92] est un modèle de conduction par saut (hopping) empirique qui décrit le comportement de la conductivité électrique dans les matériaux isolants soumis à des champs électriques élevés. Le modèle suppose que la conductivité électrique est gouvernée par l'émission de porteurs à partir de pièges localisés dans le matériau, comme le montre la figure 2.29. Le modèle de Poole-Frenkel est souvent utilisé pour expliquer le phénomène d'émission de porteurs de

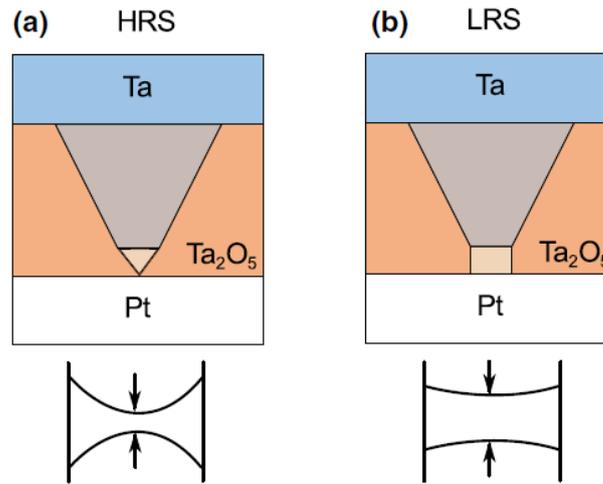


FIGURE 2.27 – Schéma récapitulatif du modèle de QPC : (a) HRS (b) LRS. Source : [15]

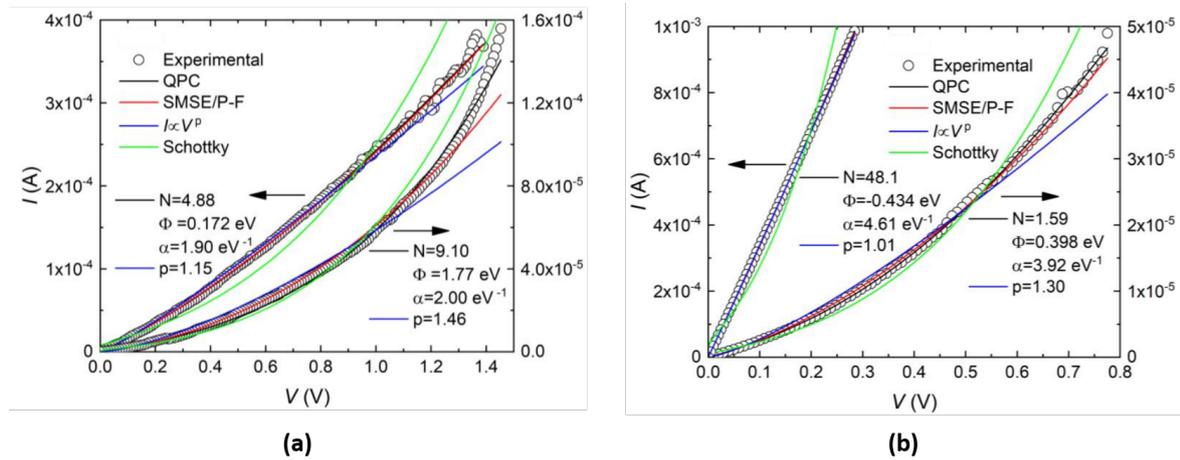


FIGURE 2.28 – Courbe courant en fonction de la tension en : (a) HRS (b) LRS. Source : [15]

charge depuis des pièges profonds dans les matériaux isolants, qui est renforcé par les champs électriques élevés [93].

L'équation de courant dans le modèle de Poole-Frenkel est donnée par :

$$I = I_0 E \exp\left(\frac{\beta_{PF} \sqrt{E} - \phi_t}{k_B T}\right) \quad (2.13)$$

où I est la densité de courant, E est le champ électrique appliqué, I_0 est une constante liée à la concentration de pièges et à la probabilité d'émission, $\beta_{PF} = \sqrt{\frac{q^3}{\pi \epsilon \epsilon_0}}$, q est la charge électronique, ϵ_0 est la constante diélectrique du vide, ϵ est la constante diélectrique relative, T est la température, k_B est la constante de Boltzmann et ϕ_t est la profondeur du puits potentiel de pièges.

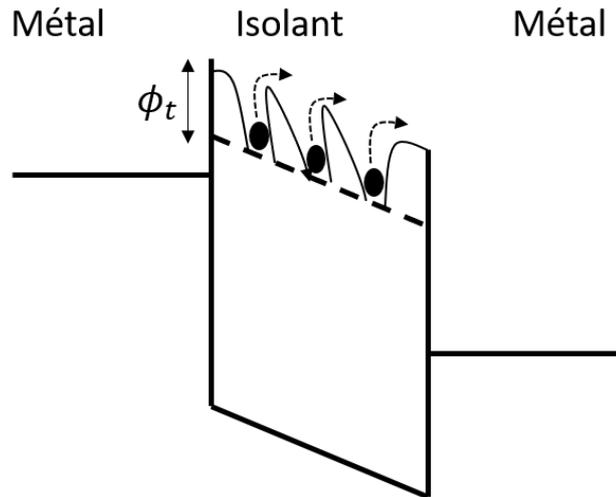


FIGURE 2.29 – Schéma du mécanisme Poole-Frenkel. Source : [16]

L'équation décrit une relation non linéaire entre le champ électrique appliqué et la densité de courant résultante, avec une dépendance exponentielle de la force du champ. À mesure que le champ électrique augmente, la probabilité que les porteurs de charge s'échappent des pièges et contribuent au courant augmente de manière exponentielle. Ceci est dû au fait que l'émission de porteurs de charge est limitée par la barrière d'énergie associée aux pièges, qui diminue avec l'augmentation du champ.

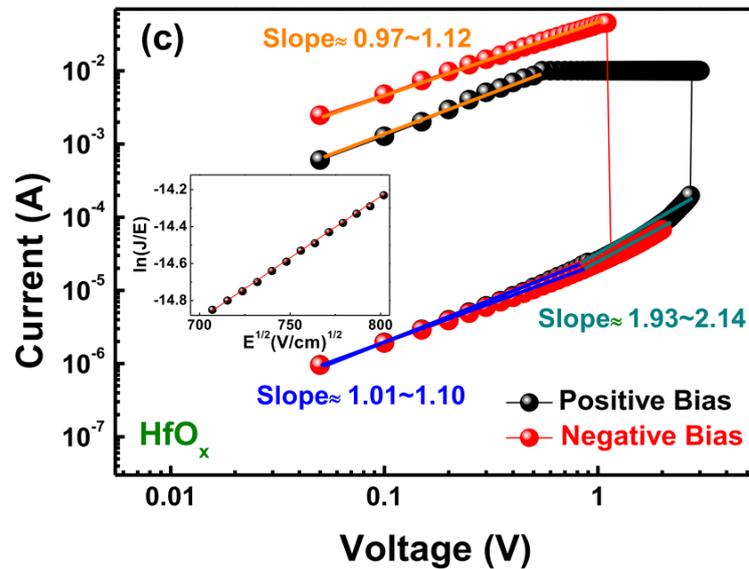


FIGURE 2.30 – Courbe courant en fonction de la tension issu de données de simulation en HRS et LRS. Source : [17]

Dans Jang et al.[17], le modèle de Poole-Frenkel a été utilisé pour modéliser la conduction dans les OxRAM. L'état LRS est caractérisé par une forte relation

linéaire entre le courant et la tension, suggérant que la conduction ohmique est le mécanisme principal, comme le montre le graphique I-V linéaire dans la Figure 2.30. En revanche, l'état HRS est régi par la conduction ohmique à faible champ, tandis que l'effet non linéaire prédomine à des tensions élevées ($V > 0,5$ V). Cette région non linéaire est bien décrite par le modèle P-F ($\ln\left(\frac{I}{V}\right) \propto V^{1/2}$), comme indiqué dans la Figure 2.30.

2.6 Objectifs et problématique de la thèse

Dans les sections précédentes, nous avons présenté les figures de mérite et les effets parasites qui ont une influence sur la performance de l'OxRAM. Bien que les OxRAMs commencent à être industrialisées, leur production de masse n'a pas encore été pleinement réalisée. Des améliorations sont encore nécessaires pour répondre aux exigences de performance et de fiabilité à grande échelle. En effet, certains critères tels que l'endurance et le temps d'accès montrent des résultats très prometteurs, par contre la variabilité en courant est le principal problème de l'OxRAM. Il est donc crucial de comprendre l'origine et les facteurs influents sur cet effet parasite. C'est un enjeu important dans l'industrie et la Recherche et Développement des NVM émergentes. La variabilité est principalement liée au mécanisme de commutation, comme l'a montré Wong et *al.* [6], et pour mieux comprendre ce phénomène, il est nécessaire de comprendre et de modéliser ce mécanisme. Cependant, pour étudier le comportement global de l'OxRAM, il est également important de comprendre les deux autres mécanismes impliqués dans son fonctionnement : le mécanisme de forming, qui permet de construire le filament conducteur du dispositif pour la première fois, et le mécanisme de conduction, qui décrit le mouvement des électrons dans l'oxyde métallique.

Dans cette thèse, notre attention s'est focalisée sur le mécanisme de conduction de l'OxRAM. Dans les sections précédentes, divers modèles ont été exposés afin d'élucider ce mécanisme. Il convient de souligner que plusieurs modèles de conduction ont été utilisés pour simuler le comportement électrique d'un OxRAM. Notre revue de la littérature a également révélé une distinction de modèles entre les deux états de conduction. En effet, en raison de la densité élevée de pièges dans le filament conducteur, la conduction en LRS est considérée comme ohmique, tandis que pour le HRS, plusieurs modèles ont été avancés. Cette différenciation est faite sur les modèles de conduction, bien que le dispositif modélisé reste le même, sans changement de nature dans l'oxyde. En effet, comme mentionné dans [56, 59, 58], la commutation entre l'état HRS et LRS se produit par la destruction ou la reconstruction du filament conducteur, qui résulte de la migration des lacunes d'oxygène sous l'effet d'une différence de potentiel. La question que nous abordons dans cette thèse est de savoir pourquoi deux modèles sont nécessaires pour expliquer le comportement élec-

trique d'un OxRAM. Sachant que la conduction dans ce type de composant est une conduction dans une chaîne d'états finie, ne serait-il pas envisageable de modéliser une chaîne d'états avec un seul modèle quelque soit le nombre d'états ?

Bien que certaines réponses aient été fournies dans la littérature, notamment avec le modèle QPC de J. Andrade [15], qui est capable de reproduire le comportement d'une OxRAM non seulement en LRS mais aussi en HRS, une compréhension approfondie du mécanisme de conduction est nécessaire pour répondre complètement à notre problématique.

Pour tenter de résoudre cette problématique, nous envisageons d'aborder, dans un premier temps, les deux modèles développés pour reproduire les caractéristiques électriques de l'OxRAM à la fois en HRS et en LRS. Cette partie sera traitée dans le prochain chapitre, intitulé "Modélisation de la conduction".

Afin de tester la robustesse de notre modèle, nous avons également décidé d'effectuer des mesures courant-tension en fonction de la température sur des structures OxRAM. Cette démarche nous permettra de comparer nos données expérimentales à celles déjà existantes dans la littérature. En outre, nous avons entrepris d'obtenir une compréhension claire du comportement de ce composant en fonction de la température, car la littérature rapporte des conclusions divergentes à ce sujet. Cette partie sera traitée dans le dernier chapitre, intitulé "Étude expérimentale".

Chapitre 3

Modélisation de la conduction

3.1 Introduction

Dans le chapitre précédent, plusieurs modèles ont été présentés, chacun apportant une contribution à la compréhension de la conduction dans les dispositifs OxRAM. Le modèle Trap Assisted Tunelling (TAT) de Larcher et *al.* a partiellement réussi à ajuster les différentes caractéristiques du dispositif. Cependant, ses limites principales résident dans l'utilisation d'un nombre très important de pièges pour atteindre les niveaux de courant attendus [11]. En effet, pour obtenir les niveaux de LRS, un nombre de pièges supérieur à 100 [11] sur une épaisseur d'oxyde de 5 nm est nécessaire, ce qui est peu réaliste. De plus, au lieu d'avoir un seul modèle de conduction pour modéliser les deux états de résistance, Larcher et *al.* considèrent que le HRS est régi par un modèle TAT, tandis que le LRS est géré par un modèle ohmique, car la conduction y est considérée comme métallique. En analysant le modèle, on constate que l'équation de continuité, c'est-à-dire la condition de non-accumulation de charge, n'est pas nécessairement respectée chez Larcher car elle n'est pas prise en compte dans son ensemble d'équations. À aucun moment, il ne spécifie que le courant entre les pièges doit être le même quel que soit le piège considéré.

En revanche, le modèle de Ielmini respecte la condition de continuité et la considère même comme la pierre angulaire de son modèle [13]. Cependant, le seul inconvénient de ce modèle est que Ielmini n'a présenté qu'une version à deux pièges. Bien qu'il ait mentionné que son modèle soit généralisable à N pièges avec $N > 2$, il ne l'a jamais démontré.

Ainsi, dans cette section, nous présenterons une méthode permettant de respecter la condition de continuité indépendamment du modèle physique sous-jacent. Tout d'abord, nous introduirons un formalisme général valable pour n'importe quel modèle, montrant que pour respecter la condition de continuité, il sera nécessaire de résoudre les niveaux de Fermi des pièges. Ensuite, nous expliquerons comment nous avons appliqué ce formalisme pour généraliser le modèle de Ielmini et comment

nous l'avons utilisé sur un modèle peu présent dans la littérature sur les OxRAM : le modèle de hopping, dit modèle de Miller et Abrahams. Dans ce chapitre, nous aborderons également deux approches différentes permettant de calculer le courant, à savoir l'approche Nearest Neighbour (NN) et l'approche Variable Range (VR). À la fin de chaque formulation de modèle, nous inclurons une section dédiée à la validation, afin de vérifier la robustesse du modèle proposé. Nous finirons ce chapitre par faire une comparaison entre les données expérimentales et le modèle qui sera le plus cohérent et le plus apte à reproduire le comportement électrique d'une OxRAM

3.2 Méthode numérique pour résolution des niveaux de Fermi

Dans cette section, nous allons d'abord établir le cadre de la simulation en définissant les hypothèses qui ont été prises pour mener à bien ces simulations. Nous expliquerons également comment l'équation de continuité est liée à la résolution des niveaux de Fermi. Ensuite, nous présenterons le formalisme qui permet de respecter l'équation de continuité et de résoudre les niveaux de Fermi en utilisant deux approches différentes pour calculer le courant : l'approche Nearest Neighbour et l'approche Variable Range.

3.2.1 Cadre de simulation

Le figure 3.1 représente le diagramme de bande considérée pour un dispositif OxRAM. On suppose que les électrodes métalliques (à savoir l'anode et la cathode) agissent toutes deux en tant que réservoirs infinis de porteurs sous l'équilibre thermodynamique. Une couche d'oxyde de quelques nanomètres d'épaisseur, généralement HfO_2 , est située entre les deux contacts métalliques. En supposant que l'oxyde n'est plus dans un état vierge [94] et qu'un forming s'est déjà produit, un certain nombre de pièges sont donc disponibles dans l'ensemble de l'oxyde. Les pièges sont caractérisés par leur niveau d'énergie (E_t dans la figure 3.1), leur niveau de Fermi (E_f dans la figure 3.1), leur coordonnée transversale (x dans la figure 3.1) et leur coordonnée longitudinale (dénommée la coordonnée y dans la suite). Dans l'absolu, la répartition de ces pièges se fait dans 3 dimensions, toutefois dans le cadre de la simulation, une approche 1D du problème a été adopté. On considéra ainsi dans ce qui suit une chaîne simple de piège en 1D réparti en x , mais pas en y , ni en z . Nous verrons plus tard dans ce chapitre que la répartition des pièges en 3D est en effet prise en compte par notre algorithme.

Sur l'ensemble des simulations qui seront faites un certains nombres de paramètres seront utilisés. Ils sont résumés dans le tableau 3.1.

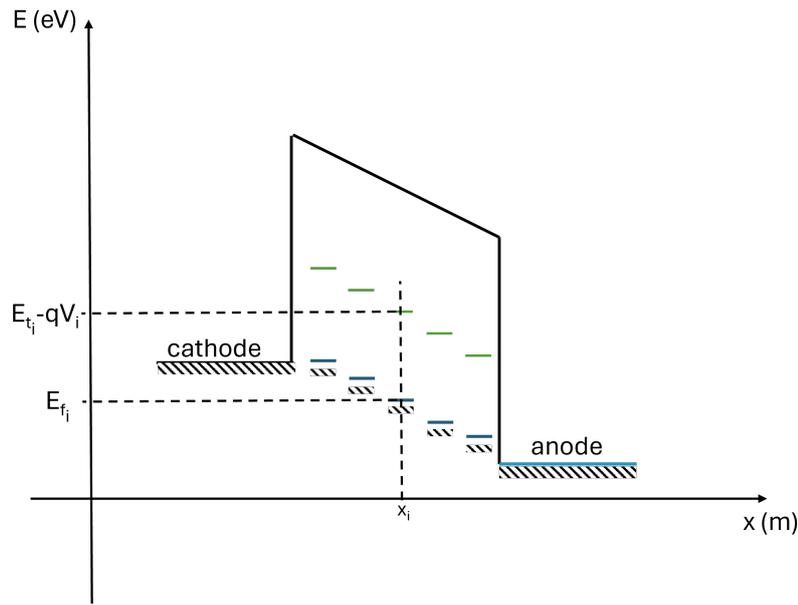


FIGURE 3.1 – Schéma du diagramme de bande de l'OxRAM. x_i représente la position du piège le long de l'oxyde. E_{f_i} représente le niveau de Fermi du piège. E_{t_i} représente l'énergie du piège. V_i représente la tension appliquée sur le piège.

3.2.2 Nécessité de résoudre les niveaux de Fermi

Dans le modèle de Ielmini, il est stipulé que le courant arrivant vers un piège est égale au courant partant de celui-ci ce qui peut se traduire pour un piège i quelconque par : $J_{i,i+1}(E_{f_i}, E_{f_{i+1}}) = J_{i-1,i}(E_{f_{i-1}}, E_{f_i})$. Cet énoncé est la formulation de l'équation de continuité, il s'agit de considérer qu'en régime permanent, il n'y a ni accumulation ni disparition d'électrons le long du chemin de conduction. En d'autres termes, si l'on prend un volume de contrôle dans un chemin de conduction, on constatera que le nombre d'électrons qui y entre est égal à celui qui en sort. Cette équation bilan permet ainsi de garantir la continuité du courant.

Afin de formuler l'équation de continuité, on utilise un opérateur différentiel bien connu, à savoir la divergence. Pour exprimer la conservation du courant, on peut reformuler cela en indiquant que la divergence du courant est égale à zéro. Cette condition se traduit par l'équation (3.1), où le terme $div(J)$ représente la divergence du courant. L'équation prend en compte les courants $J_{i,i+1}$ et $J_{i-1,i}$ à deux points distincts de la chaîne de conduction, ainsi que les niveaux de Fermi associés à ces points, à savoir E_{f_i} , $E_{f_{i+1}}$ et $E_{f_{i-1}}$. Il est important de souligner que la seule variable pertinente dans nos courants est le niveau de Fermi et que ce sera notre unique point d'ajustement pour les courants.

En d'autres termes, l'équation de continuité du courant sert à équilibrer les flux de courant à différents points de la chaîne de conduction, en ajustant les niveaux de Fermi de manière à ce que la divergence du courant soit nulle. Le niveau de Fermi est

Paramètre	Dénomination	Valeur
t_{ox}	Épaisseur d'oxyde	5 nm
Φ_l	Hauteur de Barrière du HfO_2	3.1 eV
T	Température	298 K
N_t	Nombre de piège dans l'oxyde	Variable
E_t	Energie du piège	0.2 eV
E_f	Niveau de Fermi du piège	Variable
x	Position du piège	Variable
m_{eff}	Masse effective	$0.5m_0$
α	Largeur de densité d'état (paramètre modèle Miller et Abrahams)	$10^{-10}m$
w_0	Fréquence maximale de sauts (paramètre modèle Miller et Abrahams)	$10^{17}s^{-1}$

TABLE 3.1 – Tableau récapitulatifs des principaux paramètres de simulations

donc la seule variable non fixée à l'avance dans cette équation, et il est utilisé comme inconnu pour garantir la continuité du courant dans les composants électroniques.

$$div(J) = 0 \Leftrightarrow J_{i,i+1}(E_{f_i}, E_{f_{i+1}}) - J_{i-1,i}(E_{f_{i-1}}, E_{f_i}) = 0 \quad (3.1)$$

Une fois que la condition de continuité respectée, il reste néanmoins à établir le système d'équations qui en découle et à le résoudre. Dans la plupart des cas, le système à résoudre est non linéaire, ce qui rend impossible la séparation des variables et la résolution par inversion de matrice. Par conséquent, il est nécessaire de recourir à une méthode itérative pour résoudre ce système. La section suivante abordera la méthode de Newton-Raphson, une approche couramment utilisée pour résoudre ce type de système non linéaire.

3.2.3 La méthode de Newton-Raphson

Cette partie a été largement inspiré du livre de S. Mazumder sur les méthodes numérique [95].

La méthode de Newton-Raphson est une méthode numérique itérative, utilisée pour résoudre des équation non linéaires, qui convergent rapidement vers la solution pour trouver les zéros d'une fonction différentiable, mais elle peut ne pas converger ou converger lentement dans certaines situations. Pour que la méthode converge, il est important que la fonction soit différentiable et continue sur l'intervalle considéré, que l'estimation initiale soit suffisamment proche de la solution, que la fonction ait une dérivée continue et ne change pas de signe à proximité du zéro, et que les équations du système d'équations non linéaires soient bien conditionnées [96].

Tout d'abord, considérons une seule équation de la forme $f(x) = 0$, où x est une variable scalaire. La méthode de Newton-Raphson commence par une supposition initiale x_0 pour la racine de l'équation, puis affine itérativement la supposition en utilisant la formule :

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)} \quad (3.2)$$

Ici, $f'(x^k)$ est la dérivée de $f(x)$ par rapport à x , évaluée en x^k . Géométriquement, cette formule signifie que nous prenons la supposition actuelle x^k et la déplaçons horizontalement d'une distance de $\frac{f(x^k)}{f'(x^k)}$, pour obtenir la prochaine supposition x^{k+1} . Ce processus est répété jusqu'à ce qu'un niveau de précision souhaité soit atteint, généralement en vérifiant si le changement de x entre les itérations est plus petit qu'une certaine tolérance.

Pour résoudre un système d'équations, nous pouvons étendre la méthode de Newton-Raphson aux fonctions vectorielles. Considérons un système de n équations avec n variables, qui peut être écrit sous la forme :

$$F(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_i(x) \\ \vdots \\ f_n(x) \end{pmatrix} \quad (3.3)$$

Ici, $x = (x_1(x), \dots, x_i(x), \dots, x_n(x))^T$ est le vecteur de variables, et $f_i(x)$ est la i^{eme} équation. Nous voulons trouver un vecteur x qui satisfait $F(x) = 0$.

La méthode de Newton-Raphson pour les systèmes d'équations commence par une supposition initiale x^0 pour la solution, puis affine itérativement la supposition en utilisant la formule :

$$x^{k+1} = x^k - \mathcal{J}^{-1}(x^k)F(x^k) \quad (3.4)$$

Ici, $\mathcal{J}(x_n)$ est la matrice jacobienne de $F(x)$ évaluée en x_n , donnée par :

$$\mathcal{J}(x_n) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_i} & \dots & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \frac{\partial f_i}{\partial x_1} & \dots & \frac{\partial f_i}{\partial x_i} & \dots & \dots & \frac{\partial f_i}{\partial x_n} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_i} & \dots & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} \quad (3.5)$$

La jacobienne est une matrice dont l'entrée (i,j) est la dérivée partielle de $f_i(x)$ par rapport à x_j . L'inverse de la jacobienne, $\mathcal{J}^{-1}(x^k)$, est utilisé pour résoudre le changement de x entre les itérations. Géométriquement, cette formule signifie que nous prenons la supposition actuelle x^k et la déplaçons d'une distance de $\mathcal{J}^{-1}(x^k)F(x^k)$, qui correspond à la solution du système linéaire $J(x^k)dx = -F(x^k)$, où $dx = x^{k+1} - x^k$

est la variation de x entre les itérations. Pour implémenter la méthode de Newton-Raphson pour les systèmes d'équations, nous pouvons suivre ces étapes :

1. Choisir une supposition initiale x^0 pour la solution.
2. Calculer $F(x^0)$ et $\mathcal{J}^{-1}(x^0)$.
3. Résoudre le système linéaire $dx = -\mathcal{J}^{-1}(x^0)F(x^0)$ pour dx .
4. Calculer $x^1 = x^0 + dx$.
5. Répéter les étapes 2-4 avec x^k remplacé par x^{k+1} jusqu'à ce qu'un niveau de précision souhaité soit atteint, c'est à dire un dx satisfaisant.

Comme dans le cas d'une seule équation, la méthode de Newton-Raphson pour les systèmes d'équations présente certaines limitations. Si la supposition initiale n'est pas suffisamment proche de la solution, la méthode peut converger vers une solution différente ou ne pas converger du tout.

Une fois la méthode mathématique qui permettra la résolution du système qui découle de l'équation de continuité, passons maintenant à l'exposé de ce formalisme en utilisant les deux approches qui nous permettent de calculer les courant : l'approche Nearest Neighbour et l'approche Variable Range.

3.2.4 Approche Nearest Neighbour

Dans l'approche Nearest Neighbour [97], on suppose que le courant ne se produit que entre les pièges les plus proches. En régime permanent, le courant $J_{i,i+1}$ entre les pièges i et $i + 1$ et $J_{i-1,i}$, tels qu'illustré dans la figure 3.2, sont égaux d'après la continuité du courant. Les deux courants sont des fonctions des niveaux de Fermi des pièges, respectivement $E_{f_{i-1}}$, E_{f_i} et $E_{f_{i+1}}$. Dans la figure 3.2, $i, i - 1$ et $i + 1$ représentent l'indice des positions des pièges. N_t représente l'indice du dernier piège et donc le nombre total de piège.

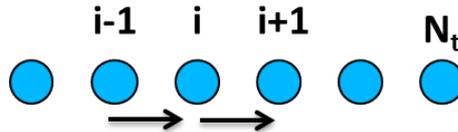


FIGURE 3.2 – Schéma de l'approche Nearest Neighbour.

En considérant tous les pièges situés entre les deux électrodes, l'équation (3.1) fournit un ensemble d'équations dont les inconnues sont les niveaux de Fermi des pièges. Certaines hypothèses pourraient être faites pour simplifier et linéariser ce système (par exemple en ne considérant que la distribution de Maxwell-Boltzmann), mais dans le cas le plus général, le système est non linéaire et nécessite une méthode

itérative telle que la méthode de Newton-Raphson pour être résolu. Ce schéma numérique est formulé dans l'équation 3.6.

$$E_f^{(k+1)} = E_f^{(k)} - \mathcal{J}^{-1} F(E_f^{(k)}) \quad (3.6)$$

$$\text{avec } \begin{cases} E_f^{(k+1)} & : \text{ Vecteur ligne des niveaux de Fermi à l'itération } k+1. \\ E_f^{(k)} & : \text{ Vecteur ligne des niveaux de Fermi à l'itération } k+1 \text{ k.} \\ \mathcal{J} & : \text{ Matrice Jacobienne du système.} \\ F & : \text{ Vecteur représentant } \textit{div}(J) \end{cases}$$

F représente l'ensemble des équations qui devrait être nuls, c'est-à-dire la différence entre $J_{i,i+1}$ et $J_{i-1,i}$, fonction du vecteur de E_f^i , et a la forme suivante :

$$F = \begin{pmatrix} J_{1,2}(E_{f_1}, E_{f_2}) - J_{c,1}(E_{f_c}, E_{f_1}) \\ \vdots \\ J_{i,i+1}(E_{f_i}, E_{f_{i+1}}) - J_{i-1,i}(E_{f_{i-1}}, E_{f_i}) \\ \vdots \\ J_{N_t,a}(E_{f_{N_t}}, E_{f_a}) - J_{N_t-1,N_t}(E_{f_{N_t-1}}, E_{f_{N_t}}) \end{pmatrix} \quad (3.7)$$

Dans l'équation (3.7), les conditions aux limites sont définies en fixant E_{f_c} et E_{f_a} , qui sont respectivement les niveaux de Fermi de la cathode et de l'anode, séparé d'une différence proportionnelle à la tension appliquée. La matrice jacobienne \mathcal{J} est issue de l'équation 3.8 et est une matrice tridiagonale dans le cas de l'approche Nearest Neighbour.

$$\mathcal{J} = \begin{pmatrix} \frac{\partial F_1}{\partial E_{f_1}} & \frac{\partial F_2}{\partial E_{f_1}} & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \frac{\partial F_i}{\partial E_{f_{i-1}}} & \frac{\partial F_i}{\partial E_{f_i}} & \frac{\partial F_i}{\partial E_{f_{i+1}}} & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & \frac{\partial F_{N_t}}{\partial E_{f_{N_t-1}}} & \frac{\partial F_{N_t}}{\partial E_{f_{N_t}}} \end{pmatrix} \quad (3.8)$$

L'équation (3.7) est mise en œuvre dans une boucle convergente, qui commence par une estimation initiale $E_f^{(k=0)}$ correspondant à une distribution linéaire des niveaux de Fermi entre la cathode et l'anode. La convergence est atteinte lorsque la différence quadratique moyenne entre $E_f^{(k)}$ et $E_f^{(k+1)}$ est inférieure à une valeur seuil, fixé à 0,01 meV.

3.2.5 Approche Variable Range

Dans le cas du Variable Range, le courant est supposé possible entre chaque piège, comme illustré dans la figure 3.3. Cependant, le principe reste le même que dans le cas du Nearest Neighbour, il consiste à construire un ensemble d'équations basées sur la continuité du courant.

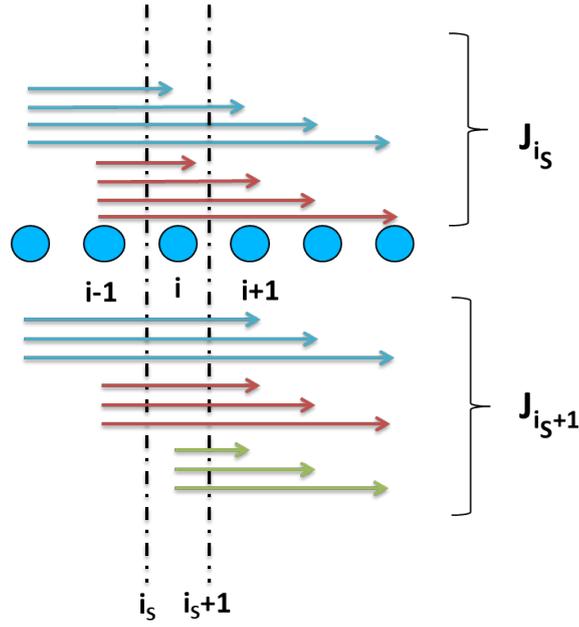


FIGURE 3.3 – Schéma de l’approche Variable Range sur une chaîne de pièges : Tout les pièges avant la section i_s contribuent au courant.

Pour ce faire, une section entre les pièges est considérée, à travers celle-ci, toutes les contributions de courant s’ajoutent. Dans la figure 3.3, la section i_s est traversée par un courant émis par chaque piège situé avant le piège i (exclu), vers tous les pièges situés après $i - 1$ (également exclu). Cela conduit à une double sommation exprimée dans l’équation 3.9 :

$$J_{i_s} = \sum_{i=0}^{i-1} \sum_{j=i}^{Nt} J_{i,j} \quad (3.9)$$

Cette procédure est ensuite répétée pour la section suivante i_{s+1} et le courant $J_{i_{s+1}}$. Comme dans le cas du voisinage le plus proche, en raison de la continuité du courant, la différence entre $J_{i_{s+1}}$ et J_{i_s} est nulle. La différence entre les deux doubles sommes donne :

$$\begin{aligned} J_{i_{s+1}} - J_{i_s} &= \sum_{i=0}^i \sum_{j=i+1}^{Nt} J_{i,j} - \sum_{i=0}^{i-1} \sum_{j=i}^{Nt} J_{i,j} \\ &= \sum_{i=0}^{i-1} \sum_{j=i+1}^{Nt} J_{i,j} - \sum_{i=0}^{i-1} \sum_{j=i}^{Nt} J_{i,j} + \sum_{j=i+1}^{Nt} J_{i,j} = 0 \\ &\Leftrightarrow \text{div}(J) = 0 \end{aligned} \quad (3.10)$$

Le schéma de Newton-Raphson est presque inchangé. F est toujours donné par l’équation 3.7. Seulement \mathcal{J} est modifié, et est maintenant une matrice pleine donnée

par l'équation 3.11.

$$\mathcal{J} = \begin{pmatrix} \frac{\partial F_1}{\partial E_{f_1}} & \cdots & \frac{\partial F_1}{\partial E_{f_i}} & \cdots & \frac{\partial F_1}{\partial E_{f_{N_t-1}}} & \frac{\partial F_1}{\partial E_{f_{N_t}}} \\ \vdots & \ddots & & \ddots & & \vdots \\ \frac{\partial F_i}{\partial E_{f_1}} & \cdots & \frac{\partial F_i}{\partial E_{f_i}} & \cdots & \frac{\partial F_i}{\partial E_{f_{N_t-1}}} & \frac{\partial F_i}{\partial E_{f_{N_t}}} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \frac{\partial F_{N_t}}{\partial E_{f_1}} & \cdots & \frac{\partial F_{N_t}}{\partial E_{f_i}} & \cdots & \frac{\partial F_{N_t}}{\partial E_{f_{N_t-1}}} & \frac{\partial F_{N_t}}{\partial E_{f_{N_t}}} \end{pmatrix} \quad (3.11)$$

Ainsi, dans les parties consacrées aux approches NN et VR, nous avons présenté un formalisme permettant de résoudre les niveaux de Fermi, sans toutefois nécessiter une expression explicite du courant. Cela signifie que nous avons développé un schéma de résolution suffisamment général, permettant d'explorer divers modèles de transport du courant entre les pièges, en exprimant la dépendance du courant J par rapport aux niveaux de Fermi des pièges.

3.3 Conduction assistée par piège

Le modèle développé par Ielmini et *al.* [13, 14] est un modèle qui convient bien à la description des courants de fuite dans les transistors et les mémoires Flash, mais qui est également intéressant car il traite de la conduction inter-pièges dans les oxydes. De plus, ce modèle peut être généralisé pour inclure N-pièges, ce qui en fait un modèle prometteur pour la modélisation de la conduction dans les dispositifs OxRAM.

Dans la partie modélisation, nous allons expliquer comment généraliser ce modèle en N-pièges en utilisant l'équation de Tsu-Esaki [87]. Tout d'abord, nous allons faire une hypothèse forte sur les fonctions d'occupation en considérant qu'elles sont des fonctions de Boltzmann. Ensuite, nous allons examiner ce qui se passe lorsque les pièges sont dans un état dégénéré, c'est-à-dire lorsque le niveau de Fermi est à un niveau supérieur de celui du piège. Enfin, nous allons présenter une approche plus générale en considérant que les fonctions d'occupation sont des fonctions de Fermi-Dirac, qui permettent de prendre en compte la dégénérescence des pièges. Cependant, cette approche conduit à un système non-linéaire, nécessitant l'utilisation de la méthode de Newton-Raphson, que nous avons précédemment expliquée.

Dans la partie validation, nous débuterons en établissant un protocole de vérification de la robustesse des modèles. Tout d'abord, nous examinerons si l'équation de continuité est respectée en analysant le comportement du courant et des niveaux de Fermi en fonction de la position. Ensuite, nous évaluerons la forme des courbes I-V ainsi que les niveaux de courant en HRS et en LRS. Pour conclure cette phase de vérification, nous observerons le comportement en température du modèle, en examinant les courbes I-V à différentes températures.

3.3.1 Formulation du modèle de TE (Tsu-Esaki) en Boltzmann

Modélisation

En reprenant la Figure 3.1, qui représente une barrière remplie de N pièges avec des niveaux d'énergie E_t supérieurs à E_f , on considère les N pièges comme étant N réservoirs thermodynamiques possédant un continuum d'états 3D. Dans ce contexte, la cathode et l'anode sont nos conditions aux limites avec des niveaux de Fermi fixe et connue. Considérons une séquence de trois pièges : $i-1$, i et $i+1$. En utilisant l'équation de Tsu-Esaki 3.35, et en remplaçant les fonctions d'occupations par les fonctions d'occupation de Fermi-Dirac des pièges, l'expression du courant entre deux pièges dans l'approche Nearest Neighbour sera donnée par :

$$J_{i-1,i} = \alpha \int_{\max(E_{t_{i-1}} - V_{t_{i-1}}, E_{t_i} - V_{t_i})}^{\Phi_l - V_{t_{i-1}}} \mathcal{T}(E_x) \ln \left(1 + \exp \left(-\frac{E_x - E_{f_{i-1}}}{k_B T} \right) \right) - \mathcal{T}(E_x) \ln \left(1 + \exp \left(-\frac{E_x - E_{f_i}}{k_B T} \right) \right) dE_x \quad (3.12)$$

Où \mathcal{T} est la transparence WKB de la barrière entre le piège i et $i-1$ considéré comme trapézoïdale, explicité par la formule suivante [85] :

$$\mathcal{T} = \exp \left(-\frac{4}{3} \left(\frac{\Phi_1}{h_\theta} \right)^{\frac{3}{2}} + \frac{4}{3} \left(\frac{\Phi_2}{h_\theta} \right)^{\frac{3}{2}} \right) \quad (3.13)$$

$$\text{Avec } \begin{cases} \alpha = \frac{4\pi q m_{eff}}{h^3} \\ \Phi_1 = \Phi_l - V_{t_{i-1}} \\ \Phi_2 = \Phi_l - V_{t_i} \\ h_\theta = \left(\frac{qh \left(\frac{qV_{t_i}}{t_{ox}} \right)^2}{2m_{eff}} \right)^{\frac{1}{3}} \\ V_{t_i} = V \frac{x_i}{t_{ox}} \end{cases} \quad m_{eff} \text{ est une masse de densité d'état dans le piège}$$

Puisque $E_{t_i} > E_{f_i}$ donc $E_x > E_{f_{i-1}}$ ou $E_x > E_{f_i}$.

Il en résulte que $\ln \left(1 + \exp \left(-\frac{E_x - E_{f_i}}{k_B T} \right) \right) \sim \exp \left(-\frac{E_x - E_{f_i}}{k_B T} \right)$.

L'équation 3.12 peut s'écrire de la manière suivante :

$$\begin{aligned} J_{i-1,i} &= \alpha \int_{\max(E_{t_{i-1}} - V_{t_{i-1}}, E_{t_i} - V_{t_i})}^{\Phi_l - V_{t_{i-1}}} \mathcal{T}(E_x) \left(\exp \left(-\frac{E_x - E_{f_{i-1}}}{k_B T} \right) - \exp \left(-\frac{E_x - E_{f_i}}{k_B T} \right) \right) dE_x \\ &= \alpha \int_{\max(E_{t_{i-1}} - V_{t_{i-1}}, E_{t_i} - V_{t_i})}^{\Phi_l - V_{t_{i-1}}} \left(\mathcal{T}(E_x) \exp \left(\frac{-E_x}{k_B T} \right) \right) \left(\exp \left(\frac{E_{f_{i-1}}}{k_B T} \right) - \exp \left(\frac{E_{f_i}}{k_B T} \right) \right) dE_x \end{aligned} \quad (3.14)$$

L'équation ainsi obtenue est l'équation de courant entre deux piège dans l'approxi-

mation de Boltzmann. La reformulation de cette équation donnera :

$$\begin{aligned}
 J_{i-1,i} &= \alpha \int_{\max(E_{t_{i-1}} - V_{t_{i-1}}, E_{t_i} - V_{t_i})}^{\Phi_l - V_{t_{i-1}}} \left(\mathcal{T}(E_x) \exp\left(\frac{-E_x}{k_B T}\right) \right) \left(\exp\left(\frac{E_{f_{i-1}}}{k_B T}\right) - \exp\left(\frac{E_{f_i}}{k_B T}\right) \right) dE_x \\
 &= \alpha \int_{\max(E_{t_{i-1}} - V_{t_{i-1}}, E_{t_i} - V_{t_i})}^{\Phi_l - V_{t_{i-1}}} \left(\mathcal{T}(E_x) \exp\left(\frac{-E_x}{k_B T}\right) \right) (X_{i-1} - X_i) dE_x \\
 &= A_{i-1,i} X_{i-1} - A_{i-1,i} X_i
 \end{aligned} \tag{3.15}$$

$$\text{Avec } \begin{cases} X_i = \exp\left(\frac{E_{f_i}}{k_B T}\right) \\ A_{i-1,i} = \alpha \int_{\max(E_{t_{i-1}} - V_{t_{i-1}}, E_{t_i} - V_{t_i})}^{\Phi_l - V_{t_{i-1}}} \left(\mathcal{T}(E_x) \exp\left(\frac{-E_x}{k_B T}\right) \right) dE_x \end{cases}$$

Il est à noter que les bornes de l'intégrale sont entre le $\max(E_{t_{i-1}} - V_{t_{i-1}}, E_{t_i} - V_{t_i})$, c'est à dire le premier niveau d'énergie où un état est disponible. Ces bornes sont schématisés sur les figures 3.4 et 3.5, pour les deux cas à savoir quand $E_{t_i} < E_{t_{i-1}}$ et quand $E_{t_i} > E_{t_{i-1}}$.

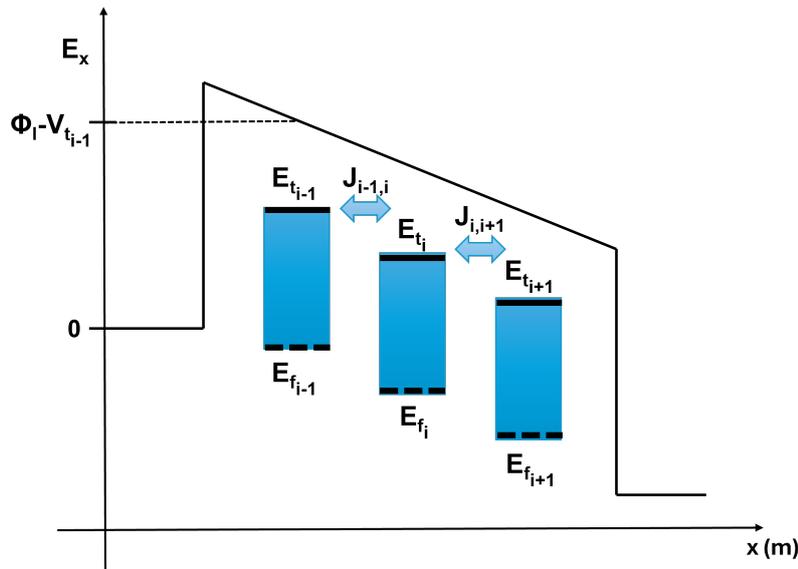


FIGURE 3.4 – Cas où $E_{t_i} < E_{t_{i-1}}$: Diagramme de bande simplifié sur trois pièges expliquant le modèle TE simplifié. Φ_l représente la hauteur de barrière, E_{f_i} (traits en pointillés) est le niveau de Fermi au niveau du piège i , E_{t_i} (traits plein) est le niveau d'énergie au niveau du piège i , V_{t_i} représente la tension au niveau du piège i .

Une fois cette expression obtenue, il sera important de noter que l'inconnu dans cette équation est le niveaux de Fermi et donc au travers du paramètre X . Pour tenter de trouver les niveaux de Fermi, il sera nécessaire d'étudier la continuité du courant. En utilisant le fait qu'on soit dans un état stationnaire, et donc que $\text{div}(j) = 0$, une relation entre le courant entre le piège i , le piège qui le suit $i+1$ et piège précédent $i-1$ peut être établie. Cette relation peut se formuler de telle sorte que $j_{i,i+1} = j_{i,i-1}$.

$$j_{i-1,i} = j_{i,i+1} \implies A_{i-1,i} X_{i-1} - A_{i-1,i} X_i = A_{i,i+1} X_i - A_{i,i+1} X_{i+1} \tag{3.16}$$

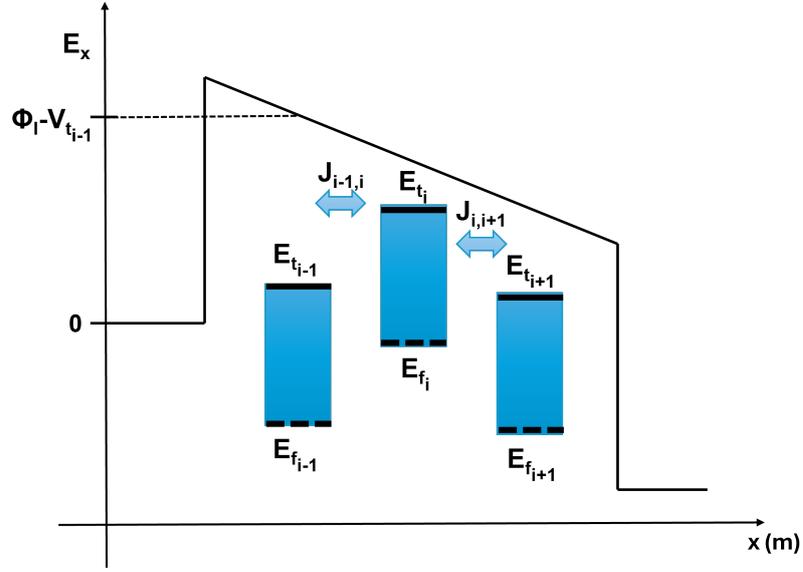


FIGURE 3.5 – Cas où $E_{t_i} > E_{t_{i-1}}$: Diagramme de bande simplifié sur trois pièges expliquant le modèle TE simplifié. Φ_i représente la hauteur de barrière, E_{f_i} (traits en pointillés) est le niveau de Fermi au niveau du piège i, E_{t_i} (traits pleins) est le niveau d'énergie au niveau du piège i, V_{t_i} représente la tension au niveau du piège i.

Ce qui permettra d'écrire pour chaque piège i l'équation suivante :

$$A_{i-1,i}X_{i-1} - (A_{i-1,i} + A_{i,i+1})X_i + A_{i,i+1}X_{i+1} = 0 \quad (3.17)$$

Ainsi contrairement au cas générale, le système est linéaire et donc solvable sans passer par une méthode de résolution numérique itérative. Il suffira de fixer les conditions aux limites du problèmes afin de pouvoir résoudre complètement le système. Connaissant les niveaux de Fermi de l'anode et de la cathode, il sera aisé de calculer les paramètres F correspondant (noté respectivement X_a et X_c). Ces deux derniers paramètres sont nos conditions aux limites fixes. L'équation 3.17, pourra s'écrire sous forme matriciel de la manière suivante :

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & A_{i-1,i} & -(A_{i-1,i} + A_{i,i+1}) & A_{i,i+1} & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_c \\ \vdots \\ X_i \\ \vdots \\ X_a \end{pmatrix} = \begin{pmatrix} X_c \\ \vdots \\ 0 \\ \vdots \\ X_a \end{pmatrix} \quad (3.18)$$

Une fois cette matrice tridiagonale inversée, le vecteur X est obtenu. Il sera ainsi assez simple d'extraire les niveaux de Fermi pour chaque piège et de les réinjecté dans l'équation 3.12 et ainsi obtenir le courant entre deux piège pour l'approche Nearest Neighbour avec le modèle de Tsu-Esaki. En effet, une fois le vecteur X

obtenu, il suffit d'appliquer la formule suivante :

$$E_f = k_B T \ln(X) \quad (3.19)$$

Dans l'approche Variable Range, la conduction est divisée en différentes sections S (voir figure 3.3). Dans chaque section S considérée, on suppose que le courant est la somme de toutes les contributions qui passent d'un côté à l'autre de cette section. Un système linéaire de taille $(N_{trap} + 2) \times (N_{trap} + 2)$ est utilisé pour résoudre les niveaux de Fermi (équation 3.20) dans cette approche. Pour calculer le courant, il faut additionner toutes les contributions qui précèdent la section S. D'un point de vue mathématique le système reste le même, à la différence près que la matrice contenant les coefficient $A_{i,n}$, ne sera plus une matrice tridiagonale mais une matrice pleine. Le système pourra ainsi s'écrire de la manière suivante :

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ A_{0,i} & \ddots & A_{i-1,i} & -\sum_{n=0}^{n=N_{trap}+2} A_{i,n} & A_{i,i+1} & \ddots & A_{i,N_{trap}+2} \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_c \\ \vdots \\ X_i \\ \vdots \\ X_a \end{pmatrix} = \begin{pmatrix} X_c \\ \vdots \\ 0 \\ \vdots \\ X_a \end{pmatrix} \quad (3.20)$$

Après avoir présenté la partie modélisation, examinons maintenant les résultats obtenus une fois le code implémenté.

Validation modèle

La figure 3.6 illustre la densité de courant en fonction de la position pour une tension de $V=0,5$ V, en utilisant l'approche NN, avec 20 pièges. On observe que le courant reste constant, montrant ainsi que la condition de continuité du courant a été vérifié.

En revanche, les niveaux de Fermi (E_f) ne présentent pas une relation linéaire avec la position (figure 3.7), contrairement à l'hypothèse courante dans de nombreux modèles TAT, tels que celui de l'université de Modena [12]. Cela met en évidence la nécessité de résoudre les niveaux de Fermi, comme démontré par Ielmini [13].

La figure 3.8 (a) offre un aperçu du courant en fonction de la tension pour six pièges, représentant l'état HRS. La courbe adopte une forme exponentielle, suggérant une conduction de type Trap-Assisted Tunneling (TAT). En revanche, la figure 3.8 (b) présente le courant en fonction de la tension pour 25 pièges, correspondant à l'état LRS. Dans ce cas, le courant est huit ordres de grandeur supérieur à celui de la figure 3.8 (a), et la courbe adopte une allure linéaire, suggérant une conduction de type ohmique.

Grâce à ce modèle, le comportement en I-V d'une OxRAM a été reproduit, tant en HRS qu'en LRS, avec une dépendance exponentielle en HRS et linéaire en LRS.

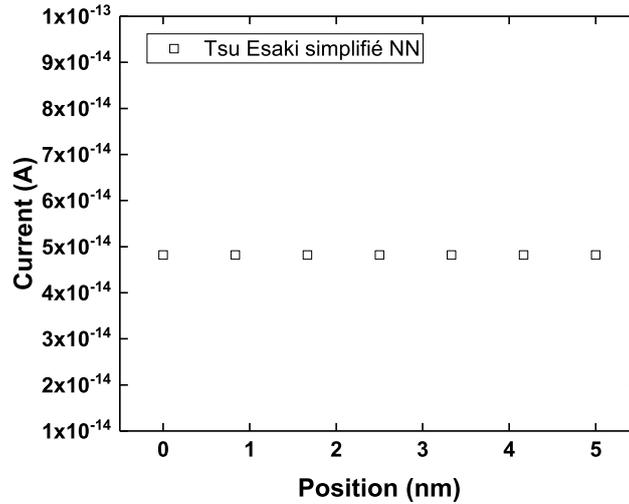


FIGURE 3.6 – Courant en en fonction de la position des pièges en utilisant le modèle TE simplifié Boltzmann. Les paramètres utilisés : $N_t=5$, $E_t=0.2$ eV et $V=0.5$ V.

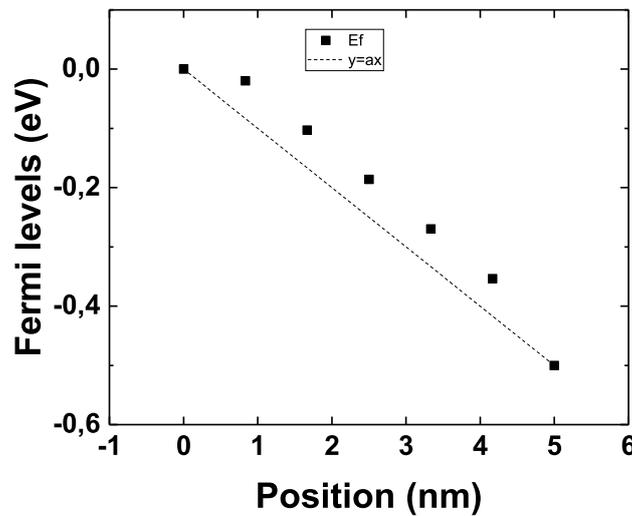


FIGURE 3.7 – Niveau de Fermi résolu en fonction de la position des pièges. Les paramètres utilisés : $N_t = 5$, $E_t = 0.2$ eV et $V=0.5$ V

Contrairement aux modèles existants dans la littérature, qui suggèrent généralement de modéliser le LRS par un modèle de conduction métallique et le HRS par un TAT, utilisant ainsi deux formalismes de transport distincts, le modèle développé au cours de cette thèse ne nécessite pas une telle contrainte. Un seul formalisme de transport explique efficacement les deux régimes.

Cependant, il est crucial de noter que les valeurs de courant obtenues restent considérablement inférieures aux valeurs typiques observées dans les OxRAM, au-

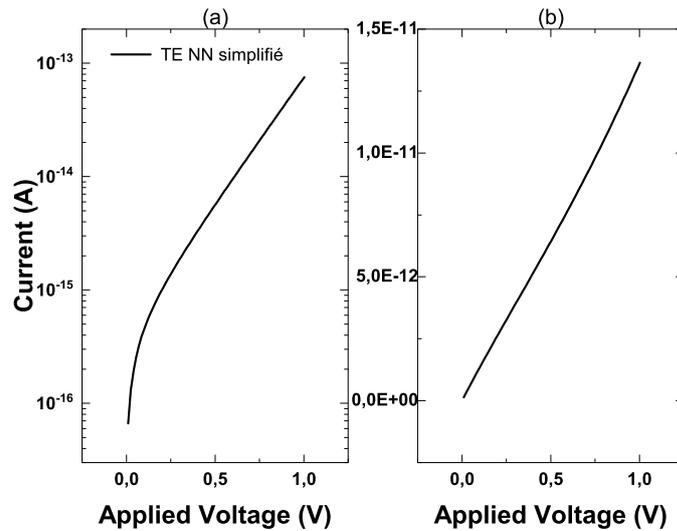


FIGURE 3.8 – (a) Courant en fonction de la tension avec 6 piège (HRS) (b) Courant en fonction de la tension avec 25 pièges (LRS). Les paramètres utilisée : $E_t = 0.2$ eV

tour de 10^{-6} A pour l'état HRS et 10^{-4} A pour l'état LRS. Bien que l'ajustement raisonnable des paramètres physiques puisse accroître ces courants, cela ne suffit pas à atteindre les valeurs de courant recherchées. Par conséquent, une alternative a été envisagée, consistant à ne pas restreindre le transport aux pièges les plus proches, mais à permettre le transport vers des pièges plus éloignés. C'est l'approche Variable Range qui a déjà été expliquée. Comme illustré sur la figure 3.9 et la figure 3.10, avec l'application de l'approche VR, les niveaux de courant atteints sont certes plus significatifs en LRS, mais on observe également un effet non linéaire à tension élevée. La courbe du courant traité en VR se situe au-dessus de celle traitée en NN. Cependant, on observe qu'à 0.8 V pour le VR et à 1 V pour le NN, le courant LRS présente une déviation par rapport à une ligne droite, adoptant une dépendance parabolique. Ce phénomène a déjà été documenté dans la littérature, notamment dans [18] (voir figure 3.11), où une mesure en LRS sans transistor de sélection a révélé une non-linéarité de la courbe à partir de 0.8 V.

La présence de cette non-linéarité s'explique par le fait qu'à de faibles tensions entre deux pièges, particulièrement lorsque de nombreux pièges sont présents le long du chemin de conduction (donc pour de très petits dV entre pièges), le développement limité à l'ordre 1 en tension demeure linéaire (voir équation 3.21). En revanche, à des polarisations plus élevées et toujours entre deux pièges, l'ordre 2 devient non négligeable, induisant ainsi l'effet quadratique (voir équation 3.22).

$$J(V) = \alpha \int \mathcal{T}(E)f(E)dE \sim \exp\left(\frac{qV}{k_B T}\right) \sim 1 + \frac{qV}{k_B T} \quad (3.21)$$

$$J(V) = \alpha \int \mathcal{T}(E)f(E)dE \sim \exp\left(\frac{qV}{k_B T}\right) \sim 1 + \frac{qV}{k_B T} + \frac{1}{2} \left(\frac{qV}{k_B T}\right)^2 \quad (3.22)$$

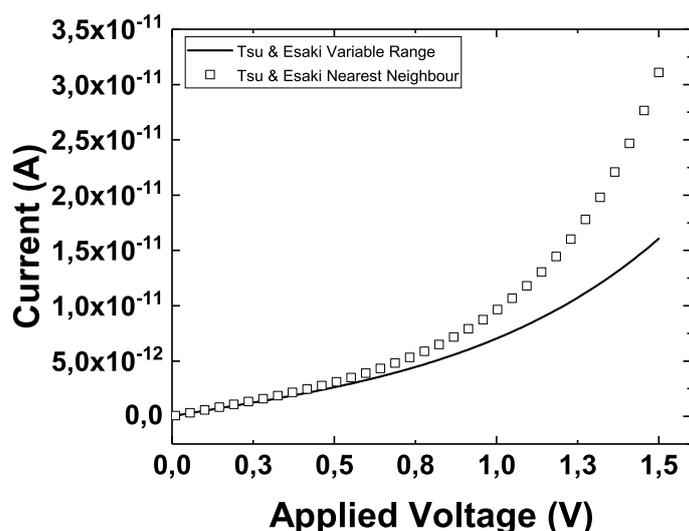


FIGURE 3.9 – Courbe courant en fonction de la tension pour le modèle TE simplifié Boltzmann avec l’approche NN et VR en LRS. Les paramètres utilisés : $N_t=20$, $E_t=0.2$ eV

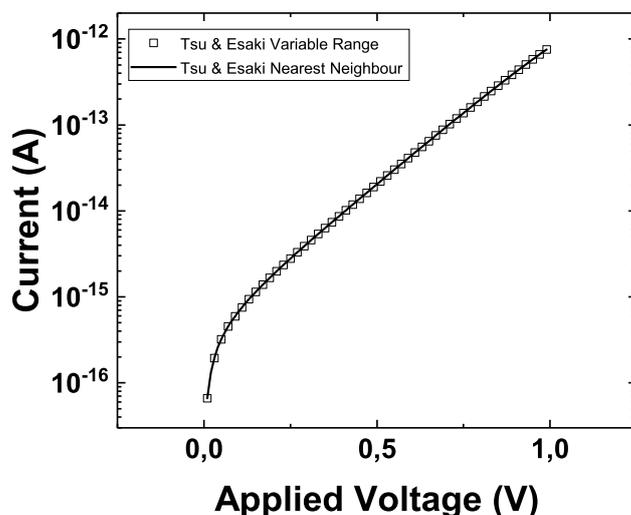


FIGURE 3.10 – Courbe courant en fonction de la tension pour le modèle TE simplifié Boltzmann avec l’approche NN et VR en HRS. Les paramètres utilisés : $N_t=6$, $E_t=0.2$ eV

En HRS, comme le montre la courbe 3.10, aucun de ces effets n’est visible et les deux courbes NN et VR sont identiques. Ceci est due au fait qu’en HRS, l’approche VR n’apporte pas une contribution significative au vu le faible nombre de pièges dans cet état de résistance. En plus des validations habituelles effectuées sur ce modèle, nous avons également tenté de modifier certains paramètres, tels que la

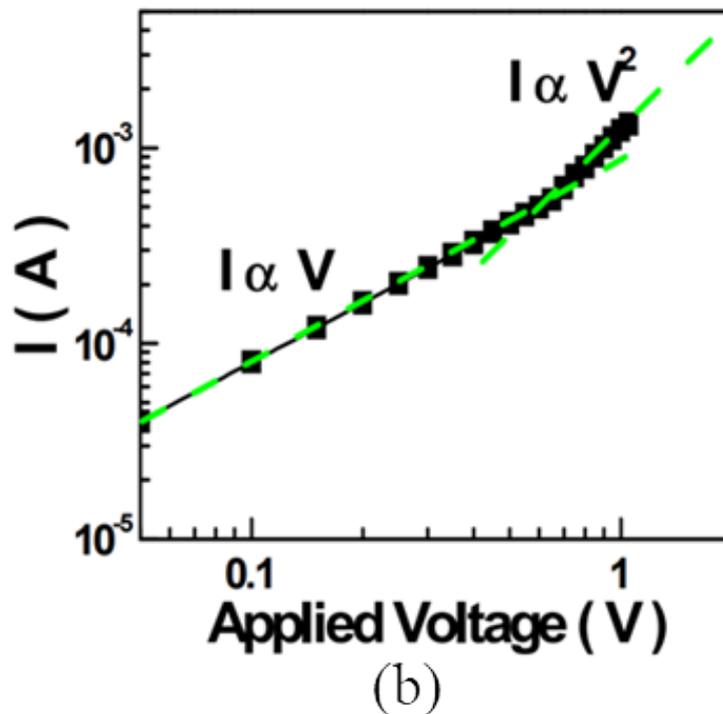


FIGURE 3.11 – Courbe courant en fonction de la tension en LRS sur un dispositif 1R mesurée issu de [18]

hauteur de la barrière (figure 3.14), l'épaisseur de l'oxyde (figure 3.12) et l'énergie des pièges (figure 3.12), afin de déterminer si nous pouvons atteindre les niveaux de courant requis pour une OxRAM. Les résultats obtenus pour l'état LRS avec 25 pièges sont présentés. On constate que, même en ajustant ces paramètres, les niveaux de courant restent inférieurs aux attentes. Malgré toutes les tentatives déployées pour obtenir des courants représentatifs des mesures effectuées dans la littérature, les courants modélisés demeurent encore trop faibles. Par conséquent, il devient impératif d'explorer un autre modèle.

3.3.2 Formulation du modèle de TE en dégénéré

Modélisation

Afin d'inclure la dégénérescence des porteurs, on reprend l'approche TE décrite dans la section précédente et on considère que les niveaux de Fermi sont au-dessus des niveaux des pièges ($E_f > E_t$). On considère que la fonction de distribution est la fonction de Fermi-Dirac tout en faisant l'hypothèse de dégénérescence. L'approximation de Boltzmann qui a été faite précédemment est impossible car dans ce cas $E_f > E_t$, tandis que pour que l'approximation de Boltzmann soit vraie il est nécessaire que $E_f \ll E_t$. En reprenant l'équation de Tsu-Esaki, avec comme fonction d'occupation la fonction de Fermi-Dirac, le courant entre deux pièges $i-1$ et i sera

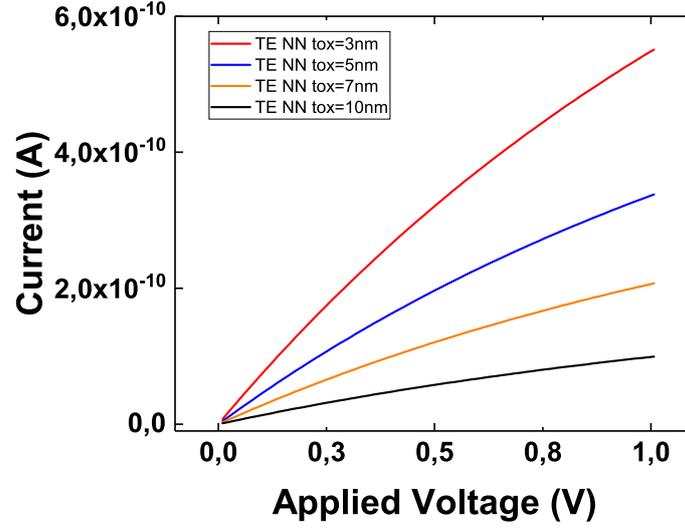


FIGURE 3.12 – Courbe courant en fonction de la tension en LRS pour différents t_{ox} . Les paramètres utilisés : $E_t=0.2$ eV, $N_t=25$, $\Phi_l=3.1$ eV.

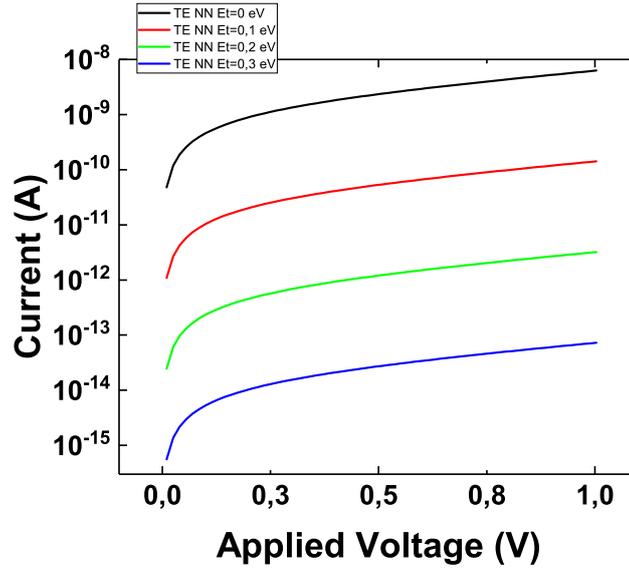


FIGURE 3.13 – Courbe courant en fonction de la tension en LRS pour différents E_t . Les paramètres utilisés : $\Phi_l=3.1$ eV, $N_t=25$, $t_{ox}=5$ nm

donné par :

$$\begin{aligned}
 J_{i-1,i} = & \alpha \int_{E_{t_{i-1}}}^{E_{f_{i-1}}} \mathcal{T}(E_x) \ln \left(1 + \exp \left(-\frac{E_x - E_{f_{i-1}}}{k_B T} \right) \right) dE_x \\
 & - \alpha \int_{E_{t_{i-1}}}^{E_{f_i}} \mathcal{T}(E_x) \ln \left(1 + \exp \left(-\frac{E_x - E_{f_i}}{k_B T} \right) \right) dE_x
 \end{aligned} \tag{3.23}$$

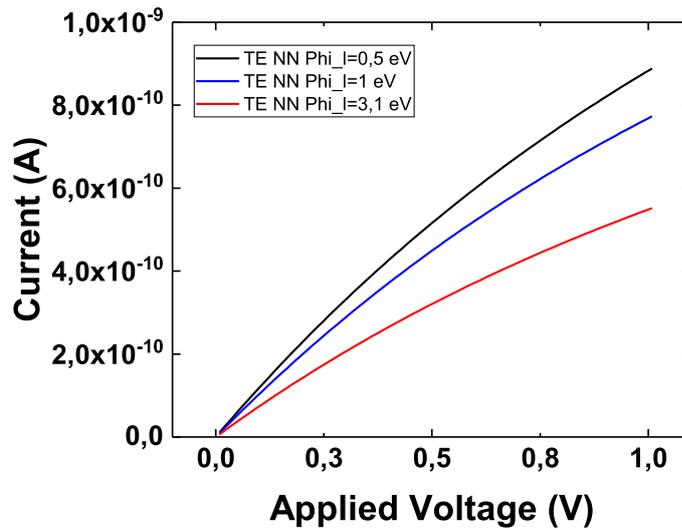


FIGURE 3.14 – Courbe courant en fonction de la tension en LRS pour différents Φ_l . Les paramètres utilisés : $E_t=0.2$ eV et $N_t=25$, $t_o x=5$ nm

Sachant que $E_f > E_t$ car niveau des pièges dégénéré, il en résulte que $-\frac{E_x - E_{f_{i-1}}}{k_B T} \gg 1$ et $-\frac{E_x - E_{f_i}}{k_B T} \gg 1$. Il est à noter que qu'un certains nombre de nombre de contribu-

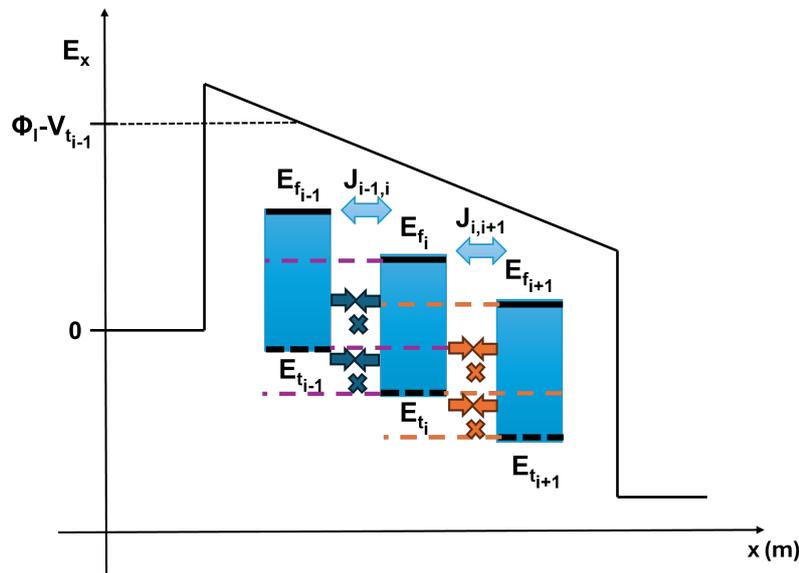


FIGURE 3.15 – Diagramme de bande simplifié sur trois pièges expliquant le modèle TE dégénéré. Φ_l représente la hauteur de barrière, E_{f_i} (traits en pointillés) est le niveau de Fermi au niveau du piège i , E_{t_i} (traits pleins) est le niveau d'énergie au niveau du piège i , V_{t_i} représente la tension au niveau du piège i .

tions va s'annuler, comme le montre le diagramme de bande 3.15. Ainsi, l'équation

3.23 peut se réécrire :

$$J_{i-1,i} = \frac{\alpha}{k_B T} \int_{E_{t_{i-1}}}^{E_{f_{i-1}}} \mathcal{T}(E_x)(-E_x + E_{f_{i-1}})dE_x - \frac{\alpha}{k_B T} \int_{E_{t_{i-1}}}^{E_{f_i}} \mathcal{T}(E_x)(-E_x + E_{f_i})dE_x \quad (3.24)$$

Pour tenter d'avoir une solution analytique et aussi une meilleure approximation du courant, le potentiel entre deux pièges est considéré comme parabolique. Donc la transparence $\mathcal{T}(E_x)$ qui va être considéré sera l'approximation WKB d'une barrière parabolique. Comme ce calcul n'est pas tout à fait courant dans la littérature, il sera exposé succinctement dans ce qui suit.

Transparence WKB d'une barrière parabolique

La transparence WKB (Wentzel-Kramers-Brillouin) d'une barrière est une mesure de la probabilité qu'une particule traverse une barrière de potentiel dans le cadre de la mécanique quantique. La méthode WKB est une approximation semi-classique qui permet d'estimer de manière très correcte la transparence d'une barrière. Pour calculer la transparence WKB d'une barrière, on considère l'équation de Schrödinger pour la particule se déplaçant dans le potentiel de la barrière. L'idée est d'approximer la solution de cette équation en utilisant des solutions classiques dans les régions où le potentiel est élevé et en appliquant des conditions aux limites appropriées. La transparence WKB dépend de plusieurs facteurs, tels que la hauteur de la barrière, l'énergie de la particule, la largeur de la barrière et la masse de la particule. En général, plus la barrière est haute et large, plus la transparence WKB est faible, ce qui signifie qu'il est moins probable que la particule traverse la barrière.

La formule générale pour la transparence WKB est donnée par :

$$\mathcal{T}(E) = \exp\left(\frac{-2}{\hbar} \int_{x_1}^{x_2} \sqrt{2m(V(x) - E)}dx\right) \quad (3.25)$$

où \mathcal{T} est la transparence, m est la masse de la particule, $V(x)$ est l'énergie de la barrière, E est l'énergie de la particule, x_1 et x_2 sont les positions classiques de rebond de la particule sur les bords de la barrière et \hbar est la constante réduite de Planck.

Dans notre cas V est connu, et est l'énergie potentiel d'une barrière parabolique et a la forme suivante :

$$V(x) = \Phi_l - \beta(x - x_{1/2})^2 - \frac{qV_{app}}{2x_{1/2}}x \quad (3.26)$$

Avec Φ_l la hauteur de barrière, $x_{1/2}$ la position de la hauteur maximum de la barrière, V_{app} la tension appliqué et β est la courbure du potentiel. Après quelques lignes de calculs, on arrive au résultat suivant :

$$\mathcal{T}(E) = \mathcal{T}_0 \exp\left(\frac{\pi}{\hbar} \sqrt{\frac{2m}{\beta}} E\right) \quad (3.27)$$

Avec \mathcal{T}_0 un paramètre dépendant de la polarisation, de la hauteur de barrière et de la courbure du potentiel. Dans l'approche Nearest Neighbour, le courant pourra alors s'écrire de la manière suivante :

$$J_{i-1,i} = \gamma \exp(\kappa E_{t_i})(E_{f_i} - E_{f_{i-1}}) + \frac{\gamma}{\kappa}(\exp(\kappa E_{f_{i-1}}) - \exp(\kappa E_{f_i})) \quad (3.28)$$

Avec $\kappa = \frac{\sqrt{2m}}{\sqrt{\alpha\hbar}}$ et $\gamma = \frac{\alpha\mathcal{T}_0}{\kappa k_B T}$

Le processus de calcul de courant est similaire à celui utilisé pour le modèle TE non dégénéré. Nous supposons un état stationnaire où le courant est constant le long de l'oxyde. Afin de calculer ce courant, il est nécessaire de déterminer les niveaux de Fermi. Pour ce faire, nous examinons la continuité du courant, ce qui implique de calculer la différence entre $J_{i-1,i}$ et $J_{i,i+1}$ et de constater que cette différence est nulle. Ce qui permettra d'écrire pour chaque piège i l'équation suivante :

$$\begin{aligned} f(E_{f_i}) = J_{i,i+1} - J_{i-1,i} = & \frac{-2}{\kappa} \exp(\kappa E_{f_i}) + (\exp(\kappa E_{t_{i-1}}) + \exp(\kappa E_{t_i})) E_{f_i} \\ & - E_{f_{i-1}} \exp(\kappa E_{t_{i-1}}) - E_{f_{i+1}} \exp(\kappa E_{t_i}) + \frac{1}{\kappa} (\exp(\kappa E_{f_{i-1}}) + \exp(\kappa E_{f_{i+1}})) \end{aligned} \quad (3.29)$$

En regardant de près ce système d'équations, il est facilement remarquable que ce dernier est non-linéaire. Il sera donc nécessaire d'utiliser la méthode de Newton Raphson précédemment décrite afin de résoudre le système et trouver enfin les niveaux de Fermi. Par conséquent, le vecteur F qui contient tous les éléments qui doivent être nuls prendra la forme suivante :

$$F = \begin{pmatrix} J_{1,2}(E_{f_1}, E_{f_2}) - J_{c,1}(E_{f_c}, E_{f_1}) \\ \vdots \\ J_{i,i+1}(E_{f_i}, E_{f_{i+1}}) - J_{i-1,i}(E_{f_{i-1}}, E_{f_i}) \\ \vdots \\ J_{N_t,a}(E_{f_{N_t}}, E_{f_a}) - J_{N_t-1,N_t}(E_{f_{N_t-1}}, E_{f_{N_t}}) \end{pmatrix} \quad (3.30)$$

avec $J_{i,j}$ l'expression du courant donnée par l'équation 3.28.

Après avoir calculé le vecteur F qui contient toutes les équations à résoudre pour le système, la prochaine étape consiste à calculer la matrice jacobienne du système.

Pour calculer la matrice jacobienne, il est nécessaire de calculer les dérivées partielles de la fonction F par rapport à chaque variable inconnue. Dans ce cas, les variables inconnues sont $E_{f_{i-1}}$, E_{f_i} et $E_{f_{i+1}}$.

Après quelques lignes de calcul, on arrive aux expressions suivantes :

$$\begin{aligned} \frac{\partial f(E_{f_i})}{\partial E_{f_i}} &= -2 \exp(\kappa E_{f_i}) + \exp(\kappa E_{t_{i-1}}) + \exp(\kappa E_{t_i}) \\ \frac{\partial f(E_{f_i})}{\partial E_{f_{i-1}}} &= \exp(\kappa E_{f_{i-1}}) - \exp(\kappa E_{t_{i-1}}) \\ \frac{\partial f(E_{f_i})}{\partial E_{f_{i+1}}} &= \exp(\kappa E_{f_{i+1}}) - \exp(\kappa E_{t_i}) \end{aligned} \quad (3.31)$$

Maintenant que nous avons les expressions des dérivées partielles de la fonction F par rapport à chaque variable inconnue, nous pouvons utiliser l'expression de la matrice jacobienne en Nearest Neighbour 3.5 pour calculer la matrice jacobienne complète. Une fois cette dernière calculer, il suffit d'appliquer l'algorithme de Newton-Raphson et enfin déterminer les niveaux de Fermi, qui permettront par la suite de calculer les différents courants. Pour l'approche Variable Range, on adopte le même processus que dans l'approche Nearest Neighbour. La seule différence est l'expression de la forme du courant et donc de l'expression de la continuité du courant mais aussi de la matrice jacobienne. Ainsi, en reprenant l'expression du courant en Nearest Neighbour, et en appliquant l'approche Variable Range, le courant en dégénéré dans l'approche Variable Range aura la forme suivante :

$$J_{is} = \gamma \sum_{i=0}^{i-1} \sum_{j=i}^{Nt} \left(\exp(\kappa E_{t_i}) (\exp(\kappa E_{f_j}) - \exp(\kappa E_{f_i})) + \frac{\exp(\kappa E_{f_j}) + \exp(\kappa E_{f_i})}{\kappa} \right) \quad (3.32)$$

De la même manière qu'avec l'approche Nearest Neighbour, l'étude de la continuité du courant est nécessaire. Ainsi, la continuité du courant qui sera donné par la fonction f , pourra s'écrire de la manière suivante :

$$f(E_{f_i}) = \sum_{j=0}^i \left(\exp(\kappa E_{t_i}) (E_{f_{i+1}} - E_{f_j}) - \frac{\exp(\kappa E_{f_j}) - \exp(\kappa E_{f_{i+1}})}{\kappa} \right) \quad (3.33)$$

Après avoir calculé le vecteur F qui contient toutes les équations à résoudre pour le système, la prochaine étape consiste à calculer la matrice jacobienne du système.

Pour calculer la matrice jacobienne, il est nécessaire de calculer les dérivées partielles de la fonction F par rapport à chaque variable inconnue. Dans ce cas, les variables inconnues sont $E_{f_{i-1}}$, E_{f_i} et $E_{f_{i+1}}$.

Après quelques lignes de calcul, on arrive aux expressions suivantes :

$$\frac{\partial f(E_{f_i})}{\partial E_{f_i}} = \begin{cases} \exp(\kappa E_{t_i}) + \exp(\kappa E_{f_i}) & \text{si } i \leq j \\ \exp(\kappa E_{t_i}) - \exp(\kappa E_{f_{j+1}}) & \text{si } i = j + 1 \\ 0 & \text{sinon} \end{cases} \quad (3.34)$$

Validation modèle

Dans les modélisations précédentes du Tsu-Esaki, le niveau d'énergie des pièges est généralement bien au-dessus des niveaux de Fermi, correspondant ainsi à un régime non dégénéré du gaz de porteurs, c'est-à-dire le régime de Boltzmann. La tentative de modélisation basée sur le modèle de Tsu-Esaki N-piège a donc impliqué l'hypothèse selon laquelle le niveau de Fermi des pièges est largement supérieur au niveau du piège, entraînant ainsi le gaz de porteurs dans un régime dégénéré.

Pour évaluer si l'équation de continuité a été correctement respectée, le courant en fonction de la position a été tracé (figure 3.17). Comme souligné précédemment

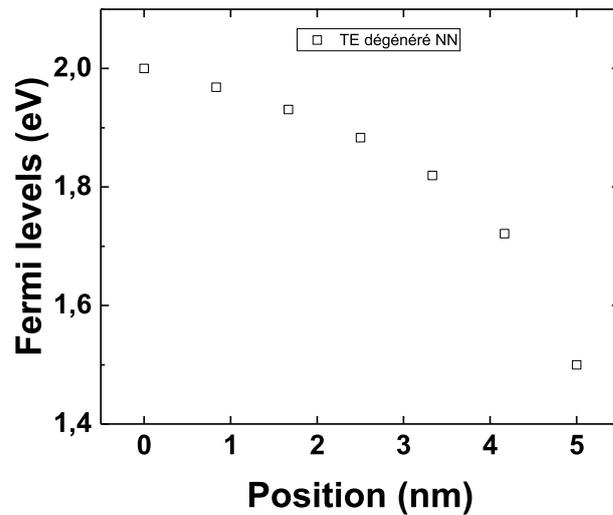


FIGURE 3.16 – Courbe niveaux de Fermi en fonction de la position en TE dégénéré en utilisant l’approche NN. Les paramètres utilisées : $N_t=5$, $E_t=0.2$ eV, $V=0.5$ V

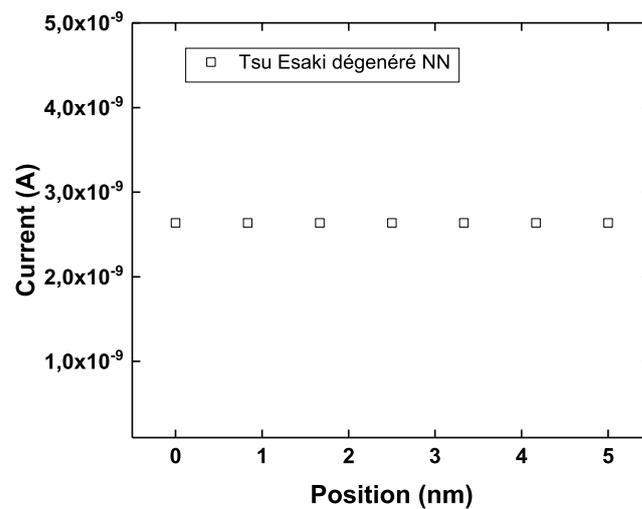


FIGURE 3.17 – Courbe courant en fonction de la position en TE dégénéré en utilisant l’approche NN. Les paramètres utilisés : $N_t=5$, $E_t=0.2$ eV, $V=0.5$ V

pour le modèle non dégénéré, le courant en fonction de la position reste constant, confirmant ainsi le respect de l’équation de continuité. Les niveaux de Fermi des pièges ont également été représentés en fonction de la position des pièges (voir figure 3.16), mettant en évidence que l’allure des E_f en fonction de la position n’est pas linéaire. Cette observation justifie une nouvelle fois la nécessité de résoudre des niveaux de Fermi plutôt que de les considérer comme linéaire en fonction de la position, hypothèse courante dans de nombreuses publications [12, 83, 84].

Afin de vérifier si la transition entre l'état HRS et LRS est assurée par le modèle dégénéré, les courbes courant-tension ont été tracées pour le LRS et le HRS, comme illustré en figure 3.18.

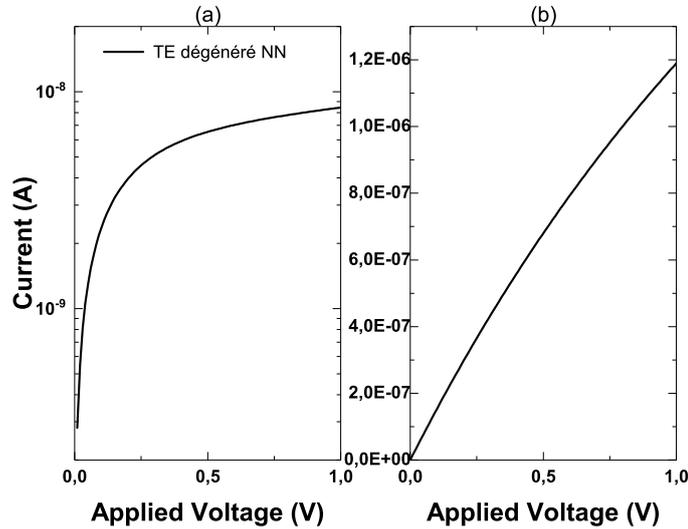


FIGURE 3.18 – Courbe courant en fonction de la tension pour le modèle dégénéré en (a) HRS et en (b) LRS. Les paramètres utilisés : (a) $N_t=6$, $E_t=0.2$ eV et (b) $N_t=25$, $E_t=0.2$ eV

On remarque que dans l'état HRS, la courbe présente une allure exponentielle, tandis que dans l'état LRS, l'allure est nettement linéaire. Par conséquent, on peut conclure que la transition entre l'état HRS et LRS est assurée par le modèle non dégénéré. Une fois que le modèle a démontré qu'il respecte les hypothèses préliminaires, il devient nécessaire de procéder à une comparaison entre les niveaux de courant en régime dégénéré et non dégénéré.

La figure 3.19 illustre les courbes de courant en fonction de la tension pour le modèle dégénéré et non dégénéré. On remarque qu'avec 6 pièges, représentant ainsi l'état HRS, le modèle dégénéré permet de gagner 4 ordres de grandeur par rapport au modèle non dégénéré. .

Le graphique montrant la courbe I-V pour les modèles dégénéré et non dégénéré est présentée dans la figure 3.20. On observe qu'avec l'inclusion de 25 pièges, symbolisant l'état LRS, le modèle dégénéré affiche une amélioration de 4 ordres de grandeur par rapport au modèle non dégénéré. Cette observation revêt une importance significative en ce qui concerne l'ajustement des données expérimentales. Afin d'évaluer si les niveaux de courant escomptés sont atteints, il est impératif d'explorer l'approche VR pour le modèle dégénéré. La figure 3.21 illustre la courbe courant en fonction de la tension en NN et VR pour le HRS. En HRS, on remarque que les niveaux de courant sont presque les mêmes pour l'approche VR et NN à basse tension.

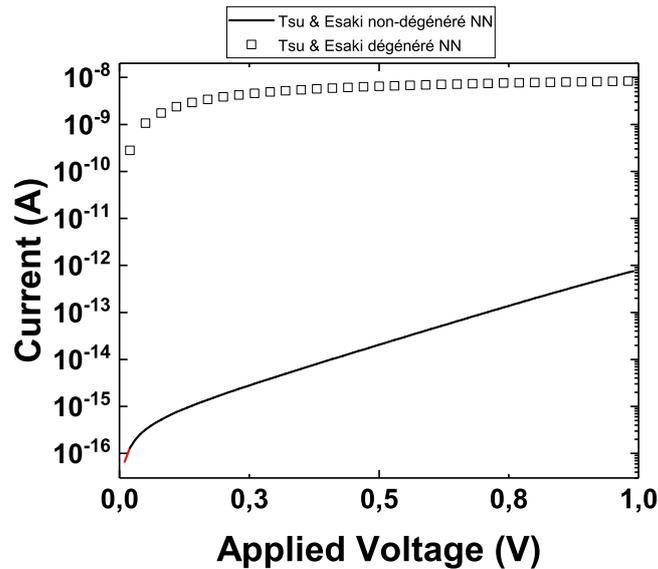


FIGURE 3.19 – Courbe courant en fonction de la tension pour le modèle dégénéré et non dégénéré en HRS. Les paramètres utilisées : $N_t=6$, $E_t=0.2$ eV

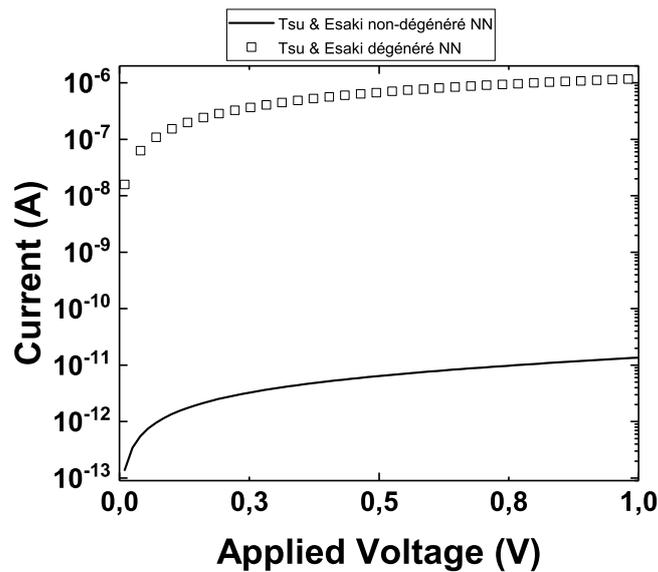


FIGURE 3.20 – Courbe courant en fonction de la tension pour le modèle dégénéré et non dégénéré en LRS. Les paramètres utilisées : $N_t=25$, $E_t=0.2$ eV

Cependant, à haute tension, le VR présente une forte augmentation, permettant au courant d'atteindre le seuil des 10^{-8} A.

Concernant le LRS qui est illustré en figure 3.22, on observe initialement une disparité importante de courant entre le NN et le VR, principalement due à la manière dont les courants sont calculés. En effet, le VR prend en compte le transport entre plusieurs pièges, de sorte que plus le nombre de pièges est important, plus le courant est élevé. À l'inverse, l'approche NN est insensible au nombre de pièges, car

elle considère uniquement le transport avec les voisins les plus proches. Ainsi, en LRS, on parvient à atteindre un niveau de courant significatif, à savoir 10^{-6} A.

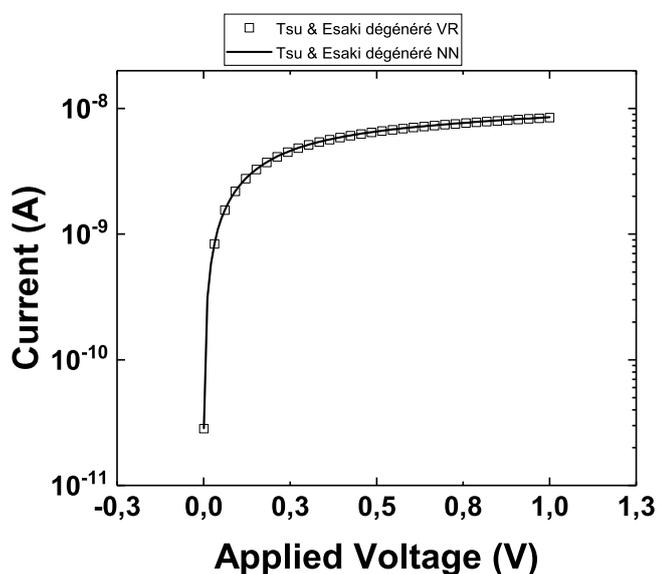


FIGURE 3.21 – Courbe courant en fonction de la tension pour le modèle dégénéré en HRS en utilisant l’approche VR. Les paramètres utilisés : $N_t=6$, $E_t=0.2$ eV

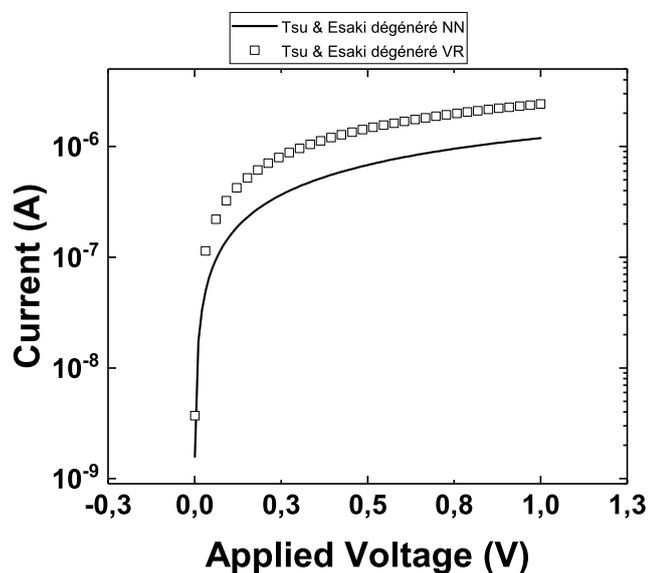


FIGURE 3.22 – Courbe courant en fonction de la tension pour le modèle dégénéré en LRS en utilisant l’approche NN et l’approche VR. Les paramètres utilisés : $N_t=25$, $E_t=0.2$ eV

Un autre aspect essentiel à examiner est le comportement en fonction de la température du modèle. Ainsi, le courant en HRS (3.23) et LRS (3.24) a été tracé en fonction de la tension à différentes températures, tant pour le modèle non dégé-

né (représenté par des étoiles sur les deux courbes) que pour le modèle dégénéré (représenté par une ligne sur les deux courbes).

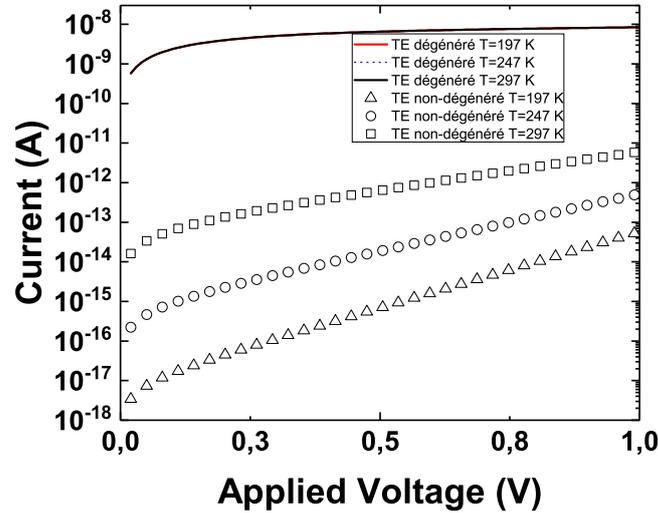


FIGURE 3.23 – Courbe courant en fonction de la tension pour le modèle dégénéré et non dégénéré en HRS en fonction de la température en utilisant l’approche NN. Les paramètres utilisés : $N_t=6$, $E_t=0.2$ eV

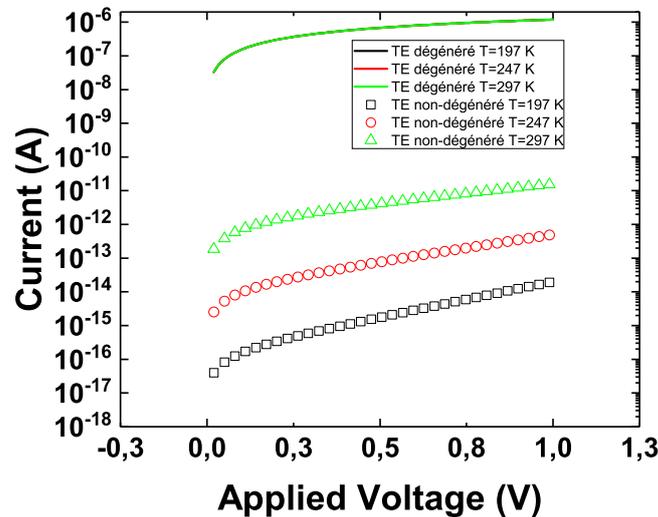


FIGURE 3.24 – Courbe courant en fonction de la tension pour le modèle dégénéré et non dégénéré en LRS en fonction de la température en utilisant l’approche NN. Les paramètres utilisés : $N_t=25$, $E_t=0.2$ eV

On observe que le modèle non dégénéré présente une activation en température, c’est-à-dire qu’entre 197 K et 298 K, il existe une différence d’un ordre de grandeur en

termes de courant. En revanche, pour le modèle dégénéré, le courant reste insensible aux variations de température, une caractéristique qui est également confirmée par l'équation 3.28 ne montrant aucune dépendance en température. Cette caractéristique revêt une importance cruciale, car la littérature indique que tant en LRS qu'en HRS, le courant présente une activation en température [19, 90, 98]. Par conséquent, on peut conclure que le modèle dégénéré ne peut reproduire le comportement I-V d'une OxRAM, étant donné qu'il ne peut pas prendre en compte la dépendance en température du courant.

Malgré toutes les tentatives déployées pour modéliser en utilisant la fonction de Boltzmann en tant que fonction d'occupation, il devient évident qu'il est impossible d'aboutir à une solution satisfaisante pour obtenir une modélisation précise de la caractéristique I-V. Par conséquent, il devient impératif d'adopter la fonction de Fermi-Dirac comme fonction d'occupation. Cette décision est motivée par le fait que, en plus de prendre en compte la forme dégénérée du modèle, la fonction de Fermi-Dirac intègre également la forme non dégénérée, offrant ainsi une approche plus complète et représentative.

3.3.3 Comparaison en fonction du nombre de piège des modèles TE Boltzmann

Dans la figure 3.25, deux courbes sont représentées, illustrant le courant en fonction du nombre de pièges pour le modèle dégénéré et le modèle non-dégénéré. Une comparaison entre les deux modèles révèle que les deux saturent au-delà d'un certain nombre de pièges. Cette saturation suggère l'existence d'un optimum en termes de nombre de pièges. En outre, il est important de souligner que l'idée d'avoir 100 pièges sur une épaisseur d'oxyde de 5 nm n'est pas réaliste, vu les densités de pièges annoncées dans la littératures [59, 58].

Pour la suite de cette thèse, le nombre de piège optimale en LRS sera considéré autour des 25 pièges.

3.3.4 Formulation du modèle TE dans le cadre générale

Modélisation

Le modèle de courant de tunneling de Tsu-Esaki [87] est un modèle de courant de tunneling entre deux continuums d'états 3D. Le modèle implémenté dans notre cas est une généralisation en N-pièges du modèle présenté par Ielmini et *al.* dans [13][14].

Dans le cadre général de la densité d'états 3D, la courant de Tsu-Esaki est donnée

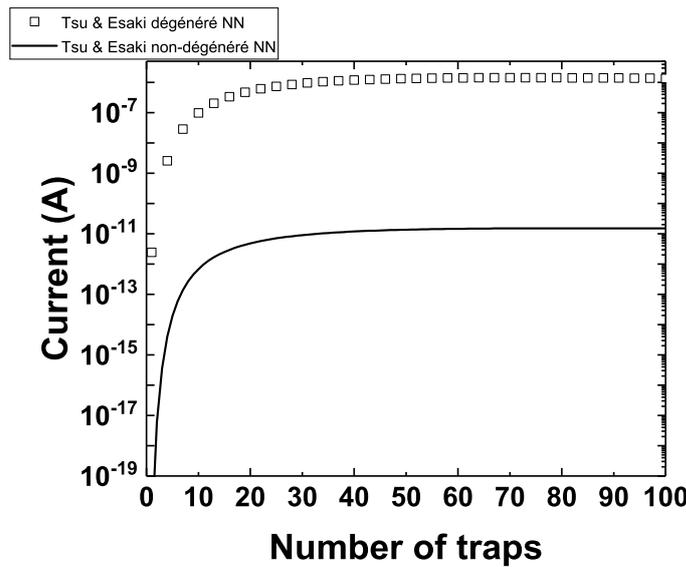


FIGURE 3.25 – Courbe courant en fonction du nombre de pièges en dégénéré et non dégénéré. Les paramètres utilisées : $E_t=0.2$ eV, $V = 1V$

par [87] :

$$J_{i-1,i} = \alpha \int_0^\infty \frac{1}{2\pi^3 \hbar} \frac{\partial E(k_x)}{\partial k_x} (f_{i-1} - f_i) \mathcal{T}(k_x) d^3 k_x \quad (3.35)$$

où $\alpha = \frac{4\pi q m_e f f}{\hbar^3}$, $\mathcal{T}(k_x)$ est la transparence de la barrière, $E(k_x)$ est la structure de bande, f_{i-1} est la fonction d'occupation du piège de départ, f_i est la fonction d'occupation du piège d'arrivée et k_x est le vecteur d'onde dans la direction x, de la cathode à l'anode.

Dans le cas actuellement traité, certaines simplifications sont effectuées. Les fonctions d'occupation sont toujours considérées comme des fonctions de Fermi-Dirac. La transparence est dérivée de l'approximation WKB sur une barrière trapézoïdale [99]. Le courant est donc donné par :

$$J_{i-1,i} = \alpha \int_a^b \mathcal{T}(E_x) \ln \left(\frac{1 + \exp\left(-\frac{E_x - E_{f_{i-1}}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)} \right) dE_x \quad (3.36)$$

Où $E_{f_{i-1}}$ est le niveau de Fermi du piège $i-1$, $\mathcal{T}(E_x)$ est la transparence WKB de la barrière trapézoïdale entre les deux pièges $i-1$ et i , $a = \Phi_l - qV \frac{x_i}{t_{ox}}$ et $b = \Phi_l - qV \frac{x_{i-1}}{t_{ox}}$, qui sont les limites de l'intégrale. Φ est la hauteur de la barrière et x_i est la position du piège.

Comme présenté précédemment, l'inconnu dans ces équations est le niveau de Fermi des pièges. Pour trouver les niveaux de Fermi en question, il sera nécessaire de commencer par étudier la continuité du courant, afin de trouver une solution en utilisant l'approche Nearest Neighbour. Dans cette approche, l'équation du courant sera l'équation 3.36. En réutilisant l'équation 3.1, et en injectant l'équation de courant

en Tsu-Esaki 3.36, on se retrouvera avec :

$$\begin{aligned}
 f(E_{f_i}) = J_{i,i+1} - J_{i-1,i} &= \alpha \int_c^a \mathcal{T}(E_x) \ln \left(\frac{1 + \exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_{i+1}}}{k_B T}\right)} \right) dE_x \\
 &- \alpha \int_a^b \mathcal{T}(E_x) \ln \left(\frac{1 + \exp\left(-\frac{E_x - E_{f_{i-1}}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)} \right) dE_x \\
 &= 0
 \end{aligned} \tag{3.37}$$

Avec $c = \Phi_l - qV \frac{x_{i+1}}{t_{ox}}$ Étant donné que le système $J_{i,i+1} - J_{i-1,i} = 0$ est non linéaire, il est nécessaire d'utiliser la méthode de Newton-Raphson. Par conséquent, le vecteur F qui contient tous les éléments qui doivent être nuls prendra la forme suivante :

$$F = \begin{pmatrix} J_{1,2}(E_{f_1}, E_{f_2}) - J_{c,1}(E_{f_c}, E_{f_1}) \\ \vdots \\ J_{i,i+1}(E_{f_i}, E_{f_{i+1}}) - J_{i-1,i}(E_{f_{i-1}}, E_{f_i}) \\ \vdots \\ J_{N_t,\alpha}(E_{f_{N_t}}, E_{f_\alpha}) - J_{N_t-1,N_t}(E_{f_{N_t-1}}, E_{f_{N_t}}) \end{pmatrix} \tag{3.38}$$

avec $J_{i,j}$ l'expression du courant donnée par l'équation 3.37.

Après avoir calculé le vecteur F qui contient toutes les équations à résoudre pour le système, la prochaine étape consiste à calculer la matrice jacobienne du système.

Pour calculer la matrice jacobienne, il est nécessaire de calculer les dérivées partielles de la fonction F par rapport à chaque variable inconnue. Dans ce cas, les variables inconnues sont $E_{f_{i-1}}$, E_{f_i} et $E_{f_{i+1}}$.

Après quelques lignes de calcul, on arrive aux expressions suivantes :

$$\begin{aligned}
 \frac{\partial f(E_{f_i})}{\partial E_{f_i}} &= \frac{-\alpha}{k_B T} \int_a^b \mathcal{T}(E_x) \frac{\exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)} dE_x - \frac{\alpha}{k_B T} \int_c^a \mathcal{T}(E_x) \frac{\exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)} dE_x \\
 \frac{\partial f(E_{f_i})}{\partial E_{f_{i-1}}} &= \frac{\alpha}{k_B T} \int_a^b \mathcal{T}(E_x) \frac{\exp\left(-\frac{E_x - E_{f_{i-1}}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_{i-1}}}{k_B T}\right)} dE_x \\
 \frac{\partial f(E_{f_i})}{\partial E_{f_{i+1}}} &= \frac{\alpha}{k_B T} \int_c^a \mathcal{T}(E_x) \frac{\exp\left(-\frac{E_x - E_{f_{i+1}}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_{i+1}}}{k_B T}\right)} dE_x
 \end{aligned} \tag{3.39}$$

Maintenant que nous avons les expressions des dérivées partielles de la fonction f par rapport à chaque variable inconnue, nous pouvons utiliser l'expression de la matrice jacobienne en Nearest Neighbour 3.5 pour calculer la matrice jacobienne complète.

Pour cela, il suffit de remplacer les expressions des dérivées partielles de la matrice jacobienne en Nearest Neighbour par celles de l'équation 3.39.

Une fois que nous avons la matrice jacobienne, nous pouvons continuer avec l'algorithme de Newton-Raphson pour résoudre les niveaux de Fermi et ensuite réinjecter ces valeurs dans l'équation 3.36 pour calculer les courants en fonction de la tension.

En résumé, le calcul du courant va consister à :

1. Définir une estimation initiale pour les niveaux de Fermi E_{f_i}
2. Utiliser l'équation de continuité 3.37 pour construire le vecteur F qui contient toutes les équations à résoudre pour le système.
3. Calculer la matrice jacobienne complète en utilisant l'expression de la matrice jacobienne en Nearest Neighbour 3.5 et en remplaçant les expressions des dérivées partielles par celles de l'équation 3.39.
4. Utiliser l'algorithme de Newton-Raphson itérativement pour trouver les niveaux de Fermi qui satisfont toutes les équations du système.
5. Réinjecter les valeurs des niveaux de Fermi trouvés dans l'équation 3.36 pour calculer les courants en fonction de la tension.

Pour l'approche Variable Range, la méthode est exactement la même que pour l'approche Nearest Neighbour, à une différence près. L'expression du courant est légèrement modifiée, ce qui a pour conséquence la modification de l'expression de la continuité du courant et de la matrice jacobienne. En Variable Range, le courant aura la forme suivante :

$$J_{iS} = \alpha \sum_{i=0}^{i-1} \sum_{j=i}^{Nt} \int_a^b \mathcal{T}(E_x) \ln \left(\frac{1 + \exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_j}}{k_B T}\right)} \right) dE_x \quad (3.40)$$

Pour la continuité du courant, après quelques lignes de quelques on se retrouve avec l'expression suivante :

$$J_{iS+1} - J_{iS} = \sum_{i=0/i \neq i+1}^{Nt} \alpha \int_a^b \mathcal{T}(E_x) \ln \left(\frac{1 + \exp\left(-\frac{E_x - E_{f_{j+1}}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)} \right) dE_x \quad (3.41)$$

Pour la matrice jacobienne, les différentes dérivées auront la forme suivante :

$$\frac{\partial f(E_{f_i})}{\partial E_{f_i}} = \frac{\alpha}{k_B T} \begin{cases} - \int_a^b \mathcal{T}(E_x) \frac{\exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_i}}{k_B T}\right)} dE_x & \text{si } i \leq j \\ \sum_{i=0/i \neq i+1}^{Nt} \int_a^b \mathcal{T}(E_x) \frac{\exp\left(-\frac{E_x - E_{f_{j+1}}}{k_B T}\right)}{1 + \exp\left(-\frac{E_x - E_{f_{j+1}}}{k_B T}\right)} dE_x & \text{si } i = j + 1 \end{cases} \quad (3.42)$$

Validation modèle

Dans ce chapitre, nous avons préalablement exposé la partie théorique du modèle. À présent, nous procéderons à la vérification de la fonctionnalité du modèle. En premier lieu, nous examinerons le courant en fonction de la position à une $V=1V$, en nous basant sur la figure 3.26, qui représente le courant en fonction de la position pour le modèle Tsu Esaki utilisant la fonction d'occupation de Fermi-Dirac avec les approches NN et VR. On remarque que le courant reste constant avec la position, démontrant ainsi que le schéma de Newton-Raphson satisfait l'hypothèse centrale du modèle, à savoir la continuité du courant. Une légère différence de courant entre NN et VR est notée, attribuable au nombre limité de pièges. Cependant, on remarque

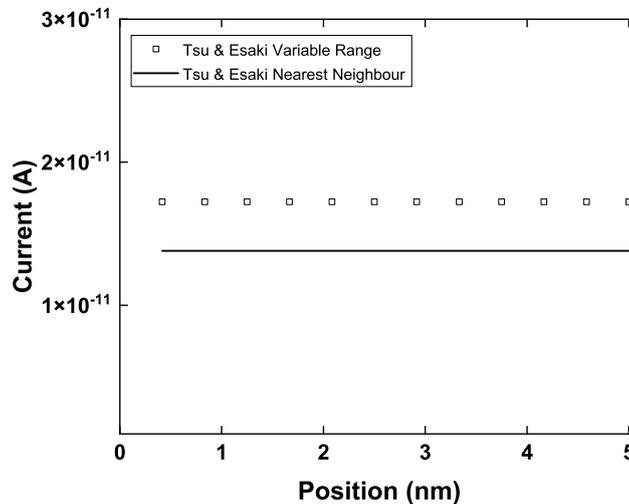


FIGURE 3.26 – Courant en fonction de la position à $V = 1 V$ pour le modèle de Tsu-Esaki en NN et VR. Les paramètres utilisés : $N_t=12$, $E_t=0.2 eV$

que la relation entre les niveaux de Fermi en fonction de la position (figure 3.27 (a)) n'est pas linéaire. Près des électrodes, une légère non-linéarité peut être observée, en particulier lors de la simulation d'un petit nombre de pièges, ce qui justifie la résolution choisie pour les niveaux de Fermi. Dans la figure 3.27 (b), le niveau de Fermi d'un unique piège situé au milieu de l'oxyde en fonction de la tension est présenté pour les approches NN et VR. On remarque que l'abaissement du niveau de Fermi est linéaire par rapport à la tension appliquée. Aucune différence significative n'est observée entre NN et VR sur les deux figures.

Afin de déterminer si le modèle prend correctement en compte la transition HRS-LRS, il est crucial de tracer les courbes I-V pour différents nombres de pièges et d'observer si la transition entre l'allure exponentielle (HRS) et la allure linéaire (LRS) se produit. Ces courbes permettront également de vérifier si les niveaux de courant attendus pour les deux états de résistance ont été atteints.

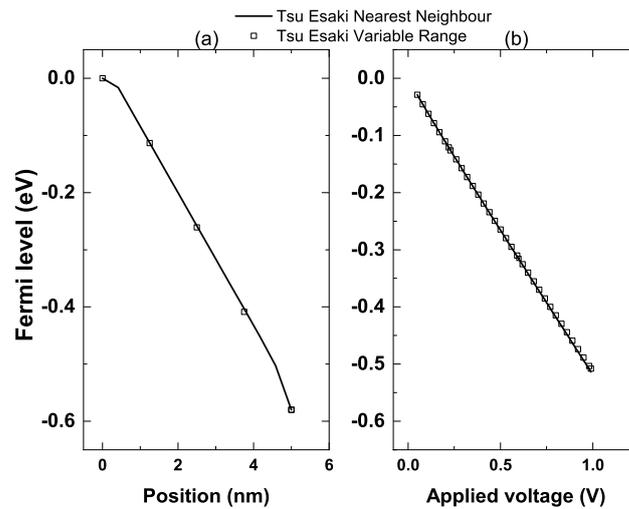


FIGURE 3.27 – (a) Niveau de Fermi en fonction de la position à $V = 0,5$ V pour le modèle de Tsu-Esaki en NN et VR. (b) Niveau de Fermi d'un piège au milieu de l'oxyde en fonction de la tension pour le modèle de Tsu-Esaki en NN et VR. Les paramètres utilisés : $N_t=12$, $E_t = 0.2$ eV

Sur la figure 3.28, on remarque que les courbes présentent une allure exponentielle en HRS et une allure linéaire en LRS pour le modèle TE. Cependant, en ce qui concerne les valeurs de courant, le modèle TE affiche des valeurs bien en deçà de celles attendues dans les deux états de résistance, que ce soit pour NN ou VR. En effet, on atteint au maximum 10^{-11} A en LRS et à peine 10^{-12} A en HRS, alors que les valeurs attendues sont plutôt de l'ordre de 10^{-4} A pour le LRS et 10^{-6} A pour le HRS.

Concernant les approches VR et NN, en HRS, aucune différence n'est visible étant donné que le nombre de pièges n'est pas important. En revanche, en LRS, une différence d'un facteur 2 subsiste entre la valeur du courant en VR et en NN, car le nombre élevé de pièges et la communication entre eux permet d'obtenir un niveau de courant plus élevé en VR.

Une fois la forme et le niveau de courant vérifiés, deux caractéristiques restent à examiner pour conclure sur la robustesse du modèle : le comportement du modèle dans le cas dégénéré et son comportement en température. Ces deux caractéristiques seront exploitées pour confirmer que le modèle simplifié et le modèle complet concordent, que ce soit dans le cas dégénéré ou non dégénéré.

Les figures 3.29 et 3.30 illustrent la relation entre le courant et la tension pour différentes températures, en utilisant deux modèles distincts. Le premier modèle adopte la fonction de Fermi-Dirac comme fonction d'occupation, tandis que le second modèle utilise la fonction de Boltzmann dans l'approche NN. Les résultats sont présentés pour le HRS dans la figure 3.29 et le LRS dans la figure 3.29.

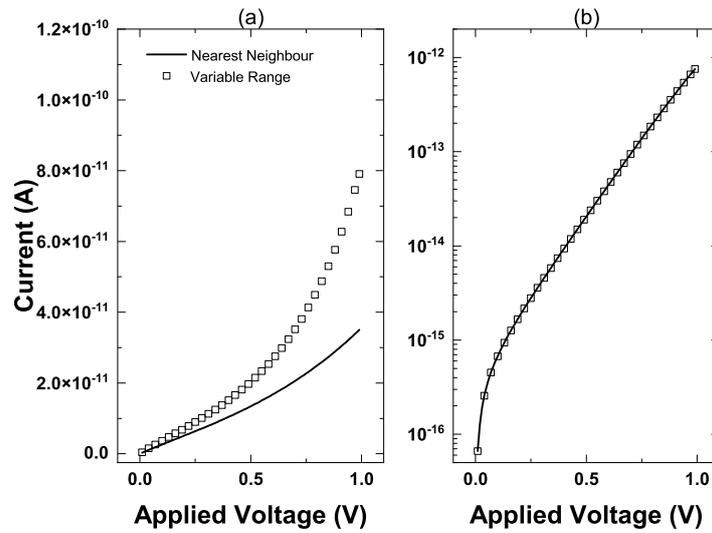


FIGURE 3.28 – (a) Courbe I-V avec le modèle Tsu-Esaki en LRS (25 pièges). (b) Courbe I-V avec le modèle Tsu-Esaki en HRS (6 pièges). Les paramètres utilisés : $E_t = 0.2\text{eV}$

Pour le HRS et le LRS, on observe une concordance entre les deux modèles pour toutes les températures. Ces observations confirment que le modèle basé sur la fonction de Fermi-Dirac capture efficacement la dépendance en température et la dégénérescence des porteurs pour les deux états de résistance. Pour conclure

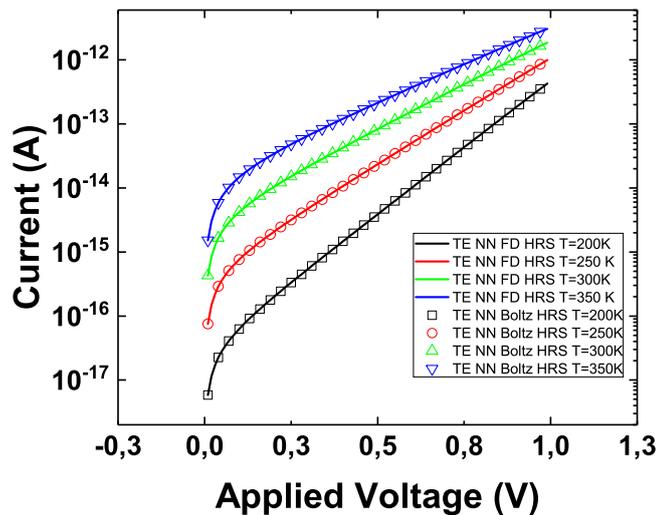


FIGURE 3.29 – Courbe I-V en fonction de la température pour le modèle Tsu-Esaki NN en HRS. Les paramètres utilisés : $E_t = 0.2\text{ eV}$ et $N_t = 6$

l'exploration de ce modèle, examinons maintenant l'approche VR. Les figures 3.31 et

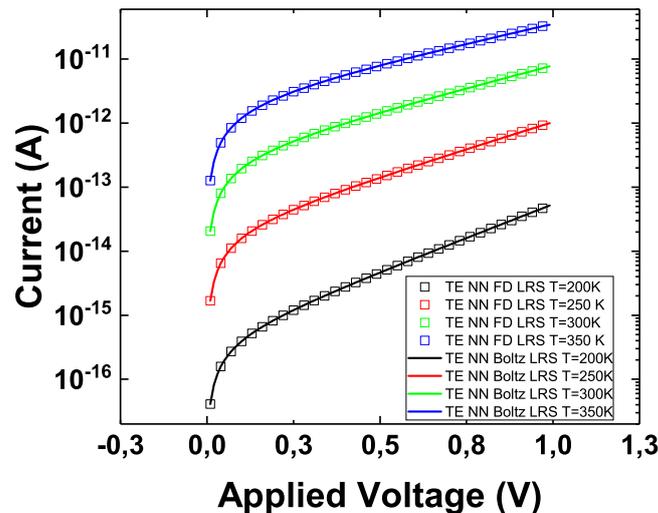


FIGURE 3.30 – Courbe I-V en fonction de la température pour le modèle Tsu-Esaki NN en LRS. Les paramètres utilisés : $E_t = 0.2$ eV et $N_t = 25$

3.32 illustrent la relation entre le courant et la tension pour différentes températures, en utilisant deux modèles distincts. Le premier modèle utilise la fonction de Fermi-Dirac comme fonction d'occupation, tandis que le second modèle utilise la fonction de Boltzmann dans l'approche VR. Les résultats sont présentés pour le HRS dans la figure 3.31 et en LRS dans la figure 3.32.

Pour le HRS, un bon accord entre les deux modèles est observée, indépendamment de la température, ce qui rejoint l'observation faite avec l'approche NN. En ce qui concerne le LRS, aucun effet de la dégénérescence n'est visible sur les courbes. Les deux modèles présentent une concordance pour toutes les températures, suggérant que l'état LRS semble moins sensible à la dépendance en température et à la dégénérescence des porteurs.

En conclusion, les mêmes observations qui ont été faites sur le modèle simplifié de Tsu Esaki, concernant la densité de courant insuffisante pour ajuster une courbe I-V d'une OxRAM, pourraient également être valables pour le modèle complet qui utilise la fonction de Fermi-Dirac comme fonction d'occupation. Cette constatation nous a incités à explorer une autre gamme de modèles, à savoir le modèle de hopping. Nous allons maintenant examiner les résultats de cette exploration.

3.4 Conduction par saut : Hopping

En complément au modèle de conduction TAT discuté précédemment, il convient d'explorer un autre mécanisme connu sous le nom de "hopping", dont la pertinence se manifeste de manière significative dans les matériaux dépourvus d'une structure

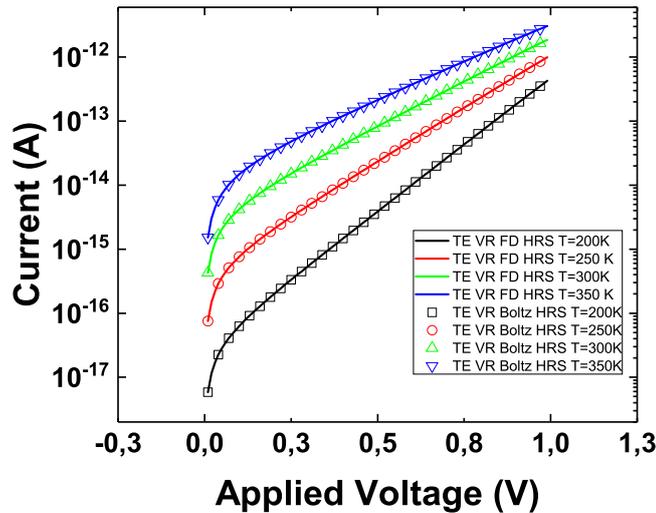


FIGURE 3.31 – Courbe I-V en fonction de la température pour le modèle Tsu-Esaki VR en HRS. Les paramètres utilisés : $E_t = 0.2$ eV et $N_t = 6$

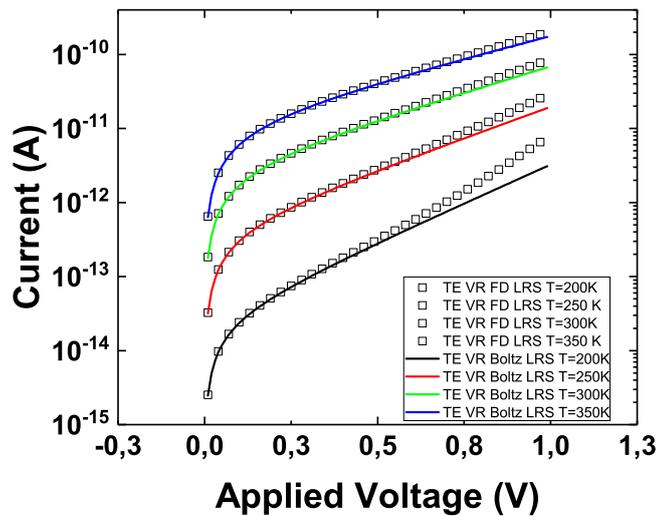


FIGURE 3.32 – Courbe I-V en fonction de la température pour le modèle Tsu-Esaki VR en LRS. Les paramètres utilisés : $E_t = 0.2$ eV et $N_t = 6$

crystalline régulière. En effet, dans de tels matériaux amorphes ou désordonnés, l'organisation atomique ne suit pas un motif cristallin défini, mais adopte plutôt une disposition aléatoire. Cette disposition aléatoire engendre la formation de zones où la densité d'électrons diffère, avec certaines régions présentant une densité électronique plus élevée que d'autres.

Dans le contexte de notre étude, les pièges créés au sein de l'oxyde jouent le rôle de sites électroniques, mais se distinguent en tant qu'états liés plutôt que constituant

un continuum d'états. Ces pièges, en tant qu'états liés, contribuent à la complexité du mécanisme de conduction, offrant ainsi une perspective distincte dans la compréhension du transport électronique dans les matériaux désordonnés. Cette considération des états liés, plutôt que d'un continuum d'états, représente une nuance importante dans le transport électronique dans ce type de matériaux particuliers. Il est important de souligner que des travaux antérieurs ont exploré le mécanisme de conduction par hopping en employant des approches similaires, notamment en utilisant des conditions aux limites périodiques. Ces études, bien qu'intéressantes, diffèrent de notre approche actuelle car elles traitaient des chaînes unidimensionnelles infinies [100]. En revanche, notre modèle se focalise sur l'étude de chaînes finies, ce qui constitue une distinction clé dans notre méthodologie. Tout d'abord, dans la section dédiée à la modélisation, une analyse de ce modèle ainsi que son application à l'OxRAM seront présentées. Par la suite, dans la section de validation du modèle, une série de tests sera réalisée après son implémentation afin de vérifier sa robustesse.

3.4.1 Modélisation

Modèle de Miller et Abrahams

Dans cette section, nous focalisons notre attention sur l'étude du transport électronique dans un matériau isolant spécifique, à savoir un oxyde métallique, dont la principale caractéristique est la possibilité de créer et d'annihiler des pièges électroniques. L'objectif sous-jacent à cette section réside dans l'établissement d'une expression du courant qui traverse l'oxyde, tout en adoptant des hypothèses le moins contraignantes possible. Les pièges dans l'oxyde seront modéliser sous forme chaîne d'électron comme illustré dans la figure 3.33. Cette approche s'avère cruciale pour prendre en compte la nature discrète des pièges électroniques et évaluer leur incidence sur les propriétés de transport au sein du matériau isolant considéré. Pour

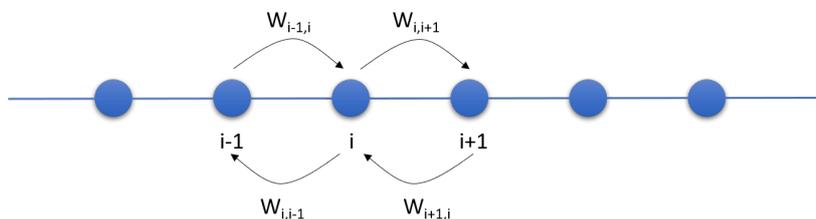


FIGURE 3.33 – Schéma explicatif du Hopping

formuler l'équation de courant, on adoptera une approche probabiliste. A un instant t_0 , on suppose que $P_i(t)$ et $P_j(t)$ sont les probabilités pour qu'une particule donnée soit dans l'état i ou j , à l'instant t . Évidemment, elles doivent satisfaire $P_i(t) + P_j(t) = 1$. Nous allons maintenant construire une équation différentielle pour

décrire $P_i(t)$. Nous le faisons en reliant les probabilités à deux instants, t et $t + dt$. Ce faisant, nous utilisons implicitement l'hypothèse selon laquelle le taux de transition d'un état à l'autre pendant l'intervalle de temps $(t, t + dt)$ ne dépend pas du tout de l'historique précédent, mais seulement de l'état au temps t . La probabilité $P_i(t + dt)$ que la particule soit dans l'état i au temps $t + dt$ a deux contributions : celle d'être dans l'état i au temps t et de ne pas avoir sauté à l'état j pendant l'intervalle $(t, t + dt)$, et celle d'être dans l'état j au temps t et d'avoir fait un saut de j à i dans l'intervalle $(t, t + dt)$. En additionnant tous ces cas et en utilisant les règles de la probabilité conditionnelle, nous avons :

$$P_i(t + dt) = P_i(t)Prob(\text{rester en } i) + P_j(t)Prob(\text{sauter de } j \text{ vers } i) \quad (3.43)$$

En définissant le taux de transition la probabilité de passer de l'état j à l'état i dans l'intervalle de temps $(t, t + dt)$ par $\omega_{j,i}dt$ tandis que la probabilité de rester dans l'état i est égale à 1 moins la probabilité de quitter l'état i dans le même intervalle de temps, soit $1 - w_{i,j}dt$. Cela conduit à :

$$P_i(t + dt) = P_i(t)(1 - w_{i,j}dt) + P_j(t)\omega_{j,i}dt \quad (3.44)$$

En réarrangeant les termes et en prenant la limite $dt \rightarrow 0$, on peut écrire :

$$\frac{dP_i}{dt}(t) = -P_i(t)w_{i,j} + P_j(t)\omega_{j,i} \quad (3.45)$$

En ayant le même raisonnement on arrive à un résultat similaire pour l'état j :

$$\frac{dP_j}{dt}(t) = P_i(t)w_{i,j} - P_j(t)\omega_{j,i} \quad (3.46)$$

Les équations 3.43, 3.44 sont un exemple de la Master equation : équations qui permet de calculer la probabilité qu'une particule stochastique peut "sauter" entre différents états à un instant t . Ces équations différentielles peuvent être résolues à condition de connaître les conditions initiales à savoir $P_i(t_0)$ et $P_j(t_0)$. Pour mieux appréhender ces quantités, on peut définir un courant de probabilité $I(i \longleftrightarrow j)$ comme suit :

$$I(i \longleftrightarrow j) = P_i(t)w_{i,j} - P_j(t)\omega_{j,i} \quad (3.47)$$

En nous remettant dans le contexte, où les particules en question sont des électrons et que ces derniers passent d'un piège à l'autre, en supposant que les pièges ne portent pas de charge et que la probabilité de passer d'un piège à un autre dépend uniquement de son état à l'instant t , l'expression de la probabilité d'être dans le site i est déterminée par la probabilité d'être sur le site i sachant que le piège en j est vide, multiplié par la probabilité que le site j soit vide, ce qui se résume à :

$$P_i(t) = f_i(1 - f_j) \quad (3.48)$$

Un raisonnement similaire conduirait à :

$$P_j(t) = f_j(1 - f_i) \quad (3.49)$$

Avec f_i et f_j les fonctions d'occupations des pièges i et j . $w_{i,j}$ désigne le taux de transition du piège i vers le piège j . En réinjectant les équations 3.48 3.49, dans l'équation du courant de probabilité 3.47 on arrive à l'équation suivante :

$$I_{i,j} = f_i(1 - f_j)w_{i,j} - f_j(1 - f_i)w_{j,i} \quad (3.50)$$

En multipliant l'équation 3.50, par la charge élémentaire q , on arrive à avoir une équation de courant comme suit :

$$J_{i,j} = q (f_i(1 - f_j)w_{i,j} - f_j(1 - f_i)w_{j,i}) \quad (3.51)$$

Une fois l'expression du courant posé, définissons maintenant les différents termes de cette équation.

Définition des fonctions d'occupation

En nous inspirant de l'approche de Miller et Abrahams, nous choisissons d'adopter les fonctions d'occupation sous la forme des fonctions de Fermi-Dirac. Cette sélection s'avère particulièrement pertinente étant donné que les particules considérées sont des fermions. Cependant, l'hypothèse cruciale réside dans la considération d'un niveau de Fermi pour le piège, bien que ce dernier demeure inconnu. Nous envisageons de déterminer ce niveau en nous basant sur la conservation du courant. L'expression des fonctions d'occupation est formulée comme suit :

$$f_i = \frac{1}{\exp\left(\frac{E_{t_i} - E_{f_i}}{k_B T}\right) + 1} \quad (3.52)$$

où E_{t_i} représente l'énergie associée au piège, et E_{f_i} est le niveau de Fermi attribué au piège.

Définition du taux de transition

Le taux de transition $w_{i,j}$ désigne la probabilité intrinsèque qu'un électron, initialement situé sur le site i , effectue une transition vers le site j . Pour déterminer cette probabilité de transition entre deux états propres distincts, $|\psi_i\rangle$ et $|\psi_j\rangle$, nous faisons appel à la règle d'or de Fermi, comme illustré de manière explicite dans la figure 3.34. En se référant à l'hamiltonien H détaillé dans l'article de Miller et Abrahams, l'expression précise du taux de transition est formulée comme suit :

$$w_{i,j} = \frac{2\pi}{\hbar} |\langle \psi_j | H | \psi_i \rangle|^2 \quad (3.53)$$

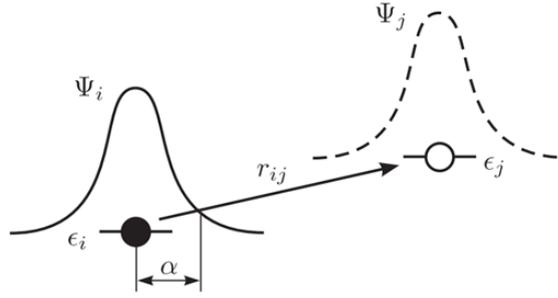


FIGURE 3.34 – Schéma explicatif de la transition entre deux états .

Source : [101]

L'hamiltonien H se compose des termes suivants : l'énergie cinétique T , l'énergie potentielle V , les contributions électrostatiques liées aux distances des pièges à la cathode ($\frac{e^2}{\epsilon r_i} + \frac{e^2}{\epsilon r_j}$), où ϵ est la constante diélectrique et r_i représente la distance du piège i à la cathode.

En considérant l'hypothèse que la transition se produit exclusivement par effet tunnel entre l'état i et l'état j , et que le milieu est suffisamment anisotrope pour assimiler les différentes directions cristallographiques comme équivalentes, l'intégrale dans l'expression du taux de transition peut être simplifiée de la manière suivante ¹ :

$$w_{i,j} = w_0 \exp\left(\frac{-2r_{i,j}}{\alpha}\right) \begin{cases} \exp\left(-\frac{E_j - E_i}{k_B T}\right) & \text{si } E_j - E_i > 0 \\ 1 & \text{sinon} \end{cases} \quad (3.54)$$

Ici, $r_{i,j} = |r_i - r_j|$, $\alpha = 10^{-10} m$, et $w_0 = 10^{17} s^{-1}$.

Les contributions de la transition tunnel et de l'activation thermique deviennent encore plus explicites dans ce modèle que dans le modèle de Tsu-Esaki. En effet, le premier terme dans l'expression rend compte de la contribution du saut tunnel. Le second terme rend quant à lui compte, de l'activation thermique qui est introduite via un terme de Boltzmann. Lorsque le saut implique une augmentation d'énergie c'est à dire que l'état initial i a une énergie inférieure à l'état final j ($\Delta E_{i,j} > 0$), le processus est pénalisé de manière exponentielle. En effet, pour qu'une telle transition puisse se produire, des phonons doivent intervenir pour fournir l'énergie nécessaire. Cette décroissance exponentielle reflète donc l'influence des phonons dans le modèle. Cependant, le saut est dans la direction opposée ($\Delta E_{i,j} < 0$) le terme de Boltzmann, est remplacé par 1, cette transition est considéré facile car elle n'implique aucun gain en énergie.

w_0 symbolise la fréquence maximale de sauts, parfois qualifiée de fréquence de tentative d'évasion. Le paramètre de saut, α , délimite le rayon de localisation, indiquant dans quelle mesure les porteurs de charge peuvent se frayer un chemin à travers la distance $r_{i,j}$ séparant les sites i et j , comme le montre le schéma 3.34.

1. Les calculs ont été fait dans la thèse de Miller [102]

α désigne ici le rayon de localisation (localization radius en anglais), c'est à dire la largeur de la densité d'état. Dans le cadre de Miller-Abrahams, les spécificités moléculaires sont reléguées au second plan. Au lieu de traiter formellement les intégrales de recouvrement, une approximation est faite en estimant la valeur de w_0 . En outre, l'énergie d'ionisation est rapportée aux différences énergétiques entre les sites, déduites d'une distribution de densité d'états, souvent supposée gaussienne.

Mise en équation dans le cadre d'une OxRAM

Dans cette section, nous entreprenons une adaptation du modèle de Miller et Abrahams spécifiquement pour les OxRAM. Nous aborderons la résolution des niveaux de Fermi de chaque piège en fonction de la tension appliquée. Comme la résolution analytique n'est pas toujours possible, nous mettrons en place une résolution numérique. Avant d'aborder directement la résolution, nous explorerons quatre étapes de résolution.

Premièrement, nous simplifierons au maximum le modèle, en supposant que la probabilité de non-occupation du piège suivant ($1 - f_j = 1$) est égale à 1, et que le préfacteur $w_{i,j}$ sera réduit à $w_0 \exp\left(\frac{-2r_{i,j}}{\alpha}\right)$.

Ensuite, le préfacteur $w_{i,j}$ sera intégré dans sa formulation intégrale à la résolution dans une deuxième étape de résolution.

Pour le troisième stade de résolution, les fonctions d'occupation du piège suivant seront prises en compte et considérées comme des fonctions de Boltzmann. Il est important de noter que, pour toutes les simplifications, seule la formulation en Nearest Neighbour sera donnée, car elle offre une bonne estimation lors de la résolution sans les complications liées à l'approche Variable Range.

Dans le dernier stade, lorsque le modèle sera traité dans son intégralité, l'approche Variable Range sera bien sûr abordée. À chaque stade de résolution, deux modes seront proposés : soit une approche utilisant la méthode de Newton-Raphson, soit une approche utilisant directement une formulation linéaire du problème, lorsque cela est possible. Il a été décidé d'exposer ces deux approches quand cela est possible pour un but pédagogique qui est entre autre de préparer le lecteur à l'application de la méthode de NR dans sur un système non linéaire.

a) Simplifications : étape 1 Dans cette partie, notre objectif principal consiste à simplifier au maximum le modèle de Miller et Abrahams, représenté par l'équation 3.51. Cette simplification implique deux étapes clés : premièrement, le terme $(1 - f_j)$ qui rend compte de la non occupation du piège j sachant que l'électron est au niveau du piège i , et le terme $(1 - f_i)$ qui rend compte de la non occupation du piège i sachant que l'électron est au niveau du piège j , seront constamment supposés égal à 1. Deuxièmement, le coefficient $w_{i,j}$ sera systématiquement traité comme étant de valeur 1. En adoptant ces simplifications, nous établirons une approche plus épurée

du modèle. En reprenant les même étapes que pour le modèle Tsu-Esaki, nous allons tout d'abord réécrire l'équation 3.51 en intégrant les simplifications cités ci-haut :

$$J_{i,j} = f_i - f_j \quad (3.55)$$

Pour l'approche Nearest Neighbour on pourra remplacer les indices i et j par $i-1, i$ et $i+1$. Comme présenté précédemment, l'inconnu dans ces équations est le niveau de Fermi des pièges. Pour trouver les niveaux de Fermi en question, il sera nécessaire de commencer par étudier la continuité du courant. Commençons donc par essayer de trouver une formulation de l'équation de continuité en utilisant l'approche Nearest Neighbour. Dans cette approche, l'équation du courant sera l'équation 3.55. En réutilisant l'équation 3.1, et en injectant l'équation de courant en Miller et Abrahams simplifié 3.55, on se retrouvera avec :

$$F_i(E_{f_i}) = J_{i,i+1} - J_{i-1,i} = 2f_i - f_{i-1} - f_{i+1} = 0 \quad (3.56)$$

En utilisant l'approximation de Boltzmann $f_i = \exp\left(-\frac{E_{t_i} - E_{f_i}}{k_B T}\right)$, l'équation de continuité pourra s'écrire :

$$\begin{aligned} F_i(E_{f_i}) &= 2\exp\left(-\frac{E_{t_i} - E_{f_i}}{k_B T}\right) - \exp\left(-\frac{E_{t_{i-1}} - E_{f_{i-1}}}{k_B T}\right) - \exp\left(-\frac{E_{t_{i+1}} - E_{f_{i+1}}}{k_B T}\right) \\ &= 2\exp\left(-\frac{E_{t_i}}{k_B T}\right) \exp\left(\frac{E_{f_i}}{k_B T}\right) - \exp\left(-\frac{E_{t_{i-1}}}{k_B T}\right) \exp\left(\frac{E_{f_{i-1}}}{k_B T}\right) \\ &\quad - \exp\left(-\frac{E_{t_{i+1}}}{k_B T}\right) \exp\left(\frac{E_{f_{i+1}}}{k_B T}\right) = 0 \end{aligned} \quad (3.57)$$

Le système $f(E_{f_i}) = J_{i,i+1} - J_{i-1,i}$ étant linéaire en posant $X_i = \exp\left(\frac{E_{f_i}}{k_B T}\right)$, il n'est pas nécessaire d'utiliser une approche de type Newton-Raphson pour pouvoir résoudre ce système. Comme dit en introduction de cette partie, le choix d'exposer les deux méthodes a été fait. Commençons donc par la résolution linéaire des niveaux de Fermi.

Système linéaire Comme mentionné précédemment, le système représenté par l'équation 3.57 présente une propriété de linéarité qui offre la possibilité de démêler les variables E_{f_i} des constantes E_{t_i} . Pour illustrer davantage ce point, en tenant compte des conditions aux limites spécifiques $E_{f_{cathode}} = 0$ et $E_{f_{anode}} = -qV_{app}$, le système peut être réexprimé de la manière suivante :

$$M \times B = C \quad (3.58)$$

avec

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -\exp\left(-\frac{Et_0}{k_B T}\right) & 2\exp\left(-\frac{Et_1}{k_B T}\right) & -\exp\left(-\frac{Et_2}{k_B T}\right) & 0 & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & 0 & -\exp\left(-\frac{Et_{i-1}}{k_B T}\right) & 2\exp\left(-\frac{Et_i}{k_B T}\right) & -\exp\left(-\frac{Et_{i+1}}{k_B T}\right) \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$B = \left(1 \quad X_1 \quad \cdots \quad X_{N_t} \quad \exp\left(\frac{-qV_{app}}{k_B T}\right)\right)^T$$

$$C = \left(1 \quad 0 \quad \cdots \quad 0 \quad \exp\left(\frac{-qV_{app}}{k_B T}\right)\right)$$
(3.59)

Par conséquent, en effectuant une opération de matrice inverse, on parvient à déterminer le vecteur B , qui joue le rôle de solution pour les termes $\exp\left(\frac{Ef_i}{k_B T}\right)$.

Newton-Raphson Comme mentionné dans l'introduction de cette section, la méthode NR n'est pas nécessaire dans ce cas particulier, étant donné que le système est linéaire. Son utilisation ici vise principalement un objectif pédagogique, en simplifiant sa compréhension dans un contexte plus accessible. Dans notre contexte, nous nous intéressons à la résolution de l'équation $f(E_f) = \text{div}(j) = 0$. Pour trouver les solutions de cette équation, nous appliquons la méthode de Newton-Raphson (NR). Mathématiquement, cela se traduit par l'équation suivante :

$$E_f^{(k+1)} = E_f^{(k)} - \mathcal{J}^{-1} \times F(E_f^{(k)})$$
(3.60)

Dans cette équation, $F(E_f^{(k)})$ représente la valeur de $\text{div}(j)$ calculée à partir de l'équation 3.56. Le vecteur F prend la forme suivante :

$$F = \begin{pmatrix} 2\exp\left(-\frac{Et_1 - Ef_1}{k_B T}\right) - \exp\left(-\frac{Et_{cathode} - Ef_{cathode}}{k_B T}\right) - \exp\left(-\frac{Et_2 - Ef_2}{k_B T}\right) \\ \vdots \\ 2\exp\left(-\frac{Et_i - Ef_i}{k_B T}\right) - \exp\left(-\frac{Et_{i-1} - Ef_{i-1}}{k_B T}\right) - \exp\left(-\frac{Et_{i+1} - Ef_{i+1}}{k_B T}\right) \\ \vdots \\ 2\exp\left(-\frac{Et_{N_t} - Ef_{N_t}}{k_B T}\right) - \exp\left(-\frac{Et_{N_t-1} - Ef_{N_t-1}}{k_B T}\right) - \exp\left(-\frac{Et_{anode} - Ef_{anode}}{k_B T}\right) \end{pmatrix}$$
(3.61)

Le symbole \mathcal{J} représente la matrice jacobienne du problème, définie comme suit :

$$\mathcal{J} = \begin{pmatrix} \frac{\partial F_1}{\partial E_{f_1}} & \frac{\partial F_1}{\partial E_{f_2}} & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \frac{\partial F_i}{\partial E_{f_{i-1}}} & \frac{\partial F_i}{\partial E_{f_i}} & \frac{\partial F_i}{\partial E_{f_{i+1}}} & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & \frac{\partial F_{N_t}}{\partial E_{f_{N_t-1}}} & \frac{\partial F_{N_t}}{\partial E_{f_{N_t}}} \end{pmatrix} \quad (3.62)$$

où les dérivées partielles sont données par l'équation 3.63 :

$$\begin{cases} \frac{\partial F_i}{\partial E_{f_{i-1}}} = \frac{-1}{k_B T} \exp\left(-\frac{Et_{i-1} - Ef_{i-1}}{k_B T}\right) \\ \frac{\partial F_i}{\partial E_{f_i}} = \frac{2}{k_B T} \exp\left(-\frac{Et_i - Ef_i}{k_B T}\right) \\ \frac{\partial F_i}{\partial E_{f_{i+1}}} = \frac{-1}{k_B T} \exp\left(-\frac{Et_{i+1} - Ef_{i+1}}{k_B T}\right) \end{cases} \quad (3.63)$$

Cette matrice jacobienne permet d'incorporer l'information sur les variations de la fonction F par rapport aux variations des niveaux de Fermi adjacents. En combinant ces éléments, la méthode de Newton-Raphson est en mesure de trouver les solutions recherchées pour les niveaux de Fermi.

Après avoir déterminé les niveaux de Fermi pour les deux méthodes, il sera simple de calculer les niveaux de courant. En effet, il sera possible d'injecter ces niveaux de Fermi dans l'équation de courant indiquée par 3.57.

b) Simplification étape 2 Dans cette section, notre objectif est d'introduire une complexité supplémentaire à l'équation 3.51. Nous allons prendre en considération un terme additionnel : plus précisément, le coefficient $w_{i,j}$ ne sera plus supposé égal à 1 de manière générale. D'un point de vue physique, cela signifie qu'il sera désormais possible pour un électron ayant une énergie E_i inférieure à celle d'un piège E_{t_j} de passer de l'état i à l'état j . Cette modification contribuera également à accroître la sophistication de notre équation. L'équation du courant pourra alors s'écrire de la façon suivante :

$$J_{i,j} = w_{i,j} f_i - w_{j,i} f_j \quad (3.64)$$

Comme présenté précédemment, l'inconnu dans ces équations est le niveau de Fermi des pièges. Pour trouver les niveaux de Fermi en question, il sera nécessaire de commencer par étudier la continuité du courant. Commençons donc par essayer de trouver une formulation de l'équation de continuité en utilisant l'approche Nearest Neighbour. Dans cette approche, l'équation du courant sera l'équation 3.64. En réutilisant l'équation 3.1, et en injectant l'équation de courant en Miller et Abrahams

simplifié 3.64, on se retrouvera avec :

$$\begin{aligned}
 F_i(E_{f_i}) &= J_{i,i+1} - J_{i-1,i} = (w_{i,i-1} + w_{i,i+1})f_i - w_{i-1,i}f_{i-1} - w_{i+1,i}f_{i+1} \\
 &= (w_{i,i-1} + w_{i,i+1})\exp\left(-\frac{E_{t_i} - E_{f_i}}{k_B T}\right) - w_{i-1,i}\exp\left(-\frac{E_{t_{i-1}} - E_{f_{i-1}}}{k_B T}\right) \\
 &\quad - w_{i+1,i}\exp\left(-\frac{E_{t_{i+1}} - E_{f_{i+1}}}{k_B T}\right) = 0
 \end{aligned} \tag{3.65}$$

Une fois l'équation de continuité établie, il suffit maintenant de résoudre le système pour trouver les différents niveaux de Fermi. Vu la forme de cette équation, elle peut être résolue de deux différentes manières, soit directement en utilisant un système linéaire, soit en utilisant la méthode de Newton-Raphson. Comme le choix a été fait d'exposer les deux méthodes, commençons par l'utilisation du système linéaire.

Système linéaire L'équation 3.65 permet de séparer les E_{f_i} et les E_{t_i} . En prenant comme conditions aux limites $E_{f_{cathode}} = 0$ et $E_{f_{anode}} = -qV_{app}$, le système pourra s'écrire comme suit :

$$M \times B = C \tag{3.66}$$

avec

$$\begin{aligned}
 M &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ M_{1,0} & M_{1,1} & M_{1,2} & 0 & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & 0 & M_{i,i-1} & M_{i,i} & M_{i,i+1} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
 B &= \left(1 \quad X_1 \quad \cdots \quad X_{N_t} \quad \exp\left(\frac{-qV_{app}}{k_B T}\right)\right)^T \\
 C &= \left(1 \quad 0 \quad \cdots \quad 0 \quad \exp\left(\frac{-qV_{app}}{k_B T}\right)\right)
 \end{aligned} \tag{3.67}$$

où :

$$\begin{cases} M_{1,0} = -w_{0,1}\exp\left(-\frac{E_{t_0}}{k_B T}\right) \\ M_{1,1} = (w_{1,0} + w_{1,2})\exp\left(-\frac{E_{t_1}}{k_B T}\right) \\ M_{1,2} = -w_{2,1}\exp\left(-\frac{E_{t_2}}{k_B T}\right) \\ M_{i,i-1} = -w_{i-1,i}\exp\left(-\frac{E_{t_{i-1}}}{k_B T}\right) \\ M_{i,i} = (w_{i,i-1} + w_{i,i+1})\exp\left(-\frac{E_{t_i}}{k_B T}\right) \\ M_{i,i+1} = -w_{i+1,i}\exp\left(-\frac{E_{t_{i+1}}}{k_B T}\right) \end{cases} \tag{3.68}$$

et $X_i = \exp\left(\frac{E_{f_i}}{k_B T}\right)$. Ainsi, le vecteur B sera le vecteur solution des $\exp\left(\frac{E_{f_i}}{k_B T}\right)$. Pour obtenir les différents E_{f_i} , il suffit d'utiliser une fonction logarithmique sur le vecteur B et ensuite multiplier le résultat par $k_B T$

Newton-Raphson Pour la résolution avec la méthode NR, la même méthode que dans la partie simplification étape 1 sera suivie. En reprenant l'équation 3.60, il sera aisé de remarquer que les deux termes manquants pour faire fonctionner cet algorithme itératif sont le vecteur F (vecteur continuité du courant établie par l'équation 3.65) et la matrice jacobienne. Dans le cadre de l'approche Nearest Neighbour, la matrice jacobienne va contenir l'ensemble des dérivées partielles du vecteur f par rapport à l'énergie des pièges adjacents, traduisant ainsi les variations de la fonction F par rapports à ces différentes variables.

\mathcal{J} désigne la matrice jacobienne et aura la forme suivante.

$$\mathcal{J} = \begin{pmatrix} \frac{\partial F_1}{\partial E_{f_1}} & \frac{\partial F_1}{\partial E_{f_2}} & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \frac{\partial F_i}{\partial E_{f_{i-1}}} & \frac{\partial F_i}{\partial E_{f_i}} & \frac{\partial F_i}{\partial E_{f_{i+1}}} & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & \frac{\partial F_{N_t}}{\partial E_{f_{N_t-1}}} & \frac{\partial F_{N_t}}{\partial E_{f_{N_t}}} \end{pmatrix} \quad (3.69)$$

où :

$$\begin{cases} \frac{\partial F_i}{\partial E_{f_{i-1}}} = \frac{-w_{i-1,i}}{k_B T} \exp\left(-\frac{E_{t_{i-1}} - E_{f_{i-1}}}{k_B T}\right) \\ \frac{\partial F_i}{\partial E_{f_i}} = \frac{w_{i,i-1} + w_{i,i+1}}{k_B T} \exp\left(-\frac{E_{t_i} - E_{f_i}}{k_B T}\right) \\ \frac{\partial F_i}{\partial E_{f_{i+1}}} = \frac{-w_{i+1,i}}{k_B T} \exp\left(-\frac{E_{t_{i+1}} - E_{f_{i+1}}}{k_B T}\right) \end{cases} \quad (3.70)$$

Une fois les niveaux de Fermi déterminé, il sera simple de calculer les niveaux de courant. Il suffira d'injecter ces niveaux de Fermi dans l'équation de courant 3.64.

c) Simplification étape 3 Dans cette étape, le modèle sera étudié dans son ensemble, la seule simplification qui sera prise en compte sera l'utilisation de la fonction de Boltzmann en tant que fonction d'occupation. Le courant entre deux pièges i et j prendra la forme de l'équation 3.51. Pour calculer la conservation du courant, il sera nécessaire de reprendre l'équation 3.56 et d'y injecter la nouvelle forme de courant. Ceci permettra d'écrire :

$$\begin{aligned} F_i(E_{f_i}) &= J_{i,i+1} - J_{i-1,i} \\ &= -(1 - f_i)(w_{i-1,i}f_{i-1} + w_{i+1,i}f_{i+1}) + f_i(w_{i,i+1}(1 - f_{i+1}) + w_{i,i-1}(1 - f_{i-1})) \\ &= 0 \end{aligned} \quad (3.71)$$

La forme de la conservation de courant montre que la résolution de cette équation ne peut se faire au travers un système linéaire du fait qu'il y ait des termes croisés entre

f_i , f_{i+1} et f_{i-1} . Il sera donc nécessaire d'utiliser la méthode de Newton-Raphson pour résoudre ce système.

Dans cette partie sera exposée la méthode de NR pour la résolution des niveaux de Fermi avec la simplification mentionné ci-dessus. Comme les expressions deviennent assez compliqué, dans la suite de ce document, les expressions des fonctions d'occupations ne seront explicités que si cela est nécessaire. Pour rappel, la fonction d'occupation sera la fonction de Boltzmann $f_i = \exp\left(-\frac{E_{t_i} - E_{f_i}}{k_B T}\right)$. En examinant l'équation 3.60, on constate que les termes qui ne sont pas encore déterminés sont la fonction F et la matrice jacobienne. Comme expliqué précédemment, la fonction F est représentée par le vecteur de continuité du courant. Ainsi, l'équation 3.71 définira la fonction F. En ce qui concerne la matrice jacobienne, elle sera calculée à partir des dérivées partielles des plus proches voisins. \mathcal{J} désigne la matrice jacobienne du problème et aura la forme suivante :

$$\mathcal{J} = \begin{pmatrix} \frac{\partial F_1}{\partial E_{f_1}} & \frac{\partial F_1}{\partial E_{f_2}} & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \frac{\partial F_i}{\partial E_{f_{i-1}}} & \frac{\partial F_i}{\partial E_{f_i}} & \frac{\partial F_i}{\partial E_{f_{i+1}}} & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & \frac{\partial F_{N_t}}{\partial E_{f_{N_t-1}}} & \frac{\partial F_{N_t}}{\partial E_{f_{N_t}}} \end{pmatrix} \quad (3.72)$$

où :

$$\begin{cases} \frac{\partial F_i}{\partial E_{f_{i-1}}} = & \frac{-w_{i-1,i}}{k_B T} f_{i-1}(1-f_i) - \frac{w_{i,i-1}}{k_B T} f_i f_{i-1} \\ \frac{\partial F_i}{\partial E_{f_i}} = & \frac{f_i}{k_B T} (w_{i+1,i} f_{i+1} + w_{i-1,i} f_{i-1} + w_{i,i+1}(1-f_{i+1}) + w_{i,i-1}(1-f_{i-1})) \\ \frac{\partial F_i}{\partial E_{f_{i+1}}} = & \frac{-w_{i+1,i}}{k_B T} f_{i+1}(1-f_i) - \frac{w_{i,i+1}}{k_B T} f_i f_{i+1} \end{cases} \quad (3.73)$$

d) Modèle sans simplification Dans cette partie, il sera convenu de ne prendre en compte aucune simplification. Le modèle sera considéré dans sa forme originel. La fonction de distribution sera la fonction de Fermi Dirac $f_i = \frac{1}{\exp\left(\frac{E_{t_i} - E_{f_i}}{k_B T}\right) + 1}$, et le coefficient $w_{i,j}$ sera aussi pris en compte dans son intégralité, et prendra la forme de l'équation 3.54. Ainsi l'équation de courant dans l'approche Nearest Neighbour pourra s'écrire, à l'aide des deux termes précédents de la manière suivante :

$$J_{i,j} = f_i(1-f_j)w_{i,j} - f_j(1-f_i)w_{j,i} \quad (3.74)$$

Une fois l'équation de courant établie, il sera nécessaire de calculer la conservation du courant pour pouvoir ensuite résoudre les niveaux de Fermi. En reprenant l'équation

du courant 3.74, la conservation du courant pourra s'écrire de la manière suivante :

$$\begin{aligned}
 F_i(E_{f_i}) &= J_{i,i+1} - J_{i-1,i} \\
 &= -(1 - f_i)(w_{i-1,i}f_{i-1} + w_{i+1,i}f_{i+1}) + f_i(w_{i,i+1}(1 - f_{i+1}) + w_{i,i-1}(1 - f_{i-1})) \\
 &= 0
 \end{aligned} \tag{3.75}$$

Il sera aisé de remarquer tout comme dans le paragraphe **simplification étape 3**, que l'obtention d'une forme linéaire à partir de l'équation de continuité du courant 3.75 est impossible vu que le système est fortement non linéaire. Pour tenter de résoudre ce système, il sera donc nécessaire d'utiliser une méthode itérative telle que la méthode de Newton Raphson.

Dans la méthode de Newton-Raphson, le vecteur F est constitué des éléments de l'équation de continuité du courant définie dans l'équation 3.75. Une fois ce vecteur établi, il faut calculer la matrice jacobienne du système. Pour ce faire, il faudra calculer les dérivées partielles des composantes du vecteur F en fonction des différentes énergies de pièges. Dans l'approche Nearest Neighbour, les dérivées partielles non nulles sont celles avec les plus proches voisins donc la matrice jacobienne prendra la forme suivante :

$$\mathcal{J} = \begin{pmatrix} \frac{\partial F_1}{\partial E_{f_1}} & \frac{\partial F_1}{\partial E_{f_2}} & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \frac{\partial F_i}{\partial E_{f_{i-1}}} & \frac{\partial F_i}{\partial E_{f_i}} & \frac{\partial F_i}{\partial E_{f_{i+1}}} & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & \frac{\partial F_{N_t}}{\partial E_{f_{N_t-1}}} & \frac{\partial F_{N_t}}{\partial E_{f_{N_t}}} \end{pmatrix} \tag{3.76}$$

Les différentes dérivées partielles ont été calculés et auront les expressions suivantes :

$$\left\{ \begin{aligned} \frac{\partial F_i}{\partial E_{f_{i-1}}} &= \frac{1}{k_B T} \left(-w_{i-1,i}\beta(1 - f_i)f_{i-1}^2 - w_{i,i-1}\beta f_i f_{i-1}^2 \right) \\ \frac{\partial F_i}{\partial E_{f_i}} &= \frac{1}{k_B T} \left(w_{i-1,i}f_{i-1} + w_{i+1,i}f_{i+1} + w_{i,i+1}(1 - f_{i+1}) + w_{i,i-1}(1 - f_{i-1})\beta f_i^2 \right) \\ \frac{\partial F_i}{\partial E_{f_{i+1}}} &= \frac{1}{k_B T} \left(-w_{i+1,i}\beta(1 - f_i)f_{i+1}^2 - w_{i,i+1}\beta f_i f_{i+1}^2 \right) \end{aligned} \right. \tag{3.77}$$

avec :

$$\left\{ \begin{aligned} f_{i-1} &= \frac{1}{\exp\left(\frac{E_{t_{i-1}} - E_{f_{i-1}}}{k_B T}\right) + 1} \\ f_i &= \frac{1}{\exp\left(\frac{E_{t_i} - E_{f_i}}{k_B T}\right) + 1} \\ f_{i+1} &= \frac{1}{\exp\left(\frac{E_{t_{i+1}} - E_{f_{i+1}}}{k_B T}\right) + 1} \\ \beta &= \exp\left(\frac{E_{t_i} - E_{f_i}}{k_B T}\right) \end{aligned} \right. \tag{3.78}$$

En résumé, le calcul du courant suit les mêmes étapes que celles décrites pour le modèle Tsu-Esaki en Nearest Neighbour (NN). Pour l'approche Variable Range, la méthode est presque identique à celle de l'approche Nearest Neighbour mis à part l'expression du courant qui est modifiée. Cette modification entraîne des changements dans l'expression de la continuité du courant ainsi que dans la matrice jacobienne. En Variable Range, le courant aura la forme suivante :

$$J_{iS} = \alpha \sum_{i=0}^{i-1} \sum_{j=i}^{Nt} f_i(1 - f_j)w_{i,j} - f_j(1 - f_i)w_{j,i} \quad (3.79)$$

Pour la matrice jacobienne, les différentes dérivées auront la forme suivante :

$$\frac{\partial F(E_{f_i})}{\partial E_{f_i}} = \begin{cases} w_{i-1,i}f_{i-1}^2\beta(1 - f_i) + w_{i,i-1}\beta f_i f_{i-1}^2 & \text{si } i \leq j \\ \sum_{i=0/i \neq i+1}^{Nt} w_{i-1,i}f_{i-1}\beta f_i^2 + w_{i,i-1}\beta f_i^2(1 - f_{i-1}) & \text{si } i = j + 1 \end{cases} \quad (3.80)$$

3.4.2 Validation modèle

Dans cette section, nous présenterons les résultats de simulations obtenus à partir du modèle de hopping. La partie théorique de ce modèle a été discutée précédemment dans ce chapitre. À présent, nous allons examiner et commenter les divers résultats obtenus.

Tout d'abord, comme nous l'avons fait avec le modèle Tsu-Esaki (TE), nous allons vérifier que le modèle respecte bien l'équation de continuité. Pour ce faire, nous devons tracer la courbe du courant en fonction de la position des pièges à une tension donnée et vérifier que le courant reste constant en fonction de la position des pièges.

La densité de courant en fonction de la position a été tracée dans la figure 3.35 pour une tension appliquée de 1 V, en utilisant à la fois l'approche NN et l'approche VR. Comme illustré, le courant demeure constant en fonction de la position, ce qui confirme que le schéma de résolution numérique basé sur l'équation de Newton-Raphson satisfait l'hypothèse fondamentale du modèle, à savoir la continuité du courant.

Une légère différence de courant est observée entre l'approche NN et l'approche VR. Cette disparité peut être expliquée par le fait que le nombre important de pièges favorise la communication entre eux, entraînant ainsi une densité de courant plus élevée en utilisant l'approche VR par rapport à l'approche NN.

La deuxième caractéristique à examiner, découlant directement de la continuité du courant, concerne la forme des niveaux de Fermi. Dans la figure 3.36 (a), le niveau de Fermi des pièges est représenté en fonction de leur position, en utilisant les approches VR et NN. En ce qui concerne l'allure cette courbe, une légère non-linéarité est remarquée près des électrodes, surtout lorsque le nombre de pièges simulés est réduit. Cette constatation soutient donc la décision de résoudre les niveaux de Fermi de manière précise.

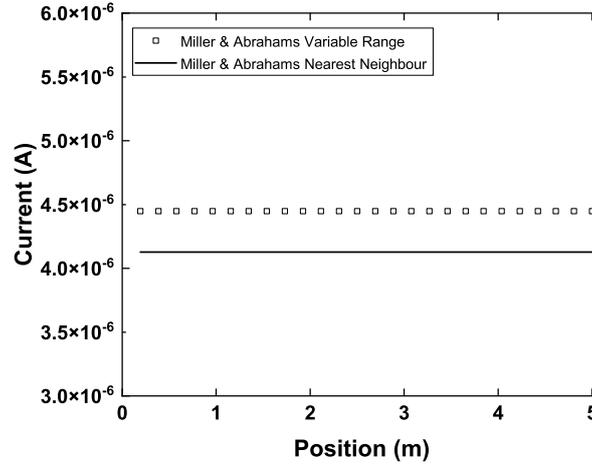


FIGURE 3.35 – Courant en fonction de la position à $V = 1$ V pour le modèle de Miller et Abrahams en utilisant l’approche Nearest Neighbour et Variable range. Les paramètres utilisés : $E_t = 0.2\text{eV}$, $N_t = 25$, $\alpha = 10^{-10}\text{m}$, $w_0 = 10^{17}\text{s}^{-1}$

Dans la figure 3.36 (b), le niveau de Fermi d’un seul piège situé au milieu de l’oxyde est représenté en fonction de la tension, selon les approches NN et VR. On peut remarquer que la diminution du niveau de Fermi est linéaire avec la tension appliquée. Aucune différence significative n’est perceptible entre l’approche NN et l’approche VR sur ces deux figures. Cela est tout à fait normal, car bien que les deux approches diffèrent dans leur manière de calculer le courant, elles partagent la même méthode pour déterminer les niveaux de Fermi.

Ces observations diverses conduisent à la conclusion que la résolution précise des niveaux de Fermi est essentielle pour respecter la condition de continuité du courant. Supposer une évolution linéaire des niveaux de Fermi en fonction de la position aurait été une erreur, comme démontré dans nos résultats. Pour vérifier entièrement la robustesse du modèle, il est essentiel d’analyser la conformité des courbes I-V en HRS et en LRS, ainsi que d’évaluer la capacité du modèle à reproduire la transition entre ces deux états de conduction. De plus, il convient de vérifier si le modèle génère un courant suffisant pour ces deux états de résistance. Dans la figure 3.37 (a), la caractéristique I-V est présentée en utilisant le modèle de Miller et Abrahams avec 25 pièges. Avec l’approche NN, on observe une courbe quasi-linéaire sur toute la plage de tension. Le niveau de courant atteint environ 10^{-4} A, indiquant que ce modèle en NN parvient enfin à reproduire de manière satisfaisante la forme et le niveau de courant attendus pour le LRS dans une OxRAM. Concernant l’approche VR, à basse tension, aucune différence significative n’est observée par rapport à l’approche NN. En revanche, à des tensions plus élevées, l’approche VR montre une tendance parabolique distincte avec un niveau de courant légèrement supérieur à

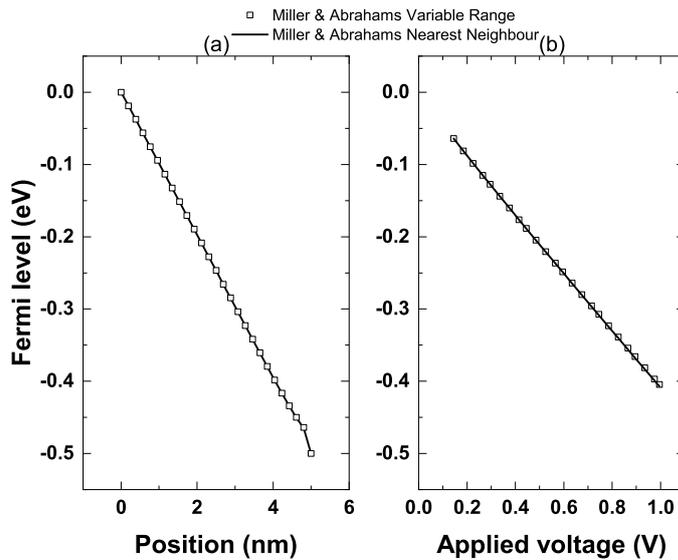


FIGURE 3.36 – (a) Niveau de Fermi en fonction de la position à $V = 0.5$ V pour le modèle de Miller et Abrahams en NN et VR. (b) Niveau de Fermi d'un piège en fonction de la tension appliqué pour le modèle de Miller et Abrahams en NN et VR. Les paramètres utilisés : $E_t = 0.2$ eV, $N_t = 25$, $\alpha = 10^{-10}$ m, $w_0 = 10^{17}$ s $^{-1}$

celui de l'approche NN. Ceci a déjà été observé pour le modèle de Tsu-Esaki.

La figure 3.37 (b) représente une caractéristique I-V en utilisant le modèle de Miller et Abrahams avec seulement 6 pièges. On peut noter que la courbe présente une allure exponentielle, et les niveaux de courant sont de l'ordre de 10^{-6} A. Cela signifie que ce modèle parvient également à reproduire de manière adéquate la forme et le niveau de courant attendus pour un HRS dans une OxRAM. En ce qui concerne les approches NN et VR, aucune différence notable n'est observée entre les deux, car le nombre limité de pièges rend improbable toute influence de la communication entre pièges distants sur la valeur du courant.

Le principal paramètre utilisé pour ajuster les courbes expérimentales est la fréquence maximale de saut. Ce paramètre joue un rôle crucial car il détermine la fréquence à laquelle les électrons peuvent surmonter les barrières d'énergie et effectuer des transitions entre états. Selon la littérature, la fréquence maximale de saut varie généralement entre 10^{14} et 10^{18} s $^{-1}$ [103].

Pour conclure la validation approfondie de la robustesse du modèle, il est crucial d'explorer son comportement en fonction de la température. La question qui subsiste à résoudre est de déterminer la capacité de ce modèle à reproduire fidèlement les caractéristiques I-V de l'OxRAM en réponse aux variations de température. Afin de répondre à cette interrogation, il est impératif de tracer des courbes I-V en LRS et en HRS pour différentes températures.

La figure 3.38 illustre la caractéristique I-V à différentes températures en LRS. On

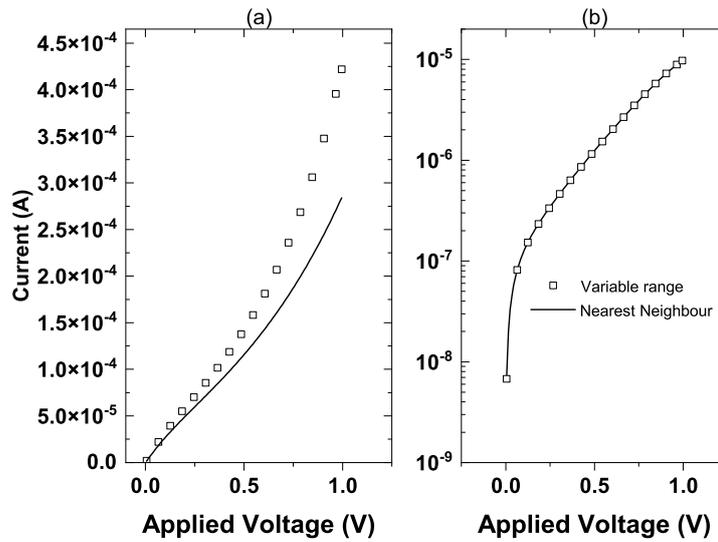


FIGURE 3.37 – (a) Courbe I-V pour le modèle de Miller et Abrahams en LRS (avec 25 pièges). (b) Courbe I-V pour le modèle de Miller et Abrahams en HRS (avec 6 pièges). Les paramètres utilisés : $E_t = 0.2\text{eV}$, $\alpha = 10^{-10}\text{m}$, $w_0 = 10^{17}\text{s}^{-1}$

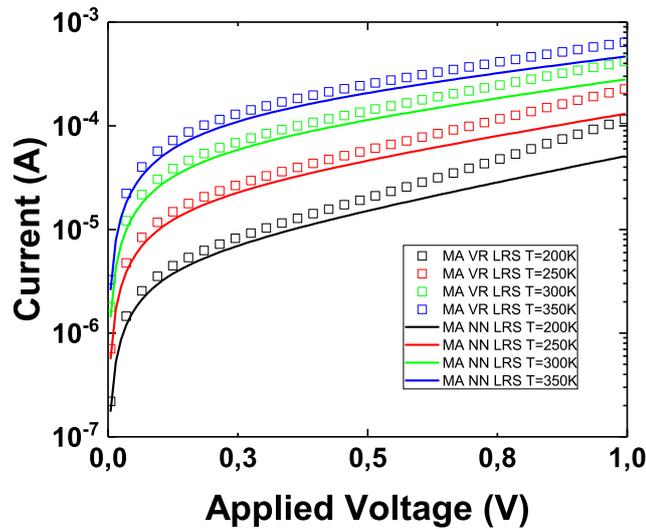


FIGURE 3.38 – (a) Courbe I-V à différentes températures pour le modèle de Miller et Abrahams en LRS (avec 25 pièges). Les paramètres utilisés : $E_t = 0.2\text{eV}$, $\alpha = 10^{-10}\text{m}$, $w_0 = 10^{17}\text{s}^{-1}$

observe clairement que le modèle réagit à la variation de température, indiquant une sensibilité thermique de la caractéristique I-V. Des écarts significatifs sont notables entre les trois températures simulées. Cependant, pour les approches NN et VR, aucune différence majeure n'est à signaler. En effet, les variations sont généralement d'un facteur 2, ce qui n'est pas notable sur une échelle logarithmique. Concernant

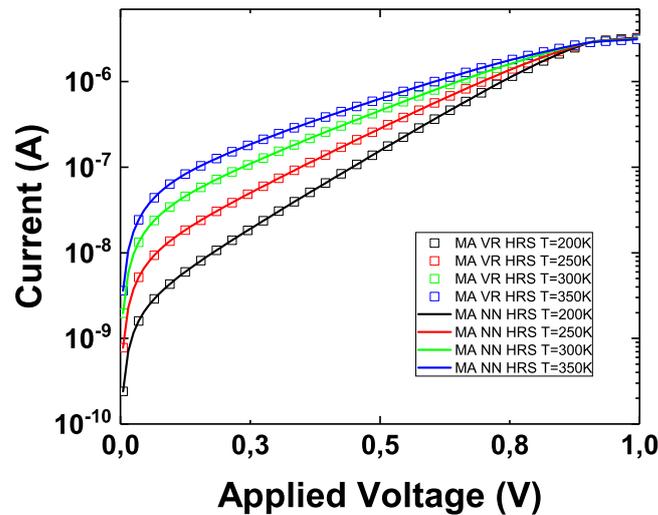


FIGURE 3.39 – Courbe I-V à différentes températures pour le modèle de Miller et Abrahams en HRS (avec 6 pièges). Les paramètres utilisés : $E_t = 0.2\text{eV}$, $\alpha = 10^{-10}\text{m}$, $\omega_0 = 10^{17}\text{s}^{-1}$

la caractéristique I-V en HRS à différentes températures, présentée dans la figure 3.39, on observe également une sensibilité du modèle à la température. Cependant, les écarts de courant entre les différentes températures sont moins significatifs que pour le LRS. Pour les méthodes NN et VR, aucune variation significative du courant n'est constatée.

En résumé, le modèle de Miller et Abrahams se distingue par sa capacité à reproduire à la fois la transition entre le HRS et le LRS, ainsi que la forme des courbes I-V associées à ces états. De plus, le modèle parvient à atteindre les niveaux de courant attendus pour les deux états de résistance, tout en utilisant un nombre raisonnable de pièges, soit 25 pièges pour le LRS et 6 pièges pour le HRS, sur un oxyde d'une épaisseur de 5 nm.

Une caractéristique importante est la sensibilité du modèle à la température, tant en HRS qu'en LRS. Cette capacité à réagir thermiquement permettra peut être de modéliser le comportement d'une OxRAM en fonction de la température. L'ensemble de ces résultats suggère que le modèle de Miller et Abrahams constitue une approche solide et appropriée pour la modélisation précise de la conduction dans les dispositifs OxRAMs.

3.5 Prise en compte des simulations 3D

Les pièges présents dans l'oxyde sont répartis dans les trois dimensions de l'espace. Cependant, ce qui a été présenté jusqu'à présent se concentre uniquement sur

une approche unidimensionnelle (1D). Cette limitation pourrait amener le lecteur à penser que le modèle n'est pas compatible avec une approche tridimensionnelle (3D). En réalité, la prise en compte des pièges dans les trois dimensions est correctement intégrée dans le modèle.

Lors de la commutation, si plusieurs pièges sont situés à la même coordonnée x , c'est-à-dire qu'ils partagent la même position le long de l'axe de l'épaisseur de l'oxyde tout en ayant des coordonnées transversales différentes, cette situation est illustrée par la figure 3.40. Dans cet exemple, un certain nombre de pièges ont été générés pour $x = 2$ nm. Leur contribution au courant est alors simplement additionnée, en considérant que leurs niveaux de Fermi sont nécessairement identiques. Cette

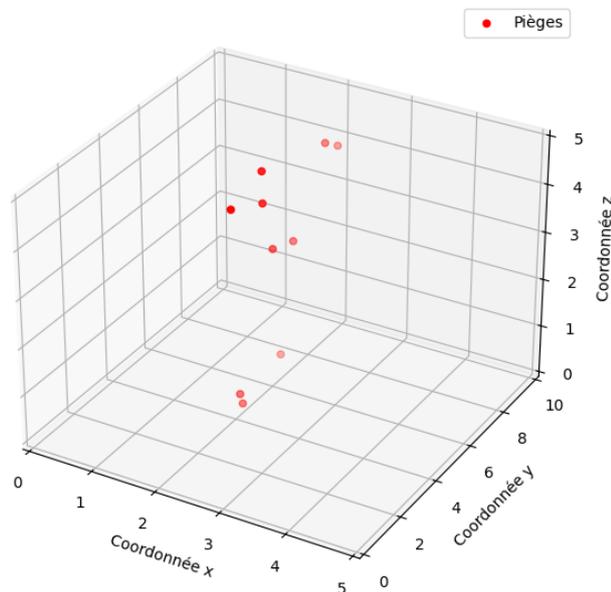


FIGURE 3.40 – Schéma représentant la situation où plusieurs pièges sont générés pour un $x = 2$ nm dans un oxyde de 5 nm d'épaisseur.

approche simplifiée considérablement le modèle par rapport à ceux de la littérature, tels que le modèle de Larcher, qui nécessitent une identification complexe du chemin le plus conducteur parmi les pièges. En évitant cette complexité, notre modèle propose une méthode plus directe pour intégrer les simulations en 3D.

Avant de nous plonger directement dans la comparaison avec les données expérimentales, il serait judicieux d'entreprendre une démarche visant à distinguer le modèle le plus adapté à reproduire le comportement électrique d'une OxRAM. Pour ce faire, nous introduirons deux grandeurs cruciales : l'énergie d'activation et la *voltage power function*. L'énergie d'activation nous permettra de mieux appréhender

le comportement électrique en fonction de la température, tandis que la *voltage power function* nous fournira des indices sur la conformité de la forme de la courbe, en particulier pour l'état LRS.

3.6 Exploitation des résultats

Dans cette section, nous avons l'intention d'analyser minutieusement les résultats issus des divers modèles. Cette section sera divisée en trois grandes parties : la première tentera de discriminer entre les différents modèles, La deuxième partie sera dédiée à faire le bilan de l'ensemble des modèles développés, tandis que la troisième se concentrera sur la comparaison de ces modèles avec les données expérimentales.

Dans la première partie, nous examinerons deux paramètres cruciaux : l'énergie d'activation et la fonction du facteur de puissance. En utilisant ces deux paramètres ainsi que les conclusions établies précédemment, notre objectif sera de déterminer quel modèle est le mieux adapté pour modéliser la caractéristique courant-tension dans une OxRAM.

Dans la deuxième partie, un bilan global sera réalisé sur l'ensemble des modèles et des approches développés dans cette thèse, afin d'identifier quel modèle s'avère être le plus performant.

Dans la troisième partie, nous procéderons à une comparaison des courbes I-V simulés avec des données expérimentales déjà publiées. L'objectif principal sera d'évaluer le comportement du modèle en le confrontant directement à des données expérimentales. Nous chercherons à déterminer si un ajustement satisfaisant peut être obtenu, permettant ainsi de valider la pertinence du modèle en question.

3.6.1 Discrimination des modèles de courants

Dans cette partie, une analyse comparative entre les deux modèles sera effectué. Nous procéderons donc, à la simulation de deux paramètres accessibles expérimentalement. L'objectif est d'évaluer la possibilité de distinguer les deux modèles. Dans un premier temps, nous examinerons l'énergie d'activation thermique du courant, puis nous nous pencherons sur la *voltage power function*, noté γ , de la caractéristique I-V. Cette approche détaillée nous permettra de mieux appréhender les différences subtiles entre les deux modèles et d'évaluer leur adéquation avec les données expérimentales.

Energie d'activation

Pour approfondir l'analyse de la variation du courant en fonction de la température, il est essentiel d'introduire un nouveau paramètre permettant une meilleure compréhension des courbes I-V en relation avec la température. C'est ainsi que

l'énergie d'activation a été intégrée à l'étude. En effet, une énergie d'activation élevée indique une forte dépendance du courant à la température, tandis qu'une faible énergie d'activation suggère que le courant n'est que faiblement affecté par les variations de température. Une énergie d'activation nulle signifie que le courant reste constant indépendamment de la température. Ainsi, ce paramètre se révèle être une ressource précieuse pour identifier les mécanismes de transport électronique. En effet, une énergie d'activation élevée, de l'ordre de l'eV ou proche de l'énergie de gap, est indicative d'un courant de diffusion. En revanche, un mécanisme de recombinaison Shockley-Read-Hall (SRH) se caractérise par une E_a proche de la moitié du gap. Enfin, un courant tunnel se manifeste par une E_a de l'ordre de la dizaine de meV. Il devient donc intéressant d'observer le comportement de ces deux modèles en fonction de la température afin de les différencier. L'énergie d'activation est définie en prolongement de la loi d'Arrhenius, comme le montre l'expression suivante :

$$E_a = - \frac{d \ln(I)}{d \left(\frac{q}{k_B T} \right)} \quad (3.81)$$

où E_a représente l'énergie d'activation, I le courant et T la température. Pour calculer l'énergie d'activation, des courbes I-V ont été générées, à partir du modèle de Miller et Abrahams et du modèle de Tsu-Esaki, à différentes températures, en maintenant constant le nombre de pièges à 6 pour le HRS et à 25 pour le LRS. La figure 3.41 (a) illustre l'évolution de l'énergie d'activation en fonction de la tension dans le HRS selon le modèle de Miller et Abrahams. On observe que E_a commence légèrement en dessous de 0,2 eV, correspondant à la valeur fixée pour l'énergie des pièges par rapport au niveau de Fermi de la cathode dans les simulations. À mesure que la tension appliquée augmente, E_a chute significativement pour atteindre quelques dizaines de meV. Cette tendance indique que, à une tension élevée, le modèle prévoit une faible activation en température du courant. La figure 3.41 (b) présente l'énergie d'activation dans le LRS, toujours en utilisant le modèle de Miller et Abrahams. Contrairement au cas du HRS, une légère variation de l'énergie d'activation est observée en fonction de la tension appliquée. Ceci suggère que, pour cette plage de température spécifique, le courant demeure activé en température malgré l'augmentation de la tension. Dans la figure 3.42 (a) et (b), l'évolution de l'énergie d'activation en fonction de la tension est représentée respectivement en HRS et en LRS en utilisant le modèle de Tsu-Esaki. La variation de l'énergie d'activation présente des similitudes avec celle observée dans le modèle de Miller et Abrahams. Dans le HRS, une chute significative de l'énergie d'activation est notable, tandis que dans le LRS, la variation est minimale. Les approches de NN et de VR ne génèrent que peu de variations.

En conclusion, il apparaît que la distinction entre ces deux modèles n'est pas réalisable en se basant sur l'énergie d'activation thermique, étant donné qu'ils présentent des comportements globalement similaires. De plus, les approches de NN et de VR

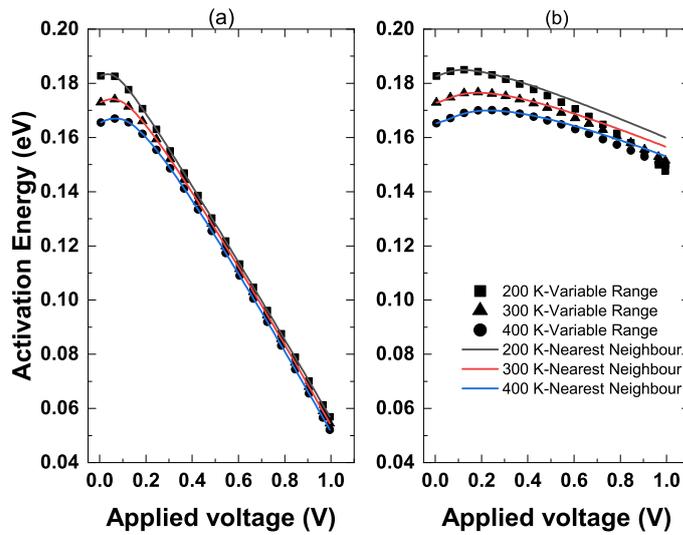


FIGURE 3.41 – Courbe de l'énergie d'activation en fonction de la tension appliquée pour le modèle de (a) Miller et Abrahams en HRS (6 pièges) et (b) en LRS (25 pièges).

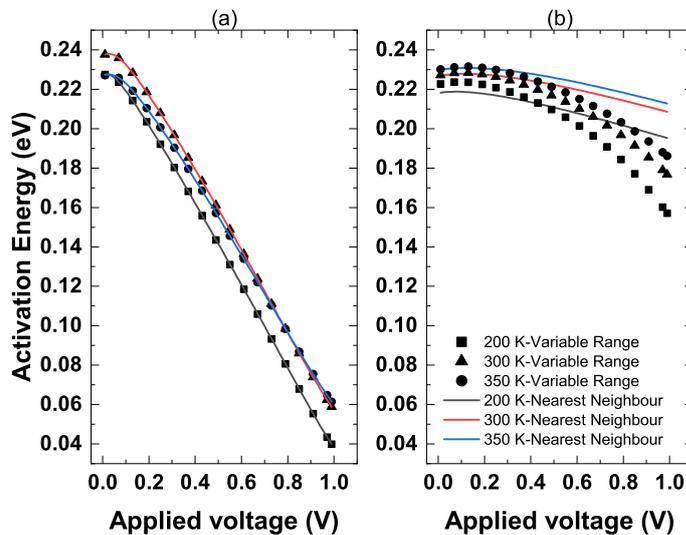


FIGURE 3.42 – Courbe de l'énergie d'activation en fonction de la tension appliquée pour le modèle de Tsu-Esaki en HRS (6 pièges) et $E_t=0.2$ eV et (b) en LRS (25 pièges) et $E_t=0.2$ eV.

ne révèlent aucune différence significative dans les résultats. Nous explorerons par la suite si l'utilisation de la *voltage power function* pourrait nous offrir une voie permettant de différencier ces deux modèles.

Voltage power factor function

Un deuxième paramètre susceptible de nous permettre de différencier les deux modèles est la *Voltage Power function*. Cette fonction est fréquemment utilisée pour des courbes présentant une variation de type $f(x) = x^\gamma$. Elle offre des informations sur la variation du facteur de puissance en fonction de la variable principale. Cet aspect de la fonction est particulièrement utile pour analyser l'allure des courbes I-V, notamment en LRS. En LRS, on s'attend à ce que les courbes aient une allure quasi linéaire. Par conséquent, l'exploration de la variation de ce paramètre nous permettra de comprendre comment le courant évolue en fonction de la tension. On anticipe, en outre, que le paramètre γ soit proche de 1 dans cette configuration. En revanche, l'utilisation de cette fonction en HRS n'a pas de sens car la courbe courant-tension présente une allure exponentielle. Ainsi, dans cette section, notre attention se portera sur l'analyse de la variation de la power voltage function (γ) pour les deux modèles en LRS. L'extraction du paramètre γ à partir des simulations se fera à travers la formule suivante :

$$\gamma = \frac{d \ln(I)}{d \ln(V)} \quad (3.82)$$

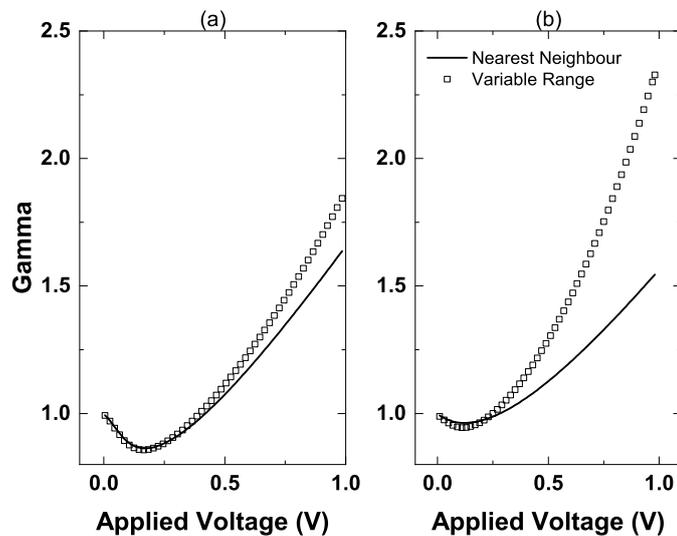


FIGURE 3.43 – (a) Gamma en fonction de la tension appliquée en LRS pour (a) le modèle de Miller et Abrahams and (b) modèle de Tsu-Esaki. 25 pièges ont été utilisés pour les deux simulations et $E_t=0.2$ eV.

La figure 3.43 illustre la variation de la fonction γ en fonction de la tension appliquée, présentant d'une part le modèle basé sur Miller et Abrahams, et d'autre part, le modèle de Tsu-Esaki, dans les approches NN et VR.

Pour le modèle de Miller et Abrahams (figure 3.43 (a)), on observe que γ commence autour de 1 et augmente progressivement jusqu'à atteindre 2. Cette évolution

traduit une allure linéaire à faible tension et une allure quasi quadratique à haute tension. Ces résultats sont cohérents avec les observations expérimentales effectuées sur une structure 1R (structure de test sans transistor de limitation de courant) [18]. En ce qui concerne les approches NN et VR, la principale distinction réside dans la "vitesse" à laquelle γ tend vers 2, cette tendance étant plus rapide en VR, ce qui peut s'expliquer par la méthode de calcul du courant.

Dans le cas du modèle de Tsu-Esaki (figure 3.43 (b)), pour l'approche NN, γ suit une tendance similaire à celle du modèle de Miller et Abrahams. En revanche, en VR, le paramètre γ atteint des valeurs supérieures à 2,5, indiquant une allure de courbe I-V quadratique.

3.6.2 Bilan sur les modèles de conduction

À ce stade, nous pouvons faire le bilan des deux modèles implémentés. Comme exposé précédemment et résumé dans le tableau 3.44, on constate que les allures des courbes en HRS et en LRS présentent des formes presque identiques pour les deux modèles. Il en va de même pour les variations de température, qui montrent un comportement similaire dans les deux cas, comme l'indiquent les courbes d'énergie d'activation. En ce qui concerne la densité de courant, le modèle de Miller et Abrahams se distingue en étant capable d'atteindre les densités requises pour ajuster les courbes I-V d'une OxRAM. Ainsi, il est difficile de discriminer entre les deux

	Allure des I-V	Densité de courant	Energie d'activation	Power voltage function
Modèle de Tsu-Esaki	<ul style="list-style-type: none"> Linéaire en LRS → OK Exponentielle en HRS → OK 	<ul style="list-style-type: none"> 10^{-12} A en HRS 10^{-11} A en LRS 	<ul style="list-style-type: none"> HRS → Chute importante de E_a vs V LRS → Faible variation vs V 	<ul style="list-style-type: none"> $V < 0.25$ V → $\gamma \approx 1$ $V > 0.25$ V → $\gamma \nearrow 2.5$
Modèle de Miller et Abrahams	<ul style="list-style-type: none"> Linéaire en LRS → OK Exponentielle en HRS → OK 	<ul style="list-style-type: none"> 10^{-6} A en HRS 10^{-4} A en LRS 	<ul style="list-style-type: none"> HRS → Chute importante de E_a vs V LRS → Faible variation vs V 	<ul style="list-style-type: none"> $V < 0.5$ V → $\gamma \approx 1$ $V > 0.5$ V → $\gamma \nearrow 2$

FIGURE 3.44 – Bilan sur les deux modèles de conduction

modèles ici, car ils présentent tous deux des tendances générales similaires. Aucune différence notable n'est observée dans les énergies d'activation.

Par conséquent, le seul argument en faveur d'un modèle par rapport à l'autre est que le modèle de Miller et Abrahams parvient à atteindre les niveaux de courant attendus en HRS et en LRS avec un nombre raisonnable de pièges. Ainsi, pour la section suivante, qui consistera à comparer les courbes I-V en HRS et en LRS avec des données expérimentales, seul le modèle de Miller et Abrahams sera retenu.

Dans cette étude, plusieurs simplifications ont été proposées pour chaque modèle, et deux approches ont été développées pour calculer le courant : l'approche Nearest Neighbour (NN) et l'approche Variable Range (VR). Ces simplifications et ces approches de calcul ne sont pas seulement utiles sur le plan pédagogique pour comprendre les phénomènes physiques, mais elles ont également démontré que le temps d'exécution est un facteur déterminant dans le choix des méthodes, comme l'illustrent les graphes 3.45 et 3.46. Ces graphes présentent le temps d'exécution nécessaire pour obtenir une courbe I-V avec 100 points en fonction des différents modèles proposés dans cette étude pour 5 pièges et une énergie $E_t = 0.2$ eV.

Les résultats montrent que les modèles utilisant l'approche NN ont un temps d'exécution nettement inférieur à ceux utilisant l'approche VR. Par exemple, le modèle TE sans simplification met environ 19 secondes à s'exécuter en NN, contre 51 secondes en VR, soit plus du double. De plus, les modèles les plus simplifiés, tels que le modèle TE simplifié qui utilise un système linéaire pour résoudre les niveaux de Fermi, nécessitent environ 100 fois moins de temps d'exécution que le modèle TE complet sans simplification, lequel utilise la méthode de Newton-Raphson.

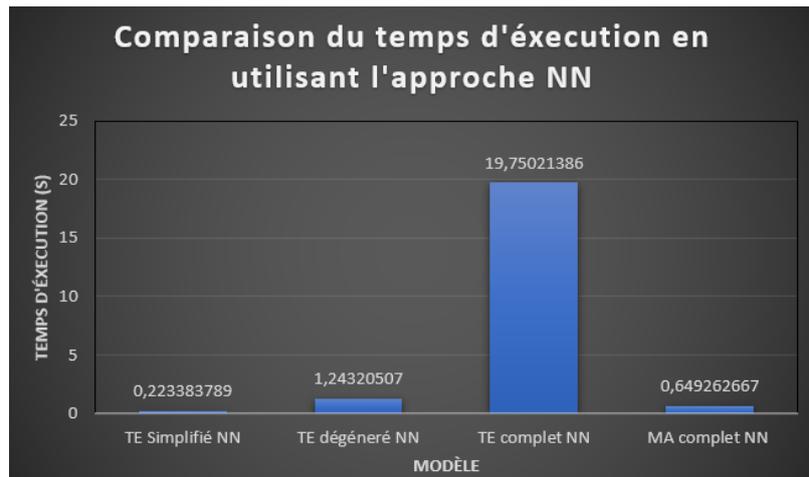


FIGURE 3.45 – Temps d'exécution des des codes utilisant l'approche NN

3.6.3 Comparaison avec des données expérimentales

Dans cette section, nous comparerons les données issues du modèle de Miller et Abrahams avec des résultats expérimentaux représentatifs provenant de la littérature, en utilisant deux types de structures de test. Nous commencerons d'abord par la structure la plus simple, une structure 1R, où le courant est mesuré directement sur l'OxRAM sans transistor de limitation de courant en série. Ensuite, nous examinerons le comportement des structures 1T1R, qui sont des structures de mémoires typiques avec un transistor en série pour protéger l'OxRAM des fluctuations de courant trop importantes, que la compliance de l'instrument de mesure ne peut pas

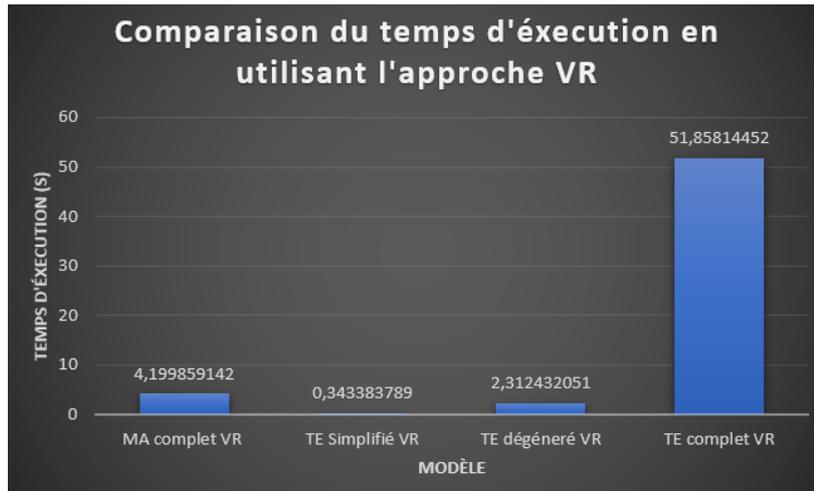


FIGURE 3.46 – Temps d'exécution des des codes utilisant l'approche VR

supporter.

Commençons par examiner le comportement en HRS. La figure 3.47 représente le courant en fonction de la tension en HRS : (a) en utilisant des valeurs expérimentales issues de structures sans transistor et (b) en utilisant des valeurs expérimentales issues de structures avec transistor. Les points représentent les données expérimentales, et la ligne représente les données simulées. On observe qu'avec seulement 6 pièges, un accord presque parfait est obtenu entre les données expérimentales et la simulation. Ainsi, grâce au modèle de Miller et Abrahams, l'allure expérimentale des courbes I-V en HRS est bien reproduite avec un nombre raisonnable de pièges. Il est également important de noter que pour le HRS, le transistor de limitation de courant n'a aucune influence significative. En effet, pour les courants maximums atteints en HRS, le transistor reste dans sa région de conduction, et sa saturation se situe plutôt autour de $100 \mu\text{A}$.

La figure 3.48 présente les courants en fonction de la tension en LRS, à partir de données expérimentales et de simulations, avec (figure 3.48 (b)) et sans (figure 3.48 (a)) transistor MOS en série. Les simulations sont réalisées avec 33 pièges pour la figure 3.48 (a) et 16 pièges pour la figure 3.48 (b).

D'une part, une excellente concordance est observée entre toutes les courbes en dessous de $0,75 \text{ V}$. D'autre part, au-delà de $0,75 \text{ V}$, dans la figure 3.48 (b), la simulation brute indique un courant supérieur à la valeur expérimentale. Cette disparité peut s'expliquer par la présence du transistor en série avec l'OxRAM, conçu pour limiter le courant pendant la programmation. Afin de tenir compte de la présence de ce MOSFET en série, il est nécessaire de déterminer le courant passant par l'OxRAM lorsque le transistor est en saturation. Pour ce faire, il faut résoudre

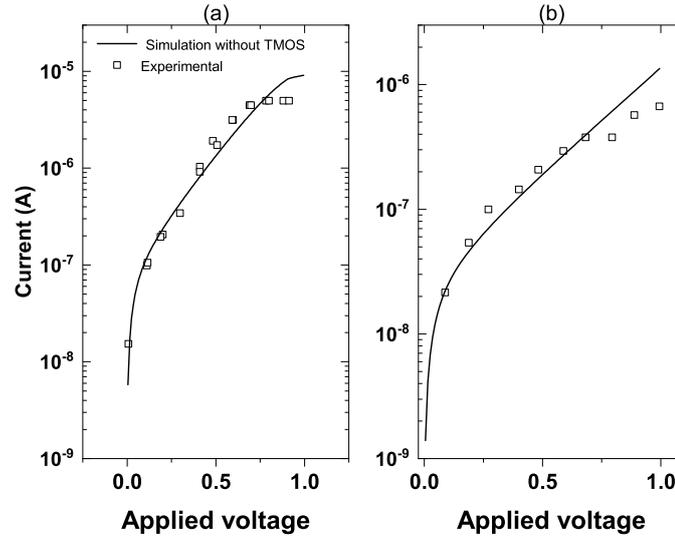


FIGURE 3.47 – Comparaison entre les courbes I-V issues du modèle de Miller et Abrahams model et les données expérimentales issues de : (a) Source : [18] et (b) Source : [19], en HRS. Les paramètres utilisées : (a) 6 pièges, $w_0 = 3.5 \times 10^{18} \text{ s}^{-1}$, $\alpha = 0.15 \times 10^{-10} \text{ m}$, $E_t = 0.13 \text{ eV}$ et (b) 6 pièges, $w_0 = 10^{17} \text{ s}^{-1}$, $\alpha = 10^{-10} \text{ m}$, $E_t = 0.2 \text{ eV}$

l'équation de courant $I_{OxRAM} = I_d$, formulée par l'équation 3.83.

$$I_d = \mu_n C_{ox} \frac{W}{L} \begin{cases} (V_g - V_T - \frac{V_d}{2}) V_d & \text{si } V_d < V_{dsat} \\ \frac{1}{2} (V_g - V_T)^2 & \text{si } V_d \geq V_{dsat} \end{cases} \quad (3.83)$$

$$= I_{OxRAM}$$

Où $V_{dsat} = V_g - V_T$.

Une approche relativement simple pour résoudre cette équation consiste à effectuer une dichotomie sur la variable principale V_d . Une fois cette tension obtenue, il est possible de déterminer le courant à partir de cette tension, car elle correspond au courant de drain du transistor. Ainsi, une fois cette équation résolue, on observe un bien meilleur accord avec les données expérimentales. Il devient alors évident que le transistor MOS a une influence significative sur la courbe de courant, comme illustré dans la figure 3.48 (b) avec TMOS. En effet, une fois l'équation 3.83 résolue, on constate que la courbe avec TMOS devient constante (saturation du transistor) au-delà d'une certaine tension, tandis que la courbe représentant le courant sans TMOS sur la même figure suit une tendance plus linéaire.

En conclusion, le modèle de Miller et Abrahams parvient à reproduire de manière précise la caractéristique I-V dans une OxRAM, aussi bien dans l'état HRS que dans le LRS, et ce, quel que soit le type de structure, que ce soit en 1T ou en 1T1R. Comme discuté précédemment lors de l'analyse de l'énergie d'activation et de la

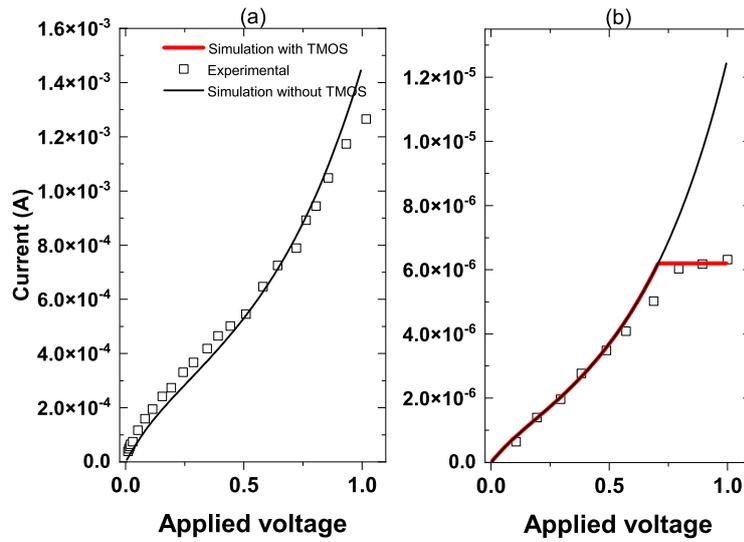


FIGURE 3.48 – Comparaison entre les courbes I-V issues du modèle de Miller et Abrahams et les données expérimentales issues de : (a) [18] et (b) [19], en LRS. Les paramètres utilisées : (a) 33 pièges, $w_0 = 3.5 \times 10^{18} \text{ s}^{-1}$, $\alpha = 0.15 \times 10^{-10} \text{ m}$, $E_t = 0.13 \text{ eV}$ et (b) 16 pièges, $w_0 = 10^{17} \text{ s}^{-1}$, $\alpha = 10^{-10} \text{ m}$, $E_t = 0.2 \text{ eV}$

power voltage function, le modèle de Miller et Abrahams ainsi que le modèle de Tsu-Esaki présentent une tendance générale similaire. Cependant, en ce qui concerne les niveaux de courant, le modèle de Miller et Abrahams offre une meilleure concordance avec les résultats expérimentaux, tout en utilisant des paramètres de simulation raisonnables.

Chapitre 4

Etude experimentale

4.1 Introduction

Dans cette thèse, qui se concentre principalement sur un travail de modélisation, le choix a été fait d'inclure une partie expérimentale, assez différente de ce qui a été exposé jusqu'à présent. Cependant, il est crucial de souligner que l'objectif de cette thèse est de modéliser la conduction dans les OxRAM pour les deux états de résistance. Afin de comprendre en profondeur les mécanismes des modèles de courants, il est nécessaire d'avoir une vision claire de la variation du courant en fonction de la température. Ceci implique la réalisation de mesures précises en température et donc l'obtention de courbes d'énergie d'activation expérimentales pour vérifier l'adéquation du modèle aux mesures.

Cependant, les mesures de courants en température sur ce type de dispositif sont complexes en raison de leur variabilité intrinsèque. Par conséquent, ces mesures peuvent parfois conduire à des conclusions contradictoires. Par exemple, dans la référence [104], il est rapporté que l'énergie d'activation (E_a) est quasi nulle voire négative en LRS et positive en HRS, tandis que dans une étude similaire [105], il est rapporté que E_a est nulle en HRS et en LRS.

Afin de clarifier cette contradiction observée dans la littérature et d'obtenir des mesures fiables capables de valider ou d'invalider les modèles proposés, des mesures de courant en température ont été réalisées. Ces mesures permettront d'obtenir une compréhension précise de la variation du courant en fonction de la température dans les deux états de résistance. En cas d'impossibilité d'obtenir des résultats concluants sur cette variation du courant, l'accent sera mis sur la quantification de la variabilité de l'énergie d'activation thermique.

Pour ce faire, des mesures en température ont été effectuées sur une trentaine de dispositifs en HRS et en LRS, en faisant varier la température. De plus, une étude de la variabilité cycle à cycle a été entreprise en faisant commuter les dispositifs une centaine de fois à chaque température. Cette analyse permettra de comprendre

comment l'énergie d'activation varie après ces cycles.

Grâce à ces différentes mesures, l'objectif est d'obtenir des données de courant en fonction de la température pour tester au maximum le modèle proposé et de vérifier sa robustesse.

4.2 Description du protocole expérimentale

4.2.1 Présentation de l'Elite 300

La station de traitement Elite 300 Cascade dont nous disposons est une station sous pointe semi-automatique industrielle, conçue pour le traitement de wafer d'une dimension maximale de 12 pouces et minimale de 4 pouces. Elle est équipée d'une platine unique qui permet un déplacement avec une précision inférieure à 1 μm . De plus, elle dispose d'un module de température pour le traitement thermique des échantillons allant de 25 à 150°C, ce qui permet de réaliser des mesures à haute température. La station est également équipée d'un double blindage Attoguard + pour minimiser les interférences électromagnétiques. Enfin, elle permet la mesure de courant jusqu'au fA ou des capacités de fF. La principale avantage de ce type de station par rapport à une station sous pointe manuelle réside dans sa capacité à automatiser les mesures répétitives sur un ensemble de dispositifs. En effet, grâce à la possibilité de cartographier le wafer et de créer un programme de test, les opérations peuvent être exécutées de manière automatique. Une fois les pointes positionnées initialement, la station est capable de se déplacer sans nécessiter d'intervention humaine, ce qui simplifie considérablement le processus de mesure.

4.2.2 Présentation de l'Agilent B1500

Le dispositif de mesure est composé du B1500A, un analyseur de paramètres de dispositifs à semi-conducteurs de la marque Keysight. Ce système permet la réalisation de différentes courbes telles que $I(V)$, $C(V)$, $I(V)$ pulsé/dynamique, etc. Il est équipé de quatre Source Monitor Unit (SMU) pour la mesure des signaux, ainsi que de deux autres unités SMU pour la mesure des puissances moyennes (MPSMU). Le B1500A est capable de générer les deux tensions de commande requises pour alimenter la structure 1T1R et mesure le courant traversant l'OxRAM. En d'autres termes, il fournit les tensions nécessaires à la manipulation de la structure 1T1R et enregistre le courant qui passe à travers l'OxRAM.

En associant le B1500A à la station sous pointe l'Elite 300, et en utilisant son outil de création de programme de test, nous avons pu élaborer un programme de test spécifique. Ce programme nous a permis de contrôler la station sous pointe, d'effectuer nos mesures, et de récupérer les données de mesure dans le format désiré

pour un traitement ultérieur.

4.2.3 Dispositif testé

Nous proposons une caractérisation électrique des cellules mémoires OxRAM intégrées présentant une architecture 1T1R. Cette architecture exploite les caractéristiques d'un transistor MOS intégré en série avec un élément résistif afin d'assurer un contrôle efficace du courant lors des opérations d'écriture et d'effacement (voir figure). Par conséquent, par rapport aux cellules 1R, le fonctionnement de cette architecture mémoire est influencé par les caractéristiques et la polarisation du transistor MOS, comme nous l'avons démontré dans le chapitre précédent.

Dans notre étude, nous avons travaillé avec un empilement OxRAM TiN/HfO₂/Ti/TiN. L'OxRAM a été intégré dans le BEOL (Back-End-Of-Line) d'un processus CMOS de 28 nm sur des wafers de 300 mm (voir figure 4.1 (a)). Comme illustré sur les figures, l'OxRAM est connecté au drain du transistor FEOL (Front-End-Of-Line). En termes de structure, l'OxRAM est constitué d'un film d'oxyde de HfO₂ de 5 nm déposé par dépôt ALD à 300 °C, d'une électrode inférieure (BE) en TiN, et d'une électrode supérieure (TE) sous forme d'une structure bicouche Ti/TiN, déposée par PVD à 55 °C.

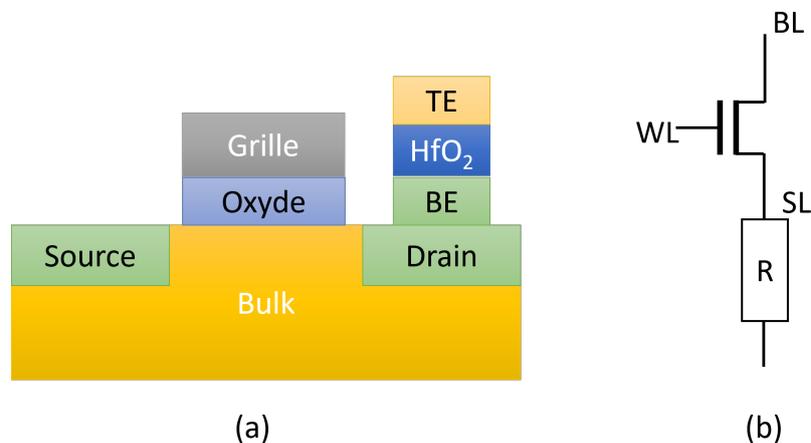


FIGURE 4.1 – (a) Schéma de la structure de test 1T1R. (b) Schéma équivalent de la structure de test

La figure 4.1 (b) représente le circuit équivalent de la structure 1T1R. La grille du transistor est définie comme la WorldLine (WL), la source comme la SourceLine (SL), et l'électrode supérieure (TE) de l'OxRAM est définie comme la borne BitLine (BL).

	FORMING	SET	RESET
V_{WL}	4.5	2.5	4.5
V_{BL}	3	2	0
V_{SL}	0	0	2.6

TABLE 4.1 – Tableau récapitulatif des valeurs de tension de Forming, SET et RESET

4.2.4 Plan de mesure

Pour débiter les mesures concernant les caractéristiques de l'OxRAM, il est nécessaire de choisir les tensions appropriées pour la grille, le drain et la source afin de limiter le courant traversant l'OxRAM. Le tableau 4.1 montre les tensions V_{WL} , V_{BL} , V_{SL} à utiliser pour éviter les pics de courant susceptibles d'endommager l'OxRAM. Ces spécifications nous ont été transmises par nos collaborateurs du CEA-LETI. En raison d'un module de courant pulsé défectueux, les opérations de FORMING, SET et RESET sont effectuées en mode balayage en tension plutôt qu'en mode pulsé, ce qui est plus conventionnel pour ce type de mesure. Pour obtenir nos courbes I-V en fonction de la température et par la suite nos énergies d'activation, nous proposons le protocole suivant :

1. Effectuer un FORMING à 25°C en appliquant une tension de grille constante sur le transistor $V_{WL} = 4.5V$, en effectuant un balayage en tension pour V_{BL} de 0 à 3V sur les 30 dispositifs.
2. Effectuer un balayage en tension sur V_{BL} allant de 0 à 2V pour obtenir une courbe courant-tension en LRS.
3. Répéter l'étape 2 sur les 30 dispositifs.
4. Chauffer le prober de 5°C et laisser reposer pendant 30 minutes.
5. Répéter les étapes 2 et 3 à cette température.

Une fois les mesures à 45°C réalisées, laisser le prober refroidir à 25°C. Une fois revenu à 25°C :

6. Effectuer un RESET à 25°C en appliquant une tension constante de 4.5 V sur V_{WL} et un balayage en tension de 2.6 V sur V_{SL} .
7. Effectuer un balayage en tension sur V_{BL} allant de 0 à 2 V pour obtenir une courbe courant-tension en HRS.
8. Répéter l'étape 7 sur les 30 dispositifs.
9. Chauffer le prober de 5°C et laisser reposer pendant 30 minutes.
10. Répéter les étapes 7 et 8 à cette température.

Une fois les mesures à 45°C réalisées, laisser le prober refroidir à 25°C. Une fois revenu à 25°C :

11. Effectuer un SET en appliquant une tension constante V_{WL} de 2.5 V et en effectuant un balayage V_{BL} allant de 0 à 2 V.
12. Une fois ce SET effectué, reprendre à partir de l'étape 2 et effectuer 100 cycles.

Ce protocole expérimentale a été résumé par la figure 4.2, qui reprend d'un point schématiquement le plan exposé ci avant.

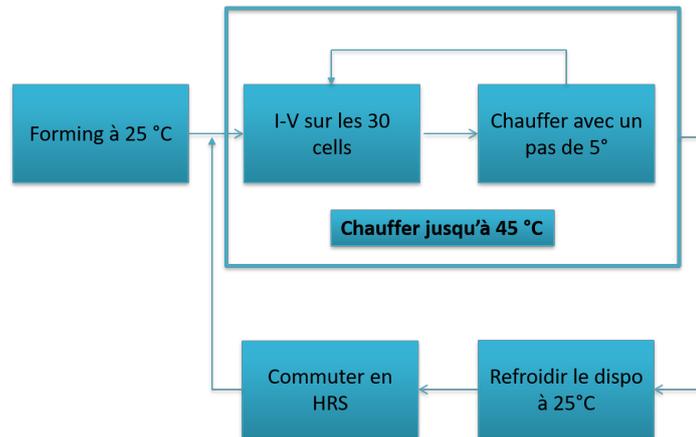


FIGURE 4.2 – Schéma du plan d'expérience

Il convient de souligner que le choix de ce protocole n'a pas été fait de manière arbitraire. En effet, la commutation (SET et RESET) est réalisée à une température constante de 25°C, car l'objectif principal de ces mesures est d'analyser la variation du courant en fonction de la température, et non de la commutation en elle-même. De plus, la décision d'effectuer 100 cycles sur 30 dispositifs différents a été prise dans le but d'étudier le comportement de l'énergie d'activation d'un cycle à l'autre et d'un dispositif à l'autre.

Initialement, la variation de température prévue pour l'expérience devait s'étendre de 25 à 150 °C. Cependant, un problème est survenu : les dispositifs étaient endommagés dès que la température dépassait 45 °C. En conséquence, nous avons été contraints de limiter la plage de température à 45 °C pour éviter tout d'endommager les dispositifs.

4.3 Exploitation des données

Cette section vise à examiner de manière approfondie les résultats obtenus à partir de la campagne de mesure. Nous allons analyser les différentes données recueillies et explorer leurs implications.

Les figures 4.3 et 4.4 illustrent les courbes I-V à différentes températures, respectivement en LRS et en HRS. Il est observé que les fluctuations de courant caractéristiques des OxRAM [106], rendent pratiquement impossible toute interprétation.

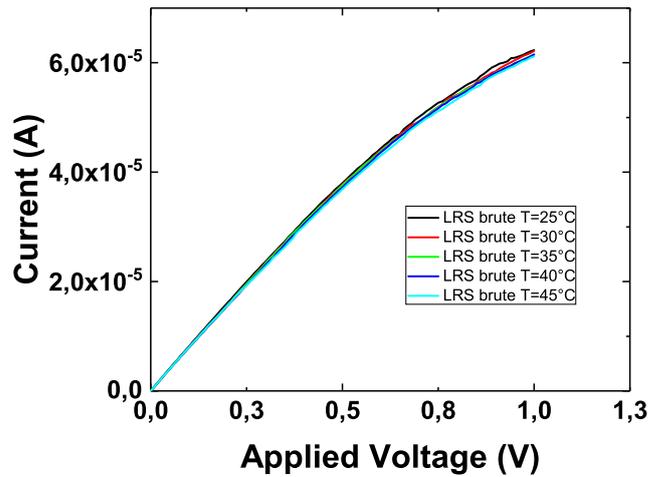


FIGURE 4.3 – Courbe I-V en LRS issue des données expérimentale brute.

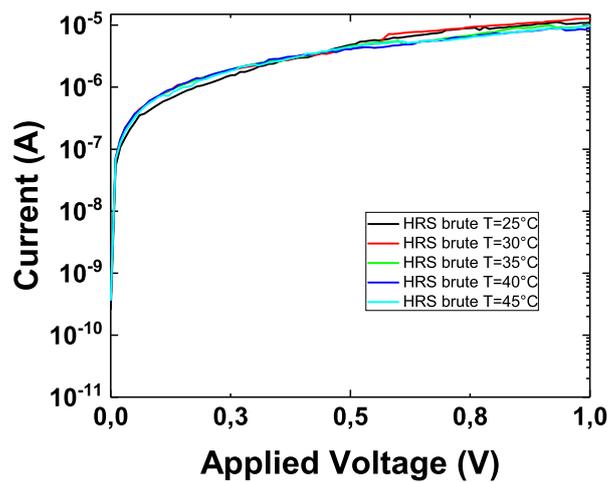


FIGURE 4.4 – Courbe I-V en HRS issue des données expérimentale brute

Il est important de mentionner que nous avons initialement prévu de réaliser 100 cycles de mesure. Cependant, nous n'en avons finalement présenté qu'un seul. En effet, après seulement 10 cycles, nos composants ont commencé à se détériorer de manière significative, rendant impossible la réalisation des 100 cycles prévus. C'est pourquoi nous avons choisi de ne montrer qu'un seul cycle de mesure dans nos résultats.

De même, les courbes d'énergie d'activation extraites des données brutes, présentées dans les figures 4.7 (a) et 4.8 (a), présentent les mêmes difficultés d'interprétation. Pour tenter de faciliter l'analyse des courbes obtenues expérimentalement, nous avons décidé de limiter la tension à 0,4 V et d'appliquer un processus de lissage aux différentes courbes. Ce lissage a été effectué en utilisant la méthode de la moyenne

glissante. Cette technique implique de remplacer chaque valeur du tableau par la moyenne des valeurs environnantes. Cette moyenne est dite glissante car elle est recalculée de manière continue, en utilisant à chaque fois un sous-ensemble d'éléments où un nouvel élément remplace le plus ancien.

Les résultats de ce processus pour le LRS sont présentés sur la figure 4.5 et pour le HRS sur la figure 4.6. Chaque graphique représente un seul dispositif.

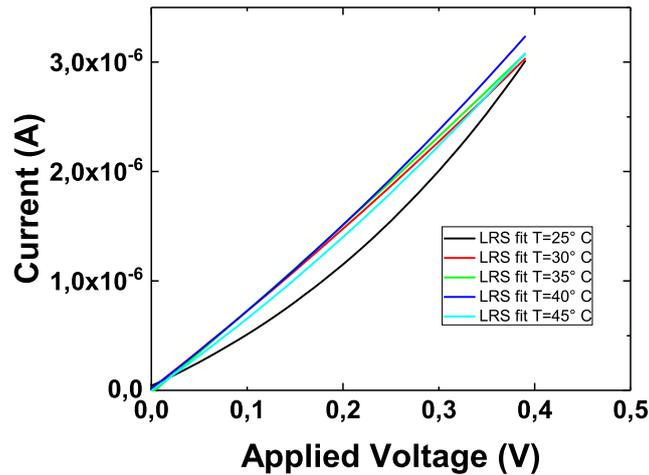


FIGURE 4.5 – Courbe I-V en LRS issue des données expérimentale ajustées.

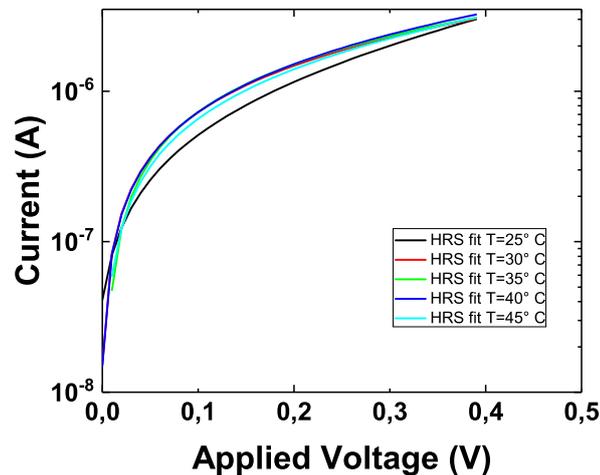


FIGURE 4.6 – Courbe I-V en HRS issue des données expérimentale ajustées.

On remarque qu'en fonction de la température, le courant en LRS ne varie quasiment pas. On remarque le même effet sur la totalité des autres cellules. Dans la Figure 4.4, la courbe I-V en fonction de la température est représentée pour une seule cellule. Contrairement au LRS, on observe que le courant est plus sensible à

la température, montrant une variation significative entre les différentes températures. La fluctuation en courant est très importante, comme le confirment les deux graphiques d'énergies d'activation, 4.7 (b) et 4.8 (b).

En effet, les figures 4.8 (b) et 4.7 (b) représentent l'énergie d'activation en fonction de la tension à différentes températures, respectivement en LRS et en HRS. On remarque qu'aucune tendance claire ne se dégage à la fois en HRS et en LRS. L'analyse de ces différentes courbes ne fournit donc aucune information pertinente sur le comportement du courant en fonction de la température.

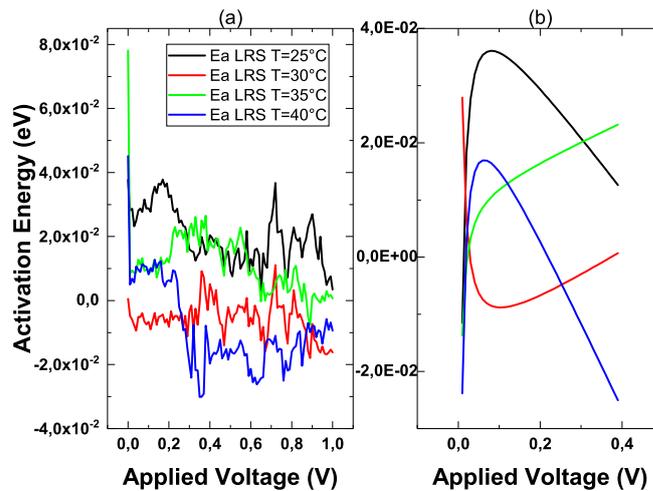


FIGURE 4.7 – Courbe E_a -V en LRS issue des données expérimentale : (a) Brute (b) Ajustées

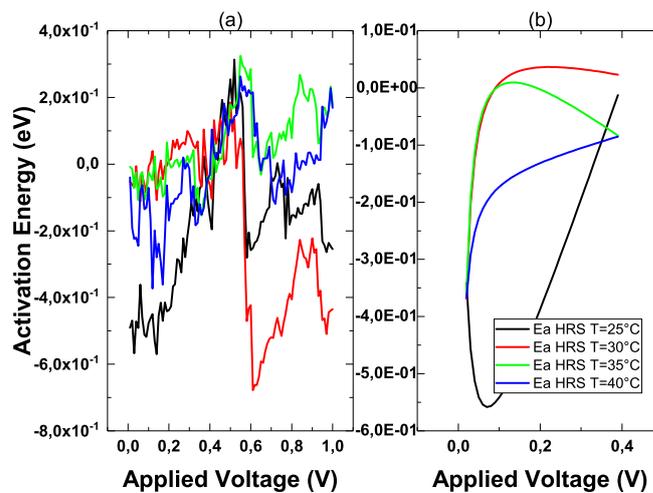


FIGURE 4.8 – Courbe E_a -V en HRS issue des données expérimentale : (a) Brute (b) Ajustées

À noter également que la courbe I-V à 25°C croise celle d'une température supérieure, ce qui complique considérablement l'interprétation des résultats.

Il semble que l'incapacité à exploiter les résultats soit largement attribuable à la méthode utilisée pour les mesures. Dans le plan de mesure, il est noté que les commutations sont effectuées avec un balayage en tension, en raison de la défaillance du module permettant de créer des impulsions sur le banc de mesure. Malheureusement, ce module n'a pu être réparé qu'après la conclusion de la thèse. Par conséquent, le choix de procéder aux commutations par balayage en tension a probablement contribué à endommager les cellules. De plus, il a été observé que la variation du courant en fonction de la température est moindre par rapport à la variabilité et aux fluctuations observées dans l'oxRAM. Cela suggère qu'une plage de température plus étendue aurait dû être étudiée pour obtenir des résultats plus significatifs. Toutefois, en raison de la limitation à 45°C, il serait nécessaire de prolonger les mesures à des températures plus basses. Il est à noter qu'une autre stratégie expérimentale est proposée dans la section dédiée aux perspectives.

Chapitre 5

Conclusion

Les technologies de mémoire telles que les NVM émergentes, et plus spécifiquement les OxRAMs, suscitent un grand intérêt en raison de leur caractéristiques intéressantes et prometteuses. Cependant, l'un des principaux défauts des OxRAMs réside dans leur variabilité importante en courant. Malgré des recherches intensives visant à modéliser cette variabilité, aucun modèle connu ne parvient à reproduire de manière précise cette caractéristique. Ainsi, pour aborder cette variabilité, il est essentiel de disposer de modèles capables de représenter fidèlement le comportement électrique global de l'OxRAM. Cela implique de modéliser à la fois les mécanismes de forming, de commutation et de conduction. Dans cette thèse, l'accent a été exclusivement mis sur le mécanisme de conduction. Bien que de nombreux modèles aient été proposés dans la littérature, la plupart d'entre eux reposent sur l'utilisation de deux modèles pour reproduire le comportement électrique attendu.

Dans cette thèse, l'objectif principal a été de développer un modèle de conduction innovant capable de reproduire le comportement électrique d'une OxRAM dans ses deux états de résistance. Plusieurs contributions ont été apportées dans ce cadre. Tout d'abord, contrairement aux approches antérieures de la littérature, l'accent a été mis sur l'équation de continuité. Une décision a été prise pour réexaminer cette équation et la résoudre afin d'obtenir des valeurs des niveaux de Fermi permettant de maintenir un courant constant entre les pièges. Ainsi, un nouveau formalisme a été élaboré, permettant de résoudre les niveaux de Fermi pour tous les pièges de manière indépendante de l'équation de courant utilisée. Une fois ces niveaux de Fermi obtenus, ils pourront être injectés dans les équations de courant. Le formalisme développé se décline en deux approches distinctes : l'approche du Nearest Neighbour, qui ne considère que les voisins les plus proches, et l'approche Variable Range, qui prend en compte la conduction entre l'ensemble des pièges. Ce formalisme a ensuite été appliqué en utilisant les deux approches mentionnées à deux modèles de conduction distincts. Le premier modèle est le modèle TAT basé sur l'équation de Tsu Esaki. Dans ce modèle, une hypothèse forte est sous-entendue : il considère que les

pièges forment un continuum d'états. Bien que ce modèle ait permis d'obtenir des courbes I-V présentant les bonnes tendances en LRS et en HRS, tout en respectant la continuité du courant, il s'est avéré incapable d'ajuster précisément les courbes courant-tension en raison du faible courant de sortie à la fois en LRS et en HRS. Le deuxième modèle est un modèle de hopping, basé sur l'équation de Miller et Abrahams. Ce dernier considère que les pièges sont des états liés. Grâce à ce modèle, il a été possible d'ajuster les courbes courant-tension dans les deux états de conduction en utilisant un nombre de pièges raisonnable par rapport à ce qui est rapporté dans la littérature. Dans l'implémentation de ces deux modèles, plusieurs variantes ont été explorées, avec des versions simplifiées proposées avant d'aboutir à la version complète de chaque modèle. Pour le modèle TAT, différentes fonctions d'occupation ont été testées, notamment la fonction de Boltzmann et la fonction de Fermi-Dirac. Une version avec un niveau de Fermi dégénéré a également été proposée. Pour le modèle de Miller et Abrahams, les simplifications ont consisté à ne pas prendre en compte tous les effets de couplage dès le début ou à considérer les fonctions d'occupation comme des fonctions de Boltzmann. Une fois ces implémentations réalisées, un travail critique des résultats a été entrepris, mettant en lumière les limites de chaque modèles.

Afin de tenter de discriminer entre les deux modèles, des paramètres expérimentaux tels que l'énergie d'activation thermique et la fonction de puissance tension ont été représentés avec les données de simulation. Ces deux paramètres ont montré que le même comportement global se dégageait des courbes produites par ces deux modèles, et qu'il n'était pas possible de distinguer les deux modèles à travers l'énergie d'activation thermique et la fonction de puissance tension. La seule manière de distinguer les deux modèles réside dans le fait que le modèle de Miller et Abrahams atteint les niveaux de courant escomptés dans les deux états de conduction, tandis que le modèle TAT basé sur l'équation de Tsu-Esaki ne parvient pas à le faire.

Pour évaluer la robustesse de notre modèle, il était essentiel de le confronter à des données expérimentales. En utilisant le modèle de Miller et Abrahams, nous avons réussi à reproduire les courbes I-V des OxRAMs en HRS et en LRS avec un nombre de pièges raisonnable. Cela nous a confortés dans l'idée que le modèle de Miller et Abrahams était le plus à même de modéliser la caractéristique électrique d'une OxRAM.

Enfin, des mesures I-V en fonction de la température ont été réalisées afin de mieux comprendre comment le courant varie en fonction de la température et de vérifier si notre modèle était capable de reproduire cette variation des caractéristiques I-V. Cependant, les résultats obtenus n'ont pas permis de tirer des conclusions claires sur la variation du courant en fonction de la température. Cette difficulté est principalement due au fait que l'amplitude de la variation de température était trop faible par rapport à la variabilité inhérente aux composants et aux fluctuations observées,

ce qui a limité notre capacité à discerner des tendances significatives.

Chapitre 6

Perspectives

Dans le cadre de notre étude, il est essentiel de souligner quelques éléments qui nécessitent des ajustements ou des améliorations : la prise en compte du bruit RTN, la réalisation de nouvelles simulations avec des distributions de pièges plus réalistes, et la reprise de la caractérisation en utilisant un plan de mesure plus adéquat.

6.1 Prise en compte du bruit RTN

la prise en compte du bruit RTN (Random Telegraph Noise). Ce phénomène joue un rôle essentiel dans le fonctionnement des dispositifs, et sa modélisation est basée sur la description du comportement aléatoire des pièges présents dans le matériau. Chaque piège a une certaine probabilité de capturer une charge, une probabilité qui est influencée par la barrière énergétique et la température ambiante. Le processus de capture et d'émission des charges par ces pièges est généralement représenté à l'aide de distributions exponentielles, comme l'a démontré Puglisi *et al.* [20] (voir figure 6.1).

Pour étudier de manière appropriée le bruit RTN dans notre modèle, il serait pertinent d'activer et de désactiver un piège afin de simuler la variation de courant δI induite. Cette approche permettrait de vérifier si la réponse simulée correspond aux observations expérimentales. En outre, il serait envisageable d'établir une corrélation entre les variations de courant δI et la densité de pièges présents dans le filament.

6.2 Simulation Réaliste

Dans l'ensemble des simulations réalisées, nous avons considéré une distribution uniforme des pièges à travers l'épaisseur de l'oxyde. Par exemple, pour simuler un état HRS avec 5 pièges, ces pièges ont été répartis uniformément, avec un piège à 1 nm, un autre à 2 nm, et ainsi de suite. Cependant, en examinant les résultats des simulations KMC (Kinetic Monte Carlo), il apparaît que la distribution des

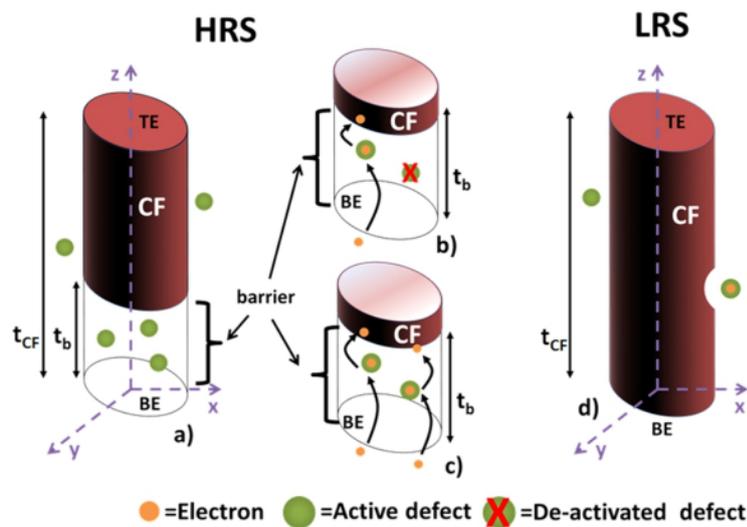


FIGURE 6.1 – Schéma du processus de suppression de piège pour modélisation du bruit RTN. Source : [20].

pièges, en particulier dans l'état HRS, n'est pas uniforme. La majorité des pièges se concentre généralement vers le sommet du filament conducteur (coté électrode supérieur), tandis que la région proche de l'électrode inférieure est presque dépourvue de pièges.

Compte tenu de cette observation, il serait pertinent de refaire les simulations, notamment pour l'état HRS, en utilisant des distributions de pièges plus réalistes, alignées sur celles observées dans les résultats des simulations KMC. Cela permettrait d'obtenir des résultats plus fidèles à la réalité physique des dispositifs.

6.3 Nouveau plan de mesure

Comme montré dans la partie expérimentale, le plan de mesure précédemment élaboré n'a malheureusement pas permis d'obtenir des résultats concluants, notamment en ce qui concerne l'activation en température des courants dans l'OxRAM. Le principal problème rencontré était la dégradation des cellules OxRAM lors de l'augmentation de la température. En effet, nous n'avons pu effectuer qu'une variation de 20°C, ce qui était insuffisant pour observer un effet notable de la température sur le courant.

Pour remédier à ce problème, nous proposons un nouveau plan de mesure, résumé par la figure 6.2. Ce plan consistera à inverser l'approche en réduisant la température, plutôt qu'en l'augmentant. Des mesures cryogéniques seront réalisées en abaissant la température par paliers de 10°C jusqu'à 100K, avec des courbes I-V mesurées à chaque palier. Les I-V seront obtenues en appliquant un balayage en tension de 0 à 2V sur V_d pour l'état LRS, et de 0 à 2V sur V_s pour l'état HRS, tout en maintenant

une tension de grille fixe à $V_g = 2.5\text{V}$. Les opérations de commutation et de forming seront effectuées à 25°C , en utilisant le module de mesures pulsées.

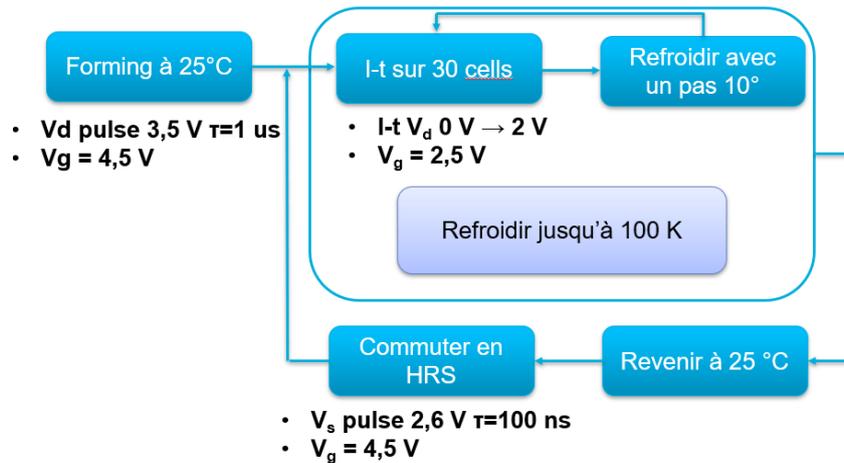


FIGURE 6.2 – Nouveau plan de mesure pour caractérisation I-V en température.

Avec ce nouveau plan de mesure, nous espérons être en mesure de capturer l'activation du courant en fonction de la température. Cependant, il est important de noter que cette approche présente deux inconvénients majeurs :

- Le mécanisme de conduction du courant pourrait différer entre les basses et les hautes températures. Cela signifie que bien que nous puissions observer une activation en fonction de la température, celle-ci pourrait ne pas correspondre à celle que l'on aurait observée à haute température. En d'autres termes, les mécanismes de transport dominants pourraient changer, ce qui compliquerait l'interprétation des résultats en termes de comportement thermique du courant.
- Avec le matériel disponible au laboratoire, il faut environ 5 heures pour effectuer des mesures sur 10 dispositifs avec une variation de température allant de 300 K à 100 K. Si l'objectif est de réaliser cette manipulation sur plus de 100 dispositifs, cela nécessiterait environ une semaine de mesures continues. Il faudra évaluer la faisabilité de cette approche, compte tenu des coûts élevés que cela impliquerait.

Bibliographie

- [1] <http://www-cs-faculty.stanford.edu/uno/abcde.html>, “How much data is generated every day (2024),” 2016. Accessed on 01.09.2016.
- [2] C. Balestra, Graef and Gargini, “The dawn of the new computer industry,” in *International Roadmap for Devices and Systems (IRDS) 2021 Edition*, pp. 43–49, IEEE, 2021.
- [3] S. Kawashima and J. S. Cross, *FeRAM*, pp. 279–328. Boston, MA : Springer US, 2009.
- [4] C. Marchand, I. O’Connor, M. Cantan, E. T. Breyer, S. Slesazeck, and T. Mikolajick, “A fefet-based hybrid memory accessible by content and by address,” *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits*, vol. 8, no. 1, pp. 19–26, 2022.
- [5] F. M. Simanjuntak, D. Panda, K.-H. Wei, and T.-Y. Tseng, “Status and prospects of zno-based resistive switching memory devices,” *Nanoscale research letters*, vol. 11, pp. 1–31, 2016.
- [6] H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. T. Chen, and M.-J. Tsai, “Metal–oxide rram,” *Proceedings of the IEEE*, vol. 100, no. 6, pp. 1951–1970, 2012.
- [7] D. Garbin, E. Vianello, O. Bichler, Q. Rafhay, C. Gamrat, G. Ghibaudo, B. De-Salvo, and L. Perniola, “Hfo 2-based oxram devices as synapses for convolutional neural networks,” *IEEE Transactions on Electron Devices*, vol. 62, no. 8, pp. 2494–2501, 2015.
- [8] D. A. Robayo, G. Sassine, Q. Rafhay, G. Ghibaudo, G. Molas, and E. Nowak, “Endurance statistical behavior of resistive memories based on experimental and theoretical investigation,” *IEEE Transactions on Electron Devices*, vol. 66, no. 8, pp. 3318–3325, 2019.
- [9] Y. Y. Chen, M. Komura, R. Degraeve, B. Govoreanu, L. Goux, A. Fantini, N. Raghavan, S. Clima, L. Zhang, A. Belmonte, *et al.*, “Improvement of data retention in hfo 2/hf 1t1r rram cell under low operating current,” in *2013 IEEE International Electron Devices Meeting*, pp. 10–1, IEEE, 2013.

-
- [10] S. Hamdioui, H. A. Du Nguyen, M. Taouil, A. Sebastian, M. Le Gallo, S. Pande, S. Schaafsma, F. Catthoor, S. Das, F. G. Redondo, *et al.*, "Applications of computation-in-memory architectures based on memristive devices," in *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 486–491, IEEE, 2019.
- [11] L. Vandelli, A. Padovani, L. Larcher, R. Southwick, W. Knowlton, and G. Bersuker, "A physical model of the temperature dependence of the current through sio₂/hfo₂ stacks," *IEEE Transactions on Electron Devices*, vol. 58, no. 9, pp. 2878–2887, 2011.
- [12] L. Larcher, "Statistical simulation of leakage currents in mos and flash memory devices with a new multiphonon trap-assisted tunneling model," *IEEE Transactions on Electron Devices*, vol. 50, no. 5, pp. 1246–1253, 2003.
- [13] D. Ielmini, A. S. Spinelli, A. L. Lacaita, A. Martinelli, and G. Ghidini, "A recombination-and trap-assisted tunneling model for stress-induced leakage current," *Solid-State Electronics*, vol. 45, no. 8, pp. 1361–1369, 2001.
- [14] D. Ielmini, A. Spinelli, A. L. Lacaita, and A. Modelli, "A new two-trap tunneling model for the anomalous stress-induced leakage current (silc) in flash memories," *Microelectronic engineering*, vol. 59, no. 1-4, pp. 189–195, 2001.
- [15] J. M. Andrade, C. M. Rosário, S. Menzel, R. Waser, and N. A. Sobolev, "Application of the quantum-point-contact formalism to model the filamentary conduction in ta₂o₅-based resistive switching devices," *Physical review applied*, vol. 17, no. 3, p. 034062, 2022.
- [16] E. W. Lim and R. Ismail, "Conduction mechanism of valence change resistive switching memory : a survey," *Electronics*, vol. 4, no. 3, pp. 586–613, 2015.
- [17] K.-H. Jang, S.-M. Oh, H.-M. An, and W.-J. Cho, "Solution-processed high-k thin films as a resistive switching for rram applications," *Current Applied Physics*, vol. 14, no. 3, pp. 462–466, 2014.
- [18] H. Lee, P. Chen, T. Wu, Y. Chen, C. Wang, P. Tzeng, C. Lin, F. Chen, C. Lien, and M.-J. Tsai, "Low power and high speed bipolar switching with a thin reactive ti buffer layer in robust hfo₂ based rram," in *2008 IEEE International Electron Devices Meeting*, pp. 1–4, IEEE, 2008.
- [19] Y. Y. Chen, R. Degraeve, B. Govoreanu, S. Clima, L. Goux, A. Fantini, G. S. Kar, D. J. Wouters, G. Groeseneken, and M. Jurczak, "Postcycling lrs retention analysis in hfo₂/hf rram 1t1r device," *IEEE Electron Device Letters*, vol. 34, no. 5, pp. 626–628, 2013.
- [20] F. M. Puglisi, N. Zagni, L. Larcher, and P. Pavan, "Random telegraph noise in resistive random access memories : Compact modeling and advanced circuit design," *IEEE Transactions on Electron Devices*, vol. 65, no. 7, pp. 2964–2972, 2018.

- [21] B. Govoreanu, G. S. Kar, Y. Chen, V. Paraschiv, S. Kubicek, A. Fantini, I. Radu, L. Goux, S. Clima, R. Degraeve, *et al.*, “10× 10nm² hf/hfo x crossbar resistive ram with excellent performance, reliability and low-energy operation,” in *2011 International Electron Devices Meeting*, pp. 31–6, IEEE, 2011.
- [22] C. Nail, G. Molas, P. Blaise, G. Piccolboni, B. Sklenard, C. Cagli, M. Bernard, A. Roule, M. Azzaz, E. Vianello, *et al.*, “Understanding rram endurance, retention and window margin trade-off using experimental results and simulations,” in *2016 IEEE International Electron Devices Meeting (IEDM)*, pp. 4–5, IEEE, 2016.
- [23] L. Goux, “1 - oxram technology development and performances,” in *Advances in Non-Volatile Memory and Storage Technology (Second Edition)* (B. Magyari-Köpe and Y. Nishi, eds.), Woodhead Publishing Series in Electronic and Optical Materials, pp. 3–33, Woodhead Publishing, second edition ed., 2019.
- [24] D. Efnusheva, A. Cholakovska, and A. Tentov, “A survey of different approaches for overcoming the processor-memory bottleneck,” *International Journal of Computer Science and Information Technology*, vol. 9, no. 2, pp. 151–163, 2017.
- [25] M. Qazi, M. Sinangil, and A. Chandrakasan, “Challenges and directions for low-voltage sram,” *IEEE design & test of computers*, vol. 28, no. 1, pp. 32–43, 2010.
- [26] K.-H. Yang, H.-J. Tsai, C.-Y. Li, P. Jendra, M.-F. Chang, and T.-F. Chen, “etag : Tag-comparison in memory to achieve direct data access based on edram to improve energy efficiency of dram cache,” *IEEE Transactions on Circuits and Systems I : Regular Papers*, vol. 64, no. 4, pp. 858–868, 2016.
- [27] S. S. Kim, S. K. Yong, W. Kim, S. Kang, H. W. Park, K. J. Yoon, D. S. Sheen, S. Lee, and C. S. Hwang, “Review of semiconductor flash memory devices for material and process issues,” *Advanced Materials*, p. 2200659, 2022.
- [28] L. Perniola, G. Molas, G. Navarro, E. Nowak, V. Sousa, E. Vianello, and B. De Salvo, “Universal signatures from non-universal memories : clues for the future...,” in *2016 IEEE 8th International Memory Workshop (IMW)*, pp. 1–3, IEEE, 2016.
- [29] <https://www.lesnumeriques.com/cpu-processeur/intel-grave-en-3-nm-et-revient-dans-la-course-a-la-miniaturisation-n223423.html>, “Intel grave en 3 nm et veut en remonter à son concurrent tsmc,” 2024. accessed on 7.7.2024.
- [30] A. P. Jacob, R. Xie, M. G. Sung, L. Liebmann, R. T. Lee, and B. Taylor, “Scaling challenges for advanced cmos devices,” *International Journal of High Speed Electronics and Systems*, vol. 26, no. 01n02, p. 1740001, 2017.

-
- [31] R. K. Ratnesh, A. Goel, G. Kaushik, H. Garg, M. Singh, B. Prasad, *et al.*, “Advancement and challenges in mosfet scaling,” *Materials Science in Semiconductor Processing*, vol. 134, p. 106002, 2021.
- [32] J. S. Meena, S. M. Sze, U. Chand, and T.-Y. Tseng, “Overview of emerging nonvolatile memory technologies,” *Nanoscale research letters*, vol. 9, pp. 1–33, 2014.
- [33] N. Menou, *Technologie FeRAM : fiabilité et mécanismes de défaillance de condensateurs ferroélectriques élémentaires et intégrés*. PhD thesis, Université du Sud Toulon Var, 2004.
- [34] T. Mikolajick, U. Schroeder, P. Lomenzo, E. Breyer, H. Mulaosmanovic, M. Hoffmann, T. Mittmann, F. Mehmood, B. Max, and S. Slesazeck, “Next generation ferroelectric memories enabled by hafnium oxide,” in *2019 IEEE International Electron Devices Meeting (IEDM)*, pp. 15–5, IEEE, 2019.
- [35] A. Khvalkovskiy, D. Apalkov, S. Watts, R. Chepulskii, R. Beach, A. Ong, X. Tang, A. Driskill-Smith, W. Butler, P. Visscher, *et al.*, “Basic principles of stt-mram cell operation in memory arrays,” *Journal of Physics D : Applied Physics*, vol. 46, no. 7, p. 074001, 2013.
- [36] J. C. Slonczewski, “Conductance and exchange coupling of two ferromagnets separated by a tunneling barrier,” *Physical Review B*, vol. 39, no. 10, p. 6995, 1989.
- [37] J. C. Slonczewski, “Current-driven excitation of magnetic multilayers,” *Journal of Magnetism and Magnetic Materials*, vol. 159, no. 1-2, pp. L1–L7, 1996.
- [38] S. Raoux, G. W. Burr, M. J. Breitwisch, C. T. Rettner, Y.-C. Chen, R. M. Shelby, M. Salinga, D. Krebs, S.-H. Chen, H.-L. Lung, *et al.*, “Phase-change random access memory : A scalable technology,” *IBM Journal of Research and Development*, vol. 52, no. 4.5, pp. 465–479, 2008.
- [39] H. Cheng, J. Wu, R. Cheek, S. Raoux, M. BrightSky, D. Garbin, S. Kim, T. Hsu, Y. Zhu, E. Lai, *et al.*, “A thermally robust phase change memory by engineering the ge/n concentration in (ge, n) x sb y te z phase change material,” in *2012 International Electron Devices Meeting*, pp. 31–1, IEEE, 2012.
- [40] Q. Hubert, C. Jahan, A. Toffoli, G. Navarro, S. Chandrashekar, P. Noe, D. Blachier, V. Sousa, L. Perniola, J.-F. Nodin, *et al.*, “Lowering the reset current and power consumption of phase-change memories with carbon-doped ge₂sb₂te₅,” in *2012 4th IEEE International Memory Workshop*, pp. 1–4, IEEE, 2012.
- [41] T. Hickmott, “Low-frequency negative resistance in thin anodic oxide films,” *Journal of Applied Physics*, vol. 33, no. 9, pp. 2669–2682, 1962.

- [42] J. Gibbons and W. Beadle, "Switching properties of thin nio films," *Solid-State Electronics*, vol. 7, no. 11, pp. 785–790, 1964.
- [43] P. Nielsen and N. Bashara, "The reversible voltage-induced initial resistance in the negative resistance sandwich structure," *IEEE Transactions on Electron Devices*, vol. 11, no. 5, pp. 243–244, 1964.
- [44] W. Hiatt and T. Hickmott, "Bistable switching in niobium oxide diodes," *Applied Physics Letters*, vol. 6, no. 6, pp. 106–108, 1965.
- [45] K. Aratani, K. Ohba, T. Mizuguchi, *et al.*, "Electron devices meeting, 2007. iedm 2007," in *IEEE International*, vol. 783, 2007.
- [46] H. Lee, P. Chen, T. Wu, Y. Chen, C. Wang, P. Tzeng, C. Lin, F. Chen, C. Lien, and M. Tsai, "December. low power and high speed bipolar switching with a thin reactive ti buffer layer in robust hfo2 based rram," in *2008 IEEE International Electron Devices Meeting*, pp. 1–4, IEEE, 2008.
- [47] H. S. Yoon, I.-G. Baek, J. Zhao, H. Sim, M. Y. Park, H. Lee, G.-H. Oh, J. C. Shin, I.-S. Yeo, and U.-I. Chung, "Vertical cross-point resistance change memory for ultra-high density non-volatile memory applications," in *2009 Symposium on VLSI Technology*, pp. 26–27, IEEE, 2009.
- [48] M. Jeong, P. M. Solomon, S. Laux, H.-S. Wong, and D. Chidambarrao, "Comparison of raised and schottky source/drain mosfets using a novel tunneling contact model," in *International Electron Devices Meeting 1998. Technical Digest (Cat. No. 98CH36217)*, pp. 733–736, IEEE, 1998.
- [49] X. Yuan, S. Kim, J. Juyon, M. D'Urbino, T. Bullmann, Y. Chen, A. Stettler, A. Hierlemann, and U. Frey, "A microelectrode array with 8,640 electrodes enabling simultaneous full-frame readout at 6.5 kfps and 112-channel switch-matrix readout at 20 ks/s," in *2016 IEEE Symposium on VLSI Circuits (VLSI-Circuits)*, pp. 1–2, IEEE, 2016.
- [50] S. Yu, "Resistive random access memory (rram)," *Synthesis lectures on emerging engineering technologies*, vol. 2, no. 5, pp. 1–79, 2016.
- [51] S. Long, L. Perniola, C. Cagli, J. Buckley, X. Lian, E. Miranda, F. Pan, M. Liu, and J. Suñé, "Voltage and power-controlled regimes in the progressive unipolar reset transition of hfo2-based rram," *Scientific reports*, vol. 3, no. 1, p. 2929, 2013.
- [52] J. Guy, G. Molas, E. Vianello, F. Longnos, S. Blanc, C. Carabasse, M. Bernard, J. Nodin, A. Toffoli, J. Cluzel, *et al.*, "Investigation of the physical mechanisms governing data-retention in down to 10nm nano-trench al 2 o 3/cutege conductive bridge ram (cbram)," in *2013 IEEE International Electron Devices Meeting*, pp. 30–2, IEEE, 2013.

-
- [53] E. Vianello, G. Molas, F. Longnos, P. Blaise, E. Souchier, C. Cagli, G. Palma, J. Guy, M. Bernard, M. Reyboz, *et al.*, “Sb-doped ges 2 as performance and reliability booster in conductive bridge ram,” in *2012 International Electron Devices Meeting*, pp. 31–5, IEEE, 2012.
- [54] R. Waser, R. Dittmann, G. Staikov, and K. Szot, “Redox-based resistive switching memories—nanoionic mechanisms, prospects, and challenges,” *Advanced materials*, vol. 21, no. 25-26, pp. 2632–2663, 2009.
- [55] I. Valov, R. Waser, J. R. Jameson, and M. N. Kozicki, “Electrochemical metalization memories—fundamentals, applications, prospects,” *Nanotechnology*, vol. 22, no. 25, p. 254003, 2011.
- [56] G. Bersuker, D. Gilmer, D. Veksler, J. Yum, H. Park, S. Lian, L. Vandelli, A. Padovani, L. Larcher, K. McKenna, *et al.*, “Metal oxide rram switching mechanism based on conductive filament microscopic properties,” in *2010 International Electron Devices Meeting*, pp. 19–6, IEEE, 2010.
- [57] Z. Fang, X. P. Wang, J. Sohn, B. B. Weng, Z. P. Zhang, Z. X. Chen, Y. Z. Tang, G.-Q. Lo, J. Provine, S. S. Wong, *et al.*, “The role of ti capping layer in hfo x-based rram devices,” *IEEE Electron Device Letters*, vol. 35, no. 9, pp. 912–914, 2014.
- [58] Y. Y. Chen, L. Goux, S. Clima, B. Govoreanu, R. Degraeve, G. S. Kar, A. Fantini, G. Groeseneken, D. J. Wouters, and M. Jurczak, “Endurance/retention trade-off on cap 1t1r bipolar rram,” *IEEE Transactions on electron devices*, vol. 60, no. 3, pp. 1114–1121, 2013.
- [59] L. Courtade, C. Turquat, C. Muller, J. Lisoni, L. Goux, D. Wouters, D. Gougenheim, P. Roussel, and L. Ortega, “Oxidation kinetics of ni metallic films : formation of nio-based resistive switching structures,” *Thin solid films*, vol. 516, no. 12, pp. 4083–4092, 2008.
- [60] R. Waser, R. Dittmann, G. Staikov, and K. Szot, “Redox-based resistive switching memories—nanoionic mechanisms, prospects, and challenges,” *Advanced materials*, vol. 21, no. 25-26, pp. 2632–2663, 2009.
- [61] S. Yu, B. Lee, and H.-S. P. Wong, “Metal oxide resistive switching memory,” *Functional Metal Oxide Nanostructures*, pp. 303–335, 2011.
- [62] S. Tirano, L. Perniola, J. Buckley, J. Cluzel, V. Jousseau, C. Muller, D. Delleruyelle, B. De Salvo, and G. Reimbold, “Accurate analysis of parasitic current overshoot during forming operation in rrams,” *Microelectronic engineering*, vol. 88, no. 7, pp. 1129–1132, 2011.
- [63] J. Cayo, I. Vourkas, and A. Rubio, “A circuit-level spice modeling strategy for the simulation of behavioral variability in reram,” in *2022 IFIP/IEEE 30th International Conference on Very Large Scale Integration (VLSI-SoC)*, pp. 1–4, IEEE, 2022.

- [64] X. Guan, S. Yu, and H.-S. P. Wong, "On the switching parameter variation of metal-oxide rram—part i : Physical modeling and simulation methodology," *IEEE Transactions on electron devices*, vol. 59, no. 4, pp. 1172–1182, 2012.
- [65] H. Lee, Y. Chen, P. Chen, T. Wu, F. Chen, C. Wang, P. Tzeng, M.-J. Tsai, and C. Lien, "Low-power and nanosecond switching in robust hafnium oxide resistive memory with a thin ti cap," *IEEE Electron Device Letters*, vol. 31, no. 1, pp. 44–46, 2009.
- [66] S. Ikegawa, F. B. Mancoff, J. Janesky, and S. Aggarwal, "Magnetoresistive random access memory : Present and future," *IEEE Transactions on Electron Devices*, vol. 67, no. 4, pp. 1407–1419, 2020.
- [67] O. Golonzka, J.-G. Alzate, U. Arslan, M. Bohr, P. Bai, J. Brockman, B. Buford, C. Connor, N. Das, B. Doyle, *et al.*, "Mram as embedded non-volatile memory solution for 22ffl finfet technology," in *2018 IEEE International Electron Devices Meeting (IEDM)*, pp. 18–1, IEEE, 2018.
- [68] K. Lee, K. Yamane, S. Noh, V. Naik, H. Yang, S. Jang, J. Kwon, B. Behin-Aein, R. Chao, J. Lim, *et al.*, "22-nm fd-soi embedded mram with full solder reflow compatibility and enhanced magnetic immunity," in *2018 IEEE Symposium on VLSI Technology*, pp. 183–184, IEEE, 2018.
- [69] Y. Song, J. Lee, S. Han, H. Shin, K. Lee, K. Suh, D. Jeong, G. Koh, S. Oh, J. Park, *et al.*, "Demonstration of highly manufacturable stt-mram embedded in 28nm logic," in *2018 IEEE International Electron Devices Meeting (IEDM)*, pp. 18–2, IEEE, 2018.
- [70] W. J. Gallagher, E. Chien, T.-W. Chiang, J.-C. Huang, M.-C. Shih, C. Wang, C. Bair, G. Lee, Y.-C. Shih, C.-F. Lee, *et al.*, "Recent progress and next directions for embedded mram technology," in *2019 Symposium on VLSI Circuits*, pp. T190–T191, IEEE, 2019.
- [71] L. Wu, M. Taouil, S. Rao, E. J. Marinissen, and S. Hamdioui, "Survey on stt-mram testing : Failure mechanisms, fault models, and tests," *arXiv preprint arXiv :2001.05463*, 2020.
- [72] T. Endoh, H. Koike, S. Ikeda, T. Hanyu, and H. Ohno, "An overview of nonvolatile emerging memories—spintronics for working memories," *IEEE journal on emerging and selected topics in circuits and systems*, vol. 6, no. 2, pp. 109–119, 2016.
- [73] Z. Song, D. Cai, X. Li, L. Wang, Y. Chen, H. Chen, Q. Wang, Y. Zhan, and M. Ji, "High endurance phase change memory chip implemented based on carbon-doped ge 2 sb 2 te 5 in 40 nm node for embedded application," in *2018 IEEE International Electron Devices Meeting (IEDM)*, pp. 27–5, IEEE, 2018.

- [74] F. Arnaud, P. Zuliani, J. Reynard, A. Gandolfo, F. Disegni, P. Mattavelli, E. Gomiero, G. Samanni, C. Jahan, R. Berthelon, *et al.*, “Truly innovative 28nm fdsoi technology for automotive micro-controller applications embedding 16mb phase change memory,” in *2018 IEEE International Electron Devices Meeting (IEDM)*, pp. 18–4, IEEE, 2018.
- [75] J. Izraelevitz, J. Yang, L. Zhang, J. Kim, X. Liu, A. Memaripour, Y. J. Soh, Z. Wang, Y. Xu, S. R. Dullloor, *et al.*, “Basic performance measurements of the intel optane dc persistent memory module,” *arXiv preprint arXiv :1903.05714*, 2019.
- [76] <https://www.fujitsu.com/jp/group/fsm/en/products/reram/>, “Non-volatile memory with very small operating current,” 2023. accessed on 31.03.2023.
- [77] https://www.infineon.com/cms/en/about-infineon/press/market-news/2022/INFATV202211_031.html, “Infineon and tsmc to introduce rram technology for automotive aurix™ tc4x product family,” 2022. accessed on 25.11.2022.
- [78] <https://www.fujitsu.com/jp/group/fsm/en/products/feram/>, “Non-volatile memory with high performance and high reliability,” 2023. accessed on 28.08.2023.
- [79] <https://www.infineon.com/cms/en/product/memories/f-ram-ferroelectric-ram/>, “Highly reliable and low-power data logging memory with virtually unlimited endurance,” 2023. accessed on 31.03.2023.
- [80] B. Li, B. Yan, and H. Li, “An overview of in-memory processing with emerging non-volatile memory for data-intensive applications,” *IEEE*, pp. 381–386, 2019.
- [81] G. Molas and E. Nowak, “Advances in emerging memory technologies : From data storage to artificial intelligence,” *Applied Sciences*, vol. 11, no. 23, p. 11254, 2021.
- [82] <https://www.mouser.com/pdfDocs/AN000002-Avalanche-STT-MRAM-Device-Characteristics-and-Capabilities.pdf>, “Endurance, data retention and field immunity of stt-mram,” 2023. Accessed on 01.09.2023.
- [83] G. Bersuker, J. Yum, L. Vandelli, A. Padovani, L. Larcher, V. Iglesias, M. Porti, M. Nafria, K. McKenna, A. Shluger, *et al.*, “Grain boundary-driven leakage path formation in hfo2 dielectrics,” *Solid-State Electronics*, vol. 65, pp. 146–150, 2011.
- [84] R. Degraeve, A. Fantini, S. Clima, B. Govoreanu, L. Goux, Y. Y. Chen, D. Wouters, P. Roussel, G. S. Kar, G. Pourtois, *et al.*, “Dynamic ‘hour glass’ model for set and reset in hfo 2 rram,” in *2012 Symposium on VLSI Technology (VLSIT)*, pp. 75–76, IEEE, 2012.

- [85] M. Herrmann and A. Schenk, "Field and high-temperature dependence of the long term charge loss in erasable programmable read only memories : Measurements and modeling," *Journal of applied physics*, vol. 77, no. 9, pp. 4522–4540, 1995.
- [86] F. M. Puglisi, L. Larcher, G. Bersuker, A. Padovani, and P. Pavan, "An empirical model for rram resistance in low-and high-resistance states," *IEEE Electron Device Letters*, vol. 34, no. 3, pp. 387–389, 2013.
- [87] R. Tsu and L. Esaki, "Tunneling in a finite superlattice," *Applied Physics Letters*, vol. 22, no. 11, pp. 562–564, 1973.
- [88] R. Landauer, "Spatial variation of currents and fields due to localized scatterers in metallic conduction," *IBM Journal of research and development*, vol. 1, no. 3, pp. 223–231, 1957.
- [89] M. Büttiker, "Four-terminal phase-coherent conductance," *Physical review letters*, vol. 57, no. 14, p. 1761, 1986.
- [90] X. Lian, M. Wang, M. Rao, P. Yan, J. J. Yang, and F. Miao, "Characteristics and transport mechanisms of triple switching regimes of taox memristor," *Applied Physics Letters*, vol. 110, no. 17, p. 173504, 2017.
- [91] E. A. Miranda, C. Walczyk, C. Wenger, and T. Schroeder, "Model for the resistive switching effect in mim structures based on the transmission properties of narrow constrictions," *IEEE Electron Device Letters*, vol. 31, no. 6, pp. 609–611, 2010.
- [92] J. Frenkel, "On pre-breakdown phenomena in insulators and electronic semiconductors," *Physical Review*, vol. 54, no. 8, p. 647, 1938.
- [93] K.-C. Chang, T.-M. Tsai, T.-C. Chang, Y.-E. Syu, S.-L. Chuang, C.-H. Li, D.-S. Gan, and S. M. Sze, "The effect of silicon oxide based rram with tin doping," *Electrochemical and Solid-State Letters*, vol. 15, no. 3, p. H65, 2011.
- [94] G. Bersuker, D. Gilmer, D. Veksler, P. Kirsch, L. Vandelli, A. Padovani, L. Larcher, K. McKenna, A. Shluger, V. Iglesias, *et al.*, "Metal oxide resistive memory switching mechanism based on conductive filament properties," *Journal of Applied Physics*, vol. 110, no. 12, p. 124518, 2011.
- [95] S. Mazumder, *Numerical methods for partial differential equations : finite difference and finite volume methods*. Academic Press, 2015.
- [96] A. Galántai, "The theory of newton's method," *Journal of Computational and Applied Mathematics*, vol. 124, no. 1, pp. 25–44, 2000. Numerical Analysis 2000. Vol. IV : Optimization and Nonlinear Equations.
- [97] S. Kirkpatrick, "Percolation and conduction," *Reviews of modern physics*, vol. 45, no. 4, p. 574, 1973.

-
- [98] A. Chen, “A review of emerging non-volatile memory (nvm) technologies and applications,” *Solid-State Electronics*, vol. 125, pp. 25–38, 2016.
- [99] T. Feuchtwang, W. Leipold, and R. C. Martino, “Theory of resonant tunneling through the insulator conduction band of a thin film metal-insulator-metal (mim) heterojunction,” *Surface Science*, vol. 64, no. 1, pp. 109–130, 1977.
- [100] F. Jansson, “Charge transport in disordered materials : simulations, theory, and numerical modeling of hopping transport and electron-hole recombination,” 2011.
- [101] S. Li, X. Jin, Y. Xuan, X. Zhou, W. Chen, Y.-X. Wang, and X. Yan, “Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting,” *Advances in neural information processing systems*, vol. 32, 2019.
- [102] A. Miller and E. Abrahams, “Impurity conduction at low concentrations,” *Physical Review*, vol. 120, no. 3, p. 745, 1960.
- [103] Y. Lee, S. Jung, A. Plews, A. Nejm, O. Simonetti, L. Giraudet, S. D. Baranovskii, F. Gebhard, K. Meerholz, S. Jung, *et al.*, “Parametrization of the gaussian disorder model to account for the high carrier mobility in disordered organic transistors,” *Physical Review Applied*, vol. 15, no. 2, p. 024021, 2021.
- [104] C. Cagli, J. Buckley, V. Jousseau, T. Cabout, A. Salaun, H. Grampeix, J. Nodin, H. Feldis, A. Persico, J. Cluzel, *et al.*, “Experimental and theoretical study of electrode effects in hfo 2 based rram,” in *2011 International Electron Devices Meeting*, pp. 28–7, IEEE, 2011.
- [105] T. Cabout, *Optimisation technologique et caractérisation électrique de mémoires résistives OxRRAM pour applications basse consommation*. PhD thesis, Aix-Marseille, 2014.
- [106] L. Reganaz, E. Esmanhotto, N. Ait Abdelkader, J. Minguet Lopez, N. Castellani, Q. Rafhay, D. Deleruyelle, L. Grenouillet, F. Aussenac, E. Vianello, *et al.*, “Elucidating postprogramming relaxation in multilevel cell-resistive random access memory by means of experimental and kinetic monte carlo simulation data,” *physica status solidi (a)*, vol. 219, no. 13, p. 2100753, 2022.