



HAL
open science

Structures de protection innovantes contre les décharges électrostatiques dédiées aux entrées/sorties hautes tensions de technologies SmartPower

Amaury Gendron

► **To cite this version:**

Amaury Gendron. Structures de protection innovantes contre les décharges électrostatiques dédiées aux entrées/sorties hautes tensions de technologies SmartPower. Micro et nanotechnologies/Microélectronique. Université Paul Sabatier - Toulouse III, 2007. Français. NNT: . tel-00362699

HAL Id: tel-00362699

<https://theses.hal.science/tel-00362699>

Submitted on 19 Feb 2009

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.

THESE

en vue de l'obtention du

DOCTORAT DE L'UNIVERSITE DE TOULOUSE

délivré par l'Université Paul Sabatier

Ecole Doctorale : GEET

Discipline : Microélectronique

présentée et soutenue

par

Amaury Gendron

le 29 mars 2007

**Structures de protections innovantes contre les décharges
électrostatiques, dédiées aux Entrées/Sorties hautes
tensions de technologies SmartPower**

Directrice de thèse :

Mme. Marise Baffleur, Directrice de recherche - LAAS-CNRS

Codirecteur de thèse :

M. Nicolas Nolhier, Professeur - Université Paul Sabatier

JURY

M. Christian Schaeffer, Professeur Institut National Polytechnique de Grenoble,
Président

M. Herman Maes, Professeur Université Catholique de Louvain, Rapporteur

M. Dionyz Pogany, Assistant professeur Université de Technologie de Vienne

M. Philippe Renaud, Ingénieur de recherche Freescale Toulouse

M. David Tremouilles, Chercheur IMEC Louvain

M. Michel Zecri, Ingénieur de recherche Freescale Crolles

Remerciements

Le travail de thèse présenté dans ce mémoire a été effectué dans le cadre d'une "Convention Industrielle de Formation par la REcherche" (CIFRE) établie avec la société Freescale semiconducteur SAS et du "Laboratoire d'Architecture et d'Analyse des Systèmes" (LAAS) du "Centre National de la Recherche Scientifique" (CNRS). Au LAAS-CNRS, j'ai intégré le groupe d'"Intégration de Systèmes et Gestion de l'Energie" (ISGE) et à Freescale la division "Technology Solution Organization" (TSO), entre lesquels un partenariat de recherche sur trois ans avait été conclu pour former le "Laboratoire d'Intégration des Systèmes de Puissance Avancée" (LISPA).

A ce titre, je remercie Monsieur Jean-Louis CHAPTAL, directeur de recherche et développement à la société Freescale semiconducteur, ainsi que Messieurs Malik GHALLAB et Raja CHATILA, directeurs successifs du LAAS-CNRS, pour m'avoir accueilli accueillis au sein de leurs équipes. Je tiens à remercier également Monsieur Jean-Louis Sanchez et Madame Marise BAFLEUR, responsables successifs du groupe ISGE.

C'est avec une profonde sincérité que je remercie une nouvelle fois Madame Marise BAFLEUR, directrice de thèse, qui par ses remarquables compétences et ses grandes qualités humaines a grandement contribué à l'aboutissement des ces travaux dans les meilleures conditions. Je remercie Nicolas NOLHIER, co-directeur de thèse, dont l'implication dans ce sujet a été particulièrement bénéfique. J'adresse mes remerciements à Messieurs Michel ZECRI et Phillipe RENAUD, ingénieurs à Freescale, qui ont assuré successivement l'encadrement côté industrielle, pour leur apport scientifique à ces travaux, la confiance qu'ils m'ont témoignée, leurs soutiens permanents.

Je remercie vivement les personnes extérieures ayant acceptées de participer au jury de thèse :

- Monsieur Herman MAES, professeur à l'université catholique de Louvain en Belgique et senior vice-président de l'IMEC, et Monsieur Christian SCHAEFFER, professeur à l'institut national polytechnique de Grenoble et directeur du CIME Nanotech, qui, en qualité de rapporteurs, se sont fortement investis pour émettre un jugement sur la qualité du mémoire.

- Messieurs Dioniz POGANY, assistant professeur à l'université de Vienne en Autriche et David TREMOUILLES, chercheur à l'IMEC, examinateurs ayant porté une grande attention à mes travaux.

Je tiens à assurer ma gratitude à Monsieur Alain DERAM, mon manager direct à Freescale, extrêmement efficace et dont l'implication pour la mise en place des conditions de mon embauche a été très appréciable. J'en profite pour remercier une seconde fois Phillipe RENAUD et Jean-Louis CHAPTAL qui m'ont aidé à mettre en avant mon travail dans les équipes de Freescale et ainsi contribués de manière décisive à mon embauche.

Ce travail a été motivé par la demande du groupe de design automobile et produits standards (TSPG) de Freescale à Toulouse. Je remercie Monsieur Erwan HEMON, responsable de ce groupe, ainsi que les designers Patrice BESSE et Phillipe GIVELIN, pour leur collaboration et leurs conseils avisés.

Nicolas MAURAN, technicien au LAAS, s'est investi dans les mesures réalisées au LAAS, de même que David LAFFARGUE pour celles réalisées à Freescale. A tous deux, un grand merci.

Je tiens à exprimer ma reconnaissance aux thésards du LAAS qui m'ont précédé et que j'ai trouvé très disponibles pour me faire part de leur expérience. Je pense à nouveau à Patrice BESSE (mon "grand frère") et David TREMOUILLES ainsi qu'à Christophe SALAMERO et Stéphane ALVES.

Je remercie Isabelle NOLHIER, secrétaire du groupe ISGE, et les secrétaires successives des laboratoires communs entre le LAAS et Freescale, Catherine LEFEUVRE, Caroline MALE et, la dernière mais pas la moindre, Catherine COURET, toutes très efficaces.

Je pense également aux collègues de Freescale Bertrand V. (mon sauveur pour word et powerpoint), Christophe L. (le contremaître), Jean-Michel R., Evgueniy S., René E., Cécile D. (la terreur de MSN), Isabelle B. (reine de l'anacoluthie), Adan S., Céline B, Léna S., Jean-Baptiste S., Yann W., Béatrice B., Jean-Phillipe L., Laurence M., Aurélie L., Marlène N., Delphine (une compatriote de Bretagne),

Pierre L. (discussion cinématographiques passionnantes), Sébastien C. (spécialiste de Kierkegaard?), Raphaël M. (un filon de blagues inépuisable), Ahmed C. (le petit nouveau), encore une fois Philippe R. (pour Napoléon, Victor Hugo, Herman Hess et les blagues du vendredi après-midi) et les anciens forçats du bungalow 5 qui se reconnaîtront.

Je n'oublie pas les collègues du LAAS Christian C., Wasim H., Taoufik E., Yuan G., Nicolas L., Nicolas G., Frank J., Fabrice C., Rodolphe D., Frédéric M, Eric I., Abdelhakim B., Sandrine A.

De manière générale, je salue toutes les personnes que j'ai eu l'occasion de fréquenter pendant ces trois années de thèse et avec qui mes rapports furent aussi divers qu'enrichissants.

Je prie Sisteron ddl2 et Brax d'accepter mes excuses pour les avoir soumis sans pitié à des simulations, les forçant à travailler même la nuit et le week-end.

Enfin, une pensée émue va à mes proches et ma famille pour le soutien et la confiance qu'ils m'ont toujours accordés.

Sommaire

<i>par</i>	Error! Bookmark not defined.
Soutenue le 29 Mars 2007, devant le jury :	Error! Bookmark not defined.
Remerciements	3
Sommaire	5
Introduction générale	9
Chapitre1 Problématique et état de l'art de la protection des technologies SmartPower contre les décharges électrostatiques (ESD)	13
1.1. Introduction.....	13
1.2. Spécifications des protections contre les décharges électrostatiques	13
1.2.1. Robustesse	14
1.2.1.1. Test HBM	14
1.2.1.2. Test MM	16
1.2.1.3. Test "pistolet"	18
1.2.1.4. Test CDM	19
1.2.1.5. Méthodologie retenue pour la caractérisation de la robustesse	21
1.2.2. Définition de la fenêtre de conception.....	22
1.2.2.1. Intervalle autorisé pour la tension à l'état passant.....	22
1.2.2.2. Résistance à l'état passant (R_{ON}) maximale	23
1.2.3. Spécifications des protections dédiées aux E/S "hautes tensions" en technologies SmartPower.....	23
1.2.3.1. Fenêtre de conception et propriétés de la structure de protection	23
1.2.3.2. Difficultés scientifiques et techniques.....	24
1.2.3.3. Protection "haute tension" pour une application ethernet	26
1.3. Spécificités de la physique des protections ESD.....	27
1.3.1. Etude du comportement physique d'une protection.....	28
1.3.1.1. Densités de courant extrêmement élevées.....	28
1.3.1.2. Forte élévation de température.....	29
1.3.2. Méthodes de caractérisation	30
1.3.2.1. Caractérisation TLP	30
1.3.2.2. Caractérisation vfTLP	31
1.3.2.3. Mesures TIM	31
1.4. Etat de l'art des protections ESD "hautes tensions" en technologies SmartPower	32
1.4.1. Principes de base des circuits de protection	32
1.4.1.1. Protections centralisées	33
1.4.1.2. Protections localisées	33
1.4.2. Structures de protection intégrées sur silicium	34
1.4.2.1. Transistors bipolaires autopolarisés	35
1.4.2.2. Structures PNPN	39
1.4.2.3. Transistors à effet de champ	42
1.4.3. Protection des E/S 80 Volts dans la technologie SmartMOS 8 MV de Freescale	45
1.4.3.1. Solution conventionnelle.....	45
1.4.3.2. Type de protection retenu pour une solution alternative	46
1.5. Conclusion.....	47
Chapitre2 Etude théorique du comportement des transistors bipolaires autopolarisés pendant une ESD	49
2.1. Introduction.....	49
2.2. Description du mode autopolarisé d'un transistor bipolaire.....	50
2.2.1. Développement analytique	50
2.2.2. Caractéristique électrique	51
2.3. Comportement à faible courant	53
2.3.1. Suppression du repliement en polarisation base flottante	53

2.3.2.	Tensions de maintien élevées	54
2.3.2.1.	Gain.....	54
2.3.2.2.	Facteur de multiplication par avalanche.....	56
2.4.	Résistance à l'état passant /Fortes densités de courant	57
2.4.1.	Notion de forte densité de courant.....	57
2.4.2.	Chute du gain en courant.....	58
2.4.3.	Modification de la ZCE base-collecteur par l'injection d'émetteur.....	59
2.4.4.	Effet du courant de polarisation sur la modification de la ZCE base-collecteur.....	61
2.5.	Etude analytique du fonctionnement aux densités de courant générées par les ESD	62
2.5.1.	Description du rapport d'injection aux densités de courant générées par les ESD	62
2.5.1.1.	Formalisation du problème	62
2.5.1.2.	Calcul des charges stockées	64
2.5.1.3.	Formule analytique du rapport d'injection.....	66
2.5.1.4.	Etude asymptotique aux fortes densités de courant.....	68
2.5.1.5.	Implications sur le R_{ON} et la tension de maintien.....	69
2.5.2.	Modulation de la ZCE base-collecteur	70
2.5.2.1.	Calcul de la concentration de charge en fonction de la densité de courant	70
2.5.2.2.	Etude des phénomènes induits par la modulation de la ZCE base-collecteur	71
2.5.3.	Confrontation avec la simulation numérique	74
2.5.3.1.	Cas du transistor NPN.....	75
2.5.3.2.	Cas du Transistor PNP.....	76
2.6.	Règles pour la conception de protections "hautes tensions" à faible R_{ON}	78
2.6.1.	Transistors NPN	78
2.6.1.1.	Stratégie d'optimisation	78
2.6.1.2.	Règles de dessin.....	79
2.6.2.	Transistors PNP.....	79
2.6.2.1.	Stratégie d'optimisation	79
2.6.2.2.	Règles de dessin.....	80
2.7.	Résistance à l'état passant / Effets thermiques.....	80
2.7.1.	Résistances d'accès	80
2.7.2.	Multiplication par avalanche	82
2.7.3.	Modulation de la ZCE base-collecteur	83
2.7.4.	Recombinaison-génération SRH dans la ZCE base-collecteur	83
2.7.5.	Synthèse des modifications du champ électrique de la ZCE base-collecteur.....	84
2.8.	Conclusion.....	85
Chapitre3 Développement de protections ESD à base de transistors bipolaires PNP		87
3.1.	Introduction.....	87
3.2.	Simulation pour la conception de protections ESD.....	87
3.2.1.	Simulation du procédé technologique.....	88
3.2.1.1.	Profils SIMS	89
3.2.1.2.	Mesures SCM	90
3.2.2.	Simulation électrothermique.....	90
3.3.	Réduction du R_{ON} des transistors PNP latéraux autopolarisés	92
3.3.1.	Règles de dessin	93
3.3.1.1.	Configuration d'émetteur	93
3.3.1.2.	Configuration de collecteur.....	97
3.3.1.3.	Synthèse des règles de dessin.....	100
3.3.2.	Caractérisations	100
3.3.2.1.	Description des structures testées.....	100
3.3.2.2.	Variation du profil de dopage base-collecteur	102
3.3.2.3.	Variation de la longueur de collecteur	105
3.4.	Couplage d'un transistor PNP latéral avec une diode verticale.....	107
3.4.1.	Principe de fonctionnement	107
3.4.1.1.	Description théorique et choix d'une diode à avalanche verticale comme nouvelle source de courant	107
3.4.1.2.	Etude en simulation.....	108
3.4.1.3.	Variante favorable au couplage.....	110
3.4.2.	Etude des diodes en inverse pendant une ESD.....	110
3.4.2.1.	Simulation comparative d'une diode à avalanche et d'un transistor NPN autopolarisé.....	111

3.4.2.2.	Description des mécanismes physiques.....	112
3.4.2.3.	Résultats de mesure.....	115
3.4.3.	Règles d'optimisation.....	116
3.4.3.1.	Tensions de déclenchement de la diode et du transistor PNP.....	116
3.4.3.2.	Dimension de la diode.....	117
3.4.3.3.	Contribution des différentes composantes du courant.....	118
3.4.4.	Caractérisations.....	120
3.4.4.1.	Validation des règles de dessin pour l'optimisation du couplage.....	120
3.4.4.2.	Structure de protection pour les E/S 80 Volts.....	123
3.5.	Conclusion.....	125
Chapitre4 Protections ESD "hautes tensions" et à faible R_{ON} à base de transistors bipolaires NPN.....		127
4.1.	Introduction.....	127
4.2.	Mise à profit des spécificités technologiques pour la réduction du facteur de mérite du R_{ON}	127
4.2.1.	Réalisation technologique.....	128
4.2.1.1.	Rappel des règles de dessin définies au Chapitre1.....	128
4.2.1.2.	Présentation de la structure.....	128
4.2.2.	Etude du fonctionnement.....	129
4.2.2.1.	Simulation.....	129
4.2.2.2.	Caractérisations.....	130
4.2.2.3.	Tentative d'explication du R_{ON} nul à fort courant.....	132
4.2.2.4.	Analyse du fort repliement.....	134
4.2.3.	Réduction de la tension de déclenchement.....	135
4.2.3.1.	Stratégie de déclenchement.....	135
4.2.3.2.	Caractérisation.....	136
4.3.	Structures à région flottante.....	139
4.3.1.	Claquage statique.....	139
4.3.1.1.	Tension de claquage en fonction de la position de la région flottante.....	140
4.3.1.2.	Modèle théorique unidimensionnel.....	141
4.3.1.3.	Modèle précis incluant les effets 2D.....	142
4.3.2.	Fonctionnement ESD pour une région flottante dans la base.....	146
4.3.2.1.	Présentation des résultats de simulation.....	146
4.3.2.2.	Mécanismes physiques.....	147
4.3.3.	Fonctionnement ESD pour une région flottante dans le collecteur.....	152
4.3.3.1.	Simulation du comportement ESD.....	152
4.3.3.2.	Mécanismes physiques et stratégies d'optimisation.....	153
4.3.3.3.	Résultats de mesures.....	157
4.4.	Conclusion.....	160
Conclusion générale.....		163
	Bibliographie.....	167
	Liste des symboles.....	173
	Contributions scientifiques.....	176

Introduction générale

Une décharge électrostatique, ou "electrostatic discharge" (ESD) en anglais, est un phénomène courant dans la vie de tous les jours. Les "châtaignes" qu'il vous arrive de ressentir, notamment quand vous introduisez la clé dans la serrure de votre voiture, ou les étincelles accompagnées de crépitements quand vous enlevez un vêtement en sont des exemples. Dans certains cas, ses manifestations peuvent être spectaculaires, telle que la foudre lors d'un orage. En Physique, une ESD correspond à un transfert de charges électriques lors d'un rééquilibrage de potentiel entre deux corps. Ce transfert suppose qu'initialement au moins l'un des deux corps a accumulé des charges électriques, ce qui a pu résulter de différents phénomènes. L'échange d'électrons lors du frottement entre deux matériaux, ou triboélectrification, en est l'un des plus répandus. Par exemple, c'est ce phénomène qui est en jeu dans l'expérience de physique amusante consistant à frotter une règle en plastique avec un pull en laine afin qu'elle puisse ensuite attirer des objets légers tels que des bouts de papier. Une personne peut également se charger par triboélectrification en marchant sur une moquette, en s'asseyant sur certains sièges automobiles... Un corps pourra aussi être chargé par induction, s'il est placé dans un champ électrique, ou par conduction, si les charges sont apportées par l'intermédiaire d'un autre corps préalablement chargé.

Dans le domaine de la microélectronique, un circuit intégré est susceptible de subir une ESD, soit lors de sa manipulation au cours des étapes industrielles (fabrication, test et assemblage), soit une fois monté dans l'application finale. Etant donné ses très petites dimensions, il peut être dégradé, y compris par des ESD qu'une personne ne sentirait même pas. A titre de comparaison, l'énergie dissipée par unité de masse est équivalente à celle pour un arbre sur lequel s'abat la foudre. Afin de limiter les risques, les industriels prennent des précautions, telles que l'utilisation de bracelets antistatiques, la mise à la masse des surfaces conductrices pouvant recevoir les circuits intégrés ou encore l'ionisation de l'air pour neutraliser les charges. Cependant, certains risques d'exposition aux ESD, essentiellement dans l'application finale, ne peuvent pas être supprimés. Prenons l'exemple de l'électronique embarquée dans une voiture. Une ESD peut l'atteindre si, au cours d'une réparation, la batterie a été débranchée ou si un moteur électrique (lève-vitre, essuie-glace) a été retiré pour être changé. Bien évidemment, les mêmes précautions que dans l'industrie microélectronique ne pourront pas être appliquées dans un garage automobile. Aussi, pour garantir la fiabilité d'un circuit intégré, il s'avère nécessaire de le protéger directement, afin qu'il puisse supporter une ESD sans être détruit. Pour cela, une méthode consiste à ajouter des composants, intégrés ou non, spécifiquement conçus pour absorber l'énergie de la décharge lorsque celle-ci survient.

L'évolution des procédés technologiques rend les circuits intégrés de plus en plus vulnérables face aux ESD. En particulier, la réduction des dimensions lithographiques conduit à des oxydes de grille très minces et à des jonctions très peu profondes, particulièrement fragiles. Dans le même temps, les clients des industries de microélectronique exigent des produits de plus en plus fiables, correspondant à des spécifications extrêmement contraignantes. Ces tendances sont particulièrement critiques dans le cas des technologies SmartPower. Une technologie SmartPower permet de réaliser à la fois des éléments de logique et des composants de puissance, capables de commuter des courants de plusieurs Ampères et de tenir des tensions élevées, typiquement jusqu'à 100 Volts. Ainsi, il est possible d'intégrer sur la même puce de silicium des fonctionnalités de plus en plus complexes tout en réduisant les coûts. Cependant, du point de vue des ESD, les tenues en tension élevées posent de sérieuses difficultés, car les énergies dissipées sont d'autant plus importantes. De plus, le risque pour

que le circuit intégré subisse une ESD est très élevé dans certains des domaines d'application, tels que les applications domestiques (imprimantes, domotiques) ou les transports (automobile). Concernant l'automobile, les fonctions sécuritaires (Airbag, correcteur de trajectoire, assistance au freinage) comprennent également des modules réalisés en technologie SmartPower. Dans ce cas, les spécifications de fiabilité sont draconiennes. Aujourd'hui, les ESD sont la première cause de non-conformité d'un circuit intégré nécessitant des modifications de conception, et elles représentent une proportion significative des retours clients de produits jugés défectueux.

L'objectif de la thèse est de développer des protections intégrées sur silicium pour les Entrée/Sortie (E/S) "hautes tensions" (80-40 Volts) des technologies SmartPower. Ces protections devront remplir les spécifications ESD requises, tout en permettant de réduire la surface de la puce, et donc son coût. Afin de réaliser ces protections, nous avons à disposition le nombre relativement important de diffusions de dopage offertes par une technologie SmartPower. Ainsi, pour faire face à des spécifications très sévères, la technologie nous donne une grande liberté pour innover, ce qui établit un contexte extrêmement motivant.

Le premier chapitre fixe le cadre dans lequel s'insèrent les travaux de la thèse. Tout d'abord, les formalismes permettant de définir les spécifications des structures de protection sont présentés. Après avoir décrit les standards industriels des tests de robustesse, la fenêtre de conception est définie, en combinant les spécifications de robustesse et de fiabilité. Des considérations sur les fenêtres de conception des E/S "hautes tensions" permettent de déterminer les caractéristiques des protections à développer. Notons dès à présent que les défis majeurs consistent à supprimer ou du moins limiter le repliement, c'est-à-dire la diminution de la tension au déclenchement, et à atteindre une robustesse élevée et une résistance à l'état passant (R_{ON}) faible, alors que la protection doit fonctionner à "haute tension". L'un des principaux objectifs industriels, la protection des E/S 80 Volts d'une application ethernet, est détaillée, donnant une évaluation quantitative des performances à atteindre. Ensuite, les particularités du comportement physique d'une protection ESD sont abordées. Il apparaît des densités de courant extrêmement élevées, ne se retrouvant dans aucune autre application, et une forte élévation de température, susceptible d'initier la défaillance. Pour caractériser un tel comportement en mesure, il existe des outils spécifiques permettant d'obtenir des caractéristiques électriques représentatives de la réponse à une ESD (mesures TLP et vfTLP), ou d'acquérir des images de la distribution thermique (mesures TIM). Enfin, nous rendons compte de l'état de l'art des protections des E/S "hautes tensions" en technologie SmartPower. Les circuits et structures de protections sont décrits tout en identifiant les solutions déjà brevetées. Une des protections conventionnelles pour protéger les E/S 80 Volts est également détaillée, mettant en évidence la nécessité de solutions alternatives. Nous avons choisi de développer cette solution alternative à partir de transistors bipolaires autopolarisés, sur lesquels l'ensemble des travaux s'est focalisé.

Au deuxième chapitre, une étude théorique des transistors bipolaires autopolarisés est menée. Dans un premier temps, une équation de polarisation est proposée en faisant abstraction des spécificités des protections ESD. A partir de cette équation, plusieurs stratégies d'optimisation sont définies, la plus significative concernant la possibilité de supprimer le repliement en laissant la base flottante. Dans un second temps, les effets des densités de courant élevées sont pris en compte. Aux valeurs atteintes, de nombreux aspects du fonctionnement d'un transistor bipolaire étaient encore dans l'ombre, aussi une étude analytique approfondie s'est avérée indispensable. Elle a permis de montrer pour la première fois certains phénomènes fondamentaux et a pu être mise à profit pour définir des règles d'optimisation portant, pour la plupart, sur l'obtention simultanée d'un faible R_{ON} à d'une tension élevée. Enfin, nous décrivons les modifications de la caractéristique électrique suite à l'élévation de température, ce qui permet de préciser la pertinence de certaines règles d'optimisation. Au final, il apparaît qu'au regard de la protection des E/S "hautes tensions", les points forts ou points faibles sont différents suivant qu'il s'agit d'un transistor PNP ou NPN. Par la suite, ces deux types seront étudiés séparément.

Le troisième chapitre traite du développement de protections "hautes tensions" à base de transistors bipolaires PNP. Dans la suite des travaux, nous nous sommes basés en grande partie sur la simulation électrothermique des protections à l'aide d'outils TCAD (Tools Computer Aided Design), cette approche étant extrêmement puissante pour comprendre leurs fonctionnements et les optimiser. En

préliminaire, les méthodes suivies pour calibrer les simulations sont exposées. Ensuite, les stratégies pour réduire le R_{ON} de transistors PNP latéraux sont complétées par des règles sur le layout. Les réalisations sur silicium montrent de bonnes performances, telles que la suppression du repliement, des tensions de fonctionnement élevées et des propriétés en termes de R_{ON} meilleures que celle de la protection usuelle. Afin d'améliorer encore le R_{ON} , une structure très innovante est développée en couplant le transistor PNP latéral avec une diode verticale contribuant à son courant de polarisation. Les particularités de cette structure amènent à définir de nouvelles règles d'optimisation. Leur mise en pratique permet d'obtenir des R_{ON} extraordinairement faibles et ainsi de développer une protection vérifiant les spécifications pour les E/S 80 Volts présentées dans le premier chapitre.

Dans le quatrième chapitre, nous passons au cas des transistors NPN. La principale difficulté consiste alors à obtenir une tension de fonctionnement élevée, les performances en termes de R_{ON} étant moins problématiques. Dans un premier temps, une structure originale est développée en mettant à profit les spécificités de la technologie, comme les tranchées d'isolation profondes, de manière à appliquer au mieux les règles définies dans le deuxième chapitre. Les mesures montrent d'excellentes propriétés telles qu'une tension supérieure à 40 Volts et un R_{ON} quasi-nul. Cependant, la tension au déclenchement monte au-dessus de 100 Volts et le temps de charge est trop long compte tenu de la rapidité d'une ESD. Aussi, nous proposons une stratégie de déclenchement utilisant un transistor PNP, sa caractérisation sur silicium montrant des propriétés compatibles avec la protection vis-à-vis des ESD. Dans un second temps, nous étudions des transistors avec une région flottante dans la base ou le collecteur. Le comportement en statique est décrit en tenant compte des effets bidimensionnels, puis la réponse à une ESD est analysée par la simulation. Les cas d'une région flottante dans la base ou dans le collecteur présentent tous les deux des propriétés intéressantes pour la protection des E/S "hautes tensions". Cependant, un transistor à région flottante est plus simple à mettre en œuvre et assure la suppression du repliement sans dégrader les performances du transistor NPN en termes de R_{ON} . Ces propriétés sont validées par des mesures, toutefois le temps a manqué pour développer des protections basées sur ce principe.

Chapitre1 Problématique et état de l'art de la protection des technologies SmartPower contre les décharges électrostatiques (ESD)

1.1. Introduction

Ce chapitre situe le contexte de la thèse de manière à mettre en évidence les principaux défis scientifiques et techniques. La motivation première est liée à l'évolution des spécifications requises pour les protections ESD, dans des secteurs industriels tels que l'automobile ou les applications domestiques (imprimantes, domotique...). Or, certaines de ces nouvelles spécifications exigent le développement de solutions innovantes, les protections courantes n'étant pas adaptées pour y faire face. Dans un premier temps, les contraintes auxquelles est soumise une protection ESD et les spécifications correspondantes sont présentées, permettant de définir la fenêtre de conception, c'est-à-dire la région du plan courant-tension (I-V) dans lequel doit être incluse la caractéristique de la protection à l'état passant. Les problématiques liées aux protections des E/S "hautes tensions" (allant de 40 à 80 Volts) en technologie SmartPower sont ensuite mises en évidence, à partir de considérations sur l'évolution de leurs fenêtres de conception. Dans un deuxième temps, les spécificités physiques à prendre en compte pour la conception d'une structure de protection ESD sont introduites, ainsi que les méthodes mises en œuvre pour caractériser ces modes de fonctionnement particuliers. La dernière partie dresse un état de l'art des stratégies et des structures de protection contre les ESD, pour le cas particulier des E/S "hautes tensions" en technologie SmartPower.

1.2. Spécifications des protections contre les décharges électrostatiques

Dans le domaine des ESD, une des propriétés essentielles d'un circuit intégré est sa robustesse, c'est-à-dire son aptitude à supporter une décharge sans être dégradé. Plus précisément, la robustesse est définie pour chacune des broches par le niveau de décharge auquel elle peut être soumise sans que les caractéristiques du circuit sortent des spécifications. Les différentes normes de tests industriels sont décrites en faisant ressortir les contraintes générées par chacune d'elles et les types de défaillance

qu'elles induisent. Ensuite, en combinant les spécifications en robustesse avec des contraintes liées à la fiabilité, la fenêtre de conception d'une structure de protection est définie. En particulier, il apparaît un intervalle de tension à l'intérieur duquel la caractéristique à l'état passant doit être comprise et une valeur maximale de la résistance à l'état passant (R_{ON}). Enfin, le cas particulier des E/S "hautes tensions" en technologie SmartPower est étudié, ce qui nous amène à présenter les objectifs de la thèse concernant les propriétés des protections à développer. Les défis scientifiques et techniques pour atteindre ces propriétés sont également présentés.

1.2.1. Robustesse

Les tests de robustesse consistent à reproduire les décharges que risque de subir le circuit intégré, de sa fabrication à son application finale dans un produit. Les normes associées, apparues dans les années 1970-1980, sont maintenant bien établies, garantissant en général une bonne reproductibilité des tests. Dans la plupart des cas, l'ESD est approchée par la décharge d'une capacité, dont la polarisation initiale permet de fixer la quantité de charges mises en jeu et donc le niveau de courant. Suivant les conditions de décharge de la capacité, différents types d'ESD seront reproduits. Quatre tests parmi les plus courants sont présentés ici : le test HBM modélisant la décharge d'une personne, le test MM modélisant la décharge d'une machine, le test "pistolet" correspondant à des décharges très énergétiques pouvant survenir dans l'application finale et le test CDM modélisant la décharge du composant lui-même chargé. Au préalable, il est important de préciser que les tests HBM et MM permettent de caractériser aussi bien un circuit intégré qu'une structure de protection seule, alors que les tests "pistolet" et CDM s'appliquent au niveau du système.

1.2.1.1. Test HBM

Le test HBM [1] [2] [3] consiste à appliquer sur une broche du circuit une forme d'onde en courant représentant la décharge produite par une personne debout qui la toucherait du doigt, les autres broches étant généralement reliées à la masse. La forme d'onde est reproduite via un circuit électrique modélisant le corps humain, et dont les éléments de base sont une capacité C_{HBM} de 100 picoFarads et une résistance R_{HBM} de 1500 Ohms montées en série (Figure 1(a)). La tension initiale de charge de la capacité est fixée par la spécification en robustesse, allant typiquement de 2 à 8 10^3 Volts. Dans la suite, cette tension de charge sera donnée en kiloVolt, correspondant à 10^3 Volts et noté kV. Lors de la décharge de la capacité, l'influence du composant testé sur la forme d'onde en courant peut le plus souvent être négligée, son impédance étant faible devant celle du circuit équivalent.

La forme d'onde ne dépend pas seulement de la capacité C_{HBM} et de la résistance R_{HBM} , mais aussi des éléments parasites de l'appareil de test, essentiellement l'inductance L_p et les capacités C_{p1} et C_{p2} (Figure 1(b)). Aussi, ces éléments parasites doivent être calibrés, ce qui a conduit à définir des gabarits pour la forme d'onde. Par exemple, la norme [2] stipule que, pour une impédance de charge nulle, le temps de montée, de 10 % à 90 % du courant maximal, doit être compris entre 2 et 10ns et que la constante de temps de la décroissance doit être comprise entre 130 nanosecondes et 170 nanosecondes. Ces valeurs correspondent aux facteurs contrôlant la réponse temporelle du circuit équivalent de l'appareil de test. Si les capacités parasites C_{p1} et C_{p2} sont négligeables, hypothèse justifiée dans la plupart des cas, une expression analytique relativement simple décrit la réponse temporelle.

$$I(t) = \frac{V_{HBM}}{R_{HBM}} \left(1 - \exp\left(-\frac{R_{HBM}}{L_p} t\right) \right) \exp\left(-\frac{t}{R_{HBM} C_{HBM}}\right) \quad (1)$$

Où V_{HBM} est la tension initiale aux bornes de la capacité C_{HBM} . En première approximation le temps de montée est donné par :

$$t_m = \ln 9 \frac{L_p}{R_{HBM}} = 2,2 \frac{L_p}{R_{HBM}} \quad (2)$$

Pour respecter le gabarit [2], la valeur de l'inductance parasite devra typiquement être comprise entre 2 et 5 mH. Concernant La constante de temps de la décroissance, elle est donnée par :

$$\tau = R_{HBM} C_{HBM} \quad (3)$$

Cette formule permet de fixer la précision sur les valeurs de la résistance R_{HBM} et de la capacité C_{HBM} . D'après la norme [2], R_{HBM} doit être ajustée à plus ou moins 1 % et C_{HBM} à plus ou moins 10 %, ce qui permet d'être en accord avec le gabarit.

La formule (1) permet également de déterminer le courant maximal :

$$I_{Max} = \frac{V_{HBM}}{R_{HBM}} \quad (4)$$

En pratique, le courant maximal peut être évalué rapidement à partir de la proportionnalité avec la tension de charge ($V_{HBM}=1 \text{ kV} \Leftrightarrow I_{Max}=670 \text{ mA}$).

La forme d'onde simulée pour une tension de charge V_{HBM} de 2 kV est présentée sur la Figure 2, où l'inductance série L_p est égale à 5 mH, et les capacités parasites C_{p1} et C_{p2} sont négligées. Les valeurs des principales caractéristiques de la forme d'onde sont listées dans le Tableau 1, où elles sont comparées avec celles calculées à partir de l'équation (1).

Concernant les modes de défaillance, l'énergie dissipée lors d'une décharge HBM peut être suffisante pour faire fondre localement le silicium. Le plus souvent cette fusion apparaît au niveau d'une jonction en inverse, où le fort champ électrique induit une dissipation énergétique par effet Joule élevée. Il en résulte une redistribution des dopants, qui se manifeste par une modification du courant de fuite, voire de la tension de claquage. Si la fusion du silicium atteint un contact métallique, il se forme un alliage entre les deux matériaux pouvant court-circuiter la jonction. Un second mode de défaillance survient si la conduction de la décharge induit une différence de potentiel aux bornes d'un oxyde supérieure à sa tension de claquage. Il se crée alors un chemin de conduction entraînant un perçage quasi-instantané et irréversible de l'oxyde. Néanmoins, ce mode de défaillance est peu fréquent et résulte généralement d'une mauvaise conception du circuit. Pour une structure de protection seule, le test HBM permet principalement d'évaluer l'énergie que celle-ci peut supporter.

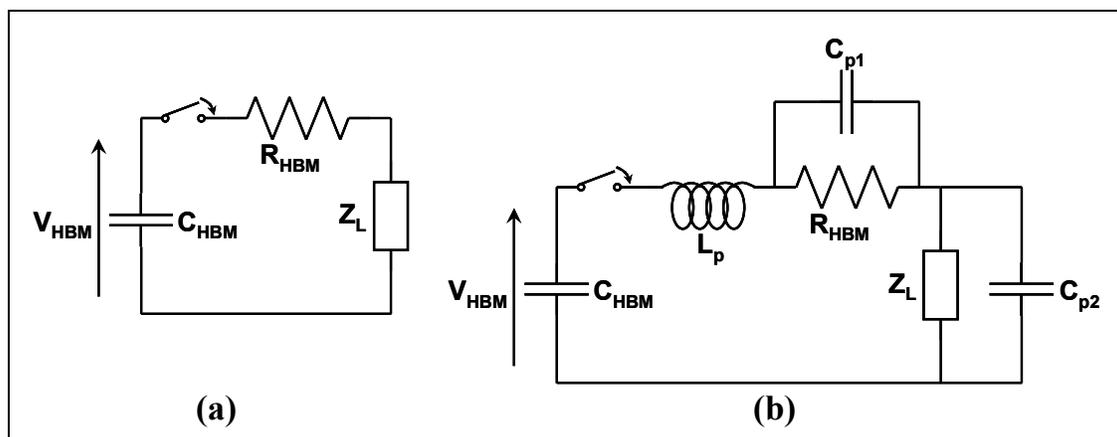


Figure 1: Modèles électriques du corps humain (a) et de l'appareil de test correspondant (b).

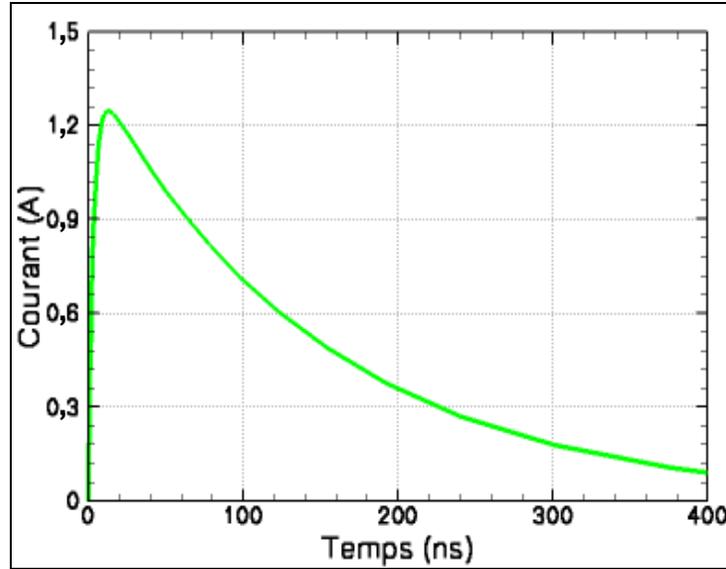


Figure 2: Forme d'onde en courant d'une décharge de 2 kV HBM.

Tableau 1: Principales caractéristiques d'une décharge de 2 kV HBM (Figure 2).

	I_{Max} (A)	τ (ns)	t_m (ns)
Simulation	1,25	150	6,1
Calcul	1,3	150	7,3

1.2.1.2. Test MM

Le principe du test MM [4] [5] [6] est le même que celui du test HBM, à la différence que la forme d'onde du courant représente la décharge produite d'une machine, par exemple un robot ou un outil métallique tenu par une personne. Le circuit électrique équivalent est formé d'une capacité C_{MM} de 200 picoFarads en série avec une résistance dont la valeur doit être inférieure à 10 Ohms, correspondant généralement à la résistance parasite R_p de l'appareil de test (Figure 3(a)). Le modèle MM peut être considéré comme un pire cas du modèle HBM, lorsque le courant maximal n'est plus limité que par les résistances parasites. Aussi, les tensions de charge de la capacité sont plus faibles, typiquement de quelques centaines de Volts. Par ailleurs, la forme d'onde présente des oscillations entraînant une inversion du sens de conduction à une fréquence de plusieurs megaHertz.

Les appareils de test sont calibrés en fonction de gabarits normalisés de la forme d'onde. Selon la norme [5], une décharge de 200 Volts sur un court-circuit doit conduire à un maximum de courant et une fréquence de résonance compris respectivement entre 2,8 et 3,8 Ampères et entre 11 et 16 MHz. Des expressions analytiques du maximum I_{Max} en courant et de la fréquence de résonance f_0 peuvent être déterminées à partir de la réponse en courant du modèle d'appareil de test (Figure 3(b)), dans le cas où les capacités parasites C_{p1} et C_{p2} sont négligeables :

$$I(t) = Ae^{-\frac{t}{\tau}} \sin(2\pi f_0 t) \quad (5)$$

Avec :

$$A = V_{MM} \sqrt{\frac{C_{MM}}{L_p}} \approx I_{Max} \quad (6)$$

$$\tau = \frac{2L_p}{R_p} \quad (7)$$

$$f_0 = \frac{1}{2\pi} \sqrt{\frac{1}{L_p C_{MM}}} \quad (8)$$

Où V_{MM} est la tension initiale aux bornes de la capacité C_{MM} . Pour respecter les contraintes relatives à I_{Max} et f_0 [5], la précision sur la capacité C_{MM} doit être de 10 % et la valeur de l'inductance parasite L_p doit être comprise entre 0,5 et 1,1 microHenry.

La constante de temps de la décroissance correspond au paramètre τ . Le temps de montée de 10 % à 90 % du courant maximal est donné par :

$$t_m = \frac{1}{2\pi f_0} (\text{Arc sin}(0,9) - \text{Arc sin}(0,1)) \approx \frac{1}{2\pi f_0} \quad (9)$$

La forme d'onde simulée pour une tension de charge V_{MM} de 200 Volts est présentée sur la Figure 4, où la résistance R_p et l'inductance L_p sont respectivement de 10 Ohms et 1 microHenry, et où les capacités parasites C_{p1} et C_{p2} sont négligées. Les valeurs des principales caractéristiques de la forme d'onde sont listées dans le Tableau 2, où elles sont comparées avec celles calculées à partir de l'équation (5).

Les tests MM induisent des modes de défaillance similaires à ceux des tests HBM, c'est-à-dire une fusion du silicium ou un claquage d'oxyde. Toutefois, une bonne robustesse HBM n'implique pas nécessairement une bonne robustesse MM. Le test MM permet d'évaluer simultanément les robustesses en direct et en inverse, contrairement au test HBM qui ne met en jeu qu'un seul sens de conduction [7]. De plus, la commutation d'un sens de conduction à l'autre doit être suffisamment rapide pour suivre les oscillations. En particulier, des dégradations peuvent survenir si le temps de recouvrement des diodes en direct est trop long [8]. Ainsi, le test MM est un complément pertinent du test HBM. Cependant, il est relativement peu pratiqué car sa reproductibilité reste problématique, malgré les normes existantes.

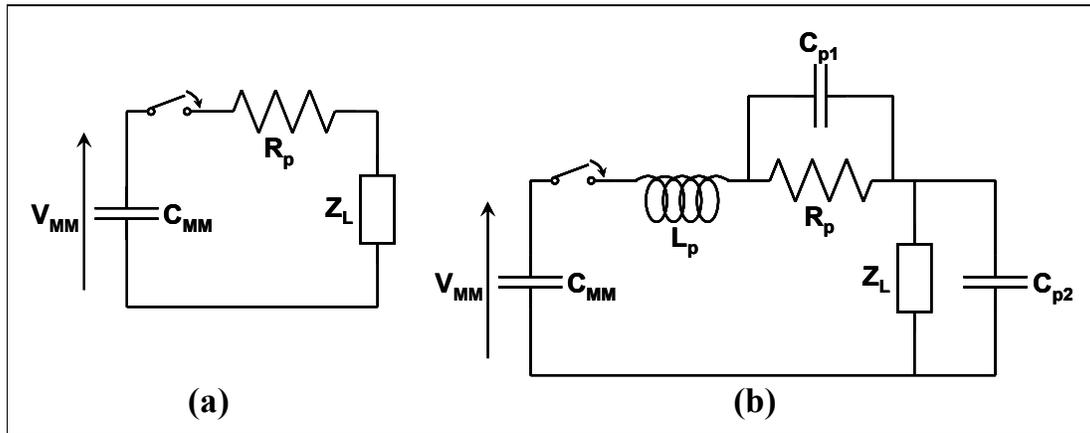


Figure 3: Modèles électriques d'une machine (a) et de l'appareil de test correspondant (b).

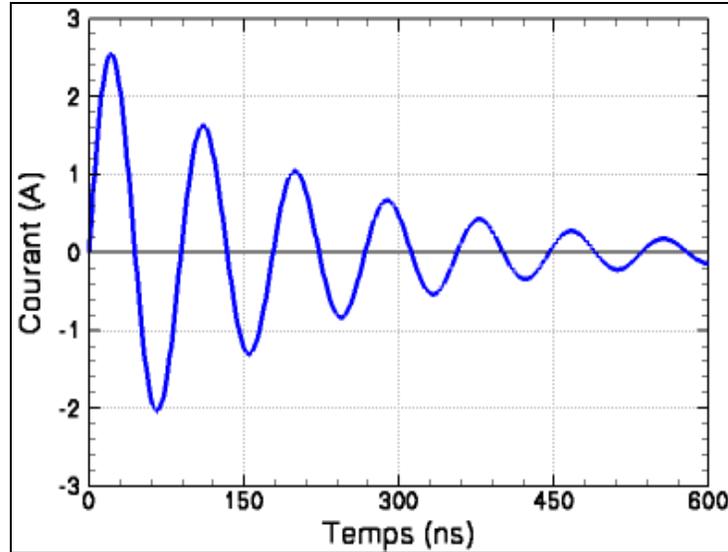


Figure 4: Forme d'onde en courant d'une décharge MM de 200 Volts.

Tableau 2: Principales caractéristiques d'une décharge MM de 200 Volts (Figure 4).

	I_{Max} (A)	f_0 (MHz)	τ (ns)	t_m (ns)
Simulation	2,5	11	200	12,5
Calcul	2,8	11	200	15

1.2.1.3. Test "pistolet"

Le test "pistolet" [9] [10] concerne des applications dans lesquelles des décharges très énergétiques peuvent survenir. Par exemple, dans le domaine automobile, il permet de représenter des décharges issues de moteurs électriques tournants (essuie-glace, lève-vitre...), ayant accumulé des quantités de charges très importantes. Ce test est effectué au niveau de l'application finale, ce qui donne la possibilité d'insérer des protections discrètes. Néanmoins, il est parfois plus économique d'intégrer la protection sur la puce du circuit à protéger.

En ESD, un pistolet est un appareil manipulable manuellement, permettant de venir appliquer une décharge en différents points d'une carte électronique (Figure 5). Il comprend une capacité et une résistance en série de valeurs respectives 150 ou 330 picoFarads et 330 Ohms ou 2 kiloOhms. La polarisation de la capacité est typiquement comprise entre 8 et 25 kV. A 25 kV, la charge stockée est de 3,75 μC pour la capacité de 150 picoFarads et de 8,25 μC pour la capacité de 330 picoFarads. A titre de comparaison, la charge stockée pour un test HBM de 8 kV est de 0,8 μC .

Le contrôle de la forme d'onde en courant est moins critique que pour les tests HBM et MM, le paramètre essentiel du test "pistolet" étant la quantité de charges injectée et l'énergie associée. Une forme d'onde typique est illustrée par la simulation d'une décharge de 15 kV, dans le cas d'une capacité de 330 picoFarads et 330 Ohms (Figure 6). Celle-ci met en évidence deux pics de courant, dont les caractéristiques sont résumées dans le Tableau 3. Le premier s'étend sur quelques nanosecondes en début d'impulsion, il est caractérisé par courant maximal I_{Max1} particulièrement élevé (vingt cinq fois supérieur au courant maximal d'une décharge HBM de 2 kV), atteint en un temps $t(I_{Max1})$ inférieur à une nanoseconde. Le second pic présente une dynamique nettement plus lente, il s'étend sur plus de cent nanosecondes, cependant son courant maximal I_{Max2} reste très élevé (seize fois plus élevé que pour une décharge HBM de 2 kV). Ainsi, les charges sont majoritairement conduites par le second pic, au cours duquel la dissipation énergétique sera prédominante.

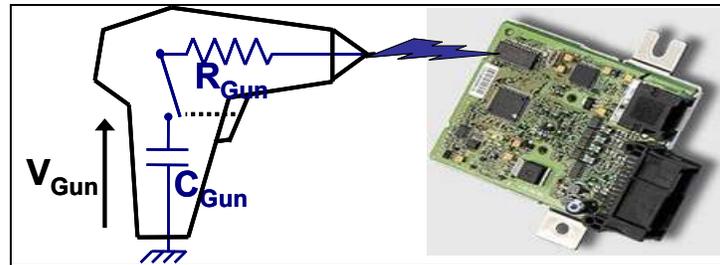


Figure 5: Représentation schématique d'un test "pistolet".

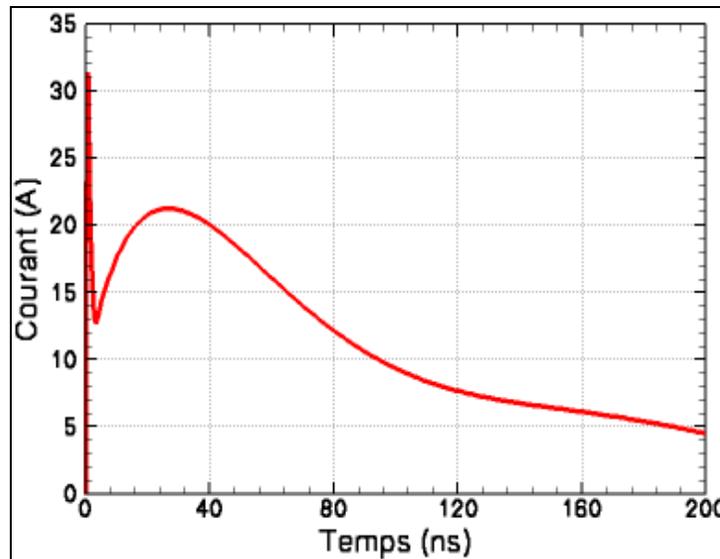


Figure 6: Forme d'onde en courant d'une décharge "pistolet" de 15 kV (330 pF et 330 Ohms).

Tableau 3: Principales caractéristiques d'une décharge "pistolet" de 15 kV (Figure 6).

I_{Max1} (A)	$t(I_{Max1})$ (ns)	I_{Max2} (A)	$t(I_{Max2})$ (ns)	τ (ns)
31	0,7	21	26	71

Deux catégories de modes de défaillance peuvent être distinguées suivant qu'elles sont induites par le premier ou le second pic. Concernant le premier pic, le risque majeur est lié à la rapidité de la montée du courant. La structure de protection doit alors présenter un temps de déclenchement suffisamment court pour éviter l'apparition d'une surtension qui conduirait au claquage des oxydes des E/S à protéger. Etant donné que l'énergie mise en jeu reste limitée, une dégradation thermique de la structure de protection est peu probable. Au cours du second pic, la structure de protection doit supporter une très forte dissipation énergétique. Dans ces conditions, les défaillances par fusion du silicium sont les plus courantes. Ces défaillances sont du même type que celles observées en HBM, en plus étendues du fait des très fortes énergies des tests "pistolets". Par ailleurs, il faut préciser que le premier pic peut très bien être filtré par les éléments parasites du système testé. Ainsi, le test "pistolet" se rapprocherait d'un test HBM, la différence n'étant que l'ordre de grandeur du courant de décharge.

1.2.1.4. Test CDM

Le modèle CDM [11] [12] correspond à la décharge lorsqu'une broche d'un composant préalablement chargé touche un plan de masse. Aussi, le principe du test CDM ne consiste pas à appliquer une décharge de forme d'onde prédéfinie entre deux broches, comme c'est le cas pour les tests HBM et MM, mais à reproduire les conditions de charge et de décharge du circuit intégré. La

méthodologie la plus courante consiste à charger le composant par induction en le plaçant sur l'armature de polarisation positive d'un condensateur, puis à le décharger en approchant de la broche testée une pointe reliée à l'armature de masse. La tension du condensateur est fixée par la spécification requise, généralement de plusieurs centaines de Volts. La forme d'onde générée lors d'un test CDM dépend, d'une part, de l'appareil utilisé [13] via l'inductance de la pointe de test L_p et la capacité du condensateur C_{GP} [14], et, d'autre part, du composant testé [15] via sa capacité C_{DUT} vis-à-vis de l'armature du condensateur, dépendant à la fois de la puce et du type de boîtier (Figure 7(a)(b)). La résistance la plus significative est celle de l'arc électrique entre la pointe de test et la broche du composant.

La quantité de charges stockées pour un test CDM est beaucoup plus faible que pour les tests HBM et MM, les valeurs des capacités C_{GP} et C_{DUT} étant typiquement d'une dizaine de picoFarads alors que C_{HBM} et C_{MM} sont respectivement de 100 picoFarads et 200 picoFarads. Cette faible quantité de charges stockées se traduit par une décharge moins énergétique et plus courte. En revanche, un temps de montée très court et un courant maximal très élevé sont attendus du fait des faibles valeurs de l'inductance L_p et de la résistance R_p . Afin d'évaluer précisément ces tendances, la forme d'onde en courant est simulée pour une polarisation de 250 Volts (Figure 8), en prenant comme valeurs des éléments du schéma équivalent celles listées dans le Tableau 4. La durée de cette décharge n'est que de quelques nanosecondes, cependant le courant maximal est élevé (six fois supérieur à celui d'une décharge HBM de 2 kV), et le temps de montée extrêmement court (la pente est sensiblement la même que pour une décharge "pistolet"). De plus, la forme d'onde présente des oscillations entraînant une inversion du courant à une fréquence de 540 MHz (bien supérieure à celle d'une décharge MM qui n'est seulement que d'une dizaine de MégaHertz).

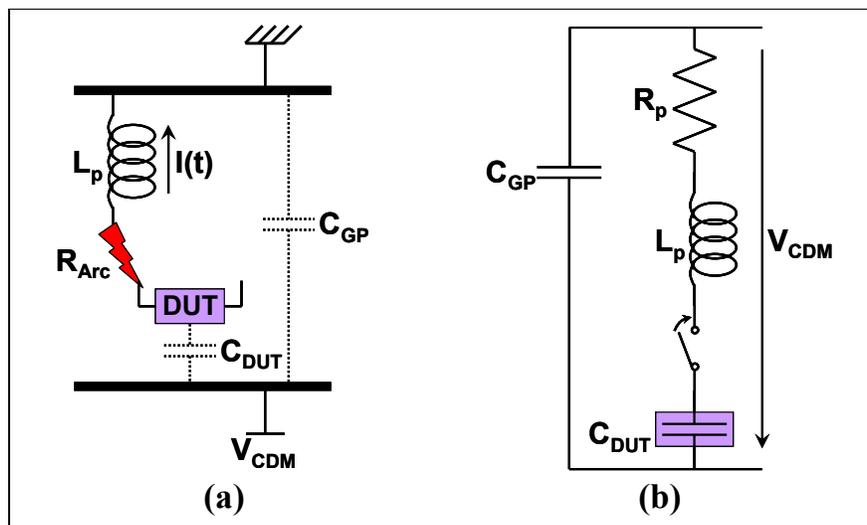


Figure 7: Modélisation d'un appareil de test CDM (a) et schéma électrique équivalent (b).

Tableau 4: Valeurs typiques pour les éléments du schéma électrique équivalent de la Figure 7(b).

C_{DUT} (pF)	R_p (Ω)	L_p (μ H)	C_{GP} (pF)
25	10	6,5	25

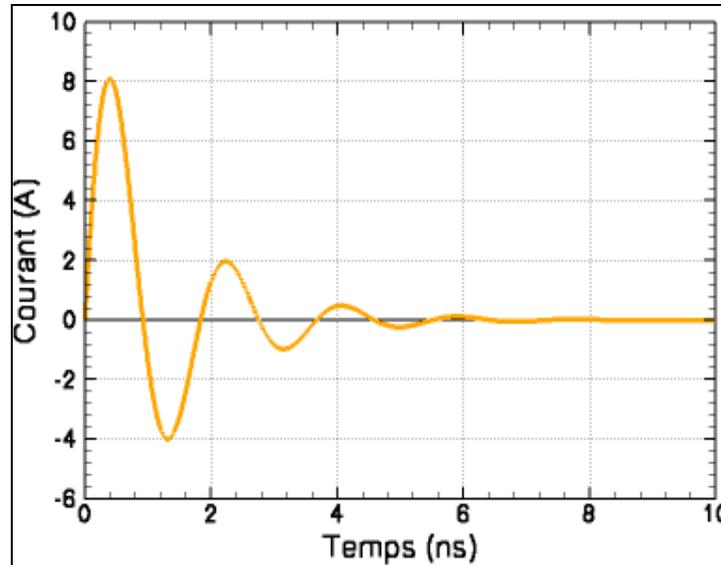


Figure 8: forme d'onde en courant d'une décharge CDM de 250 Volts.

Tableau 5: Principales caractéristiques de la décharge CDM de 250 Volts présentée sur la Figure 8.

I_{Max} (A)	f_0 (MHz)	τ (ns)	t_m (ns)
8,1	540	1,3	250

L'énergie dissipée lors d'une décharge CDM est généralement beaucoup trop faible pour entraîner une fusion du silicium. Par contre, la structure de protection doit se déclencher rapidement pour éviter que la tension atteigne le claquage des oxydes. Cette problématique est la même que pour le premier pic d'une décharge "pistolet". Par ailleurs, les changements de sens de conduction peuvent également conduire à des défaillances si la commutation de la structure de protection est trop lente pour suivre les oscillations. Dans de nombreux cas, le test CDM pourra donc se substituer au test MM, lorsqu'il s'agit d'évaluer l'influence sur la robustesse des inversions du sens de conduction. Enfin, les charges se déplaçant de l'intérieur de la puce vers l'extérieur, la stratégie de protection au niveau du circuit est déterminante. Même si la structure de protection est adaptée aux contraintes des décharges CDM, elle sera totalement inopérante si la dégradation survient avant que le courant ne l'ait atteinte.

1.2.1.5.Méthodologie retenue pour la caractérisation de la robustesse

Le travail de thèse porte sur le développement de structures de protection. Aussi, les caractérisations sont effectuées pour des structures seules et non pour des circuits complets. Dans ces conditions, seuls les tests HBM et MM sont applicables. Pour chacun d'eux, les tensions de charge nécessaires pour dégrader la protection sont mesurées, avec l'appareil de test du laboratoire de fiabilité de Freescale à Toulouse. En revanche, les tests pistolet et CDM, portant sur des circuits montés en boîtier, ne s'appliquent pas. Les propriétés des protections vis-à-vis de ces tests sont évaluées indirectement. En particulier, cela nécessite caractériser la dynamique de déclenchement et l'uniformité de la conduction du courant.

1.2.2. Définition de la fenêtre de conception

La fenêtre de conception délimite les régions du plan I-V autorisées pour le fonctionnement d'une structure de protection ESD (Figure 9). Elle constitue une synthèse des principales spécifications que doit vérifier la caractéristique à l'état passant.

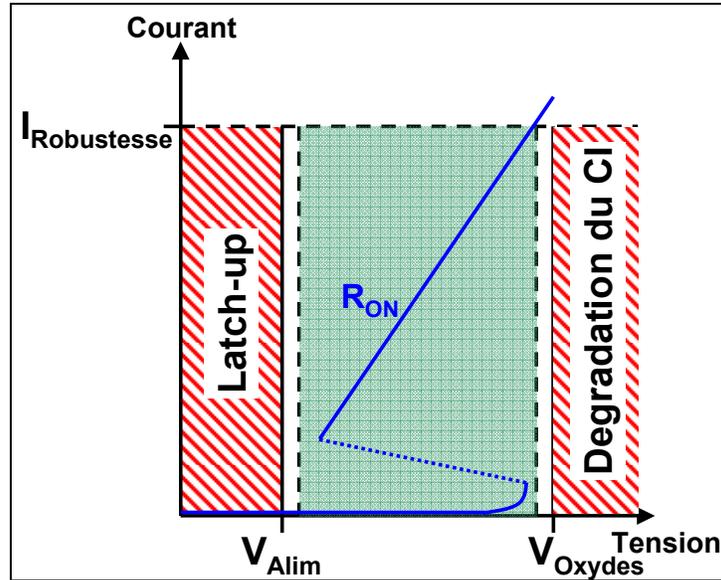


Figure 9: Fenêtre de conception typique d'une structure de protection ESD.

1.2.2.1. Intervalle autorisé pour la tension à l'état passant

La protection ne joue son rôle que si la tension reste inférieure aux valeurs pour lesquelles le circuit serait dégradé. Typiquement, cette limite est donnée par la tension de claquage des oxydes (V_{Oxydes}), qui dépend uniquement des caractéristiques de la technologie. En particulier, la réduction de l'épaisseur des oxydes de grille avec la réduction des dimensions lithographiques entraîne une diminution de cette tension de claquage pour les nouvelles générations technologiques.

La limite inférieure de la tension est fixée de manière à ce que la protection ne puisse pas être polarisée à l'état passant sous une tension continue. Dans le cas contraire, la protection risquerait de partir en latch-up, c'est-à-dire qu'elle resterait en permanence à l'état passant sans pouvoir repasser à l'état bloqué. Or, une protection ESD étant dimensionnée pour conduire le courant en transitoire et non continu, elle serait inmanquablement dégradée. Cette dégradation est susceptible d'avoir des conséquences dramatiques pour la fiabilité, pouvant aller jusqu'à un départ de feu. En général, la limite est donnée par la tension d'alimentation de l'application (V_{Alim}). Cette limite correspond au cas où la protection est déclenchée lorsque le circuit est en fonctionnement, par exemple par un pic de tension parasite. Elle devra alors repasser à l'état bloqué avant que l'alimentation se stabilise à sa valeur nominale, sans quoi elle sera maintenue à l'état passant. Si des protections externes sont rajoutées pour faire face à des "décharges ISO" très énergétiques [16], il faudra prendre pour limite la tension de fonctionnement maximale de ces protections. Compte tenu de la durée d'une "décharge ISO", la protection ESD verrait une telle décharge comme une polarisation continue. Pour la norme [16], la durée moyenne est de 2 ms, soit un rapport de 10^4 comparé à une ESD de type HBM.

L'intervalle de tension autorisé doit prendre en compte des marges de sûreté vis-à-vis de ces deux limites. Ces marges sont déterminées en fonction des variations avec la température et la dispersion du procédé technologique. D'une part, les spécifications dans le domaine automobile doivent généralement être vérifiées typiquement entre -40°C et 200°C . D'autre part, la plupart des règles de fiabilité exigent une tolérance de 6σ sur la tension de claquage, où σ est l'écart type de la gaussienne.

1.2.2.2. Résistance à l'état passant (R_{ON}) maximale

Une protection ESD doit présenter un R_{ON} suffisamment faible de manière à ce que la tension de dégradation du circuit intégré ne soit pas dépassée. Les contraintes sur le R_{ON} seront d'autant plus sévères que l'intervalle de tension autorisé est étroit et que le maximum de courant atteint au cours de la décharge ($I_{Robustesse}$) est élevé. Ce maximum de courant est directement lié à la spécification en robustesse. Pour l'exemple d'une spécification HBM, il est proportionnel à la charge initiale de la capacité ($V_{HBM}=1 \text{ kV} \Leftrightarrow I_{Max}=670 \text{ mA}$). Ainsi, la robustesse conduit à deux exigences concernant la protection ESD, l'une sur le niveau de la décharge qu'elle doit supporter et l'autre sur son R_{ON} .

1.2.3. Spécifications des protections dédiées aux E/S "hautes tensions" en technologies SmartPower

Maintenant que les éléments de la spécification d'une protection ESD ont été décrits, il reste à déterminer comment ils seront influencés pour le cas traité au cours de cette thèse, c'est-à-dire celui des E/S "hautes tensions" (40 Volts – 80 Volts) en technologie SmartPower. Ces considérations permettront de définir les caractéristiques des protections à développer. Ensuite, les principaux défis à relever pour obtenir ces caractéristiques seront identifiées et décrits avec précision. Enfin, les performances à atteindre seront illustrées par un exemple particulièrement critique d'une spécification pour une application ethernet. Entre autres, cet exemple amènera à présenter la technologie qui nous est imposée pour les réalisations sur silicium.

1.2.3.1. Fenêtre de conception et propriétés de la structure de protection

Dans une application en technologie SmartPower, la tension d'alimentation peut aller de quelques Volts à près de cent Volts, suivant que le bloc concerne la partie commande ou puissance. Ces variations se répercutent directement sur la limite inférieure de la fenêtre de conception (Figure 10). Ainsi, la première propriété de la protection d'une E/S "haute tension" est de commuter de l'état bloqué à l'état passant pour une tension élevée.

Par ailleurs, la limite supérieure de la fenêtre de conception ne peut pas dépasser la tension de claquage des oxydes, dont la valeur est fixée pour une technologie donnée. En conséquence, les fenêtres de conception des protections des E/S "hautes tensions" sont très étroites (Figure 10). Dans cette condition, la caractéristique électrique ne doit pas se replier, ou alors le repliement doit être faible. Ce repliement correspond à la diminution de la tension lors de la commutation à l'état passant. Il sera étudié en détail dans la partie sur l'état de l'art (1.4). Pour le moment, il suffit de préciser que c'est une propriété quasi-générale des protections ESD.

Enfin, l'ordre de grandeur de la spécification en robustesse est indépendant de l'E/S. Aussi, une E/S "haute tension" peut avoir à tenir des niveaux de décharges élevés. Le courant maximal sera également élevé, ce qui, combiné avec l'intervalle de tension étroit, impose un R_{ON} de faible valeur.

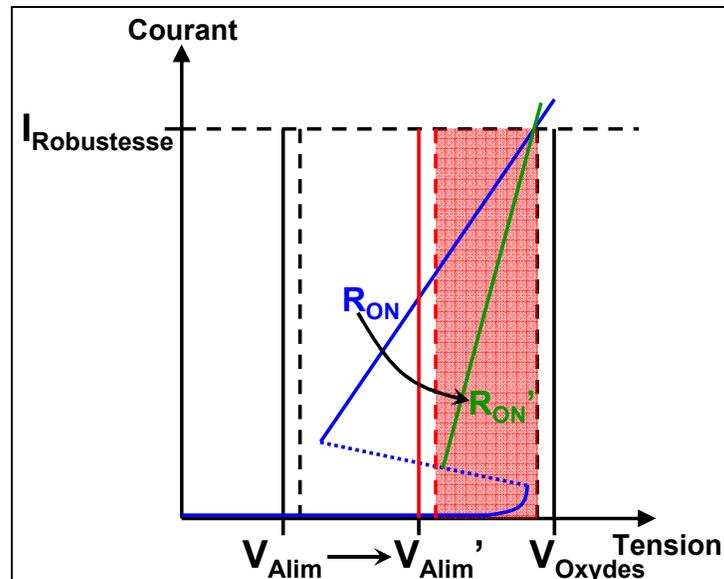


Figure 10: Modification de la fenêtre de conception suite à une augmentation de la tension d'alimentation.

1.2.3.2. Difficultés scientifiques et techniques

Les trois principaux défis sont de supprimer ou réduire le repliement tout en gardant une tension à l'état passant élevée, de garantir une bonne robustesse et d'atteindre un faible R_{ON} .

Réduction ou suppression du repliement

Comme nous le verrons dans la partie 1.4, la plupart des protections ESD usuelles présentent un fort repliement. C'est-à-dire que même si, au déclenchement, la tension est élevée, elle retombera aussitôt à des valeurs relativement faibles. Or, un tel comportement est intimement lié aux mécanismes physiques survenant dans la protection. Pour s'en affranchir, de nouveaux concepts basés sur des composants innovants devront être développés.

Robustesse élevée

La tension de fonctionnement élevée a une influence négative sur la robustesse. En effet, l'énergie dissipée dans la protection est égale à l'intégrale sur la durée de la décharge du produit du courant par la tension. La forme d'onde du courant est imposée par la spécification en robustesse, indépendamment du comportement de la protection. Par contre, la forme d'onde de la tension est fixée par la réponse de la protection. En conséquence, une tension de fonctionnement élevée entraîne une forte dissipation énergétique, et donc une forte élévation de température risquant de conduire à une fusion prématurée du silicium.

Dans de nombreux cas, il suffit d'agrandir la protection pour augmenter sa robustesse. Si la dissipation énergétique se répartit sur l'ensemble de la structure, localement l'énergie et la température seront d'autant plus faibles que la surface est grande. Cependant, la protection doit être la plus petite possible, pour minorer son coût. Aussi, c'est l'aptitude d'une protection à supporter une ESD qui devra être considéré.

Afin de pouvoir comparer les protections entre elles indépendamment de leurs tensions de fonctionnement et de leurs surfaces, nous proposons de définir un facteur de mérite de la robustesse :

$$F_{I2} = \frac{I_{I2} V_{I2}}{S} \quad (10)$$

Où I_{t2} et V_{t2} sont respectivement le courant et la tension auxquels la protection est dégradée et S sa surface. L'unité usuelle de F_{t2} est le $\text{mW} \cdot \mu\text{m}^{-2}$. Concernant le courant I_{t2} et la tension V_{t2} , ils sont relevés sur la caractéristique TLP. Cette caractéristique TLP permet de représenter le fonctionnement d'une protection pendant une ESD, elle sera introduite dans la partie 1.2.1.2. Si le courant est réparti uniformément, I_{t2} est proportionnel à S (densité de courant à la défaillance constante) et donc F_{t2} est le même quel que soit S . Si plusieurs structures sont mises en série pour augmenter la tension, alors à la fois V_{t2} et S sont multipliés par le nombre de structures, de sorte que F_{t2} est constant. Enfin, de bonnes propriétés en robustesse correspondent à une valeur élevée de F_{t2} .

Les simulations ne permettent pas d'obtenir le point (I_{t2} , V_{t2}) et donc de calculer le facteur de mérite F_{t2} . Pour se faire une idée de l'aptitude à supporter des énergies élevées, un facteur de mérite thermique F_T est défini à partir des résultats de simulation :

$$F_T = \frac{IV}{ST_{Max}} \quad (11)$$

Où T_{Max} est la température maximale à l'intérieur de la structure. L'unité usuelle de F_{t2} est le $\text{mW} \cdot \mu\text{m}^{-2} \cdot \text{K}^{-1}$. La température n'étant pas proportionnelle à l'énergie dissipée, F_T sera toujours évaluée au voisinage de 600 degrés Kelvins, limite jusqu'à laquelle sont validés les modèles du simulateur.

Faible R_{ON}

La difficulté pour atteindre un faible R_{ON} résulte de la tension de fonctionnement élevée. En technologie SmartPower, le R_{ON} dépend principalement des variations de tension induites par la modification du champ électrique au voisinage des jonctions en inverse. Afin de mettre en évidence la relation entre le R_{ON} et la tension de fonctionnement, la surtension induite par une augmentation de champ maximal est étudiée dans le cas d'un profil triangulaire (Figure 11). La pertinence de cet exemple sera justifiée par les conclusions du chapitre 2. Il y sera montré, d'une part, que l'accroissement du courant entraîne une augmentation du maximum du champ électrique et, d'autre part, que le profil de dopage optimal vis-à-vis du R_{ON} conduit à un profil de champ triangulaire. La tension étant égale à l'intégrale du champ, son expression pour un champ maximal E et une extension W est donnée par :

$$V = \frac{1}{2} E W \quad (12)$$

Pour une augmentation du champ maximal ΔE entraînant un élargissement de l'extension ΔW , le calcul de la surtension ΔV donne :

$$\Delta V = \frac{1}{2} (E \Delta W + W \Delta E + \Delta E \Delta W) \quad (13)$$

De manière générale, la tension soutenue dépend seulement de l'extension W , la valeur du champ maximal E étant fixée par le courant. D'après (12), plus l'extension est grande, plus la tension est élevée. Mais, d'après (13), la surtension pour un même ΔE sera également plus élevée. Par exemple, si l'on veut doubler la tenue en tension, l'extension sera deux fois plus grande. De même, l'élargissement résultant de l'augmentation du champ maximal est deux fois plus grand, ce qui peut être montré par des considérations géométriques élémentaires. Au final, la surtension aura doublé, ce qui équivaut en première approximation à une dégradation du R_{ON} d'un facteur deux. Par ailleurs, le R_{ON} peut être réduit par une augmentation de la surface de la structure. Cependant, cette solution conduit à un surcroît de coût. De même que pour la robustesse, un facteur de mérite permettant d'évaluer les propriétés en termes de R_{ON} s'avère nécessaire.

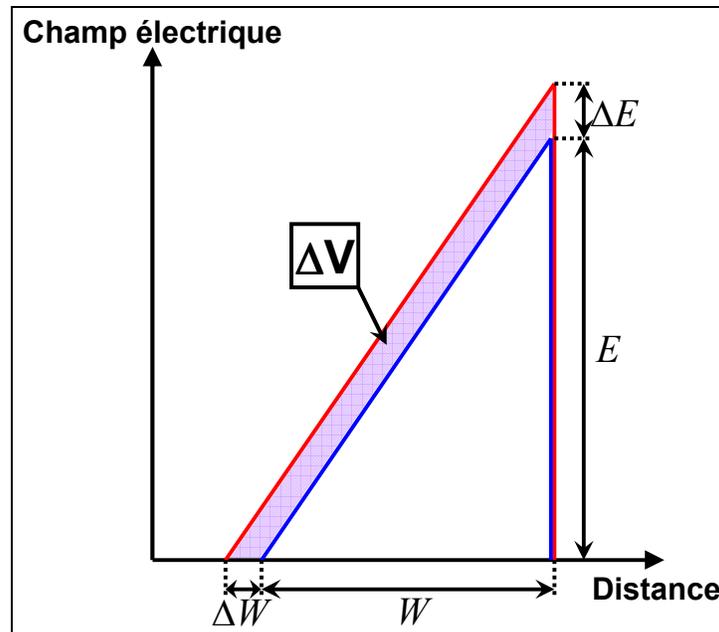


Figure 11: Effets de l'augmentation de la valeur maximale d'un champ électrique de profil triangulaire.

Le facteur de mérite du R_{ON} (F_{RON}) est défini par :

$$F_{RON} = \frac{R_{ON} S}{V^2} \quad (14)$$

L'unité usuelle de F_{RON} est le $\mu\text{m}^2 \cdot \text{W}^{-1}$. Dans le cas où le courant est réparti uniformément sur la protection, R_{ON} est proportionnel à S , donc F_{RON} est indépendant de S . Si plusieurs structures sont mises en série pour augmenter la tension, les trois paramètres R_{ON} , S et V sont multipliés par le nombre de structures, de sorte que F_{RON} reste constant. Les propriétés en termes de R_{ON} seront d'autant meilleures que F_{RON} est faible.

1.2.3.3. Protection "haute tension" pour une application ethernet

Une partie du travail de thèse a consisté à développer des protections dédiées pour cette application. Les spécifications donnent un ordre de grandeur des performances à atteindre pour un cas particulièrement contraignant.

Spécifications

Les spécifications requises sont résumées dans le Tableau 6, où V_{\min} et V_{\max} sont respectivement les bornes inférieures et supérieures de la fenêtre de conception, V_{HBM} la tension de charge pour le test de robustesse HBM et S_{Max} la surface maximale autorisée.

Tableau 6: Spécifications requises.

V_{\min} (V)	80
V_{\max} (V)	100
V_{HBM} (kV)	2
S_{Max} (μm^2)	150*150

Ces données permettent de calculer le R_{ON} , la puissance à tenir P_{I_2} et les facteurs de mérite F_{RON} et F_{I_2} (Tableau 7). La correspondance entre la charge du test HBM et le courant maximal atteint donne 1,3 Ampères pour 2 kV. Le courant I_{I_2} sera approché par cette valeur, pour le calcul de P_{I_2} et F_{I_2} .

Tableau 7: Performances à atteindre.

R_{ON} (Ω)	15
P_{I_2} (W)	130
F_{RON} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	42
F_{I_2} ($\text{mW} \cdot \mu\text{m}^{-2}$)	5,8

La technologie pour réaliser cette protection est imposée, il s'agit de la technologie SmartMOS 8 MV de Freescale.

Présentation de la technologie

La technologie SmartMOS 8 MV sera utilisée non seulement pour ce cas particulier, mais aussi pour toutes les réalisations sur silicium au cours de la thèse.

Cette technologie est de type SmartPower, permettant de réaliser à la fois des composants dédiés à la logique et des composants dédiés à la puissance. La bibliothèque offre un large choix de transistors MOS (NMOS et PMOS), de transistors bipolaires (NPN et PNP) et de diodes, pour une gamme de tensions de fonctionnement allant de quelques Volts à environ 60 Volts. Le minimum lithographique est de 0,18 μm et la tenue en tension est garantie jusqu'à 100 Volts. Les composants sont réalisés dans une épitaxie P faiblement dopée et isolés du substrat fortement dopé P par une couche enterrée N (NBL). Ils sont isolés les uns des autres par des tranchées profondes oxydées puis remplies de polysilicium. La NBL peut être polarisée de l'extérieur via une implantation N profonde. Des tranchées d'oxyde en surface permettent de séparer les régions entre lesquelles les écarts de potentiel en fonctionnement sont importants. Une dizaine d'implantations de types N et P sont disponibles. En jouant avec ces différentes implantations, on a la liberté de faire varier les profils de dopages et d'ajuster les tensions de claquage. Une implantation N permet de compenser l'épitaxie P jusqu'à la NBL de manière à s'approcher d'une épitaxie N faiblement dopée. Deux implantations de surface fortement dopées, respectivement N et P, permettent de prendre des contacts ohmiques. La surface des plots d'E/S est de 100*100 μm^2 , ce qui, dans beaucoup d'applications, donne l'ordre de grandeur acceptable pour une protection ESD. Pour des raisons de confidentialité industrielle, les dimensions spatiales des paramètres technologiques seront données dans une unité arbitraire, notée udm (unité de mesure).

Un exemple de composant, réalisé en technologie SmartMOS 8 MV, est donné sur la Figure 12. Il s'agit d'un transistor bipolaire NPN vertical. Sa jonction base-collecteur correspond à la jonction plane entre une implantation P^+ profonde et la couche enterrée N^+ . Le courant transite jusqu'au collecteur via l'implantation N^+ profonde, contactant la couche enterrée. L'émetteur est formé par l'implantation très fortement dopée N^{++} de surface, permettant en même temps de prendre un contact ohmique.

1.3. Spécificités de la physique des protections ESD

La conception d'une protection ESD nécessite au préalable de bien connaître son comportement physique et les spécificités qui y sont liées. Il est montré que la densité de courant atteint des niveaux extrêmement élevés, ne se retrouvant dans aucun autre domaine d'application, et que la température peut s'élever au point d'entraîner la dégradation du composant. Compte tenu de ces spécificités des méthodes de caractérisation appropriées ont dû être développées. Des exemples de méthodes utilisées pour la mesure de la réponse électrique et l'étude du fonctionnement interne sont présentés.

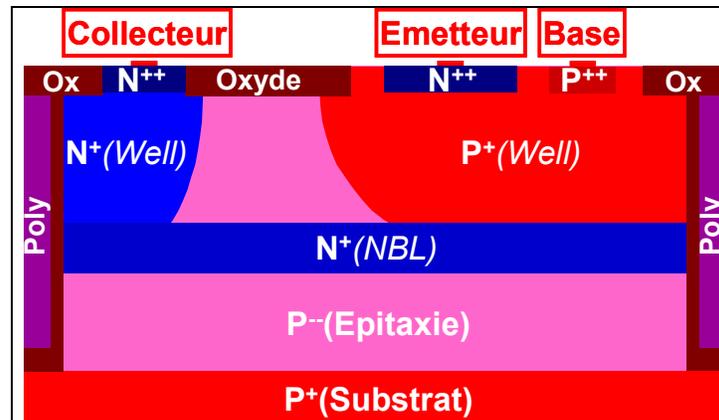


Figure 12: Transistor NPN vertical dans la technologie SmartMOS 8 MV de Freescale.

1.3.1. Etude du comportement physique d'une protection

Les particularités du comportement physique d'une protection ESD résultent à la fois des contraintes sur la surface de silicium, qui doit être la plus faible possible, et des durées très courtes des ESD, typiquement de l'ordre de la centaine de nanosecondes. La surface d'une protection n'est généralement limitée que par sa robustesse ESD, et non la dégradation de ses performances fonctionnelles. Or, sur la durée d'une ESD, une protection sur silicium peut supporter de très fortes puissances, de plusieurs centaines de Watts, bien au-delà de la limite en continu définie par l'hyperbole de dissipation. Dans ces conditions, l'on s'attend à des densités de courant très élevées et à une forte élévation de température.

1.3.1.1. Densités de courant extrêmement élevées

La surface d'une protection ESD conçue pour supporter plusieurs Ampères est typiquement de $100 \times 100 \mu\text{m}^2$, ce qui donne un courant par unité de surface de l'ordre de 10^4 A.cm^{-2} . Cette valeur est une borne inférieure pour la densité de courant atteinte dans la structure. Celle-ci n'est pas accessible en mesure, elle ne peut être évaluée que par une simulation bidimensionnelle du comportement électrique de la structure (Figure 13). Les niveaux de densités de courant relevées sont de l'ordre de 10^5 , voire 10^6 A.cm^{-2} . De telles valeurs sont extraordinairement élevées et ne se retrouvent dans aucun autre domaine d'application. Aussi, l'on peut s'attendre à des modes de fonctionnement se rencontrant uniquement dans des protections ESD. Ces effets des fortes densités de courant seront étudiés en détail au Chapitre2, pour le cas de transistors bipolaires autopolarisés.

A titre de comparaison, la densité de courant maximale dans un transistor de puissance tel qu'un VDMOS 45 Volts est de l'ordre de 150 A.cm^{-2} . Dans les applications RF, de fortes densités de courant peuvent être mises en jeu, cependant elles sont de l'ordre de 10^4 A.cm^{-2} , soit encore une à deux décades en dessous de celles mises en jeu dans une protection ESD.

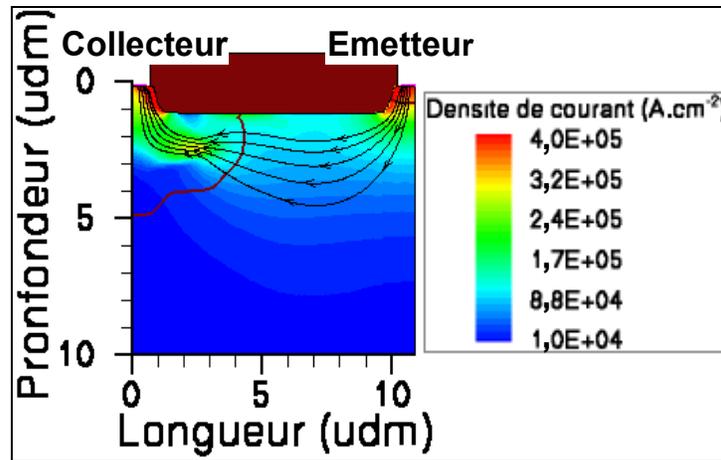


Figure 13: Distribution de densité de courant dans un transistor PNP latéral pour un courant surfacique de $3,2 \cdot 10^4 \text{ A.cm}^{-2}$, valeur typique lors d'une ESD.

1.3.1.2. Forte élévation de température

L'élévation de température dans une protection ESD résulte de l'autoéchauffement par effet Joule, proportionnel au produit de la densité de courant par le champ électrique. Or, si la protection est correctement optimisée, la robustesse est limitée par des dégradations d'origine thermique. Afin d'obtenir la plus petite surface possible tout en respectant la spécification en robustesse, la température s'approchera au plus près de la valeur à laquelle sont initialisées ces dégradations, sans jamais la dépasser. Dans ces conditions, l'élévation de température est très forte et influence le comportement électrique via la modification des paramètres électriques du silicium (durées de vie, mobilités, vitesses de saturation, coefficients d'ionisation par impact...). Ainsi, il apparaît un fort couplage électrothermique qu'il faudra prendre en compte dans l'étude des protections ESD.

Dans le cas d'une jonction en avalanche, les dégradations d'origine thermique sont induites par le second claquage ou claquage thermique. Celui-ci survient lorsque le courant des porteurs générés thermiquement est suffisant pour suppléer au courant d'avalanche. Il résulte une diminution de la valeur du champ électrique, contrôlant l'avalanche, et donc une diminution de la tension [17] [18]. L'apparition d'une résistance dynamique négative implique un comportement électrique instable, conduisant à la formation de filaments de courant à l'intérieur desquels la température augmente rapidement jusqu'à la dégradation du composant par fusion du silicium (à 1687 degrés Kelvins) [19]. La température au second claquage dépend de la géométrie et des niveaux de dopage du composant. L'ordre de grandeur, pour des dopages allant de 10^{15} et 10^{18} cm^{-3} , est compris entre 700 et 1250 degrés Kelvins.

Une autre particularité est le fonctionnement thermique quasi-adiabatique pendant la très courte durée d'une ESD. Une application numérique simple permet d'évaluer la longueur de diffusion de la chaleur l_T , donnée par :

$$l_T = \sqrt{D_T t_{ESD}} \quad (15)$$

Où D_T est le coefficient de diffusion thermique du silicium (égal à $0,09 \mu\text{m}^2.\text{ns}^{-1}$) et t_{ESD} la durée de la décharge. Pour une impulsion de 100ns, la longueur de diffusion est seulement de $3 \mu\text{m}$. Ainsi, la distribution de température est très inhomogène et présente généralement un "point chaud" au voisinage de la région où la génération de chaleur est maximale. Dans le cas d'un transistor bipolaire, le "point chaud" est localisé au niveau de la ZCE de la jonction base-collecteur en inverse, du fait de la conjonction d'un champ électrique et d'une densité de courant élevés (Figure 14). En conséquence, les effets thermiques seront plus ou moins intenses suivant qu'ils surviennent dans des régions voisines ou éloignées du "point chaud".

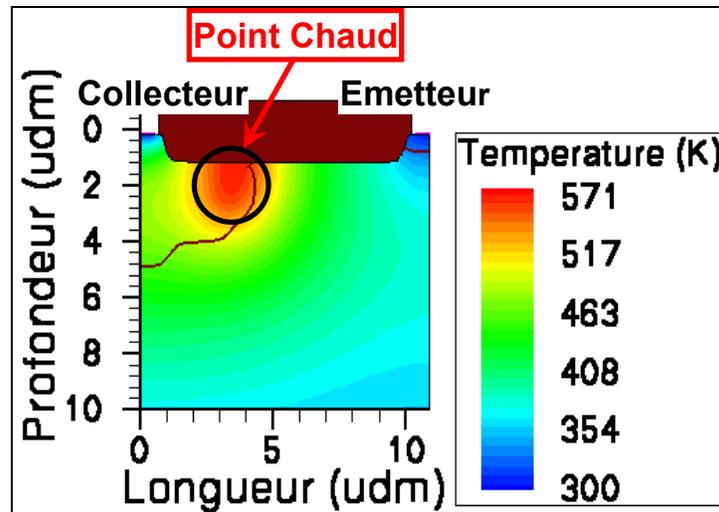


Figure 14: Distribution thermique dans un transistor PNP latéral à la fin d'une impulsion rectangulaire de 100 nanosecondes reproduisant les caractéristiques d'une ESD.

1.3.2. Méthodes de caractérisation

Des caractéristiques électriques I-V représentatives du fonctionnement au cours d'une ESD peuvent être mesurées par les méthodes TLP (Transmission Line Pulsing) et vfTLP (very fast TLP). Pour étudier le comportement interne d'une structure de protection, la méthode TIM (Transient Interferometric Mapping), permettant de cartographier la température au cours d'une ESD, est à ce jour la plus efficace.

1.3.2.1. Caractérisation TLP

L'acquisition des formes d'ondes temporelles du courant et de la tension lors de tests de robustesse pose des problèmes très ardues de précision, compte tenu des hautes fréquences (plusieurs GHz) et des fortes puissances [20]. De plus, les appareils de tests industriels ne sont pas adaptés pour permettre ce type de mesure. Aussi, des méthodes alternatives ont été développées pour caractériser le comportement électrique des protections ESD. La principale contrainte à respecter est de reproduire de manière réaliste le couplage électrothermique, la difficulté étant que la distribution thermique n'a pas le temps de se stabiliser pendant une ESD. En particulier, une mesure statique n'est pas envisageable, la protection serait détruite avant d'atteindre les courants requis, de plusieurs Ampères.

Actuellement, la méthode de caractérisation TLP [21] est la plus largement utilisée. Son principe consiste à soumettre le composant testé à des impulsions rectangulaires d'énergies comparables à celle d'une ESD. Pour chaque impulsion, le courant et la tension sont relevés en effectuant une moyenne sur les dernières nanosecondes, afin de reconstruire point par point une caractéristique électrique I-V. Ces mesures sont quasi-statiques, le courant et la tension à la fin d'une impulsion étant stabilisés et ne variant plus que sous l'effet de l'élévation de température. Afin que les résultats puissent être corrélés avec le test HBM, les durées des impulsions sont typiquement de 100 nanosecondes et le temps de montée compris entre 2 et 10 nanosecondes.

Le schéma de principe d'un banc de mesure TLP est représenté sur la Figure 15. Une impulsion de 100 nanosecondes est générée par la décharge d'une ligne de transmission d'impédance caractéristique 50 Ohms et longue de 10 mètres. Cette ligne est chargée à une tension V_{Charge} à travers une résistance de forte valeur, puis la commutation d'un relais permet sa décharge dans le circuit de test. Une résistance d'adaptation de 50 Ohms est placée en parallèle afin d'éviter les réflexions vers la ligne. Le courant dans le composant sous test (DUT) est contrôlé par une résistance en série (R_S)

typiquement de 500 Ohms, indispensable pour obtenir une mesure précise à bas courant. Enfin, une résistance en série (L_s) permet d'ajuster le temps de montée.

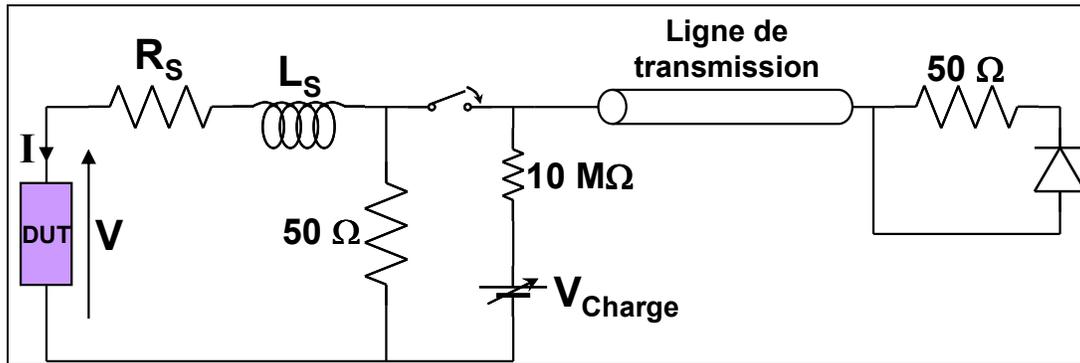


Figure 15: Schéma de principe d'un banc de mesure TLP.

Les protections développées seront systématiquement caractérisées en TLP. Les caractéristiques I-V obtenues permettront de vérifier si la fenêtre de spécification est respectée.

1.3.2.2. Caractérisation vfTLP

La méthode de caractérisation vfTLP est basée sur le même principe que la méthode TLP, à la différence que les impulsions générées reproduisent les spécificités d'une forme d'onde de type CDM [22]. Leurs durées sont typiquement comprises entre 1,2 et 10 nanosecondes, et leurs temps de montée entre 100 picosecondes et 2 nanosecondes. En particulier, la comparaison entre les caractéristiques TLP et vfTLP permettra d'évaluer si la dynamique de déclenchement est compatible avec les tests CDM. Il faut préciser que les caractérisations vfTLP donnent des informations sur le comportement d'une protection ESD vis-à-vis du test CDM, mais ne peuvent pas être directement corrélées avec les résultats de robustesse. En effet, lors d'un test CDM, le circuit chargé se décharge à travers une seule broche, alors qu'une impulsion vfTLP est appliquée entre deux broches.

1.3.2.3. Mesures TIM

La méthode TIM est un outil extrêmement puissant pour l'analyse du comportement interne d'une protection ESD. Elle permet d'acquérir des cartographies de la température avec une résolution temporelle de 5 nanosecondes et une résolution spatiale de 2 micromètres [23]. Ainsi, les points chauds peuvent être localisés précisément et il est possible de suivre leurs évolutions pendant une ESD, en particulier d'éventuels étalements ou déplacements [24] [25] [26]. Actuellement, les autres méthodes ne permettent pas d'avoir accès à des informations aussi précises. Elles donnent soit une moyenne sur plusieurs décharges, comme en mesure EMMI (EMission Microscopy), ou une simple évaluation qualitative.

L'idée consiste à exploiter les variations de l'indice de réfraction optique avec la température [27]. Pour cela, une mesure basée sur le principe de l'interférométrie holographique est mise en œuvre (Figure 16). La face arrière du composant testé est balayée par un faisceau laser. L'onde optique se propage à travers le silicium, se réfléchit sur la face avant (métallisations, oxydes...), puis retransverse le silicium en suivant le même chemin qu'à l'aller. Dans la zone active du composant, la phase de l'onde est affectée suivant les variations de l'indice optique associées à la température. Par interférence de l'onde réfléchie avec une onde de référence, ne transitant pas dans le silicium, les variations de phase sont transformées en des variations d'intensité lumineuse. Ensuite, une transformée de Fourier spatiale est effectuée sur l'image acquise pour remonter aux variations de phase, elles-mêmes proportionnelles aux variations de l'intégrale de la température sur le chemin de

l'onde optique. Enfin, il suffit de synchroniser cette mesure avec une impulsion TLP pour suivre l'évolution de la distribution thermique.

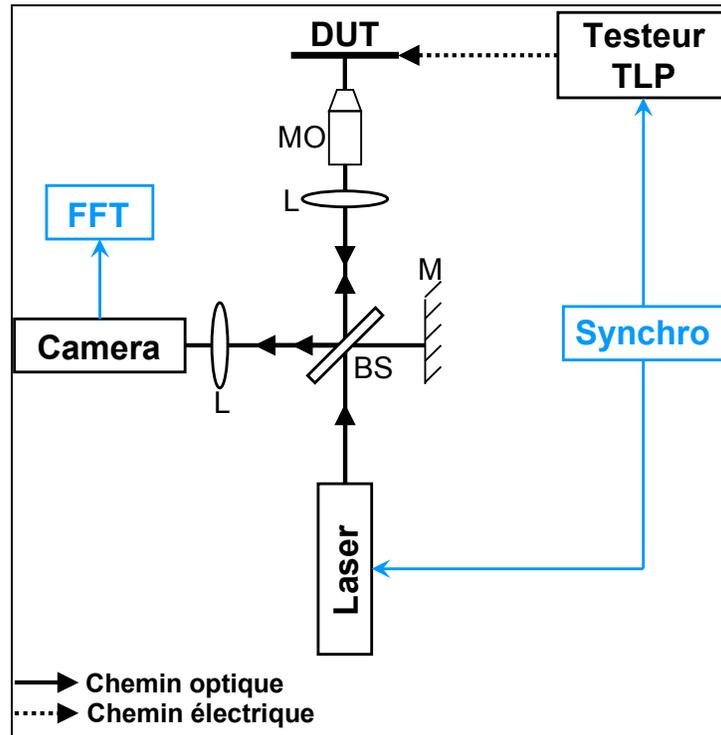


Figure 16: Schéma de principe de la méthode TIM (M : Miroir, L : Lentille, MO : Objectif de microscope, BS : Séparateur de faisceau).

1.4. Etat de l'art des protections ESD "hautes tensions" en technologies SmartPower

Pour évaluer l'intérêt industriel des protections qui sont développées, leurs performances doivent être comparées à celles des solutions existantes. De plus, la plupart de ces solutions étant brevetées, il faut pouvoir s'en démarquer, ce qui suppose de s'appuyer sur des idées très innovantes. Pour ces raisons, un état de l'art aussi complet que possible doit être établi. Tout d'abord, les deux principales stratégies de protection d'un circuit intégré, centralisée ou localisée, sont décrites. Ensuite, les structures de protections dédiées aux E/S "hautes tensions" en technologie SmartPower sont présentées. Pour finir, un cas particulier est traité, celui des E/S 80 Volts dans la technologie SmartMOS 8 MV de Freescale.

1.4.1. Principes de base des circuits de protection

Si les composants du circuit ne sont pas suffisamment robustes pour supporter eux-mêmes l'ESD, la stratégie de protection consiste à dévier le courant du plot d'E/S vers la masse, via le bus V_{SS} . Ainsi, le cœur du circuit intégré n'a pas à supporter l'énergie de la décharge. Deux stratégies sont disponibles suivant que plusieurs E/S sont protégées par la même structure, protection centralisée, ou qu'une structure est dédiée spécifiquement à une E/S, protection localisée. Suivant le cas, la structure

de protection ne fonctionnera pas selon le même principe, des méthodes de développement différentes devront être mises en œuvre.

1.4.1.1. Protections centralisées

Le schéma électrique d'une protection centralisée est représenté sur la Figure 17. La structure de protection P_{ESD} est placée entre les bus V_{DD} et V_{SS} , et l'E/S est reliée à chacun de ces bus par une diode, appelée diode d'isolation. Pendant le fonctionnement normal de l'application, les diodes d'isolation sont bloquées. Lors d'une ESD négative, la diode vis-à-vis de V_{SS} devient passante et évacue la décharge sans que la structure de protection soit impliquée. Si l'ESD est positive, la diode vis-à-vis de V_{DD} devient passante, ce qui permet à la décharge d'atteindre V_{SS} via le bus V_{DD} et la protection. L'établissement de ce chemin de conduction suppose un déclenchement actif de la protection sur une ESD, impliquant un circuit de commande capable de détecter une ESD.

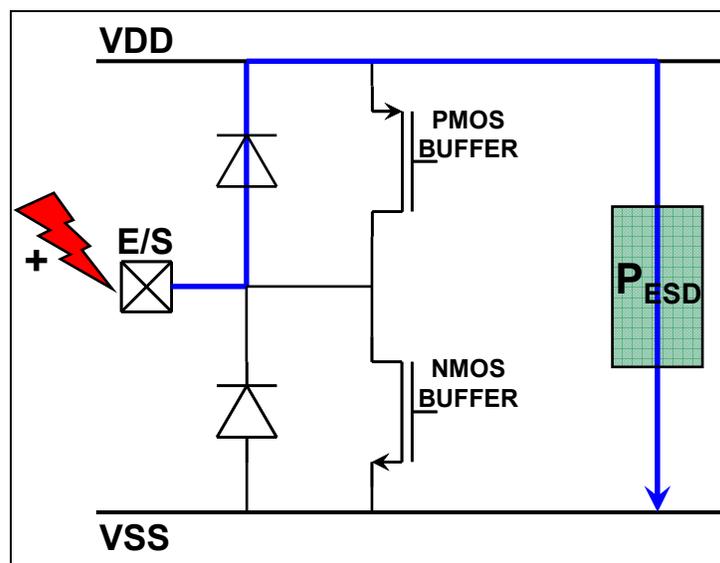


Figure 17: Schéma de principe d'une protection centralisée montrant le chemin de courant pour une ESD positive arrivant sur le plot d'E/S.

Cette stratégie de protection est basée sur une approche de type circuit, ce qui permet d'effectuer facilement des simulations, avec des simulateurs électriques de type Spice, et de transposer les solutions entre différents procédés technologiques. En mettant à profit ces avantages, les circuits intégrés peuvent passer les spécifications ESD dès les premières qualifications sur silicium. Toutefois, cette approche n'est efficace que si un grand nombre d'E/S présentent les mêmes spécifications. Ainsi, elle est bien adaptée aux technologies CMOS, pour lesquelles elle donne d'excellents résultats. Par contre, son utilisation sera plus limitée en technologies SmartPower. De plus, la chute de potentiel dans les résistances de la diode en direct et du bus V_{DD} réduit la fenêtre de conception, alors que ce paramètre est très critique pour les protections "hautes tensions". Au final, les surfaces requises sont le plus souvent réhivitoires au regard des spécifications visées.

1.4.1.2. Protections localisées

L'alternative aux protections centralisées est basée sur des protections localisées au niveau des plots des E/S à protéger, suivant le schéma présenté sur la Figure 18. Dans ce cas, la structure de protection P_{ESD} est placée directement entre l'E/S et le bus V_{SS} . Cette structure doit être conçue pour rester bloquée tant que le circuit est en fonctionnement normal, et pour commuter à l'état passant pendant

une ESD, lorsque la tension passe au-dessus de V_{DD} ou au-dessous de V_{SS} . La plupart des structures usuelles sont déclenchées par le claquage d'une jonction inverse, dans le cas d'une décharge positive, et par la polarisation d'une jonction en direct, dans le cas d'une décharge négative. Le déclenchement par claquage s'accompagne le plus souvent d'un repliement, lié au fonctionnement intrinsèque de la structure. Aussi, les protections localisées sont également appelées protections à repliement.

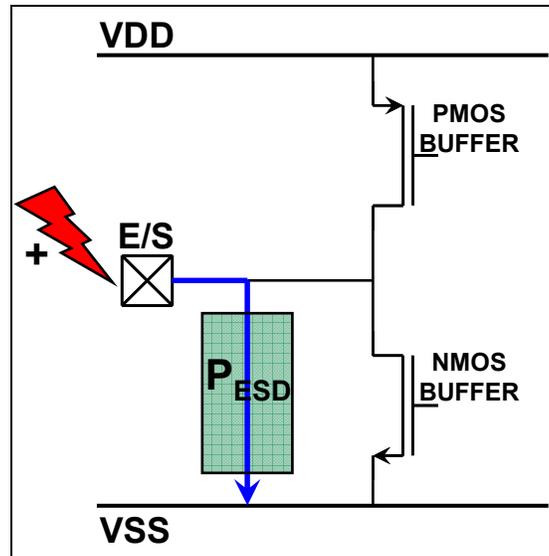


Figure 18: Schéma de principe d'une protection localisée montrant le chemin de courant pour une ESD positive survenant sur le plot d'E/S.

Le comportement électrothermique interne doit être simulé pour pouvoir évaluer le fonctionnement d'une protection localisée. Ce type de simulation est très lourd en temps de calcul et d'un calibrage délicat. De plus, le travail doit être refait pour chaque nouvelle technologie, entre autres, parce que les tensions de claquage dépendent des profils de dopage. En conséquence, le développement d'une protection localisée est plus ardu et plus long que celui d'une protection centralisée. Par contre, nous disposons de plus de degrés de liberté dans l'utilisation de la technologie, ce qui permet d'obtenir des protections plus efficaces, en particulier de plus petites tailles. Si la protection est suffisamment petite pour tenir sous un plot d'E/S, aucune augmentation de la surface totale de la puce ne sera nécessaire.

1.4.2. Structures de protection intégrées sur silicium

L'état de l'art proposé des structures de protection ESD concerne spécifiquement celles utilisées pour les E/S "hautes tensions" en technologie SmartPower. Ces structures ont été classées suivant trois grands ensembles : les transistors bipolaires autopolarisés, les structures PNP et les transistors à effet de champ [28].

Les diodes à avalanche en inverse, qui semblent intéressantes en raison de leurs tensions de claquage élevées et de l'absence de repliement, ne sont pas utilisées comme protections intégrées en technologie SmartPower, du fait d'un comportement à l'état passant trop résistif. Toutefois, elles sont parfois mises en œuvre en tant que protections discrètes dans des applications très énergétiques, cas des diodes "foudres" ou des diodes TRANSIL par exemple.

1.4.2.1. Transistors bipolaires autopolarisés

Le principe de fonctionnement d'un transistor bipolaire est décrit, en faisant abstraction des effets spécifiques aux protections ESD (fort courant et température élevée), qui sont étudiés en détail au chapitre 2. Ensuite, les méthodes mises en œuvre pour adapter ces composants à la protection des E/S "hautes tensions" sont présentées. Seul le transistor NPN est considéré, celui-ci étant généralement préféré au transistor PNP.

Principe de fonctionnement

En mode autopolarisé, l'émetteur et la base du transistor bipolaire sont reliés, soit directement par un court-circuit soit à travers une résistance. Ainsi, le transistor fonctionne en dipôle, en particulier le courant de polarisation ne peut plus être amené de l'extérieur via le contact de base. Le schéma électrique est donné Figure 19(a), où la résistance de base R_B correspond à la résistance d'accès interne, plus une éventuelle résistance externe.

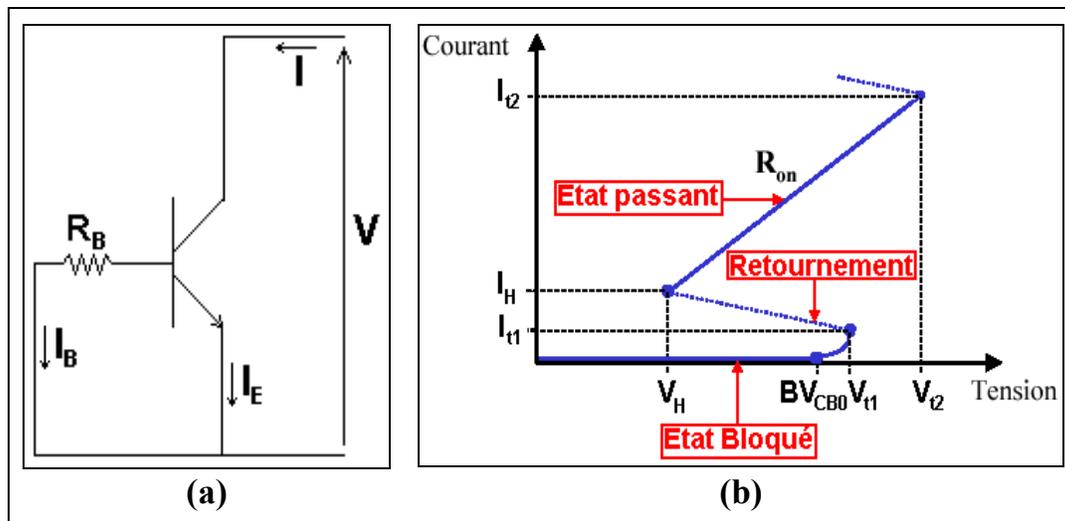


Figure 19: Schéma électrique d'un transistor bipolaire autopolarisé (a) et sa caractéristique TLP (b).

La caractéristique électrique à fort courant, de type TLP, d'un transistor bipolaire autopolarisé est représentée sur la Figure 19(b). La jonction base-collecteur étant en inverse, le transistor reste à l'état bloqué tant que la tension de claquage par avalanche de cette jonction BV_{CB0} n'est pas atteinte. De manière simplifiée, le claquage par avalanche survient lorsque le maximum du champ électrique atteint une valeur critique E_C . Les porteurs transitant dans la zone dépeuplée acquièrent alors suffisamment d'énergie pour que la génération de paires électron-trou par ionisation par impact s'emballent. Il en résulte une rapide augmentation du courant. Ce courant transite à travers la résistance de base R_B , créant une différence de potentiel ayant tendance à polariser la jonction émetteur-base en direct. Lorsque cette différence de potentiel devient suffisante pour que l'injection d'émetteur soit significative, typiquement à des valeurs de l'ordre de 0,5 Volts, la tension aux bornes du transistor diminue. Ce phénomène, appelé repliement, est dû à la contribution à l'avalanche des électrons injectés, lorsque ceux-ci traversent la zone dépeuplée. Plus précisément, cette contribution permet à la probabilité qu'un porteur génère une paire électron-trou de diminuer, et donc au champ électrique de diminuer, tout en maintenant le même courant d'avalanche, nécessaire à la polarisation en direct de la jonction émetteur. Au point I-V où est initialisé le repliement de la caractéristique électrique, la terminologie adoptée est courant de déclenchement et tension de déclenchement, notés respectivement I_{t1} et V_{t1} . Ce repliement est un phénomène instable, c'est-à-dire que les points I-V correspondants ne sont pas accessibles par la mesure TLP. Au point où le comportement se stabilise, le courant et la tension sont appelés courant de maintien et tension de maintien, notés respectivement

I_H et V_H . Le transistor se trouve alors dans un nouveau mode de polarisation, équivalent à une configuration base flottante, comme nous le démontrerons au chapitre 2. Le fonctionnement dans ce mode de polarisation est schématisé Figure 20. Il est principalement caractérisé par la résistance dynamique R_{ON} , et est limité à fort courant par la destruction du transistor, typiquement par second claquage thermique au point $(I_{t2}; V_{t2})$. Suivant cette description, la caractéristique électrique I-V d'un transistor bipolaire autopolarisé résulte du couplage entre deux phénomènes de multiplication du courant, l'effet bipolaire et l'avalanche.

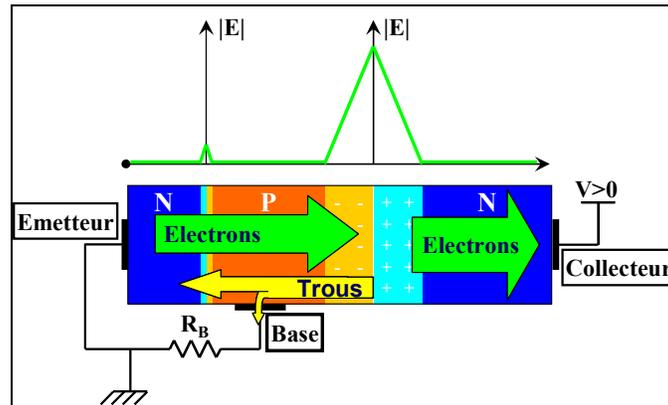


Figure 20: Représentation du champ électrique et des flux de porteurs dans un transistor bipolaire autopolarisé à l'état passant.

Structures latérales

Une structure est dite latérale (Figure 21) lorsque le flux de courant dans la base est latéral. L'avantage de cette configuration est de permettre un réglage de la tension de déclenchement V_{t1} et de la tension de maintien V_H , via les paramètres géométriques de la structure [29]. Pour régler V_{t1} , il faut ajuster la distance $D1$ entre les bords des masques des implantations P^+ de base et N^+ de collecteur, de manière à obtenir une région faiblement dopée P^- plus ou moins étendue. En contrôlant ainsi le dopage au voisinage de la jonction base-collecteur, on contrôle la tension de claquage BV_{CB0} , dont V_{t1} est proche. Pour régler V_H , il faut ajuster la distance $D2$ entre les bords des masques des implantations N^{++} d'émetteur et P^+ de base, ce qui revient à contrôler la largeur de base interne. Un élargissement de la base interne permet d'augmenter V_H , via la diminution du gain, comme il sera montré au chapitre 2. En contrepartie, les propriétés des structures latérales vis-à-vis de la robustesse sont généralement médiocres. Le courant circulant en surface, d'une part les densités de courant correspondantes ont tendance à être élevées et d'autre part, les oxydes de surface ayant une faible conductivité thermique ($0,014 \text{ W.cm}^{-1}.\text{K}^{-1}$ pour le SiO_2 , à comparer à $1,5 \text{ W.cm}^{-1}.\text{K}^{-1}$ pour le Si), la chaleur reste dans le silicium, ce qui a pour effet d'augmenter la température du point chaud. En conséquence, les structures latérales n'ont des robustesses acceptables que dans les cas des faibles tensions de maintien, pour lesquelles la puissance dissipée reste limitée. A des tensions de maintien élevées, le bas niveau de robustesse est en général réhibitoire pour une protection ESD.

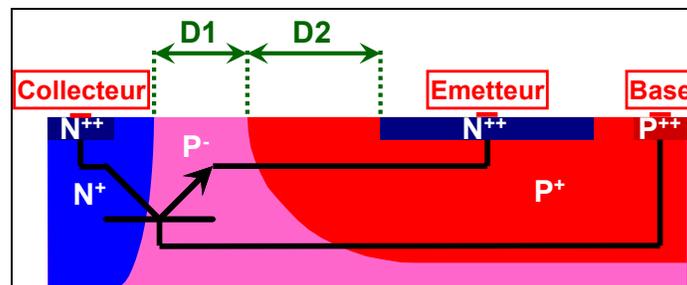


Figure 21: Coupe technologique simplifiée d'un transistor NPN latéral.

Structures verticales

Une structure est dite verticale (Figure 22) lorsque le flux de courant dans la base est vertical. Pour y parvenir, le collecteur est formé par une couche enterrée N placée sous la diffusion d'émetteur. Dans cette configuration, le profil de dopage vertical sous l'émetteur est uniforme et la largeur de base W_B constante, ce qui conduit à une répartition du courant très uniforme. De plus, le point chaud est situé en profondeur, dans le volume du silicium, ce qui est optimal pour la diffusion de la chaleur. Ces deux propriétés permettent d'obtenir des robustesses très élevées avec des structures verticales.

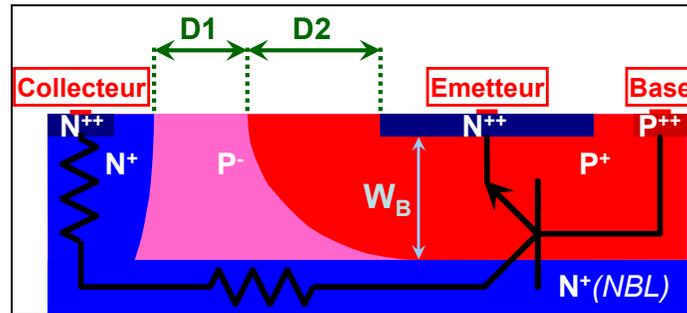


Figure 22: Coupe technologique simplifiée d'un transistor NPN vertical.

La tension de déclenchement V_{t1} d'une structure verticale peut être ajustée par un claquage latéral [30] [31]. Les trous générés par avalanche à la jonction latérale sont collectés au contact de base, et polarisent ainsi la diode émetteur-base en direct. Le flux d'électrons injectés par l'émetteur transitera dans le transistor vertical, pourvu que son gain soit meilleur que le gain du transistor latéral. Pour garantir cette condition, il suffit d'élargir la base de ce dernier, accessible avec le paramètre D2 (Figure 22). Ainsi, la structure va se replier sur la tension de maintien du transistor vertical. Si celle-ci est inférieure à la tension de claquage latéral, seul le chemin de courant vertical persiste, assurant une robustesse élevée. Le réglage de la tension de claquage latérale peut être effectué soit par l'ajustement de la distance D1 (Figure 22), selon le principe décrit pour une structure latérale, soit par l'insertion d'une diffusion de surface fortement dopée au voisinage de la jonction (Figure 23). La seconde méthode est en général privilégiée, car elle permet un réglage plus précis, les diffusions de surface étant bien contrôlées en technologie. Elle n'est pas utilisable dans le cas d'une structure latérale, où elle entraînerait une très forte concentration du courant en surface, réductrice vis-à-vis de la robustesse.

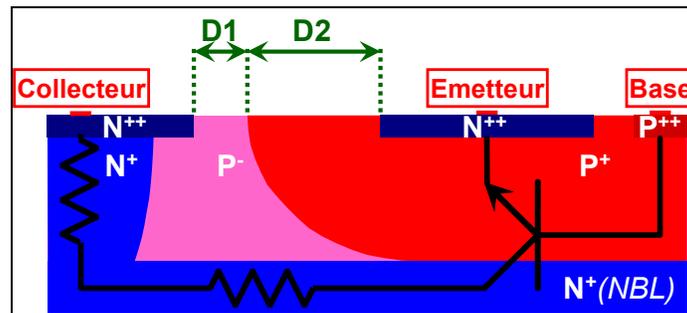


Figure 23: Coupe technologique d'un transistor NPN vertical déclenché par un claquage latéral.

Par contre, le réglage de la tension de maintien V_H est beaucoup plus problématique. Celle-ci dépend de caractéristiques technologiques telles que l'épaisseur d'épitaxie, déterminant la largeur de base interne, et le profil du dopage de la couche enterrée. Ainsi, la tension de maintien est une constante pour une technologie donnée, et sa valeur peut être particulièrement faible si cette technologie a été développée sans prendre en compte les problématiques liées aux ESD. Une méthode

pour contourner cette difficulté consiste à supprimer une partie de la couche enterrée afin d'obtenir une séparation latérale avec l'émetteur (Figure 24), ce qui permet à nouveau d'ajuster la largeur de base interne. Cette méthode est très efficace, une augmentation de V_H de 20 Volts à 40 Volts est reportée dans la publication [32]. La structure n'est plus rigoureusement verticale, néanmoins la robustesse reste élevée. En contrepartie, la structure n'est plus isolée du substrat, d'où des risques d'injection de porteurs minoritaires et de latch-up. Aussi, il n'est pas toujours possible de mettre en pratique cette solution. Pour garder la couche enterrée sous toute la structure, une seconde méthode consiste à la diluer sous l'émetteur [32] (Figure 25). Pour cette dilution, le masque d'implantation est formé de petites ouvertures les unes à côté des autres, permettant de réduire la dose implantée. Les diffusions lors des cycles thermiques dans la suite du procédé technologique assurent l'homogénéisation du dopage sur la surface correspondante. Cette méthode revient également à augmenter la largeur de base effective, la génération du courant de polarisation par avalanche étant plus intense dans la région non diluée que dans la région diluée. L'augmentation de V_H est cependant moins efficace qu'avec la première méthode, mais n'en reste pas moins très significative (de 20 Volts à 30 Volts dans la publication [32]).

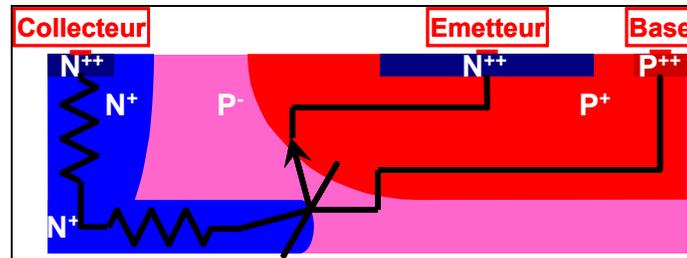


Figure 24: Coupe technologique d'un transistor NPN vertical avec couches enterrées partielle.

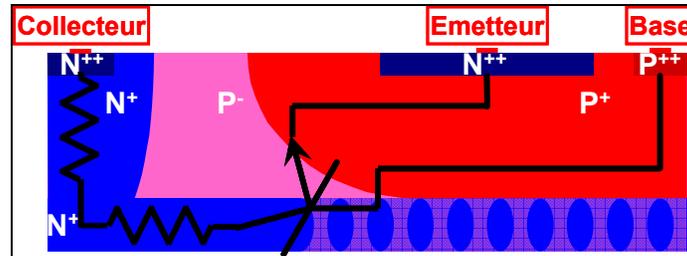


Figure 25: Coupe technologique d'un transistor NPN vertical avec couches enterrées diluée.

Structures en série

Pour augmenter V_H , les deux méthodes basées sur la modification de la couche enterrée ont été brevetées par une entreprise concurrente de Freescale (Infineon) [33] [34]. La seule possibilité restante est de mettre en série les transistors (Figure 26(a)), et ainsi additionner les tensions. Cependant, les surfaces et les résistances passantes R_{ON} sont également additionnées [35] (Figure 26(b)). Pour préserver le même R_{ON} qu'une structure seule, il faut multiplier la surface par le carré du nombre de protections en série. Le coût d'une telle augmentation de la surface peut être pénalisant pour la compétitivité d'un produit, de sorte qu'il existe une forte demande pour le développement de solutions alternatives.

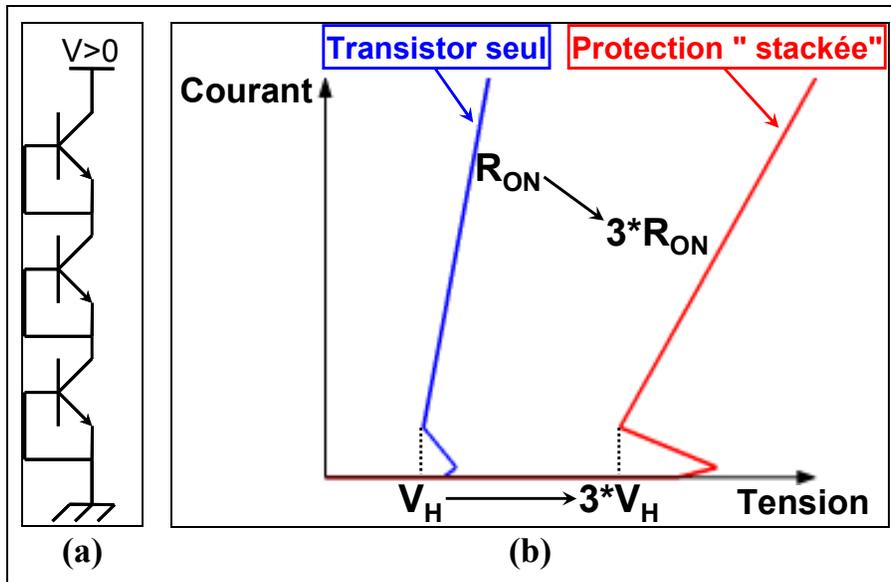


Figure 26: Schéma électrique de trois transistors NPN en série (a), et caractéristique I-V résultante (b).

Focalisation du courant

Malgré l'uniformité du dopage sous l'émetteur, le courant peut se focaliser sous l'effet d'instabilités dues à une résistance dynamique négative aux bornes de la ZCE base-collecteur [36]. Cette condition est remplie pendant le repliement et éventuellement au cours de l'état passant. La focalisation du courant dégrade la robustesse, l'élévation de température étant plus forte que si la conduction était uniforme. La solution classique pour uniformiser le courant consiste à augmenter la valeur des résistances de ballast dans l'émetteur et le collecteur [19]. Cependant, de fortes résistances de ballast ne garantissent pas toujours une conduction uniforme. Dans ce cas, seul un déplacement du filament de courant permet d'obtenir une température relativement uniforme sur toute la longueur du composant, et ainsi de limiter la dégradation de la robustesse [37]. Pour favoriser ce déplacement, la règle fondamentale consiste à élargir la base interne afin d'atténuer le couplage entre le courant d'émetteur et la génération pour avalanche. Cette règle ne pourra être appliquée que sur des structures pseudo-verticales, comme celles à couches enterrées partielles ou diluées, et non sur des structures rigoureusement verticales, pour lesquelles la base interne est fixée par la technologie. Une seconde règle consiste à augmenter le dopage de base afin d'homogénéiser la tension émetteur-base.

1.4.2.2. Structures PNPN

Une structure PNPN correspond à un transistor NPN et un transistor PNP imbriqués de telle sorte que le collecteur du NPN soit la base du PNP et le collecteur du PNP soit la base du NPN (Figure 27). Dans cette configuration, le courant de collecteur du transistor NPN est vu par le transistor PNP comme un courant de polarisation.

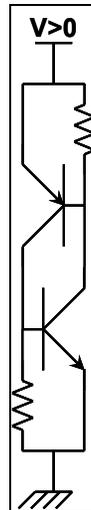


Figure 27: Schéma électrique équivalent d'une structure PNPN.

Principe de fonctionnement

La caractéristique électrique d'une structure PNPN est similaire à celle d'un transistor bipolaire autopolarisé. Elle montre un état bloqué et un état passant avec un repliement en tension lors du déclenchement. A l'état bloqué, la tension est soutenue à la jonction base-collecteur en inverse commune aux deux transistors. Lorsque la tension de claquage de cette jonction est atteinte, le courant généré polarise les deux transistors NPN et PNP, suivant le même principe que pour un transistor bipolaire autopolarisé. Toutefois, le repliement n'est pas dû seulement à la participation à l'avalanche des courants d'émetteur, mais aussi à la contribution réciproque aux courants de polarisation entre les transistors. Deux types de fonctionnement devront être différenciés suivant la valeur de la somme des gains en base commune α_{NPN} et α_{PNP} (ou du produit des gains en base commune β_{NPN} et β_{PNP}) [38]. Si $\alpha_{\text{NPN}} + \alpha_{\text{PNP}} < 1$ (ou $\beta_{\text{NPN}}\beta_{\text{PNP}} < 1$), la contribution réciproque aux courants de polarisation n'est pas suffisante pour maintenir la structure à l'état passant sans l'apport d'un courant d'avalanche. La tension de maintien est alors fixée par le champ électrique nécessaire à la génération de ce courant d'avalanche. Dans ce cas, on parle de deux transistors, NPN et PNP, couplés. Si $\alpha_{\text{NPN}} + \alpha_{\text{PNP}} > 1$ (ou $\beta_{\text{NPN}}\beta_{\text{PNP}} > 1$), l'état passant nécessite uniquement la contribution réciproque aux courants de polarisation. Le champ électrique à la jonction en inverse s'annule, conduisant à une tension de maintien très faible, de l'ordre du Volt, égale au potentiel des deux jonctions émetteur-base en direct. Dans ce cas, le fonctionnement est identique à celui d'un thyristor.

Transistors bipolaires NPN et PNP couplés

Une tension de maintien identique à celle d'un transistor seul peut être préservée, en particulier sa valeur reste compatible avec les protections "hautes tensions" [39] [40]. D'une part, cette structure est mise à profit pour améliorer la robustesse d'un transistor NPN vertical, en lui couplant un transistor PNP latéral (Figure 28) [39]. Le dédoublement des chemins de courant conduit à deux points chauds de températures inférieures à celle du point chaud d'un transistor seul. De plus, le courant du PNP s'oppose à la focalisation du courant du NPN, ce qui permet d'obtenir une très bonne uniformité de la conduction. Enfin, le couplage peut être optimisé pour obtenir de très faibles résistances à l'état passant R_{ON} [40]. Ces stratégies sont particulièrement intéressantes pour la protection des E/S "hautes tensions". Cependant, elles sont brevetées par des compagnies concurrentes de Freescale [41] [42] (respectivement Infineon et AMI Semiconductor).

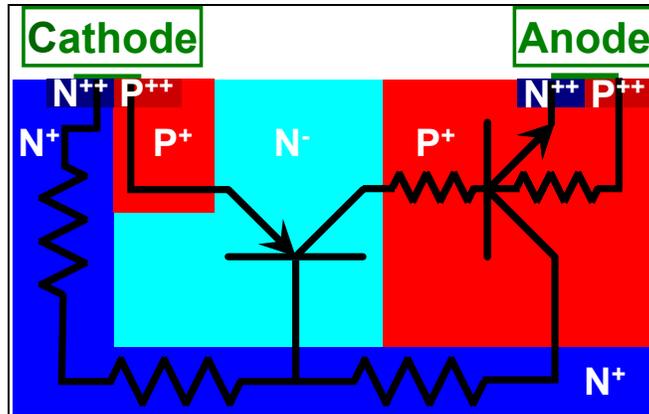


Figure 28: Coupe technologique simplifiée d'un transistor NPN vertical couplé à un transistor PNP latéral.

Thyristors

Malgré la tension de maintien de l'ordre d'un Volt, les structures PNPN fonctionnant en thyristor peuvent être utilisées comme protections des E/S "hautes tensions". Pour y parvenir, des stratégies doivent être mises en œuvre afin d'éviter le latch-up de la protection lorsque le produit est en fonctionnement. Les deux principales méthodes publiées reposent, pour l'une, sur un circuit forçant la structure à l'état bloqué si l'application est alimentée [43], et, pour l'autre, sur la mise à profit de la limitation en courant de l'alimentation [43] [44].

Pour garantir qu'une structure PNPN montée entre un plot d'E/S et le bus de masse V_{SS} ne parte pas en latch-up, la stratégie consiste à connecter la base du transistor PNP au bus d'alimentation V_{DD} (Figure 29). En fonctionnement normal, la tension sur l'E/S est inférieure à la tension d'alimentation V_{DD} , de sorte que la jonction émetteur-base du transistor PNP est polarisée en inverse. Dans cette configuration, même si une perturbation transitoire entraîne une élévation de tension sur l'E/S suffisante pour déclencher le thyristor, celui-ci se coupe dès que la tension redevient inférieure à V_{DD} . Ainsi, la protection ne peut pas être polarisée à l'état passant en continu par l'alimentation. Lors d'un test ESD entre l'E/S et le bus V_{SS} , le bus V_{DD} sera flottant, ce qui permet le repliement du thyristor. Compte tenu de la très faible tension de maintien, il est aisé d'obtenir une robustesse élevée et un faible R_{ON} (1.2.3.2). Cette structure présente également d'autres avantages. La tension de déclenchement V_{th} peut être ajustée à l'aide de la résistance $R1$, en particulier à des valeurs inférieures à la tension de claquage de la jonction en inverse. La stratégie de protection est indépendante de la tension d'alimentation, et donc facilement transposable entre différentes applications ou différentes technologies. Le brevet de cette protection [45] est détenu par Sarnoff Corporation.

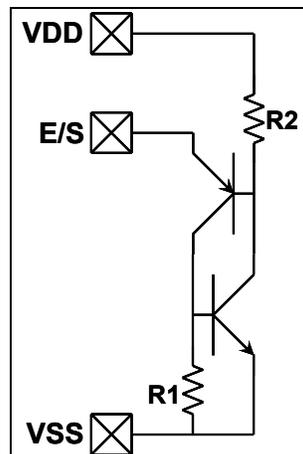


Figure 29: Polarisation d'une structure PNPN la maintenant à l'état bloqué si le circuit est alimenté.

Si le courant après repliement ne peut pas être fourni par l'alimentation, il n'y a aucun risque de latch-up, même si la tension est inférieure à la tension d'alimentation. Cette considération permet d'agrandir la fenêtre de conception en incluant une nouvelle région à basse tension et fort courant (Figure 30(a)). Pour en tirer parti, une stratégie consiste à adjoindre à la structure PNPN un composant externe qui assurera la conduction à bas courant avant de passer le relais au thyristor. Ce composant est placé en parallèle avec le transistor NPN et est relié à l'émetteur du transistor PNP via une résistance externe R_{Ext} (Figure 30(b)). Ce circuit comprend un second chemin de décharge, qui sera activé en premier si le composant externe se déclenche avant le thyristor. Ensuite, la chute de potentiel à travers la résistance R_{Ext} permettra de polariser en direct la jonction émetteur-base du transistor PNP et ainsi déclencher le thyristor. En contrôlant la valeur de cette résistance R_{Ext} , le courant après repliement du thyristor peut être ajusté au-dessus du maximum que peut débiter l'alimentation. Des protections utilisant un PMOS [43] ou un ggNMOS (gate-grounded NMOS) [44] comme composant externe ont été publiées. Le PMOS semble être mieux adapté aux protections "hautes tensions", car il peut être déclenché sans repliement à des tensions élevées, contrairement au ggNMOS qui se replie à des tensions faibles. Enfin, il faut souligner que l'opportunité de cette stratégie de protection dépend du type d'application visée. Notamment, elle n'est pas applicable en automobile, où l'alimentation provient de la batterie, capable de fournir des pics de courant d'intensités bien supérieures à celles des ESD.

1.4.2.3. Transistors à effet de champ

En tant que protections ESD, les transistors à effet de champ offrent deux modes de fonctionnement. D'une part, l'ESD peut être conduite par le transistor bipolaire parasite, dont l'émetteur, la base et le collecteur correspondent respectivement à la source, au substrat et au drain (Figure 31). D'autre part, l'effet MOS peut être mis à profit, avec une conduction de l'ESD dans le canal sous la grille.

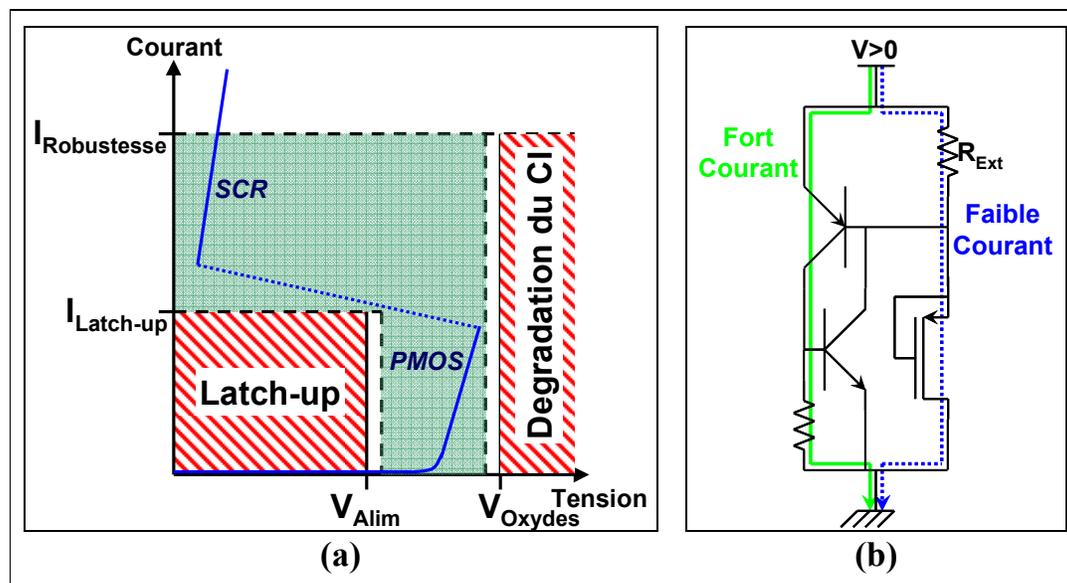


Figure 30: Fenêtre de conception intégrant la limite en courant de l'alimentation ($I_{Latch-up}$) et caractéristique électrique vérifiant cette spécification (a) pour le cas d'une protection dont la conduction est assurée successivement par un transistor PMOS et un thyristor (SCR) (b).

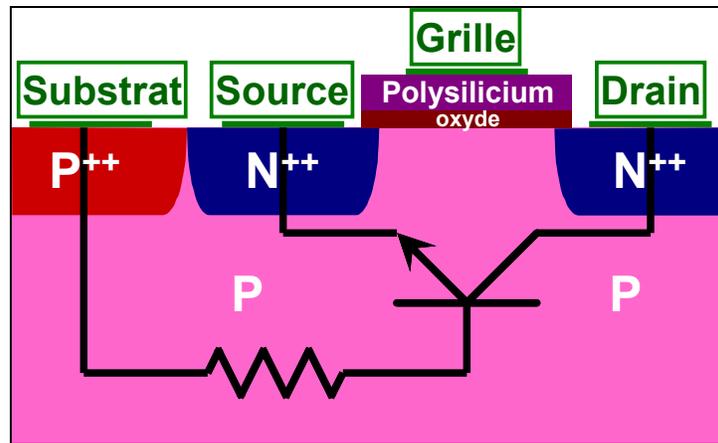


Figure 31: Coupe technologique simplifiée d'un transistor NMOS montrant le transistor NPN parasite.

Bipolaires parasites

Dans une protection ESD basée sur le transistor bipolaire parasite, les contacts de source, de substrat et de grille sont reliés à la masse pour un NMOS (gate-grounded NMOS) ou au potentiel positif pour un PMOS (gate-up PMOS) (Figure 32). Ainsi, le transistor MOS restera bloqué, un canal sous la grille ne pouvant pas se former dans ces conditions de polarisation, et le transistor bipolaire se retrouvera en mode autopolarisé.

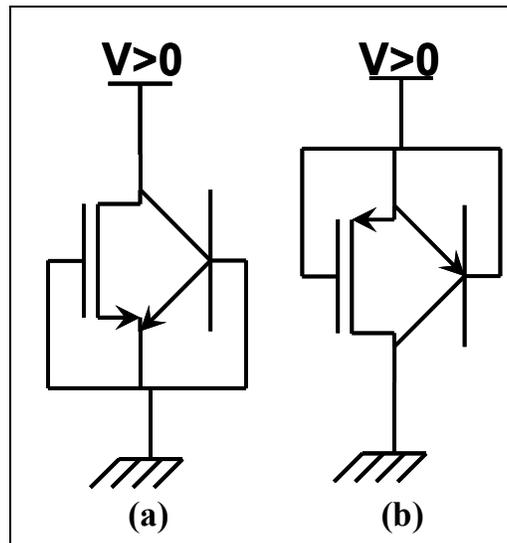


Figure 32: Schémas électriques équivalents d'un ggNMOS (a) et d'un guPMOS (b).

Les problématiques sont donc les mêmes que celles exposées au paragraphe 1.4.2.1 pour les transistors bipolaires autopolarisés latéraux. L'énergie dissipée ne pourra être supportée que si les tensions de maintien sont modérées. En technologies HV CMOS (High Voltage CMOS), elles sont typiquement comprises entre 5 et 10 Volts [24] [46]. Pour protéger des E/S "hautes tensions", à plusieurs dizaines de Volts, il s'avère donc indispensable de mettre en série plusieurs composants [46]. Par ailleurs, certaines règles doivent être respectées pour que le composant ne casse pas dès le repliement. La résistance dynamique négative entraînant une focalisation du courant, il faut garantir que la densité de courant dans le filament puisse être supportée. D'une part, les résistances de ballast dans l'émetteur (source) et le drain (collecteur) doivent être élevées, pour s'opposer aux instabilités induites par la résistance dynamique négative, et, d'autre part, la résistance de base (substrat) doit être

suffisante pour uniformiser la polarisation de la jonction émetteur-base. Après repliement, le filament s'étend à densité de courant constante, de sorte que, si le composant arrive à la tension de maintien sans être détruit, une conduction uniforme et donc une bonne robustesse seront atteintes [19]. A titre de remarque, ces considérations expliquent la mauvaise tenue en ESD des technologies CMOS avancées, où les siliciures déposés en surface réduisent les résistances de ballast, et où le substrat fortement dopé est très peu résistif [47].

Effet MOS

L'intérêt des protections basées sur l'effet MOS réside dans la possibilité d'un contrôle actif, via un circuit polarisant la grille lors d'une ESD. Cette propriété permet d'en tirer parti pour développer des protections centralisées. Par contre, le canal est généralement trop résistif pour que le R_{ON} soit acceptable dans le cas d'une protection localisée. Pour les applications "hautes tensions", des transistors de type DMOS (Double diffused MOS) sont utilisés, leurs dopages de drain étant optimisés afin de tenir des tensions élevées. Concernant le circuit de polarisation, il doit maintenir le MOS à l'état passant pendant toute la durée de l'ESD, tout en assurant qu'il reste bloqué en fonctionnement normal. De plus, la tension doit rester inférieure à la tension de déclenchement du transistor bipolaire parasite. S'il se déclenchait, la focalisation du courant suite au repliement entraînerait sa destruction immédiate. Etant donné le niveau de tension élevé, l'énergie dissipée dans le filament de courant serait au-delà de la limite supportable par le composant.

Le contrôle de la polarisation de grille peut être réalisé en la reliant au drain via une chaîne de diodes en inverse (Figure 33) [48]. Lorsque la tension sur le drain dépasse la tension de claquage de ces diodes, le potentiel de la grille augmente, entraînant la formation du canal et le déclenchement du MOS. Pour uniformiser la polarisation, il est recommandé d'insérer une résistance de quelques Ohms entre la grille et la source. Accessoirement, une diode en inverse entre la grille et la source limite la tension sur l'oxyde, afin de le protéger.

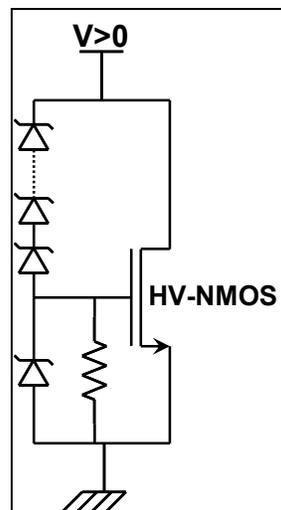


Figure 33: Circuit de polarisation de la grille d'un transistor NMOS à base de diodes en inverse.

Une autre option consiste à déclencher le MOS par effet capacitif, en insérant une capacité C_{HV} et une résistance R respectivement entre le drain et la grille, et entre la grille et la source (Figure 34(a)). Lors d'une ESD, la capacité doit être passante afin que la chute de potentiel dans la résistance polarise la grille. Cette condition sera garantie si la constante de temps RC_{HV} est supérieure à la durée d'une ESD. Afin d'assurer également que le transistor soit bloqué lorsque l'alimentation est stabilisée à sa tension nominale, la constante de temps sera de l'ordre d'une microseconde. En technologies HV CMOS, les valeurs typiques de la capacité C_{HV} et de la résistance R sont respectivement de quelques picoFarads et de quelques centaines de kiloOhms. Du point de vue de la surface totale de la

protection, une capacité de cette valeur est pénalisante. Etant donné les épaisseurs d'oxyde requises pour tenir les "hautes tensions", les capacités surfaciques des technologies HV CMOS sont très faibles. Pour limiter la surface du circuit de polarisation, le courant dans la capacité est multiplié à l'aide d'un miroir de courant (Figure 34(b)) [46]. Ce montage ne nécessite que des transistors, MN0 et MN1 de petites dimensions, et permet de diviser la surface de la capacité par le facteur de multiplication.

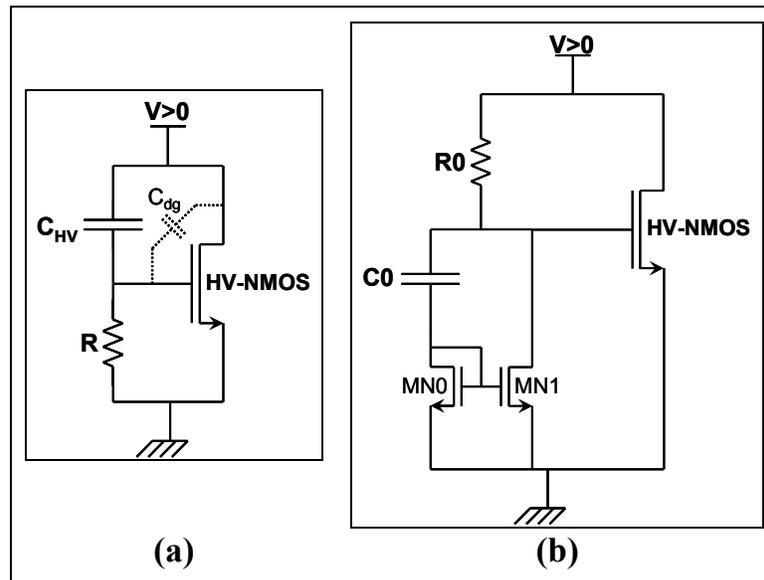


Figure 34: Circuit de polarisation de la grille d'un transistor NMOS par couplage capacitif (a) et amélioration basée sur un miroir de courant (b).

1.4.3. Protection des E/S 80 Volts dans la technologie SmartMOS 8 MV de Freescale

Un des principaux défis est le développement d'une protection répondant à la spécification des E/S 80 Volts d'une application ethernet, détaillée au paragraphe 1.2.3.3. Dans un premier temps, une solution standard pour protéger les E/S de la technologie SmartMOS 8 MV à ces niveaux de tension est décrite. Par la suite, elle servira de référence afin d'évaluer les performances des protections développées. Dans un second temps, l'option retenue pour une solution alternative est présentée.

1.4.3.1. Solution conventionnelle

Pour les E/S 80 Volts, une des solutions standard utilisées à Freescale est basée sur une stratégie centralisée, dont la structure de protection comprend un LDMOS (Lateral DMOS) et des transistors bipolaires verticaux mis en série. Comme toute protection centralisée, son efficacité dépend du nombre d'E/S à protéger, sa surface équivalente étant d'autant plus petite que ce nombre est important. La Figure 35 représente la caractéristique TLP avec un courant ramené par unité de surface équivalente, pour dix E/S protégées. Ce cas est particulièrement optimiste, une puce en technologie SmartPower présentant en général moins d'E/S "hautes tensions". Le Tableau 8 donne les performances correspondantes. Le facteur de mérite du R_{ON} est calculé pour un courant allant jusqu'à 1,3 Ampères, soit le courant maximal d'une décharge HBM de 2 kV.

Ces performances ne vérifient pas les spécifications de l'application ethernet présentées au paragraphe 1.2.3.3 (Tableau 6). La tension de maintien V_H est bien supérieure à 80 Volts et le facteur

de mérite de la robustesse F_{t2} est nettement supérieur aux $5,8 \text{ mW} \cdot \mu\text{m}^2$ requis. Mais, le facteur de mérite du $R_{ON} F_{RON}$ dépasse la limite de $42 \mu\text{m}^2 \cdot \text{W}^{-1}$, et ne peut être ramené à une valeur acceptable qu'au prix d'une augmentation de la surface. De plus, la tension de déclenchement V_t est supérieure de quelques Volts à la tension de dégradation (100 Volts), ce qui nécessite l'ajout d'un circuit de déclenchement externe.

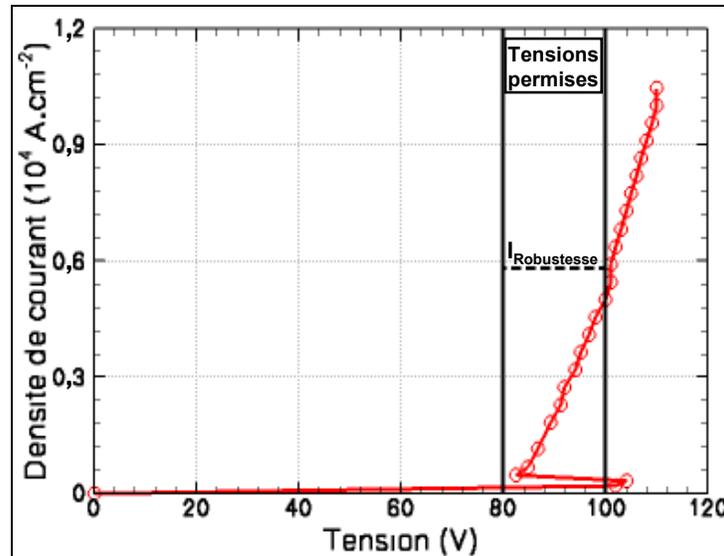


Figure 35: Caractéristique TLP (avec le courant ramené à la surface équivalente) d'une protection centralisée conventionnelle des E/S 80 Volts.

Tableau 8: Performances de la solution conventionnelle pour dix E/S protégées.

V_t (V)	104
V_H (V)	82,5
F_{RON} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	52
F_{t2} ($\text{mW} \cdot \mu\text{m}^2$)	11,5

1.4.3.2. Type de protection retenu pour une solution alternative

La solution conventionnelle ne remplissant pas les spécifications visées, le développement d'une solution alternative s'impose. Une stratégie à base de protections localisées est retenue. Si la spécification sur la surface maximale ($150 \times 150 \mu\text{m}^2$) est respectée, alors la protection peut être insérée au voisinage d'un plot d'E/S sans pénaliser la surface totale de la puce. Concernant la structure de protection, nous avons choisi de la développer à partir de transistors bipolaires autopolarisés. Une tension de maintien élevée devrait être plus facile à atteindre que dans le cas d'une structure PNPN, et les propriétés en termes de R_{ON} sont meilleures que celles des transistors à effet de champ. Afin de réaliser des structures de protection "haute tension", nous nous attacherons à définir des règles de dessin pour l'optimisation du comportement vis-à-vis des trois défis majeurs : limitation ou suppression du repliement, forte robustesse à tension élevée et faible R_{ON} à tension élevée. Au regard de ces conditions, nous évaluerons les opportunités offertes à la fois par les transistors NPN et par les transistors PNP. Du point de vue de la réalisation technologique, des structures latérales et verticales seront disponibles pour les transistors NPN, alors que seules des structures latérales seront disponibles pour les transistors PNP, l'absence de couche enterrée P ne permettant pas la réalisation de transistors PNP verticaux.

1.5. Conclusion

L'étude des tests de robustesse des circuits intégrés vis-à-vis des ESD a permis de mettre en évidence différentes contraintes concernant les structures de protection. En transitoire, sur quelques centaines nanosecondes, elles doivent pouvoir supporter des puissances de plusieurs centaines de Watt. Le circuit ne sera effectivement protégé que si elles limitent la tension en dessous de la tension de claquage des oxydes. Enfin leur temps de réponse doit être compatible avec des variations de plusieurs dizaines d'Ampères sur moins d'une nanoseconde. Concernant la fiabilité lorsque le circuit est en fonctionnement, la prise en compte des risques de latch-up de la protection impose une limite inférieure pour la tension à l'état passant, généralement la tension d'alimentation. Ces contraintes, à l'exception de celle sur les temps de réponse, peuvent être regroupées sous la forme d'une fenêtre de conception, correspondant à la région du plan I-V autorisée pour la caractéristique à l'état passant. Dans le cas particulier des E/S "hautes tensions", la fenêtre de conception est étroite et couvre des valeurs de tension élevées. En conséquence, la structure de protection doit se déclencher à une tension élevée, sa résistance à l'état passant R_{ON} doit être faible et le repliement de la caractéristique électrique doit être limité, voire supprimé. Dans ces conditions, les défis scientifiques et techniques à relever consistent, d'une part, à contrôler le repliement, et d'autre part à obtenir une protection robuste et un faible R_{ON} alors que la tension est élevée et que, pour des raisons de coût, la surface disponible est très limitée.

Du point de vue des effets physiques mis en jeu, une structure de protection présente des spécificités ne se retrouvant dans aucun autre domaine d'application. Sa surface devant être réduite au maximum, l'énergie dissipée au cours d'une ESD sera confinée dans un volume tellement restreint que le fonctionnement se situe à la limite de la destruction. Or, sur une durée de quelques centaines nanosecondes, les densités de courant restent acceptables jusqu'à des valeurs exceptionnellement élevées, de l'ordre de 10^6 A.cm⁻². Si la protection est optimisée, la destruction résulte du claquage thermique, ou second claquage thermique, survenant entre 700 et 1250 degrés Kelvins. Aussi, une protection ne pourra être correctement optimisée que si, au préalable, son comportement dans de telles conditions a bien été compris. Pour cette étude, nous pourrions nous appuyer sur des caractérisations spécifiques au domaine des ESD, telles que les caractérisations TLP et vfTLP, permettant d'appréhender la réponse électrique à une ESD, et les caractérisations TIM, donnant accès à la distribution de température à l'intérieur de la structure de protection. Le LAAS dispose d'appareils de mesure TLP et vfTLP, auxquels nous avons eu facilement accès tout au long de la thèse. Par contre, nous n'avons pas réalisé de mesures TIM, à l'heure actuelle un seul appareil ayant été développé, à l'université de Vienne en Autriche. Toutefois, les travaux auxquels il a donné lieu ont conduit à de nombreuses publications, sur lesquelles nous avons pu nous appuyer.

Les importants enjeux industriels génèrent un effort considérable en recherche et développement sur la protection des circuits intégrés vis-à-vis des ESD. Ce domaine étant très concurrentiel, de nombreux brevets ont été déposés, restreignant le champ d'investigation pour de nouvelles protections. Il s'avère donc indispensable de disposer au préalable d'un état de l'art aussi complet que possible. Après avoir décrit le principe des stratégies de protection, centralisées et localisées, les structures dédiées aux E/S "hautes tensions" en technologie SmartPower ont été recensées. Essentiellement, trois types de composants sont utilisés : les transistors bipolaires autopolarisés, les structures PNP et les transistors à effet de champ. Pour les E/S 80 Volts dans la technologie SmartMOS 8 MV de Freescale, la solution retenue est basée sur une stratégie centralisée, avec comme structure de protection un transistor DMOS et des transistors bipolaires verticaux, en série. Or, la surface nécessaire à un R_{ON} acceptable pénalise le coût de la puce, de sorte qu'il existe une forte demande pour des solutions alternatives de plus petites dimensions. Le développement de telles solutions a constitué l'un des principaux objectifs de ce travail. Pour cela, nous avons fait le choix d'une protection localisée à base de transistors bipolaires autopolarisés. C'est sur ce type protection que porte l'ensemble des travaux effectués au cours de la thèse.

Chapitre2 Etude théorique du comportement des transistors bipolaires autopolarisés pendant une ESD

2.1. Introduction

L'objectif est de mettre en place le cadre théorique nécessaire pour développer des protections ESD permettant de répondre à une fenêtre de conception étroite à "haute tension" (40 Volts – 80 Volts). Face ce type de spécification, des protections localisées à base de transistors bipolaires autopolarisés sont envisagées. Elles devront présenter une tension de fonctionnement élevée, une faible résistance à l'état passant (R_{ON}) et un repliement réduit, si celui-ci n'est pas supprimé. En se basant sur une approche analytique, les principaux phénomènes physiques seront décrits, ce qui permettra de définir des règles d'optimisation en s'appuyant sur la compréhension des mécanismes mis en jeu. Dans un premier temps, le comportement des transistors bipolaires autopolarisés sera étudié en faisant abstraction des spécificités du comportement physique des protections ESD, de manière à identifier les paramètres électriques caractéristiques de ce mode de polarisation, et d'en déterminer leurs influences respectives. A ce stade, il sera possible de décrire la partie "bas courant" de la caractéristique électrique I-V d'une protection ESD, c'est-à-dire le repliement et, en première approximation, la tension de maintien. En revanche, la description de l'état passant, en particulier du R_{ON} , nécessite de prendre en compte les effets des fortes densités de courant et des températures atteintes dans les protections ESD (Chapitre1). Cette étude sera abordée par l'évaluation des modifications induites sur les paramètres contrôlant la polarisation. Concernant les densités de courant, la quasi-totalité des études publiées sur les transistors bipolaires ne sont pas valables aux niveaux atteints pendant les ESD, et un important effort de conceptualisation s'est donc avéré nécessaire. Concernant l'augmentation de la température, elle induit une dérive des phénomènes physiques, due à la modification des valeurs des paramètres électriques du silicium. La dépendance thermique de ces paramètres étant décrite avec précision sur une large plage de températures (au moins jusqu'à 600 degrés Kelvins), il est possible de déterminer les tendances des dérives. Au cours de cette étude, les performances des transistors NPN et PNP seront évaluées et comparées entre elles, de manière à déterminer les atouts et les faiblesses de chacun d'entre eux.

2.2. Description du mode autopolarisé d'un transistor bipolaire

2.2.1. Développement analytique

Le cas traité est celui d'un transistor NPN autopolarisé dont la base et l'émetteur sont court-circuités. Le choix du NPN est arbitraire, les résultats seraient similaires pour un PNP. En se basant sur la description du mode de fonctionnement au Chapitre1 (1.4.2.1), un schéma électrique équivalent en est établi, faisant apparaître les particularités de ce mode de polarisation (Figure 36). Etant donné que le transistor fonctionne en dipôle, il n'est pas nécessaire de faire apparaître explicitement les trois électrodes, la tension et le courant seront notés simplement V et I . L'effet bipolaire est modélisé par la diode émetteur-base D_{EB} , traversée par le courant d'émetteur I_E , en série avec une source de courant $\alpha * I_E$ où α est le gain du transistor bipolaire en base commune donné par :

$$\alpha = \frac{I_C}{I_E} \quad (16)$$

Avec I_C courant de collecteur. Le gain en base commune est compris entre 0 et 1, et tend vers 1 dans le cas idéal. Pour la technologie SmartMOS 8 MV de Freescale, les recombinaisons dans la base peuvent être négligées, du fait des durées de vie (de l'ordre de 10^7 s) très supérieures aux temps de transit dans la base (de l'ordre de 10^{-9} s pour des distances de quelques micromètres). Le gain est alors équivalent au rapport d'injection de la jonction émetteur-base en direct donné par :

$$\gamma = \frac{I_{e-}}{I_{h+}} \quad (17)$$

Avec I_{e-} le courant d'électrons injectés de l'émetteur vers la base et I_{h+} le courant de trous injecté de la base vers l'émetteur. Les paramètres α et γ sont liés par la relation :

$$\alpha = \frac{\gamma}{1 + \gamma} \quad (18)$$

La génération par avalanche à la jonction base-collecteur est modélisée par une source de courant $(M-1) * (\alpha * I_E)$, où M est le facteur de multiplication par avalanche donné par :

$$M = \frac{I}{\alpha I_E} \quad (19)$$

Où αI_E représente le courant des porteurs entrant dans la zone de charge d'espace (ZCE) base-collecteur, avant avalanche, et I le courant total, prenant en compte l'avalanche. M est un paramètre supérieur à 1, qui diverge lorsque la tension s'approche de la tension de claquage de la jonction base-collecteur. Enfin, la résistance R_B représente la résistance de base, comprenant la résistance interne, vue par les trous entre la ZCE base-collecteur et le contact de base, plus la résistance externe, insérée entre les contacts d'émetteur et de base.

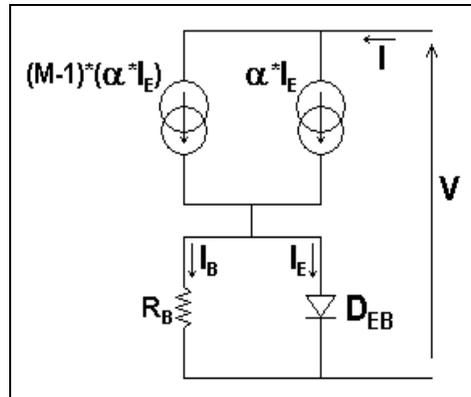


Figure 36: Schéma électrique équivalent d'un transistor NPN autopolarisé

En combinant la définition de M (19) avec l'équation :

$$I = I_E + I_B \quad (20)$$

On obtient une équation très synthétique décrivant la polarisation d'un transistor bipolaire autopolarisé :

$$\alpha M = 1 + \frac{I_B}{I_E} \quad (21)$$

Cette équation exprime le fait que la génération par avalanche (représentée par M) fournit à la fois le courant de trous injecté de la base vers l'émetteur (représenté par α) et le courant de base nécessaire à l'établissement d'une chute de potentiel dans la résistance de base suffisante pour polariser en direct la diode émetteur-base (représenté par $\frac{I_B}{I_E}$).

2.2.2. Caractéristique électrique

Pour tracer la forme générale de la caractéristique électrique I-V, il faut tout d'abord déterminer les variations des termes de l'équation (21) avec le courant et la tension. Le facteur de multiplication M est une fonction croissante de la tension. Pour une simple jonction en inverse, il existe une expression empirique [49] :

$$M = \frac{1}{1 - \left(\frac{V}{BV}\right)^n} \quad (22)$$

Avec BV tension de claquage de la jonction et n paramètre d'ajustement compris entre 2 et 4. Cette expression est une bonne approximation de la multiplication à la jonction base-collecteur d'un transistor NPN. En revanche, elle n'est pas applicable dans le cas d'un transistor PNP, où la multiplication est initialisée par les trous, dont le coefficient d'ionisation est très inférieur à celui des électrons. Concernant le gain α , les effets de fort courant et de température n'étant pas pris en compte, il est justifié de le supposer constant. Concernant les variations du rapport $\frac{I_B}{I_E}$, elles peuvent être déterminées à partir de l'équation de la jonction émetteur-base en direct :

$$I_E = I_S \left[\exp\left(\frac{qR_B I_B}{kT}\right) - 1 \right] \quad (23)$$

Avec I_S courant de saturation inverse. Cette relation exponentielle implique qu'après d'éventuels effets bas courant, le rapport $\frac{I_B}{I_E}$ décroît avec le courant et tend à devenir négligeable.

De ces considérations sur les termes de l'équation (21), on conclut à une diminution du facteur de multiplication, et donc de la tension, lorsque le courant augmente, ce qui se traduit par un repliement de la caractéristique électrique. La tension de maintien est atteinte lorsque $\frac{I_B}{I_E}$ est négligeable. Sous

cette condition, le courant de base n'apparaît plus dans l'équation (21), ce qui indique que le transistor fonctionne comme si la base était flottante. En conséquence, la tension de claquage émetteur-collecteur, base ouverte, BV_{CE0} , donne une première approximation de la tension de maintien, dont la valeur peut être calculée par l'équation implicite :

$$\alpha M = 1 \quad (24)$$

Dans le cas d'un transistor NPN, une formule explicite est obtenue à l'aide de (22) :

$$V_H = BV_{CE0} = BV_{CB0} \sqrt[3]{1 - \alpha} \quad (25)$$

Si le courant continue d'augmenter, la tension restera constante, ce qui se traduit par une caractéristique verticale et un R_{ON} nul. Concernant la tension de déclenchement, le formalisme n'est pas suffisamment précis pour la faire apparaître. On peut montrer que le rapport $\frac{I_B}{I_E}$, exprimé à partir

de la formule (23), décroît quel que soit le courant, ce qui est en contradiction avec l'augmentation de la tension lorsque la jonction est bloquée. Pour étudier le comportement avant repliement, il faudrait tenir compte des phénomènes de génération et de recombinaison, et de probables effets bidimensionnels. Cependant, cette étude n'a pas d'intérêt pratique, toutes les mesures montrant que la tension de déclenchement est proche de la tension de claquage de la jonction base-collecteur, pourvu que la résistance de base soit suffisante (de l'ordre de quelques centaines d'Ohms). Au final, on obtient la caractéristique électrique représentée sur la Figure 37.

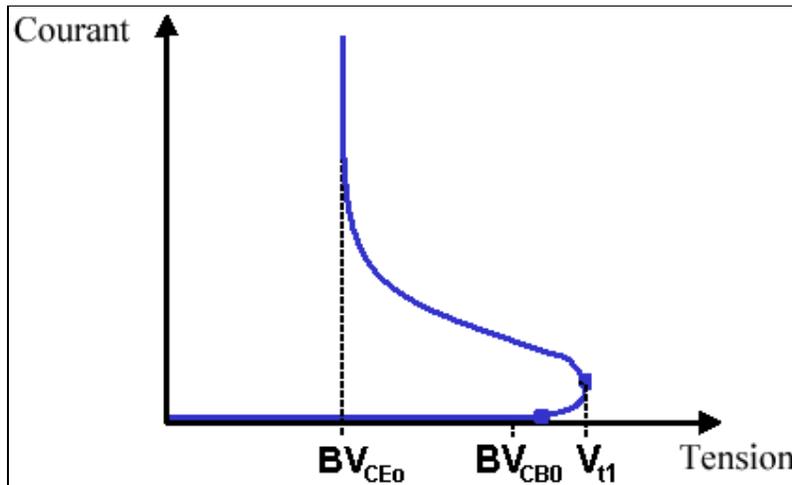


Figure 37: Caractéristique électrique obtenue à partir du développement analytique.

Il est important de souligner que le R_{ON} intrinsèque à un transistor bipolaire autopolarisé est nul, ce qui est particulièrement intéressant pour développer une protection à faible R_{ON} .

2.3. Comportement à faible courant

Les résultats de l'étude du mode autopolarisé sont directement applicables pour décrire le repliement de la caractéristique TLP d'une protection ESD. En particulier, ils permettent de donner une approximation de la tension de maintien. Aux niveaux de courant correspondants, la forte injection et l'élévation de température peuvent être négligées en première analyse.

2.3.1. Suppression du repliement en polarisation base flottante

Pour supprimer le repliement de la caractéristique I-V d'un bipolaire autopolarisé, il suffit que la base soit flottante, et non pas connectée à l'émetteur (Figure 38). Ainsi, dès le claquage (à la tension BV_{CE0}), le transistor se trouve dans le même mode de fonctionnement qu'à fort courant. Dans ce cas, l'avalanche fournit uniquement les porteurs injectés de la base vers l'émetteur, le courant à travers la résistance de base étant supprimé. Le potentiel aux bornes de la jonction émetteur-base en direct est fixé par la charge de la capacité de jonction, et non plus par le point de fonctionnement entre cette capacité et la résistance de base.

La suppression du contact de base est un moyen pour s'affranchir de la surtension au déclenchement, sans influencer sur les stratégies d'optimisation de l'état passant. La seule contrepartie est une augmentation du courant de fuite, le courant de saturation inverse de la jonction base-collecteur étant multiplié par le gain du transistor. Il sera donc préférable de garder le contact de base dans des applications où le courant de fuite est critique, ou pour les transistors dont le gain est particulièrement élevé. Enfin, même si la base est flottante, un repliement peut être induit sous l'effet des fortes densités de courant, conduisant à une tension de maintien inférieure à BV_{CE0} [50] [51] [52]. Nous reviendrons sur ce phénomène dans la section 2.4 dédiée aux effets en fort courant.

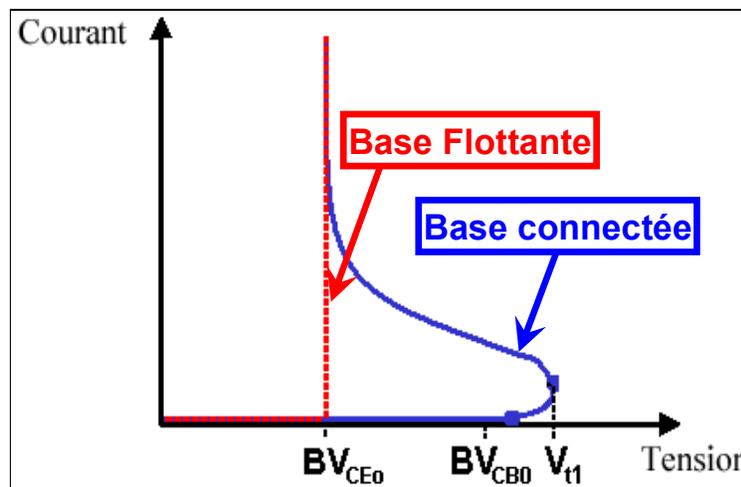


Figure 38: Caractéristiques électriques en configurations base flottante et base connectée à l'émetteur.

2.3.2. Tensions de maintien élevées

D'après la formule (24), la tension de maintien est définie par l'interaction entre l'effet bipolaire et l'avalanche à la jonction base-collecteur. Pour augmenter sa valeur, deux méthodes sont envisageables en fonction du phénomène sur lequel elles se basent. La première méthode consiste à diminuer le gain du transistor bipolaire, et ainsi augmenter le facteur de multiplication et donc la tension. La seconde méthode consiste à modifier la dépendance en tension du facteur de multiplication, de telle sorte que, sa valeur étant fixée, la tension correspondante soit plus élevée.

2.3.2.1. Gain

Pour déterminer des règles de réduction du gain, l'expression en émetteur commun est préférée à celle en base commune, car plus simple de formulation. En l'absence de recombinaison, le gain en émetteur commun β d'un transistor NPN (Figure 39) s'exprime par :

$$\beta = \frac{I_C}{I_B} = \frac{D_{nB} n_{iB}^2 N_{DE} W_E}{D_{pE} n_{iE}^2 N_{AB} W_B} \quad (26)$$

Avec I_C courant de collecteur, I_B courant de base, D_{nB} coefficient de diffusion des électrons dans la base, D_{pE} coefficient de diffusion des trous dans l'émetteur, n_{iB} concentration intrinsèque de porteurs dans la base, n_{iE} concentration intrinsèque de porteurs dans l'émetteur, N_{DE} concentration de dopants donneurs dans l'émetteur, N_{AB} concentration de dopants accepteurs dans la base, W_E largeur d'émetteur et W_B largeur de base intrinsèque.

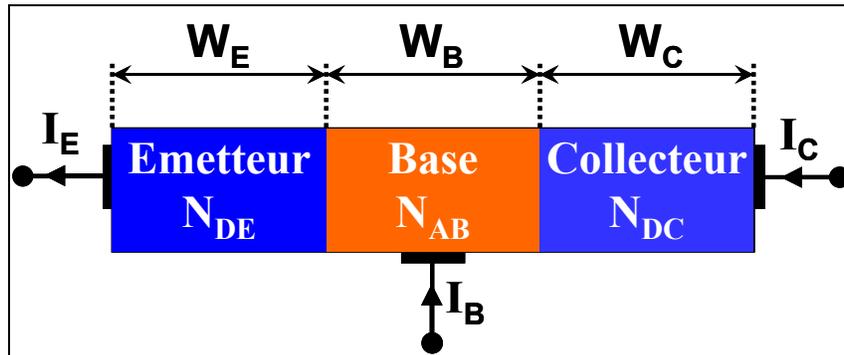


Figure 39: Représentation simplifiée d'un transistor NPN faisant apparaître les notations retenues.

Pour obtenir un gain faible, la formule (26) montre que :

- La largeur de base doit être grande devant celle d'émetteur (rapport $\frac{W_E}{W_B}$ minimum)
- Le dopage de base doit être élevé devant celui d'émetteur (rapport $\frac{N_{DE}}{N_{AB}}$ minimum)

L'évaluation des variations du gain doit également prendre en compte l'effet des coefficients de diffusions et des concentrations intrinsèques de porteurs, dont les valeurs dépendent du dopage. Ainsi, la seconde règle sur le rapport des dopages est susceptible d'être affinée, voire corrigée. Les coefficients de diffusion diminuent avec le dopage (Figure 41), de sorte qu'un rapport $\frac{N_{DE}}{N_{AB}}$ faible

implique un rapport $\frac{D_{nB}}{D_{pE}}$ faible, ce qui amplifie la réduction du gain. Par contre, le phénomène de

dégénérescence des forts dopages conduit à une augmentation de la concentration intrinsèque de porteurs n_i (Figure 40), qui se traduit par une réduction des variations apparentes du dopage [53]. Ce phénomène commence à se manifester pour des concentrations de l'ordre de $5 \cdot 10^{18} \text{ cm}^{-3}$, et provoque une réduction du dopage apparent à partir de $2 \cdot 10^{19} \text{ cm}^{-3}$. En conséquence, la concentration de dopants dans la base ne doit pas dépasser cette valeur, pour ne pas pénaliser la réduction du gain souhaitée.

Par ailleurs, le gain d'un transistor PNP est intrinsèquement plus faible que celui d'un transistor NPN, aussi, dans l'optique de la réduction du gain, les transistors PNP sont plus appropriés que les transistors NPN. Cette différence est due à la plus faible valeur du coefficient de diffusion des trous comparée à celle des électrons (dans un rapport trois aux faibles dopages (Figure 41)).

Il est important de préciser que les règles ainsi définies pour la réduction du gain devront être évaluées en fonction de leur influence sur le R_{ON} . Elles ne seront définitivement retenues que si elles tendent aussi à diminuer le R_{ON} .

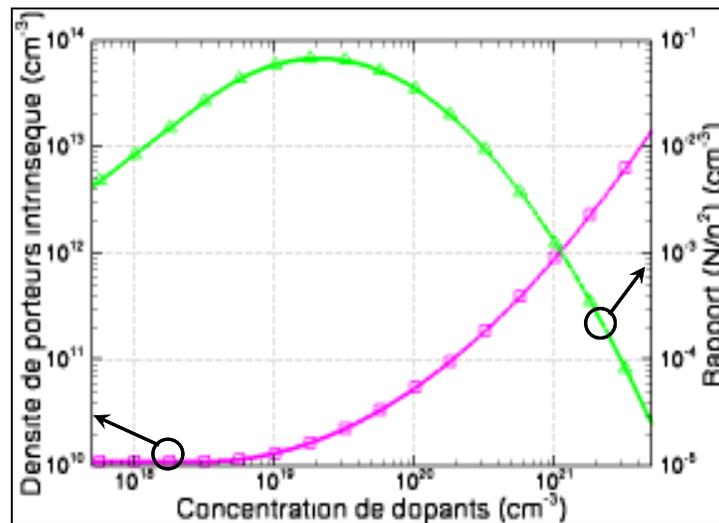


Figure 40: Densité intrinsèque de porteurs d'après le modèle de Bennett-Wilson [54] et rapport de la concentration de dopants sur la densité intrinsèque au carré

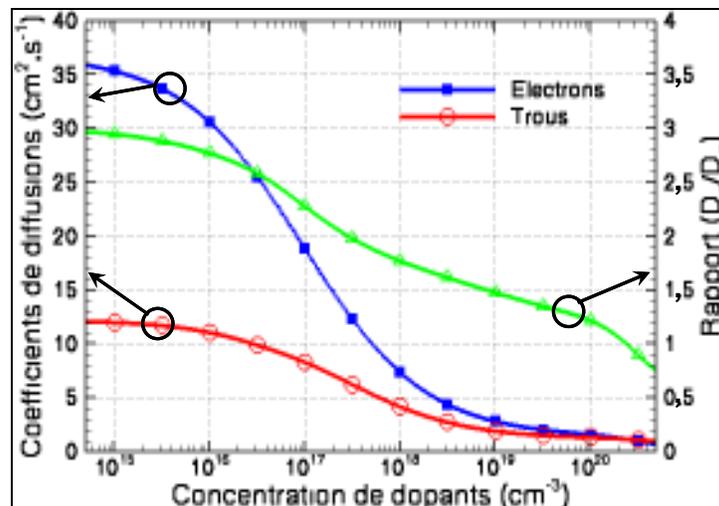


Figure 41: Coefficients de diffusion des électrons et des trous d'après le modèle de Masetti [55], et leur rapport.

2.3.2.2. Facteur de multiplication par avalanche

Le facteur de multiplication par avalanche dépend des coefficients d'ionisation par impact, qui sont fonctions du champ électrique (Figure 42). Ainsi, pour le contrôler, il faudrait travailler sur la forme du champ électrique dans la ZCE base-collecteur, c'est-à-dire sur le profil de dopage. Cependant, il est extrêmement difficile de prédire avec précision l'effet du profil de dopage sur les variations du facteur de multiplication. Entre le dopage et le facteur de multiplication, il faut passer par deux paramètres intermédiaires dont les variations sont relativement complexes : le champ électrique et les coefficients d'ionisation. De plus, le facteur de multiplication est lié aux coefficients d'ionisation par une expression mettant en jeu l'intégrale d'ionisation, particulièrement ardue à manipuler. Dans le cas d'une multiplication initialisée par des électrons, cette relation est exprimée par :

$$M_n = \frac{1}{1 - I_n} = \frac{1}{1 - \int_0^W \alpha_n \exp\left[-\int_x^W (\alpha_n - \alpha_p) dx'\right] dx} \quad (27)$$

Avec I_n l'intégrale d'ionisation des électrons, W l'extension de la ZCE base-collecteur, α_n le coefficient d'ionisation par impact des électrons et α_p le coefficient d'ionisation par impact des trous. Dans la littérature [56] [57], les résolutions analytiques de cette expression supposent une jonction abrupte et des dopages constants. Sous ces conditions et dans le cas d'une jonction plane, l'intégrale d'ionisation peut être exprimée en fonction de la tension et du dopage :

$$I_n = 0,92 \left[\exp\left(6,6 \cdot 10^{-35} \left(\frac{q N_A N_D}{\varepsilon(N_A + N_D)}\right)^3 V^4\right) - 1 \right] \quad (28)$$

Où V est la tension, N_D la concentration de dopants N et N_A la concentration de dopants P. Pour réduire le facteur de multiplication, cette expression conduit uniquement à la règle triviale consistant à réduire le dopage, ce qui revient à augmenter la tension de claquage.

Il reste à comparer le facteur de multiplication entre un transistor NPN et un transistor PNP. Dans un transistor bipolaire autopolarisé, l'avalanche est initialisée par les porteurs injectés de l'émetteur (des électrons dans un NPN et des trous dans un PNP). Or, le coefficient d'ionisation des trous étant nettement inférieur à celui des électrons (Figure 42), l'avalanche dans un transistor PNP est défavorisée. Pour compenser la faiblesse de l'avalanche et assurer la génération du courant de polarisation, le champ électrique, et donc la tension, doivent augmenter. Ainsi, pour l'obtention de tensions de maintien élevées, l'effet de la multiplication conduit à privilégier les transistors PNP plutôt que les transistors NPN, ce qui va dans le même sens que l'effet du gain. Entre ces deux effets, le plus prononcé est celui de la multiplication. En pratique, un transistor PNP a une tension de maintien élevée, même si son gain est fort. Dans les publications [58] [59] et [60], des comparaisons en mesure et en simulation entre le NPN parasite d'un NMOS et le PNP parasite d'un PMOS confirment que les transistors PNP fonctionnent à des tensions plus élevées que les transistors NPN. De plus, les caractéristiques de certains PMOS ne se replient pas, bien que le substrat (base) soit connecté à la source (émetteur).

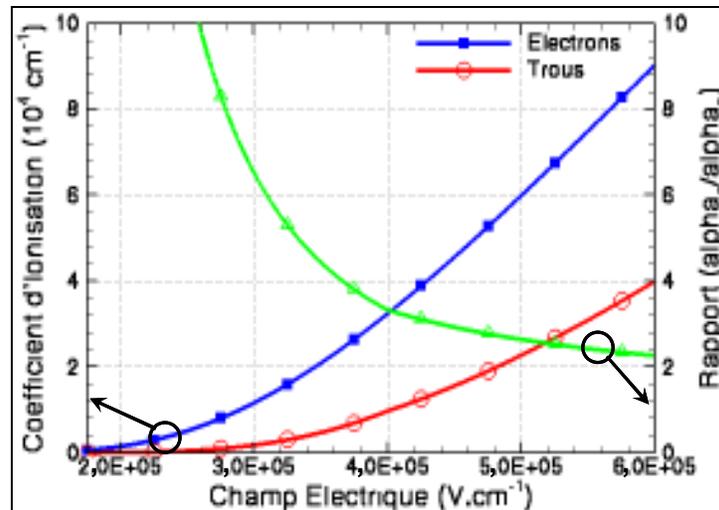


Figure 42: Coefficients d'ionisation des électrons et des trous d'après le modèle de Van Overstraeten - De Man [61], et leur rapport.

2.4. Résistance à l'état passant /Fortes densités de courant

L'étude théorique du comportement des transistors bipolaires autopolarisés fait apparaître un R_{ON} nul (2.2.2). Cependant, le cas traité est très simplifié, comparé à une protection ESD réelle. Les résistances d'accès ont été négligées, alors que leurs valeurs sont typiquement de quelques Ohms, et l'influence des fortes densités de courant et de l'autoéchauffement n'ont pas été prises en compte. Si les résistances d'accès sont une contrainte inévitable, par contre, on peut espérer définir des stratégies d'optimisation en se basant sur le contrôle des paramètres influencés par les fortes densités de courant et la température. Les effets s'y rattachant seront étudiés successivement, en commençant par ceux dus aux fortes densités de courant.

2.4.1. Notion de forte densité de courant

Les effets des fortes densités de courant apparaissent lorsque la charge induite par le courant devient comparable au dopage. Une formule simple permet de calculer la concentration de porteurs induite par le courant, pour des électrons :

$$c_n = \left| \frac{J_{e^-}}{qv_n} \right| \quad (29)$$

Où J_{e^-} est la densité de courant d'électrons et v_n la vitesse des électrons.

Afin d'évaluer les niveaux de dopage affectés lors d'une ESD, des applications numériques sont effectuées pour une densité de courant de 10^6 A.cm^{-2} . Pour la vitesse des porteurs, elle dépend du type de conduction. S'il s'agit d'une conduction par diffusion, par exemple le courant de porteurs minoritaires dans la base, elle est donnée par :

$$v_n = \frac{2D_{nB}}{W_B} \quad (30)$$

Où D_{nB} est le coefficient de diffusion dans la base et W_B la largeur de base. Pour une base de quelques micromètres, la vitesse est de l'ordre de 10^5 cm.s^{-1} , ce qui induit une concentration pouvant atteindre 10^{20} cm^{-3} . En technologie SmartPower, la base n'est jamais aussi fortement dopée, elle sera systématiquement affectée par la densité de courant. S'il s'agit d'une conduction par dérive sous un fort champ électrique, par exemple le courant à travers la ZCE base-collecteur, le flux de porteurs se déplace à la vitesse de saturation, voisine de 10^7 cm.s^{-1} . Dans ce cas, le courant peut induire jusqu'à une concentration de porteurs de 10^{18} cm^{-3} . Dans les protections "hautes tensions", les tensions de claquage élevées imposent un dopage à la jonction base-collecteur, relativement faible, inférieur à cette valeur. Aussi, les effets des fortes densités de courant se manifesteront également à la jonction base-collecteur.

2.4.2. Chute du gain en courant

La chute du gain en courant est l'un des principaux effets dus aux fortes densités de courant dans la base. Très succinctement, la concentration des porteurs majoritaires dans la base augmente pour compenser celle des porteurs minoritaires et maintenir la neutralité électrique. Le produit des concentrations de trous et d'électrons étant constant à travers la jonction émetteur-base, les porteurs minoritaires dans l'émetteur augmentent également. Il en résulte un accroissement du courant de porteurs minoritaires dans l'émetteur, c'est-à-dire un accroissement du courant de polarisation, ce qui fait chuter le gain.

Concernant le R_{ON} , cette chute du gain est préjudiciable. D'après la formule (24) décrivant l'état passant d'un transistor bipolaire autopolarisé, une diminution du gain entraîne une augmentation du facteur de multiplication et donc une augmentation de la tension. Il en résulte une résistance dynamique positive, tendant à dégrader le R_{ON} . De ce fait, une des stratégies d'optimisation consiste à limiter la diminution du gain. Pour définir des règles précises, une description quantitative est nécessaire. Il existe une formule simple décrivant les variations du gain lorsqu'apparaît la condition de forte densité de courant dans la base [62] [63] [64] :

$$\beta = \frac{J_C}{J_B} = \frac{\beta_0}{1 + \frac{J_C}{J_H}} \quad (31)$$

Où J_C est la densité de courant de collecteur, J_B la densité de courant de base, β_0 le gain en courant maximal et J_H la densité de courant critique caractérisant le début de la chute du gain. Dans le cas d'un transistor NPN, cette densité de courant critique s'exprime en fonction des paramètres de la base :

$$J_H = \frac{2qD_{nB}N_B}{W_B} \quad (32)$$

Où N_B est la concentration de dopants dans la base, W_B la largeur de base effective et D_{nB} le coefficient de diffusion des électrons dans la base. La question qui se pose est de savoir si la formule (31) est applicable aux très fortes densités de courant générées par une ESD. Pour y répondre, une application numérique est effectuée avec des valeurs typiques d'une protection ESD en technologie SmartPower :

$$\begin{aligned} \beta_0 &= 200 \\ N_B &= 1 \cdot 10^{16} \text{ cm}^{-3} \\ W_B &= 3 \text{ } \mu\text{m} \\ D_{nB} &= 30 \text{ cm}^2\text{s}^{-1} \end{aligned}$$

Le calcul de la densité de courant critique donne :

$$J_H = 320 \text{ A.cm}^{-2}$$

L'application de la formule (31) pour une densité de courant de collecteur J_C égale à 10^6 A.cm^{-2} donne :

$$\beta(10^6 \text{ A.cm}^{-2}) = 0,06$$

Pour un gain aussi faible, l'effet bipolaire serait négligeable, de sorte que le fonctionnement serait équivalent à celui d'une diode à avalanche, ce qui est en contradiction avec les mesures [28]. La formule (31) n'est donc pas applicable à un bipolaire autopolarisé durant un stress ESD, certaines des hypothèses qu'elle suppose ne sont manifestement plus valides. Cependant, aucune référence n'a été trouvée sur les variations du gain dans de telles conditions. L'étude correspondante devra donc être menée (2.5), et seulement ensuite des règles pour limiter la diminution du gain pourront être définies.

Par ailleurs, la chute du gain affecte aussi la tension de maintien. La densité de courant correspondante est de l'ordre de 10^4 A.cm^{-2} , soit deux décades au-dessus de la valeur de densité de courant critique J_H obtenue par application numérique. Ainsi, le gain a fortement diminué lorsque la tension de maintien est atteinte, ce qui conduit à deux effets très bénéfiques. D'une part, la faible valeur du gain est favorable à l'obtention d'une tension de maintien élevée. D'autre part, toute la dynamique de la chute du gain n'est pas supportée pendant l'état passant. Seule la partie finale est à prendre en compte, ce qui limite la dégradation du R_{ON} . Concernant la validité de l'expression (31) au niveau de la tension de maintien, l'application numérique donne :

$$\beta(10^4 \text{ A.cm}^{-2}) = 6,2$$

Cette valeur n'est pas incohérente, comparée par exemple aux applications de puissance où le gain est généralement de l'ordre de 10. Cependant, elle est suffisamment faible pour qu'il soit probable qu'elle soit entachée d'une erreur importante.

2.4.3. Modification de la ZCE base-collecteur par l'injection d'émetteur

A la ZCE base-collecteur d'un bipolaire, la modulation de la charge induite par des fortes densités de courant entraîne des variations du champ électrique, et donc de la tension. Cette dépendance de la tension avec le courant crée une résistance dynamique équivalente qui contribue au R_{ON} . La modulation de la charge étant différente suivant qu'elle est induite par le courant injecté de l'émetteur ou le courant de polarisation généré par avalanche, les effets de ces courants seront traités séparément, respectivement dans la présente section 2.4.3 et dans la section suivante 2.4.4.

Les variations du champ électrique sont régies par l'équation de Gauss. En une dimension, celle-ci exprime la proportionnalité entre la pente du champ électrique E et la concentration de charge ρ :

$$\frac{dE}{dx} = \frac{\rho}{\varepsilon} \quad (33)$$

Or, la charge des porteurs injectés par l'émetteur s'ajoute à celle des dopants dans la base et se retranche à celle des dopants dans le collecteur (Figure 43). Il en résulte une augmentation de la pente du champ électrique dans la ZCE côté base et une diminution côté collecteur (Figure 44). Si la charge induite par les porteurs injectés compense celle du dopage de collecteur, alors la charge totale ne s'inverse plus à la jonction métallurgique, mais à une jonction effective qui tend à se déplacer vers le collecteur (Figure 44). Ces phénomènes sont usuellement désignés par "effet Kirk", du nom de l'auteur qui les décrit le premier. Néanmoins, cette dénomination est abusive : à proprement parler, l'effet Kirk désigne la chute de la fréquence de coupure du transistor, qui est une conséquence de la modification du champ électrique.

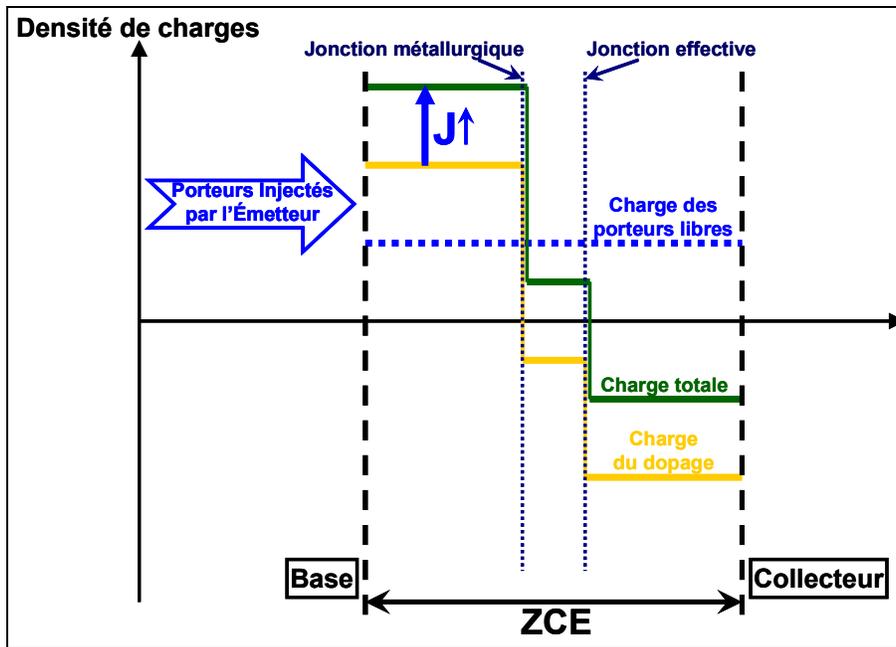


Figure 43: Effet de la forte injection sur la densité de charge dans la ZCE base-collecteur.

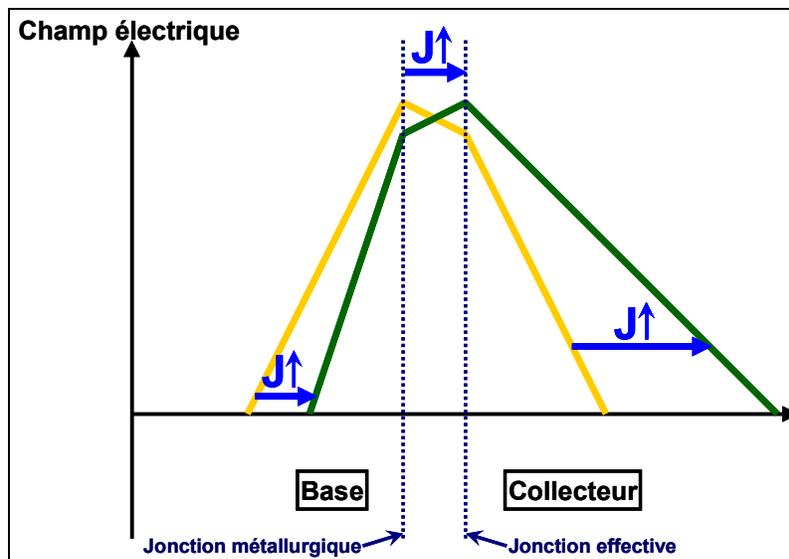


Figure 44: Effet de la forte injection sur champ électrique dans la ZCE base-collecteur.

Afin d'évaluer les variations induites sur le R_{ON} , la valeur maximale du champ électrique sera considérée constante. L'avalanche n'étant significative qu'au voisinage du maximum de champ électrique, cette hypothèse est justifiée tant que l'on fait abstraction des modifications du facteur de multiplication. Sous cette condition, l'augmentation de la pente du champ électrique côté base tend à diminuer l'extension de la ZCE et sa diminution côté collecteur ainsi que le déplacement de la jonction effective tendent à augmenter cette même extension (Figure 44). Or, pour que l'effet sur le R_{ON} soit bénéfique, il faut que l'extension totale de la ZCE diminue de manière à créer une résistance dynamique négative. La stratégie d'optimisation consistera donc à favoriser la modulation du dopage de base et à éviter celle du dopage de collecteur. Ainsi, la concentration de dopants dans la base doit être inférieure à la concentration de porteurs induite par le courant lors d'une ESD. De plus, il est préférable qu'elle soit suffisamment élevée pour ne pas être négligeable dès le repliement, ce qui induirait une diminution de la tension de maintien. Typiquement, le dopage de base doit être compris

entre 10^{16} et 10^{17} cm^{-3} . Concernant le dopage de collecteur, sa valeur doit être élevée, typiquement supérieure à $5 \cdot 10^{18}$ cm^{-3} . Enfin, pour limiter le déplacement de la jonction effective, le profil du collecteur doit être le plus abrupt possible.

Par ailleurs, l'application de ces règles entraîne un élargissement de la base effective (comprise entre les ZCE émetteur-base et base-collecteur), appelé "base push-out" en anglais. Cet élargissement de la base tend à réduire le gain, ce qui est défavorable pour le R_{ON} , d'après les considérations de la section 2.4.2. Cependant, aucune conclusion définitive ne peut être établie, tant que l'on ne dispose pas d'une description du gain aux densités de courant générées par les ESD.

Il résulte de la règle sur le dopage de base une conséquence fondamentale concernant la méthode de protection des E/S "hautes tensions". Le faible dopage entraînant une tension de claquage élevée, il est préférable de limiter la mise en série de composants, si possible d'en utiliser qu'un seul, plutôt que de mettre un grand nombre de composants "basses tensions" en série.

2.4.4. Effet du courant de polarisation sur la modification de la ZCE base-collecteur

L'effet du courant de polarisation sur la ZCE se manifeste si le gain du transistor bipolaire est faible. Sous cette condition, la charge induite par le courant de polarisation compense en partie celle induite par le courant d'injection. Pour une étude précise, il faut connaître les variations du gain lorsque sa chute en forte injection conduit à de très faibles valeurs. En attendant de disposer de la théorie adéquate, nous nous limiterons à une description qualitative.

Les modifications du champ électrique sont schématisées sur la Figure 45. Le courant de polarisation d'un transistor bipolaire autopolarisé étant généré par l'avalanche au voisinage de la jonction effective, seule la charge située du côté de la base est affectée, et non celle située du côté du collecteur. D'une part, la charge contribuant au dopage de base est affaiblie, ce qui est néfaste pour le R_{ON} . D'autre part, le déplacement de la jonction effective est freiné, ce qui est bénéfique pour le R_{ON} . Si la structure de protection est optimisée pour limiter ce déplacement, il faut s'attendre à ce que l'effet global soit défavorable.

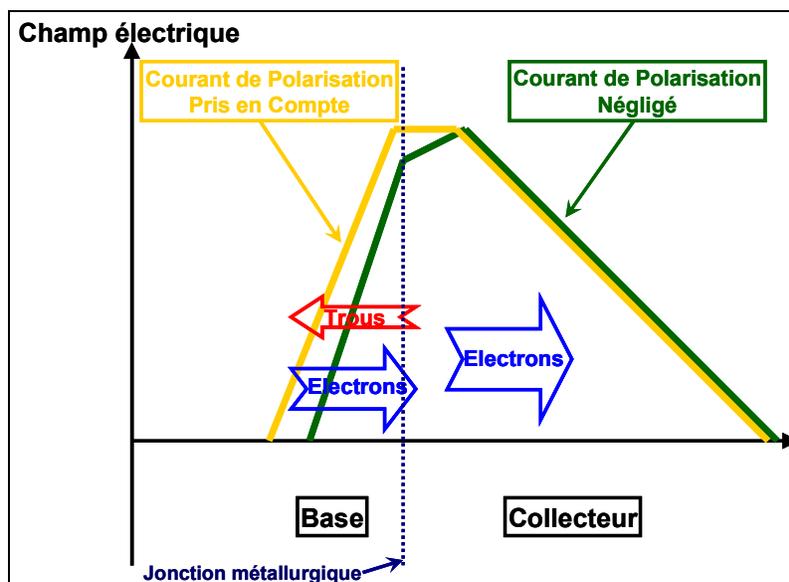


Figure 45: Effet du courant de polarisation (trous) sur le champ électrique dans la ZCE base.

2.5. Etude analytique du fonctionnement aux densités de courant générées par les ESD

Le comportement des transistors bipolaires a été peu étudié pour les densités de courant extrêmement élevées atteintes dans les protections ESD, allant jusqu'à 10^6 A.cm^{-2} . La littérature sur ce sujet est très incomplète, en particulier, les variations du gain n'ont jamais été décrites. Cette lacune est due au fait qu'aucun autre domaine d'application ne présente des densités de courant aussi élevées. Même pour des composants de puissance ou haute fréquence, la différence est de plusieurs décades.

L'objectif de l'étude est d'établir une formule analytique du gain applicable aux densités de courant générées par les ESD. A l'aide de cette formule, les questions précédemment restées en suspens, faute d'une description théorique adaptée, pourront être traitées. Ces questions concernent les effets sur la caractéristique électrique I-V de la chute du gain (2.4.2), du "base push-out" (2.4.3) et de la contribution du courant de polarisation à la modification du champ électrique (2.4.4). Les réponses devront permettre de définir des règles d'optimisation, en particulier pour la réduction du R_{ON} .

Cette étude s'appuiera sur celle présentée dans la thèse de D. Trémouilles [65]. Son approche a consisté à établir un modèle analytique unidimensionnel d'un transistor bipolaire autopolarisé aux densités de courant générées par les ESD. Ce modèle lui a permis de tracer les caractéristiques électriques I-V, par simulation numérique. Toutes les simulations sur les transistors NPN ont montré que, lorsque le courant augmente, la tension aux bornes de la ZCE diminue. Dans ce cas, la résultante des effets en fort courant est bénéfique pour le R_{ON} . Par contre, les simulations de certains transistors PNP ont montré une augmentation de la tension avec le courant, désavantageuse pour le R_{ON} .

2.5.1. Description du rapport d'injection aux densités de courant générées par les ESD

2.5.1.1. Formalisation du problème

Modélisation du transistor

La structure étudiée est un transistor NPN décrit en une dimension. Pour simplifier les calculs, les dopages sont constants et les jonctions abruptes. La description unidimensionnelle ne permettant pas une polarisation de base réaliste, celle-ci est laissée flottante. De plus, les recombinaisons seront systématiquement négligées, ce qui revient à considérer les courants d'électrons et de trous constants en dehors des zones d'ionisation par impact.

Les notations retenues dans la suite de cette section sont listées ci-dessous :

J	Densité de courant totale
J_e	Densité de courant d'électrons
J_{h+}	Densité de courant de trous
N_E	Concentration de dopants dans l'émetteur
N_B	Concentration de dopants dans la base
W_E	Largeur d'émetteur
W_B	Largeur de base effective
μ_{nE}	Mobilité des électrons dans l'émetteur
μ_{pE}	Mobilité des trous dans l'émetteur
μ_{nB}	Mobilité des électrons dans la base
μ_{pB}	Mobilité des trous dans la base
D_{nE}	Coefficient de diffusion des électrons dans l'émetteur
D_{pE}	Coefficient de diffusion des trous dans l'émetteur

D_{nB}	Coefficient de diffusion des électrons dans la base
D_{pB}	Coefficient de diffusion des trous dans la base
$v_{sat,n}$	Vitesse de saturation des électrons
$v_{sat,p}$	Vitesse de saturation des trous
n_{iE}	Concentration intrinsèque de porteurs dans l'émetteur
n_{iB}	Concentration intrinsèque de porteurs dans la base

Les résultats avec un transistor PNP seront équivalents à ceux avec transistor NPN, il suffira d'invertir les indices n et p employés respectivement pour les électrons et les trous.

Mise en équation

Puisque la base est flottante, le gain bipolaire n'est plus défini. L'étude portera sur un paramètre équivalent : le rapport d'injection γ , égal au rapport du courant d'électrons injecté de l'émetteur vers la base J_e sur le courant de trous injectés de la base vers l'émetteur J_{h+} , tel que défini par (17).

Dans les publications traitant de la chute du gain, notamment [62], le gain est supposé élevé. Or, l'application numérique (2.4.2) a montré qu'aux densités de courant générées par les ESD, la formule proposée conduit à des valeurs extrêmement faibles (au point de ne plus être réalistes). L'un des points clés de l'étude à mener consistera donc à s'affranchir de cette hypothèse de gain élevé. Dans cette optique, l'approche présentée dans la thèse de J.P. Bailbé [64] semble particulièrement bien adaptée. Elle est très générale, et n'introduit d'hypothèses ni sur la valeur du gain, ni sur le niveau de courant. En particulier, il n'est pas nécessaire que l'approximation de Boltzmann soit vérifiée, ce qui était indispensable, vu que son domaine de validité ne couvre pas les très forts courants. La formule du rapport d'injection qui a été proposée s'écrit :

$$\gamma = \frac{\int_E^{BC} \frac{n}{D_p n_i^2} dx}{\int_E^{BC} \frac{p}{D_n n_i^2} dx} = \frac{\frac{Q_E}{D_{pE} n_{iE}^2} + \frac{Q_{SB}}{D_{pB} n_{iB}^2} + \frac{Q_{SE}}{D_{pE} n_{iE}^2}}{\frac{Q_B}{D_{nB} n_{iB}^2} + \frac{Q_{SB}}{D_{nB} n_{iB}^2} + \frac{Q_{SE}}{D_{nE} n_{iE}^2}} \quad (34)$$

Où Q_E est la charge des dopants ionisés dans l'émetteur, Q_B la charge des dopants ionisés dans la base, Q_{SB} la charge induite par le courant d'électrons dans la base et Q_{SE} la charge induite par le courant de trous dans l'émetteur. Les bornes d'intégration E et BC correspondent respectivement au contact d'émetteur et à la limite entre la région de base quasi-neutre et la ZCE base-collecteur.

Le terme lié à la charge des dopants ionisés dans la base Q_B sera systématiquement négligé, étant donné que le rapport d'injection commence à diminuer quand Q_B devient négligeable devant la charge induite par le courant d'électrons dans la base Q_{SB} . Concernant les autres termes, aucun ne peut être négligé a priori. D'une part, si le rapport d'injection est faible, les courants d'électrons et de trous sont du même ordre de grandeur, impliquant que les charges Q_{SB} et Q_{SE} sont également du même ordre de grandeur, à moins d'une différence importante entre les largeurs de base W_B et d'émetteur W_E . D'autre part, les termes liés à Q_{SB} et à Q_{SE} ne doivent pas être négligés devant celui lié à la charge des dopants ionisés Q_E , le passage de la faible à la forte injection dans l'émetteur se produisant généralement aux densités de courant générées par les ESD. En anticipant sur les résultats du paragraphe suivant, une application numérique est effectuée pour un émetteur de largeur $0,25 \mu\text{m}$ et dopé à 10^{18} cm^{-3} . Le calcul de la charge Q_E (36) donne $4 \cdot 10^{-6} \text{ C.cm}^{-2}$. En prenant le coefficient de diffusion des trous égal à $4 \text{ cm}^2 \cdot \text{s}^{-1}$, la charge Q_{SE} induite pour une densité de courant de 10^5 A.cm^{-2} est $8 \cdot 10^{-6} \text{ C.cm}^{-2}$, sous la condition de faible injection (38) (39), et $4 \cdot 10^{-6} \text{ C.cm}^{-2}$, sous la condition de forte injection (38) (40). Q_E et Q_{SE} sont bien du même ordre de grandeur, pour une densité de courant typique d'une ESD. En tenant compte de ces remarques, il faudra considérer l'expression du rapport d'injection :

$$\gamma = \frac{\frac{Q_E}{D_{pE} n_{iE}^2} + \frac{Q_{SB}}{D_{pB} n_{iB}^2} + \frac{Q_{SE}}{D_{pE} n_{iE}^2}}{\frac{Q_{SB}}{D_{nB} n_{iB}^2} + \frac{Q_{SE}}{D_{nE} n_{iE}^2}} \quad (35)$$

2.5.1.2. Calcul des charges stockées

En pré requis au calcul du rapport d'injection, les charges stockées doivent être exprimées en fonction des paramètres du transistor.

Charge liée aux dopants ionisés dans l'émetteur

Le calcul de cette charge est trivial :

$$Q_E = qN_E W_E \quad (36)$$

Formules générales des charges induites par le courant

Les charges induites par le courant sont données par :

$$Q_{SB} = \tau_B J_{e-} \quad (37)$$

$$Q_{SE} = \tau_E J_{h+} \quad (38)$$

Où τ_B est le temps de transit des électrons dans la base et τ_E le temps de transit des trous dans l'émetteur. La difficulté est qu'il n'existe pas de formule analytique générale de ces temps de transit. Différentes formules approchées doivent être prises en compte, suivant les conditions de fonctionnement du transistor. Les formules en faible et forte injection sont données respectivement par [64] :

$$\text{Faible injection : } \tau_B = \frac{W_B^2}{2D_{nB}} \quad (39)$$

$$\text{Forte injection : } \tau_B = \frac{W_B^2}{4D_{nB}} \quad (40)$$

Les expressions des temps de transit dans l'émetteur sont équivalentes. La formule (40) suppose le rapport d'injection élevé ($J_{e-} \gg J_{h+}$), et n'est donc pas applicable pendant une ESD.

Les calculs des temps de transit dans la base et dans l'émetteur devront être séparé. Dans la base, il est légitime de considérer l'hypothèse de forte injection satisfaite. Par contre, dans l'émetteur, les hypothèses de faible injection et de forte injection sont susceptibles d'être validées successivement, au cours de la même décharge.

Charge induite par le courant dans la base

La différence du temps de transit en forte injection (40) par rapport celui en faible injection (39), est due à la contribution du champ électrique dans la région quasi-neutre de base, tendant à accélérer les électrons. Ce courant de dérive est modélisé en le comparant au courant de diffusion, de manière à en déduire un coefficient de diffusion effectif $D_{nB\text{effectif}}$ tenant compte de ces deux courants. Dans des conditions de forte injection et de rapport d'injection élevé, les contributions de la dérive et de la

diffusion sont égales, ce qui conduit à un coefficient de diffusion effectif de $2D_{nB}$ [62], d'où le rapport un demi entre les formules (39) et (40).

Pour déterminer le temps de transit dans des conditions de forte injection et de rapport d'injection faible, il faut revenir aux équations de continuité, qui expriment explicitement les courants de dérive et de diffusion :

$$J_{e-} = q \left(n \mu_{nB} E + D_{nB} \frac{dn}{dx} \right) \quad (41)$$

$$J_{h+} = q \left(p \mu_{pB} E - D_{pB} \frac{dp}{dx} \right) \quad (42)$$

Où n est la concentration d'électrons, p la concentration de trous et E le champ électrique (dans la région quasi-neutre de base). En forte injection, la neutralité dans la base impose $p=n$ (et $\frac{dp}{dx} = \frac{dn}{dx}$).

La densité de courant de trous (42) peut alors s'exprimer sous la forme :

$$J_{h+} = q \left(n \mu_{pB} E - D_{pB} \frac{dn}{dx} \right) \quad (43)$$

En combinant les expressions des densités de courant d'électrons J_{e-} (41) et de trous J_{h+} (43) pour supprimer le champ électrique, on obtient :

$$J_{e-} = \frac{\mu_{nB}}{\mu_{pB}} J_{h+} + 2qD_{nB} \frac{dn}{dx} \quad (44)$$

J_{h+} est ensuite supprimé à l'aide de la définition (17) du rapport d'injection γ :

$$J_{e-} = \frac{2qD_{nB}}{1 - \frac{\mu_{nB}}{\mu_{pB}\gamma}} \frac{dn}{dx} \quad (45)$$

Cette formule exprime la densité de courant d'électrons sous la forme d'un courant de diffusion, donnant le coefficient de diffusion effectif :

$$D_{nB, \text{effectif}} = \frac{2D_{nB}}{1 - \frac{\mu_{nB}}{\mu_{pB}\gamma}} \quad (46)$$

Le coefficient de diffusion effectif diminue avec le rapport d'injection, ce qui correspond à une augmentation de la contribution du courant de dérive. De plus, sa valeur est supérieure à $2D_{nB}$, c'est-à-dire que le courant de dérive est supérieur au courant de diffusion.

On peut maintenant exprimer le temps de transit des électrons dans la base et la charge induite par le courant, en fonction des paramètres du transistor et du rapport d'injection :

$$\tau_B = \frac{\left(1 - \frac{\mu_{nB}}{\mu_{pB}\gamma} \right) W_B^2}{4D_{nB}} \quad (47)$$

$$Q_{SB} = \frac{\left(1 - \frac{\mu_{nB}}{\mu_{pB}\gamma} \right) W_B^2 J_{e-}}{4D_{nB}} \quad (48)$$

Charge induite par le courant dans l'émetteur

La contribution au rapport d'injection de la charge Q_{SE} est très difficile à décrire avec précision, du fait de l'éventualité du passage, au cours de la décharge, de faible à forte injection. En effet, il a été établi [64] qu'il n'existe pas une expression unique du temps de transit en fonction des paramètres du transistor directement accessibles, applicable à la fois en faible et en forte injection. Afin de pallier à cette difficulté, le rapport d'injection sera calculé pour les deux expressions du temps de transit, en faible injection (39), et en forte injection avec un rapport d'injection élevé (40), conduisant respectivement aux charges Q_{SE1} et Q_{SE2} :

$$Q_{SE1} = \frac{W_E^2}{2D_{pE}} J_{h+} \quad (49)$$

$$Q_{SE2} = \frac{W_E^2}{4D_{pE}} J_{h+} \quad (50)$$

Malgré l'absence de description précise, cette méthode permettra du moins d'identifier les principales tendances induites par les termes liés à Q_{SE} et d'en donner une évaluation quantitative. Les informations obtenues ne pouvant être qu'approximatives, il n'est pas nécessaire de traiter le cas de la forte injection avec un rapport d'injection faible.

2.5.1.3. Formule analytique du rapport d'injection

Expression générale

L'expression (49) de Q_{SE} sera utilisée pour les calculs. Pour obtenir le résultat correspondant à l'expression (50), il suffira de multiplier par deux les termes appropriés. En insérant les expressions des charges stockées (36), (48) et (49) dans la formule du rapport d'injection (35), on obtient une équation donnant le rapport d'injection en fonction des paramètres du transistor et des densités de courants d'électrons et de trous:

$$\gamma = \frac{\frac{qN_E W_E}{D_{pE} n_{iE}^2} + \frac{\left(1 - \frac{\mu_{nB}}{\mu_{pB} \gamma}\right) W_B^2 J_{e-}}{4D_{nB} D_{pB} n_{iB}^2} + \frac{W_E^2 J_{h+}}{2D_{pE}^2 n_{iE}^2}}{\frac{\left(1 - \frac{\mu_{nB}}{\mu_{pB} \gamma}\right) W_B^2 J_{e-}}{4D_{nB}^2 n_{iB}^2} + \frac{W_E^2 J_{h+}}{2D_{nE} D_{pE} n_{iE}^2}} \quad (51)$$

Pour faciliter l'interprétation physique, il est pratique de faire apparaître la densité de courant critique J_{Cr} et les rapports Γ_{E1} et Γ_{E2} définis par :

$$J_{Cr} = \frac{2qD_{nB}^2 n_{iB}^2 N_E W_E}{D_{pE} n_{iE}^2 W_B^2} \quad (52)$$

$$r_{E1} = \frac{D_{nB}^2 n_{iB}^2 W_E^2}{D_{nE} D_{pE} n_{iE}^2 W_B^2} \quad (53)$$

$$r_{E2} = \frac{2D_{nB}^2 n_{iB}^2 W_E^2}{D_{pE}^2 n_{iE}^2 W_B^2} \quad (54)$$

Tous calculs faits, on obtient la formule du rapport d'injection recherchée :

$$\gamma = \frac{J_{Cr}}{J_{e-}} + \left(\frac{\mu_{nB}}{\mu_{pB}} - r_{E1} \right) + \sqrt{\left[\frac{J_{Cr}}{J_{e-}} + \left(\frac{\mu_{nB}}{\mu_{pB}} - r_{E1} \right) \right]^2 - \left(\frac{\mu_{nB}}{\mu_{pB}} \right)^2} + r_{E2} \quad (55)$$

La densité de courant d'électrons apparaît toujours normalisée par rapport à la densité de courant critique J_{Cr} . Celle-ci fixe donc l'ordre de grandeur des densités de courant auxquelles se manifestent les effets des faibles rapports d'injection. Une formule simple lie J_{Cr} à J_H (32), densité de courant critique caractérisant l'apparition de la forte injection dans la base :

$$J_{Cr} = \gamma_0 J_H \quad (56)$$

Où l'expression γ_0 est celle du maximum du rapport d'injection :

$$\gamma_0 = \frac{D_{nB} n_{iB}^2 N_E W_E}{D_{pE} n_{iE}^2 N_B W_B} \quad (57)$$

L'effet de la charge Q_{SE} est représenté par r_{E1} et r_{E2} . Pour obtenir le résultat correspondant à Q_{SE2} (50), ces termes doivent être divisés par deux :

$$\gamma = \frac{J_{Cr}}{J_{e-}} + \left(\frac{\mu_{nB}}{\mu_{pB}} - \frac{r_{E1}}{2} \right) + \sqrt{\left[\frac{J_{Cr}}{J_{e-}} + \left(\frac{\mu_{nB}}{\mu_{pB}} - \frac{r_{E1}}{2} \right) \right]^2 - \left(\frac{\mu_{nB}}{\mu_{pB}} \right)^2} + \frac{r_{E2}}{2} \quad (58)$$

Cette formule montre une réduction de l'effet de Q_{SE} avec la forte injection dans l'émetteur. Par ailleurs, les faibles rapports d'injection amplifiant l'effet de forte injection, via une contribution plus importante du courant de dérive, il est probable que leur prise en compte contribuerait à réduire encore plus fortement l'influence de Q_{SE} .

Expression approchée

Dans les cas où la largeur d'émetteur W_E est faible devant la largeur de base W_B , la charge induite par le courant dans l'émetteur Q_{SE} peut être négligée devant celle induite dans la base Q_{SB} . Il est alors possible de supprimer les termes liés à Q_{SE} dans l'expression (55), ce qui conduit à une nouvelle formule :

$$\gamma = \frac{J_{Cr}}{J_{e-}} + \frac{\mu_{nB}}{\mu_{pB}} + \sqrt{\left[\frac{J_{Cr}}{J_{e-}} + \frac{\mu_{nB}}{\mu_{pB}} \right]^2 - \left(\frac{\mu_{nB}}{\mu_{pB}} \right)^2} \quad (59)$$

Pour la majorité des composants qui seront étudiés dans les chapitres 3 et 4, l'émetteur est formé par une implantation de surface, de telle sorte que W_E est effectivement très faible devant W_B . Aussi, par la suite, la formule (59) sera préférentiellement utilisée. Suivant cette formule, les variations du rapport d'injection en fonction de la densité de courant normalisée sont tracées pour les transistors

NPN et PNP (Figure 46). La densité de courant normalisée d'un NPN correspond au rapport $\frac{J_{e-}}{J_{Cr}}$, et

celle d'un PNP au rapport $\frac{J_{h+}}{J_{Cr}}$. Le rapport des mobilités $\frac{\mu_{nB}}{\mu_{pB}}$ est pris égal à 3 pour un NPN et à

0,33 pour un PNP. Ces valeurs sont celles pour de faibles niveaux de dopage (10^{15} cm^{-3}) et une température de 300 K.

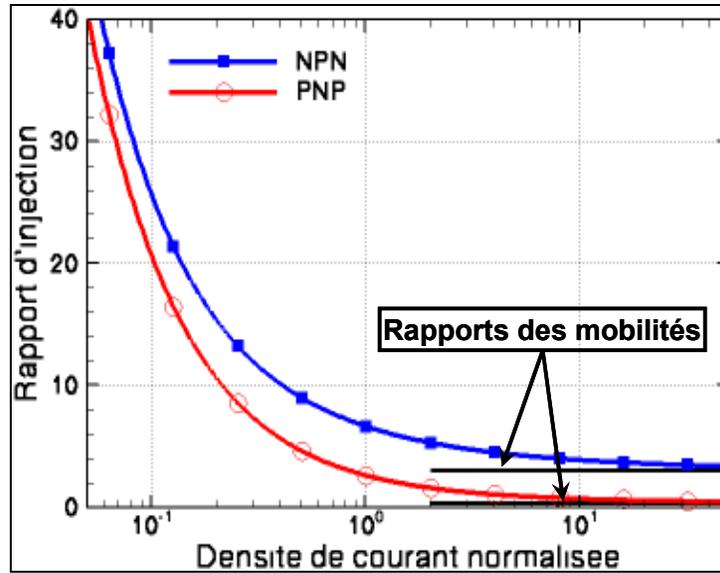


Figure 46: Rapports d'injection en suivant la densité courant normalisée (J/J_{Cr}), d'après la formule (59).

2.5.1.4. Etude asymptotique aux fortes densités de courant

Borne inférieure

La formule (55), pour le cas d'une faible injection dans l'émetteur, fait apparaître la borne inférieure du rapport d'injection :

$$\gamma_L = \left(\frac{\mu_{nB}}{\mu_{pB}} - r_{E1} \right) + \sqrt{\left(\frac{\mu_{nB}}{\mu_{pB}} - r_{E1} \right)^2 - \left(\frac{\mu_{nB}}{\mu_{pB}} \right)^2} + r_{E2} \quad (60)$$

Pour passer au cas d'une forte injection dans l'émetteur, les deux paramètres r_{E1} et r_{E2} doivent être divisé par deux.

Si la charge induite par le courant dans l'émetteur Q_{SE} est négligée, la borne inférieure devient, d'après (59) :

$$\gamma_l = \frac{\mu_{nB}}{\mu_{pB}} \quad (61)$$

Il est particulièrement remarquable que cette expression soit indépendante des caractéristiques du transistor, telles que les dimensions et les dopages. Seul le dopage de base a une influence indirecte, les mobilités étant fonction du dopage (Figure 41).

Pour que le rapport d'injection tende vers la borne inférieure γ_L ou γ_l , il faudra que la condition $J_e \gg J_{Cr}$ soit vérifiée.

Description physique

L'existence d'une borne inférieure assure la persistance de l'effet bipolaire dans les protections ESD à base de transistors bipolaires autopolarisés. La chute du rapport d'injection ne risque pas de conduire à des valeurs si faibles que le fonctionnement serait équivalent à celui d'une diode à avalanche, contrairement à ce que montrait l'application numérique avec la formule (31) (2.4.2).

La condition $J_e \gg J_{Cr}$ doit être évaluée afin de savoir si le rapport d'injection tend vers la borne inférieure γ_l . Cette évaluation n'est pas triviale car la densité de courant critique J_{Cr} dépend du courant

via la largeur de base effective W_B (52). Plus précisément, W_B varie avec la modulation de la ZCE base-collecteur sous l'effet en fort courant (2.4). En général, il en résulte un élargissement de la base effective, le "base push-out". Dans ce cas, J_{Cr} diminue avec le courant, et le fonctionnement du transistor ne s'oppose pas à ce que la condition $J_e \gg J_{Cr}$ soit vérifiée. Le rapport d'injection tendra donc vers la borne inférieure γ_1 , si l'ordre de grandeur de J_{Cr} n'est pas supérieur à celui des densités de courant générées par les ESD. Cependant, D. Trémouilles a reporté une inversion du phénomène de "base push-out" dans certains transistors PNP, c'est-à-dire une réduction de W_B [65]. Dans ce cas, J_{Cr} augmente avec le courant, et il n'est pas garanti que la condition $J_e \gg J_{Cr}$ puisse être vérifiée. La réduction de W_B peut sembler paradoxale, étant donné qu'elle correspond à une diminution de la charge de la ZCE base-collecteur alors que le courant augmente. Néanmoins, pour des rapports d'injection de l'ordre de grandeur de la borne inférieure (61), le courant de polarisation est significatif et sa contribution tend effectivement à réduire la charge de la ZCE (2.4.4). Les mécanismes physiques conduisant à une inversion du "base push-out" seront détaillés par la suite, en s'appuyant sur une expression analytique des variations de la concentration de charges avec le courant.

Dans le cas où le rapport d'injection tend vers le rapport des mobilités γ_1 (Q_{SE} négligée), la formule (35) est particulièrement explicite quant au fonctionnement du transistor. En supprimant les termes liés à Q_{SE} , on obtient :

$$\gamma = \frac{\frac{Q_E}{D_{pE} n_{iE}^2} + \frac{Q_{SB}}{D_{pB} n_{iB}^2}}{\frac{Q_{SB}}{D_{nB} n_{iB}^2}} = \frac{\frac{Q_E}{D_{pE} n_{iE}^2}}{\frac{Q_{SB}}{D_{nB} n_{iB}^2}} + \frac{\mu_{nB}}{\mu_{pB}} \quad (62)$$

Ainsi, le régime asymptotique est atteint lorsque la charge liée au courant dans la base devient prépondérante devant la charge des dopants ionisés dans l'émetteur ($Q_{SB} \gg Q_E$). Une seconde implication apparaît en comparant les équations de continuité (41) et (42). Pour que le rapport d'injection tende vers γ_1 , il faut que les courants de dérive des électrons et des trous dans la base soient prépondérants devant les courants de diffusion. Cette différence est fondamentale comparée à des densités de courant modérées (au niveau du plateau de la courbe du gain), où la diffusion des porteurs minoritaires est prépondérante, et comparée à la forte injection aux gains élevés (au début de la chute du gain), où la dérive et la diffusion sont égales.

2.5.1.5. Implications sur le R_{ON} et la tension de maintien

Dans le cas où le rapport d'injection tend vers la limite (61), des réponses peuvent être données à certaines des questions restées en suspens, faute d'une description à des densités de courant extrêmement élevées. Concernant l'influence de la chute du rapport d'injection sur le caractéristique I-V, il est possible de s'affranchir de l'effet défavorable sur le R_{ON} si la limite est atteinte dès la tension de maintien, c'est-à-dire si la densité de courant correspondante est inférieure à J_{Cr} . A titre de rappel (2.4.2), une diminution du rapport d'injection entraîne une augmentation du facteur de multiplication (24) et donc une augmentation de la tension et du R_{ON} . De plus, la faible valeur de cette limite permet d'obtenir une tension de maintien élevée. Concernant le "base push-out", la base effective W_B n'apparaissant pas dans l'expression de la limite (61), il n'influe pas sur la valeur du rapport d'injection. Seule la question relative aux effets liés au courant de polarisation (2.4.4) reste toujours sans réponse, le calcul de la charge induite dans la ZCE étant indispensable à une analyse précise. Cette analyse est essentielle, en particulier c'est elle qui conditionne l'hypothèse selon laquelle le rapport d'injection tend vers la borne inférieure (61), et donc la validité des conclusions de ce paragraphe.

2.5.2. Modulation de la ZCE base-collecteur

De manière à déterminer les tendances des modifications du champ électrique sous l'effet des fortes densités de courant, il est indispensable de faire des hypothèses sur le profil de dopage. Selon les règles d'optimisation du R_{ON} définies en 2.4.3, le dopage de collecteur sera considéré élevé et de profil abrupt de telle sorte qu'il ne soit pas modifié par la charge résultant du courant, et le dopage de base sera considéré suffisamment faible pour être négligeable devant cette même charge. Ainsi, la charge côté collecteur peut être approchée par celle du dopage, et seule la charge du côté base doit être calculée en prenant en compte les courants d'injection et de polarisation. Sous ces conditions et si le champ électrique maximum est supposé constant, les variations de la charge induite par le courant côté base reflètent directement les variations de l'extension de la ZCE, c'est-à-dire les variations de l'extension de la région de fort champ électrique. La tension étant l'intégrale du champ électrique, elle suit les mêmes variations, qui au final sont aussi celles du R_{ON} . Le calcul de l'extension de la ZCE peut ainsi être évité : ce calcul est difficile (résolution de l'intégrale d'ionisation (27)) et ne facilite pas la description des phénomènes physiques.

Par la suite, la concentration de charge induite par le courant du côté base de la jonction sera nommée simplement "concentration de charge", afin d'alléger la rédaction.

2.5.2.1. Calcul de la concentration de charge en fonction de la densité de courant

Dans la ZCE base-collecteur, les électrons et les trous transitent à leurs vitesses de saturation respectives, de sorte que la concentration de charge résultante peut s'écrire :

$$\rho = \frac{J_{h+}}{v_{sat,p}} - \frac{J_{e-}}{v_{sat,n}} \quad (63)$$

Ce que l'on cherche à déterminer, c'est une formule de la concentration de charge en fonction du courant total J , et non en fonction de J_{e-} et J_{h+} . Tout d'abord J_{h+} est supprimé, en utilisant la définition de J :

$$J = J_{e-} + J_{h+} \quad (64)$$

Ce qui donne une expression de la concentration de charge en fonction du courant total et du courant d'électrons :

$$\rho = \frac{1}{v_{sat,p}} J - \left(\frac{1}{v_{sat,n}} + \frac{1}{v_{sat,p}} \right) J_{e-} \quad (65)$$

Il reste à exprimer J_{e-} en fonction de J . Pour cela la définition du courant total (64) et les formules du rapport d'injection (17) et (59) sont combinées de manière à obtenir l'équation du second ordre en J_{e-} , dont les coefficients sont fonction de J :

$$(1 + r_1)^2 J_{e-}^2 + 2(J_{Cr} - r_1(1 + r_1)J)J_{e-} + r_1^2 J^2 - 2J_{Cr}J = 0 \quad (66)$$

$$\text{avec } r_1 = \frac{\mu_{nB}}{\mu_{pB}} \quad (67)$$

La résolution de cette équation donne :

$$J_{e-} = \frac{r_1(1 + r_1)J - J_{Cr} + \sqrt{2(1 + r_1)J_{Cr}J + J_{Cr}^2}}{(1 + r_1)^2} \quad (68)$$

En substituant à J_e l'expression (68) dans la formule (65), on obtient la formule recherchée de la concentration de charge en fonction de la densité de courant totale :

$$\rho = \frac{1}{v_{sat,p}} J - \left(\frac{1}{v_{sat,n}} + \frac{1}{v_{sat,p}} \right) \frac{r_1(1+r_1)J - J_{Cr} + \sqrt{2(1+r_1)J_{Cr}J + J_{Cr}^2}}{(1+r_1)^2} \quad (69)$$

2.5.2.2. Etude des phénomènes induits par la modulation de la ZCE base-collecteur

Représentation graphique de la concentration de charge en fonction de la densité de courant

Pour représenter graphiquement les variations de la concentration de charge avec le courant, l'équation (69) est modifiée de manière à normaliser les densités de courant et la concentration de charge, respectivement par rapport à J_{Cr} et à une concentration de charge critique définie par :

$$\rho_{Cr} = J_{Cr} \left(\frac{1}{v_{sat,n}} + \frac{1}{v_{sat,p}} \right) \quad (70)$$

On obtient la nouvelle formulation de la concentration de charge en fonction de la densité de courant totale :

$$\frac{\rho}{\rho_{Cr}} = r_2 \frac{J}{J_{Cr}} - \frac{r_1(1+r_1)\frac{J}{J_{Cr}} - 1 + \sqrt{2(1+r_1)\frac{J}{J_{Cr}} + 1}}{(1+r_1)^2} \quad (71)$$

$$\text{avec } r_2 = \frac{v_{sat,n}}{v_{sat,n} + v_{sat,p}} \quad (72)$$

De même que pour la représentation graphique du rapport d'injection, le rapport des mobilités r_1 est pris égal à 3 pour un NPN et à 0,33 pour un PNP. Le paramètre r_2 , qui ne dépend pas du dopage, est de 0,56 pour un NPN et 0,44 pour un PNP. La Figure 47 présente les courbes obtenues avec ces valeurs numériques.

Tant que la densité de courant est faible devant la densité de courant critique J_{Cr} , les concentrations de charge dans un transistor NPN et dans un transistor PNP sont semblables. En revanche, dès que la densité de courant dépasse J_{Cr} , leurs comportements diffèrent fortement. Pour le NPN, la pente s'infléchit d'un rapport deux (difficilement discernable du fait de l'échelle logarithmique), néanmoins la concentration de charge continue d'augmenter. Pour le PNP, l'inflexion de la pente est beaucoup plus marquée, au point d'entraîner une diminution de la concentration de charge jusqu'à ce qu'elle s'annule, à une densité de courant normalisée voisine de vingt. La contribution du courant de polarisation est à l'origine de ces modifications. En effet, lorsque la densité de courant s'approche de J_{Cr} , la valeur du rapport d'injection est seulement de quelques unités (59), de sorte que la polarisation n'est plus négligeable devant l'injection.

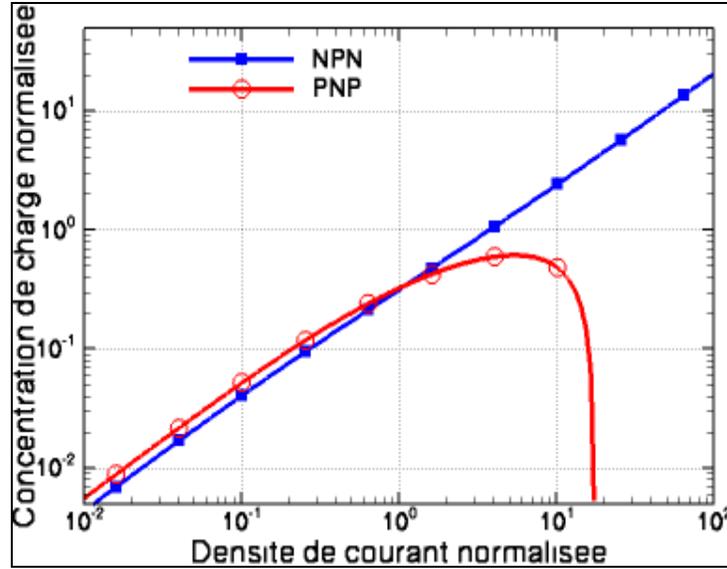


Figure 47: Concentration de charge normalisée (ρ/ρ_{Cr}) suivant la densité de courant normalisée (J/J_{Cr}), d'après la formule (71).

Calcul des variations de la concentration de charge

Pour caractériser quantitativement les tendances observées sur la Figure 47, les variations de la concentration de charge sont appréhendées en dérivant la formule (69) :

$$\frac{d\rho}{dJ} = \frac{1}{v_{sat,p}} - \frac{1}{1+r_1} \left(\frac{1}{v_{sat,n}} + \frac{1}{v_{sat,p}} \right) \left(r_1 + \frac{J_{Cr}}{\sqrt{2(1+r_1)J_{Cr}J + J_{Cr}^2}} \right) \quad (73)$$

Le sens de variation s'inverse si cette formule a des racines. Les calculs montrent que c'est le cas si et seulement si :

$$(1+r_1)r_2 - r_1 \geq 0 \quad (74)$$

Les cas des transistors NPN et PNP doivent être distingués, les évaluations numériques de cette condition étant différentes.

- Transistors NPN

Les rapports r_1 et r_2 calculés avec les valeurs numériques retenues pour un transistor NPN sont rappelés :

$$r_1 = \frac{\mu_{nB}}{\mu_{pB}} = 3,0$$

$$\frac{v_{sat,n}}{v_{sat,p}} = 1,25 \Rightarrow r_2 = 0,56$$

L'évaluation de la condition (74) donne :

$$(1+r_1)r_2 - r_1 = -0,76 \leq 0$$

La dérivée (73) n'a pas de racine réelle. La concentration de charge induite par le courant dans la ZCE est donc croissante quelle que soit la densité de courant, comme le montre la Figure 47.

Il est maintenant possible de conclure quant à la vérification de l'inégalité $J_e \gg J_{Cr}$, conditionnant la limite du rapport d'injection. L'accroissement de la concentration de charge induit une réduction de l'extension de la ZCE, et donc une augmentation de la largeur de base effective W_B . Ainsi, J_{Cr} est une

fonction décroissante de la densité de courant (52), ce qui montre que le comportement du transistor ne s'oppose pas à la vérification de la condition $J_c \gg J_{Cr}$. En conséquence, il suffit que l'ordre de grandeur de J_{Cr} soit compatible avec les densités de courant générées par les ESD, pour que le rapport d'injection tende vers le rapport des mobilités.

Concernant l'effet de la densité de courant de polarisation sur la concentration de charge, les dérivées sont calculées pour $J \ll J_{Cr}$ (courant de polarisation négligeable) et $J \gg J_{Cr}$ (contribution maximale du courant de polarisation) :

$$\left| \frac{d\rho}{dJ} \right| \xrightarrow{J \ll J_{Cr}} \frac{1}{v_{Sat,n}} = l_{1,n} \quad (75)$$

$$\left| \frac{d\rho}{dJ} \right| \xrightarrow{J \gg J_{Cr}} \frac{1}{(1+r_1)} \left(r_1 \frac{1}{v_{Sat,n}} - \frac{1}{v_{Sat,p}} \right) = l_{2,n} \quad (76)$$

Le rapport de ces limites donne :

$$\frac{l_{2,n}}{l_{1,n}} = \frac{1}{(1+r_1)} \left(r_1 - \frac{v_{Sat,p}}{v_{Sat,n}} \right) = 0,44 \quad (77)$$

La réduction par deux (0,44) des variations de la concentration de charge, observée Figure 47, est bien retrouvée. Ce ralentissement des modifications de la ZCE est significatif. Toutefois, l'effet sur le R_{ON} reste bénéfique, si les règles adéquates définies en 2.4.3 sont appliquées.

- Transistors PNP

Les rapports r_1 et r_2 calculés avec les valeurs numériques retenues pour un transistor PNP sont rappelés :

$$r_1 = \frac{\mu_{pB}}{\mu_{nB}} = 0,33$$

$$r_2 = \frac{v_{sat,p}}{v_{sat,p} + v_{sat,n}} = 0,44$$

L'évaluation de la condition (74) donne :

$$(1+r_1)r_2 - r_1 = 0,26 \geq 0$$

Dans ce cas, la dérivée (73) a une racine réelle, ce qui implique une inversion des variations de la concentration de charge, comme observé sur la Figure 47. La densité de courant de courant correspondant à cette inversion peut être calculée :

$$J_I = \frac{1 - ((1+r_1)r_2 - r_1)^2}{2(1+r_1)[(1+r_1)r_2 - r_1]^2} J_{Cr} \quad (78)$$

L'application numérique donne :

$$J_I = 5,40 J_{Cr} \quad (79)$$

Par la suite, J_I sera nommée "densité de courant d'inversion".

Une décroissance de la concentration de charge entraîne une augmentation de l'extension de la ZCE et donc une diminution de la largeur de base effective W_B . Ainsi, J_{Cr} augmente avec la densité de courant, de sorte qu'il est envisageable que la condition $J_{ht} \gg J_{Cr}$ ne puisse pas être vérifiée. Cependant, les informations disponibles ne sont pas encore suffisantes pour conclure. L'analyse du comportement asymptotique sera approfondie au paragraphe suivant, de manière à déterminer la limite du rapport d'injection. Concernant le R_{ON} , la décroissance du rapport d'injection est très pénalisante.

La modification de la ZCE ne peut plus être mise à profit, au contraire elle induit une augmentation du R_{ON} .

Tant que le gain est suffisamment faible pour que l'influence du courant de polarisation soit négligeable, la variation de la concentration de charge peut être décrite par la limite de la dérivée (73) pour $J \ll J_{Cr}$.

$$\left| \frac{d\rho}{dJ} \right| \xrightarrow{J \ll J_{Cr}} \frac{1}{v_{Sat,p}} = l_{1,p} \quad (80)$$

Limite du rapport d'injection des transistors PNP et explication physique de l'inversion du "base push-out"

La description du comportement asymptotique d'un transistor PNP nécessite de prendre en compte le fait que la charge du côté base de la ZCE doit rester positive. Cette nouvelle contrainte assure la cohérence des variations du champ électrique. L'expression (63) (avec les indices n et p intervertis) permet de la traduire par l'inégalité :

$$\gamma = \frac{J_{h+}}{J_{e-}} \geq \frac{v_{sat,p}}{v_{sat,n}} \quad (81)$$

Le rapport des vitesses de saturation est également une borne inférieure du rapport d'injection.

Dans un transistor PNP, la limite du rapport d'injection est le rapport des vitesses de saturation, dont la valeur (≈ 0.8) est supérieure à celle du rapport des mobilités (≈ 0.33). Concernant les conséquences de cette nouvelle limite sur le fonctionnement du transistor, la contribution de la diffusion au courant total ne peut être inférieure à 40 %, contrairement au cas du transistor NPN où elle devient négligeable. Cette contribution significative de la diffusion est maintenue grâce à la réduction de la base effective, conduisant à un accroissement du gradient de porteurs. Quant à la réduction de la base effective, elle est due à une augmentation plus rapide de la charge induite par le courant de polarisation que celle induite par le courant d'injection.

2.5.3. Confrontation avec la simulation numérique

Afin d'évaluer la précision de l'étude analytique, ces résultats sont confrontés à ceux d'une étude par simulation numérique unidimensionnelle. Deux transistors bipolaires seront simulés : un transistor NPN (Figure 48) et un transistor PNP complémentaires l'un de l'autre (de dopages N et P inversés). Comme supposé pour l'étude analytique, les dopages sont constants et les jonctions abruptes. Les dimensions des régions d'émetteur et de base sont choisies de manière à limiter l'influence de la charge induite par le courant dans l'émetteur Q_{SE} : la largeur d'émetteur est $0,25 \mu\text{m}$ et celle de base $3 \mu\text{m}$. La largeur de collecteur est fixée arbitrairement à $1 \mu\text{m}$. Pour les niveaux de dopages, la concentration dans l'émetteur est de 10^{18}cm^{-3} (abstraction faite d'un dopage près de la surface plus élevé, nécessaire pour une prise de contact ohmique), dans la base de 10^{16}cm^{-3} et dans le collecteur de 10^{20}cm^{-3} . La différence de deux décades entre les dopages d'émetteur et de base assure un fort gain maximal et donc une dynamique de variation importante. Le profil de dopage base-collecteur remplit les hypothèses adoptées pour l'étude des effets du courant sur la ZCE (faible dopage de base, dopage de collecteur élevé et jonction abrupte) (2.5.2). Par ailleurs, les recombinaisons ne seront pas prises en compte.

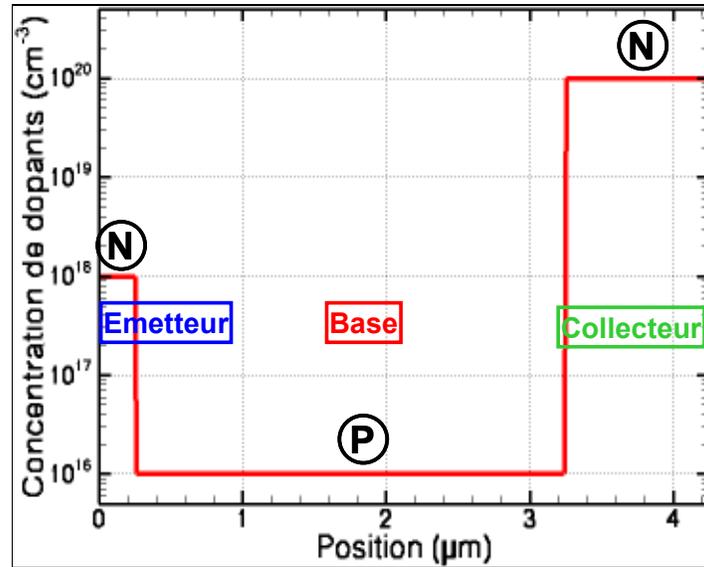
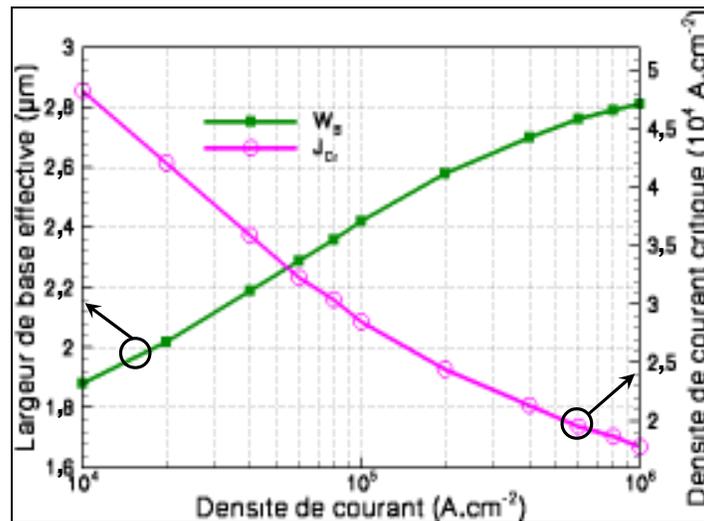


Figure 48: Profil de dopage du transistor NPN simulé.

2.5.3.1. Cas du transistor NPN

Les simulations sont effectuées pour des densités de courant allant de 10^4 A.cm^{-2} , où les hypothèses retenues pour l'étude analytique commencent à être vérifiées, jusqu'à 10^6 A.cm^{-2} , soit le maximum atteint dans les protections ESD. Les largeurs de bases effectives W_B sont relevées et les densités de courant critiques J_{Cr} (52) sont calculées, pour plusieurs densités de courant (Figure 49). Comme attendu, ces deux paramètres augmentent continûment. L'ordre de grandeur de J_{Cr} étant de 10^4 A.cm^{-2} , le comportement asymptotique sera atteint avant la fin de la simulation à 10^6 A.cm^{-2} .

Figure 49: Largeur de base effective W_B et densité de densité de courant critique J_{Cr} du transistor NPN pour des densités de courant allant de 10^4 à 10^6 A.cm^{-2} .

Le rapport d'injection est calculé à partir du relevé des densités de courants d'électrons et de trous dans la base (Figure 50). Comme attendu, il tend effectivement vers le rapport des mobilités. Ensuite, le calcul est effectué avec les expressions (59), où la charge dans l'émetteur Q_{SE} est négligée, et (58), où Q_{SE} est pris en compte dans le cas d'une forte injection dans l'émetteur (Figure 50). L'accord avec les résultats issus directement de la simulation est acceptable, en particulier pour les fortes densités de

courant, auxquelles les hypothèses retenues pour l'étude analytique sont pleinement vérifiées. Par ailleurs, la modification induite par Q_{SE} est très faible, et il est justifié de la négliger.

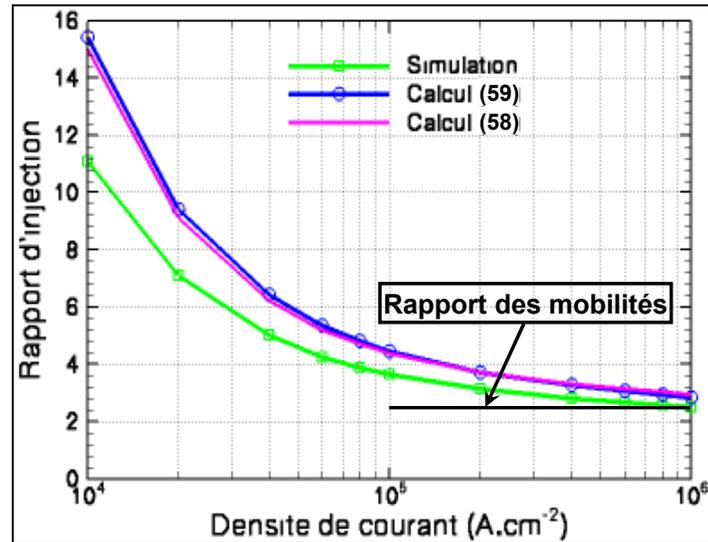


Figure 50: Rapport d'injection du transistor NPN donné par le rapport des densités de courant simulées ou par les expressions (59) et (58) pour des densités de courant allant de 10^4 à 10^6 A.cm⁻².

Concernant la précision des formules analytiques, elle est principalement limitée par l'hypothèse de mobilités constantes dans la région quasi-neutre de base, alors que celles-ci diminuent fortement dans la zone proche de la ZCE base-collecteur (Figure 51). Les calculs ont été effectués avec les valeurs maximales, en limite de la ZCE émetteur-base.

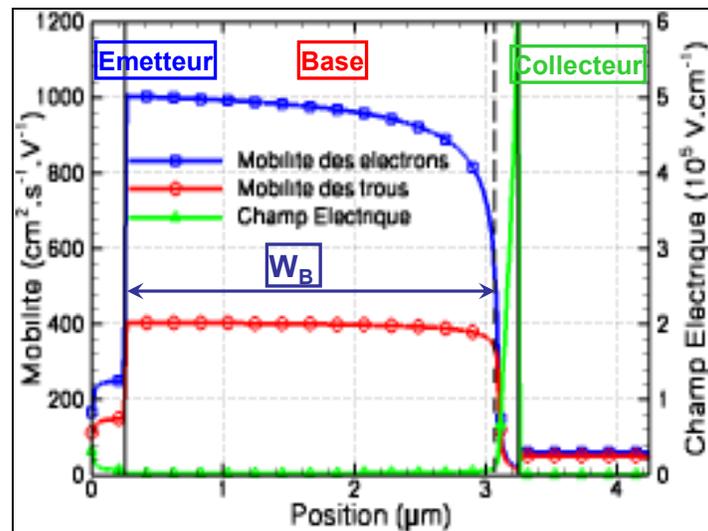


Figure 51: Profils des mobilités des électrons et des trous et profil du champ électrique dans le transistor NPN à 10^6 A.cm⁻².

2.5.3.2. Cas du Transistor PNP

Les références des expressions analytiques seront données pour les résultats obtenus avec le transistor NPN. Dans le cas du transistor PNP, les indices désignant les électrons et les trous doivent être intervertis.

Les simulations sont effectuées de 10^3 A.cm⁻², où les hypothèses en très fort courant deviennent acceptables, jusqu'à $2 \cdot 10^5$ A.cm⁻², valeur au-dessus de laquelle les résultats sont difficilement exploitables. La largeur de base W_B est relevée et la densité de courant critique J_{Cr} (52) est calculée, pour différentes densités de courant (Figure 52). Ces résultats valident, sans ambiguïté, le rétrécissement de la base à fort courant, prédit par la théorie. La comparaison des profils de champ électrique au cours de l'élargissement et au cours du rétrécissement de la base (Figure 53) montre une diminution de la pente, c'est-à-dire une réduction de la concentration de charge. La densité de courant d'inversion J_I , calculée avec l'expression (78), est de $4,1 \cdot 10^4$ A.cm⁻², en parfait accord avec la simulation. Ainsi, cette expression est exploitable en vue de la définition de règles visant à éviter l'effet négatif sur le R_{ON} du rétrécissement de la base.

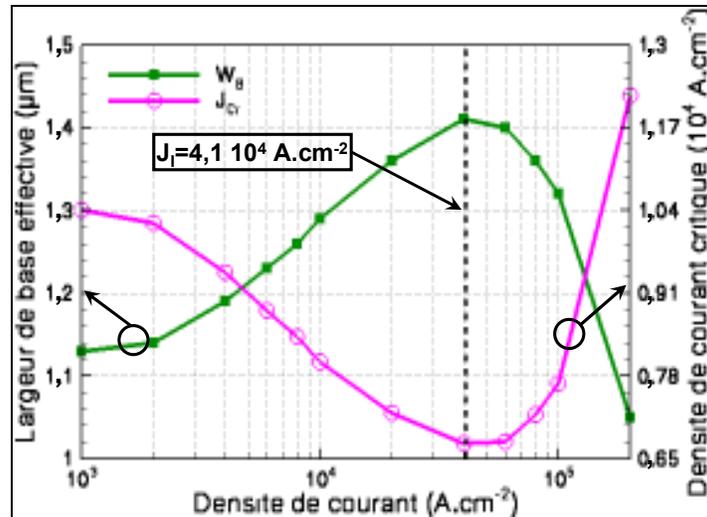


Figure 52: Largeur de base effective W_B et densité de courant critique J_{Cr} du transistor PNP pour des densités de courant allant de 10^3 à $2 \cdot 10^5$ A.cm⁻².

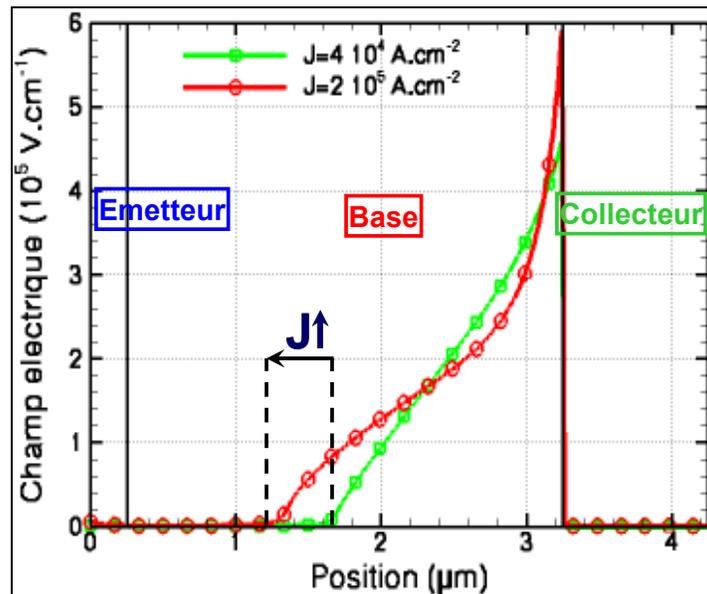


Figure 53: Profils de champ électrique dans un transistor PNP à $4 \cdot 10^4$ A.cm⁻² et $2 \cdot 10^5$ A.cm⁻².

Comme pour le transistor NPN, le rapport d'injection est calculé directement avec les densités de courant des trous et des électrons, et en utilisant les expressions analytiques (59) et (58) (Figure 54).

En accord avec l'analyse théorique, la limite n'est pas le rapport des mobilités, mais le rapport des vitesses de saturation. La précision de l'expression (59) est acceptable à partir de $4 \cdot 10^3 \text{ A.cm}^{-2}$. De plus, elle peut être significativement améliorée par la prise en compte de la charge Q_{SE} , conduisant à la formule (58).

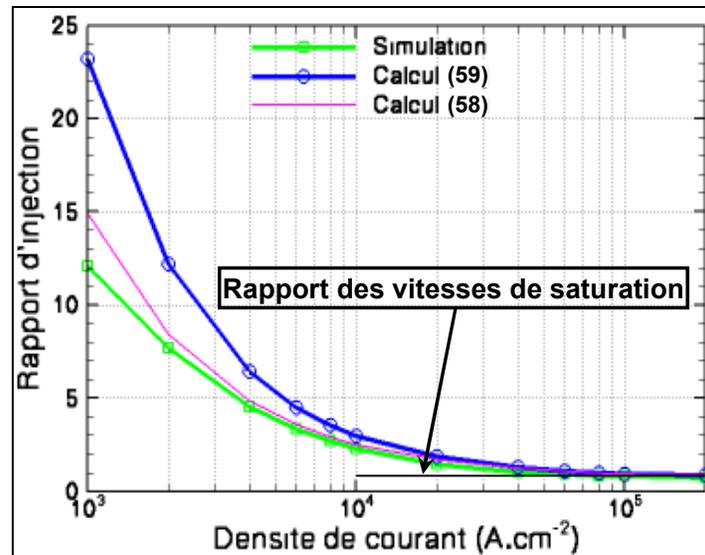


Figure 54: Rapport d'injection du transistor PNP donné par le rapport des densités de courant simulées ou par les expressions (59) et (58) pour des densités de courant allant de 10^3 à $2 \cdot 10^5 \text{ A.cm}^{-2}$.

2.6. Règles pour la conception de protections "hautes tensions" à faible R_{ON}

Tous les phénomènes électriques majeurs affectant la tension de maintien et le R_{ON} ont maintenant été décrits. Il reste à mettre à profit cette étude pour définir les stratégies d'optimisation et les règles de dessin à respecter afin d'atteindre des tensions de fonctionnement élevées et avec une résistance passante, R_{ON} , faible.

2.6.1. Transistors NPN

2.6.1.1. Stratégie d'optimisation

L'étude des mécanismes contrôlant la tension de maintien (2.3.2) a montré que la caractéristique I-V des transistors NPN a tendance à se replier de faibles tensions, à cause d'un gain généralement élevé et du fort coefficient d'ionisation des électrons. Aussi, l'augmentation de la tension de maintien est une des principales difficultés du développement de protections "hautes tensions" à base de transistors NPN. Dans cette optique, la mise à profit de la limite du rapport d'injection semble très prometteuse. Comme évoqué au paragraphe 2.5.1.5, si cette limite est atteinte au cours du repliement, sa faible valeur peut conduire à des tensions de maintien élevées. Pour cela, la densité de courant critique J_{Cr} doit être ajustée de manière à être inférieure à la densité de courant correspondant à la tension de

maintien. Il est donc inutile d'appliquer les règles de dessin visant à réduire la valeur maximale du gain (2.3.2.1), seuls les paramètres contrôlant J_{Cr} sont à considérer.

Concernant le R_{ON} , le fonctionnement asymptotique dès la tension de maintien permet de s'affranchir des effets pénalisants liés à la chute du gain. En revanche, le ralentissement de la modulation de la ZCE base-collecteur, due à la compensation induite par la charge du courant de polarisation (2.5.2.2), entraîne une atténuation des effets bénéfiques sur le R_{ON} . Bien qu'il soit envisageable de déterminer les densités de courant où le compromis entre ces deux effets est optimal, les calculs pour y parvenir seraient assez lourds, et les résultats d'une analyse aussi fine risqueraient de ne pas être directement applicables, du fait des hypothèses simplificatrices (dopages constants et jonctions abruptes). Aussi, la stratégie d'optimisation adoptée se limitera à réduire la densité de courant critique J_{Cr} , ce qui garantit une contribution des effets électriques bénéfique pour le R_{ON} , même si elle n'est pas optimale.

2.6.1.2. Règles de dessin

La formule (52) du courant critique est rappelée :

$$J_{Cr} = \frac{2qD_{nB}^2 n_{iB}^2 N_E W_E}{D_{pE} n_{iE}^2 W_B^2}$$

Les règles de dessin pour réduire le courant critique J_{Cr} d'un transistor bipolaire NPN sont :

- Un émetteur étroit (faible W_E)
- Une base large (W_B élevée)
- Un faible dopage d'émetteur (faible N_E)

Ces règles vont dans le sens de celles définies pour la réduction du gain maximal (2.3.2.1). Pour ajuster les valeurs des dopages, il faut considérer leur influence sur les coefficients de diffusion et les concentrations intrinsèques. Si le dopage d'émetteur N_E est faible, le coefficient de diffusion des trous dans l'émetteur D_{pE} aura tendance à être élevé, ce qui réduit également J_{Cr} . Pour tirer pleinement avantage de cette contribution, N_E doit être inférieur à 10^{17} cm^{-3} (Figure 41). A ce niveau de dopage n_{iE} n'a pas d'influence, la concentration intrinsèque n'augmentant qu'à partir de $5 \cdot 10^{18} \text{ cm}^{-3}$ (Figure 40). Cependant, cette règle sur le dopage d'émetteur risque d'être difficile à réaliser, car ne permettant pas de prendre un contact ohmique. Concernant le choix du dopage de base, le seul critère est de minimiser le produit $D_{nB}^2 n_{iB}^2$. D'une part, N_B doit être supérieur à 10^{18} cm^{-3} , pour profiter de la diminution de D_{nB} avec le dopage (Figure 41), et d'autre part, il doit être inférieur à $5 \cdot 10^{18} \text{ cm}^{-3}$, pour éviter l'augmentation de n_{iB} .

2.6.2. Transistors PNP

2.6.2.1. Stratégie d'optimisation

Pour un transistor PNP, l'obtention d'une tension de maintien élevée n'est pas critique. Les faibles valeurs des coefficients d'ionisation des trous assurent un faible repliement, même si le gain est élevé. La difficulté majeure concerne la réduction du R_{ON} . Contrairement au cas du transistor NPN, le R_{ON} ne sera pas amélioré en évitant la chute du rapport d'injection par la mise à profit de sa limite, car, simultanément, l'inversion du "base push-out" a une influence extrêmement négative (Figure 47). Pour avoir une chance d'obtenir un faible R_{ON} , il est impératif d'éviter cette inversion du "base push-out", et pour cela la densité de courant doit rester inférieure à J_{Cr} . Le cas optimal serait que J_{Cr} soit égale au maximum de densité de courant atteint au cours de la décharge, ce qui permettrait de limiter

la dégradation liée à la chute du gain. Cependant, un tel ajustement est très délicat, et les règles de dessin seront définies uniquement en vue d'augmenter J_{Cr} . Il faudra garder à l'esprit que ces règles permettent d'éviter l'inversion du "base push-out", mais peuvent être loin du cas optimal. Ainsi, un important travail en simulation est nécessaire pour optimiser le R_{ON} d'une protection à base de transistors PNP.

2.6.2.2. Règles de dessin

Dans un transistor bipolaire PNP, la densité de courant critique J_{Cr} est donnée par :

$$J_{Cr} = \frac{2qD_{pB}^2 n_{iB}^2 N_E W_E}{D_{nE} n_{iE}^2 W_B^2} \quad (82)$$

Les règles de dessin à appliquer pour augmenter J_{Cr} sont :

- Un émetteur large (W_E élevée)
- Une base étroite (faible W_B)
- Un dopage d'émetteur élevé (N_E élevé)

Les considérations sur les variations des coefficients de diffusion et des concentrations intrinsèques orientent le choix des dopages d'émetteur N_E et de base N_B vers des niveaux respectivement compris entre 10^{18} et 10^{20} cm^{-3} , et inférieurs à 10^{17} cm^{-3} .

Les règles sur les largeurs d'émetteur et de base effective tendent à augmenter la charge induite par le courant dans l'émetteur Q_{SE} comparativement à celle induite dans la base Q_{SB} . Le cas où Q_{SE} n'est plus négligeable n'étant pas couvert dans l'analyse des mécanismes physiques, des phénomènes non étudiés peuvent survenir, et il faut donc être prudent dans l'application de ces règles.

2.7. Résistance à l'état passant / Effets thermiques

Les phénomènes électriques étant maintenant décrits, il reste à étudier les phénomènes thermiques, et ainsi compléter la description des mécanismes physiques survenant dans un bipolaire lors d'une ESD. Seules les conséquences de l'élévation de température sur le R_{ON} seront à caractériser. Les conséquences sur le repliement, y compris la tension de maintien, sont négligeables, celui-ci survenant au début de la décharge alors que la température n'a que faiblement augmenté. L'étude portera sur les effets induits par la dépendance en température des mobilités, des coefficients d'ionisation par impact, des vitesses de saturation et de la densité intrinsèque de porteurs. Les variations de ces paramètres seront déterminées pour des températures allant de 300 à 600 degrés Kelvins, au-delà les modèles physiques du simulateur ISE ne sont pas validés. Cet intervalle ne couvre pas l'intégralité du fonctionnement du composant, les températures de dégradation par claquage thermique étant supérieures à 600 degrés Kelvins. Néanmoins, on peut raisonnablement supposer que les tendances identifiées resteront valables [67]. Par ailleurs, à titre de rappel (chapitre 1), la distribution thermique présente un "point chaud" au voisinage de la ZCE base-collecteur, où les effets thermiques seront plus intenses.

2.7.1. Résistances d'accès

Les effets sur les métallisations ne sont pas pris en compte, leurs fortes conductivités thermiques n'autorisant qu'une augmentation modérée de la température. Seules sont considérées les résistances d'accès internes, correspondant aux régions dans les diffusions d'émetteur et de collecteur où la conduction est ohmique. La résistivité ρ d'un semiconducteur dopé est donnée par :

$$\rho = \frac{1}{qN\mu} \quad (83)$$

Où N est la concentration de dopant et μ la mobilité des porteurs majoritaires. Dans cette formule, la mobilité chute avec la température (Figure 55) (Figure 56), impliquant un accroissement de la résistivité et donc des résistances d'accès. Cet effet tend à dégrader le R_{ON} . Néanmoins, cette dégradation devrait être limitée, dans la plupart des cas les résistances d'accès étant situées en dehors de la zone de forte température.

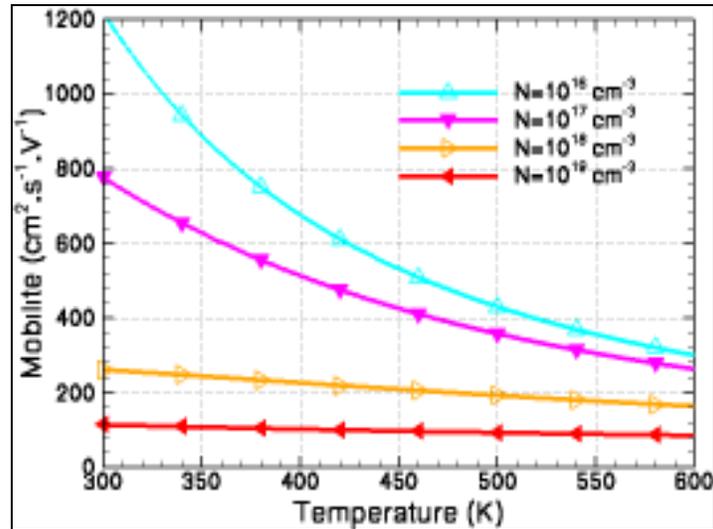


Figure 55: Mobilité des électrons suivant la température à des dopages de 10^{16} , 10^{17} , 10^{18} et 10^{19} cm^{-3} , d'après le modèle d'Arora [68].

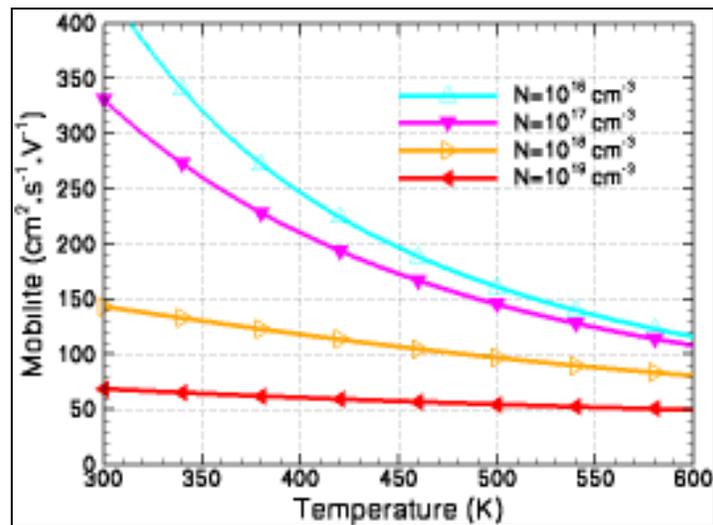


Figure 56: Mobilité des trous suivant la température à des dopages de 10^{16} , 10^{17} , 10^{18} et 10^{19} cm^{-3} , d'après le modèle d'Arora [68].

2.7.2. Multiplication par avalanche

L'élévation de température influence la génération de porteurs par avalanche via la diminution des coefficients d'ionisation par impact (Figure 57) (Figure 58). La valeur du maximum de champ électrique doit donc augmenter (Figure 61) afin de maintenir une multiplication suffisante pour fournir le courant de polarisation. La tension, correspondant à l'intégrale du champ électrique, tend à augmenter, ce qui crée une résistance dynamique positive, dégradant le R_{ON} . Ce phénomène est extrêmement défavorable, à cause, d'une part, de la forte diminution des coefficients d'ionisation (divisés par deux ou trois entre 300 et 600 degrés Kelvins) et, d'autre part, de la localisation de l'avalanche au voisinage du maximum de champ électrique, précisément où la température est la plus élevée. Pour la majorité des protections usuelles, cette contribution thermique est prépondérante de sorte que l'effet global de l'élévation de température dégrade le R_{ON} . De plus, ce phénomène n'est pas contrôlable et la seule règle pour en limiter l'effet est d'agrandir la taille de la structure, ce qui est contraire aux exigences industrielles.

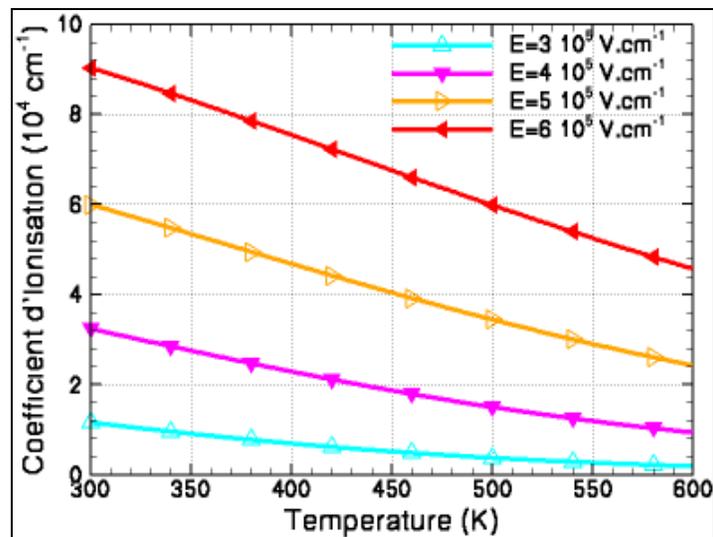


Figure 57: Coefficients d'ionisation des électrons suivant la température à des champs électriques de $3 \cdot 10^5$, $4 \cdot 10^5$, $5 \cdot 10^5$ et $6 \cdot 10^5 \text{ V.cm}^{-1}$, d'après le modèle de Van Overstraeten - De Man [61].

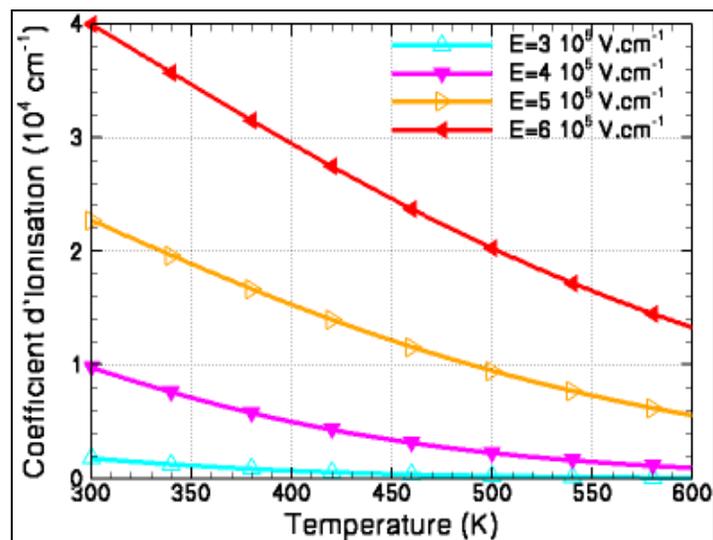


Figure 58: Coefficients d'ionisation des trous suivant la température à des champs électriques de $3 \cdot 10^5$, $4 \cdot 10^5$, $5 \cdot 10^5$ et $6 \cdot 10^5 \text{ V.cm}^{-1}$, d'après le modèle de Van Overstraeten - De Man [61].

2.7.3. Modulation de la ZCE base-collecteur

Pour caractériser l'influence de la température sur la modulation de la ZCE base-collecteur, il faut évaluer les variations de la charge résultant de la densité de porteurs induite par le courant (29). Sous l'effet de la diminution des vitesses de saturation avec la température (Figure 59), cette charge augmente, ce qui amplifie la modulation de la ZCE. A titre de rappel (2.4.3), cette modulation entraîne une augmentation de la pente du champ électrique côté base, favorable au R_{ON} , une diminution côté collecteur, défavorable au R_{ON} , et un déplacement de la jonction effective vers le collecteur, également défavorable au R_{ON} . Afin d'obtenir un effet globalement favorable, la base doit être faiblement dopée et le collecteur fortement dopé et de profil abrupt. Dans ce cas, l'amplification de la modulation de la ZCE aura tendance à compenser la résistance dynamique positive liée aux effets sur l'avalanche. La même règle sur le profil de dopage au voisinage de la jonction base-collecteur avait été définie pour tirer parti des effets en fort courant (2.4.3)

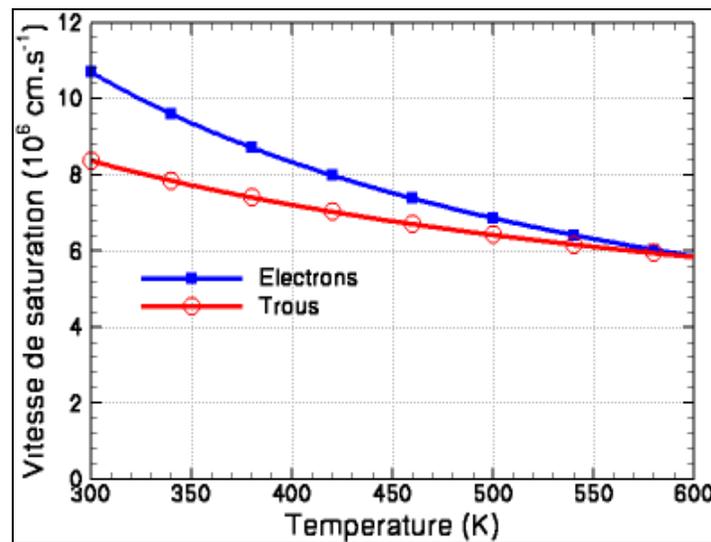


Figure 59: Vitesses de saturation suivant la température d'après le modèle de Canali [66].

2.7.4. Recombinaison-génération SRH dans la ZCE base-collecteur

La génération thermique a déjà été abordée (1.3.1.2), du fait de son rôle dans le mécanisme de défaillance par claquage thermique. Maintenant, l'objectif est de déterminer si sa contribution au courant de polarisation peut être mise à profit pour réduire le R_{ON} . Au vu des faibles niveaux de dopage requis pour une tension de claquage base-collecteur élevée, le principal mécanisme de recombinaison-génération est de type SRH, dont le taux R_{SRH} est donné par :

$$R_{SRH} = \frac{np - n_i^2}{\tau_p(n + n_i) + \tau_n(p + n_i)} \quad (84)$$

Où τ_n la durée de vie des électrons et τ_p la durée de vie des trous. Si R_{SRH} est positif, il s'agit de recombinaisons, et si R_{SRH} est négatif, de générations. Ainsi, pour que le mécanisme SRH induise une génération de porteurs, n_i^2 doit être supérieur au produit np . Bien que la concentration intrinsèque n_i augmente avec la température (Figure 60), sa valeur à 600 degrés Kelvins, égale à $4 \cdot 10^{15} \text{ cm}^{-3}$, est manifestement encore trop faible pour que la génération prenne le pas sur la recombinaison. Dans une protection "haute tension", les densités de courant pour atteindre 600 degrés Kelvins, au cours d'une

impulsion TLP, induisent des concentrations de porteurs dans la ZCE de l'ordre de 10^{17} cm^{-3} . Côté base, les concentrations d'électrons et de trous sont du même ordre de grandeur, du fait de la faible valeur du rapport d'injection, ce qui conduit à un produit np nettement supérieur à n_i^2 . Côté collecteur, la conduction n'est assurée que par un seul type de porteur, de sorte que la concentration de l'autre type de porteur chute brutalement. Ainsi, n_i^2 peut être supérieur au produit np . Cependant, si le collecteur est fortement dopé et de profil abrupt, comme préconisé pour l'optimisation du R_{ON} , la ZCE côté collecteur est trop faible pour que la génération SRH soit significative. De fait, pour les simulations des protections présentées dans les chapitres suivants, la génération SRH s'est avérée négligeable, y compris lorsque la température approche 600 degrés Kelvins. Néanmoins, pour des températures plus élevées, il a été reporté [19] que la génération thermique pouvait contribuer pour moitié au courant de polarisation. Cependant, les résultats obtenus avec le simulateur ISE ne sont plus fiables, et il est trop aléatoire d'essayer d'en déduire des stratégies pour mettre à profit la génération thermique.

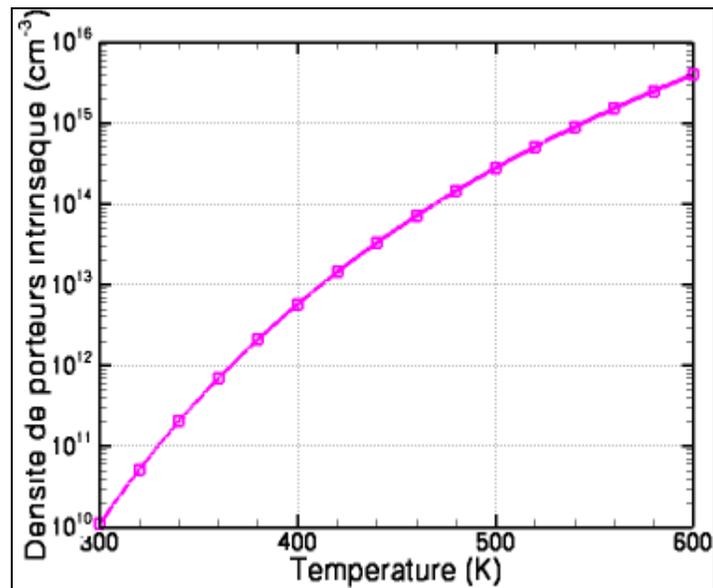


Figure 60: Concentration intrinsèque de porteurs suivant la température d'après le modèle d'Okuto [69].

2.7.5. Synthèse des modifications du champ électrique de la ZCE base-collecteur

Les coefficients d'ionisation par impact et les vitesses de saturation sont les deux paramètres dont les variations avec la température influencent le plus le champ électrique. Les effets résultants sont représentés sur la Figure 61. La diminution des coefficients d'ionisation induit une augmentation du champ électrique maximum, désavantageuse pour le R_{ON} , et d'autant plus que, pour en limiter l'effet, il n'existe pas de stratégie compatible avec les exigences industrielles. Par contre, la diminution des vitesses de saturation amplifie la modulation de la charge de la ZCE. Cet effet est bénéfique pour le R_{ON} , si le profil de dopage base-collecteur a été choisi judicieusement. Néanmoins, dans la plupart des protections usuelles, l'augmentation du maximum du champ électrique est prépondérante et le R_{ON} se dégrade avec la température.

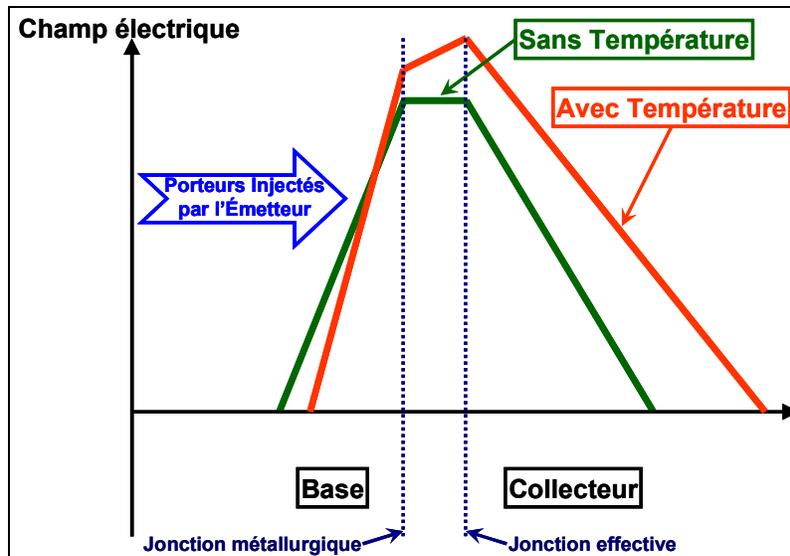


Figure 61: Modification du champ électrique dans la ZCE base-collecteur sous l'effet de la température.

2.8. Conclusion

Les mécanismes physiques dans un transistor bipolaire autopolarisé ont été décrits, dans l'optique du développement de protections "hautes tensions", à fenêtres de conception étroites. Tout d'abord, les spécificités du mode autopolarisé ont été étudiées. Les résultats les plus pertinents sont, d'une part, la possibilité de s'affranchir du repliement en laissant la base flottante, et d'autre part, la tension de maintien intrinsèquement plus élevée pour un transistor PNP que pour un transistor NPN. Ensuite, le fonctionnement en forte injection a été décrit en prenant en compte la chute du gain (ou du rapport d'injection) et la modulation de la ZCE base-collecteur. Pour affiner cette description, une étude analytique a été menée sur l'effet bipolaire aux densités de courant générées par les ESD, ce qui a permis de mettre en évidence pour la première fois certains effets de la très forte injection. En particulier, le rapport d'injection tend vers une limite égale au rapport des mobilités dans la base pour un transistor NPN et au rapport des vitesses de saturation pour un transistor PNP. Cette différence est due au fait que, dans un transistor NPN, le courant de dérive devient prépondérant par rapport au courant de diffusion, alors que, dans un transistor PNP, une contribution significative du courant de diffusion est indispensable. Par ailleurs, le calcul de la concentration de charge induite par le courant a permis d'évaluer la modulation de la ZCE. Dans un transistor NPN, cette concentration de charge croît continûment avec le courant, et la ZCE se rétrécit. Par contre, dans un transistor PNP, une diminution de la concentration de charge peut survenir, entraînant un élargissement de la ZCE. Les conséquences sur la tension avaient déjà été prédites lors de précédents travaux [65], mais l'explication physique n'avait pas été donnée. Cette étude de l'effet en fort courant a permis de définir des stratégies d'optimisation et des règles de dessin. Enfin, les modifications induites par l'élévation de température ont été présentées. Pour éviter l'influence de ces effets sur le R_{ON} , il est essentiel d'optimiser le profil de dopage base-collecteur. Cette optimisation a pour conséquence de privilégier des protections à base de composants à claquage élevé, plutôt que de mettre en série un grand nombre de composants à claquage faible.

Suivant le cas d'un transistor NPN ou PNP, les propriétés sont très différentes, aussi bien pour les niveaux de tensions de fonctionnement que le R_{ON} . Il est facile d'atteindre un fonctionnement à "haute tension" avec un PNP, alors que c'est le défi principal pour les transistors NPN. En revanche, l'optimisation du R_{ON} d'un transistor PNP résulte d'un difficile compromis entre les effets négatifs de la diminution du rapport d'injection et de l'élargissement de la ZCE. Un transistor NPN offre plus de

liberté pour régler le R_{ON} , et un tel compromis n'a pas de raison d'être. Ainsi, les deux types de bipolaires requièrent des stratégies d'optimisation très différentes. Par la suite, ils seront traités séparément, dans des chapitres dédiés.

Chapitre3 Développement de protections ESD à base de transistors bipolaires PNP

3.1. Introduction

Ce chapitre traite du développement de protections ESD à base de transistors bipolaires PNP autopolarisés. Les concepts pour l'optimisation de ces protections seront définis à partir de simulations électrothermiques, puis ils seront validés par des mesures sur silicium. En préliminaire, les problématiques liées à la simulation des protections ESD seront abordées. Ensuite, une première stratégie pour développer des protections ESD sera présentée. Elle consiste à réduire le R_{ON} de transistors bipolaires PNP latéraux. D'après l'étude du chapitre 2, les transistors PNP présentent des caractéristiques intéressantes pour la protection des E/S "hautes tensions", en particulier il est aisé d'obtenir une tension de fonctionnement élevée. Leur principal point faible est dû à des propriétés en termes de R_{ON} moins favorables que celles des transistors NPN. Des règles de dessin portant sur les configurations d'émetteur et de collecteur viendront compléter celles définies au chapitre 2, et des résultats de mesure sur des structures ainsi optimisées permettront de préciser les performances atteignables. La seconde stratégie présentée ne sera plus basée sur l'optimisation du transistor PNP seul, mais consistera à augmenter l'injection grâce à la contribution d'une diode à avalanche au courant de polarisation. Après avoir décrit le fonctionnement d'une diode à avalanche pendant une ESD, les règles de dessin pour optimiser le couplage des deux composants seront présentées et confrontées à des résultats de mesure.

3.2. Simulation pour la conception de protections ESD

Dans ces travaux, la simulation a été utilisée comme un outil d'aide à la compréhension du comportement physique (électrique et thermique) d'une protection pendant une ESD. Une telle approche est particulièrement efficace pour définir des règles d'optimisation. Par contre, la simulation ne sera pas utilisée en tant qu'outil de prédiction. Un calibrage très précis serait nécessaire, ce qui entraînerait un temps de mise en œuvre beaucoup trop long. Les simulations donneront des tendances générales et des mesures sur silicium seront indispensables pour évaluer quantitativement les résultats atteignables.

La plupart des simulations sont basées sur la même méthodologie. Dans un premier temps, une coupe 2D du profil de dopage est décrite par la simulation du procédé technologique, dans laquelle les ouvertures des masques d'implantation pour le composant étudié ont été définies au préalable.

Ensuite, cette coupe technologique sert de base à la simulation du comportement électrique ou électrothermique. Pour une protection ESD, les simulations les plus courantes sont celles du claquage statique, de la caractéristique TLP ou vTLP (la réponse à chaque impulsion d'amplitude différente étant simulée), et les réponses aux formes d'ondes standardisées (HBM, MM, IEC). Toutes ces simulations sont effectuées avec les outils TCAD de Synopsys [70].

Concernant plus spécifiquement les transistors PNP latéraux, la coupe 2D schématique d'une des structures étudiée ultérieurement est présentée sur la Figure 62(a). Le profil de dopage correspondant, obtenu par une simulation du procédé technologique SmartMOS8 MV, est présenté Figure 62(b). L'étude de cette structure servira de référence pour comprendre les effets des longueurs d'émetteur (L_E) et de collecteur (L_C), et ainsi définir les règles de dessin associées (3.3.1). La coupe 2D montre une région de collecteur formée d'une diffusion P profonde implantée à l'intérieur d'une région dopée N⁻ formant la base. La région de base s'étend jusqu'à une couche enterrée dopée N⁺ (NBL). L'émetteur est formé d'une diffusion de surface fortement dopée N⁺⁺. Deux autres diffusions de surface N⁺⁺ et P⁺⁺ permettent de prendre des contacts ohmiques respectivement pour la base et le collecteur. Enfin, une tranchée d'oxyde peu profonde sépare le contact de collecteur des contacts d'émetteur et de base, délimitant ainsi les deux zones entre lesquelles la tension sera soutenue.

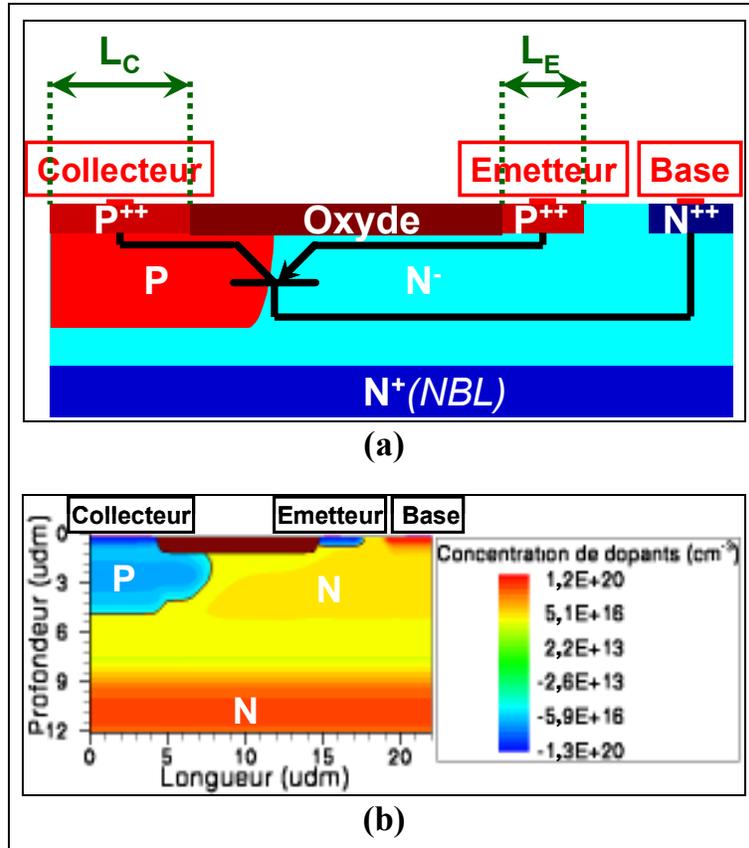


Figure 62: Coupe 2D schématique (a) et profil de dopage (b) du transistor PNP latéral retenu pour l'étude de l'influence de longueur L_E et de collecteur L_C .

3.2.1. Simulation du procédé technologique

Pour obtenir le profil de dopage 2D du composant étudié, la simulation des étapes technologiques (croissances d'épitaxies, implantations, cycles thermiques, gravures, dépôts) est préférée à une description avec des profils de dopage analytiques. Pour les technologies SmartPower avancées, telle

que la technologie SmartMOS8 MV de Freescale (Chapitre1), la description analytique des profils est très difficile, et la simulation du procédé est mieux adaptée. Cependant, une étape préliminaire de calibrage est indispensable pour obtenir des résultats réalistes [71]. Dans les fichiers de simulation disponibles, le calibrage de la plupart des étapes de la technologie SmartMOS 8 MV était déjà réalisé, à l'exception des implantations permettant de tenir les fortes tensions. Or, celles-ci sont particulièrement importantes dans le développement de protections "hautes tensions".

Les implantations sont simulées à l'aide de fonctions analytiques dont les valeurs des paramètres sont données dans des tables. L'essentiel du calibrage consiste à ajuster ces paramètres. Pour cela, on dispose de mesures des profils de dopages, obtenues par sonde ionique ou SIMS (Secondary Ion Mass Spectroscopy), et d'une évaluation de la forme des diffusions latérales, obtenue par mesure SCM (Scanning Capacitance Measurement) réalisées à l'aide d'un microscope à force atomique (AFM). Le critère pour juger de la précision du calibrage consistera à comparer la simulation des tensions de claquage avec la mesure. Dans un premier temps, l'étude porte sur les profils de dopages verticaux, pour lesquels les tensions de claquage avec la couche enterrée (NBL) seront considérées. Dans un second temps, les profils latéraux sont calibrés, et les tensions de claquage sont comparées avec celles des premiers composants réalisés sur silicium.

3.2.1.1. Profils SIMS

La sonde ionique ou SIMS est un appareil de mesure qui permet d'obtenir des profils de dopage verticaux. Très succinctement, le principe consiste à graver la surface du silicium avec un faisceau ionique, puis à séparer par filtrage en masse les différents éléments issus de la gravure. Ainsi, la concentration des dopants en fonction de la profondeur peut être relevée.

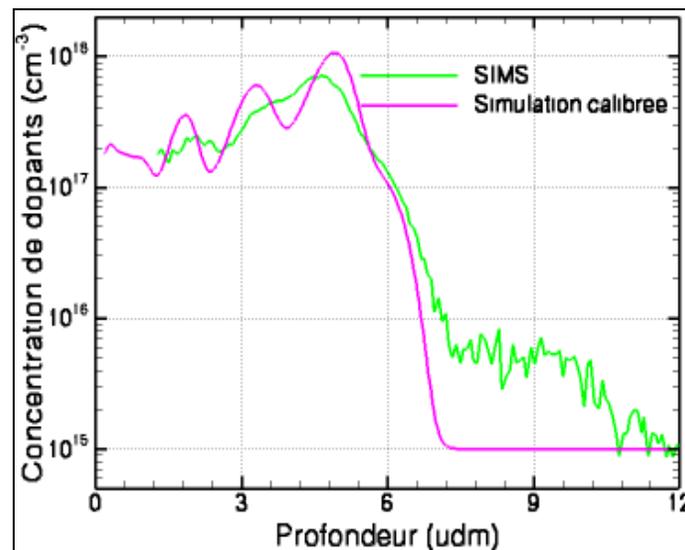


Figure 63: Profil de dopages d'une diffusion P obtenu par une mesure SIMS et la simulation du procédé.

Les mesures ont été effectuées au département de génie physique de l'INSA de Toulouse. La Figure 63 représente le résultat pour une diffusion P profonde, formant le collecteur du composant présenté sur la Figure 63. La précision est suffisante à une interprétation quantitative, néanmoins plusieurs limitations sont à prendre en considération. La mesure donne une moyenne de la concentration sur un certain intervalle en profondeur, ce qui entraîne un lissage du profil. En parallèle, le redépôt d'une fraction des atomes arrachés atténue également les variations, voire conduit à surévaluer les zones faiblement dopées. Au-dessous de 10¹⁶ cm⁻³, ce redépôt combiné avec le bruit de mesure rend le résultat inexploitable. Le substrat étant dopé approximativement à 10¹⁵ cm⁻³, la dernière décade du profil de dopage n'est donc pas mesurable. Or, l'essentiel de la ZCE (zone de charge d'espace)

s'étend dans cette région, de laquelle dépend la tension de claquage. Ainsi, le calibrage se base sur les zones les plus fortement dopées (Figure 63), l'extrapolation à faible dopage étant supposée correcte si la tension de claquage est reproduite. Pour le profil présenté, l'accord entre la simulation et la mesure du claquage avec la NBL est excellent: 33 Volts en simulation pour 34 Volts en mesure.

3.2.1.2. Mesures SCM

Le principe d'une mesure SCM consiste à déplacer la pointe d'un AFM au dessus d'une coupe verticale d'un échantillon de silicium tout en relevant la capacité correspondante. Cette capacité dépendant du type et de la concentration des dopants, une cartographie représentative du dopage est ainsi obtenue. Un des points clés pour une bonne mesure est la préparation de la coupe, en particulier sa planéité doit être la meilleure possible.

Une mesure a été effectuée avec l'AFM du LAAS, sur une structure couplant un transistor PNP et une diode à avalanche (Figure 64) du type de celle qui seront présentées dans la partie 3.4. Sur la cartographie obtenue (Figure 65), les régions de types N et P sont bien différenciées, ce qui permet de décrire finement la forme des jonctions. Il est ainsi possible d'évaluer si les jonctions sont reproduites avec précision en simulation. Par contre, la mesure ne donne pas d'informations exploitables sur les variations des concentrations de dopants. Sa sensibilité comparée au niveau de bruit est très insuffisante, elle permet à peine de distinguer des zones fortement dopées de zones faiblement dopées. Dans ces conditions, le calibrage latéral n'est pas aussi précis que le calibrage vertical. La simulation des claquages latéraux donne des écarts avec la mesure pouvant aller jusqu'à 20 %.

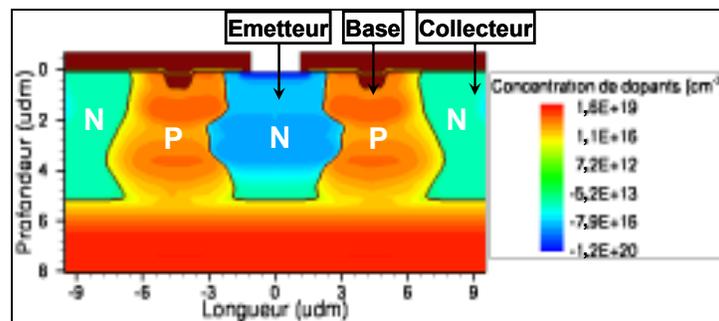


Figure 64: Coupe technologique simulée de la structure cartographiée en SCM.

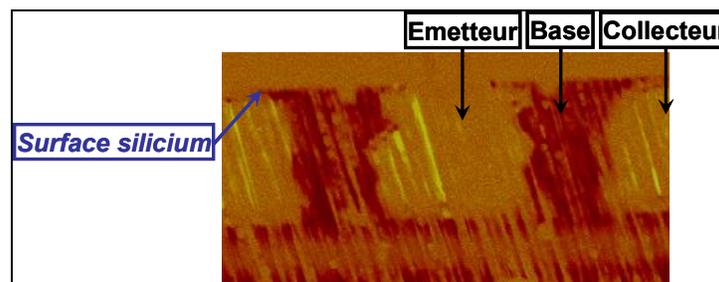


Figure 65: Cartographie SCM correspondant au profil de la Figure 64 (les régions dopées P apparaissent en clair et les régions dopées N en sombre).

3.2.2. Simulation électrothermique

La plupart des paramètres électrothermiques du silicium sont bien caractérisés, et les modèles sont fournis par le simulateur. Pour la simulation d'une protection pendant une ESD, seules les durées de vie liées aux recombinaisons SRH ont dues être calibrées. Les recombinaisons SRH étant liées à la

densité des niveaux pièges dans la bande interdite, la valeur de la durée de vie dépend fortement de la technologie, en particulier des cycles thermiques destinés à guérir les défauts cristallographiques. L'effet le plus significatif est la modification du gain, qui influence directement la caractéristique à l'état passant des transistors bipolaires autopolarisés (chapitre 2). En l'absence de calibrage, des contributions à l'erreur sur la tension de maintien allant jusqu'à 20% ont été observées.

Le calibrage s'effectue en ajustant les durées de vie de manière à ce que les courbes du gain en fonction du courant de collecteur correspondent en simulation et en mesure. Pour cet ajustement, on suppose constant le rapport de la durée de vie des électrons sur celle des trous (égal à 3,3), ce qui revient à faire varier un seul paramètre. Le gain est mesuré pour le transistor PNP de la Figure 62, dont le layout comporte huit doigts pour une surface de $100 \times 125 \mu\text{m}^2$, afin de pouvoir tenir les ESD. Pour obtenir la courbe à fort courant, une mesure en impulsion est indispensable, le transistor ne pouvant plus supporter une mesure statique sans être dégradé. Le principe consiste à polariser en continu le collecteur (par exemple avec une pile) et à appliquer une impulsion TLP sur la base. Le gain est calculé en relevant les courants de base et de collecteur à la fin de l'impulsion (Figure 66). La totalité de la courbe du gain n'a pas pu être reproduite en simulation. Les valeurs des durées de vie ne sont pas les mêmes pour simuler le gain maximal ou les variations à fort courant. Pour une simulation ESD, les valeurs les plus pertinentes sont celles à fort courant, soit 670 picosecondes pour les électrons et 300 picosecondes pour les trous (contre respectivement 10 nanosecondes et 3 nanosecondes par défaut dans le simulateur).

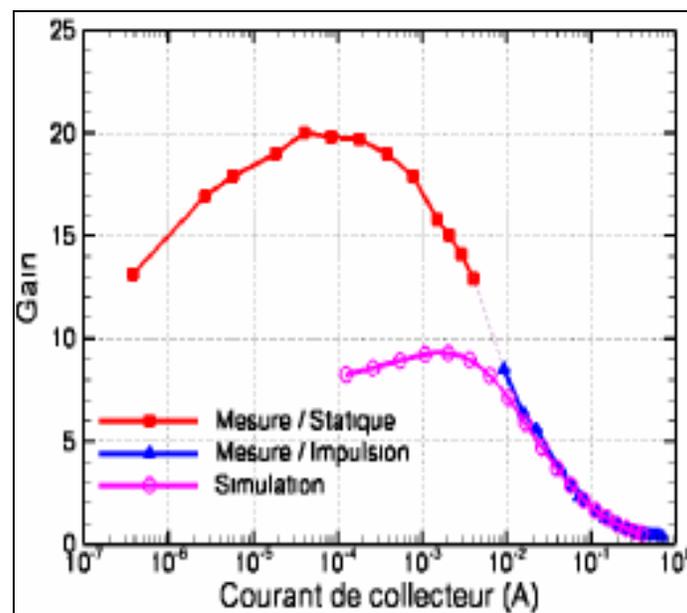


Figure 66: Mesure et simulation du gain en fonction du courant de collecteur pour le transistor PNP de la Figure 62 (mesure statique à faible courant et en impulsions à fort courant).

Les caractéristiques TLP du transistor PNP sont simulées pour les deux couples de durées de vie, extraits à partir du gain maximal où des variations à fort courant, et comparées à la mesure (Figure 67). La différence de tension entre ces deux simulations est comprise entre 3 et 5 Volts. Le calibrage des durées de vie a permis de rapprocher la simulation de la mesure, et en particulier de reproduire précisément la tension de maintien.

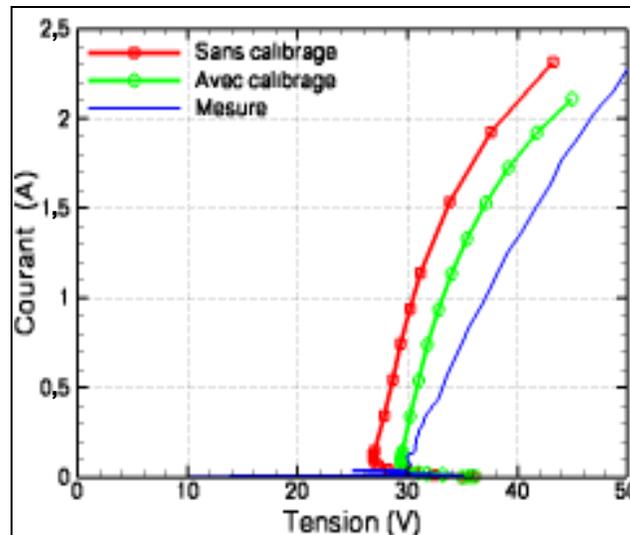


Figure 67: Caractéristique TLP mesurée et simulée (avec ou sans calibrage des durées de vie) du transistor PNP de la Figure 62 en configuration autopolarisée.

3.3. Réduction du R_{ON} des transistors PNP latéraux autopolarisés

Concernant les protections centralisées, certaines sont développées à base de transistors PMOS, présentant l'avantage d'être à la fois très compacts, car pouvant être dessinées aux dimensions minimales de la technologie, et très robustes, une bonne uniformité du courant étant garantie [72] [73]. Cependant, ces protections centralisées utilisent uniquement l'effet MOS, sans contribution de l'effet bipolaire du transistor PNP parasite. Dans le circuit de protection, les transistors PNP sont parfois mis à profit pour remplacer les diodes d'isolation [74]. D'une part, cette modification permet de réduire les risques de latch-up [75]. D'autre part, dans les technologies complètement siliciurées, la robustesse des transistors PNP reste garantie, contrairement au cas des transistors NPN et des thyristors [76]. Si, de manière standardisée, l'isolation vis-à-vis de V_{DD} est réalisée par une simple diode, la structure comprend un transistor PNP vertical parasite qui peut être mis à profit pour conduire une partie de la décharge [77] [78]. Concernant les protections localisées, l'évolution des technologies, en particulier la réduction des dimensions, est bénéfique pour les transistors PNP, au point qu'ils semblent pouvoir concurrencer les transistors NPN [58] [75] [79]. Néanmoins, seul un exemple de protection à base de transistors PNP a été identifié. Il s'agit d'une protection dédiée aux E/S "faibles tensions" d'une application pour l'ADSL, réalisée en technologies 0,35 et 0,25 μm [80]. Son développement a amené les auteurs à définir des règles d'optimisation du layout des transistors PNP [81]. Ils recommandent une structure interdigitée de longueurs d'émetteur et de collecteur minimales. Dans le cadre des applications "hautes tensions", il semble que les transistors PNP ne soient pas utilisés, ce qui laisse le champ libre pour la conception de nouvelles structures de protection.

L'étude théorique du chapitre 2 a montré que le R_{ON} est le paramètre le plus critique des protections "hautes tensions" à base de transistors PNP. Un premier ensemble de règles visant à le réduire a été défini en se basant sur un modèle unidimensionnel simple. Maintenant, l'objectif est de définir les règles vis-à-vis du layout (bi ou tridimensionnel) et des limitations technologiques. Ces aspects sont pris en compte pour les configurations de l'émetteur et du collecteur. Il faut ensuite valider les règles d'optimisation et évaluer les performances atteignables par des caractérisations sur silicium. Ces caractérisations comprennent les mesures des caractéristiques TLP et $vfTLP$, les tests de robustesses HBM et MM, et des analyses de défaillance.

3.3.1. Règles de dessin

Seules les structures latérales seront étudiées. La technologie SmartMOS 8 MV ne permet pas de réaliser des structures PNP verticales. Pour cela, il aurait fallu une couche enterrée dopée P, or seule une couche enterrée dopée N est disponible. Le principal point faible des structures latérales est leur faible robustesse si le courant se concentre près de la surface. Cependant, ce problème peut être contourné, des implantations P de forte énergie permettant d'obtenir des diffusions claquant en profondeur.

3.3.1.1. Configuration d'émetteur

Effet de la longueur d'émetteur (L_E) sur le comportement d'un transistor bipolaire latéral

La longueur de l'émetteur L_E est directement accessible par le layout, et peut ainsi être facilement ajustée. L'étude de son influence sur le comportement électrique est basée sur l'analyse du composant particulier représenté sur la Figure 62. La caractéristique TLP, pour un layout adapté aux ESD (huit doigts correspondant à une surface de $100 \times 125 \mu\text{m}^2$), est simulée pour différentes valeurs L_E (allant de 3,4 à 12,7 udm), tout en gardant inchangé le reste de la structure. A partir des caractéristiques obtenues, le R_{ON} , son facteur de mérite F_{RON} (14) et le rapport d'injection γ à la limite de la ZCE base-collecteur sont calculés au voisinage de 1,8 Ampère (Figure 68 et Figure 69). Pour le calcul du rapport d'injection, les densités de courants de trous et d'électrons sont intégrées le long d'une coupe effectuée le long de la ZCE base-collecteur.

Les résultats montrent une réduction du R_{ON} lorsque L_E diminue, sauf pour la plus faible valeur de L_E (3,4 udm) où le R_{ON} remonte. La réduction est à mettre en relation avec l'augmentation du rapport d'injection, tendant à affaiblir l'influence du courant de polarisation, influence défavorable au R_{ON} . D'autre part, la chute du rapport d'injection, également défavorable au R_{ON} , sera plus marquée si sa valeur est élevée, ce qui s'oppose à la réduction du R_{ON} et explique sa remontée pour L_E égale à 3,4 udm. Du point de vue de la surface, l'économie permise par une réduction de L_E est significative compte tenu des faibles dimensions d'un doigt. Une réduction de 12,7 à 3,4 udm représente 30 % de la surface totale. Cette tendance va dans le sens d'une diminution du facteur du mérite. Elle s'ajoute à la réduction du R_{ON} et compense l'augmentation au dernier point de sorte que le facteur de mérite diminue continûment lorsque L_E diminue.

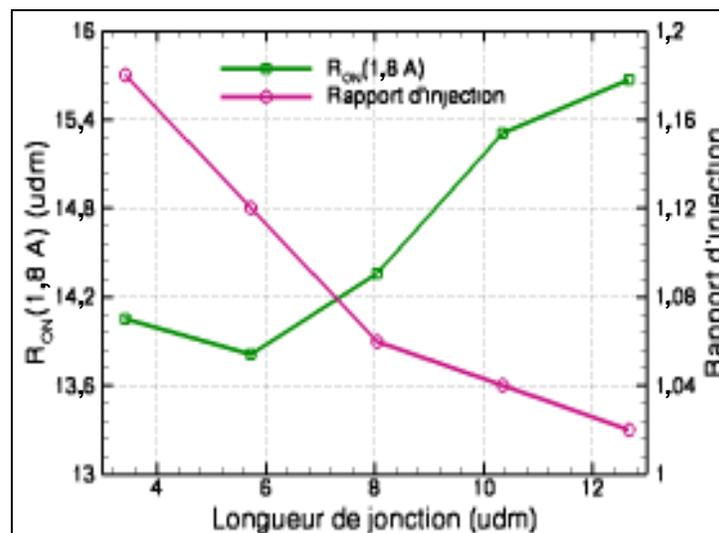


Figure 68: R_{ON} et rapport d'injection en fonction de la longueur de la jonction émetteur-base.

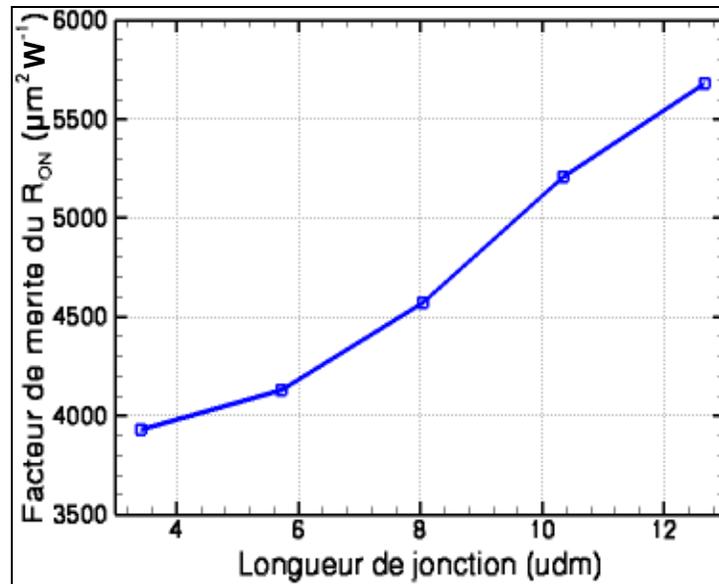


Figure 69: Facteur de mérite F_{RON} en fonction de la longueur de la jonction émetteur-base.

Pour expliquer l'augmentation du rapport d'injection avec la diminution de la longueur d'émetteur, sa valeur ainsi que celle de la densité de courant sont relevées le long de la jonction émetteur-base pour L_E égale à 3,4 (Figure 70) et 12,7 udm (Figure 71). Ces relevés montrent que le courant est réparti sur toute la jonction et que la densité de courant ne devient jamais négligeable, même si elle décroît avec l'éloignement par rapport au collecteur (qui est d'autant plus important vers la droite des figures 70 et 71). Aussi, toutes les valeurs locales du rapport d'injection, y compris celles pour les positions les plus éloignées du collecteur, contribuent significativement à la valeur totale, celle-ci étant proportionnelle à une intégrale pondérée par la densité de courant le long de la jonction.

Les relevés des figures 70 et 71 montrent que le rapport d'injection diminue avec l'éloignement par rapport au collecteur, passe par un minimum puis augmente. La diminution s'explique par un élargissement de la base interne, mis en évidence sur la Figure 72 représentant les bases internes pour trois chemins de courant issus de positions différentes sur la jonction émetteur-base. L'augmentation finale du rapport d'injection est à mettre en relation avec le taux de recombinaison élevé dans la région fortement dopée N^{++} sous la base. Aussi, elle ne survient que sur la partie de la jonction voisine du contact de base, ce qui est clairement visible sur la Figure 71. De plus, les porteurs injectés se recombinant dans la base, ils n'arrivent pas dans la ZCE base-collecteur et ne participent donc pas au R_{ON} du transistor bipolaire autopolarisé, d'après les mécanismes décrits au Chapitre 1. Dans ces conditions, la diminution du rapport d'injection le long de la jonction conduit à des valeurs locales minimales d'autant plus faibles que la longueur d'émetteur est grande. Pour L_E égale à 3,4 udm, le rapport d'injection minimal est de 1,6 à la position 1,9 udm, et pour L_E égale à 12,7 udm, il est de 1,1 à la position 8,2 udm. Cette tendance explique l'augmentation de la valeur totale du rapport d'injection avec la diminution de la longueur d'émetteur. A titre de remarque, les deux valeurs minimales relevées le long de l'émetteur sont supérieures aux valeurs totales correspondantes reportées sur la Figure 68, 1,18 et 1,02 pour L_E égale respectivement à 3,4 et 12,7 udm. Ceci s'explique par le fait que les valeurs locales sont calculées à la jonction émetteur-base et non dans la ZCE base-collecteur comme les valeurs totales. Aussi, les recombinaisons dans la base sont négligées. Or, ces recombinaisons augmentent avec la largeur de base, ce qui va également dans le sens d'une augmentation du rapport d'injection avec le rétrécissement de l'émetteur.

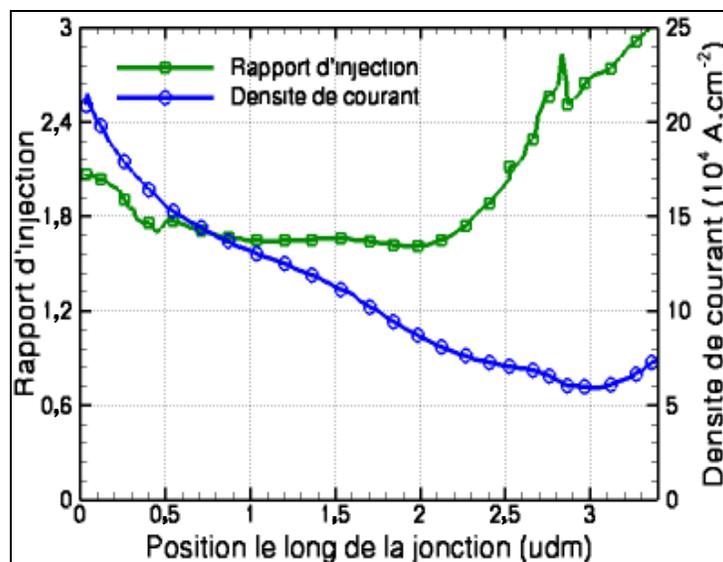


Figure 70: Rapport d'injection et densité de courant le long de la jonction émetteur-base pour un émetteur de longueur de 3,4 udm (la partie gauche du graphe est côté collecteur et la partie droite côté base).

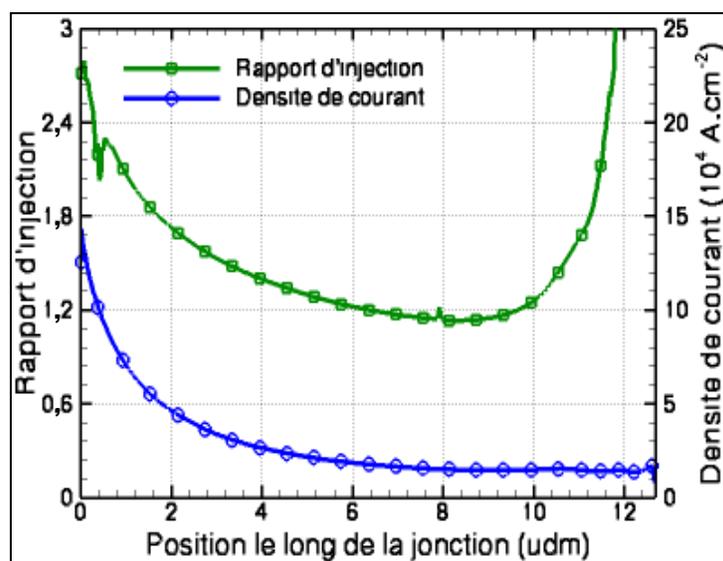


Figure 71: Rapport d'injection et densité de courant le long de la jonction émetteur-base pour un émetteur de longueur 12,7 udm.

Pour être complet, il est à noter que le rapport d'injection dans la région la plus proche du collecteur (à gauche des graphes des figures 70 et 71) est plus faible pour l'émetteur court, 2,1 de valeur maximale pour 3,4 udm contre 2,7 pour 12,7 udm. En effet, la valeur de la densité de courant étant plus élevée, $2,1 \cdot 10^5 \text{ A.cm}^{-2}$ pour 3,4 udm et $1,4 \cdot 10^5 \text{ A.cm}^{-2}$ pour 12,7 udm, la chute résultante du rapport d'injection est plus prononcée. Un tel comportement s'oppose à l'augmentation du rapport d'injection avec le rétrécissement de l'émetteur, mais sans être suffisant pour inverser cette tendance.

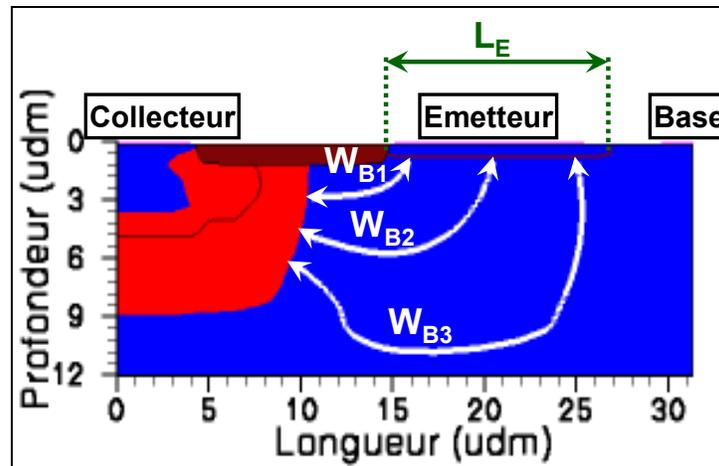


Figure 72: Distribution de la ZCE (en rouge) et bases internes le long de trois lignes de courant.

Cette étude est valable uniquement pour les transistors bipolaires latéraux. Elle ne s'applique pas aux transistors bipolaires verticaux. Pour ceux-ci, le rapport d'injection est quasi-indépendant de la longueur d'émetteur, car la largeur de base interne est constante, égale à la distance entre la diffusion d'émetteur et la couche enterrée, tel que le montre par exemple la coupe simplifiée de la Figure 22. Aussi, les règles d'optimisation du layout de l'émetteur seraient différentes.

Définition des règles de dessin et considérations sur le ballast

D'après l'étude qui vient d'être développée, la longueur d'émetteur L_E doit être minimale pour assurer les meilleures propriétés en termes de faible R_{ON} . Néanmoins, cette réduction de L_E risque de conduire à des densités de courant très élevées, qui peuvent être préjudiciables pour la robustesse, voire pour le R_{ON} si la base rétrécit (Chapitre2). Afin de limiter les densités de courant, il faut adopter un layout dont le périmètre d'émetteur en regard du collecteur est maximal, de manière à répartir le plus possible le courant. Cette règle conduira à privilégier les structures interdigitées. Les structures matricielles, faites d'une juxtaposition d'éléments hexagonaux ou circulaires, sont également bien adaptées; mais elles ne seront pas utilisées, car difficiles à dessiner. Ces règles sont en accord avec celles définies dans la référence [81] pour des technologies CMOS 0,35 et 0,25 μm , où une longueur d'émetteur minimale et une structure interdigitée sont recommandées.

Le choix d'un L_E minimal appelle à des commentaires quant à la très faible résistance de ballast qui en résulte. Cette résistance, correspondant au chemin d'accès du contact à la jonction émetteur-base, permet d'homogénéiser le courant, à la fois entre les doigts d'une structure interdigitée et le long d'un doigt. Elle joue ainsi un rôle important pour prévenir des défaillances prématurées. Dans le cas d'une structure interdigitée avec repliement, le nombre de doigts qui conduisent le courant après le repliement est aléatoire, et, le plus souvent, un seul doigt conduit. Pour assurer le déclenchement des autres doigts, il faut que la tension remonte jusqu'à la tension de déclenchement, avant la dégradation de la structure [82] [83]. Au regard de cette condition, la surtension générée par la résistance de ballast a un effet bénéfique. Cependant, les transistors PNP se repliant faiblement, la tension de déclenchement devrait facilement être atteinte, même sans résistance de ballast. De plus, si le repliement est supprimé, par exemple en utilisant une polarisation en base flottante, alors le déclenchement de tous les doigts est assuré, sans restriction. Concernant la conduction d'un seul doigt, la résistance de ballast réduit les risques liés aux instabilités thermiques [84]. Ces instabilités sont dues à l'action positive de l'autoéchauffement sur le gain. Une augmentation de la température conduit à une augmentation du gain et donc à une augmentation de la densité de courant, qui contribue à son tour à faire augmenter la température. A partir de légères inhomogénéités de la conduction, ce mécanisme peut mener à une focalisation du courant, très pénalisante pour la robustesse [37]. Le rôle de la résistance de ballast est de s'opposer à cet effet, en dépolarisant la jonction émetteur-base dans les régions où la densité de courant a tendance à augmenter. Cependant, les valeurs des gains des

transistors PNP pendant une ESD sont, en général, trop faibles pour générer ces instabilités. Ainsi, la suppression des résistances de ballast des transistors PNP ne devrait pas dégrader la robustesse, leurs caractéristiques intrinsèques permettant d'assurer à la fois le déclenchement de tous les doigts et la conduction homogène d'un doigt. Au regard de ce comportement, les transistors PNP sont plus avantageux que les transistors NPN, ces derniers nécessitant le plus souvent une résistance de ballast.

3.3.1.2. Configuration de collecteur

Considérations préliminaires sur la résistance d'accès et les effets en fort courant

Le dopage de la diffusion de collecteur doit être modéré, pour que des tensions élevées puissent être soutenues à la jonction base-collecteur. En technologie smart power, les niveaux de dopage des diffusions "hautes tensions" sont compris entre 10^{16} et 10^{18} cm^{-3} . Pour de telles valeurs, les règles de dessin devront tenir compte de la résistance d'accès et des effets en fort courant. Concernant la résistance d'accès, la résistivité en conduction ohmique est donnée par :

$$\rho = \frac{1}{qN\mu} \quad (85)$$

Où N est le niveau de dopage et μ la mobilité des porteurs. L'application numérique pour une diffusion P dopée à 10^{17} cm^{-3} donne une résistivité d'environ $0,1 \Omega \cdot \text{cm}$, soit une résistance d'accès de l'ordre de l'Ohm, pour des dimensions usuelles. Concernant les effets en fort courant, un calcul simple montre que le dopage peut être trop faible pour assurer la conduction des fortes densités de courant générées par une ESD. La densité de courant maximale que peut fournir un dopage est donnée lorsque la vitesse de saturation des porteurs est atteinte :

$$J_{Max} = qNv_{sat} \quad (86)$$

Où v_{sat} est la vitesse de saturation des porteurs. Au-delà de cette limite, une génération de porteurs est indispensable pour assurer la conduction [85]. Pour un dopage de 10^{17} cm^{-3} , l'application numérique donne des densités de courant maximales J_{Max} d'environ 10^4 $\text{A} \cdot \text{cm}^{-2}$, valeur qui sera dépassée lors d'une ESD. Afin de décrire les phénomènes apparaissant lorsque la densité de courant devient supérieure à J_{Max} , une structure unidimensionnelle $P^+/P^-/P^+$ va être étudiée à fort courant. Ensuite, cette étude sera mise à profit pour décrire le comportement de la diffusion de collecteur, et ainsi être à même de définir des règles de dessin pour la réduction du R_{ON} .

Etude à fort courant d'une structure unidimensionnelle $P^+/P^-/P^+$

Une structure représentative de la diffusion de collecteur est définie. Elle est constituée d'une région de $2 \mu\text{m}$ dopée à 10^{17} cm^{-3} , comprise entre deux régions dopées à $5 \cdot 10^{21}$ cm^{-3} . Sa caractéristique J-V simulée (Figure 73) montre trois modes de fonctionnement, correspondant à trois types de conduction distincts : un régime ohmique à bas courant, puis un régime saturé menant à une forte surtension (27 Volts) et, lorsque J_{Max} est dépassée, une conduction caractérisée par une résistance dynamique négative ou très faible. Pour identifier les mécanismes physiques mis en jeu, la distribution du champ électrique est relevée dans les trois modes de fonctionnement (Figure 74), et, pour le mode au-delà de J_{Max} , les densités d'électrons et de trous, et le taux d'ionisation par impact sont également relevés (Figure 75). Le passage du régime ohmique au régime saturé conduit à la formation d'un champ électrique sensiblement constant à travers la région P. Etant donné que le champ diverge lorsque la vitesse des porteurs s'approche de la vitesse de saturation, il augmente fortement pour les densités de courant proches de J_{Max} , d'où la surtension. Son augmentation se poursuit jusqu'à ce que le champ critique soit atteint. L'avalanche fournit alors les porteurs supplémentaires nécessaires pour assurer la

conduction du courant (Figure 75). La densité de porteurs ainsi générés est suffisante pour que la conduction ohmique se rétablisse dans une partie de la région P⁻, où le champ redevient quasi-nul. Il en résulte une diminution de la tension, d'où la résistance dynamique négative.

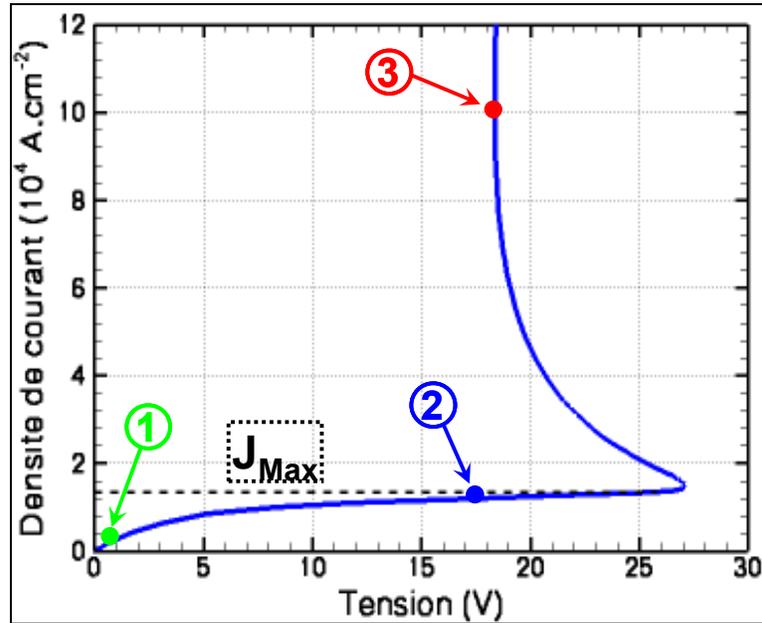


Figure 73: Caractéristique J-V de la structure P⁺/P⁻/P⁺.

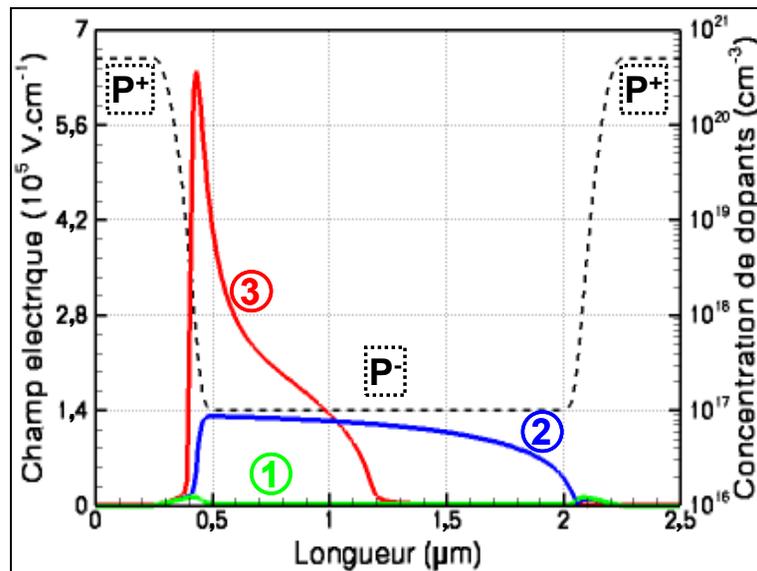


Figure 74: Distributions champ électrique en régime ohmique (1), en régime saturé (2) et pour $J > J_{Max}$ (3).

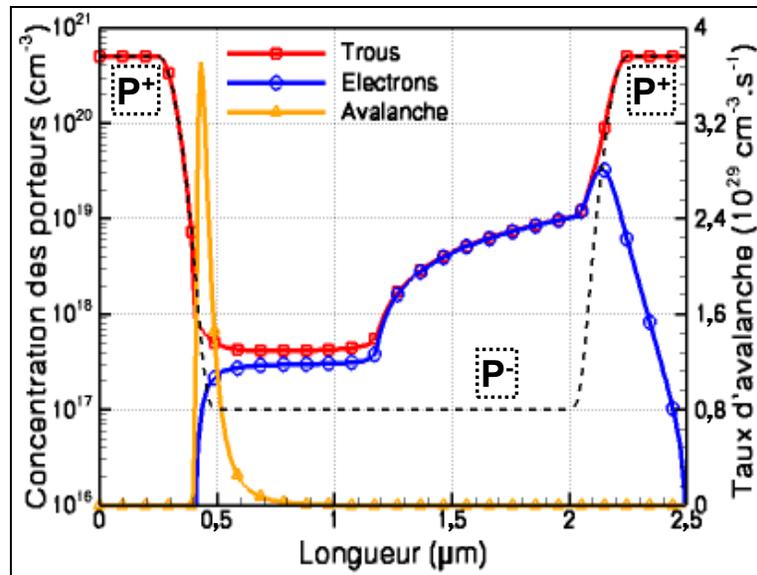


Figure 75: Distributions des concentrations de trous et d'électrons et du taux d'avalanche pour $J > J_{Max}$ (3).

Définition des règles de dessin relatives à la diffusion de collecteur

Si la conduction électrique dans la diffusion de collecteur entre en saturation, il s'y forme une nouvelle région de fort champ électrique, que nous avons nommée champ de "conduction" (Figure 76). Il en résulte une surtension, dont la valeur peut être de plusieurs dizaines de Volts, si l'on se base sur les résultats qui viennent d'être obtenus. Ce phénomène conduit à une dégradation du R_{ON} dans des proportions incompatibles avec la plupart des fenêtres de conception "hautes tensions". De plus, l'apparition d'une résistance négative rend le fonctionnement instable, ce qui crée des risques de focalisation du courant. Si une telle focalisation se produit, les très fortes densités de courant générées conduisent de manière quasi-systématique à la défaillance de la protection, soit par fusion du silicium, soit par perçage du contact de collecteur [85].

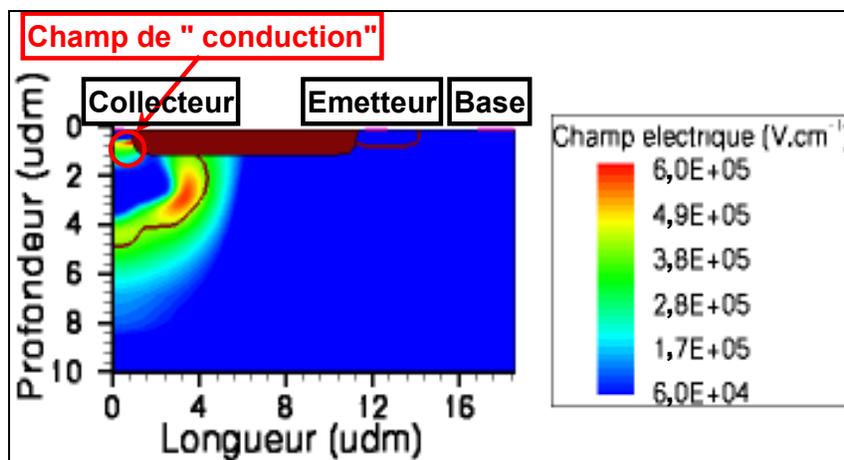


Figure 76: Distribution de champ électrique à 2,3 Ampères dans le transistor PNP de la Figure 62, mettant en évidence une zone de fort champ électrique dans la diffusion de collecteur (champ de "conduction").

Ainsi, les règles de dessin concernant le collecteur doivent permettre d'éviter l'apparition de cette nouvelle région de fort champ électrique. Une première méthode consiste à augmenter J_{Max} , en dopant au maximum la diffusion de collecteur. Pour cela, tous les dopages P offerts par la technologie seront implantés. Comme seule précaution à respecter, il faut fixer la tension de claquage avec un seul

dopage, pour avoir une bonne précision, et donc garantir une marge suffisante entre les masques de tous les autres dopages et la jonction base-collecteur. Une seconde méthode consiste à augmenter la longueur du contact de collecteur (L_C sur la Figure 62), de manière à répartir le courant sur une plus grande surface et ainsi faire baisser la densité de courant. Pour éviter d'augmenter inutilement la surface, il faut choisir la plus petite valeur de L_C assurant qu'au plus haut courant de la fenêtre de conception, la densité de courant correspondante est suffisamment faible pour ne pas générer de surtension. Par ailleurs, ces règles, sur le dopage et sur L_C , sont également favorables pour faire diminuer les résistances d'accès. Pour les technologies CMOS 0,35 et 0,25 μm traitées dans la référence [81], la règle sur la longueur de collecteur consistait simplement à la choisir minimale. Cette différence par rapport à la technologie SmartMOS 8 MV s'explique probablement par les niveaux de dopage plus élevés des technologies CMOS comparé à ceux des technologies SmartPower.

3.3.1.3.Synthèse des règles de dessin

Les règles définies aux chapitres 2 et 3, applicables aux structures latérales (du type représenté sur la Figure 62), sont listées ci-dessous.

- Emetteur
 - Implantation profonde (W_E élevé de l'étude 1D)
 - Longueur L_E égale au minimum autorisé par la technologie (fixée par l'extension de l'implantation de surface nécessaire pour pouvoir prendre un contact)
 - Périmètre d'émetteur en regard de collecteur maximal (structure interdigitée ou matricielle)
 - Dopage élevé, compris entre 10^{18} et 10^{20} cm^{-3}
- Collecteur
 - Longueur L_C suffisamment grande pour éviter la formation d'un champ électrique dans la diffusion de collecteur
 - Dopage maximal
- Base
 - Suppression du contact (fonctionnement en base ouverte)
 - Base étroite (faible W_B de l'étude 1D)
 - Dopage faible, inférieur à 10^{17} cm^{-3}
- Profil de dopage à la jonction base-collecteur
 - Fort dopage de collecteur à profil abrupt
 - Faible dopage de base

3.3.2.Caractérisations

3.3.2.1.Description des structures testées

Les caractérisations sur silicium doivent permettre d'évaluer les performances des protections à base de transistors PNP, réalisées en technologie SmartMOS 8 MV. Pour cela, les structures testées sont conçues suivant un schéma général (Figure 77) intégrant les principales règles de dessin, et à partir duquel des variations seront effectuées pour évaluer l'effet de certains paramètres. Il s'agit de structures interdigitées, composées de treize doigts couvrant une surface d'environ $100*100 \mu\text{m}^2$, soit de l'ordre de grandeur de celle d'un plot d'E/S. Concernant l'émetteur, la longueur L_E est minimale, en particulier aucun ballast n'est ajouté, le dopage est très élevé, de valeur maximale voisine de 10^{20} cm^{-3} , par contre la diffusion est peu profonde, la technologie ne disposant pas d'implantations P à la fois profonde et fortement dopée. Concernant le collecteur, toutes les implantations P disponibles sont implantées sous le contact. Concernant la base, le contact est supprimé, le dopage est faible, de l'ordre

de 10^{16} cm^{-3} , et la base effective au claquage est étroite. La largeur de base effective est fixée par la marge prise sur la base interne pour éviter le perçage entre l'émetteur et le collecteur. Si cette condition était réalisée, l'émetteur et le collecteur seraient court-circuités, et le courant augmenterait rapidement. Pour déclencher une protection ESD, le claquage par avalanche est préféré au perçage, car plus précis (la tension de claquage dépend uniquement du dopage à la jonction base-collecteur alors que la tension de perçage dépend non seulement du dopage base-collecteur mais aussi la distance entre les deux jonctions). De plus, les fuites augmentent fortement lorsque la ZCE s'approche de la jonction émetteur-base. Enfin, le niveau de tension influe fortement sur la taille de la structure, via l'extension de la ZCE base-collecteur nécessaire pour tenir des "hautes tensions" (Figure 78).

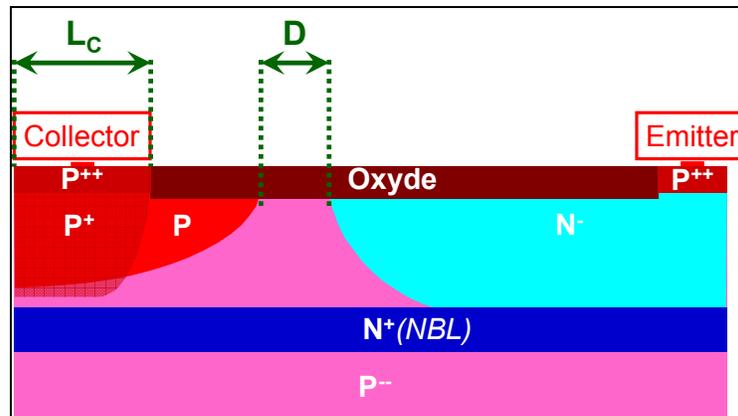


Figure 77: Coupe d'un demi-doigt de la structure type retenue pour les caractérisations sur silicium.

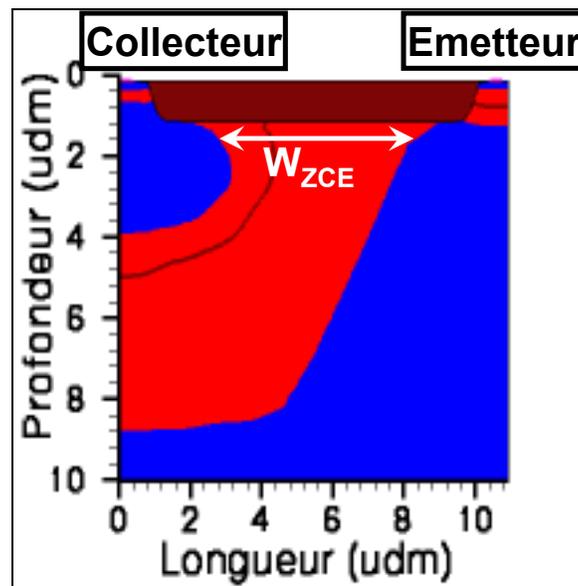


Figure 78: Distribution de la ZCE (en rouge) au claquage statique du composant DUT1.

Deux règles de dessin n'ont pas été directement appliquées : celle sur le profil de dopage base-collecteur et celle sur la longueur de collecteur. Des mesures sont indispensables pour évaluer les effets de ces deux paramètres, la précision des simulations n'étant pas suffisante pour en déterminer les valeurs optimales. Le profil de dopage base-collecteur est contrôlé par la distance D entre les bords des masques de l'implantation P de collecteur et de l'implantation N^- de base. L'idée est d'ajuster la concentration de dopage au voisinage de la jonction à l'aide de la diffusion latérale du N^- , tout en ne modifiant que légèrement le profil du dopage P , dont le niveau est nettement supérieur à celui du N^- .

On s'attend à ce que le facteur de mérite du R_{ON} augmente avec la distance D . Parallèlement, la modification du dopage entraîne des variations de la tension de claquage et, via l'extension de la ZCE, des variations de la taille de la structure. Quatre composants sont réalisés : dans le premier, nommé DUT1, l'implantation N n'est pas masquée, pour les trois autres, nommés DUT2, DUT3 et DUT4, les distances D sont respectivement 0,75 udm, 3,75 udm et 6,75 udm. Pour ces quatre composants, la longueur de collecteur L_C correspond à la taille nécessaire pour poser deux rangées de contacts. Afin de caractériser l'effet de la longueur de collecteur, un composant supplémentaire, nommé DUT5, est réalisé à partir du composant DUT4, dont la longueur L_C correspond à une seule rangée de contacts. Les surfaces de ces cinq composants sont listées dans le Tableau 9.

Tableau 9: Surfaces des composants testées.

	DUT1	DUT2	DUT3	DUT4	DUT5
S (μm^2)	$1,00 \cdot 10^4$	$1,07 \cdot 10^4$	$1,32 \cdot 10^4$	$1,57 \cdot 10^4$	$1,48 \cdot 10^4$

3.3.2.2. Variation du profil de dopage base-collecteur

Les structures étant en configuration base flottante, il a été vérifié au préalable que les augmentations des courants de fuite ne sont pas critiques. Leurs niveaux se sont révélés être inférieurs au bruit de la mesure (10 picoAmpères), soit nettement au-dessous de la plupart des spécifications, généralement de l'ordre de 1 nanoAmpère. Ce résultat est à mettre en relation avec les faibles gains de transistors PNP.

Caractéristiques TLP et v_f TLP

Les caractéristiques TLP des composants DUT1, DUT2, DUT3 et DUT4 sont présentées en Figure 79, et les principaux paramètres électriques dans le Tableau 10. Comme attendu, la polarisation en base flottante conduit à des caractéristiques sans repliement, et l'utilisation de transistors PNP permet d'atteindre des tensions de déclenchement (V_i) élevées, comprises entre 25 et 50 V. Les valeurs des courants au second claquage (I_{t2}), proches de 4,5 Ampères, indiquent une bonne robustesse, d'au moins 6 kV HBM en se basant la correspondance entre I_{t2} et le courant maximum de la forme d'onde HBM (2 kV \Leftrightarrow 1,33 A). Par ailleurs, les courants I_{t2} des quatre structures sont semblables alors que les tensions V_{t2} vont de 54 à 71 V, ce qui indique une défaillance liée au courant et non à l'énergie. L'apparition d'un champ électrique dans la diffusion de collecteur étant le principal mécanisme de dégradation dépendant uniquement du courant (3.3.1.2), il est probable qu'elle soit à l'origine de la défaillance.

Pour l'étude de la résistance à l'état passant, trois valeurs sont relevées : une valeur moyenne R_{ON} calculée entre le déclenchement et le second claquage, une valeur à bas courant (800 milliAmpères) R_{ON1} sur laquelle les effets thermiques et de fort courant ont une influence modérée et une valeur à fort courant (2,5 Ampères) R_{ON2} où les effets thermiques et de fort courant sont nettement prononcés. Les facteurs de mérite correspondants F_{RON} , F_{RON1} et F_{RON2} sont calculés, avec la formule (14), afin de pouvoir comparer les quatre structures indépendamment des variations sur les tensions et les surfaces. La tendance générale est une réduction des facteurs de mérite avec l'augmentation de la distance D , ce qui confirme l'influence bénéfique de la modification du profil de dopage base-collecteur. En particulier, le composant DUT4 a un excellent facteur de mérite pour R_{ON2} (12,5), environ trois fois inférieur à celui de DUT1. Une amélioration aussi exceptionnelle met en évidence la puissance de l'optimisation par le profil de dopage base-collecteur. Par ailleurs, les facteurs de mérite sont nettement meilleurs que celui de la protection standard pour les E/S 80 Volts (égal à 52), ce qui montre sans ambiguïté l'efficacité des règles d'optimisation définies.

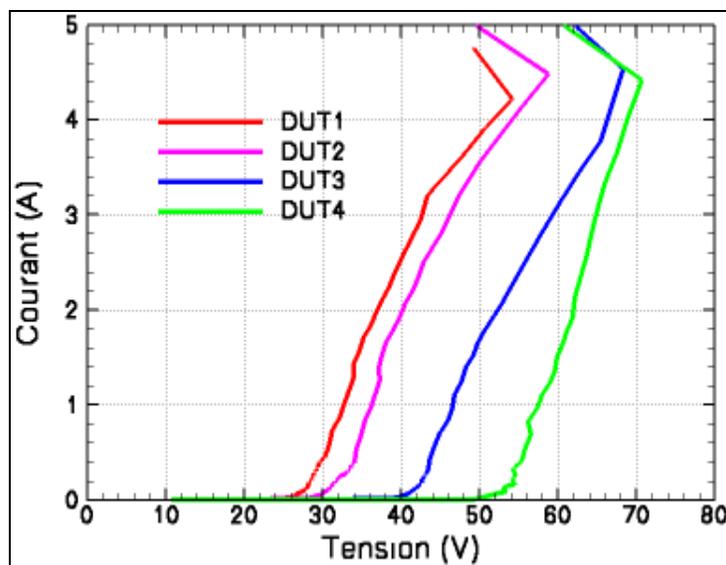


Figure 79: Caractéristiques TLP des composants DUT1-2-3-4 correspondant à une distance D croissante.

Tableau 10: Paramètres électriques extraits des caractéristiques TLP.

	DUT1	DUT2	DUT3	DUT4
V_t (V)	24	28	39	49
V_{I_2} (V)	54,2	58,8	68,3	70,7
I_{I_2} (A)	4,22	4,48	4,53	4,42
F_{I_2} (mW. μm^{-2})	22,7	24,7	23,4	19,9
R_{ON} (Ω)	7,1	6,9	6,5	4,9
F_{RON} ($\mu\text{m}^2.\text{W}^{-1}$)	46,0	39,5	29,9	21,4
R_{ON1} (Ω)	4,3	4,3	4,4	3,8
F_{RON1} ($\mu\text{m}^2.\text{W}^{-1}$)	42,8	36,7	27,6	18,8
R_{ON2} (Ω)	5,8	6,2	7,0	3,2
F_{RON2} ($\mu\text{m}^2.\text{W}^{-1}$)	35,7	36,2	30,1	12,5

Afin de déterminer le comportement des transistors PNP vis-à-vis d'agressions très brèves, telles que les décharges CDM, la caractéristique vTLP du composant DUT4 est mesurée (Figure 80), pour des impulsions de 5 nanosecondes et des temps de montée de 175 ps. Les principaux paramètres électriques sont relevés (Tableau 11), en gardant les mêmes définitions de R_{ON} , R_{ON1} et R_{ON2} que pour la caractéristique TLP. A bas courant, les caractéristiques TLP et vTLP se superposent. On en déduit que le fonctionnement électrique est stabilisé à la fin de l'impulsion du vTLP, ce qui montre que le temps de déclenchement est compatible avec les exigences d'une protection vis-à-vis des décharges CDM. La différence de quelques Volts sur les tensions de déclenchement n'est pas significative, compte tenu de l'imprécision des mesures à très bas courant. Au-delà de 1,5 Ampères, les deux caractéristiques se séparent, la résistance passante en TLP étant pénalisée par les effets thermiques qui commencent à prendre de l'ampleur. A niveau de puissance égal, le rapport des énergies dissipées entre le TLP et le vTLP correspond au rapport des durées des impulsions, soit une valeur de 20. Aussi, l'élévation de température en vTLP est toujours négligeable comparée au TLP. La dégradation du R_{ON} en TLP est de plus en plus prononcée et peut devenir très importante, pour un courant de 2,5 Ampères, la comparaison avec le facteur de mérite en vTLP fait apparaître un facteur 1,5. Concernant la défaillance, le facteur de mérite de la robustesse F_{I_2} est deux fois plus grand en vTLP. Cependant, cette amélioration reste très loin du rapport 20 entre les énergies, ce qui va dans le sens d'une défaillance initialisée par une forte densité de courant et non par une forte énergie.

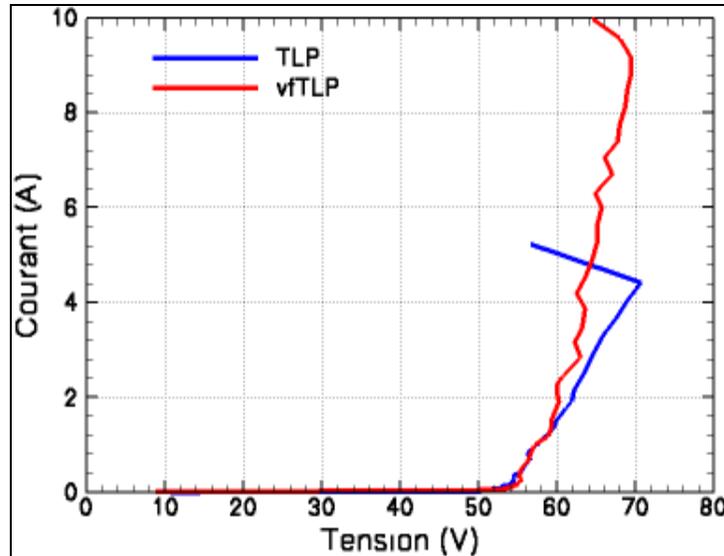


Figure 80: Comparaisons des caractéristiques TLP et vfTLP mesurées pour le composant DUT4.

Tableau 11: Paramètres électriques de la caractéristique vfTLP du composant DUT4.

	vfTLP	TLP
V_t (V)	53	49
V_{I_2} (V)	69,0	70,7
I_{I_2} (A)	9,2	4,42
F_{I_2} ($\text{mW} \cdot \mu\text{m}^{-2}$)	40,4	19,9
R_{ON} (Ω)	1,8	4,9
F_{RON} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	7,6	21,4
R_{ON1} (Ω)	3,8	3,8
F_{RON1} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	18,5	18,8
R_{ON2} (Ω)	2,1	3,2
F_{RON2} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	8,92	12,5

Tests de robustesse et analyse de défaillance

Des tests de robustesse HBM et MM ont été réalisés (Tableau 12). Les résultats HBM sont excellents, seul le composant DUT1 a été dégradé, pour les trois autres composants la limite du testeur (10 kV) a été atteinte avant la défaillance. La différence avec les 6 kV prédits à partir du I_{I_2} s'explique par une surestimation des effets thermiques en TLP, la dissipation de l'énergie au cours du temps étant défavorable comparée à une impulsion HBM. De très bons résultats MM ont également été obtenus. En particulier, les pics de courant en polarité inverse du stress MM n'entraînent pas de défaillance prématurée.

Tableau 12: Résultats des tests de robustesse HBM et MM sur les composants DUT1-2-3-4.

	DUT1	DUT2	DUT3	DUT4
V_{HBM} (kV)	9,5	≥ 10	≥ 10	≥ 10
V_{HBM}/S ($\text{mV} \cdot \mu\text{m}^{-2}$)	950	≥ 930	≥ 760	≥ 640
V_{MM} (V)	350	400	500	650
V_{MM}/S ($\text{mV} \cdot \mu\text{m}^{-2}$)	35	37	38	41

Une analyse de défaillance a été réalisée sur le composant DUT1, dégradé par le test HBM. Une simple photographie sous microscope montre une fusion localisée du contact de collecteur (Figure

81). Cette observation confirme l'hypothèse selon laquelle la défaillance est induite par le champ électrique créé dans la diffusion de collecteur. La zone de fusion étant localisée à l'extrémité d'un doigt, la terminaison de la jonction est probablement un point faible. Il serait donc possible d'améliorer la robustesse en travaillant sur cet aspect, par exemple en adoptant un layout semi-circulaire et en introduisant un ballast entre la jonction et le contact.

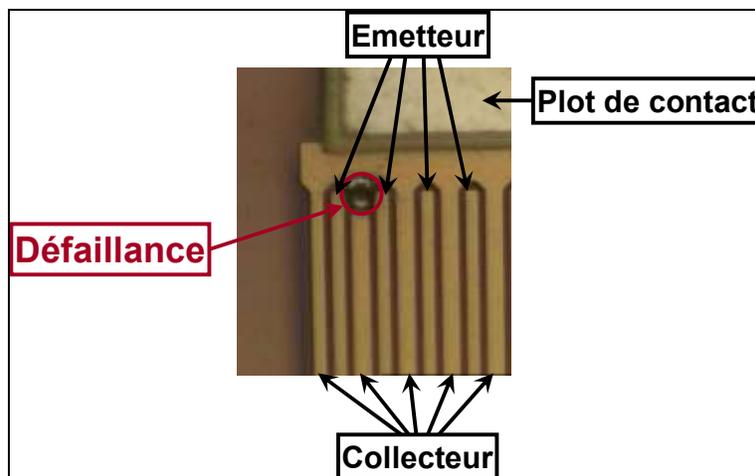


Figure 81: Photographie au microscope optique du composant DUT1 montrant une dégradation du contact de collecteur.

3.3.2.3. Variation de la longueur de collecteur

Caractéristiques TLP

Les effets des fortes densités de courant dans le collecteur sont évalués en comparant les caractéristiques TLP des composants DUT5 et DUT4 (Figure 82) (Tableau 13). La seule différence entre ces deux composants est le contact de collecteur, qui est plus étroit pour DUT5. Ainsi, celui-ci voit des densités de courant plus élevées, de sorte que l'apparition du champ électrique devrait survenir à plus faible courant. Jusqu'à un 1,7 Ampères, les deux caractéristiques TLP sont identiques, et le facteur de mérite à 800 milliAmpères, F_{RON1} , meilleur d'environ 10 % pour DUT5 comparé à DUT4, du fait du gain sur la surface du collecteur. Au-dessus de 1,7 Ampères, la résistance passante de DUT5 se dégrade fortement ce qui se traduit par une différence de 65 % sur son facteur de mérite à 2,5 Ampères, F_{RON2} , comparé à celui de DUT4. Cette dégradation de la résistance passante s'explique très bien par la surtension liée à l'apparition du champ électrique dans la diffusion de collecteur de DUT5, en accord avec la prédiction théorique. Concernant le choix pour réaliser une protection ESD, tant que le courant maximal imposé par la spécification de robustesse reste inférieur à 1,7 Ampères (correspondant à environ 3 kV HBM), DUT5 sera privilégié, par contre, à plus fort courant, il est impératif d'utiliser DUT4. Par ailleurs, le facteur de mérite sur la robustesse F_{I2} de DUT5 est nettement inférieur à celui de DUT4, d'environ 35 %, ce qui donne une confirmation supplémentaire d'une dégradation liée aux effets de fortes injections dans le collecteur.

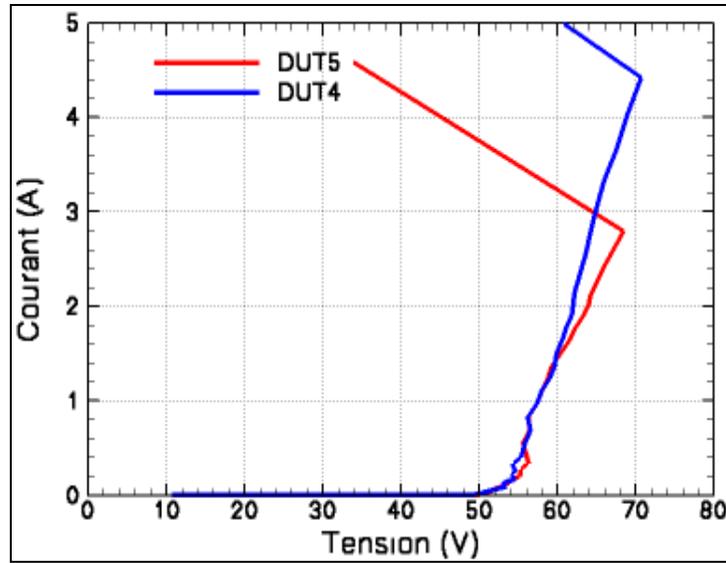


Figure 82: Caractéristiques TLP des composants DUT4 et DUT5.

Tableau 13: Paramètres électriques de la caractéristique TLP du composant DUT5.

	DUT5	DUT4
V_t (V)	50	49
V_{I2} (V)	68,4	70,7
I_{I2} (A)	2,79	4,42
F_{I2} (mW. μm^{-2})	12,9	19,9
R_{ON} (Ω)	6,6	4,9
F_{RON} ($\mu\text{m}^2.\text{W}^{-1}$)	27,9	21,4
R_{ON1} (Ω)	3,6	3,8
F_{RON1} ($\mu\text{m}^2.\text{W}^{-1}$)	16,8	18,8
R_{ON2} (Ω)	6,1	3,2
F_{RON2} ($\mu\text{m}^2.\text{W}^{-1}$)	20,7	12,5

Tests de robustesses et analyse de défaillance

Les robustesses HBM et MM de DUT5 sont mesurées (Tableau 14). Conformément à ce qui est attendu, les valeurs obtenues sont significativement inférieures à celles de DUT4. La photographie sous microscope montre à nouveau une fusion localisée du contact de collecteur à l'extrémité d'un doigt (Figure 83), indiquant un mécanisme de défaillance similaire pour les deux tailles de contacts de collecteur testées.

Tableau 14: Résultats des tests de robustesses HBM et MM sur le composant DUT5

V_{HBM} (kV)	7,0
V_{HBM}/S (mV. μm^{-2})	470
V_{MM} (V)	350
V_{MM}/S (mV. μm^{-2})	24

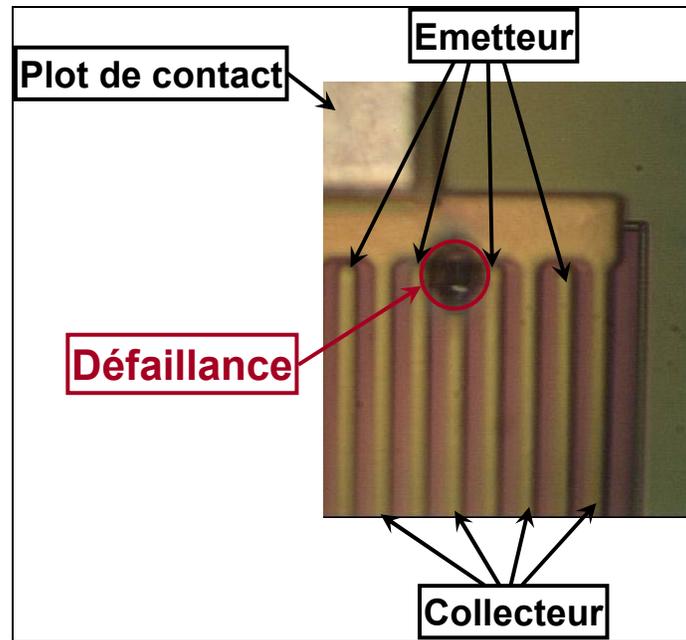


Figure 83: Photographie au microscope optique du composant DUT5 montrant une dégradation du contact de collecteur.

3.4. Couplage d'un transistor PNP latéral avec une diode verticale

Pour tenter de réduire le R_{ON} au-dessous de la limite atteignable avec un transistor PNP seul, une méthode originale consiste à coupler le transistor avec une source de courant lui fournissant une partie de son courant de polarisation. Ainsi, l'avalanche à la jonction base-collecteur peut diminuer, ce qui entraîne une décroissance du facteur de multiplication, et donc de la tension, bénéfique pour le R_{ON} . Dans un premier temps, le principe de fonctionnement sera décrit avec précision. Cette description montrera qu'une diode à avalanche verticale est particulièrement bien adaptée pour contribuer au courant de polarisation. Ensuite, le comportement des diodes à avalanche pendant une ESD sera étudié, de manière à disposer des informations indispensables pour être à même d'optimiser le couplage avec le transistor PNP. Par la suite, des règles d'optimisation concernant l'ajustement des tensions de fonctionnement et des dimensions de la diode seront établies. Enfin, des caractérisations sur silicium feront apparaître des performances exceptionnelles vis-à-vis du R_{ON} . En particulier, une structure de protection d'une E/S 80 Volts, deux fois plus petite que la solution existante, sera développée.

3.4.1.Principe de fonctionnement

3.4.1.1.Description théorique et choix d'une diode à avalanche verticale comme nouvelle source de courant

Les deux principales contraintes concernant la nouvelle source de courant sont qu'elle ne doit pas entraîner de repliement et qu'elle doit être active quand le transistor bipolaire est à l'état passant. Par

exemple, un thyristor peut être vu comme un transistor NPN et un transistor PNP couplés, le courant d'injection de l'un fournissant le courant de polarisation de l'autre. Il en résulte d'excellentes propriétés vis-à-vis du R_{ON} , par contre, le repliement conduit à des tensions très basses, de l'ordre de quelques Volts [86]. Pour éviter un tel comportement, il faut qu'un courant d'avalanche reste indispensable au fonctionnement. Dans cet esprit, il a été proposé un transistor bipolaire NPN à l'intérieur duquel deux chemins de courant, un latéral et un vertical, coexistent [29]. Cependant, si les deux chemins sont activés en même temps, les effets en fort courant suivant le chemin latéral conduisent à un repliement de la tension entraînant une réduction du courant dans le chemin vertical, qui finit par disparaître. Cette structure ne permet pas d'établir un couplage stable entre les chemins de courant, et n'est donc d'aucun intérêt pour réduire le R_{ON} ou même augmenter la robustesse. L'idée que nous avons développée s'appuie sur la possibilité de supprimer le repliement des transistors PNP. Si la source de courant contribuant au courant de polarisation est elle-même réalisée à l'aide une structure sans repliement (diode ou PNP), il est aisé de garantir que les deux courants, celui du PNP et de la source de courant, augmentent simultanément.

Concernant la réalisation de la structure, la technologie SmartMOS 8 MV offre la possibilité d'intégrer un transistor PNP latéral avec une diode verticale dans un même volume de silicium. (Figure 84). Dans cette configuration, la base et le collecteur du transistor sont respectivement l'anode et la cathode de la diode.

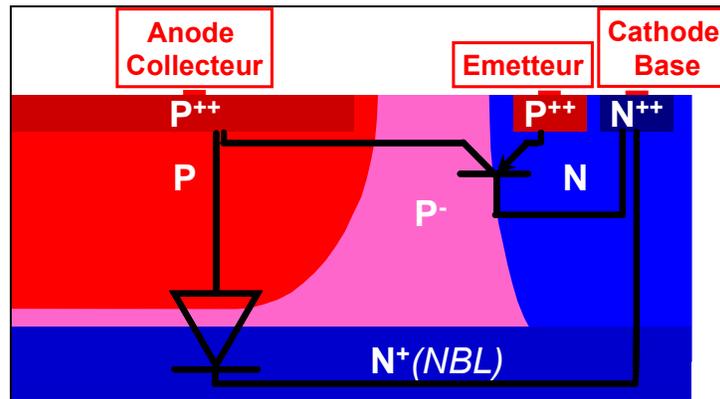


Figure 84: Coupe technologique montrant un transistor PNP latéral couplé à une diode verticale.

3.4.1.2. Etude en simulation

Une étude en simulation est effectuée afin d'illustrer le couplage entre le transistor PNP et la diode. Pour que les résultats soient représentatifs de ce fonctionnement, la structure simulée (Figure 85) est dessinée de manière à assurer une contribution significative de la diode au courant de polarisation du transistor. Dans cette optique, une tension de maintien du transistor PNP supérieure à la tension de claquage de la diode est garantie, de telle sorte que la diode ne se retrouve pas bloquée suite au repliement du transistor, et la jonction verticale est suffisamment étendue pour que le courant de diode soit important. En première approche, le fonctionnement peut être caractérisé sans qu'il soit nécessaire d'ajuster ces paramètres avec précision. Les règles de dessin pour l'optimisation du couplage seront définies ultérieurement à la section 3.4.3. Pour simuler la caractéristique TLP, la polarisation en direct de la jonction émetteur-base est favorisée par une résistance de 1 k Ω insérée entre la base et l'émetteur. La caractéristique obtenue (Figure 86) valide l'aspect de la description du principe de fonctionnement relatif au repliement, celui-ci étant très faible, de deux Volts seulement. De plus, un potentiel intéressant vis-à-vis du R_{ON} est mis en évidence. Le facteur de mérite à 1,3 Ampères est de 25 $\mu\text{m}^2 \cdot \text{W}^{-1}$, alors qu'aucune optimisation spécifique n'a été appliquée.

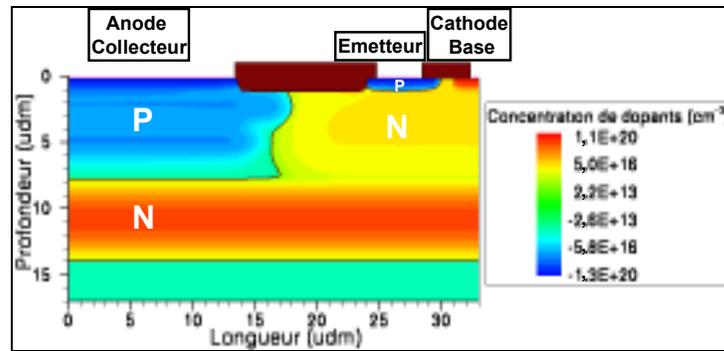


Figure 85: Profil de dopage de la structure retenue pour l'étude en simulation.

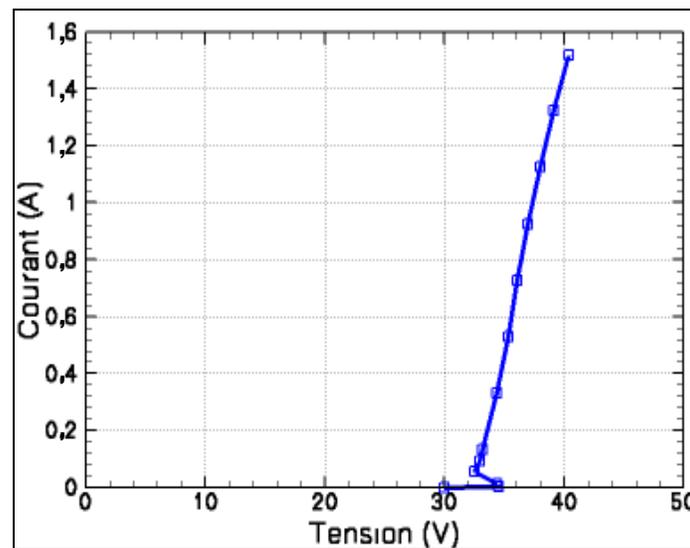


Figure 86: Caractéristique TLP simulée de la structure présentée Figure 85 (surface de $6,4 \cdot 10^4 \mu\text{m}^2$).

Les distributions de densité de courant et d'ionisation par impact sont relevées au point TLP ($V=40,4 \text{ V}$; $I=1,52 \text{ A}$), respectivement sur les figures 87 et 88. Elles montrent la coexistence de trois chemins de courant en parallèle coïncidant avec trois régions de forte ionisation par impact. Le chemin vertical sous le contact d'anode-collecteur correspond à l'avalanche de la diode en inverse. Les deux autres chemins, latéraux, sont induits par l'injection du transistor. Cette observation lève toute ambiguïté quant au fait que la diode et le transistor PNP sont tous les deux actifs lorsque la structure est passante.

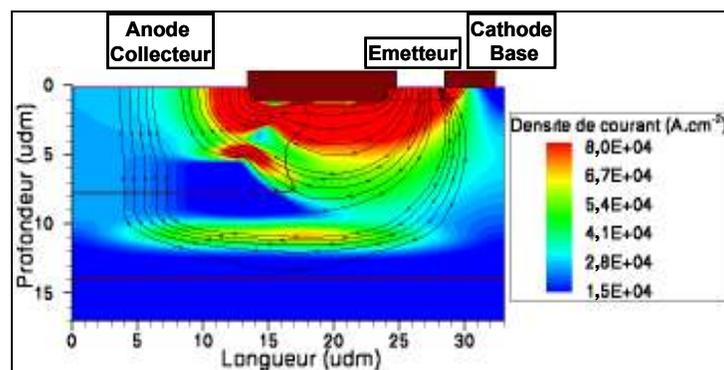


Figure 87: Distribution de densité de courant au point TLP ($V=40,4 \text{ V}$; $I=1,52 \text{ A}$).

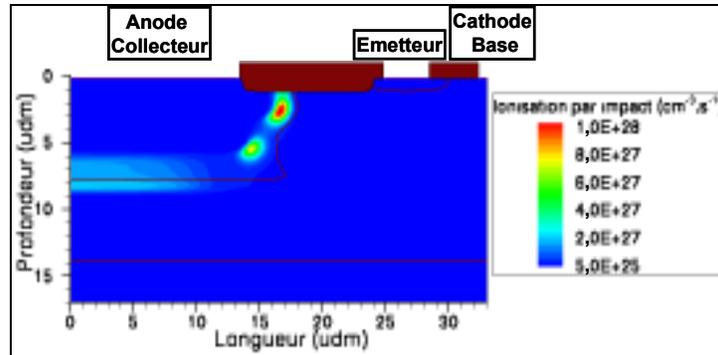


Figure 88: Distribution du taux d'ionisation par impact au point TLP ($V=40,4$ V ; $I=1,52$ A).

3.4.1.3. Variante favorable au couplage

L'efficacité du couplage entre la diode et le transistor PNP peut être optimisée en supprimant le contact de base-cathode et en remplaçant la diffusion d'émetteur en surface par une diffusion profonde (Figure 89). Bien que ces modifications transforment la diode verticale en un transistor PNP, le principe de fonctionnement n'est pas modifié. La NBL est trop fortement dopée pour que les trous puissent y transiter, de sorte que l'injection d'émetteur reste latérale et qu'il est justifié de modéliser le transistor PNP vertical par une diode. La suppression du contact de base-cathode force l'intégralité du courant de diode à passer dans l'émetteur, ce qui est le cas optimal pour la polarisation du transistor. L'utilisation d'une implantation P profonde permet de remplacer la zone de faible dopage entre le puits N et la NBL par la jonction émetteur-base et ainsi de s'affranchir de forts effets résistifs qui auraient tendance à pénaliser la diode. Cette variante devrait permettre d'améliorer encore le R_{ON} , elle sera donc privilégiée pour les réalisations sur silicium.

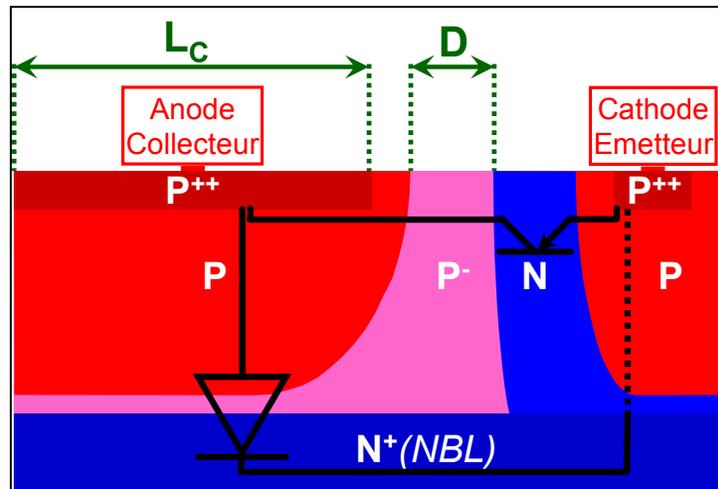


Figure 89: Coupe technologique de la structure améliorée.

3.4.2. Etude des diodes en inverse pendant une ESD

Avant de définir des règles d'optimisation, il est nécessaire de disposer d'une description précise du comportement d'une diode à avalanche au cours d'une ESD. Tout d'abord, une comparaison avec un transistor NPN autopolarisé sera effectuée en simulation, dans le but d'évaluer les performances théoriques en termes de robustesse et de R_{ON} . Ensuite, les mécanismes physiques contrôlant le

comportement à l'état passant seront décrits, en particulier la répartition spatiale de la génération thermique et les modifications de la ZCE à fort courant. Des résultats de mesures viendront compléter cette étude afin de déterminer précisément les performances atteignables en technologie SmartMOS 8 MV.

3.4.2.1. Simulation comparative d'une diode à avalanche et d'un transistor NPN autopolarisé

Pour cette étude comparative, un transistor NPN de structure verticale est retenu (Figure 90(a)). Sa base est formée par un puits P (PWell), et son collecteur par une couche enterrée (NBL) connectée via un puits N (NWell). Ce type de composant, couramment utilisé dans les protections ESD, est une référence pertinente, du fait de ses bonnes propriétés en termes de robustesse et de R_{ON} . Une diode de structure comparable est obtenue en remplaçant l'implantation N d'émetteur par une implantation P, de sorte que le contact d'émetteur du transistor NPN devient le contact d'anode de la diode, et en supprimant le contact de base (Figure 90(b)). Ainsi, la jonction de la diode est la même que la jonction base-collecteur du transistor NPN, et la configuration du contact d'anode est identique à celle du contact d'émetteur. Il faut préciser que ces deux composants n'ont pas été optimisés, notamment la surface et la résistance dans la NBL et le puits N pourraient être significativement réduites.

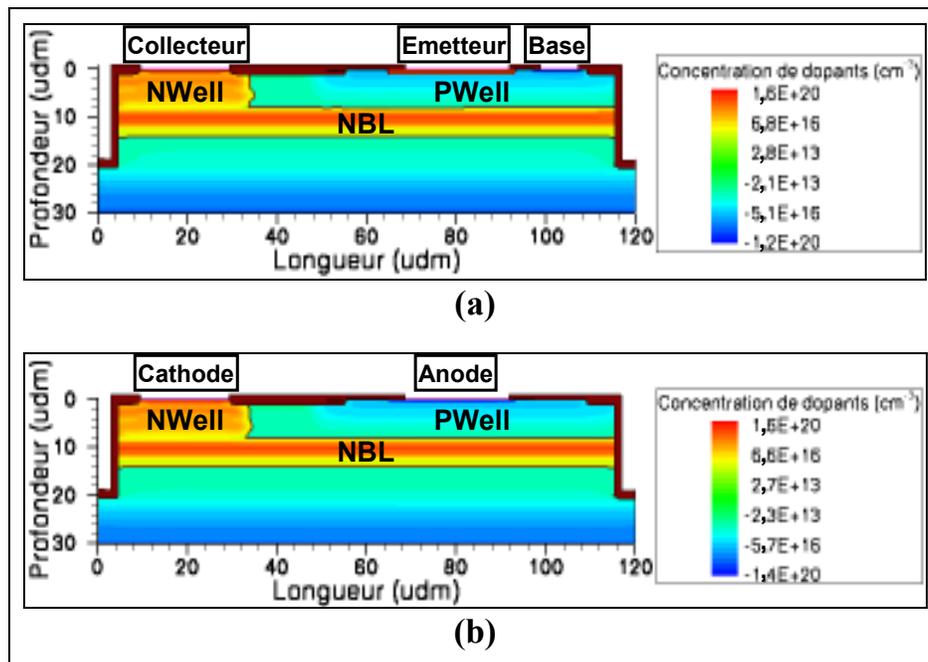


Figure 90: Profils de dopage du transistor NPN et de la diode retenus pour l'étude en simulation.

Les caractéristiques TLP de la diode et du transistor NPN sont simulées (Figure 91), pour des surfaces de $8,0 \cdot 10^3 \mu\text{m}^2$. Les deux composants commencent à conduire pour la même tension (60 Volts), correspondant au claquage de la jonction verticale. Ensuite, la tension de la diode augmente continuellement, et le NPN se replie fortement, jusqu'à une tension de maintien de l'ordre de 20 Volts. Ainsi, les tensions à l'état passant sont nettement plus élevées pour la diode que pour le NPN. Afin de ne pas biaiser la comparaison par les résistances de la NBL et du puits N, les calculs porteront sur la tension soutenue dans la ZCE de la jonction en inverse. Les puissances à des températures maximales proches (535 degrés Kelvins pour la diode et 560 degrés Kelvins pour le NPN), les facteurs thermiques associés F_T , donnés par la formule (11), correspondant à la puissance dissipée par unité de surface et pour une élévation de température de un degré Kelvin, et les R_{ON} à 700 milliAmpères sont

calculés (Tableau 15). Le facteur thermique de la diode est supérieur d'un facteur 2,4 à celui du NPN, révélant des propriétés en robustesse nettement meilleures pour la diode. Par contre, le R_{ON} de la diode est assez élevé alors que celui du NPN est négatif.

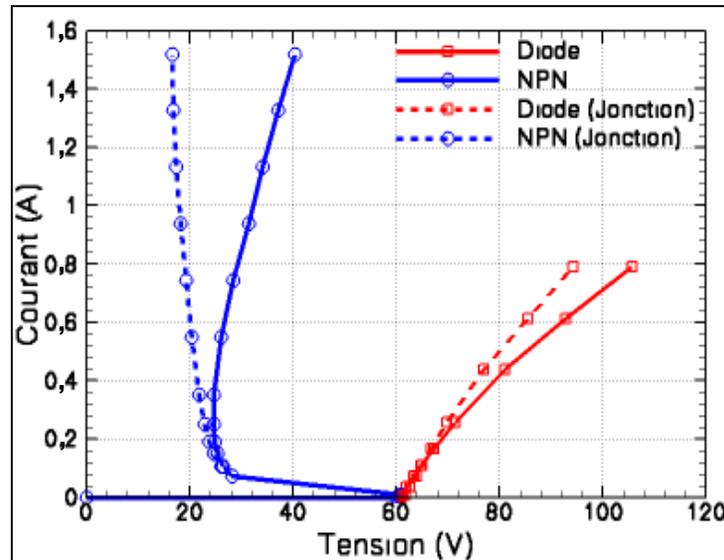


Figure 91: Caractéristiques TLP simulées de la diode et du transistor NPN, et contributions respectives des champs électriques soutenus aux jonctions inverses.

Tableau 15: Puissance dissipée et du R_{ON} pour la ZCE de la jonction en inverse de la diode et du transistor NPN.

	Diode	NPN
P (W)	83,3 (T=535 K)	38,7 (T=560 K)
F_T ($\mu\text{W} \cdot \mu\text{m}^{-2} \cdot \text{K}^{-1}$)	44,3	18,6
R_{ON} (Ω)	50,3	-5,66

3.4.2.2. Description des mécanismes physiques

Pour comprendre les bonnes propriétés de la diode vis-à-vis de la robustesse, les distributions de température (Figure 92) et de densité de courant (Figure 93) sont relevées. Dans la diode, la répartition de température est sensiblement homogène sous l'anode, alors que, dans le transistor NPN, elle est très localisée, sous forme de point chaud. Ainsi, le volume dans lequel la température augmente est plus grand pour la diode, ce qui explique qu'elle soit capable de soutenir une puissance supérieure. Ces distributions thermiques reflètent les répartitions de la conduction au voisinage de la jonction en inverse, où la dissipation de chaleur est maximale. Dans la diode, le flux de courant est homogène, la résistance dynamique positive assurant la stabilité de la conduction. Par contre, dans le NPN, le courant se concentre à l'intérieur d'un filament de dimensions réduites, sous l'effet des instabilités générées par le repliement et la résistance dynamique négative [36] [87].

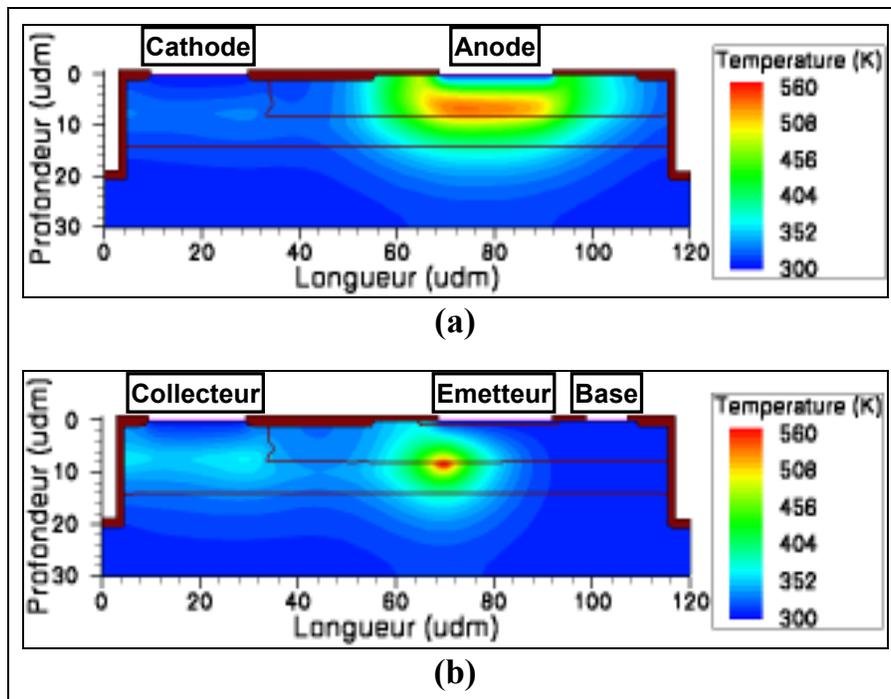


Figure 92: Distribution de la température dans la diode (a) et le transistor NPN (b) respectivement aux points TLP ($V=105,6$ V ; $I=789$ mA ; $T_{Max}=535$ K) et ($V=34,2$ V ; $I=1,13$ A ; $T_{Max}=560$ K).

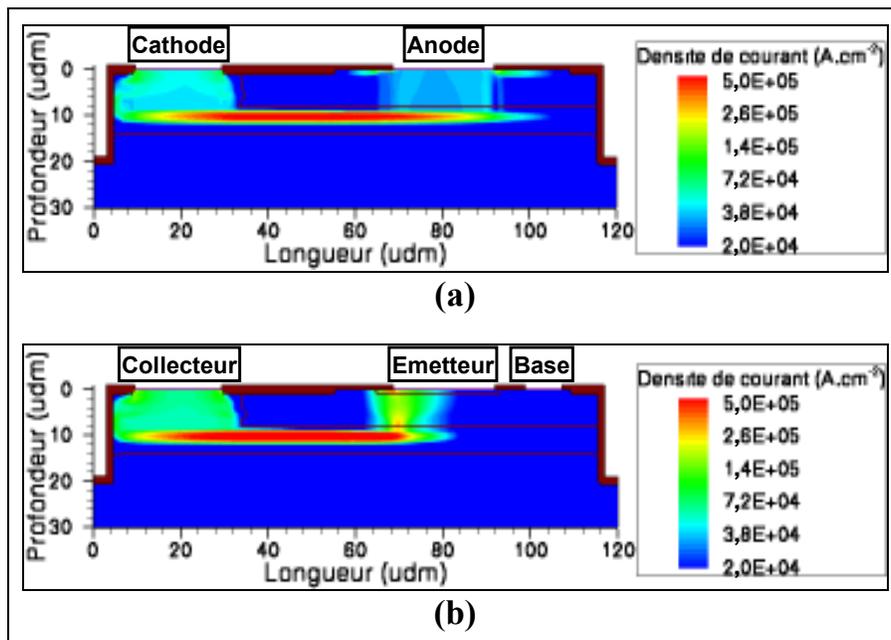


Figure 93: Distribution de la densité de courant dans la diode (a) et le transistor NPN (b) respectivement aux points TLP ($V=105,6$ V ; $I=789$ mA) et ($V=34,2$ V ; $I=1,13$ A).

Il reste maintenant à déterminer l'origine de la forte résistance dynamique de la diode. Pour cela, des coupes sont effectuées sous l'anode de la diode, au point TLP ($V=105,6$ V ; $I=789$ mA), et sous l'émetteur du transistor NPN, au point TLP ($V=34,2$ V ; $I=1,13$ A), afin de pouvoir évaluer les effets en fort courant sur les profils de champ électrique (Figure 94). A titre de référence, le profil de champ au claquage statique est également relevé. Dans la diode, le profil s'élargit, essentiellement côté anode, et deux maxima apparaissent de part et d'autre de la jonction. Le courant étant composé

majoritairement de trous côté anode et d'électrons côté cathode, la charge induite a tendance à compenser les dopants ionisés des deux côtés [88] [89] (Figure 95). Tant que le dopage est faible devant les charges induites, la charge résultante est faible et le champ varie peu. Ainsi, les effets en fort courant conduisent à un élargissement de la ZCE, tendant à augmenter la tension et le R_{ON} . C'est l'inverse de ce qui se produit dans un NPN, où la diminution de l'extension de la ZCE est bénéfique et peut être mise à profit pour obtenir un très faible R_{ON} .

Si les effets de forte injection dans une diode conduisent en général à un R_{ON} élevé, il faut cependant distinguer le cas des diodes PIN dont le dopage au voisinage de la jonction est très faible. Pour ces structures, le champ électrique entre les deux maxima peut diminuer avec le courant, résultant en un R_{ON} faible, voire négatif [88] [89]. Ainsi, les performances des diodes sont fortement liées à la technologie, suivant qu'elle permet ou non la réalisation de structures PIN de tensions de claquage et de surfaces raisonnables. Pour chaque technologie, une étape de caractérisation est donc nécessaire, afin de statuer sur l'opportunité d'utiliser les diodes comme protections ESD.

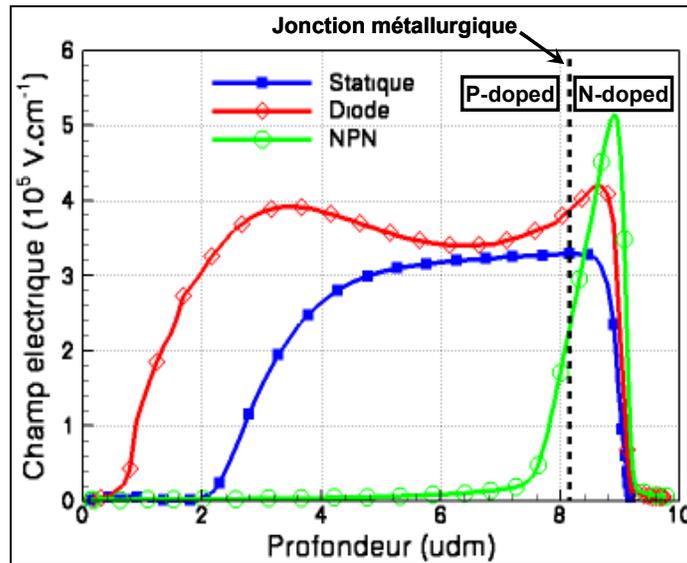


Figure 94: Champ électrique à la jonction en inverse dans la diode et le transistor NPN, au claquage statique (identique pour la diode et le NPN) et à fort courant (0,8 A pour la diode et 1,1 A pour le NPN).

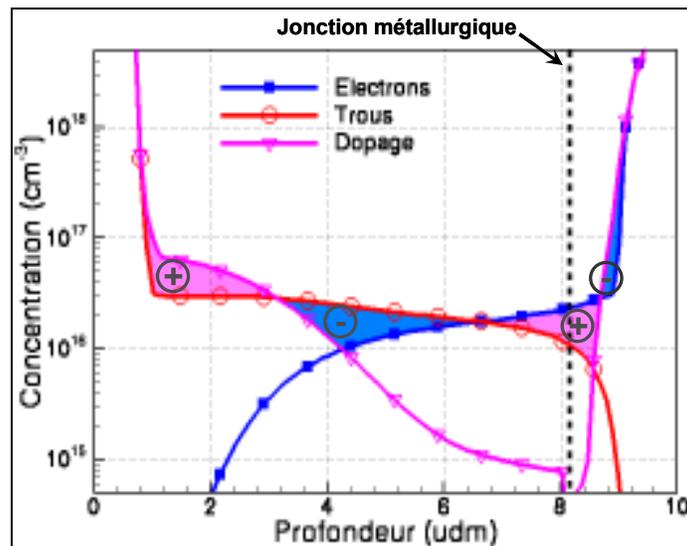


Figure 95: Distributions d'électrons, de trous et de dopants ionisés dans la diode au courant de 0,8 Ampères, faisant apparaître les charges d'espace positives (rouge) et négatives (bleu).

3.4.2.3. Résultats de mesure

Les diodes de la technologie SmartMOS 8 MV de Freescale ont été caractérisées. Les meilleures performances en termes de R_{ON} ont été obtenues pour une diode verticale de structure similaire à celle de la Figure 90 (a), et dont le profil de dopage s'apparente à celui d'une diode PIN. Sa caractéristique TLP est présentée Figure 96 et les principaux paramètres électriques sont listés dans le Tableau 16 (R_{ON} est la valeur extrapolée entre le déclenchement et la défaillance, R_{ON1} la valeur au voisinage de 250 milliAmpères et R_{ON2} la valeur au voisinage de 2,7 Ampères). Le facteur de mérite F_{RON} n'est supérieur que de 25 % à celui du transistor PNP optimisé (DUT4 présenté dans la partie 3.3.2.2). Cependant, le facteur de mérite à bas courant est très élevé (le double comparé à celui du DUT4), et c'est seulement à fort courant qu'il devient faible, alors que la chute de tension à l'état passant est déjà importante. De plus, on ne dispose d'aucun degré de liberté pour optimiser ce composant et les résultats présentés sont les meilleurs que l'on puisse atteindre. Aussi, il n'est pas envisageable d'utiliser cette diode seule comme protection ESD. Par contre, elle devrait pouvoir être mise à profit pour contribuer efficacement à la polarisation d'un transistor bipolaire. Concernant la robustesse, les excellentes propriétés de la diode sont validées, le facteur de mérite F_{I2} est très élevé, supérieur de 68 % à celui du DUT4.

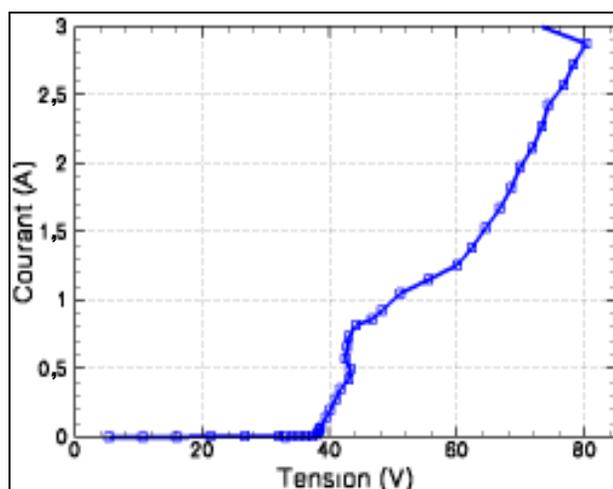


Figure 96: Caractéristique TLP de plus faible R_{ON} (surface $6,88 \cdot 10^3 \mu\text{m}^2$).

Tableau 16: Paramètres électriques de la diode (comparés à ceux du transistor PNP DUT4).

	Diode	DUT4
V_t (V)	39,5	49
V_{I2} (V)	80,2	70,7
I_{I2} (A)	2,87	4,42
F_{I2} ($\text{mW} \cdot \mu\text{m}^{-2}$)	33,5	19,9
R_{ON} (Ω)	14,2	4,9
F_{RON} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	27,3	21,4
R_{ON1} (Ω)	9,90	3,8
F_{RON1} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	40,5	18,8
R_{ON2} (Ω)	13,1	3,2
F_{RON2} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	14,8	12,5

Les résistances passantes R_{ON1} et R_{ON2} ne sont pas calculées exactement au même courant pour la diode et le transistor PNP DUT4. Cependant, dans les deux cas, elles correspondent à des comportements où les effets thermiques et de fort courant sont modérés, pour R_{ON1} , et très prononcés, pour R_{ON2} .

3.4.3. Règles d'optimisation

La stratégie consistera à optimiser le couplage entre la diode et le transistor PNP, avec pour objectif de réduire le R_{ON} tout en garantissant une robustesse acceptable. L'efficacité du couplage est liée à l'ajustement des tensions de déclenchement de la diode et du transistor l'une par rapport à l'autre, et au compromis entre le courant de la diode et sa surface. Ces deux points seront caractérisés de manière à établir des règles relatives aux paramètres les contrôlant. Cette étude sera basée sur la variante favorisant le couplage (Figure 89).

3.4.3.1. Tensions de déclenchement de la diode et du transistor PNP

La tension de claquage de la diode est fixée par le profil de dopage vertical sous l'anode. Elle ne peut prendre qu'un nombre limité de valeurs, correspondant aux implantations P disponibles dans la technologie. En revanche, la tension de claquage du transistor PNP varie continûment avec la distance D entre les masques des implantations d'anode/collecteur et de base (Figure 89). Si D augmente, le niveau de dopage au voisinage de la jonction diminue, ce qui entraîne une augmentation de la tension de claquage.

Une série de caractéristiques TLP est simulée en faisant varier D tout en gardant la même implantation d'anode, dont le claquage avec la NBL est de 34 Volts. Les résultats sont représentés sur la Figure 97 pour trois distances D correspondant aux cas où le claquage statique est uniquement latéral ($D=1,57$ μm), simultanément latéral et vertical ($D=5,98$ μm) et uniquement vertical ($D=8,92$ μm).

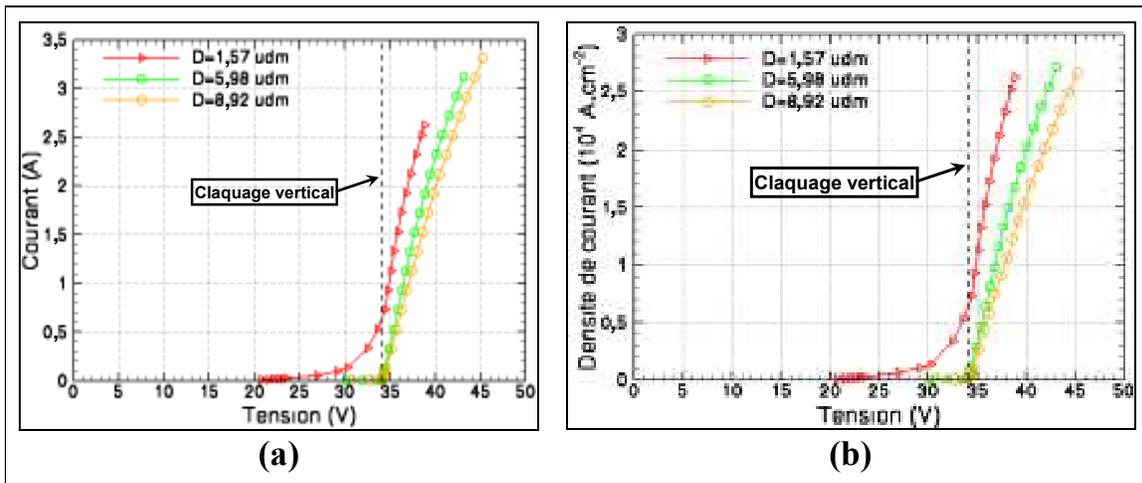


Figure 97: Caractéristiques TLP simulées pour différentes distances D , en fonction du courant (a) et du courant ramené à la surface (b).

Dans le cas du claquage latéral, le transistor PNP se déclenche en premier, et il faut attendre que la tension de claquage verticale soit atteinte pour que la contribution de la diode apparaisse. Dans l'intervalle, le transistor assure seul la conduction, de sorte que le R_{ON} est élevé. Dans le cas du claquage vertical, la diode devient passante ce qui a pour effet de polariser le transistor qui se déclenche à son tour. Ainsi, le couplage est assuré dès la commutation de la structure à l'état passant. Par contre, à 1,5 Ampères où dans tous les cas le couplage est établi, le facteur de mérite du R_{ON} se dégrade avec l'augmentation de l'espacement D (Figure 98). La valeur du R_{ON} augmente légèrement (Figure 97(a)), probablement suite à une réduction du facteur de multiplication latéral ayant effet de réduire le courant liée à l'autopolarisation du transistor PNP. Dans le même temps, l'accroissement de la surface est également pénalisant. Toutefois, les variations du facteur de mérite du R_{ON} sont de moins en moins prononcées lorsque la distance D augmente. En conclusion de cette analyse, le

comportement est optimal si les tensions de claquage en latéral et en vertical sont égales. Ainsi, le R_{ON} médiocre du transistor PNP seul est évité tout en préservant le facteur de mérite du R_{ON} le plus faible possible. Concernant le facteur thermique, ses variations en fonction de D sont très faible (Figure 98). Aussi, l'ajustement du claquage latéral devrait peu affecter la robustesse.

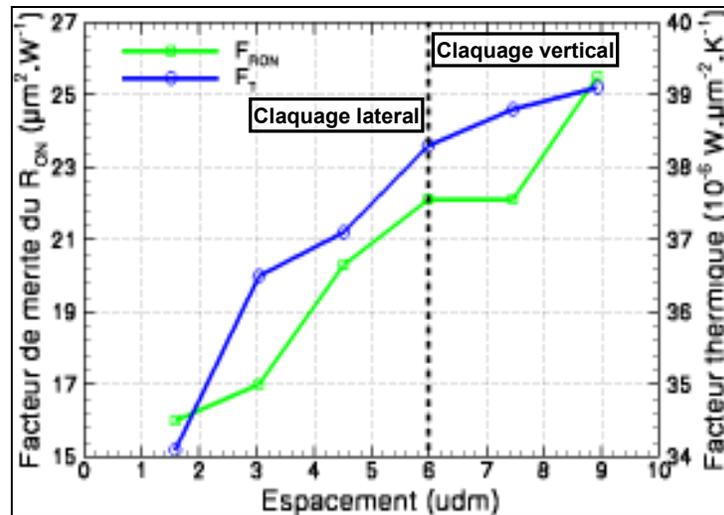


Figure 98: Facteur de mérite du R_{ON} à $1,5 \cdot 10^4 \text{ A.cm}^{-2}$ et du facteur thermique en fonction de la distance D .

3.4.3.2. Dimension de la diode

La jonction de la diode correspondant à la jonction plane verticale, sa surface est d'autant plus étendue que la longueur du contact d'anode/collecteur L_C est grande. Afin d'évaluer l'influence de ce paramètre, des caractéristiques TLP sont simulées en faisant varier sa valeur, à partir de la structure dont les tensions de claquage latérale et verticale sont égales. Les résultats montrent une diminution du R_{ON} pour une augmentation de L_C (Figure 99). Cette tendance est à mettre en relation avec l'augmentation du courant de diode, et donc de la polarisation du transistor PNP, pour une plus grande surface de la jonction plane. Cependant, l'accroissement simultané des dimensions tend à dégrader le facteur de mérite du R_{ON} . La simulation montre que sa valeur passe par un minimum (Figure 100), correspondant au meilleur compromis entre les tendances contradictoires de la réduction du R_{ON} et de l'accroissement des dimensions. Ainsi, l'optimisation du R_{ON} consistera à déterminer la longueur de collecteur pour laquelle ce minimum est atteint.

Concernant la dissipation de la chaleur, le facteur thermique diminue significativement avec L_C (Figure 100), ce qui est susceptible d'avoir des conséquences néfastes sur la robustesse. L'explication de cette tendance repose sur le fait que le couplage entraîne une forte élévation de température au niveau du transistor latéral, où la densité de courant est la plus élevée (Figure 87). Or, la surface du transistor diminue relativement à la surface de diode, si celle-ci augmente, ce qui résulte en une augmentation des densités de courant latérales, et donc en une amplification de l'élévation de température. En pratique, une diode plus grande peut conduire à réduire le nombre de doigts d'une structure interdiguée, de sorte que la densité de courant dans chaque doigt est plus élevée.

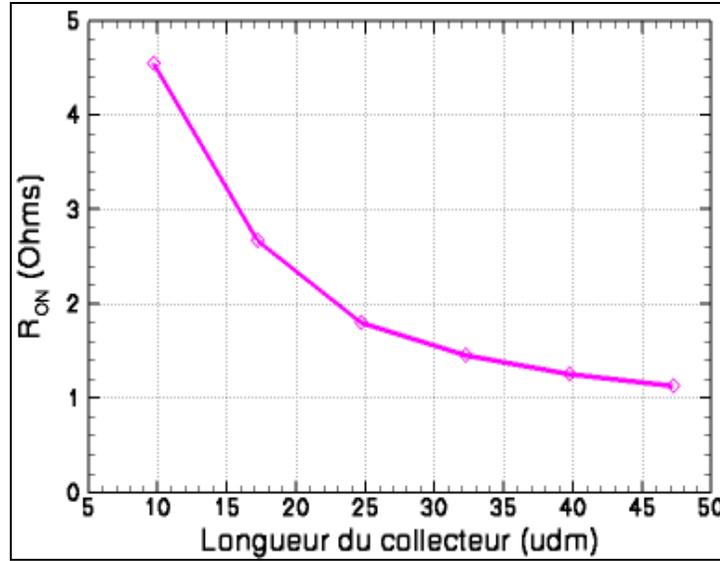


Figure 99: R_{ON} à 1,5 Ampères en fonction de la longueur de collecteur L_C .

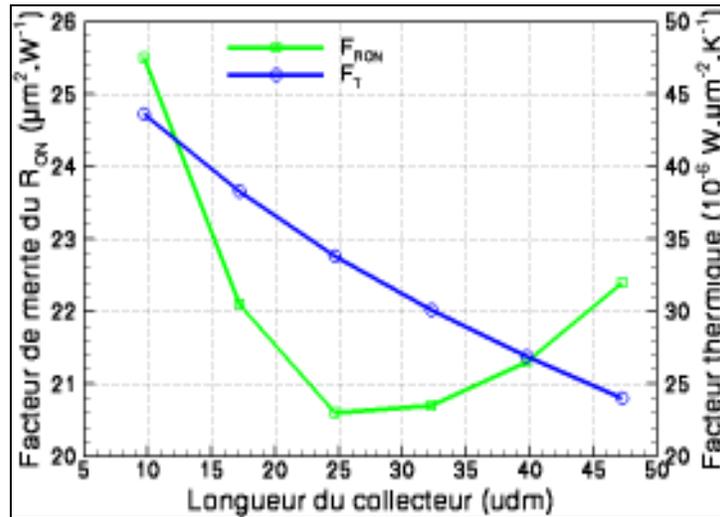


Figure 100: Facteur de mérite du R_{ON} à $1,5 \cdot 10^4 \text{ A} \cdot \text{cm}^{-2}$ et facteur thermique en fonction de la longueur de collecteur L_C .

3.4.3.3. Contribution des différentes composantes du courant

Trois composantes du courant peuvent être distinguées : le courant de la diode verticale, le courant du transistor PNP indépendamment de la polarisation induite par la diode et le courant supplémentaire du transistor PNP dû à cette polarisation. L'objectif est de séparer ces différentes composantes les unes des autres afin de pouvoir évaluer leurs contributions respectives. Pour que cette évaluation soit représentative de l'amélioration apportée par le couplage, la structure prise comme référence est celle qui présente le meilleur R_{ON} tout évitant le fonctionnement du transistor seul à bas courant (Figure 101(a)). Ensuite, cette structure est modifiée afin d'isoler le transistor (Figure 101(b)) et la diode (Figure 101(c)). Concernant le transistor PNP, la jonction plane de la NBL est supprimée et les contacts restent inchangés. En supprimant la jonction plane, il faut prendre garde de ne modifier ni le champ électrique à la jonction latérale, ni la répartition du courant dans la diffusion P du collecteur. Concernant la diode, les diffusions d'émetteur et de base sont supprimées. Le contact de cathode est

pris à l'ancien emplacement de la jonction verticale avec l'émetteur, afin de ne pas modifier la résistance d'accès dans la NBL.

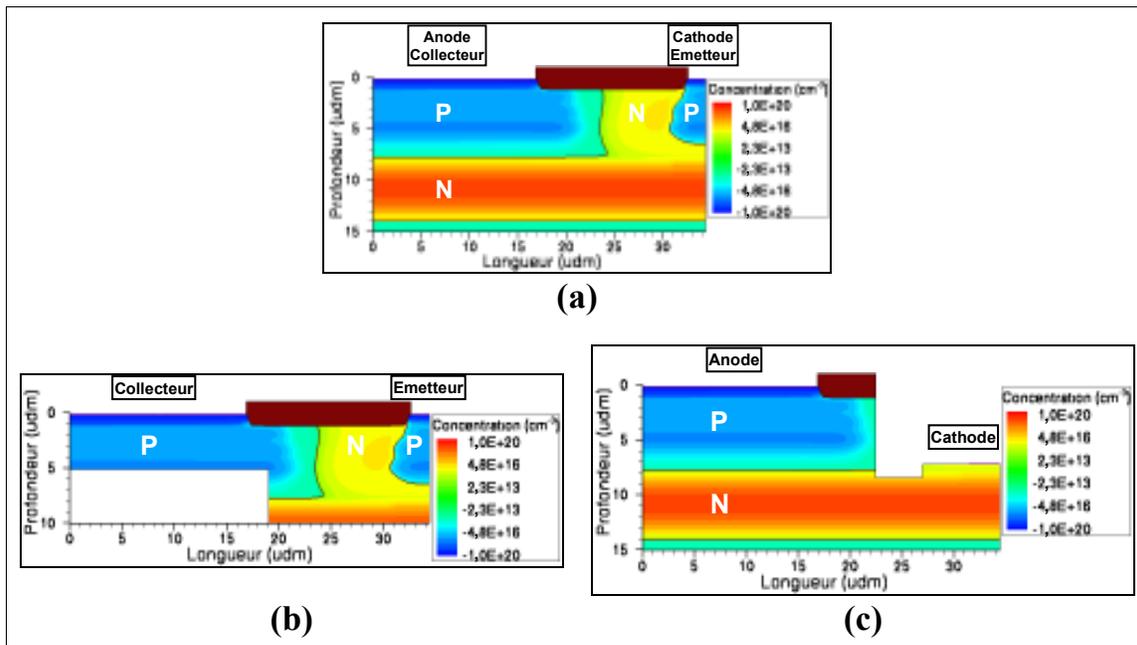


Figure 101: Profils de dopage de la structure de référence couplant un transistor PNP avec une diode (a), du transistor PNP seul (b) et de la diode seule (c).

Les caractéristiques TLP des structures modélisant le transistor PNP et la diode sont simulées. Les courants sont ensuite additionnés de manière à reproduire la caractéristique des deux composants en parallèle, sans interaction l'un avec l'autre. Le courant supplémentaire du transistor lié au couplage avec la diode est déterminé en retranchant à la caractéristique de la structure de référence (Figure 102(a)) la caractéristique des composants en parallèle. Ces résultats permettent de calculer les contributions au courant total des trois composantes précédemment identifiées (transistor PNP seul, diode seule et apport induit par couplage) (Figure 102(b)).

Sur le domaine simulé, la plus élevée des contributions est celle du courant lié au couplage, dont la valeur est comprise entre 41 % et 58 %. Ainsi, les règles définies pour optimiser le couplage sont extrêmement efficaces, et conduisent à une augmentation très importante du courant. Toutefois, sa contribution diminue lorsque la tension augmente, ce qui s'explique par une élévation de température à la jonction latérale plus rapide que dans cas du transistor seul, la densité de courant y étant plus élevée. Les effets thermiques tendant presque toujours à dégrader le R_{ON} , on s'attend effectivement à ce que leur amplification soit désavantageuse pour la contribution du couplage. Par ailleurs, la contribution de la diode est nettement supérieure à celle du transistor PNP. Cette particularité s'explique par l'amélioration du R_{ON} de la diode liée à l'insertion d'une diffusion d'émetteur profonde. Cette configuration permet de s'affranchir de la forte résistance d'accès des diffusions N généralement utilisées pour prendre le contact électrique avec la NBL, et ainsi de pouvoir profiter du profil de dopage vertical, favorable à la résistance dynamique de la diode.

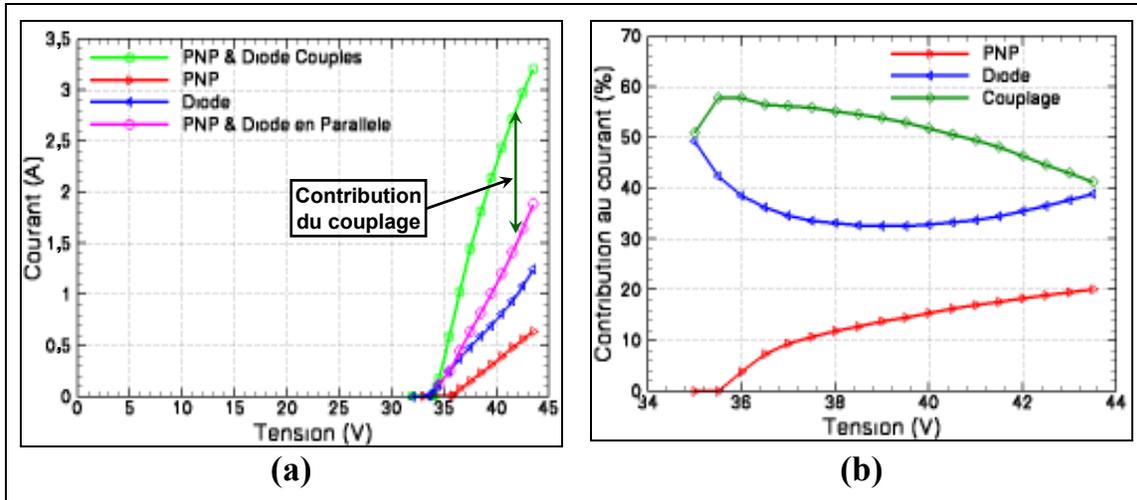


Figure 102: Caractéristiques TLP du transistor PNP, de la diode, de ces composants en parallèle et couplés dans la structure de référence (a), et les contributions au courant total de la structure de référence (b).

3.4.4. Caractérisations

Dans un premier temps, les règles d'optimisations sur les tensions de déclenchement et la surface de la diode sont confrontées avec des mesures. Ensuite, une structure sera développée pour protéger les E/S 80 Volts de l'application ethernet dont les spécifications ont été présentées au Chapitre 1 (1.2.3.3). La finalité est de pouvoir mettre en œuvre une stratégie de protection localisée qui ne sera intéressante du point de vue de la surface que la structure peut être insérée dans l'espace réservé au plot d'E/S.

3.4.4.1. Validation des règles de dessin pour l'optimisation du couplage

Des structures avec différents espacements D , entre les masques des diffusions de collecteur et de base, et différentes longueur de collecteur L_C sont réalisées sur silicium et mesurées en TLP et testées en robustesse. Ainsi, les influences des tensions de déclenchement et de la surface de la diode seront caractérisées précisément, ce qui permettra d'évaluer l'efficacité des règles d'optimisation et de guider leur mise en œuvre.

Tensions de déclenchement

Les coupes technologiques des structures réalisées sont semblables à celle de la Figure 89. Les layouts comprennent cinq doigts de 100 micromètres de long pour des surfaces totales proches de $100 \times 100 \mu\text{m}^2$, précisées dans le Tableau 17. Trois espacements D , de 7,03, 8,50 et 9,97 μm , ont été retenus. La longueur de collecteur L_C est égale à 18,3 μm (telle que définie sur la Figure 89, L_C correspond à la moitié du collecteur d'une structure interdigitées). La tension de claquage verticale entre la diffusion P de collecteur et la couche enterrée N est d'environ 59 Volts. Les résultats des caractérisations TLP sont représentés sur la Figure 103(a), et les principaux paramètres électriques relevés dans le Tableau 17.

Les tensions de déclenchement V_t pour les deux plus petits espacements (7,03 et 8,50 μm) sont inférieures à la tension de claquage verticale. Seule la structure de plus grand espacement (9,97 μm) se déclenche au claquage vertical. Comme observé en simulation, le R_{ON} à bas courant (R_{ON1}) des deux premières structures est élevé jusqu'à ce que la tension de claquage verticale soit atteinte. Les structures passent alors dans un mode faiblement résistif caractérisé par une nouvelle valeur du R_{ON} (R_{ON2}). Pour la dernière structure, elle est dans le mode faiblement résistif dès le déclenchement. Ces

comportements sont plus facilement observables sur la Figure 104, qui est focalisé sur les caractéristiques TLP à bas courant.

Concernant le mode faiblement résistif, la mesure de R_{ON2} nécessite un calibrage minutieux du banc de mesure TLP, qui a pu être réalisé jusqu'à trois Ampères. Au-delà l'ajout d'un atténuateur de courant dégrade la précision. Les trois structures présentent sensiblement le même R_{ON2} , voisin de un Ohm, correspondant à des facteurs de mérite de l'ordre de $3 \mu\text{m}^2 \cdot \text{W}^{-1}$. Pour de telles valeurs, il est probable que les contributions prépondérantes soient les résistances d'accès et de contact, et non plus les variations du champ électrique à la jonction en inverse. La simulation avait donné des facteurs de mérite significativement plus élevés, de l'ordre de $20 \mu\text{m}^2 \cdot \text{W}^{-1}$, ce qui révèle les limites de la méthodologie suivie pour prédire d'aussi faibles R_{ON} . L'augmentation du R_{ON} avec l'espacement D , observée en simulation, n'est pas retrouvée en mesure. Il est possible que l'on ne discerne pas cette tendance à cause du faible intervalle d'espacements couvert. Toutefois, nous pouvons conclure que cet effet reste limité et n'est pas critique pour l'optimisation. Seules les variations de la surface occupée tendent à différencier les facteurs de mérite du R_{ON} , mais très faiblement, comme le montre les caractéristiques TLP où le courant a été normalisé par la surface (Figure 103(b)).

Concernant la robustesse, les courants au second claquage I_{I2} sont rigoureusement les mêmes pour les trois composants, et les résultats des tests HBM et MM (Tableau 18) sont également proches les uns des autres. Comme prévu en simulation, l'espacement D peut être ajusté sans que la robustesse soit significativement affectée.

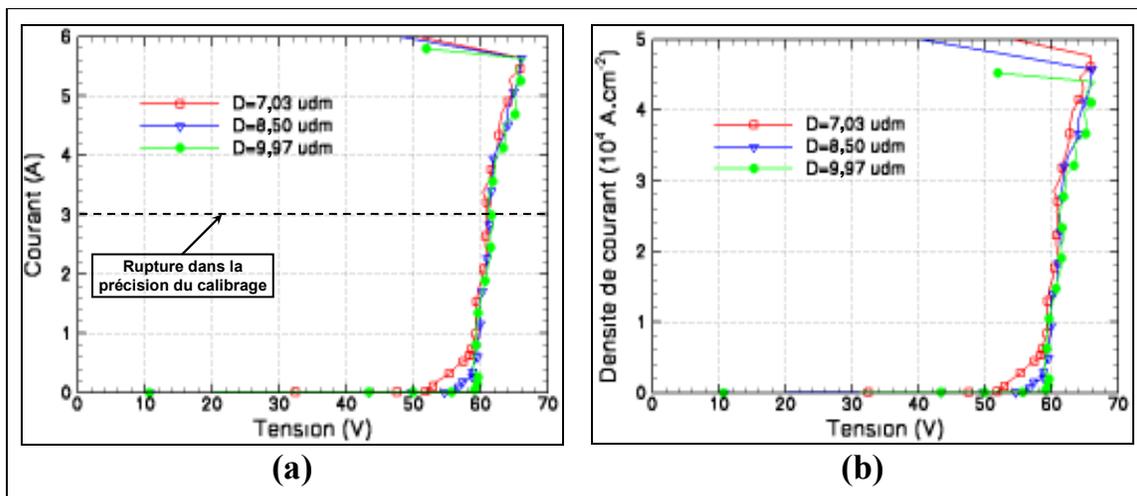


Figure 103: Caractéristiques TLP mesurées pour différentes distances D , en fonction du courant (a) et du courant ramené à la surface (b).

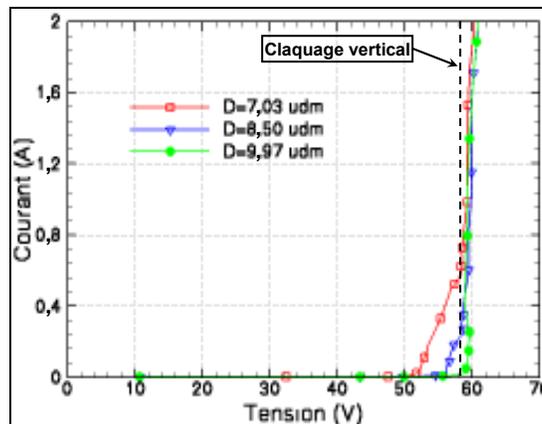


Figure 104: Caractéristiques TLP à bas courant pour différentes distances D .

Tableau 17: Paramètres des structures de test pour évaluer l'influence de la distance D.

D (udm)	7,03	8.50	9,97
S (μm^2)	$1,18 \cdot 10^4$	$1,23 \cdot 10^4$	$1,28 \cdot 10^4$
V_t (V)	52	56	59
R_{ON1} (Ω)	10	12	-
R_{ON2} (Ω)	≈ 1	≈ 1	≈ 1
F_{RON2} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	$\approx 3,2$	$\approx 3,4$	$\approx 3,6$
V_{I2} (V)	65.8	66,0	66,5
I_{I2} (A)	5,6	5,6	5,6
F_{I2} ($\text{mW} \cdot \mu\text{m}^{-2}$)	32	30	29

Tableau 18: Résultats des tests de robustesse HBM et MM.

V_{HBM} (kV)	8,0	7,0	9,5
V_{HBM}/S ($\text{mV} \cdot \mu\text{m}^{-2}$)	680	570	740
V_{MM} (V)	500	500	450
V_{MM}/S ($\text{mV} \cdot \mu\text{m}^{-2}$)	42	41	35

Surface de la diode

Des structures du type représenté Figure 89 et comprenant cinq doigts ont été réalisées pour trois longueurs de collecteur L_C (7,87, 14,80 et 21,73 udm). Ces composants avaient été les premiers réalisés sur silicium, et la tension de claquage vertical égale à 27 Volts n'avait pas été ajustée avec la tension de claquage vertical égale à 55 Volts. Les caractéristiques TLP mesurées sont représentées sur la Figure 105, et les principaux paramètres électriques relevés dans le Tableau 19.

Pour les trois structures, le transistor PNP latéral se déclenche en premier à 27 Volts. Le R_{ON} reste élevé (R_{ON1}) jusqu'au claquage de la diode verticale à 55 Volts où la structure passe en mode faiblement résistif (R_{ON2}). L'effet de la longueur de collecteur L_C doit être étudié lorsque le couplage est établi, c'est-à-dire dans le mode faiblement résistif. En accord avec la simulation, la valeur de R_{ON2} diminue avec l'augmentation de L_C . Par contre, le minimum du facteur de mérite correspondant (F_{RON2}) n'est pas observé. Il est probable que l'analyse ne soit pas assez fine pour le révéler, les mesures ne portant que sur trois valeurs de L_C comprises dans un intervalle relativement restreint. Toutefois, le facteur de mérite varie peu entre les deux plus grandes valeurs de L_C , ce qui pourrait indiquer que le minimum est atteint au voisinage de celles-ci. De plus, cette faible variation tend à indiquer que l'on dispose d'une marge importante pour ajuster L_C .

Concernant la robustesse, les variations du facteur de mérite F_{I2} sont minimes, de l'ordre de grandeur de la précision de la mesure. Ces observations semblent en contradiction avec la forte diminution du facteur thermique en simulation (Figure 100). Ce désaccord ne peut s'expliquer que si la défaillance est initialisée par un autre phénomène que l'élévation de température à la jonction latérale. Pour en savoir plus, il aurait fallu disposer d'une analyse de défaillance.

Tableau 19: Paramètres des structures de test pour évaluer l'influence de la longueur de collecteur L_C .

L_C (udm)	7,87	14.80	21.73
S (μm^2)	$0,76 \cdot 10^4$	$0,98 \cdot 10^4$	$1,20 \cdot 10^4$
V_t (V)	27	27	27
R_{ON1} (Ω)	20	20	19
R_{ON2} (Ω)	7,4	3,7	2,4
F_{RON2} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	15,5	10,0	8,3
V_{I2} (V)	65.2	64.1	62.2
I_{I2} (A)	2.5	3.41	3.92
F_{I2} ($\text{mW} \cdot \mu\text{m}^{-2}$)	21,5	22,3	20,3

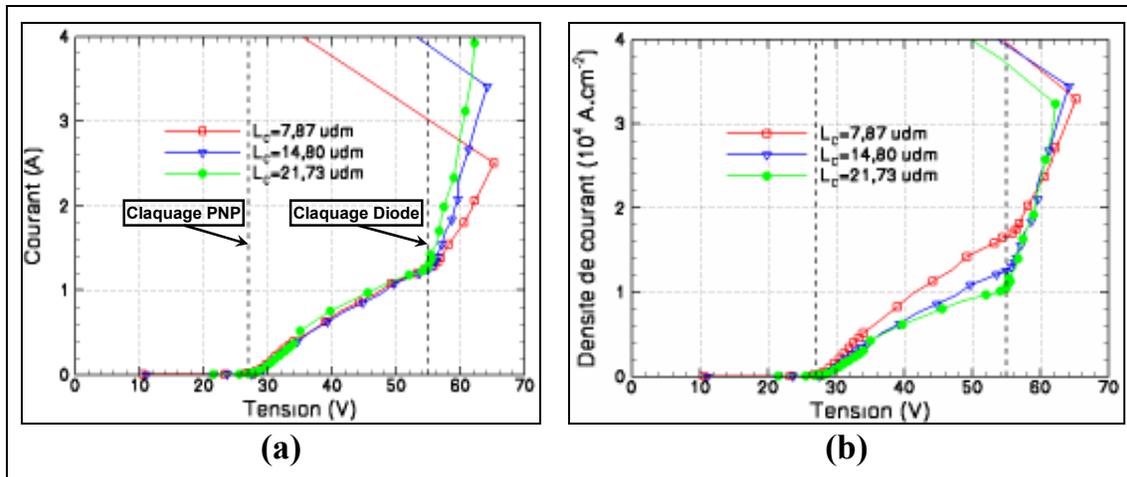


Figure 105: Caractéristiques TLP (a) et correspondance pour le courant ramené à la surface (b), de trois structures avec des longueurs de collecteur L_C différentes.

3.4.4.2. Structure de protection pour les E/S 80 Volts

Les structures étudiées sont maintenant mises à profit pour développer une protection vérifiant les spécifications présentées en 1.2.3.3, pour les E/S 80 Volts d'une application ethernet. Les mesures précédentes (3.4.4.1) indiquent que les performances en termes de R_{ON} et de robustesse devraient pouvoir être atteintes facilement. En particulier, les facteurs de mérite F_{RON2} et F_{I2} obtenus avec les structures de test vis-à-vis de l'espacement D sont bien meilleurs que ceux spécifiés.

Afin d'obtenir un déclenchement au-dessus de 80 Volts, deux structures de claquages verticaux 32 et 54 Volts, nommées respectivement "Stack1" et "Stack2", sont mises en série pour former la structure de protection, nommée "Stack". Les espacements D de "Stack1" et "Stack2" sont respectivement de 9,0 et 11,4 udm, valeurs assurant une marge importante entre le claquage latéral et le claquage vertical. Ainsi, il est garanti que les variations du procédé technologique ou des imprécisions au cours du développement n'entraîneront pas le déclenchement du transistor PNP latéral seul à bas courant. D'après les conclusions de la partie précédente (3.4.4.1), ces marges importantes ne devraient pas dégrader significativement le R_{ON} . Les longueurs de collecteur L_C sont de 14,8 udm pour les deux composants, valeur assurant un facteur de mérite du R_{ON} faible, au vu des mesures précédentes. Les layouts comprennent cinq doigts de 56 μm de long. Les surfaces, données dans le Tableau 20, incluent également un contact de base, assurant le déclenchement d'une diode en directe pour la polarité négative, plus des marges entre les extrémités des doigts et les tranchées d'isolation. Pour la structure de protection complète, une prise substrat est insérée, ce qui conduit à une surface totale de $2,02 \cdot 10^4 \mu\text{m}^2$ inférieure à la surface de $2,25 \cdot 10^4 \mu\text{m}^2$, à partir de laquelle une protection localisée est intéressante.

Les caractéristiques TLP des composants "Stack1" et "Stack2" seuls et de la protection complète "Stack" sont représentés sur la Figure 106, et les principaux paramètres électriques relevés dans le Tableau 20. La protection complète rentre parfaitement dans la fenêtre de conception. Ses tensions de déclenchement V_i et de maintien V_H sont bien comprises entre 80 et 100 Volts. A 1,3 Ampères, correspondant au courant maximal atteint durant une décharge HBM de 2 kV, la tension n'a pas dépassé 100 Volts. Le courant au second claquage I_2 correspond au moins à une robustesse de 4,5 kV HBM, bien au-dessus des 2 kV spécifiés. Les caractéristiques appellent deux remarques. Premièrement, un léger repliement est observé, alors qu'on s'attendait à ce qu'il soit supprimé. Nous supposons que cet effet est probablement dû à un temps de charge important à bas courant qui retarde le déclenchement. Deuxièmement, le R_{ON} de la protection complète est inférieur à la somme des R_{ON} de chacun des deux composants seuls. Au-dessus de 80 Volts, une contribution supplémentaire au courant est collectée au niveau de la prise substrat. La jonction en inverse entre la couche enterrée N

et le substrat P claque par avalanche, entraînant le déclenchement du transistor PNP parasite dont les électrodes d'émetteur et de base sont les mêmes que celles du transistor PNP latéral, et dont l'électrode de collecteur est la prise substrat.

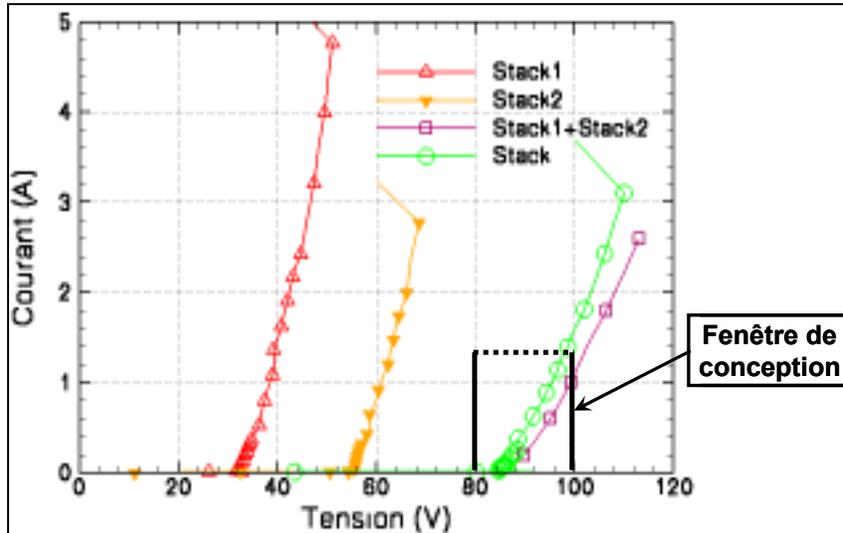


Figure 106: Caractéristiques TLP des composants "Stack1", "Stack2", et de la protection 80 Volts "Stack", et caractéristique obtenue en additionnant les tensions de "Stack1" et "Stack2".

Tableau 20: Paramètres électriques des composants "Stack1" et "Stack2" et de la protection "Stack".

	Stack1	Stack2	Stack
W*L (μm^2)	71,1*130,5	73,1*140,6	-
S (μm^2)	0,93 10^4	1,03 10^4	2,02 10^4
V _t (V)	35	55	91
V _H (V)	32	-	85
V(1,3 A) (V)	39	63	98
R _{ON} (Ω)	5,4	6,2	10
V _{I2} (V)	51	68	110
I _{I2} (A)	4,8	2,8	3,1
P _{I2} (W)	245	190	340

Sur la Figure 107 et dans le Tableau 21, les performances de la structure innovante réalisée sont comparées avec celle de la protection conventionnelle, présentée au Chapitre1 (1.4.3.1). La structure innovante a permis de diviser environ par deux le facteur de mérite. Cette avancée impressionnante ouvre des perspectives pour une forte réduction de la surface des protections "hautes tensions". De plus, la tension de déclenchement est ramenée en dessous de 100 Volts et le facteur de mérite de la robustesse est amélioré.

Tableau 21: Facteurs de mérite de la protection innovante comparés à ceux de la protection centralisée conventionnelle (valeurs calculées pour une surface équivalente à dix E/S).

	F _{RON} ($\mu\text{m}^2 \cdot \text{W}^{-1}$)	F _{I2} ($\text{mW} \cdot \mu\text{m}^{-2}$)
Structure innovante	24	17
Etat de l'art	52	11,5

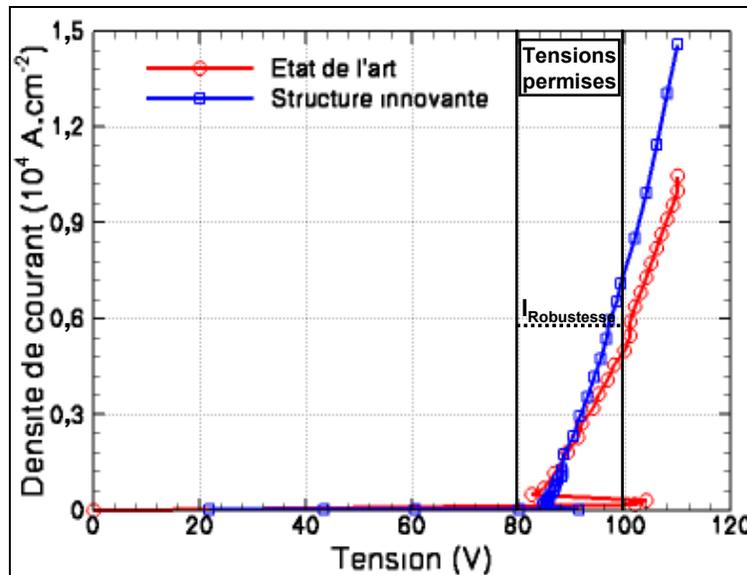


Figure 107: Caractéristiques TLP, avec le courant ramené sur la surface, de la solution conventionnelle (Etat de l'art) et de la structure innovante.

3.5. Conclusion

Les méthodes de calibrage mises en œuvre afin d'utiliser la simulation électrothermique comme outil d'optimisation des protections ESD, ont été présentées. Concernant la simulation du procédé technologique, des mesures SIMS des profils de dopage verticaux et des images SCM faisant apparaître les diffusions latérales ont fourni les références nécessaires pour ajuster les paramètres décrivant les implantations. Concernant la simulation électrothermique proprement dite, les durées de vie des électrons et des trous sont les principaux paramètres physiques du silicium qui doivent être calibrés. Pour cela, leurs valeurs sont ajustées de manière à reproduire la courbe du gain en fonction du courant de collecteur, mesurée à des courants comparables à ceux générés durant une ESD.

Une étude approfondie de structures "hautes tensions" à base de transistors PNP latéraux a été menée. L'étude du Chapitre 1 avait permis de déterminer que le repliement d'un transistor PNP autopolarisés est faible, et, qu'en conséquence, il est aisé d'obtenir une tension de maintien élevée. Par contre, un transistor PNP ne présente pas d'aussi bonnes propriétés qu'un transistor NPN en termes de faible R_{ON} . Afin de minimiser ce désavantage en réduisant au maximum le R_{ON} d'un transistor PNP, des règles de dessin sur les configurations d'émetteur et de collecteur ont été définies. La longueur d'émetteur doit être minimale, afin d'éviter que les effets 2D amplifient la chute du rapport d'injection, et le périmètre en regard du collecteur doit être maximal, afin de limiter la densité de courant. La longueur de collecteur doit être ajustée pour éviter la saturation de la vitesse des porteurs. Une telle saturation conduirait à l'apparition d'un champ électrique dans la diffusion de collecteur, ce qui dégraderait à la fois le R_{ON} et la robustesse. En combinant ces règles avec celles définies au chapitre 2, les mesures ont montré de bonnes performances vis-à-vis du R_{ON} , en particulier si le profil de dopage base-collecteur est optimisé. Dans le même temps, le comportement sans repliement des structures polarisées en base flottante a été démontré. Ainsi, un facteur de mérite du R_{ON} d'environ $20 \mu\text{m}^2.\text{W}^{-1}$ a été obtenu pour un déclenchement à 50 Volts sans repliement. Par ailleurs, cette structure présente une bonne robustesse, aussi bien vis-à-vis des tests HBM (supérieure à $640 \text{ mV}.\mu\text{m}^{-2}$) que MM ($41 \text{ mV}.\mu\text{m}^{-2}$). Concernant la dynamique de déclenchement, les caractérisations $v\text{TLP}$ ont montré que les transistors PNP sont adaptés aux spécifications CDM.

Une méthode a été développée pour réduire le R_{ON} au-dessous de la limite d'un transistor PNP seul. Elle consiste à favoriser l'injection par l'insertion d'une source de courant participant à la

polarisation. Pour réaliser cette source de courant, une diode à avalanche présente le double avantage de ne pas se replier, ce qui permet un fonctionnement à des tensions élevées, et d'être intégrable sous forme d'une structure verticale, ce qui permet de l'insérer dans le même volume de silicium qu'un transistor PNP latéral. Bien que les diodes à avalanche aient des propriétés médiocres en termes de R_{ON} , le courant généré est suffisant pour contribuer efficacement à la polarisation du transistor bipolaire. De plus, la fiabilité vis-à-vis de la défaillance est assurée par de très bonnes propriétés de dissipation de l'énergie. Des règles de dessin ont été définies pour optimiser le couplage, elles concernent l'ajustement des tensions de claquage latérales et verticales et des dimensions de la diode. L'application de ces règles permet de multiplier le courant par deux comparé, au cas où le transistor PNP et la diode fonctionneraient en parallèle, ce qui représente un gain exceptionnel pour le R_{ON} . Des caractérisations sur silicium ont permis de confirmer ces règles, tout en démontrant qu'elles sont compatibles avec des niveaux élevés de robustesse. Enfin, une structure de protection pour une fenêtre de conception comprise entre 80 et 100 Volts et une robustesse de 2 kV HBM a été développée en se basant sur ce principe. Une surface de $140 \times 150 \mu\text{m}^2$ a été atteinte, soit une réduction d'un facteur deux par rapport aux solutions précédentes. Une telle amélioration démontre de manière éclatante la puissance de l'optimisation du R_{ON} par le couplage entre le transistor PNP et la diode.

Chapitre4 Protections ESD "hautes tensions" et à faible R_{ON} à base de transistors bipolaires NPN

4.1. Introduction

Dans ce chapitre, deux types de protection "haute tension" à base de transistors NPN autopolarisés sont présentés. Tout d'abord, une structure est développée en appliquant les règles de dessin définies au Chapitre2. Son originalité consiste à tirer parti des spécificités de la technologie SmartMOS 8 MV de Freescale (tranchées d'isolation profondes, épitaxie P faiblement dopée, couche enterrée N fortement dopée...) de manière à appliquer au mieux les règles de dessin tout en limitant la surface. De plus, une stratégie de déclenchement s'avère nécessaire pour s'affranchir d'une forte surtension initiale. Le second type de protection est basé sur l'insertion d'une région flottante dans la base ou le collecteur. Cette région flottante est une diffusion de dopage complémentaire à celui de la région dans laquelle elle est insérée (diffusion N dans la base dopée P, ou diffusion P dans le collecteur dopé N), et qui n'est pas polarisée de l'extérieur par un contact. De nombreuses publications traitent de ce sujet, cependant, la plupart concerne les techniques de garde [57] [90] ou l'optimisation du compromis entre le R_{ON} et la tenue en tension des composants de puissance [91] [92] [93] [94]. Le comportement pendant une ESD est peu étudié, ce qui laisse la liberté d'imaginer des dispositifs innovants. Dans un premier temps, le fonctionnement statique est caractérisé, en particulier les effets 2D. Ensuite, les mécanismes physiques au cours d'une ESD sont approfondis dans le but de définir des stratégies de développement des protections.

4.2. Mise à profit des spécificités technologiques pour la réduction du facteur de mérite du R_{ON}

De l'étude théorique du Chapitre1, il est ressorti que l'obtention d'une tension de maintien élevée est la principale difficulté pour développer des protections "hautes tensions" à base de transistors NPN. En revanche, le R_{ON} est moins critique. Maintenant, l'objectif est de mettre en pratique les règles qui avaient été définies afin d'augmenter la tension de maintien tout en préservant le R_{ON} . Pour les appliquer sur une surface minimale, les spécificités de la technologie SmartMOS 8 MV, en particulier les tranchées profondes, sont mises à profit. Les règles appliquées ont été validées par la caractérisation TLP d'une structure sur silicium. Deux particularités appellent une analyse approfondie : un R_{ON} nul à fort courant et une tension de déclenchement élevée. Une stratégie basée

sur l'insertion d'un composant externe est testée pour réduire la tension de déclenchement, condition obligatoire pour que cette structure soit utilisée comme protection ESD.

4.2.1. Réalisation technologique

4.2.1.1. Rappel des règles de dessin définies au Chapitre1

Le développement est basé sur les règles de dessin établies au Chapitre1 :

- Emetteur
 - Implantation de surface (W_E faible)
 - Dopage faible, inférieur à 10^{17} cm^{-3}
- Base
 - Base large (W_B élevée)
 - Dopage élevé, compris entre 10^{18} et $5 \cdot 10^{18} \text{ cm}^{-3}$
- Profil de dopage à la jonction base-collecteur
 - Fort dopage de collecteur à profil abrupt
 - Faible dopage de base

Les règles sur l'émetteur et la base visent à réduire la densité de courant critique J_{Cr} , de manière à ce que le fonctionnement asymptotique soit atteint au niveau de la tension de maintien. Dans ces conditions, la faible valeur de la limite du rapport d'injection permet d'obtenir une tension de maintien élevée, et l'effet négatif de la chute du rapport d'injection sur le R_{ON} est évité. Par contre, l'influence significative de la charge liée au courant de polarisation est désavantageuse pour le R_{ON} . Néanmoins, il a été montré que cette contribution ne compense pas les effets bénéfiques de la charge liée au courant d'injection. Par ailleurs, les règles sur le dopage dans la ZCE à la jonction base-collecteur sont les règles fondamentales pour que le fonctionnement à fort courant soit favorable au R_{ON} .

4.2.1.2. Présentation de la structure

Des tranchées profondes dédiées à l'isolation des composants les uns par rapport aux autres sont disponibles dans la technologie SmartMOS 8 MV. Elles sont réalisées en oxyde et remplies de polysilicium, sur une profondeur de plusieurs micromètres. L'idée de base consiste à en insérer une entre les régions d'émetteur et de collecteur, de manière à forcer le flux de courant à circuler en profondeur. Le substrat et l'épitaxie étant dopés P, cette configuration permet d'obtenir un transistor NPN de large base interne W_B pour une surface limitée (Figure 108). Bien qu'une telle utilisation des tranchées profondes soit courante dans le domaine des composants de puissance, où elle permet d'augmenter la tenue en tension et de réduire la surface [95] [96] [97], elle est nouvelle dans le domaine des ESD. Concernant le profil de dopage à la jonction base-collecteur, la technologie offre une couche enterrée (NBL) fortement dopée N^+ avec un profil très abrupt. En l'insérant sous le collecteur, la jonction formée avec l'épitaxie P faiblement dopée remplit au mieux la règle correspondante. La contrepartie est la présence, inhérente à la technologie, d'une région faiblement dopée au-dessus de la NBL. Sa contribution tend à augmenter la résistance d'accès et, à fort courant, un champ électrique risque de s'y former, ce qui pourrait dégrader le R_{ON} (3.3.1.2). Concernant la règle sur l'émetteur, il n'existe pas d'implantation de surface faiblement dopée. Une implantation de surface sera retenue, ce choix étant guidé par l'obtention d'une base interne la plus large possible. Enfin, la base interne est essentiellement formée par l'épitaxie, dont le dopage est nettement inférieur aux niveaux optimaux. Pour compenser en partie ce désavantage, des dopages P profonds sont

implantés sous l'émetteur. En résumé, même si des compromis se sont avérés nécessaires, cette structure est extrêmement bien adaptée vis-à-vis de plusieurs règles essentielles pour l'augmentation de la tension de maintien et la réduction du R_{ON} . En particulier, il est probable que la large base interne conduise à un rapport d'injection maximal très faible, assurant une tension de maintien élevée, même si la densité de courant critique J_{Cr} n'est pas atteinte à la tension de maintien.

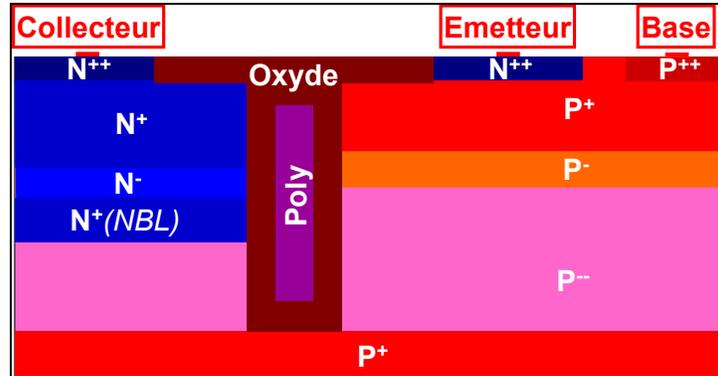


Figure 108: Coupe schématique de la structure développée.

4.2.2. Etude du fonctionnement

Une première approche en simulation est menée pour valider les principales propriétés mises à profit pour la conception. Ensuite, le comportement au cours d'une ESD est caractérisé par des mesures TLP. Deux particularités remarquables sont analysées en détail : un R_{ON} nul à fort courant et un fort repliement.

4.2.2.1. Simulation

L'objectif étant une simple validation qualitative du comportement, seul un point TLP est simulé, donnant ($V=53$ V ; $I=690$ mA). La distribution de densité de courant (Figure 109(a)) montre un flux en profondeur contournant la tranchée. La distribution de champ électrique (Figure 109(b)) est bloquée par la tranchée en latéral et se répartit principalement à la jonction base-collecteur mais aussi dans la région située au-dessus de la NBL. A la jonction base-collecteur, l'extension du champ est manifestement réduite le long du chemin de courant, ce qui confirme que le profil de dopage est favorable au R_{ON} . Le champ dans la région au-dessus de la NBL révèle que le dopage y est trop faible pour assurer une conduction ohmique des densités de courant générées lors d'une ESD. Selon l'analyse des effets survenant lorsque la vitesse des porteurs sature (3.3.1.2), il faut s'attendre à ce que ce champ électrique augmente rapidement, jusqu'à ce que le claquage soit atteint.

Afin de visualiser la base interne (Figure 110), les ZCE sont délimitées et la ligne de courant passant par le maximum de densité de courant à la jonction base-collecteur est tracée. Le chemin le long de cette ligne de courant entre les ZCE émetteur-base et base-collecteur est représentatif de la base interne.

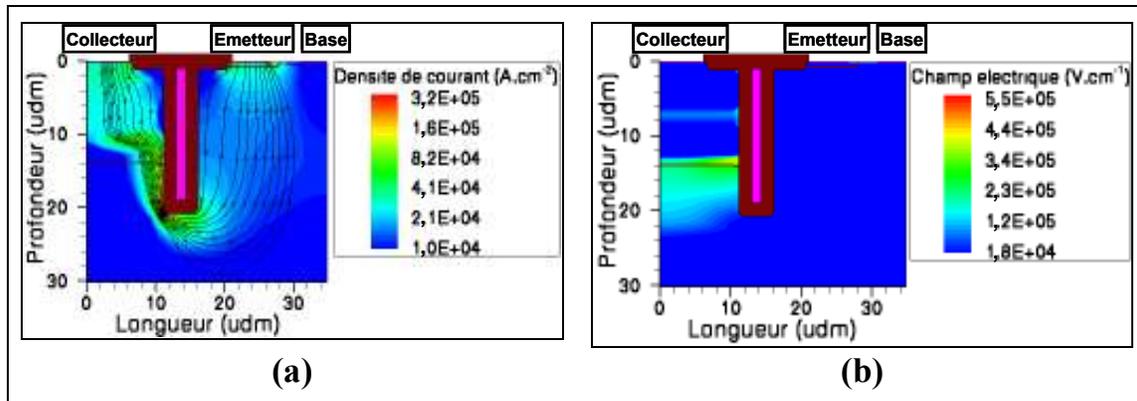


Figure 109: Distributions de la densité de courant, où les lignes de courants sont représentées, (a) et du champ électrique (b), au point TLP ($V=53\text{ V}$; $I=690\text{ mA}$).

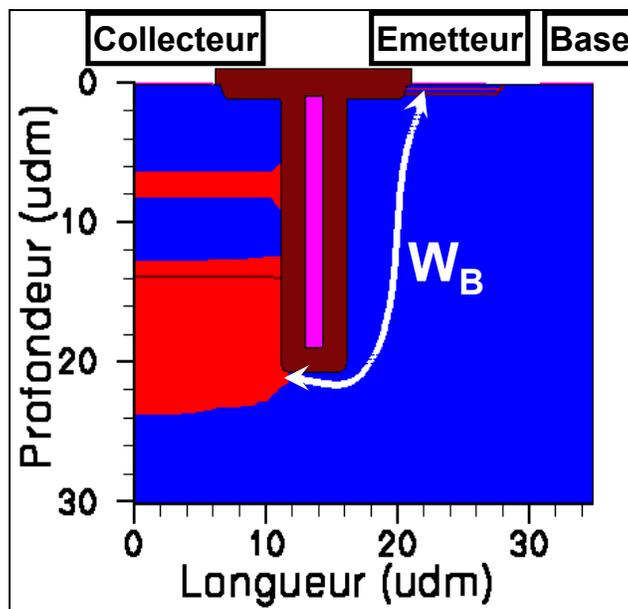


Figure 110: Répartition de la ZCE (rouge) au point TLP ($V=53\text{ V}$; $I=690\text{ mA}$), et ligne de courant dans la base représentative de la base interne W_B .

4.2.2.2. Caractérisations

Le layout de la structure testée est présenté sur la Figure 111. Il comporte quatre doigts pour une surface de $9,36 \cdot 10^3 \mu\text{m}^2$. Pour chaque doigt, les contacts d'émetteur et de base sont enserrés par des tranchées les séparant des contacts de collecteur. Une autre tranchée délimite l'ensemble du composant, l'isolant ainsi de son environnement. Une résistance en polysilicium de 860 Ohms, occupant une surface de $5,01 \cdot 10^2 \mu\text{m}^2$, est insérée entre la base et l'émetteur afin de faciliter le déclenchement.

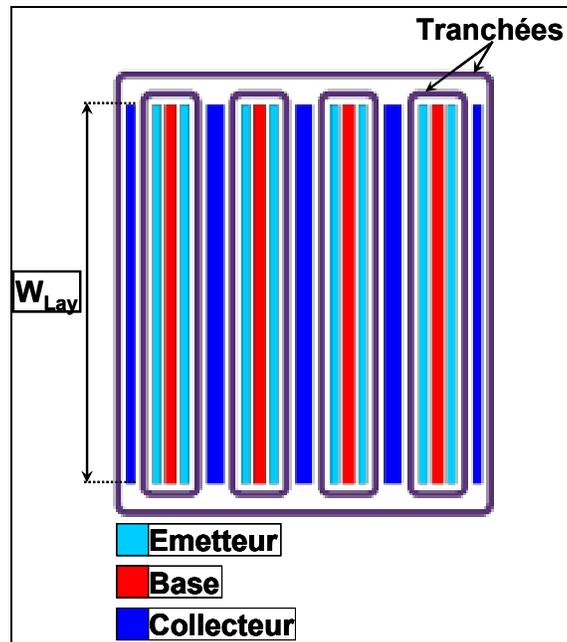


Figure 111: Layout schématique du transistor NPN testé montrant les tranchées profondes et les contacts.

Caractéristique TLP

La caractéristique TLP est mesurée (Figure 112) et les principaux paramètres électriques sont relevés (Tableau 22). Le repliement est très prononcé, avec une différence de 46 Volts entre la tension de déclenchement V_{tl} et la tension de stabilisation V_{St} (tension à laquelle le comportement électrique après repliement se stabilise, c'est-à-dire à partir de laquelle des points mesure TLP sont à nouveau obtenus). A l'état passant, deux modes de fonctionnement se succèdent. Du point de stabilisation jusqu'à un courant de 1 Ampère, le R_{ON} , noté R_{ON1} , est fortement négatif, et conduit à une chute de tension supplémentaire de 19 Volts. Lorsque le courant atteint 1 Ampère, la pente de la caractéristique s'infléchit et le R_{ON} , noté R_{ON2} , s'approche de zéro. Ce R_{ON} quasi nul se maintient jusqu'à 5 Ampères, limite en courant du testeur atteinte sans que le composant soit dégradé. Ce comportement à fort courant est extrêmement intéressant pour des protections dédiées à des fenêtres de conception étroites. Cependant, il est impératif de réduire la très forte différence entre la tension minimale V_{min} , égale à 40 volts, et la tension de déclenchement V_{tl} , égale à 109 Volts.

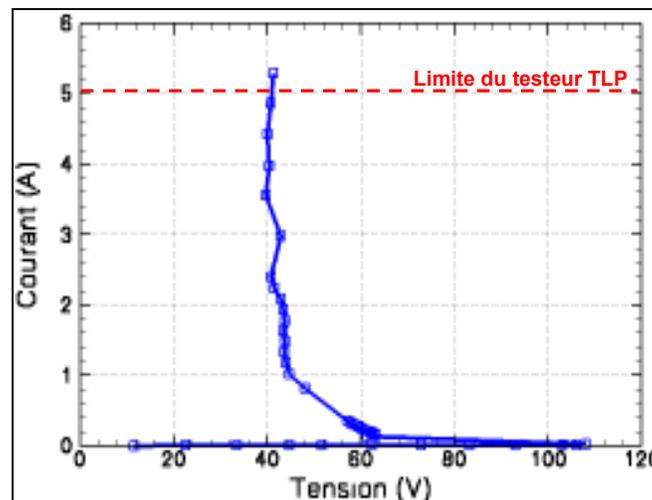


Figure 112: Caractéristique TLP mesurée.

Tableau 22: Principaux paramètres électriques.

V_{t1} (V)	V_{St} (V)	V_{min} (V)	R_{ON1} (Ω)	R_{ON2} (Ω)	I_{t2} (A)	F_{t2} (mW. μm^{-2})
109	63	40	-22	≈ 0	>5	$>20,3$

Aspects dynamiques

Des dynamiques de déclenchement particulièrement lentes sont attendues, du fait des temps de transit importants liés aux grandes dimensions de la base interne. Une application numérique simplifiée, en utilisant l'expression (39) définie au Chapitre1, donne entre 15 et 20 nanosecondes pour les électrons, et entre 50 et 60 nanosecondes pour les trous. Afin de caractériser ces aspects dynamiques, les oscillographes de la tension et du courant au cours d'une impulsion TLP sont relevés (Figure 113). Malgré le fort bruit de mesure au début de l'acquisition, le temps de charge peut être estimé à environ 50 nanosecondes. La mesure n'est stabilisée que sur les 20 dernières nanosecondes. Ces observations montrent que le composant est inadapté aux spécifications CDM. Une dynamique aussi lente peut également poser des problèmes pour la protection vis-à-vis des décharges HBM.

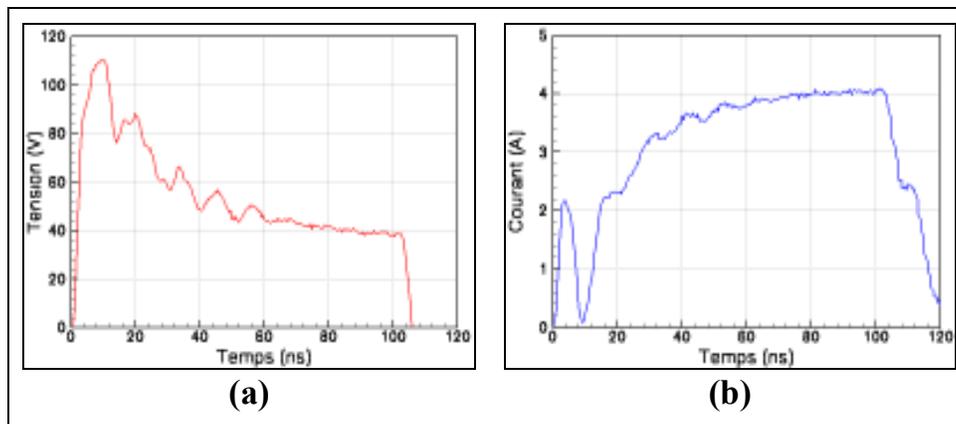


Figure 113: Oscillographes temporels de la tension (a) et de du courant (b) pendant l'impulsion TLP donnant le point ($V=40,4$ V ; $I=3,97$ A).

4.2.2.3. Tentative d'explication du R_{ON} nul à fort courant

L'application des règles d'optimisation n'explique pas à elle seule que le R_{ON} soit quasi nul entre 1 Ampère et 5 Ampères. Sur un intervalle aussi étendu, des variations significatives de la tension devraient être observées, ne serait-ce que celles induites par les résistances d'accès. Aussi, l'apparition de mécanismes tridimensionnels, non traités dans l'étude théorique du chapitre 2 (où la conduction est supposée uniforme suivant l'extension W_{Lay} (Figure 111) des doigts), est fortement suspectée.

Focalisation du courant

Une recherche bibliographique sur les structures à R_{ON} nul n'a permis d'identifier aucune publication traitant de ce sujet pour des niveaux de tension et de courant similaires à ceux du transistor NPN étudié. Néanmoins, un R_{ON} quasi-nul à bas courant a été publié pour un transistor ggNMOS (gate-grounded NMOS) en technologies CMOS 0,35 et 0,18 μm [24] et pour un transistor gcNMOS (gate-coupled NMOS) en technologie BCD 0,35 μm [25]. Les caractéristiques TLP montrent un mode de fonctionnement faiblement résistif après repliement, puis un changement franc, à un courant critique nommé I_C , vers un mode de fonctionnement de R_{ON} plus élevé (Figure 114(a)). La tension de maintien du ggNMOS en technologie CMOS est comprise entre 4 et 5 Volts, soit dix fois plus faible que celle du transistor NPN. Le R_{ON} reste quasi nul jusqu'à 2 mA. μm^{-1} ($6 \cdot 10^4$ A. cm^{-2} ,

en prenant en compte la surface totale du composant). Cette densité de courant est comparable à celle atteinte par le transistor NPN, pour lequel 5 Ampères correspondent environ à $5 \cdot 10^4 \text{ A.cm}^{-2}$. Les mêmes ordres de grandeur se retrouvent pour le gcNMOS en technologie BCD avec une tension de maintien comprise entre 5 et 6 Volts et un R_{ON} est quasi-nul jusqu'à $2,5 \text{ mA.}\mu\text{m}^{-1}$.

Afin de décrire l'origine du R_{ON} quasi-nul, les auteurs des publications [24] et [25] ont effectué des cartographies de la température à la fin de différentes impulsions TLP, à l'aide d'un appareil de mesure TIM. L'étude de ces cartographies leur a permis de mettre en évidence la formation d'un filament de courant après repliement, c'est-à-dire que la conduction n'est pas uniforme le long du composant, mais que seule une région localisée est passante. Une augmentation du courant est alors assurée par un élargissement du filament, à l'intérieur duquel la valeur de la densité de courant reste constante. Or, si la densité de courant n'est pas modifiée, la tension aux bornes du composant ne varie pas, d'où le R_{ON} quasi-nul. Pour que la densité de courant augmente, l'intégralité du composant doit être en conduction. En conséquence, le changement franc des variations du R_{ON} survient lorsque le filament s'est étendu sur tout le composant (Figure 114(b)).

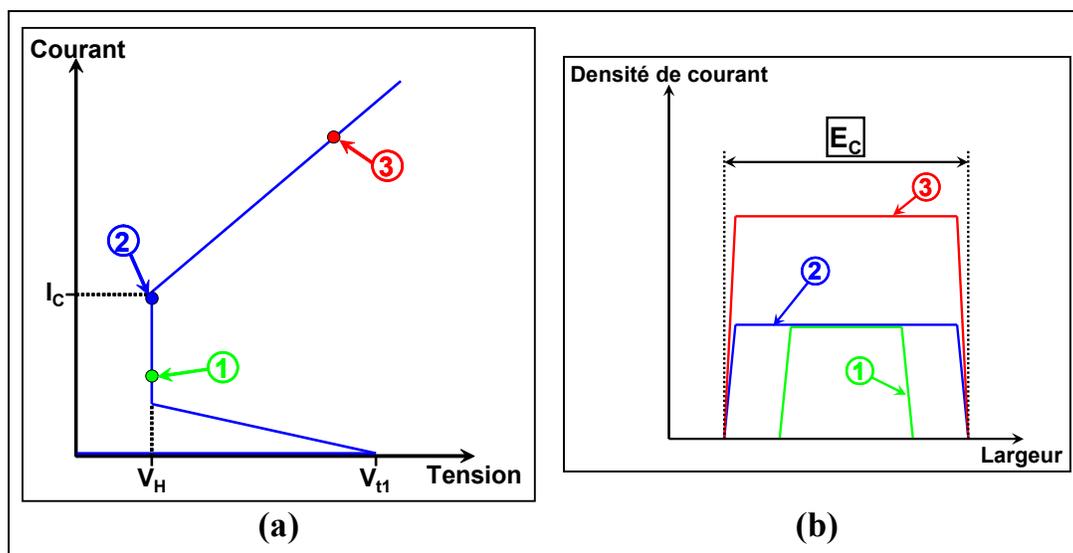


Figure 114: Caractéristiques TLP simplifiées des ggNMOS [24] et gcNMOS [25] (a), et répartitions de la densité de courant suivant la largeur d'un doigt W_{Lay} (Figure 111) en différents points TLP. (b).

Il est plausible que le R_{ON} nul du transistor NPN étudié ait pour origine une extension d'un filament de courant à densité de courant constante. D'une part, notre transistor NPN et les transistors NMOS des publications [24] et [25] présentent un R_{ON} nul à des densités de courant comparables. D'autre part, le R_{ON} du transistor étudié change pour un courant auquel une saturation de la conduction dans la région faiblement dopée de collecteur peut être attendue. Or, les phénomènes induits par cette saturation génèrent des instabilités entraînant une focalisation du courant [85]. On s'attend alors à ce que la robustesse soit dégradée, du fait de la réduction du volume dans lequel l'énergie est dissipée. Pourtant, ce n'est pas ce qui est observé : la mesure TLP montre un courant de défaillance I_{t2} élevé (Tableau 22).

Déplacement du filament de courant

Pour que la robustesse puisse rester élevée malgré la focalisation, le filament de courant doit se déplacer au cours de la décharge. Ainsi, l'énergie est dissipée sur toute l'extension du composant, ce qui permet d'éviter la formation de pics de température. Un tel déplacement du filament a été prédit en simulation [37], puis mis en évidence expérimentalement par des mesures TIM [26], réalisées sur un transistor DMOS en technologie SmartPower. Les caractéristiques TLP publiées avec ces mesures TIM indiquent une tension à l'état passant comprise entre 25 et 40 Volts, soit du même ordre de

grandeur que pour le transistor étudié. Suivant les modèles théoriques [98], ce déplacement est rendu possible par la diminution des coefficients d'ionisation par impact avec la température [61]. Ainsi, la multiplication par avalanche est d'autant plus efficace que la température est basse, de sorte que le courant a tendance à augmenter dans les régions plus froides, au détriment des régions où la présence du filament induit une température plus élevée.

4.2.2.4. Analyse du fort repliement

Une structure avec la base flottante est réalisée, afin d'abaisser la tension de déclenchement à BV_{CE0} , selon la règle établie au paragraphe 2.3.1. Sa caractéristique TLP (Figure 115) montre un déclenchement à 97 Volts, soit une réduction de seulement 12 Volts par rapport au cas où la base est reliée à l'émetteur. Une amélioration aussi limitée était prévisible étant donné le faible gain bipolaire. A fort courant, les caractéristiques avec la base flottante et la base connectée se superposent. Au total, il reste un écart de 47 Volts avec la tension minimale, ce qui est toujours beaucoup trop important pour une fenêtre de conception étroite. Un tel repliement à une tension inférieure à la tension de claquage en base flottante a été évoqué au Chapitre 1 (2.3.1 et 2.4.3). Il est lié à la modulation de la ZCE base-collecteur par le courant avant le maintien. Si cette modulation tend à réduire la tension, alors elle contribue au repliement. La structure étudiée rentre dans ce cas, étant donné le très faible niveau du dopage de base au voisinage de la jonction base-collecteur, fixé par l'épitaxie, et le profil de dopage base-collecteur ajusté précisément pour contribuer à réduire la tension. Ainsi, le fort repliement est intimement lié au dopage de la structure. Dans ces conditions, il ne sera possible de s'en affranchir qu'avec des structures de déclenchement externes.

Par la suite, l'étude portera uniquement sur la structure avec la base connectée. Quant à la structure base flottante, elle sera écartée, malgré la possibilité de réduire de 35 % la surface. Dans cette configuration, une forte robustesse n'est pas garantie, le nombre de doigts en conduction après le repliement étant aléatoire. Certains doigts se retrouvant bloqués, la densité de courant est d'autant plus élevée dans les autres, ce qui entraîne une défaillance prématurée. La connexion de la base à l'émetteur permet d'éviter ce problème. La même résistance reliant tous les contacts de base, ceux-ci se retrouvent au même potentiel, de telle sorte que si une jonction émetteur-base est polarisée en direct, les autres le sont également. Ainsi, le courant de défaillance en base flottante est mesuré à 3,5 Ampères, alors qu'il est supérieur à 5 Ampères si la base est connectée.

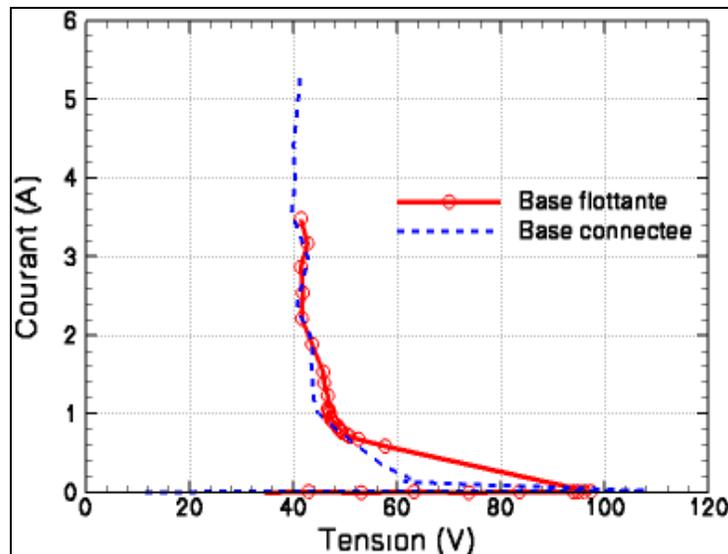


Figure 115: Comparaison des caractéristiques TLP en base flottante et avec la base connectée à l'émetteur (l'état passant en base flottante n'est décrit qu'à partir de 800 milliAmpères).

4.2.3. Réduction de la tension de déclenchement

Selon les conclusions de l'analyse du fort repliement, une stratégie basée sur un composant externe est indispensable pour limiter la tension de déclenchement. La stratégie retenue est présentée, puis validée par une application sur silicium.

4.2.3.1. Stratégie de déclenchement

Le composant externe doit se substituer au transistor NPN à bas courant, puis le déclencher pour un niveau de courant auquel sa caractéristique en tension est repliée. Ainsi, la surtension au déclenchement est éliminée tout en préservant le R_{ON} nul à fort courant. La stratégie adoptée consiste à insérer le composant externe entre la base et le collecteur (Figure 116(a)). Lorsqu'il est en conduction, il fixe la tension et contribue à polariser le transistor NPN, via l'apport d'un courant de base et la différence de potentiel dans la résistance externe. La limitation de la tension sera efficace, si, d'une part, sa tension de déclenchement est proche de la tension minimale du transistor NPN, et, d'autre part, si les caractéristiques à bas courant sont compatibles avec les contraintes imposées par une fenêtre de conception étroite (repliement faible ou nul et surtension à l'état passant limitée). Pour que le transistor NPN puisse passer en mode autopolarisé et ainsi se replier sur sa caractéristique verticale, des points I-V proches de sa caractéristique TLP doivent être atteints. En pratique, cette condition revient à s'assurer que la robustesse du composant externe est suffisante pour qu'il ne soit pas dégradé avant que les deux caractéristiques TLP se croisent (Figure 116(b)).

La structure développée ne constituera une protection efficace contre les décharges HBM que si le composant externe résout également le problème du déclenchement trop lent du transistor NPN. Pour cela, sa dynamique de déclenchement doit être compatible avec les temps de montée de ces décharges, de l'ordre de quelques nanosecondes. De plus, cette caractéristique impose une nouvelle condition sur sa robustesse. Elle doit être suffisante pour qu'il puisse absorber l'énergie au début de la décharge, le temps que le transistor NPN se déclenche.

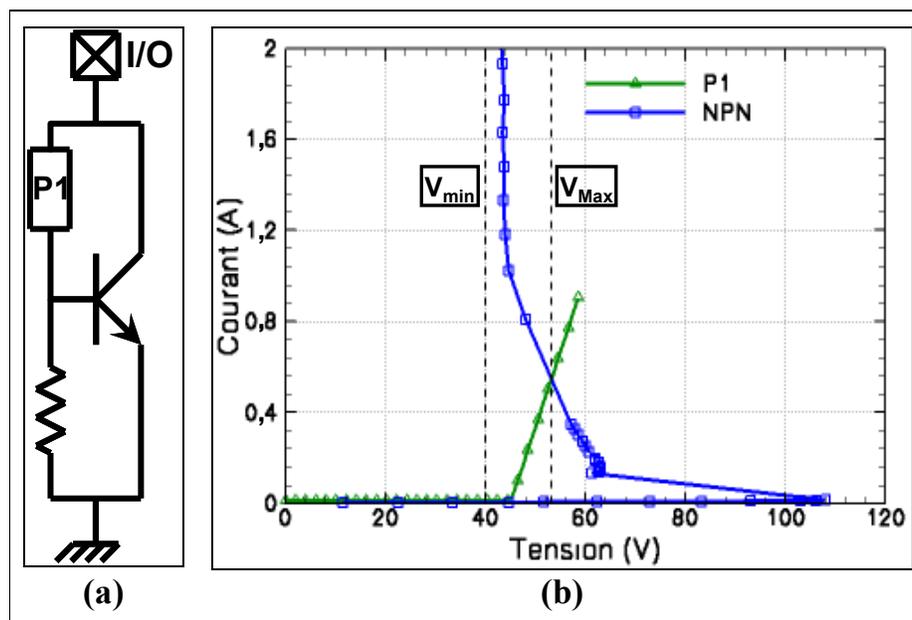


Figure 116: Schéma électrique de la stratégie de protection (a), et caractéristique TLP typique du composant P1 par rapport à celle du transistor NPN (b).

A titre de remarque, la configuration électrique est similaire à celle de la structure présentée au Chapitre 3, dans laquelle un transistor PNP est couplé avec une diode à avalanche. Le composant

externe est placé de telle manière qu'il contribue au courant de polarisation du transistor NPN, de même que la diode à avalanche contribuait au courant de polarisation du transistor PNP. Cependant, le fonctionnement est complètement différent. Dans le cas présent, il s'agit d'un circuit d'aide au déclenchement, sans influence sur la caractéristique à fort courant. En revanche, le couplage entre le transistor PNP et la diode à avalanche est mis à profit afin d'améliorer le R_{ON} .

4.2.3.2. Caractérisation

Présentation de la structure

Les transistors PNP mesurés au Chapitre3 répondent aux critères que doit vérifier le composant externe. Des tensions de déclenchement de l'ordre de grandeur de la tension minimale du transistor NPN étudié peuvent facilement être obtenues en jouant sur le profil de dopage latéral, via la position des masques d'implantation. Leurs caractéristiques TLP ne se replient pas, l'optimisation du R_{ON} a été très efficace, et les tests de robustesse ont donné de très bons résultats. De plus, la caractérisation vTLP a montré une dynamique de déclenchement compatible avec les décharges CDM, et donc a fortiori avec les décharges HBM.

Pour valider la stratégie de protection, le composant DUT4 de la section 3.3.2.2 est bien adapté. Sa tension de déclenchement est de 49 Volts, à comparer à la tension minimale de 40 Volts du transistor NPN, et les facteurs de mérite du R_{ON} et de la robustesse sont les meilleurs des transistors PNP mesurés. Seuls quatre doigts sont insérés, deux de chaque côté du transistor NPN, pour une surface de $4,83 \cdot 10^3 \mu\text{m}^2$. D'après les résultats de mesures présentés au Chapitre3, le R_{ON} pour quatre doigts devrait être de 16 Ohms, ce qui est suffisant étant donné qu'il ne contribue à la caractéristique TLP qu'à bas courant, et le I_{l2} devrait être de 1,3 Ampères, ce qui assure l'intersection des caractéristiques TLP.

Caractérisation TLP

La mesure de la caractéristique TLP est présentée sur la Figure 117 et les principaux paramètres électriques relevés dans le Tableau 23. La tension de déclenchement V_t est de 52 Volts. La différence avec les 49 Volts obtenus pour le composant DUT4 de la section 3.3.2.2 est due aux variations du procédé technologique. La tension maximale V_{Max} , égale à 59 Volts, est inférieure de 40 Volts à la tension du déclenchement du transistor NPN seul, ce qui valide la stratégie vis-à-vis de la réduction de la surtension initiale. A partir du repliement, la caractéristique se rapproche rapidement de celle du transistor NPN, dont la contribution devient prédominante. Dans un premier temps, la tension diminue selon un R_{ON} , noté R_{ON1} , de -24 Ohms, jusqu'à environ 1 Ampère, où l'on retrouve R_{ON} nul, noté R_{ON2} . La tension minimale reste d'environ 40 Volts, soit 19 Volts au-dessous de la tension maximale. Cette différence est relativement élevée, néanmoins le principe de la stratégie de déclenchement est validé. Pour obtenir un faible repliement, il suffira d'ajuster la tension de déclenchement du transistor PNP au plus près de la tension minimale du transistor NPN. Concernant la robustesse, le courant de défaillance I_{l2} n'est plus que de 2,92 Ampères comparé à une valeur supérieure à 5 Ampères pour le transistor NPN seul.

Tableau 23: Principaux paramètres électriques.

V_t (V)	V_{Max} (V)	$I(V_{Max})$ (mA)	R_{ON1} (Ω)	R_{ON2} (Ω)	I_{l2} (A)	V_{l2} (V)	F_{l2} (mW. μm^{-2})
52	59	260	-24	≈ 0	2,92	41,4	8,23

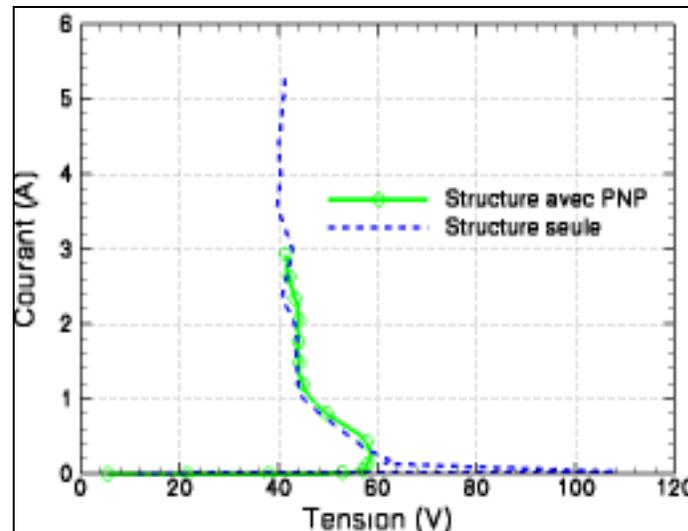


Figure 117: Caractéristique TLP de la structure avec un transistor PNP pour le déclenchement, comparée à la caractéristique TLP du transistor NPN seul.

Pour en savoir plus sur la dégradation de la robustesse, nous avons commencé une analyse des réponses temporelles complétant celle de la caractéristique TLP en apportant des informations sur les aspects dynamiques. Les oscillographes de la tension et du courant sont relevés pour l'impulsion juste avant la défaillance, soit au point TLP ($V=41,4\text{ V}$; $I=2,92\text{ A}$) (Figure 118), et pour l'impulsion où la défaillance apparaît (Figure 119). Les oscillographes avant défaillance montrent une tension et un courant constants entre 10 et 25 nanosecondes, respectivement d'environ 75 Volts et de 2,5 Ampères. Ce comportement révèle une limitation de la tension par le transistor PNP. Cette stabilisation de la tension et du courant au début de l'impulsion tend à confirmer un déclenchement rapide compatible avec la protection contre les décharges HBM, même si la mesure ne permet pas d'accéder à la réponse pendant les 10 premières nanosecondes. Jusqu'à 40 nanosecondes, la tension reste supérieure à la tension de déclenchement du transistor PNP, qui contribue donc à la conduction pendant toute cette durée. Cette observation est cohérente avec la dynamique de déclenchement particulièrement lente du transistor NPN seul (Figure 113). Sur l'oscillographe où la défaillance apparaît, celle-ci est révélée par un brusque décrochement de la tension et du courant après 20 nanosecondes, précisément pendant que le transistor PNP conduit. Ces observations montrent sans ambiguïté que la défaillance est liée à des phénomènes induits par le transistor PNP.

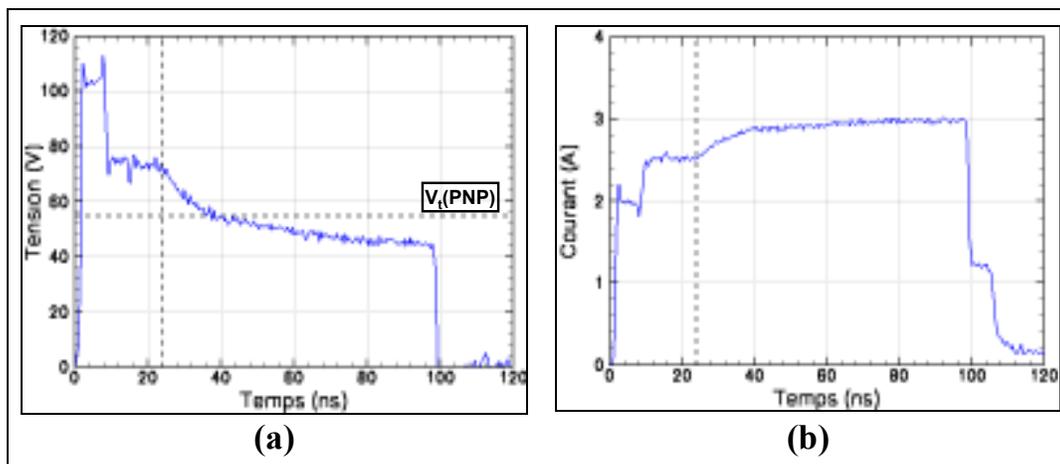


Figure 118: Oscillographes TLP de la tension (a) et du courant (b) pour le point ($V=41,4\text{ V}$; $I=2,92\text{ A}$) (mesure effective après les dix premières nanosecondes).

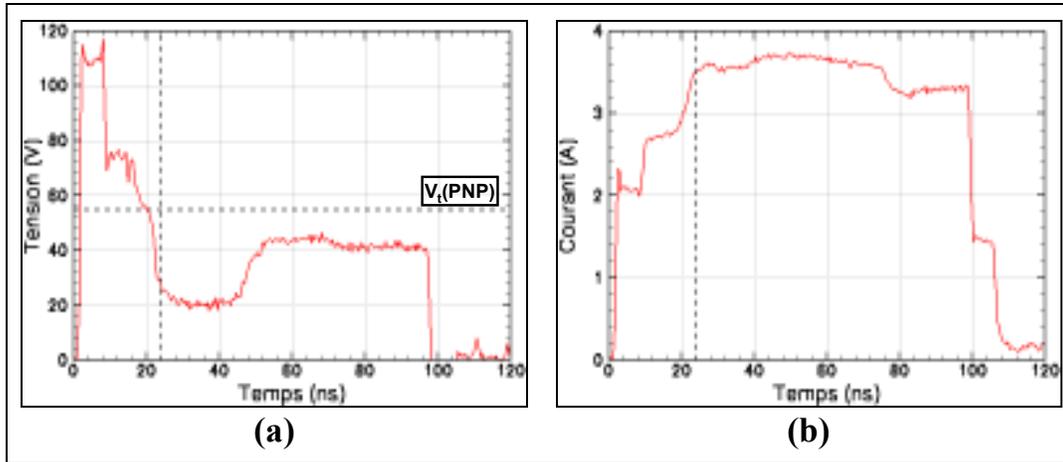


Figure 119: Oscillographes TLP de la tension (a) et du courant (b) à la défaillance.

Analyse de défaillance

L'étude de la dégradation de la robustesse a été complétée par une analyse de défaillance. Trois structures ont été testées en HBM jusqu'à la défaillance, survenue à 3,5 kV. Ensuite, le défaut a été localisé par des observations OBIRCH, technique très utilisée consistant à relever les variations de résistance sous l'effet de l'échauffement d'un laser. De manière surprenante, le défaut n'est pas situé dans le transistor PNP, comme nous aurions pu nous y attendre, mais dans le transistor NPN, ainsi que le montre la Figure 120. Le grossissement de la Figure 121 met en évidence une dégradation sous le collecteur, où se situe la jonction en inverse tenant le champ électrique pendant une ESD. En résumé, l'ajout du transistor PNP a fragilisé le transistor NPN. Pour déterminer précisément les mécanismes conduisant à cette fragilisation, la simulation de la protection complète devrait permettre de mieux comprendre l'interaction entre les deux composants. Cependant, nous n'avons pas eu le temps de mettre en œuvre cette simulation. Expérimentalement, des cartographies de température par mesures TIM pourraient également apporter des éléments de réponse.

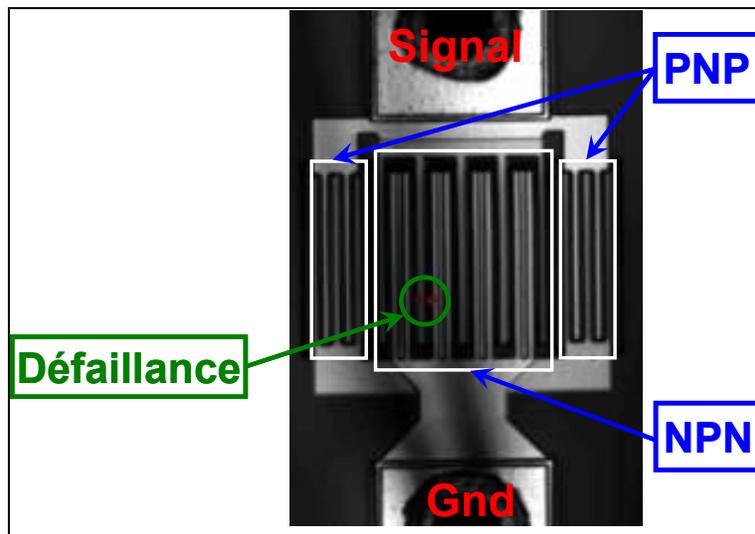


Figure 120: Localisation de la défaillance par observation OBIRCH.

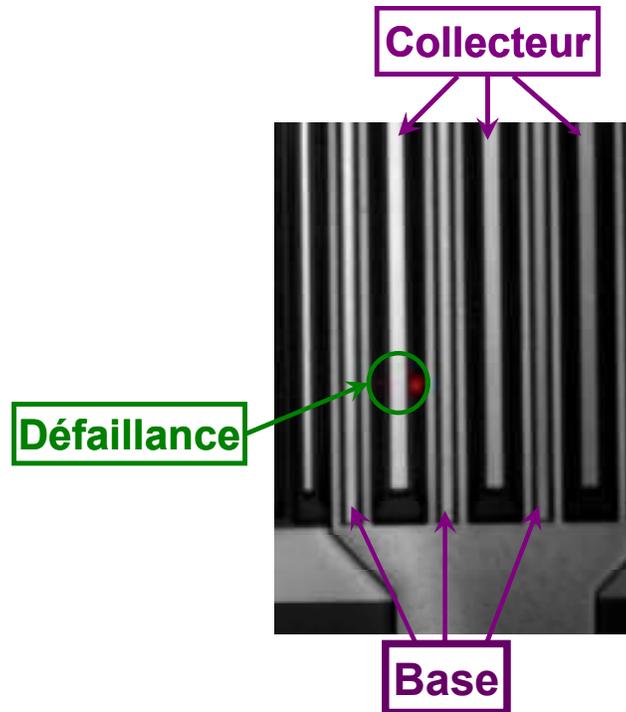


Figure 121: Zoom au voisinage du point de défaillance.

4.3. Structures à région flottante

Actuellement, la seule application connue des régions flottantes dans le domaine des ESD porte sur le réglage de la tension de déclenchement [31] [99]. Elle concerne des structures dans lesquelles le déclenchement est ajusté par un claquage latéral et la décharge est conduite à travers un transistor vertical (dont le collecteur est en général une couche enterrée). Cette configuration, précédemment décrite au Chapitre 1 (1.4.2.1), présente l'avantage d'une séparation spatiale de la partie commande (déclenchement) et de la partie puissance (conduction de l'ESD), qui peuvent être ainsi optimisées séparément [31]. La variation de la tension liée à l'insertion de la région flottante est alors mise à profit pour régler le déclenchement latéral, sans que l'état passant (tension de maintien comprise) ne soit affecté. En revanche, la présente étude traite de la conduction de l'ESD dans des structures latérales où le courant d'injection interfère avec la région flottante. L'objectif est d'évaluer les possibilités offertes pour la protection des E/S "hautes tensions", et, dans les cas les plus prometteurs, d'en déduire des stratégies de conception. Tout d'abord, le fonctionnement statique est décrit, de manière à avoir les bases théoriques pour l'étude plus complexe du fonctionnement en ESD. Ensuite, celui-ci est abordé, successivement dans les cas d'une région flottante dans la base et dans le collecteur. Les mécanismes contrôlant le repliement, la tension de maintien et le R_{ON} sont enfin approfondis.

4.3.1. Claquage statique

Les variations de la tension de claquage statique sont largement documentées dans la littérature, en particulier dans le cadre des techniques de garde, visant à augmenter la tenue en tension. Ces résultats guideront l'étude en simulation d'un transistor NPN décrit à partir des dopages de la technologie

SmartMOS 8 MV. L'objectif de cette simulation est d'identifier les principaux effets induits par la région flottante sur le comportement statique.

4.3.1.1. Tension de claquage en fonction de la position de la région flottante

Structure étudiée et terminologie

La structure considérée (Figure 122) est un transistor NPN latéral, c'est-à-dire sans couche enterrée dopée N. Il est réalisé dans une épitaxie dopée P⁻, et ses principales caractéristiques sont un collecteur dopé N⁺ et une base formée d'une région dopée P⁺ plus une région modérément dopée P au voisinage de la jonction avec le collecteur. La région flottante N⁺ est insérée dans le dopage P de base. Sa position est repérée avec la distance par rapport à la jonction base-collecteur, notée d. L'étude portera sur le claquage en configuration émetteur et base court-circuités, ce qui équivaut à considérer la diode base-collecteur.

Pour alléger la terminologie, la jonction base-collecteur sera appelée "jonction principale" et la jonction base-région flottante en regard de la région émetteur-base sera appelée "jonction secondaire".

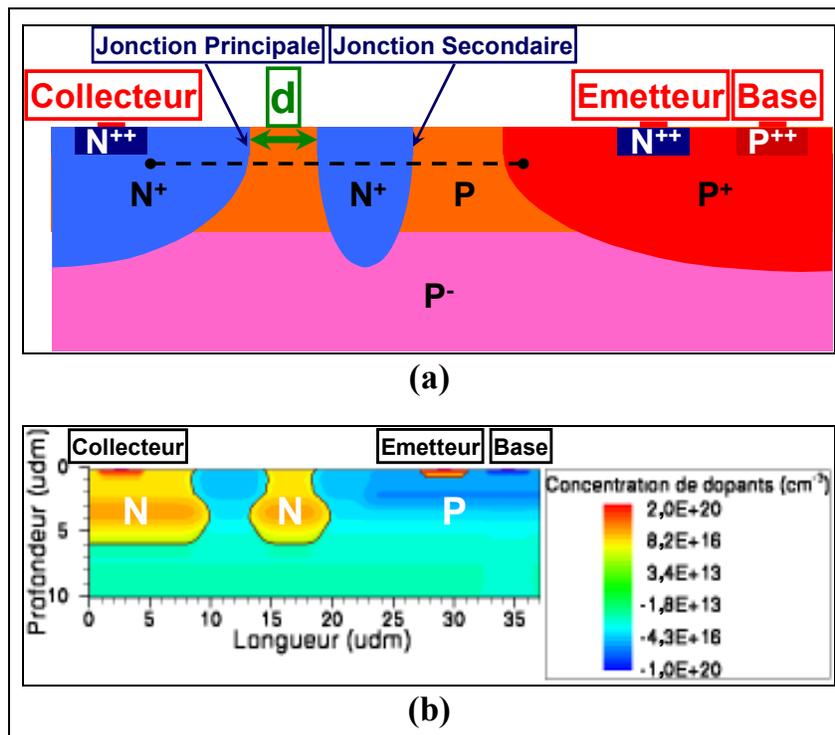


Figure 122: Coupe 2D simplifiée (a) et profil de dopage (b) du transistor NPN retenu pour l'étude statique.

Résultats de simulation

Une simulation préliminaire donne la tension de claquage, notée BV_{SR} , d'une structure sans région flottante. Ensuite, le claquage est simulé pour différentes distances d (Figure 123). La distance d est modifiée en éloignant à la fois la région flottante et la zone émetteur-base, de sorte que la configuration du côté de la jonction secondaire reste inchangée. Les résultats montrent une croissance de la tension avec l'éloignement de la région flottante jusqu'à un maximum égal à deux fois la tension de référence. Au-delà, la tension diminue graduellement et tend vers la tension de référence.

La comparaison de ce résultat avec celui obtenu par P. Besse [31] est intéressante. Les deux courbes ont la même tendance générale. Toutefois, dans son cas, le maximum n'atteint pas le double de la

tension de référence. Cette particularité s'explique par le passage d'un claquage latéral à un claquage vertical (c'est-à-dire avec une couche enterrée dopée N), contrairement à l'étude présente où le claquage reste latéral. En technologie SmartMOS 8 MV, le claquage vertical doit être évité dans les protections "hautes tensions", à cause des faibles tensions de maintien correspondantes. Aussi, seul le cas où le claquage reste latéral sera traité par la suite.

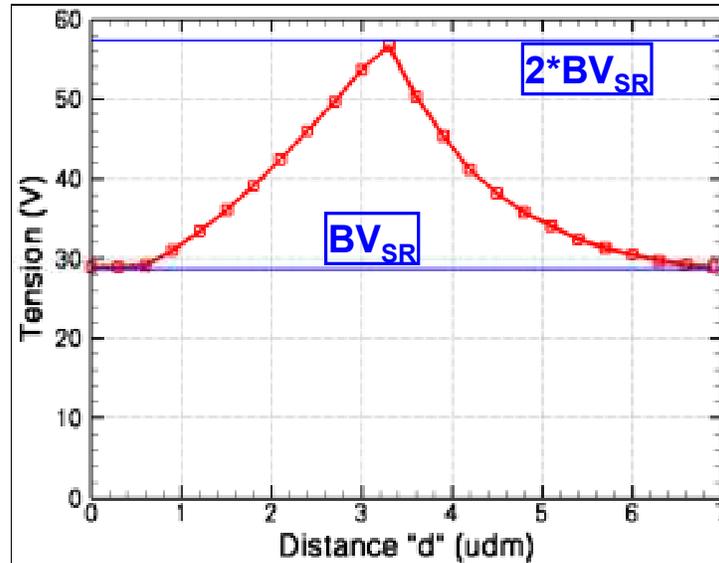


Figure 123: Tension de claquage statique en fonction de la position de la région flottante (BV_{SR} tension de claquage sans région flottante).

4.3.1.2. Modèle théorique unidimensionnel

L'approche la plus courante pour décrire l'influence de la région flottante consiste à ramener le problème à une dimension [100], en relevant le profil de dopage selon la ligne représentée par les tirets sur la Figure 122(a). L'évolution de la ZCE est alors étudiée avec l'augmentation de la tension de polarisation. Initialement, celle-ci est localisée à la jonction principale et s'étend jusqu'à être bloquée par la région flottante. Le potentiel de la région flottante est alors lié à celui du collecteur, et l'augmentation de la polarisation entraîne la formation d'une ZCE à la jonction secondaire. Cette seconde ZCE s'étend jusqu'à ce que le champ de claquage soit atteint (Figure 124). Ainsi, la tension totale est égale à la somme de la tension de claquage de la jonction secondaire plus la tension soutenue à la jonction principale. Cette contribution à la jonction principale dépend de l'extension de la ZCE correspondante, et est d'autant plus importante que la distance d est grande. Cette description relativement simple explique la croissance de la tension avec l'éloignement de la région flottante. Le maximum est atteint lorsque le champ critique est soutenu sur les deux jonctions. La tension est alors effectivement égale au double de la tension de claquage de référence.

Cependant, cette approche n'explique pas la décroissance progressive de la tension. Une chute brutale serait attendue, puisque la jonction principale claquerait avant que la ZCE ne touche la région flottante. Dans ce cas, la région flottante n'aurait plus d'influence, et la tension serait égale à la tension de claquage de référence. Cette incohérence est liée aux limites du modèle unidimensionnel, et ne peut être résolue que par la prise en compte des effets bidimensionnels [101] [102].

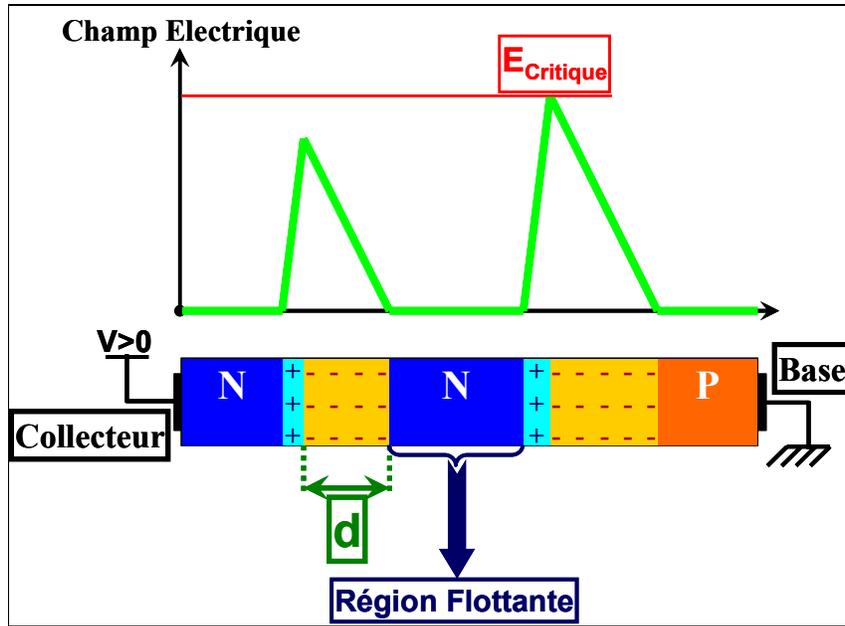


Figure 124: Répartitions unidimensionnelles de la charge et du champ électrique au claquage.

4.3.1.3. Modèle précis incluant les effets 2D

Avalanche et champ électrique dans la décroissance de la tension de claquage

Lorsque la tension atteint son maximum, l'avalanche se déplace de la jonction secondaire à la jonction principale. Pendant la décroissance, elle est exclusivement localisée à la jonction principale (Figure 125). Concernant le champ électrique, sa distribution montre un recouvrement entre les champs soutenus par la jonction du collecteur et par la jonction de la région flottante (Figure 126). Leur superposition conduit à une distribution continue englobant à la fois le collecteur et la région flottante. Une des principales limites du modèle 1D est de négliger cette interaction entre les champs électriques. Seul un modèle 2D peut la prendre en compte et décrire précisément la polarisation de la région flottante.

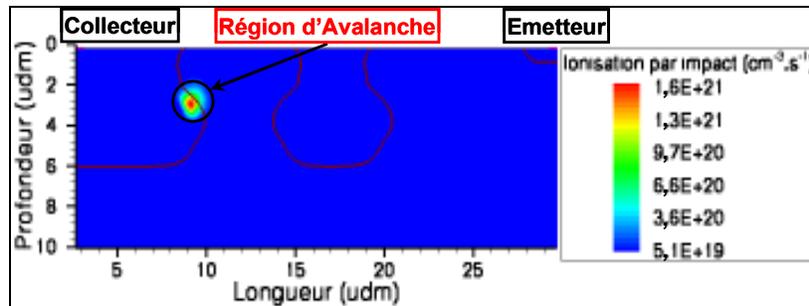


Figure 125: Distribution du taux d'ionisation par impact au claquage dans un composant de distance d 3,9 udm ($BV=45,4$ V).

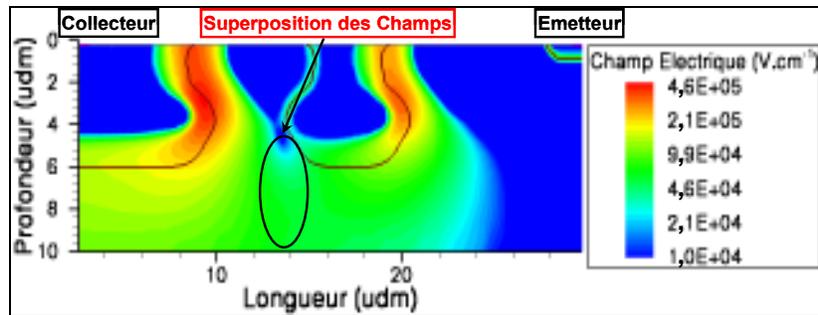


Figure 126: Distribution du champ électrique au claquage dans un composant de distance d 3,9 udm, montrant la superposition des champs soutenus aux jonctions du collecteur et de la région flottante.

Aspect du champ sans région flottante

Avant d'étudier les interactions entre les deux champs, il est nécessaire de traiter préalablement l'aspect du champ sans région flottante (Figure 127). On remarque qu'il s'étend plus loin dans les régions profondes que dans celles plus proches de la surface. Sa pente est plus faible dans le substrat P- faiblement dopé que dans l'implantation P en surface, de sorte qu'il doit être plus étendu en profondeur pour tenir une tension donnée. Par ailleurs, l'avalanche n'est pas localisée au niveau de cette zone de forte extension du champ, elle se situe clairement au-dessus. L'ionisation par impact est plus intense dans la zone de fort champ électrique, où le dopage est maximal. Dans ces conditions, même si l'espacement entre le collecteur et la région flottante est supérieur à l'extension du champ de claquage, le champ peut rester en contact avec la région flottante dans les régions profondes. La région flottante peut donc être polarisée alors que le claquage survient à la jonction principale, contrairement à ce que laissait supposer le modèle unidimensionnel. Maintenant, il reste à expliquer comment le champ à la jonction principale continue d'augmenter après avoir touché la région flottante, alors que la ZCE ne peut plus s'étendre.

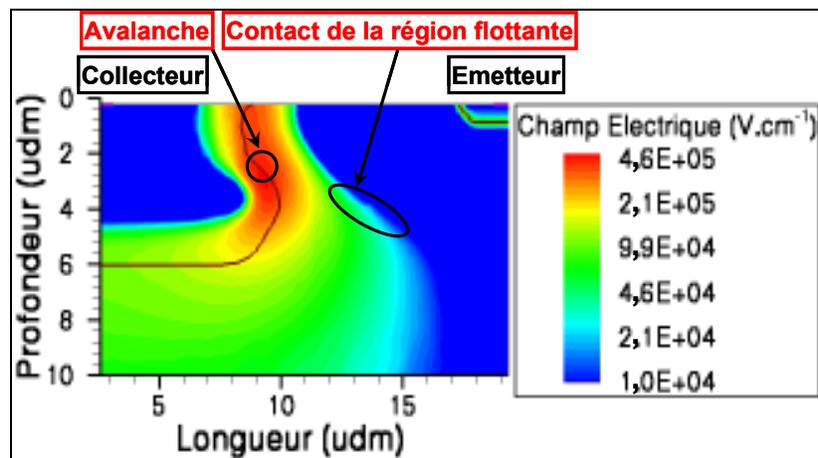


Figure 127: Distribution du champ électrique au claquage de la structure de référence (la région d'avalanche et la région où s'établit le contact avec la région flottante sont indiquées).

Interactions entre le champ à la jonction base-collecteur et de la région flottante

Lorsque le champ électrique à la jonction base-collecteur (appelé ultérieurement champ de collecteur) touche la région flottante, celle-ci se polarise et le champ électrique sur son pourtour augmente. A l'endroit où les deux champs se rencontrent, ils s'opposent l'un à l'autre, de sorte que leur résultante s'annule. Le champ de collecteur doit alors s'étendre pour garder le contact et maintenir la polarisation. Ainsi, son augmentation résulte de l'effet que lui oppose le champ induit le

long de la région flottante. Dans ces conditions, seule la zone située près de la région flottante est significativement modifiée, comme le montre la comparaison entre les distributions de champ électrique avec et sans région flottante (Figure 128).

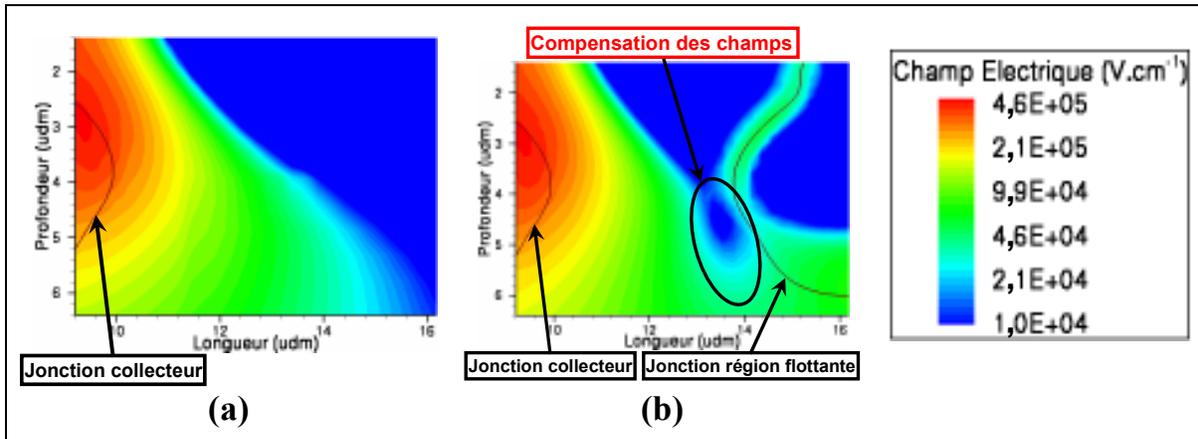


Figure 128: Grossissement des distributions de champ électrique au voisinage de la région où s'établit le contact avec la région flottante, pour la structure de référence (a) et la structure de distance d 3,9 udm (b).

L'interaction peut être caractérisée graphiquement en délimitant les zones dans lesquelles les lignes de champ issues de la région flottante supportent la polarisation ou seulement le potentiel de diffusion (Figure 129). Plus la zone supportant la polarisation est étendue, au détriment de celle supportant le potentiel de diffusion, plus le champ de la région flottante s'oppose fortement à celui du collecteur.

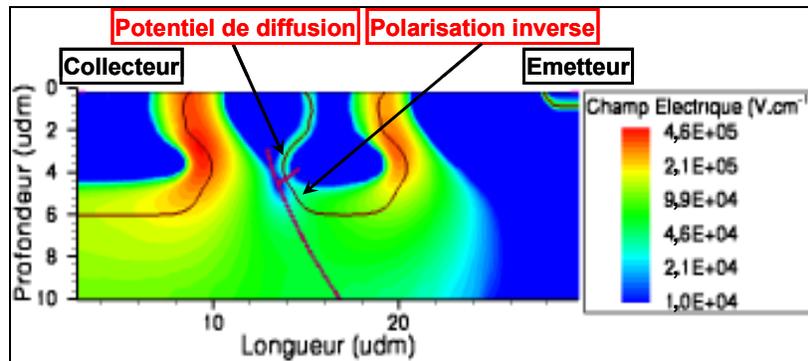


Figure 129: Distribution du champ au claquage du composant de distance d 3,9 udm, faisant apparaître les régions où les lignes de champ soutiennent soit le potentiel inverse soit le potentiel de diffusion.

Retour sur les variations de la tension de claquage

Il est maintenant possible d'expliquer les variations de la tension de claquage en fonction de la position de la région flottante (Figure 123). L'espacement d fixe le champ de collecteur au moment où se fait le contact avec la région flottante. Ce champ est d'autant plus élevé que d est grand. Une fois le contact avec la région flottante établi, les deux champs augmentent simultanément. Si le claquage survient à la jonction de la région flottante, la tension croît avec la distance d , la contribution du champ de collecteur étant de plus en plus élevée. Si le claquage survient à la jonction du collecteur, la tension décroît avec la distance d , l'effet exercé par le champ de la région flottante pour permettre au champ de collecteur d'atteindre le champ critique est de plus en plus faible. La diminution graduelle de cet effet explique que la décroissance de la tension soit progressive et non pas abrupte. L'interaction des champs en fonction du paramètre d est illustrée par le déplacement du point séparant la jonction de la région flottante entre la partie soutenant la polarisation et la partie soutenant le

potentiel de diffusion (Figure 130). L'affaiblissement de l'interaction avec l'augmentation de la valeur de d est caractérisé par la réduction de la partie soutenant la polarisation au profit de celle soutenant le potentiel de diffusion.

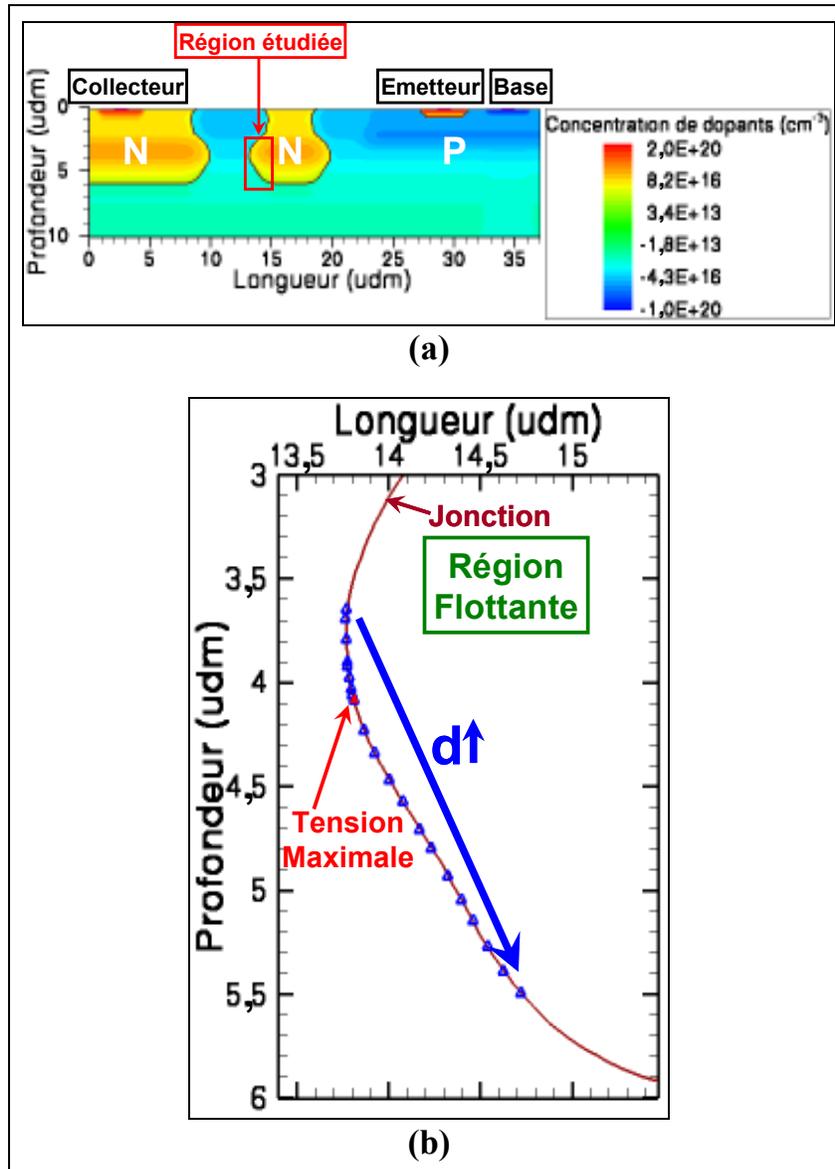


Figure 130: Profil de dopage montrant la région étudiée sur le composant de distance d 3,9 udm (a) et représentation de la jonction de la région flottante indiquant les points séparant les zones soutenant la polarisation en inverse ou en direct (b) pour des distances d allant de 0,9 à 6,9 udm par pas de 0,3 udm.

La variation simultanée de la tension des deux jonctions est défavorable pour des protections ESD "hautes tensions" et à faible R_{ON} . La diminution de la tension lors du repliement et son augmentation à l'état passant risquent d'être amplifiées du fait de cette double contribution. Cependant, ce résultat n'est pas suffisant pour conclure quant à l'effet de la région flottante lors d'une ESD. Il est nécessaire de déterminer le fonctionnement en mode autopolarisé, c'est-à-dire l'interaction entre l'injection et l'avalanche, et l'impact des fortes densités de courant. Ces effets dépendent de la position de la région flottante, dans la base ou dans le collecteur. Chaque cas doit être traité séparément.

4.3.2. Fonctionnement ESD pour une région flottante dans la base

Cette étude sera abordée par l'analyse en simulation d'un cas particulier. Les résultats obtenus permettront d'identifier des phénomènes physiques généraux. A partir de cette base théorique, les possibilités en termes de "haute tension" et de faible R_{ON} seront évaluées et, si ces possibilités s'avèrent prometteuses, des stratégies de développement idoines seront définies.

4.3.2.1. Présentation des résultats de simulation

La structure considérée est la même que celle utilisée pour l'étude du claquage statique (Figure 122). Son comportement vis-à-vis des ESD n'a pas été optimisé. Sa surface pourrait être réduite (surdimensionnement de la taille des contacts de base et d'émetteur et de la largeur de base intrinsèque), et le profil de dopage base-collecteur n'a pas été ajusté pour l'optimisation du R_{ON} . Elle n'en reste pas moins utilisable pour décrire et comprendre l'influence de la région flottante, ainsi que pour obtenir une première évaluation des performances atteignables.

Les caractéristiques TLP sont simulées pour différentes positions de la région flottante (Figure 131), (Tableau 24). La base est reliée à l'émetteur par une résistance de 1 k Ω afin de faciliter le déclenchement. Ainsi, la tension de déclenchement est proche de la tension de claquage statique. Pour des espacements d allant de 0,3 à 1,5 μm , la tension de maintien augmente et la chute de tension au repliement diminue. A 1,5 μm , le repliement est quasi-nul et la tension de maintien a augmenté de 85% par rapport à la structure de référence. A fort courant, les caractéristiques se rejoignent. Si l'on utilise la structure de référence, il est nécessaire d'en mettre deux en série pour obtenir une tension et un R_{ON} comparables à ceux des structures à régions flottantes (Figure 131). Dans ce cas, la surface est multipliée par deux, de sorte que cette structure prend plus de place que celles avec région flottante. Au total, le facteur de mérite du R_{ON} des structures avec région flottante est plus faible.

En résumé, la région flottante a permis d'améliorer les performances vis-à-vis des trois paramètres que l'on cherche à contrôler : le repliement, la tension de maintien et le R_{ON} . En revanche, le nouveau mode de fonctionnement, qui apparaît pour un espacement de 1,8 μm , est caractérisé par un fort repliement avec une tension de maintien proche de celle de la structure de référence.

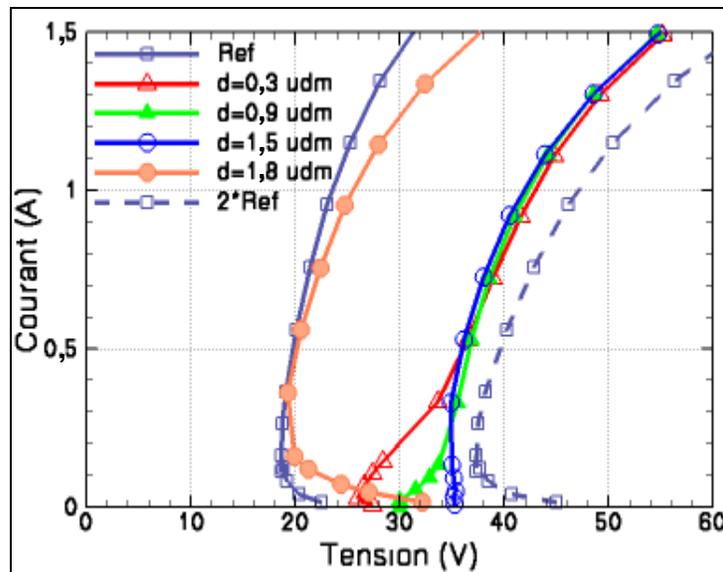


Figure 131: Caractéristiques TLP simulées pour différentes positions de la région flottante et pour la structure de référence (seule ou deux mises en série).

Tableau 24: Principaux paramètres électriques en fonction de la distance d.

d (udm)	BV (V)	V _H (V)	BV-V _H (V)	R _{ON} (Ω)	F _{RON} (μm ² .W ⁻¹)
Ref	28,7	18,7	10,0	20,6	181
0.3	28,9	26,0	2,9	31,6	118
0.9	31,0	30,0	1,0	32,2	125
1.5	36,1	34,9	1,2	33,2	130
1.8	39,1	19,3	19,8	32,7	261
2*Ref	57,4	37,4	20,0	41,2	181

Les valeurs des R_{ON} sont calculées au voisinage de 1,5 Ampères.

4.3.2.2. Mécanismes physiques

Suivant que l'avalanche est localisée à la jonction principale ou à la jonction secondaire, les mécanismes de forte injection sont différents. Pour les espacements de 0,3, 0,9 et 1,5 udm, l'avalanche est essentiellement localisée à la jonction secondaire dans la partie bas courant (respectivement jusqu'à 300, 150 et 50 milliAmpères pour les espacements de 0,3, 0,9 et 1,5 udm). A plus fort courant, en particulier lorsque les caractéristiques se rejoignent, l'avalanche s'est déplacée à la jonction principale. Pour la structure d'espacement 1,8 udm, l'avalanche est localisée à la jonction principale, quel que soit le niveau de courant.

Avalanche à la jonction secondaire

Les contributions à la tension des différentes ZCE et résistances internes sont séparées les unes des autres en chaque point de la caractéristique TLP, pour l'espacement 0,9 udm. Pour cela, une coupe est effectuée en suivant la ligne de courant passant par le point de la jonction secondaire où la densité de courant est maximale (Figure 132 (a)). Le profil de champ le long de cette coupe montre cinq contributions à la tension : les champs électriques dans les ZCE aux jonctions secondaire et principale (nommés respectivement "Champ1" et "Champ2"), le champ électrique induit par la saturation de la vitesse des porteurs dans la diffusion de collecteur (nommé "Champ3"), et les résistances d'accès dans le collecteur et la région flottante (nommées respectivement "Res1" et "Res2") (Figure 132 (b)). La tension associée à chacune de ces contributions pourra être calculée en intégrant le champ électrique sur la région correspondante.

A bas courant, lorsque l'avalanche est localisée à la jonction secondaire, seules les contributions "Champ1" et "Champ2" sont significatives. Le tracé des caractéristiques électriques I-V de ces deux contributions permet de mettre en évidence leurs influences respectives sur la caractéristique TLP (Figure 133). La tension à la jonction secondaire ("Champ1") diminue selon les mêmes mécanismes que dans un simple bipolaire autopolarisé (contribution du courant d'injection à l'avalanche et diminution de l'extension du champ dans la base sous l'effet du courant d'injection). La chute de tension résultante est importante, de plus de 10 Volts (l'étude à fort courant montrera une tension de maintien de 18,2 Volts, à comparer à la tension de déclenchement de 29,4 Volts). Cependant, la tension à la jonction principale ("Champ2") augmente, de sorte que la caractéristique TLP de la structure complète ne se replie que d'un Volt. Au premier abord, cette augmentation est surprenante. Elle est en contradiction avec la conclusion de l'étude statique selon laquelle les tensions soutenues par les deux jonctions varient dans le même sens (section 4.3.1.3). Cependant, en forte injection, cet effet est neutralisé par la modulation de la ZCE. Du côté de la région flottante, la limite de la ZCE est fixée par les conditions de polarisation. Ainsi, le long de la coupe (Figure 132 (a)), elle s'étend jusqu'à la région flottante, quelque soit le point TLP. L'augmentation du dopage effectif dans la région P entre les deux jonctions et le déplacement de la jonction effective vers le collecteur conduisent à un accroissement à la fois de valeur maximale et de l'extension du champ électrique (Figure 134), donc à un accroissement de la tension. Par comparaison au cas de la jonction base-

collecteur d'un simple transistor NPN, où l'extension de la ZCE côté base peut diminuer, les contraintes liées à la polarisation de la région flottante empêchent une telle diminution.

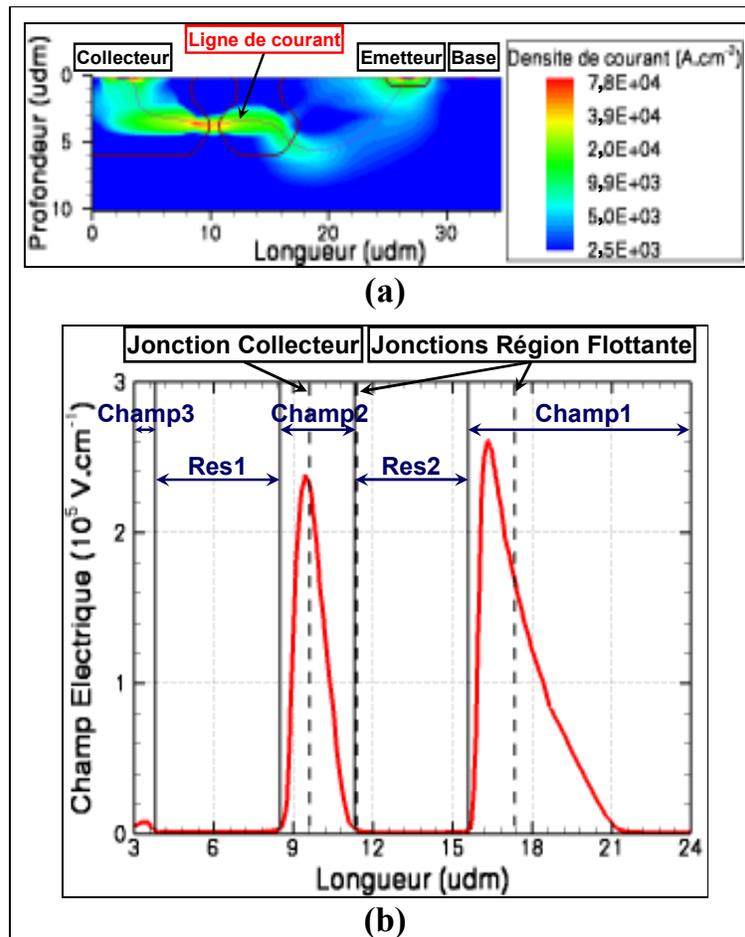


Figure 132: Distribution de la densité de courant dans le composant de distance 0,9 udm au point TLP ($V=33,7\text{ V}$; $I=132\text{ mA}$) (a), montrant la ligne de courant pour tracer le profil du champ électrique (b).

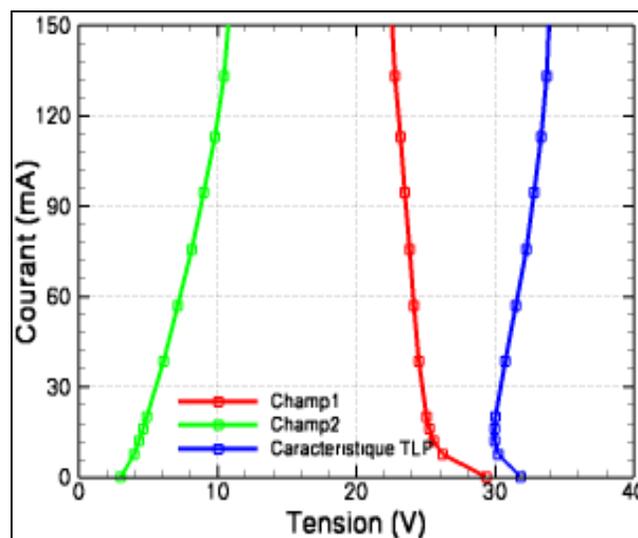


Figure 133: Caractéristique TLP bas courant et courbes I-V des contributions "Champ1" et "Champ2".

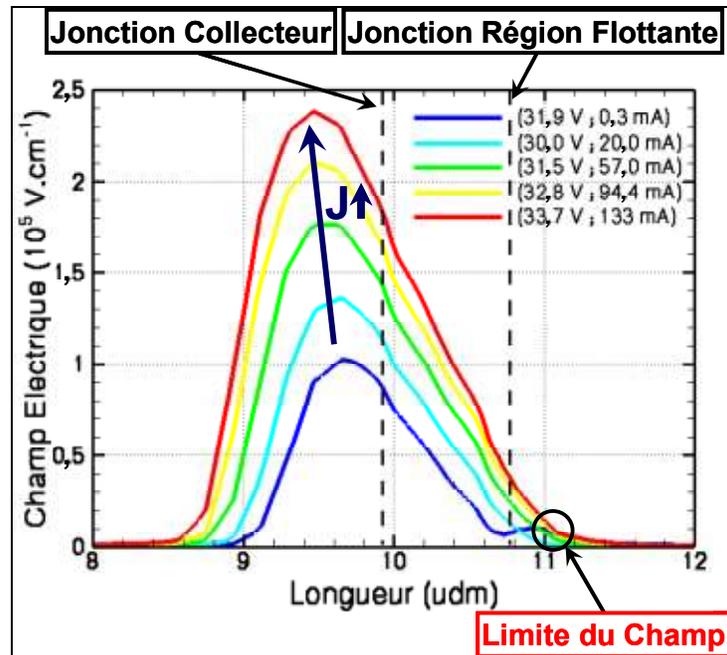


Figure 134: Profils de champ électrique à la jonction principale (Champ2) pour différents points TLP.

Avalanche à la jonction principale

A nouveau, les différentes contributions à la tension sont séparées les unes des autres, pour la structure d'espacement 0,9 udm. Cette fois-ci, les niveaux de courant sont suffisants pour que les contributions des résistances (contributions Res1 et Res2) et du champ dans la diffusion de collecteur (contribution Champ3) soient significatives (Figure 135).

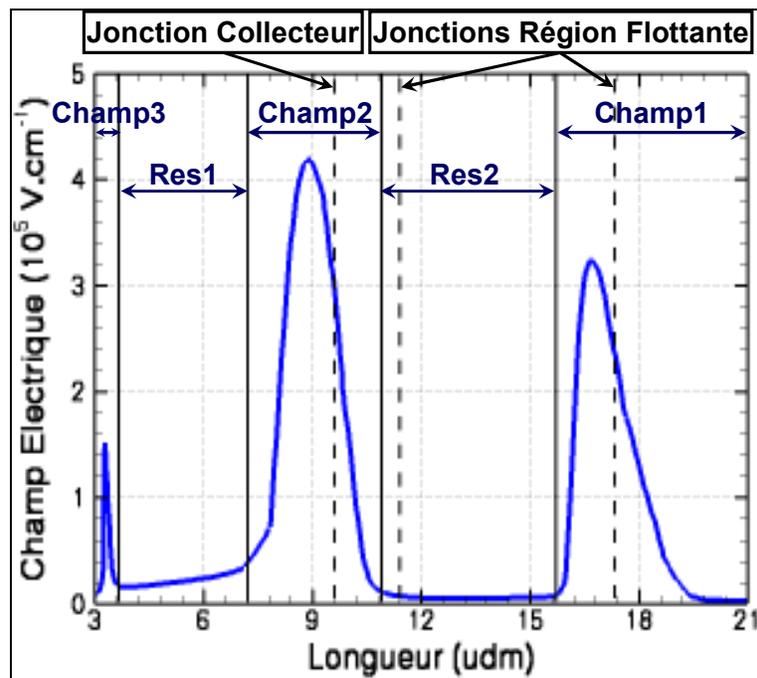


Figure 135: Profil de champ électrique suivant la ligne de courant issue du maximum de densité de courant à la jonction principale au point TLP ($V=54,6$ V ; $I=1,49$ A).

Les caractéristiques I-V des contributions de la jonction secondaire (Champ1) et de la jonction principale (Champ2) sont tracées (Figure 136) et la résistance dynamique à 1,5 Ampères de chacune des contributions est calculée (Tableau 25). Le champ à la jonction principale et celui dans la diffusion de collecteur ont les résistances dynamiques les plus importantes, correspondant respectivement à la moitié et à un tiers du R_{ON} total. Comparativement, les autres contributions sont très faibles. Toutefois, il est possible que la résistance dynamique à la jonction secondaire ne soit pas négligeable dans une structure optimisée. La suppression du champ dans la diffusion de collecteur ne devrait pas poser problème, il suffit de doper plus le collecteur et d'agrandir le contact. Concernant le contrôle des champs aux jonctions principales et secondaires, il faut au préalable comprendre les mécanismes physiques mis en jeu.

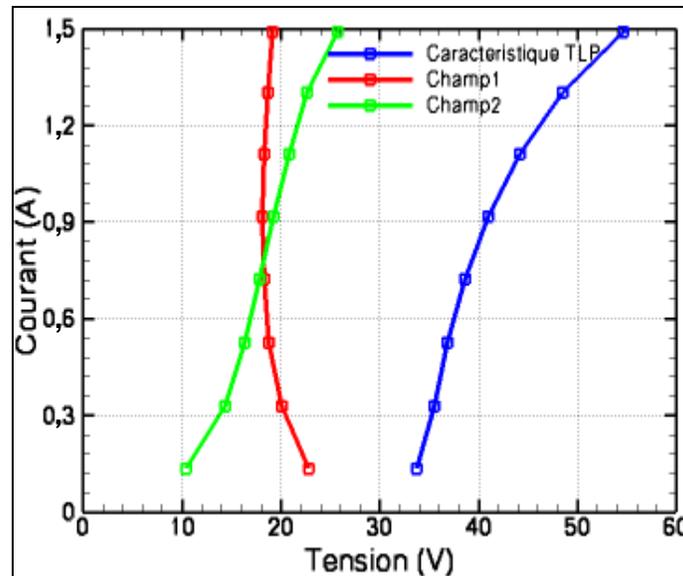


Figure 136: Caractéristique TLP fort courant et courbes I-V des contributions "Champ1" et "Champ2".

Tableau 25: R_{ON} à 1,5 Ampères proportion dans la valeur totale pour chaque contribution de la tension.

	R_{ON} (Ω)	Proportion (%)
Champ1	2,8	8,8
Champ2	16,3	50,9
Champ3	10,7	33,4
Res1	1,4	4,4
Res2	0,8	2,5
Total	32,0	-

Le champ à la jonction principale est fonction du niveau d'avalanche nécessaire pour polariser le bipolaire. Sa résistance dynamique équivalente dépend essentiellement de l'optimisation des effets en fort courant, suivant les stratégies définies au Chapitre1. Le champ à la jonction secondaire est fixé par le potentiel qu'induit le champ à la jonction principale. Pour appréhender l'effet des fortes densités de courant sur la polarisation, les distributions de champ électrique dans la structure de référence à fort courant (Figure 137) et en statique (Figure 127) sont comparées. L'effet le plus significatif est la suppression de la forte extension en profondeur. Cette nouvelle configuration ne conduit plus à une polarisation graduelle de la région flottante, de sorte que les variations de la tension sont plus franches. Cependant, la distribution du champ de référence ne décrit qu'approximativement la distribution à la jonction principale. La région flottante affecte la répartition de la densité de courant (Figure 138), ce qui influe sur la distribution spatiale des phénomènes de forte injection.

Le champ électrique à la jonction principale n'étant plus fixé par les contraintes liées à la polarisation de la région flottante, on s'attend à ce que la forte injection entraîne une diminution de son extension. Cette diminution modifie la polarisation de la région flottante, dont le potentiel devrait diminuer avec le courant. Plus l'espacement d est faible, plus la densité de courant doit être élevée pour que la modification du champ soit suffisante pour faire chuter le potentiel de la région flottante. Cette analyse permet de décrire le comportement de la structure d'espacement 1,8 μm . Par contre, elle n'explique pas les résultats pour les espacements de 0,3, 0,9 et 1,5 μm . En particulier, elle ne rend pas compte de la superposition des caractéristiques TLP à fort courant, et de l'augmentation de la tension soutenue à la jonction secondaire (Figure 136). Pour lever ces contradictions, il faudrait étudier plus précisément les modifications du champ électrique en fonction de la répartition de la densité de courant. Une étude aussi fine est ardue et les résultats risquent d'être trop complexes pour en déduire des règles d'optimisation réalistes. Aussi, l'analyse du comportement pendant une ESD ne sera pas plus amplement approfondie.

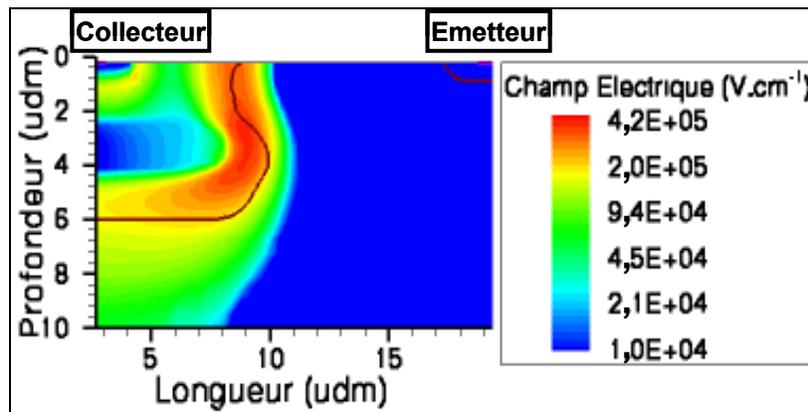


Figure 137: Distribution de champ électrique dans la structure de référence au point TLP ($V=32,1\text{ V}$; $I=1,54\text{ A}$).

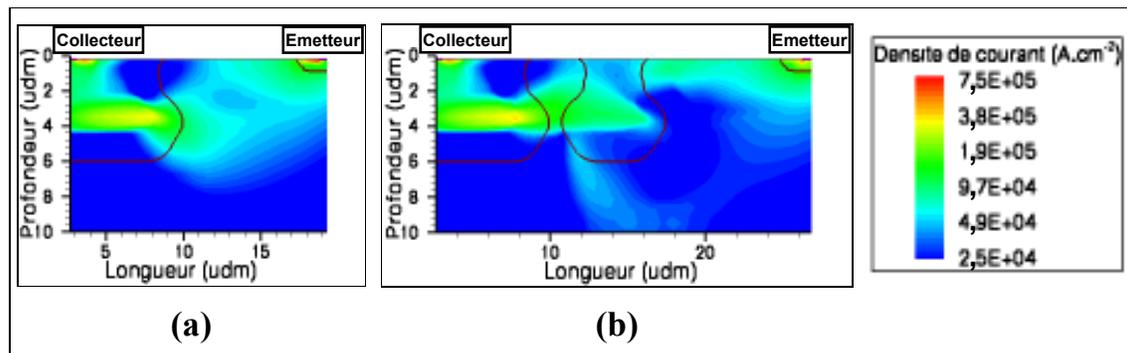


Figure 138: Distributions de densité de courant dans la structure de référence au point TLP ($V=32,1\text{ V}$; $I=1,54\text{ A}$) (a), et dans la structure de distance 0,9 μm au point TLP ($V=54,6\text{ V}$; $I=1,49\text{ A}$) (b).

Développement de protections "hautes tensions" et à faible R_{ON}

Dans le meilleur cas ($d=1,5\text{ }\mu\text{m}$), l'insertion de la région flottante permet de supprimer le repliement, d'augmenter la tension de fonctionnement, et d'améliorer le R_{ON} . Toutefois, le R_{ON} des structures étudiées reste trop élevé pour des protections ESD, et des règles d'optimisation n'ont pas pu être définies du fait des difficultés théoriques dans l'analyse des effets en fort courant. De plus, il est probable que la règle fondamentale de réduction du R_{ON} consistant à diminuer le dopage de base (définie en 2.4.3) ne soit pas applicable à la jonction principale. Cette règle vise à réduire l'extension du champ ce qui risque de poser des problèmes pour maintenir la polarisation de la région flottante et garder une tension de fonctionnement élevée. En résumé, les transistors NPN avec région flottante

dans la base sont susceptibles d'être utilisés comme protection ESD des E/S "hautes tensions". Cependant, l'absence de stratégie pour la réduction du R_{ON} rend l'optimisation très délicate.

4.3.3. Fonctionnement ESD pour une région flottante dans le collecteur

4.3.3.1. Simulation du comportement ESD

De même que pour la région flottante dans la base, l'étude est abordée par la simulation d'un cas particulier. Les résultats obtenus seront exploités pour décrire l'influence pendant une ESD de l'insertion d'une région flottante dans le collecteur du bipolaire.

Structure étudiée

La structure étudiée (Figure 139) est un transistor NPN latéral. Il est réalisé dans une épitaxie dopée P^- , et ses principales caractéristiques sont une base dopée P^+ et un collecteur formé d'une région dopée N^+ plus une région modérément dopée N au voisinage de la jonction avec la base. La région flottante P^+ est insérée dans le dopage N du collecteur. La concentration de dopants dans la région modérément dopée N est de l'ordre de 10^{16} cm^{-3} . Ultérieurement, il sera montré que cette valeur joue un rôle important dans l'optimisation.

La terminologie est la même que celle adoptée pour l'étude de la région flottante dans la base. La jonction base-collecteur sera nommée "jonction principale" et la jonction de la région flottante du côté du contact de collecteur "jonction secondaire". La position de la région flottante est repérée par l'espacement avec la jonction principale, notée d .

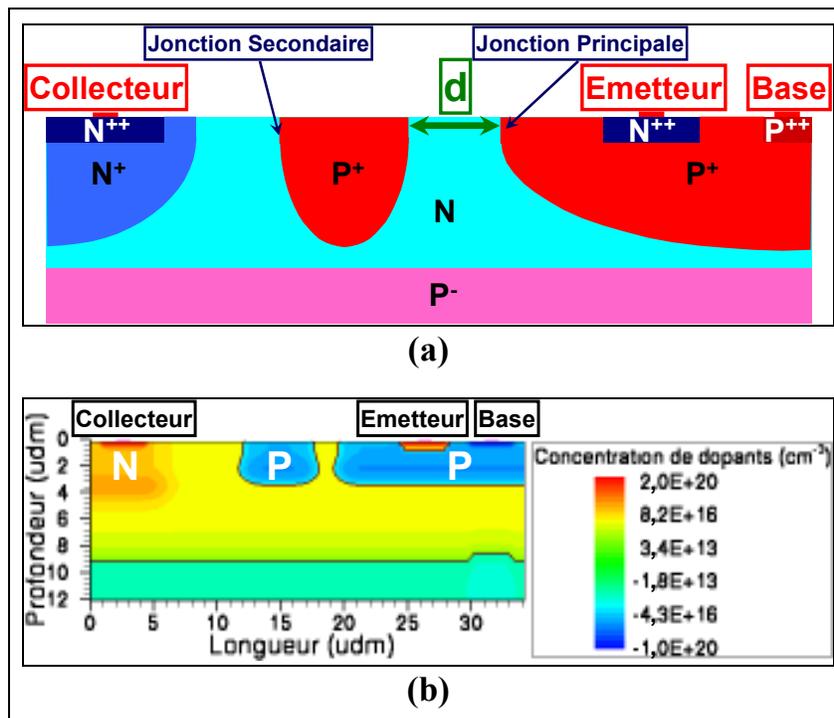


Figure 139: Coupe 2D schématique (a) et profil de dopage d'un transistor NPN latéral avec région flottante dans le collecteur de distance d 1,2 udm (b).

Résultats

Les tensions de claquage statique et les caractéristiques TLP sont simulées pour différentes distances d (Figure 140) (Tableau 26). Pour les simulations TLP, une résistance de $1\text{ k}\Omega$ est insérée entre la base et l'émetteur afin d'aider au déclenchement.

Si la valeur de d est inférieure à $1,8\text{ }\mu\text{m}$, les caractéristiques TLP ne se replient pas. La tension au déclenchement est égale à la tension de claquage statique. Pour $d=1,8\text{ }\mu\text{m}$, cette tension est plus de deux fois supérieure à la tension de maintien de la structure de référence. A bas courants (inférieurs à 300 milliAmpères), le R_{ON} est élevé. A plus fort courant, les caractéristiques TLP se rejoignent et leur R_{ON} est successivement positif, négatif puis à nouveau positif, résultant en une augmentation modérée de la tension. Ce comportement singulier rend difficile la comparaison avec la structure de référence. Néanmoins, la surtension due au fort R_{ON} à bas courant est plus importante pour la structure de référence. Sous cet angle, le comportement à l'état passant des structures avec région flottante est meilleur. Pour un espacement de $2,1\text{ }\mu\text{m}$, la caractéristique se replie à nouveau fortement, et ne présente donc plus d'intérêt pour les spécifications visées. Après ce repliement, la tension augmente très rapidement, et le comportement à fort courant est identique à celui pour des espacements inférieurs.

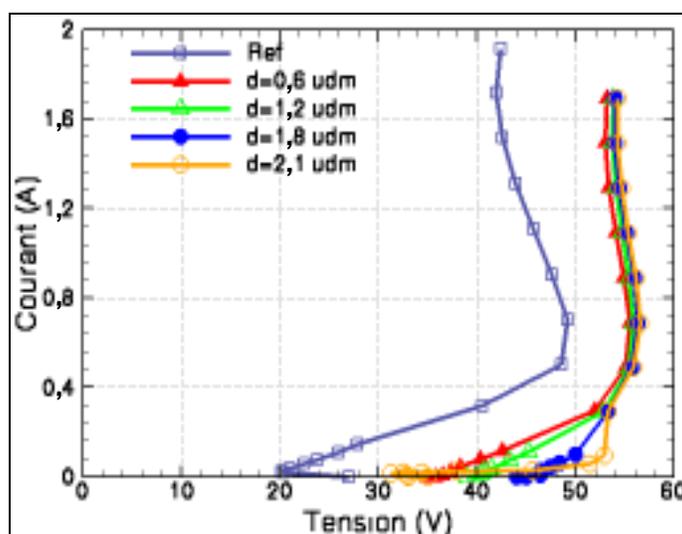


Figure 140: Caractéristiques TLP simulées pour la structure de référence et pour différentes positions de la région flottante.

Tableau 26: Paramètres électriques.

d (μm)	BV (V)	V_H (V)
Ref	33,3	20,2
0,6	34,6	-
1,2	38,0	-
1,8	43,8	-
2,1	47,5	31,2

4.3.3.2. Mécanismes physiques et stratégies d'optimisation

La suppression du repliement et les hautes tensions de fonctionnement observées en simulation constituent des résultats prometteurs. Le comportement à l'état passant semble également intéressant. Cependant, les performances atteignables sont difficiles à évaluer du fait des changements de variation du R_{ON} . L'étude des mécanismes physiques devra permettre de comprendre ces différentes propriétés et de définir les stratégies d'optimisation pour les mettre à profit.

Suppression du repliement

Les distributions du taux d'ionisation par impact (Figure 141), des densités de courants d'électrons (Figure 142) et de trous (Figure 143) sont relevées pour la structure d'espacement $d=1,2$ udm, au point TLP ($V=43,3$ V ; $I=73,4$ mA), c'est-à-dire juste après le déclenchement. Ces distributions montrent une séparation spatiale entre la région d'avalanche, localisée à la jonction secondaire de la région flottante, et le chemin du courant d'injection, contournant la région flottante. Ainsi, les porteurs injectés par l'émetteur ne contribuent pas à l'avalanche, et donc la caractéristique TLP ne se replie plus. La génération du courant de polarisation est assurée par l'augmentation du facteur de multiplication, entraînant une augmentation de la tension.

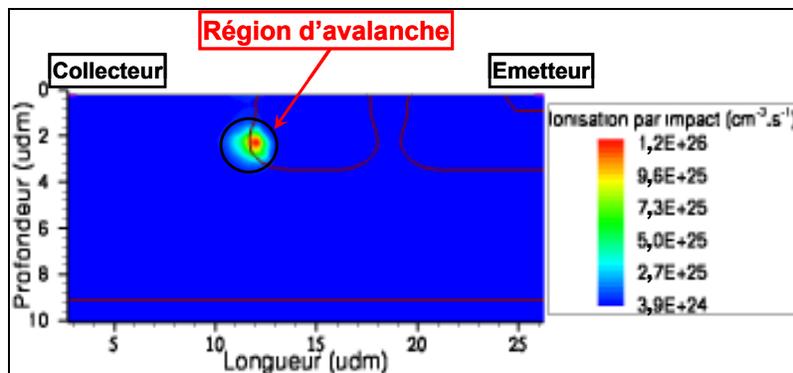


Figure 141: Distributions d'ionisation par impact dans le composant de distance $d=1,2$ udm au point TLP ($V=43,3$ V ; $I=73,4$ mA).

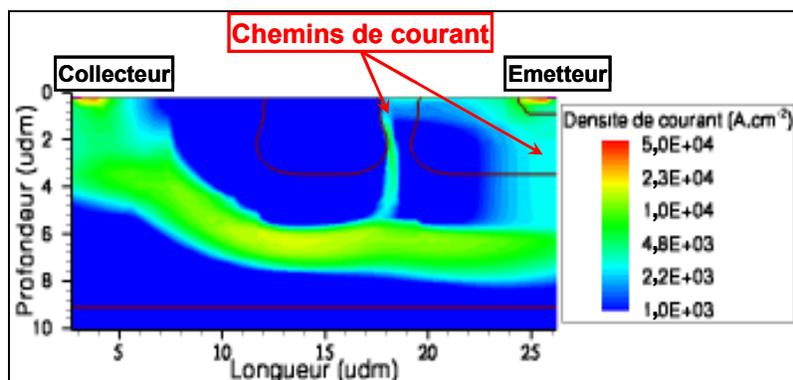


Figure 142: Distributions de densité de courant d'électrons dans le composant de distance $d=1,2$ udm au point TLP ($V=43,3$ V ; $I=73,4$ mA).

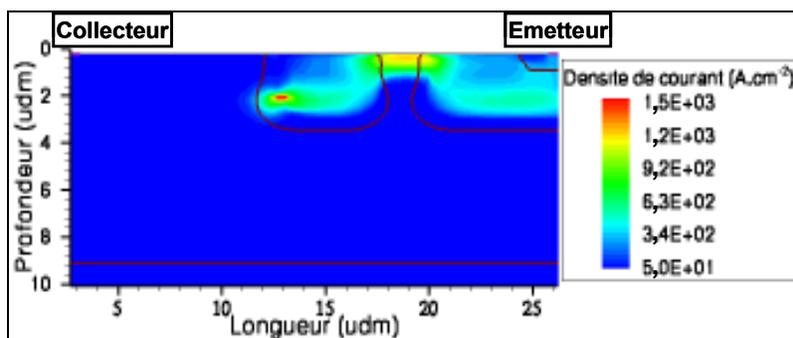


Figure 143: Distributions de la densité de courant de trous dans le composant de distance $d=1,2$ udm au point TLP ($V=43,3$ V ; $I=73,4$ mA).

Pour expliquer le contournement de la région flottante par le courant d'injection, il faut traiter séparément les flux verticaux et latéraux du courant dans la base. Concernant le chemin vertical, les électrons atteignent la région profonde dopée N à leur sortie de la ZCE base-collecteur. Ensuite, ils transitent par conduction à travers les différentes diffusions N jusqu'au contact du collecteur. En pratique, ce cas ne pourra pas être mis à profit si les tensions de maintien verticales sont inférieures à la limite basse de la fenêtre de conception. Concernant le flux latéral, l'explication du contournement est plus complexe. Une méthode consiste à analyser le fonctionnement du transistor NPN parasite inhérent à la région flottante. D'après l'étude en régime statique, la jonction de la région flottante est polarisée sur une partie en direct et sur une autre partie en inverse (Figure 129), pouvant être considérées respectivement comme les jonctions émetteur-base et base-collecteur. Ces zones polarisées en direct et en inverse se retrouvent également en régime dynamique, au cours d'une impulsion TLP (Figure 144), même si la zone polarisée en direct est peu étendue. Comme dans un transistor bipolaire autopolarisé, les trous générés par avalanche à la jonction base-collecteur (en inverse) transitent à travers la base (région flottante) jusqu'à la jonction émetteur-base (en direct). La polarisation de cette dernière permet le passage des trous par diffusion, sous l'effet du fort gradient entre la région flottante dopée P et la zone dépeuplée. Cette zone dépeuplée correspond à la ZCE de la jonction principale en contact avec la région flottante. Par contre, le gradient d'électrons est négligeable entre la zone dépeuplée, où seuls les électrons induits par le courant sont présents, et la région flottante P, où ils sont minoritaires. Les électrons ne peuvent donc pas passer la jonction en direct, et, en l'absence de courant de diffusion, ils suivent les lignes de champ qui contournent la région flottante (Figure 144).

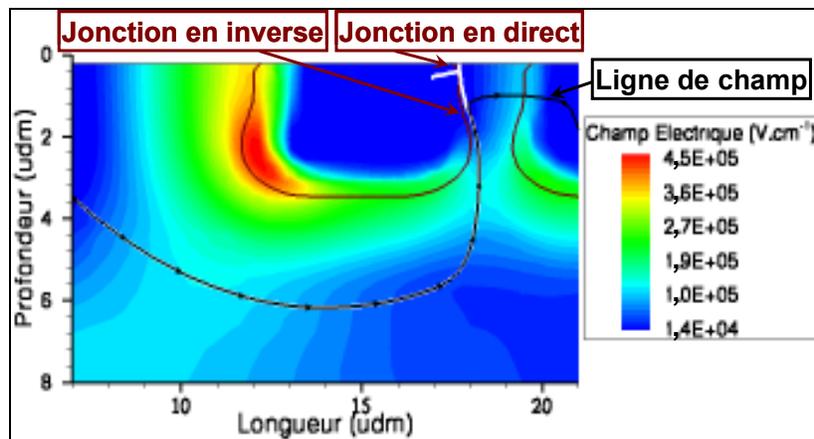


Figure 144: Distribution du champ électrique au voisinage de la région flottante au point TLP ($V=43,3\text{ V}$; $I=73,4\text{ mA}$), montrant les zones de la région flottante polarisée en direct et en inverse (séparées par les lignes blanches) et indiquant une ligne de champ suivie par les électrons (ligne noire).

Pour expliquer le repliement à partir d'un espacement de 2,1 udm, la distribution du taux d'ionisation par impact est relevée au point TLP ($V=31,2\text{ V}$; $I=15,5\text{ mA}$), correspondant à la tension de maintien (Figure 145). L'avalanche n'est plus localisée à la jonction secondaire de la région flottante, mais à la jonction verticale de la région de base. Ainsi, le courant vertical dans la base contribue à l'avalanche, d'où le repliement. On peut supposer qu'une structure, dont le courant de base est exclusivement latéral, ne devrait jamais se replier.

Les mécanismes physiques conduisant à la suppression du repliement semblent défavorables pour le R_{ON} . Du fait de la séparation spatiale entre l'avalanche et l'injection, la jonction soutenant l'avalanche se comporte comme une diode en inverse, dont le R_{ON} est généralement élevé (chapitre 3). Toutefois, seul le courant de polarisation doit être fourni par l'avalanche. Ainsi, le R_{ON} serait compris entre celui d'une diode à avalanche et celui d'un simple transistor NPN autopolarisé. Cependant, ces considérations ne prennent pas en compte les effets en fort courant dont l'influence sur le R_{ON} est, en général, déterminante.

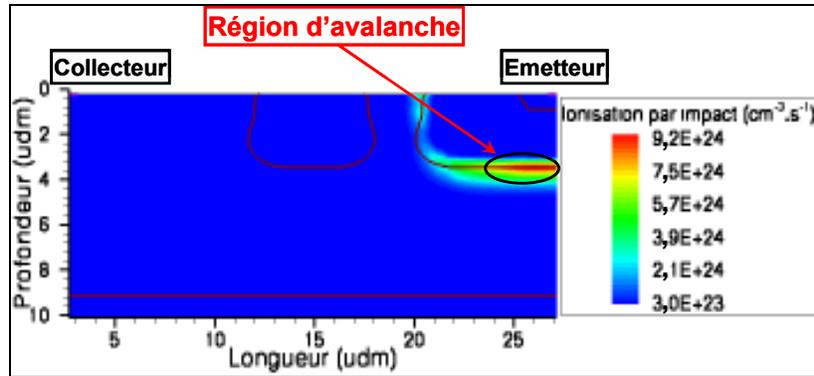


Figure 145: Distribution d'ionisation par impact au point TLP ($V=31,2$ V ; $I=15,5$ mA) de la structure d'espacement $d=2,1$ udm.

Etat passant

Les distributions du champ électrique (Figure 146), du taux d'ionisation par impact (Figure 147) et de la densité de courant (Figure 148) sont relevées pour la structure d'espacement $d=1,2$ udm, au point TLP ($V=55,4$ V ; $I=489$ mA) à partir duquel le R_{ON} diminue. Le champ électrique entre la base et la région flottante a disparu, la charge induite par le courant l'ayant repoussé vers la région dopée N^+ sous le contact de collecteur. La tension n'est plus soutenue que par la jonction secondaire, la contribution de la jonction principale tendant à s'annuler, d'où la diminution du R_{ON} . Dans certains cas, cette diminution peut être suffisamment importante pour conduire à un repliement. En parallèle, la région d'avalanche se déplace de la jonction secondaire vers une région située sur le chemin de courant. Le bipolaire passe alors en configuration autopolarisée, qui est optimale pour le R_{ON} . Ainsi, le fonctionnement en forte injection s'approche de celui d'un transistor NPN sans région flottante, mais dont la large base interne induit une tension élevée.

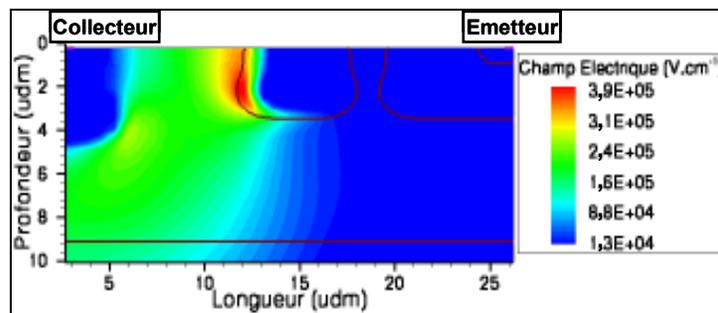


Figure 146: Distribution de champ électrique dans la structure d'espacement $1,2$ udm au point TLP ($V=55,4$ V ; $I=489$ mA).

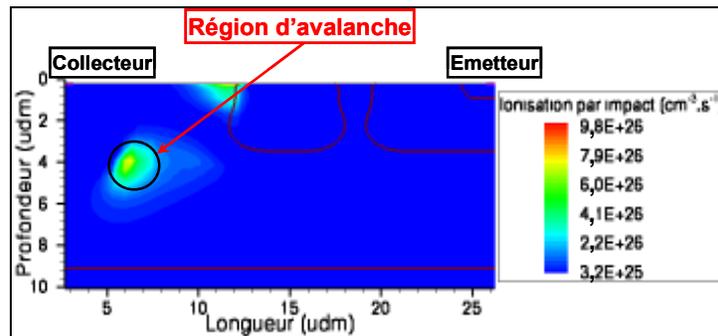


Figure 147: Distribution du taux d'ionisation par impact dans la structure d'espacement $1,2$ udm au point TLP ($V=55,4$ V ; $I=489$ mA).

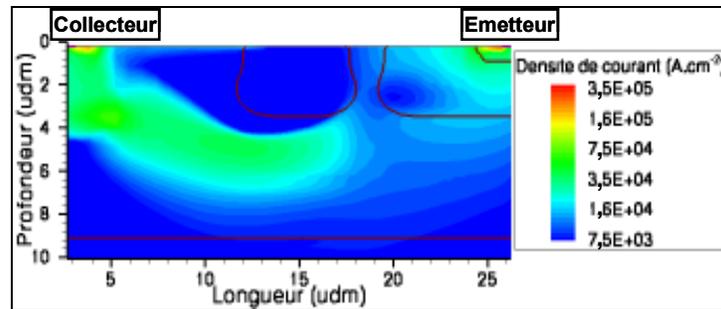


Figure 148: Distribution de la densité de courant dans la structure d'espacement 1,2 udm au point TLP ($V=55,4$ V ; $I=489$ mA).

Le niveau du dopage entre la base et la région flottante est un paramètre clé pour l'optimisation du R_{ON} . D'une part, il doit être suffisamment faible pour que le champ électrique soit rapidement repoussé sous l'effet de la forte injection, limitant ainsi la surtension à bas courant. D'autre part, s'il est trop faible, un changement prématuré de mode de fonctionnement risque de conduire à un repliement à une tension inférieure à la tension de déclenchement. Un autre paramètre très important est la diffusion N^+ sous le collecteur. Son profil doit être abrupt pour que l'évolution du champ soit bénéfique pour le R_{ON} , et sa position par rapport à la région flottante doit également être ajustée de manière à éviter que le comportement en diode aux faibles courants conduise à une trop forte surtension.

L'application de ces règles en simulation a permis de démontrer leur validité pour la protection des E/S "hautes tensions" (Figure 149). Il a été obtenu une caractéristique TLP qui ne se replie pas, dont la tension de déclenchement est de 30 Volts et le R_{ON} d'environ 3 à 4 Ohms, pour une surface de $100 \times 100 \mu\text{m}^2$. Le calcul du facteur de mérite donne une valeur comprise entre 25 et $35 \mu\text{m}^2 \cdot \text{W}^{-1}$.

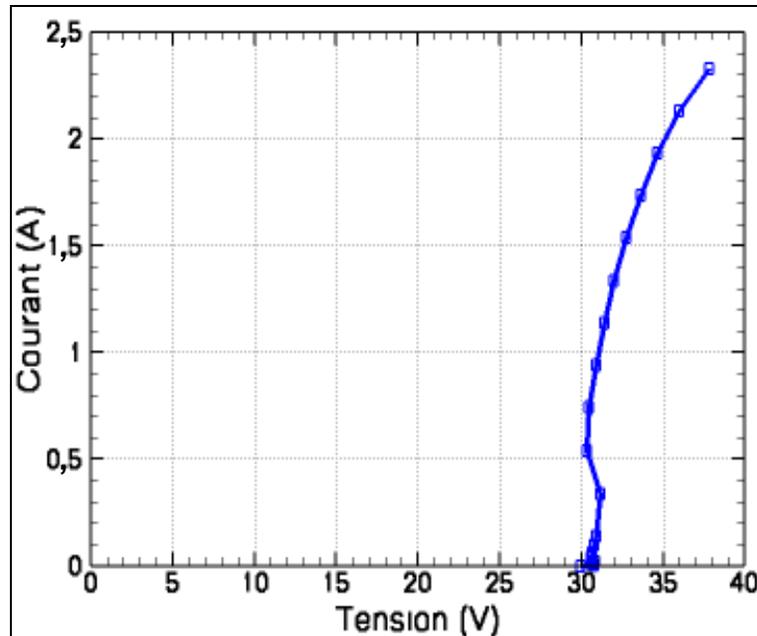


Figure 149: Caractéristique TLP simulée d'un transistor avec région flottante optimisée (surface $10^4 \mu\text{m}^2$).

4.3.3.3. Résultats de mesures

Les délais ont été trop courts pour obtenir des protections optimisées sur silicium. Néanmoins, une série de mesures préliminaires a été effectuée sur des transistors NPN latéraux avec région flottante

dans le collecteur, réalisés dans une ancienne génération de technologie smart power. L'objectif est de pouvoir valider ou infirmer le comportement observé en simulation, en particulier la suppression du phénomène de repliement. Par contre, en l'absence d'optimisation, les résultats ne permettront pas de préciser les performances atteignables.

Structures testées

La coupe technologique est identique à celle utilisée en simulation (Figure 139). Seules les caractéristiques des dopages sont différentes. Notamment, le dopage N dans lequel est insérée la région flottante est de l'ordre de $5 \cdot 10^{15} \text{ cm}^{-3}$, comparé à 10^{16} cm^{-3} pour les simulations. Le layout est composé d'un seul doigt long de $100 \mu\text{m}$ et large d'environ $20 \mu\text{m}$. Comparé aux structures multidoigts des autres composants étudiés au cours de cette thèse, des niveaux de courant nettement inférieurs sont attendus. A courant égal, la densité de courant sera environ cinq fois plus élevée que celle circulant dans une structure multidoigts de $100 \cdot 100 \mu\text{m}^2$. Sept structures ont été mesurées : six structures pour des positions de la région flottante allant de 0 à 6 udm par pas de 1,2 udm, plus une structure de référence sans région flottante.

Le R_{ON} et la robustesse n'ayant pas été optimisés, ces structures n'ont qu'un seul doigt et la longueur d'émetteur est surdimensionnée pour un composant latéral.

Caractérisations TLP

Les caractéristiques TLP de ces structures présentent toutes la même forme générale. A titre d'exemple, celle obtenue pour un espacement de 2,4 udm est représentée (Figure 150). On observe un passage à l'état passant sans repliement. En effet, la tension augmente continûment de sa valeur au déclenchement V_t , égale à 80 Volts, jusqu'au point ($V_{sp}=105 \text{ V}$; $I_{sp}=120 \text{ mA}$) auquel elle diminue brusquement. Etant donné la faible surface de la structure, un courant de 120 milliAmpères est beaucoup trop élevé pour que cette diminution de la tension corresponde au repliement classique d'un transistor bipolaire autopolarisé. Par exemple, si cinq doigts avaient été mis en parallèle pour obtenir une surface de $100 \cdot 100 \mu\text{m}^2$, plus proche de celle d'une protection ESD usuelle, le courant serait de 600 milliAmpères. Par ailleurs, un second changement brusque de la résistance dynamique est observé, au point ($V_H=60 \text{ V}$; $I_H=230 \text{ mA}$) à partir duquel elle s'approche de zéro tout en restant légèrement négative.

Ces observations confirment les principaux résultats de simulation. Le repliement au déclenchement est bien supprimé. La forte surtension à bas courant, de 80 Volts à 105 Volts, peut s'expliquer par le fait que la jonction flottante dans le collecteur se comporte comme une simple diode à avalanche, ce qui est une conséquence de la séparation spatiale entre le flux du courant d'injection et la région d'avalanche. Les changements dans les variations de la caractéristique TLP peuvent être interprétés comme résultant de la suppression du champ tenu entre la base et la région flottante, au point (V_{sp} ; I_{sp}), puis du passage en configuration autopolarisée, au point (V_H ; I_H).

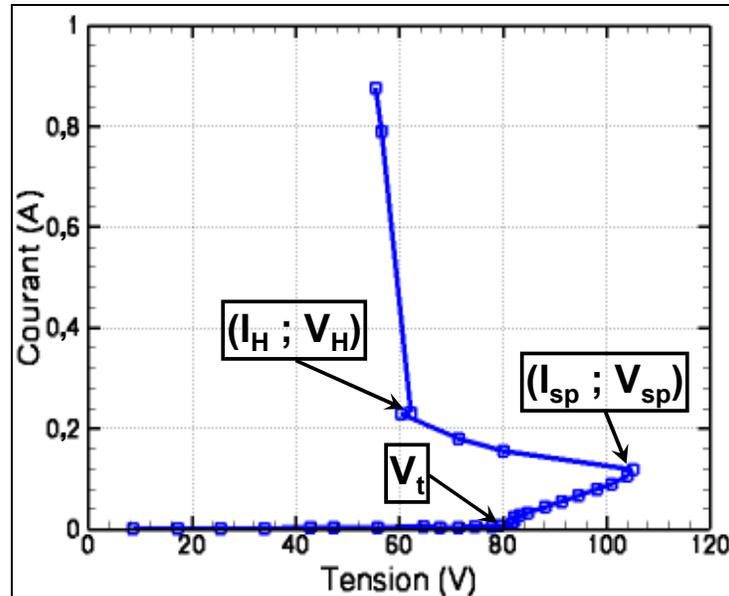


Figure 150: Caractéristique TLP pour un espacement entre la base et la région flottante de 2,4 udm.

La tension de déclenchement V_t est relevée pour les sept structures (Figure 151). Comme attendu, elle augmente significativement par rapport à la structure de référence, et peut être ajustée avec la position de la région flottante. Les couples $(I_{sp}; V_{sp})$ et $(I_H; V_H)$ correspondant aux changements de résistance dynamique sont également relevés (Tableau 27). A cause de la limitation en tension du banc TLP (110 Volts) ou de problèmes de mesures, ces valeurs n'ont pas pu être obtenues pour toutes les structures. On observe que les courants I_{sp} et I_H sont indépendants de la structure, ce qui tend à démontrer que les changements de R_{ON} sont contrôlés par le courant. Cet argument renforce l'hypothèse selon laquelle ces changements sont induits par une modification de la distribution de champ en forte injection. Par ailleurs, les tensions V_H sont proches les unes des autres, ce qui tend à confirmer qu'à fort courant, le bipolaire est dans une configuration indépendante de la région flottante.

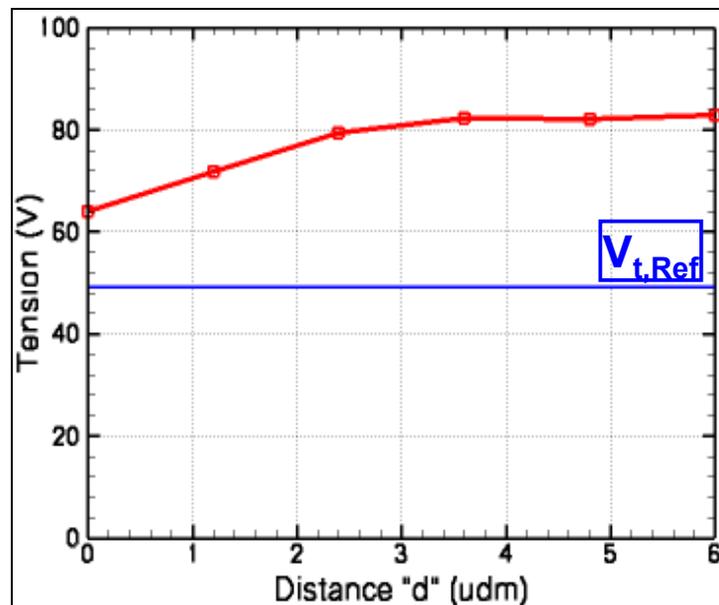


Figure 151: Tension de déclenchement en fonction de la distance d ($V_{t,Ref}$ tension de déclenchement sans région flottante).

Tableau 27: Tensions et courants aux modifications de la résistance dynamique.

d (udm)	0,0	1,2	2,4
V_{sp} (V)	97	99	105
I_{sp} (mA)	120	113	119
V_H (V)	64	60	60
I_H (mA)	200	207	229

Conclusion et perspectives

Les caractéristiques TLP obtenues en simulation ont pu être confrontées à des résultats de mesure. La suppression du repliement est maintenant acquise. Concernant les changements dans les variations du R_{ON} et le comportement aux fortes densités de courant, un ensemble d'arguments concordants tend à confirmer les hypothèses avancées, telles que la séparation spatiale de l'injection et l'avalanche à bas courant puis, lorsque les effets de fortes injection se manifestent, la suppression du champ entre la base et la région flottante, et le passage en mode autopolarisés. Dans le futur, les transistors NPN avec région flottante dans le collecteur devront être optimisés selon les critères au paragraphe 4.3.3.2, ce qui devrait permettre d'obtenir des protections efficaces des E/S "hautes tensions".

4.4. Conclusion

Un transistor NPN de large base interne a été développé de manière à assurer une tension de maintien élevée, d'après la règle définie au chapitre 2. Afin d'élargir la base tout en limitant la surface de silicium, une idée originale a été retenue, elle consiste à insérer une tranchée profonde entre les régions d'émetteur-base et de collecteur. De plus, cette configuration a permis de tirer parti de la jonction entre la couche enterrée dopée N et l'épitaxie P pour obtenir un profil de dopage base-collecteur particulièrement favorable au R_{ON} . La mesure TLP a montré des propriétés très intéressantes telles qu'une tension de maintien minimale de 40 Volts, un R_{ON} à fort courant quasi nul et un facteur de mérite de la robustesse élevé, supérieur à $20 \text{ mW} \cdot \mu\text{m}^{-1}$. Cependant, la forte surtension au déclenchement, de près de 60 Volts et le très long temps de charge, d'environ 50 nanosecondes, sont rédhibitoires pour une protection ESD. Pour résoudre ces deux problèmes, un transistor PNP a été inséré entre le collecteur et la base du transistor NPN. Une première réalisation a permis de ramener la surtension au déclenchement à 20 Volts et de supprimer les problèmes de temps de charge, tout en conservant 40 Volts de tension minimale et un R_{ON} à fort courant quasi-nul. La contrepartie est une dégradation de la robustesse, l'ajout du transistor PNP fragilisant le transistor NPN.

Dans une seconde partie, une étude a été menée sur des structures très innovantes, visant à mettre à profit l'effet d'une région flottante dans la base ou le collecteur d'un transistor NPN autopolarisé. En statique, l'insertion d'une région flottante permet de répartir le champ électrique sur deux jonctions, et ainsi d'augmenter la tension de claquage. Une description 2D de la répartition du champ électrique a montré un risque d'amplification du repliement et de dégradation du R_{ON} . Néanmoins, l'étude a été poursuivie et le comportement au cours d'une ESD analysé. Dans le cas d'une région flottante dans la base, les simulations ont montré que les effets en fort courant s'opposent aux effets statiques au point de supprimer le repliement et d'améliorer le R_{ON} . Cependant, la description des phénomènes physiques est extrêmement complexe, et des règles pour la réduction du R_{ON} n'ont pas pu être définies. Dans ces conditions, la mise en pratique des transistors avec région flottante dans la base s'avère délicate et incertaine. Dans le cas d'une région flottante dans le collecteur, le courant d'injection ne transite pas dans la région de forte avalanche, ce qui permet de supprimer facilement le repliement. La contrepartie est une dégradation des propriétés en termes de R_{ON} . En revanche, les effets en fort courant induisent un R_{ON} équivalent à celui d'un simple transistor NPN autopolarisé. Des règles d'optimisation ont été définies pour profiter au plus tôt des effets en fort courant tout en garantissant la suppression du repliement. Ces règles sont indépendantes de celles définies au chapitre

2, ces dernières n'étant pas adaptées aux effets particuliers induits par la région flottante. Le temps a manqué pour réaliser des structures silicium optimisées, toutefois, des mesures préliminaires de structures non optimisées ont permis de confirmer les principaux résultats théoriques. Au final, l'insertion d'une région flottante dans le collecteur semble une solution très prometteuse pour la protection des E/S "hautes tensions".

Conclusion générale

Afin d'améliorer la fiabilité des circuits intégrés en technologie SmartPower, des spécifications ESD de plus en plus contraignantes sont requises. Dans le cas des E/S "hautes tensions", les protections usuelles ne remplissent ces spécifications que suite à une augmentation importante de leur surface, et donc du coût total du circuit intégré. Cette évolution est d'autant plus pénalisante que, dans le même temps, les nouvelles technologies permettent de réduire la taille des circuits intégrés. Etant donné le contexte très concurrentiel de l'industrie microélectronique, la demande est très forte pour des solutions alternatives plus petites.

Une protection ESD doit offrir un chemin conducteur faiblement résistif pour évacuer le courant lorsqu'une ESD survient, et rester bloquée le reste du temps. Afin de protéger correctement le circuit et de ne pas dégrader sa fiabilité en fonctionnement, il est nécessaire que la tension à l'état passant soit au moins supérieure à la tension d'alimentation tout en restant inférieure à la tension de dégradation, en général la tension de claquage des oxydes. Ces limites doivent être respectées jusqu'à ce que soit atteint le courant maximal de la robustesse spécifiée. En reportant ces contraintes au cas des E/S "hautes tensions", trois défis majeurs ont été identifiés. Ils consistent à limiter, voire supprimer, le repliement de la caractéristique électrique, à garantir un niveau de robustesse élevé et à atteindre une faible résistance dynamique à l'état passant, ces deux derniers points étant particulièrement critiques du fait de la tension de fonctionnement élevée.

Afin de comparer la robustesse et le R_{ON} des protections indépendamment de la surface et de la tension, des facteurs de mérite F_{I2} et F_{RON} ont été définis, correspondant respectivement à la puissance maximale dissipée divisée par la surface, et au R_{ON} multiplié par la surface et divisé par la tension au carré. Au regard des défis pour protéger les E/S "hautes tensions", F_{I2} devra être élevé et F_{RON} faible. L'un de nos principaux objectifs industriels porte sur la protection des E/S 80 Volts d'une application ethernet, pour laquelle la limite inférieure de F_{I2} est $5,8 \text{ mW} \cdot \mu\text{m}^{-2}$ et la limite supérieure de F_{RON} $42 \mu\text{m}^2 \cdot \text{W}^{-1}$. Dans ce cas, la solution conventionnelle basée sur un LDMOS en série avec des transistors bipolaires verticaux ne permet pas d'atteindre la spécification relative au R_{ON} . Cependant, pour développer une solution alternative, une contrainte supplémentaire concerne l'obligation de se démarquer des brevets posés par des compagnies concurrentes de Freescale. Au préalable, il est donc impératif de bien connaître l'état de l'art des protections "hautes tensions" en technologie SmartPower. Deux stratégies de protection peuvent être mises en œuvre suivant que la structure de protection est commune à plusieurs E/S, stratégie dite centralisée, ou dédiée à une seule E/S, stratégie dite localisée. Dans l'optique de réduire la surface, une stratégie localisée est plus intéressante. En effet, si l'on arrive à faire tenir la structure de protection sous le plot d'E/S, alors la surface totale de la puce n'est pas affectée. Trois types de composants majoritairement utilisés comme structures de protection ont été identifiés : les transistors bipolaires autopolarisés, les structures PNP et les transistors à effet de champ. Nous avons choisi de focaliser notre étude sur les transistors bipolaires autopolarisés, offrant des tensions de maintien plus élevées que celles des structures PNP, et moins résistifs que les transistors à effet de champ.

Compte tenu de la courte durée d'une ESD, de quelques nanosecondes, des densités de courant extrêmement élevées peuvent transiter dans une structure de protection sans qu'elle soit dégradée suite à l'élévation de température par effet Joule. Si l'on considère en plus les limitations en termes de surface disponible, des valeurs supérieures à $10^6 \text{ A} \cdot \text{cm}^{-2}$ peuvent être brièvement atteintes. De tels niveaux de densité de courant ne se retrouvent dans aucun autre type d'application. Concernant la

température, la destruction est généralement initialisée par second claquage thermique, survenant entre 700 et 1250 degrés Kelvins. Dans ces conditions, des méthodes de caractérisation spécifiques doivent être employées, telles que les mesures TLP qui permettent d'obtenir des caractéristiques électriques représentatives d'une réponse à une ESD. Afin de définir des stratégies de développement pour une protection ESD, il s'avère indispensable de prendre en compte les effets de ces spécificités physiques sur le comportement électrique. Pour cela, nous nous sommes basés sur une étude théorique développée à partir de modèles unidimensionnels de transistors, que sont venues ensuite compléter des simulations électrothermiques bidimensionnelles, réalisées avec des outils TCAD. L'étude théorique a pour objectif de donner des règles d'optimisation générales sur les dimensions et les niveaux de dopage, alors que les simulations électrothermiques seront surtout mises à profit pour optimiser des structures particulières.

La méthodologie de l'étude théorique a consisté à décrire tout d'abord le mode autopolarisé dans le cas simplifié où les niveaux de courant sont modérés, puis à déterminer successivement les modifications induites par les fortes densités de courant et par l'élévation de température. Pour commencer, une formulation concise de la polarisation à courant modéré a été établie. Son analyse a permis de montrer, d'une part, que le repliement est supprimé en configuration base flottante, le transistor se déclenchant au niveau de sa tension de maintien, et d'autre part, que la tension de maintien des transistors PNP est intrinsèquement plus élevée que celle des transistors NPN. Par ailleurs, le R_{ON} théorique est nul, sa valeur est donc exclusivement liée aux résistances d'accès et aux spécificités du fonctionnement pendant une ESD. Concernant les fortes densités de courant, nous avons montré qu'elles influencent la caractéristique électrique via la chute du rapport d'injection et la modification du champ électrique à la jonction base-collecteur. A ce stade, une étude fondamentale des transistors bipolaires aux densités de courant générées par les ESD s'est avérée nécessaire, la littérature ne traitant que partiellement des modes de fonctionnement dans de telles conditions. Une formule analytique du rapport d'injection d'un transistor bipolaire autopolarisé a été proposée, permettant de montrer pour la première fois une limite à fort courant, égale au rapport des mobilités des électrons et des trous pour un transistor NPN, et au rapport des vitesses de saturation des trous et des électrons pour un transistor PNP. De plus, la charge induite par le courant dans la ZCE base-collecteur a été calculée. Dans le cas d'un transistor NPN, elle augmente continûment, ce qui est bénéfique pour le R_{ON} . Dans le cas d'un transistor PNP, elle décroît à partir d'une densité de courant donnée. Cette décroissance doit impérativement être évitée pour conserver un R_{ON} acceptable. A partir de ces résultats, nous avons défini deux ensembles de règles de dessin pour obtenir simultanément une tension élevée et un R_{ON} faible, soit avec des transistors NPN soit avec des transistors PNP. Concernant l'élévation de température, elle influence la caractéristique via la modification des paramètres électriques du silicium. D'une part la diminution des coefficients d'ionisation par impact et des mobilités dégrade le R_{ON} , mais, d'autre part, la diminution des vitesses de saturation l'améliore. Le plus souvent, il en résulte une augmentation du R_{ON} , cependant, il est possible de limiter cette tendance en jouant sur le profil de dopage à la jonction base-collecteur. Au terme de cette étude théorique, nous avons mis en évidence qu'une tension de maintien élevée est plus aisée à obtenir avec un transistor PNP qu'avec un transistor NPN, mais que, pour obtenir un faible R_{ON} , les propriétés d'un transistor NPN sont meilleures que celle d'un transistor PNP. Aussi, les stratégies de développement de protections "hautes tensions" sont fondamentalement différentes et ont été traitées séparément.

Concernant les transistors PNP, nous avons commencé par définir des règles de réduction du R_{ON} portant sur le layout des structures latérales. D'une part, la longueur d'émetteur doit être minimale et son périmètre en regard du collecteur maximal, ce qui conduit à privilégier des structures interdigitées avec le plus grand nombre de doigts possible. D'autre part, la longueur de collecteur doit être suffisamment grande pour ne pas risquer de dégrader la robustesse et le R_{ON} à très fort courant, en particulier il ne faut pas la prendre égale au minimum autorisé par la technologie. Des structures ont été réalisées en appliquant au mieux les diverses règles de dessin. Les mesures TLP ont montré des performances intéressantes pour la protection des E/S "hautes tensions", notamment la suppression du repliement grâce à une configuration base flottante, et des facteurs de mérite de la robustesse et du R_{ON} (Tableau 28) compatibles avec les spécifications des E/S 80 Volts de l'application ethernet à protéger. Pour améliorer encore ces performances, nous avons développé une structure innovante où

une diode verticale contribue au courant de polarisation du transistor PNP latéral. Ces deux composants étant sans repliement, le fonctionnement à l'état passant est stable. Dans ces conditions, l'accroissement du courant de polarisation, en favorisant l'injection du transistor, permet d'augmenter le courant total et donc de réduire le R_{ON} . Des règles d'optimisation du R_{ON} ont été définies, elles portent sur l'ajustement de la tension de claquage latéral par rapport à la tension de claquage vertical, et sur les dimensions de la diode relativement à celle du transistor PNP. Les réalisations sur silicium ont montré des performances exceptionnelles en termes de R_{ON} , tout en assurant une très bonne robustesse (Tableau 28). Ces structures ont pu être utilisées avec succès pour développer des protections dédiées aux E/S 80 Volts.

Concernant les transistors NPN, deux pistes ont été explorées. Dans un premier temps, nous avons développé une structure originale mettant à profit les caractéristiques technologiques pour une application optimale des règles de dessin. Une des règles fondamentales consiste à élargir la base. Pour cela, nous avons inséré une tranchée profonde entre l'électrode de collecteur et les électrodes d'émetteur et de base. Ainsi, le courant injecté de l'émetteur dans la base doit contourner cette tranchée pour arriver au collecteur, ce qui, en le forçant à circuler en profondeur, permet d'obtenir à la fois une grande base interne et une surface réduite. De plus, si la jonction base-collecteur est réalisée avec l'épitaxie P et la couche enterrée N, le profil de dopage sera particulièrement favorable au R_{ON} et une bonne robustesse devrait être atteinte grâce à un point chaud situé dans le volume du silicium. Les mesures ont montré une tension de maintien élevée, égale à 40 Volts, un R_{ON} quasi-nul et un excellent courant I_{d2} de plus de 5 Ampères. Cependant, la tension de déclenchement est trop élevée, égale à 110 Volts, et le temps de charge trop long pour que cette structure soit directement utilisable en tant que protection ESD. Aussi, nous avons ajouté un circuit externe d'aide au déclenchement, dont l'élément de base est un transistor PNP. Dans ce cas, les mesures ont montré une tension de déclenchement et un temps de charge acceptables, sans que la tension de maintien et le R_{ON} soient modifiés. Par contre, la robustesse a été dégradée, tout en restant néanmoins à un niveau élevé. Dans un second temps, nous avons étudié des structures latérales avec une région flottante dans la base ou dans le collecteur. Ces régions flottantes permettent d'augmenter la tension de claquage, typiquement jusque dans un rapport deux. Tout d'abord, nous avons décrit en détail le fonctionnement statique, en mettant en évidence l'importance des effets 2D dans les mécanismes de polarisation de la région flottante. Ensuite, le fonctionnement au cours d'une ESD a été étudié en simulation. Des propriétés intéressantes pour obtenir une tension de maintien élevée et un faible R_{ON} ont été identifiées, aussi bien pour le cas où la région est dans la base que pour celui où elle est dans le collecteur. Cependant, le développement d'une structure de protection semble nettement plus aisé pour le cas où l'anneau flottant est dans le collecteur. A bas courant, les porteurs injectés ne transitent pas dans la zone d'avalanche, ce qui assure la suppression du repliement. Lorsque le courant augmente, les effets de forte injection conduisent à une configuration équivalente à celle d'un simple transistor NPN autopolarisé, dont le R_{ON} est intrinsèquement faible. Ces phénomènes ont été validés par une première série de mesure sur des composants non optimisés, mais le temps a manqué pour les mettre à profit dans des structures de protection ESD.

Le Tableau 28 résume les performances atteintes par les différents composants étudiés au cours de la thèse, et les compare avec celles d'une protection conventionnelle des E/S 80 Volts. Cette structure de l'état de l'art est constituée d'un transistor LDMOS en série avec des transistors NPN verticaux. Les transistors PNP présentent de très bons facteurs de mérites du R_{ON} et de la robustesse, qui sont améliorés dans des proportions significatives en le couplant avec une diode verticale. Le transistor avec tranchée dans la base déclenché par un circuit externe à un excellent facteur de mérite du R_{ON} , par contre le facteur de mérite de la robustesse est moyen, suite à la dégradation causée par le circuit de déclenchement. Enfin, le facteur de mérite simulé pour un transistor NPN avec région flottante dans le collecteur donne une valeur intéressante, en tenant compte du fait que les simulations ont généralement été pessimistes par rapport aux mesures.

Tableau 28: Performances atteintes par les composants étudiés au cours de la thèse.

	$F_{RON} (\mu\text{m}^2 \cdot \text{W}^{-1})$	$F_{I2} (\text{mW} \cdot \mu\text{m}^{-2})$
Transistor LDMOS en série avec des transistors NPN verticaux (Etat de l'art)	52	11,5
Transistor PNP latéral	21	20
Transistor PNP couplé à une diode verticale	<4	29
Transistor NPN avec tranchée dans la base, déclenché par un circuit externe	≈ 0	8
Transistor NPN latéral avec région flottante dans le collecteur (Simulation)	25-35	-

Ces travaux ouvrent différentes perspectives. Les résultats théoriques sur les transistors bipolaires autopolarisés peuvent donner lieu à des études expérimentales intéressantes. Elles consisteraient à mettre en évidence des comportements singuliers, induits par les effets fondamentaux que nous avons reportés pour la première fois. Des mesures TIM sur le transistor NPN avec tranchée dans la base devraient permettre de déterminer si le R_{ON} quasi-nul provient d'effets liés à la focalisation du courant. Si tel est le cas, on peut imaginer définir des règles de dessin pour mettre à profit ce phénomène de manière systématique. Concernant les transistors NPN avec région flottante dans le collecteur, il reste à déterminer leurs performances par la mesure des structures de protection optimisées. Enfin, l'efficacité vis-à-vis des normes "pistolet" (IEC 61000-4-2 [9] et ISO 10605 [10]) des structures couplant un transistor PNP et une diode pourrait être évaluée. Les très bons facteurs de mérite offre la possibilité de réduire fortement la surface des protections standard. Cependant, il faut pouvoir garantir que la robustesse est suffisante pour supporter les très fortes énergies mises en jeu et que le temps de déclenchement est compatible avec le très court temps de montée en début de décharge.

Bibliographie

- [1] Automotive Electronic Council, Human Body Model Electrostatic Discharge Test, AEC-Q100-002-REV-C, 12P, 1998.
- [2] JESD22-A114D, “Electrostatic Discharge (ESD), Sensitivity Testing Human Body Model (HBM)”, March 2006.
- [3] ESD Association, Association Standard Test Method for Electrostatic Discharge Sensitivity Testing - Human Body Model (HBM) Component level - ESD-STM5.1-1998, 11p, 1998.
- [4] Automotive Electronic Council, Machine Model Electrostatic Discharge Test, AEC-Q100-003-REV-C, 12P, 1998.
- [5] JESD22-A115A, “Electrostatic Discharge (ESD), Sensitivity Testing Machine Model (MM)”, October 1997.
- [6] ESD Association, Association Standard Test Method for Electrostatic Discharge Sensitivity Testing - Machine Model (MM) Component level - ESD-STM5.3-1999, 12p, 1999.
- [7] H. Ishizuka, K. Okuyama, K. Kuboya, M. Komuro, Y. Hara, “A Study of ESD Protection Devices for Inputs Pins Discharge Characteristics of Diode, Lateral Bipolar Transistor, and Thyristor under MM and HBM Tests”, IEEE Trans. on Components, Packaging, and Manufacturing Technology – Part C, vol.21, n° 4, pp. 257-264, October 1998.
- [8] J. Whitfield, C. Gill, J. Yang, H. Xu, C. Zhan, B. Baumert, M. Zunino, “ESD MM Failures Resulting from Transient Reverse Currents”, in Proc. IRPS, pp. 136-139, 2006.
- [9] IEC 61000-4-2 Electromagnetic Compatibility (EMC)- Part 4-2: Testing and Measurement Techniques - Electrostatic Discharge Immunity Test. 2001
- [10] ISO 10605 Road Vehicles- Test Methods for Electrical Disturbances from Electrostatic Discharges. 2001
- [11] ESD Association, Association Standard Test Method for Electrostatic Discharge Sensitivity Testing - Charged Device Model Component Level - ESD-STM 5.3.1, 1999.
- [12] JESD22-C101B.01, “Field-Induced Charged Device Model Test Method for Electrostatic-Discharge-Withstand Thresholds of Microelectronic Component”, 2004.
- [13] H. Gieser, “Influence of Tester Parasitic in Charged Device Model Failure Thresholds”, in Proc. EOS/ESD Symposium, pp. 69-84, 1994.
- [14] C. Goëau, C. Richier, P. Salomé, J-P. Chante, H. Jaouen, “Impact of CDM Tester Ground Plane Capacitance on the DUT Stress Level”, in Proc. EOS/ESD Symposium, pp. 170-177, September 2005.
- [15] M.S.B. Sowariraj, “Role of Package Parasitics and Substrate Resistance on the Charged Device Model Failure Level”, Microelectronics Reliability, pp. 1569-1575, 2003.
- [16] ISO76371-1 Road Vehicles Electrical Disturbances by Conduction and Coupling. Part 1: Passenger cars and light commercial vehicles with nominal 12V supply voltage – Electrical transient conduction along supply lines only. 1990.
- [17] A. Amerasekera, M-C. Chang, A. Seitchik, A. Chatterjee, K. Mayaram, J-H. Chern “Self-Heating Effects in Basic Semiconductor Structures”, IEEE transactions on Electron Devices, vol. 40, n° 10, October 1993.

- [18] A. Amerasekera, J. Seitchik, "Electrothermal Behavior of Deep Submicron nMOS Transistors under High Current Snapback (ESD/EOS) Conditions", Proc. IEDM, pp. 446-449, 1994.
- [19] K. Esmark, "Device Simulation of ESD Protection Elements", Series in Microelectronics, vol. 128, ISBN 3-89649-781-2, 2002.
- [20] M. Scholz, D. Tremouilles, M.I. Natarajan, M. Sawada, T. Nakaei, T. Hasebe, G. Groeseneken, "Can HBM Tester Replace 100ns TLP Tester?", in Proc. 3rd EOS/ESD/EMI Workshop, pp. 33-35, 2006.
- [21] T. Maloney, N. Khurana, "Transmission line pulsing techniques for Circuits Modeling of ESD Phenomena", in Proc. EOS/ESD Symposium, pp. 49-55, 1985.
- [22] H. Gieser, M. Haunschild, "Very Fast Transmission Line Pulsing of Integrated Structures and the Charged Device Model", IEEE Trans. on Components, Packaging, and Manufacturing Technology – Part C, vol. 21, n° 4, pp. 278-285, October 1998.
- [23] D. Pogany, V. Dubec, S. Bychikhin, C. Furböck, M. Litzenberger, S. Naumov, G. Gross, M. Stecher, E. Gornik, "Single-Shot Nanosecond Thermal Imaging of Semiconductor Devices Using Absorption Measurements", IEEE Trans. on Device and Materials Reliability, vol. 3, n°3, pp. 85-88, September 2003.
- [24] M. Litzenberger, R. Pichler, D. Pogany, E. Gornik, K. Esmark, H. Gossner, "Influence of Layout Parameters on Triggering Behavior in 0.35 μm and 0.18 μm Process gg-nMOS ESD Protection Devices", in Proc. ESSDERC, pp. 335-338, 2001.
- [25] M. Heer, S. Bychikhin, V. Dubec, D. Pogany, E. Gornik, M. Dissegna, L. Cerati, L. Zullino, A. Andreini, A. Tazzoli, G. Meneghesso, "Analysis of the Triggering Behavior of Low Voltage BCD Single and Multi-Finger gc-NMOS ESD Protection Devices", in Proc. EOS/ESD Symposium, pp. 275-284, Sept. 2006.
- [26] M. Denison, M. Blaho, P. Rodin, V. Dubec, D. Pogany, D. Silber, E. Gornik, M. Stecher, "Moving Current Filaments in Integrated DMOS Transistors Under Short-Duration Current Stress", IEEE Trans. on Electron Devices, vol. 51, n° 10, pp. 1695-1703, Oct. 2004.
- [27] D. Pogany, S. Bychikhin, C. Furböck, M. Litzenberger, E. Gornik, G. Gross, K. Esmark, M. Stecher, "Quantitative Internal Thermal Energy Mapping of Semiconductor Devices under Short Current Stress Using Backside Laser Interferometry", IEEE Trans. on Electron Devices, vol. 49, n°11, pp. 2070-2079, November 2002.
- [28] M.P.J Mergens, M.T. Mayerhofer, J.A. Willemen, M. Stecher, "ESD Protection Considerations in Advanced High-Voltage Technologies for Automotive", in Proc. EOS/ESD Symposium, pp. 54-63, September 2006.
- [29] V. De Heyn, G. Groeseneken, B. Keppens, M. Natarajan, L. Vacaresse, G. Gallopyn, "Design and Analysis of New Protection Structures for Smart Power Technology with Controlled Trigger and Holding Voltage", in Proc. IRPS, pp. 253-258, 2001.
- [30] C. Delage, "Etude et Conception de Structures de Protection contre les Décharges Electrostatiques en Technologie BiCMOS de Puissance ", Thèse de l'INPT soutenue le 13 décembre 1999, rapport LAAS n° 99555.
- [31] P. Besse, "Tenue en énergie de structures LDMOS avancées de puissance intégrée dans les domaines temporels de la nanoseconde à la milliseconde", Thèse de l'UPS de Toulouse soutenue le 28 janvier 2004.
- [32] H. Gossner, T. Muller-Lynch, K. Esmark, M. Stecher, "Wide Range Control of the Sustaining Voltage of ESD Protection Elements Realized in a Smart Power Technology", in Proc. EOS/ESD Symposium, pp. 19-27, September 1999.

- [33] H. Gossner, M. Stecher, "Integrierte Halbleiterschaltung mit Schutzstruktur zum Schutz vor Elektrostatischer Entladung", European Patent Application EP0905782A1, 31 März 1999.
- [34] H. Gossner, "Integrated Semiconductor Circuit with Protective Structure for Protection against Electrostatic Discharge", U.S. Patent 6441437 B1, August 27, 2002.
- [35] M-D. Ker, K-H. Lin, "The Impact of Low Holding-Voltage Issue in High-Voltage CMOS Technology and the Design of Latchup-Free Power-Rail ESD Clamp Circuit for LCD Drivers ICs", *Solid-State Circuits*, vol. 40, n°8, August 2005.
- [36] B.K. Ridley, "Specific Negative Resistance in Solids", in *proc. Phys. Soc.*, vol. 82, pp. 954, 1963.
- [37] D. Trémouilles, G. Bertrand, M. Bafleur, N. Nolhier, L. Lescouzeres, "Design Guidelines to Achieve a Very High ESD Robustness in a Self-Biased NPN", in *Proc. EOS/ESD Symposium*, pp. 281-288, September 2002.
- [38] S.M. Sze, *Physics of Semiconductor Devices*, 2nd edition, Wiley-Interscience publication, 1981.
- [39] N. Jensen, G. Groos, M. Denison, J Kuzmik, D. Pogany, E. Gornik, M. Stecher, "Coupled Bipolar Transistors as Very Robust ESD Protection Devices for Automotive Applications", in *Proc. EOS/ESD Symposium*, pp. 313-318, Sept. 2003.
- [40] K. Reynders, P. Moens, "Design and Characterization of a High Voltage SCR with High Trigger Current", in *Proc. EOS/ESD Symposium*, pp. 407-412, Sept. 2005.
- [41] N. Jensen, M. Denison, "Integrated circuit", U.S. Patent 0071236 A1, April 06, 2006.
- [42] K. Reynders, P. Moens, "Hybrid ESD Clamp", WO Patent 2006/072148 A1, July 13, 2006.
- [43] B. Keppens, M.P.J Mergens, C.S. Trinh, C.C. Russ, B. Van Camp, K. G. Verhaege, "ESD Protection Solutions for High Voltage Technologies", in *Proc. EOS/ESD Symposium*, pp. 289-298, Sept. 2004.
- [44] M.P.J Mergens, C. Russ, K. Verhaege, J. Armer, P. Jozwiak, R. Mohn, "High Holding Current SCRs (HHI-SCR) for ESD Protection and Latch-up Immune IC Operation", *Microelectronics Reliability*, vol 43, pp. 993-1000, 2003.
- [45] B. Van Camp, B. Keppens, "Apparatus for ESD Protection", WO Patent 2006/033993 A1, March 30, 2006.
- [46] O. Quittard, Z. Mrcarica, F. Blanc, G. Notermans, T. Smedes, H. Van Zwol, "ESD Protections for High-Voltage CMOS Technology", in *Proc. EOS/ESD Symposium*, pp. 77-86, September 2006.
- [47] K-O. Oh, C. Duvvury, C. Salling, K. Banerjee, R.W. Dutton, "Non-uniform Conduction in Single Finger NMOS Transistors and Implications for Deep Submicron ESD Design", In *Proc. IRPS*, pp. 226-234, 2001.
- [48] C. Duvvury, F. Carvajal, C. Jones, D. Briggs, "Lateral DMOS Design for ESD Robustness", In *Proc. IEDM*, pp. 375-378, 1997.
- [49] S.L. Miller, "Avalanche Breakdown in Germanium", *Physical review*, vol. 99, n° 4, pp. 1234-1241, August 15, 1955.
- [50] G. Bertrand, C. Delage, M. Bafleur, N. Nolhier, J.M. Dorkel, Q Nguyen, N. Mauran, D. Trémouilles, P. Perdu, "Analysis and Compact Modeling of a Vertical Grounded-Base n-p-n Bipolar Transistor Used as ESD Protection in a Smart Power Technology", *Solid State Circuits*, vol. 36, n° 9, pp. 1373-1381, September 2001.
- [51] M-D. Ker, K-H. Lin, "Double Snapback Characteristics in High-Voltage nMOSFETs and the Impact to On-Chip ESD Protection Design", *IEEE Electron Devices Letters*, vol. 25, n° 9, September 2004.

- [52] M. Streibl, K. Esmark, A. Sieck, W. Stadler, W. Wendel, J. Svatkowski, H. Gossner, "Harnessing the Base-Pushout Effect for ESD Protection in Bipolar and BiCMOS Technologies", in Proc. EOS/ESD Symposium, pp. 73-82, September 2002.
- [53] M.J. Kumar, K.N. Bhat, "The Effects of Emitter Region Recombination and Bandgap Narrowing on the Current Gain and the Collector Lifetime of High-Voltage Bipolar Transistors", IEEE Trans. On Electron Devices, vol. 36, n° 9, pp. 1803-1810, September 1989.
- [54] H.S. Bennett, C.L. Wilson, "Statistical Comparison of Data on Band-Gap Narrowing in Heavily Doped Silicon : Electrical and Optical Measurements", Journal of Applied Physics, vol. 55, n° 10, pp. 3582-3587, 1984.
- [55] G. Masetti, M. Severi, S. Solmi, "Modeling of Carrier Mobility against Carrier Concentration in Arsenic-Phosphorus-and Boron-Doped Silicon", IEEE Trans. On Electron Devices, vol. ED-30, pp. 764-769, 1983.
- [56] M. Gharbi, "La Tenue en Tension et le Calibre en courant du Transistor M.O.S. dans la Gamme des Tensions Moyennes (300 à 1000 Volts)", Thèse de l'université Paul Sabatier, soutenue le 16 octobre 1985, n° d'ordre 3223.
- [57] G. Charitat, "Modélisation et Réalisation de Composants Planar Haute-Tension", Thèse d'état de l'université Paul Sabatier, soutenue le 28 septembre 1990, rapport LAAS n° 90306.
- [58] G. Boselli, J Rodriguez, C. Duvvury, V. Reddy, P.R. Chidambaram, B. Hornung, "Technology Scaling Effects on the ESD Design Parameters in Sub-100 nm CMOS Transistors", in Proc. IEDM, pp. 507-510, 2003.
- [59] A. Amerasekera, S. Ramaswamy, M.C. Chang, "Modeling MOS Snapback and Parasitic Bipolar Action for Circuit Level ESD and High Current Simulations", in Proc. Reliability Physics Symposium, pp. 318-326, 1996.
- [60] M.J.P. Mergens, W. Wilkening, S. Mettler, H. Wolf, W. Fichtner, "Modular Approach of a High Current MOS Compact Model for Circuit-Level ESD Simulation Including Transient Gate Coupling Behavior", in Proc. Reliability Physics Symposium, pp. 167-178, 1999.
- [61] R.V. Overstraeten, H.D. Man, "Measurement of the Ionization Rates in Diffused Silicon p-n Junctions", Solid State Electronics, vol. 13, pp. 583-608, 1970.
- [62] W.M. Webster, "On the Variation of Junction-Transistor Current-Amplification Factor with Emitter Current", Proceedings of the IRE, pp. 914-920, June 1954.
- [63] E.S. Rittner, "Extension of the Theory of the Junction Transistor", Physical Review, vol. 94, n° 5, pp. 1161-1171, June 1954.
- [64] J.P. Bailbé, "Contribution à l'Etude Physique des Transistors Bipolaires", Thèse d'état de l'université Paul Sabatier de Toulouse, soutenue le 8 février 1977, n° d'ordre 744.
- [65] D. Trémouilles, "Optimisation et modélisation de protections intégrées contre les décharges électrostatiques, par l'analyse de la physique mise en jeu", Thèse de l'INSA de Toulouse soutenue le 14 mai 2004.
- [66] C. Canali, G. Majni, R. Minder, G. Ottaviani, "Electron and Hole Drift Velocity Measurements in Silicon and their Empirical Relation to Electric Field and Temperature", IEEE Trans. on Electron Devices, vol. 35, n° 5, pp. 1045-1047, 1975.
- [67] C. Salamero, N. Nohier, A. Gendron, M. Bafleur, P. Besse, M. Zecri, , "TCAD Methodology for ESD Robustness Prediction of Smart Power ESD", IEEE Trans. on Device and Materials Reliability, vol. 6, pp. 399-407, September 2006.
- [68] N.D. Arora, J.R. Hauser, D.J. Roulson, "Electron and Hole Mobilities in Silicon as a Function of Concentration and temperature", Solid State Circuits, vol. 36, n° 9, pp. 1373-1381, September 2001.

- [69] Y. Okuto, C.R. Crowell, "Threshold Energy Effects on Avalanche Breakdown Voltage in Semiconductor Junctions", *Solid State Electronics*, vol. 18, pp. 161-168, 1975.
- [70] ISE TCAD Release 9.0 manual.
- [71] C. Salamero, "Méthodologie de prédiction du niveau de robustesse d'une structure de protection ESD à l'aide de la simulation TCAD", Thèse de l'UPS de Toulouse, soutenue le 12 décembre 2005.
- [72] T.J. Maloney, "Designing Power Supply Clamp For Electrostatic Discharge Protection of Integrated Circuits", *Microelectronics Reliability*, vol. 38, pp. 1691-1703, 1998.
- [73] T.J. Maloney, W. Kan, "Stacked PMOS Clamps for High Voltage Power Supply Protection", in *Proc. EOS/ESD Symposium*, pp. 70-77, September 1999.
- [74] M. Tong, R. Gauthier, V. Gross, "Study of gated PNP as an ESD Protection Device for Mixed-Voltage and Hot-Pluggable Circuit Applications", in *Proc. EOS/ESD Symposium*, pp. 280-284, September 1996.
- [75] J. Li, R. Gauthier, K. Chatty, D. Kontos, M. Muhammad, M. Woo, C. Putman, C. Russ, D. Alvarez, J. Schneider, P.T. Tan, "PMOSFET-based ESD Protection in 65 nm Bulk CMOS Technology for Improved External Latchup Robustness", in *Proc. EOS/ESD Symposium*, pp. 407-412, September 2005.
- [76] C. Duvvury, A. Amerasekera, "ESD: A Persuasive Reliability Concern for IC Technologies", in *Proc. of the IEEE*, vol. 81, n° 5, pp. 690-702, May 1993.
- [77] T. Maloney, S. Dabral, "Novel Clamp Circuits for IC Power Supply Protection", in *Proc. EOS/ESD Symposium*, pp. 1-12, September 1995.
- [78] C.A. Torres, J.W. Miller, M. Stockinger, M.D. Akers, M.G. Khazhinsky, J.C. Weldon, "Modular, Portable, and Easily Simulated ESD Protection Networks for Advanced CMOS Technologies", in *Proc. EOS/ESD Symposium*, pp. 82-95, September 2001.
- [79] G. Boselli, C. Duvvury, V. Reddy, "Efficient PNP Characteristics of PMOS Transistors in Sub-0.13 μ m ESD Protection Circuits", in *Proc. EOS/ESD Symposium*, pp. 257-266, September 2002.
- [80] M-D. Ker, W-J. Chang, W-Y. Lo "Low-Voltage-Triggered PNP Devices for ESD Protection Design in Mixed-Voltage I/O Interface with Over-VDD and Under-VSS Signal Levels", in *Proc. IEEE Int. Symp. On Quality Electronic Design*, pp. 433-438, 2004.
- [81] W-J. Chang, M-D. Ker, "Layout Optimization on Low-Voltage-Triggered PNP Devices for ESD Protection in Mixed-Voltage I/O Interfaces", in *Proc. IPFA*, pp. 213-216, 2004.
- [82] T.L. Polgreen, A. Chatterjee, "Improving the ESD Failure Threshold of Siliced n-MOS Output Transistors by Ensuring Uniform Current Flow", *IEEE Trans on Electron Device*, vol. 39, n° 2, February 1992.
- [83] M.P.J Mergens, C.C. Russ, J. Arner, P.C. Jozwiack, G. Kolluri, L.R. Avery, K.G. Verhaege, "Multi-Finger Turn-on Circuits and Design Techniques for Enhanced ESD Performance and Width-Scaling", in *Proc. EOS/ESD Symposium*, pp. 1-11, September 2001.
- [84] N. Rinaldi, V. d'Alessandro, F.M. de Paola, "Electrothermal Phenomena in Bipolar Transistors and ICs: Analysis, Modeling and Simulation", in *Proc. BCTM*, pp. 33-40, October 2006.
- [85] G. Notermans, "On the Use of N-Well Resistors for Uniform Triggering of ESD Protection Elements", in *proc. EOS/ESD Symposium*, pp. 221-29, September 1997.
- [86] M. Mergens, O. Marichal, S. Thijs, C. Russ, "Advanced SCR ESD Protection Circuits for CMS/SOI Nanotechnologies", *CICC* 2005.

- [87] P. Hower, S. Pendharkar, R. Steinhoff, J. Brodsky, J. Devore, W. Grose, , "Using Two-Dimensional Structures to Model Filamentation in Semiconductor Devices", in proc. ISPSD, pp. 385-388, 2001.
- [88] H.C. Bowers, "Space-Charged-Induced Negative Resistance in Avalanche Diodes", IEEE transactions on electron devices, vol. ED-15, n° 6, pp. 343-350, June 1968.
- [89] M. W. Muller, H. Guckel, "Negative Resistance and Filamentary Currents in Avalanche Silicon p+-i-n+ Junctions", IEEE Transactions on Electron Devices, vol. ED-15, n° 8, pp. 560-568, August 1968.
- [90] V. Boisson, "Etude de la géométrie Optimale des Périphéries des Jonctions Planar", Thèse de l'Ecole Centrale de Lyon, soutenue 23 avril 1985.
- [91] P. Rossel, F. Morancho, N. Cézac, H. Tranduc, "Schottky-Diode Semiconductor Device", U.S. Patent 0046224 A1, March 11, 2004.
- [92] A.Peyre-Lavigne, I. Pages, P. Rossel, F. Morancho, N. Cézac, "Lateral Semiconductor Device with Low On-Resistance and Method of Making the Same", U.S. Patent 0222461 A1, November 11, 2004.
- [93] N. Cézac, F. Morancho, P. Rossel, H. Tranduc, A. Peyre-Lavigne, "A New Generation of Power Unipolar Devices: the Concept of the Floating Islands MOS Transistor (FLIMOST)", in Proc. ISPSD, pp. 69-72, May 2000.
- [94] S. Alves, F. Morancho, J.-M. Reynès, J. Margheritta, I. Deram, K. Isoird, B. Beydoun, "Experimental Validation of the 'Floating Islands' Concept: 95 V Breakdown Voltage Vertical FLIDiode", in IEE Proc. Circuits Devices Syst., vol. 153, n° 1, pp. 53-60, February 2006.
- [95] M. Zitouni, F. Morancho, P. Rossel, H. Tranduc, J. Buxo, I. Pagès, "A New Concept for the Lateral DMOS Transistor for Smart Power IC's", in Proc. ISPSD, pp. 73-76, May 1999.
- [96] C.A.T. Salama, "A New Short Channel MOSFET Structure (UMOST)", Solid-State Electronics, vol. 20, pp. 1003-1010, 1977.
- [97] R. Rossel, J. Buxo, I. Pages, F. Morancho, V. Macary, "Dispositif à Semiconducteur Latéral et son Procédé de Formation", Brevet français 2770687, 04 November 1997.
- [98] D. Pogany, S. Bychikhin, M. Denison, P. Rodin, N. Jensen, G. Groos, M. Stecher, E. Gornik, "Thermally-Driven Motion of Current filaments in ESD Protection Devices", Solid-State Electronics, vol 49, pp. 421-429, 2005.
- [99] M. Zecri, P. Besse, N. Nolhier, "Arrangement and Method for ESD Protection", European Patent Application EP1396887A1, 10 March 2004.
- [100] Y.C. Kao, E.D. Wolley, "High-Voltage Planar p-n Junctions", in Proc. of the IEEE, vol. 55, n° 8, pp. 1409-1414, August 1967.
- [101] M.S. Adler, V.A.K. Temple, A.P. Ferro, R.C. Rustay, "Theory and Breakdown Voltage for Planar Devices with a Single Field Limiting Ring", IEEE Trans. on Electron Devices, vol. ED-24, n° 2, February 1977.
- [102] V. Boisson, M. Le Helley, J.P. Chante, "Analytical Expression for the Potential of Guard Rings of Diodes Operating in the Punchthrough Mode", IEEE Trans. on Electron Devices, vol. ED-32, n° 4, April 1985.

Liste des symboles

BV_{CB0}	Tension de claquage de claquage base-collecteur, avec l'émetteur à la masse
BV_{CE0}	Tension de claquage émetteur-collecteur, avec la base flottante
$c_{n,p}$	Concentration d'électrons ou de trous induites par le courant
$D_{n,p}$	Coefficient de diffusion des électrons ou des trous
D_T	Coefficient de diffusion thermique du silicium
E	Champ électrique
F_{RON}	Facteur de mérite du R_{ON}
F_T	Facteur de mérite thermique
F_{I2}	Facteur de mérite de la robustesse
I_B	Courant de base d'un transistor bipolaire
I_C	Courant de collecteur d'un transistor bipolaire
I_E	Courant d'émetteur d'un transistor bipolaire
$I_{e-,h+}$	Courant d'électrons ou de trous
I_H	Courant au maintien
I_n	Intégrale d'ionisation
I_{I1}	Courant au déclenchement
I_{I2}	Courant au second claquage
$I_{Robustesse}$	Courant maximale correspondant à la spécification de robustesse
I_S	Courant de fuite inverse d'un transistor bipolaire
J_B	Densité de courant de base d'un transistor bipolaire
J_C	Densité de courant de collecteur d'un transistor bipolaire
J_{Cr}	Densité de courant critique caractérisant le comportement asymptotique du rapport d'injection
J_E	Densité de courant d'émetteur d'un transistor bipolaire
$J_{e-,h+}$	Densité de courant d'électrons ou de trous
J_H	Densité de courant critique caractérisant le début de la chute du gain
J_I	Densité de courant d'inversion des variations de la charge dans un PNP
J_{Max}	Densité de courant à laquelle la vitesse des porteurs sature, dans une diffusion donnée
k	Constante de Boltzmann
q	Charge électrique élémentaire
Q_B	Densité de charge des dopants ionisés dans la base
Q_E	Densité de charge des dopants ionisés dans l'émetteur
Q_{SB}	Densité de charge induite par le courant dans la base
Q_{SE}	Densité de charge induite par le courant dans l'émetteur
L_C	Longueur de collecteur
L_E	Longueur d'émetteur
l_T	Longueur de diffusion thermique
M	Facteur de multiplication dans une jonction en inverse

$M_{n,p}$	Facteur de multiplication des électrons ou des trous
N_A	Concentration de dopants accepteurs
N_D	Concentration de dopants donneurs
n	Concentration d'électrons
n_i	Concentration intrinsèque de porteurs
p	Concentration de trous
P_{t2}	Puissance à la défaillance
r_{E1}, r_{E2}	Paramètres de l'expression du rapport d'injection tenant de la forte injection dans l'émetteur
R_B	Résistance de base d'un transistor bipolaire
R_{ON}	Résistance à l'état passant d'une protection ESD
R_{SRH}	Taux de recombinaisons SRH
S	Surface d'un composant
S_{Max}	Surface maximale autorisée par la spécification
T	Température
t_{ESD}	durée moyenne d'une ESD
V_{Alim}	Tension d'alimentation d'une application
V_{CDM}	Tension de charge d'un testeur CDM
V_H	Tension de maintien
V_{HBM}	Tension de charge d'un testeur HBM
$v_{n,p}$	Vitesse moyenne des électrons ou des trous
$v_{sat,n,p}$	Vitesse de saturation des électrons ou des trous
V_{Oxydes}	Tension de claquage des oxydes
V_{Max}	Tension supérieure de la fenêtre de conception
V_{min}	Tension inférieure de la fenêtre de conception
V_{MM}	Tension de charge d'un testeur MM
V_{t1} (ou V_t)	Tension de déclenchement
V_{t2}	Tension au second claquage
W_B	Largeur de base d'un transistor bipolaire
W_C	Largeur de collecteur d'un transistor bipolaire
W_E	Largeur d'émetteur d'un transistor bipolaire
W_{Lay}	Extension d'un doigt du layout
W_{ZCE}	Extension de la zone de charge d'espace d'une jonction
α	Gain en base commune d'un transistor bipolaire
$\alpha_{n,p}$	Coefficient d'ionisation par impact des électrons ou des trous
β	Gain en émetteur commun d'un transistor bipolaire
ε	Permittivité diélectrique
γ	Rapport d'injection d'un transistor bipolaire
γ_I	Limite du rapport d'injection d'un transistor bipolaire, pour la forte injection dans l'émetteur négligée
γ_L	Limite du rapport d'injection d'un transistor bipolaire, pour la forte injection dans l'émetteur prise en compte
μ_n	Mobilité des électrons

μ_p	Mobilité des trous
ρ	Densité de charge
ρ_{Cr}	Densité de charge critique vis-à-vis de l'influence du courant de polarisation
τ_B	Temps de transit des porteurs dans la base
τ_E	Temps de transit des porteurs dans l'émetteur
$\tau_{n,p}$	Durée de vie des électrons ou des trous

Contributions scientifiques

Revues

C. Salamero, N. Nolhier, A. Gendron, M. Bafleur, P. Besse, M. Zecri, “*TCAD Methodology for ESD Robustness Prediction of Smart Power ESD Devices*”, in IEEE Trans. on Device and Materials Reliability, vol. 6, issue 3, pp. 399-407, Sep. 2006.

A. Gendron, P. Renaud, M. Bafleur, N. Nolhier, “*Analytical description of the injection ratio of self-biased bipolar transistors under the very high injection conditions of ESD events*”, accepté pour publication dans Solid State Electronics.

Congrès internationaux

A. Gendron, C. Salamero, N. Nolhier, M. Bafleur, P. Renaud, P. Besse, “*Comparison of Efficient Solutions for the ESD Protection of High-Voltages I/Os in Advanced Smart Power Technologies*”, in Proc. 3rd EOS/ESD/EMI Workshop, pp. 19-22, May 2006.

A. Gendron, P. Renaud, P. Besse, C. Salamero, M. Bafleur, N. Nolhier, “*Area-Efficient Reduced and No-Snapback PNP-based ESD Protection in Advanced Smart Power Technology*”, in Proc. EOS/ESD Symposium, pp. 69-76, Sep. 2006.

A. Gendron, C. Salamero, N. Nolhier, M. Bafleur, P. Renaud, P. Besse, “*Deep Trench NPN Transistor for low-*RON* ESD Protection of High-Voltage I/Os in Advanced Smart Power Technology*”, in Proc. BCTM, pp. 150-153, Oct. 2006.

P. Renaud, A. Gendron, M. Bafleur, N. Nolhier, “*Efficient high voltage no-snapback and low *RON* ESD protection device for smart power technologies*”, accepté pour publication au 1st International Electrostatic Discharge Workshop, May 2007.

Congrès nationaux

A. Gendron, “*Nouvelles Structures de Protection contre les Décharges Electrostatiques en Technologies Smart Power*”, in Proc. JNRDM 2005, pp. 176-178, May 2005.

A. Gendron, “*Structures de protection ESD à faible retournement en technologie Smart Power*”, in Proc. Journées de l'école doctorale GEET, April 2006.

Brevets industriels

P. Renaud, A. Gendron, P. Besse, “*Semiconductor Device Structure and Integrated Circuit Therefor*”, PCT/EP2006/004032, March 13, 2006.

P. Renaud, A. Gendron, P. Besse, N. Nolhier, “*ESD Protection Device and Method of Forming an ESD Protection Devices*”, PCT/IB2007/052710, April 04, 2007.

Structures de protections innovantes contre les décharges électrostatiques, dédiées aux Entrées/Sorties hautes tensions de technologies SmartPower

Résumé

Les travaux effectués au cours de cette thèse visent à développer des structures de protections contre les décharges électrostatiques (ESD) adaptées aux Entrées/Sorties hautes tensions (40V–100V) dans le cadre d'applications automobiles et Ethernet. Pour satisfaire ce type de spécification, caractérisé par une fenêtre de conception ESD étroite, la protection doit présenter une caractéristique électrique à faible repliement, ou sans repliement, une tension de fonctionnement élevée et une résistance à l'état passant (R_{ON}) faible. De plus, la robustesse ESD requise est d'au moins 2kV HBM (Human Body Model) et peut atteindre 8kV.

Le transistor bipolaire autopolarisé a été retenu comme composant de base des protections développées, en raison de ses bonnes propriétés en termes de robustesse et de R_{ON} . Dans un premier temps, nous avons mené une étude théorique approfondie dans l'optique de déterminer les paramètres contrôlant son déclenchement et son comportement à l'état passant, sous l'effet des fortes densités de courant et des températures élevées induites par les ESD. En particulier, un nouveau modèle a été établi lorsque la diminution du rapport d'injection en forte injection conduit à de très faibles valeurs. Des règles de dessin adaptées aux spécifications visées ont ainsi pu être définies.

Dans ce travail, nous proposons différentes solutions permettant de répondre aux exigences de protection des Entrées/Sorties hautes tensions. Ainsi, quatre types de protections innovantes ont été développés : des transistors bipolaires PNP dont le R_{ON} est optimisé, des structures couplant un transistor PNP latéral avec une diode verticale, un transistor NPN dont le gain est fortement dégradé afin d'augmenter la tension de fonctionnement et des transistors NPN avec une région flottante dans la base ou dans le collecteur. Dans chaque cas, en s'appuyant sur l'analyse des mécanismes physiques à l'aide de la simulation physique 2D, nous avons défini les stratégies d'optimisation appropriées. Les résultats obtenus permettent de diviser par deux la surface des protections par rapport à la solution classique de mise en série de plusieurs protections basses tensions.

Mots clés

Décharge électrostatique, fiabilité des circuits intégrés, technologie SmartPower, simulation électro-thermique des composants (TCAD), transistor bipolaire autopolarisé, physique de la forte injection, dégradation thermique.

Innovative protection structures against electrostatic discharges, dedicated to high voltages Inputs/Outputs in SmartPower technologies

Abstract

The aim of this dissertation was to develop protection structures against electrostatic discharges (ESD) dedicated to high voltages (40V–100V) Inputs/Outputs, in the context of automotive or Ethernet applications. To fulfill this kind of specification, characterized by a narrow ESD design window, the protection should exhibit an electrical characteristic with a soft snapback, or no snapback, a high operating voltage and a low on-state resistance (R_{ON}). Beside, the required ESD robustness is at least 2kV HBM (Human Body Model), and could achieve 8 kV.

We have chosen to realize these protections with self-biased bipolar transistors, which have attractive robustness and R_{ON} properties. In a first step, we have lead a thorough theoretical study in order to determine the parameters controlling the triggering and the on-state behavior under high current densities and high temperatures conditions. In particular, a new model was proposed to describe the injection ratio variations when its value is very low values. By taking advantage of this description, design guidelines were defined.

In this work, we proposed have several solutions which satisfy the high voltage I/O requirements. Four types of innovative protections were developed: PNP bipolar transistors with an optimized R_{ON} , structures in which a lateral PNP transistor is coupled with a vertical diode, very low gain NPN transistors, and NPN transistors with a floating region in the base or the collector. In each case, the analysis of the physical mechanisms by 2D electro-thermal simulation has lead to define appropriate optimization strategies. The best results have allowed reducing by two the surface, compared to the standard solution based on a stack of several low voltages protections.

Keywords

Electrostatic discharge, integrated circuits reliability, SmartPower technology, electro-thermal devices simulation (TCAD), self-biased bipolar transistor, high injection physics, thermal degradation.

}