

Technologie et Physique de Transistors Bipolaires à Hétérojonction Si/SiGeC Auto-alignés très Hautes Fréquences

Benoit Barbalat

► To cite this version:

Benoit Barbalat. Technologie et Physique de Transistors Bipolaires à Hétérojonction Si/SiGeC Auto-alignés très Hautes Fréquences. Micro et nanotechnologies/Microélectronique. Université Paris Sud - Paris XI, 2006. Français. tel-00139028

HAL Id: tel-00139028

<https://tel.archives-ouvertes.fr/tel-00139028>

Submitted on 20 Dec 2007

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.

N° D'ORDRE : 8571



UNIVERSITÉ PARIS-SUD XI

Faculté des Sciences d'Orsay

THÈSE DE DOCTORAT

SPÉCIALITÉ : PHYSIQUE

*École Doctorale « Sciences et Technologies de l'Information des
Télécommunications et des Systèmes »*

Présentée par :

Benoît BARBALAT

Titre :

**Technologie et Physique de Transistors Bipolaires à
Hétérojonction Si/SiGeC Auto-alignés très Hautes Fréquences**

Soutenue le 22 décembre 2006 devant les membres du jury :

Pr. Frédéric Aniel	Directeur de thèse
Pr. Peter Ashburn	Rapporteur
Dr. Pascal Chevalier	Responsable de thèse
Pr. François Danneville	Examineur
Dr. Sylvain Delage	Rapporteur
Pr. Robert Plana	Président du jury
Dr. Alain Chantre	Invité
Dr. Nicolas Zerounian	Invité

*Étudiez comme si vous deviez vivre toujours,
Vivez comme si vous deviez mourir demain.*

Saint Isidore de Séville

Remerciements

Le travail de thèse que vous tenez entre les mains est le fruit d'une collaboration entre STMicroelectronics Crolles et l'Institut d'Electronique Fondamentale (IEF) de l'Université Paris-Sud Orsay. J'ai pu naviguer entre les deux structures d'accueil, tout en bénéficiant d'un tutorat motivant et passionné, ce qui a été extrêmement formateur. A ce titre, je tiens à remercier tout particulièrement les personnes qui m'ont encadré durant ces trois ans :

Je remercie tout d'abord Pascal Chevalier de STMicroelectronics, qui m'a permis de réaliser ce travail de thèse dans d'excellentes conditions. Merci pour m'avoir si vite laissé la main sur des projets de recherche, et pour les nombreuses publications que nous avons co-écrites.

Je tiens à remercier également Nicolas Zerounian, de l'IEF, pour les longues discussions que nous avons pu avoir sur le fonctionnement du composant. Merci également pour ton soutien et ton aide dans les mesures cryogéniques (parfois jusque tard le soir !).

Je remercie sincèrement Alain Chantre, chef d'orchestre de l'équipe «Bipavancés» de ST, pour l'appui qu'il m'a apporté au cours de la thèse. Merci pour ton soutien et la confiance que tu m'as apportés dans la réalisation des différents projets qui m'ont été confiés.

J'adresse mes plus profonds remerciements à Frédéric Aniel, qui a été mon directeur de thèse. Merci pour tes excellents conseils tout au long de ces trois années, et pour avoir su m'orienter lorsque j'en avais besoin.

Je tiens à remercier les rapporteurs et examinateurs de ce travail, Peter Ashburn, Sylvain Delage, Robert Plana et François Danneville pour avoir accepté de participer à ce jury de thèse, ainsi que pour le travail de relecture et d'évaluation qu'ils ont fourni.

L'étude des transistors bipolaires à hétérojonction n'est possible que si l'on a des transistors bipolaires à hétérojonction. Je tiens donc à remercier les différentes équipes de ST avec qui nous interagissons tous les jours pour réaliser des structures de TBH toutes plus complexes les unes que les autres...

Citons tout d'abord, l'équipe de Front-End Of Line R&D, dirigée par Didier Dutartre : Merci à Benoît Vandelle et Laurent Rubaldo, maîtres incontestés de l'HTF, pour la qualité des nombreux dépôts réalisés. Petite pensée aux anciens membres l'équipe FEOL, Cyril Fellous, Alexandre Talbot et Benjamin Oudet, partis vers des horizons plus ou moins lointains...

Je tiens à remercier Fabienne Judong et Claire Richard de l'équipe Patterning, pour leur disponibilité et les nombreux essais de gravure que nous avons réalisés.

Merci à Pierre Bouillon pour avoir pris soin de nos composants le week-end, lors d'étapes de photolithographie délicates.

Je remercie aussi Marina Proust et Pierre Caubet de l'équipe Métal, pour leur disponibilité, et avoir fait de l'émetteur métallique une réalité qui marche !

La réalisation des composants est une chose ; l'étude, l'analyse et la compréhension des phénomènes en est une autre. A ce titre, je tiens à remercier les équipes de caractérisation

physique et électrique, de simulation et de modélisation, qui soutiennent largement le travail de la filière. Je remercie tout particulièrement :

- Roland Pantel, Nadine Bicaïs, David Duca, Alain Margain, de l'équipe caractérisation physique, pour la qualité des nombreuses analyses qu'ils nous ont fournies.
- Fabienne Saguin et Daniel Gloria, pour les innombrables mesures hyperfréquences qui soutiennent ce travail de thèse. Les résultats fournis ont été plus que précieux tout au long de ces trois années.
- Jean-Charles Vildeuil et Anne Lachater, pour les mesures de bruit $1/f$ menées sur nos composants, et l'attention apportée à nos demandes.
- Thierry Schwartzmann, pour l'important travail de simulation réalisé, et la collaboration étroite que nous avons eue sur le sujet de la recombinaison en base neutre et de l'auto-échauffement.
- Didier Céli et Franck Pourchon, pour leur travail de modélisation. Merci tout particulièrement pour les habiles méthodes de mesure et d'extraction, qui ont été largement utilisées dans ce travail.
- La mise à disposition permanente des bancs de mesure statique est rendu possible par André Perrotin et Rudy Costanzi. Merci pour votre disponibilité et votre réactivité lorsque les systèmes de mesure tombent en panne (mais c'est rare).

Je tiens à adresser mes remerciements à Olivier Noblanc pour avoir toujours gardé un oeil sur moi durant mes quelques années passées à Crolles. Merci pour les longues discussions que nous avons eues, en stage, en thèse, ou ailleurs.

Il ne faudrait pas non plus croire que je n'ai fait *que* travailler durant trois ans. Aussi je tiens à remercier les nombreux thésards qui ont contribué à la bonne humeur ambiante, dans les couloirs ou dans un pub irlandais célèbre de la vallée du Grésivaudan. Citons notamment :

- Greg : *Rock On* toi même !
- Dorothee, la iPod-Girl
- Stéphane, avec qui nous avons passé de nombreuses heures sur les subtilités de L^AT_EX
- Carlo : BSC rules !
- Aurélie : Comment faire aboyer un chat ? On l'enduit d'essence, on allume et "Wouaf" !
- Ceux qui sont encore dedans : Julien, Pierre-Marie et Boris. Bon courage pour la rédaction !

Les doctorants des autres équipes sont trop nombreux pour être tous cités. Je voudrais néanmoins être sûr de ne pas oublier :

- Flo, Nico et Gaël, pour les bons moments dans le New-Jersey.
- Hélène : DT rules !
- Antoine, Bastien, Olivier, Axel, Siegfried, Nathalie, Marie, et toutes celles et ceux avec qui j'ai partagé de bons moments.

Un remerciement spécial à Luc avec qui je partage la même passion pour la musique. Que ce projet que nous montons tous les deux vive encore de longues années...

Puis merci à tous les membres de l'équipe R&D de Crolles 1 ou 2, Mertrand Bartinet, Dominik, Sébastien (×2), Michel (×2), Stéphanie, Laurence, Jocelyne, Germaine, Emmanuelle, Isabelle, Simon, pour les looongues pauses café...

Bonne continuation à tous !

J'ai passé ces trois années bien entouré par de vieux copains qui font presque partie de la famille maintenant. Aussi je tiens à remercier Jérôme, Alex, Rod, Jean-François, les potes de la promo 2003, la famille Molo, les JEM d'Inde 2002. Vous avez été largement responsables du maintien de ma bonne santé mentale, et je vous en remercie !

Je remercie également la famille proche, et en particulier mes grand-parents, qui ont su me communiquer cet amour de la science, et qui se sont toujours montrés intéressés par mes recherches. Merci pour votre chaleureuse attention, et votre soutien. Je vous dédie cette thèse.

Merci à frerot et sœurlette, bon courage pour la suite de vos études à vous !

Je remercie également mes parents, qui ont largement contribué à faire de moi ce que je suis aujourd'hui. Je ne pourrais jamais leur communiquer assez ma reconnaissance et ma gratitude à leur égard.

Mes dernières pensées vont à Mélanie, ma fiancée. Merci pour ton soutien, tes encouragements, ta complicité. A jamais, tu resteras pour moi associée à la réussite de cette thèse.

Table des matières

I	Le transistor bipolaire à hétérojonction Si / SiGeC	5
I.1	Introduction sur le transistor bipolaire	5
I.2	L'alliage Silicium-Germanium	8
I.2.1	Propriétés du SiGe	8
I.2.1.a	Paramètre de maille et caractéristiques de la couche	8
I.2.1.b	Épaisseur critique	8
I.2.1.c	Influence des contraintes mécaniques sur le diagramme de bandes	10
I.2.2	Hétérojonction Si / SiGe	10
I.2.2.a	Raccordement des bandes	10
I.2.2.b	Structure électronique d'un transistor bipolaire à hétérojonction	12
I.2.2.c	Introduction de carbone	13
I.2.2.c.i	Incorporation en site interstitiel ou substitutionnel	13
I.2.2.c.ii	Modification de la structure de bande	13
I.3	Fonctionnement statique du transistor bipolaire	14
I.3.1	Bilan des différents courants du transistor	14
I.3.2	Expression des courants	15
I.3.2.a	Équations fondamentales	15
I.3.2.b	Courant collecteur	16
I.3.2.b.i	Cas général sans recombinaison	16
I.3.2.b.ii	Cas avec recombinaison dans une base uniforme	18
I.3.2.b.iii	Discussion	19
I.3.2.c	Courant de base	20
I.3.3	Différents gains en courant	22
I.3.3.a	Montage base commune	22
I.3.3.b	Montage émetteur commun	23
I.3.4	Effets du second ordre	23
I.3.4.a	Courants non-idéaux à faible polarisation	24
I.3.4.a.i	Recombinaison en zone de charge d'espace	24
I.3.4.a.ii	Courant tunnel bande à bande	25
I.3.4.b	Effet Early	25
I.3.4.b.i	Effet Early direct	26
I.3.4.b.ii	Effet Early inverse	26
I.3.4.b.iii	Influence du profil graduel de SiGe	27
I.3.4.c	Tensions d'avalanche	28
I.3.4.c.i	Définition et caractérisation de l'avalanche	28
I.3.4.c.ii	BV_{CBO}	29
I.3.4.c.iii	BV_{EBO}	29
I.3.4.c.iv	BV_{CEO}	29

	I.3.4.d	Effet de haute injection dans la base (effet Webster)	31
	I.3.4.e	Effet Kirk	31
	I.3.4.f	Effet de barrière	32
I.3.5		Effet des forts dopages	32
	I.3.5.a	Réduction de bande interdite	33
	I.3.5.b	Modification des équations des courants	33
	I.3.5.c	Mobilité des porteurs	34
I.4		Fonctionnement dynamique du transistor bipolaire	36
	I.4.1	Temps de transit des porteurs	36
		I.4.1.a Temps de transit émetteur	37
		I.4.1.b Temps de transit dans la jonction émetteur/base	38
		I.4.1.c Temps de transit dans la base	38
		I.4.1.d Temps de transit dans la jonction base/collecteur	39
	I.4.2	Capacités de jonction	41
	I.4.3	Fréquences de transition f_T et f_{MAX}	41
		I.4.3.a Définition et expression de f_T	41
		I.4.3.b Expression de f_{MAX}	44
I.5		Effet des basses températures	46
	I.5.1	Comportement statique à basses températures	47
	I.5.2	Influence sur les performances dynamiques	48
	I.5.3	Un outil d'anticipation	49
I.6		Méthodes de caractérisation	49
	I.6.1	Caractérisation physique	49
		I.6.1.a Microscope électronique à balayage (MEB)	49
		I.6.1.b Microscope électronique à transmission (MET)	49
		I.6.1.c Spectroscopie de masse par ions secondaires	49
	I.6.2	Mesures électriques en régime statique	51
		I.6.2.a Idéauté, résistance d'émetteur	51
		I.6.2.b Caractérisation de l'avalanche dans la jonction B/C	51
		I.6.2.c Mesure des tensions d'Early	52
	I.6.3	Hyperfréquences	54
		I.6.3.a Intérêt des mesures de paramètres S	54
		I.6.3.b Méthode d'analyse	54
		I.6.3.b.i Adaptation	55
		I.6.3.b.ii Calibrage	55
		I.6.3.b.iii Épluchage	56
		I.6.3.c Extraction des différents paramètres	57
		I.6.3.d Mesures hyperfréquences à basses températures	58
	I.6.4	Régime impulsionnel	58
		I.6.4.a Intérêt des mesures pulsées (isothermes)	58
		I.6.4.b Mesures pulsées dynamiques	59
	I.6.5	Mesures de bruit	59
		I.6.5.a Bruit en $1/f$	60
		I.6.5.b Bruit haute fréquence	60
I.7		Applications	61
	I.7.1	Avantages du transistor bipolaire par rapport au MOS	61
	I.7.2	Utilisation dans le domaine des hautes fréquences	61

II	Technologie et fabrication des TBH SiGe	63
II.1	Introduction	63
II.2	Structures non auto-alignées développées dans le passé	63
II.2.1	Structure simple polysilicium	63
II.2.2	Structure double polysilicium	64
II.3	Développement d'une structure complètement auto-alignée	66
II.3.1	Nécessité d'une structure complètement auto-alignée	66
II.3.2	Structure auto-alignée par émetteur sacrificiel	67
II.3.2.a	Présentation de la structure inverse émetteur	67
II.3.2.b	Procédé de fabrication	67
II.3.2.c	Avantages / Inconvénients	68
II.3.3	Structure auto-alignée par épitaxie sélective de la base	68
II.3.3.a	Auto-alignement émetteur-base-collecteur	69
II.3.3.b	Épitaxie sélective de la base	70
II.3.3.c	Enchaînement des étapes de fabrication	70
II.3.3.c.i	Couche enterrée – Isolation	71
II.3.3.c.ii	Base extrinsèque – Fenêtre émetteur	71
II.3.3.c.iii	Base intrinsèque	71
II.3.3.c.iv	Espaceurs – Module émetteur	73
II.3.3.c.v	Recuit d'activation – Siliciuration – Back-End	73
II.3.4	Mise en place de la technologie BiCMOS	74
II.3.4.a	Enjeux de l'intégration	74
II.3.4.b	Schéma d'intégration BiCMOS	75
II.4	Comparaison par rapport à la concurrence	76
II.4.1	NXP (anciennement Philips)	76
II.4.2	IMEC	76
II.4.3	IHP	77
II.4.4	Jazz	78
II.4.5	Freescall (anciennement Motorola)	78
II.4.6	Hitachi	78
II.4.7	Infineon	79
II.4.8	IBM	80
II.4.9	Synthèse	81
III	Optimisation conventionnelle du TBH	83
III.1	But de l'optimisation verticale et latérale	83
III.2	Optimisation du module émetteur	83
III.2.1	Mise en place d'une méthode d'extraction de R_E	85
III.2.1.a	Méthode basée sur des mesures statiques	85
III.2.1.b	Influence de la zone d'extraction	86
III.2.1.c	Accord extraction statique / dynamique	86
III.2.2	Effet des différentes variantes technologiques	87
III.2.2.a	Effet du dopage du polysilicium	87
III.2.2.a.i	Type de dopage (Phosphore ou Arsenic)	87
III.2.2.a.ii	Niveau de dopage	88
III.2.2.b	Influence du Si-Cap	90
III.2.3	Effet de la géométrie du contact	91
III.2.4	Gain sur les fréquences de transition	92
III.3	Optimisation du profil de base	96

III.3.1	Profils de germanium	96
III.3.1.a	Taux de Germanium à l'entrée de la base	96
III.3.1.b	Épaisseur de la couche de SiGe	99
III.3.1.c	Profil de Germanium graduel ou constant	101
III.3.1.d	Synthèse de l'optimisation du profil de base	102
III.3.2	Introduction d'un nouveau profil de base	103
III.4	Optimisation du profil collecteur	107
III.4.1	Collecteur implanté sélectivement (SIC)	107
III.4.2	Résultats électriques	108
III.4.3	Influence du SIC sur les performances dynamiques	109
III.4.4	Compromis entre f_T et BV_{CEO}	110
III.5	Réduction des dimensions latérales du dispositif	111
III.5.1	Largeur d'émetteur W_E	112
III.5.1.a	Contact d'émetteur	112
III.5.1.b	Largeur de la fenêtre émetteur	113
III.5.1.c	Espaceurs internes	114
III.5.2	Largeur de la zone d'active	115
III.5.3	Largeur du polyémetteur.	116
III.5.4	Synthèse de l'optimisation latérale.	117
III.6	Synthèse de l'optimisation conventionnelle	118
III.6.1	Performances	118
III.6.2	Développement d'une filière dédiée aux applications millimétriques	119
IV	Amélioration du produit $f_T \times BV_{CEO}$	121
IV.1	Présentation de l'approche	121
IV.2	Transistor bipolaire à émetteur métallique	124
IV.2.1	Réalisation technologique	124
IV.2.1.a	Architectures possibles	124
IV.2.1.b	Détails du procédé PRETCH	125
IV.2.2	Résultats obtenus avec un émetteur fortement dopé	126
IV.2.2.a	Résultats électriques	127
IV.2.2.b	Analyse SIMS	128
IV.2.2.c	Mesures d'effet Hall	129
IV.2.2.d	Interprétation – Suite de l'étude	130
IV.2.3	Résultats obtenus avec un émetteur faiblement dopé	130
IV.2.3.a	Résultats électriques	131
IV.2.3.b	Problème d'augmentation des résistances séries	132
IV.2.4	Perspectives	133
IV.3	Insertion de SiGe dans l'émetteur	134
IV.3.1	Mise en œuvre	134
IV.3.2	Résultats électriques	135
IV.3.2.a	Résultats statiques	135
IV.3.2.b	Performances dynamiques	136
IV.3.3	Conclusion sur l'insertion de Ge dans l'émetteur	136
IV.4	Recombinaison en base neutre	138
IV.4.1	Mise en œuvre	138
IV.4.2	Résultats électriques	139
IV.4.3	Phénomène de saturation des pièges	140
IV.4.4	Évolution du bruit basse fréquence	142

IV.4.5	Influence du carbone sur le gap du matériau	143
IV.4.6	Conclusion sur la recombinaison en base neutre.	145
IV.5	Synthèse de l'optimisation	146
IV.5.1	Comparaison des différentes méthodes	146
IV.5.2	Positionnement par rapport à la concurrence	147
IV.5.3	Conclusion / Perspectives	147
V	Influence de la température	149
V.1	Introduction	149
V.2	Étude de l'auto-échauffement	149
V.2.1	Limitations induites par l'auto-échauffement	149
V.2.2	Extraction de la résistance thermique	150
V.2.2.a	Extraction en régime continu	150
V.2.2.b	Extraction en régime pulsé	153
V.2.2.c	Comparaison des différentes méthodes	154
V.2.3	Réduction de l'auto-échauffement par des variantes technologiques	154
V.2.3.a	Influence de la dimension de l'émetteur	154
V.2.3.b	Effet des tranchées profondes d'isolation	155
V.2.3.c	Émetteurs fractionnés ou multi-doigts	158
V.2.3.d	Effet du type de substrat	160
V.2.3.e	Élaboration d'un modèle simple	162
V.3	Étude du TBH à basse température	166
V.3.1	Modification des paramètres statiques à basse température	166
V.3.1.a	Evolution des niveaux de courant	166
V.3.1.b	Variation du gain en courant avec la température	167
V.3.1.c	Évolution des autres paramètres statiques	168
V.3.2	Performances dynamiques à basse température	169
V.3.2.a	Fréquences f_T et f_{MAX} aux températures cryogéniques	169
V.3.2.b	Étude des temps de transit	172
V.3.3	Synthèse de l'étude cryogénique	174
V.4	Conclusion / Perspectives	175
	Conclusion générale	177
	Publications de l'auteur	181
	Références bibliographiques	183
	Liste des figures	193
	Liste des tableaux	201

Introduction

Jusqu'à il y a une dizaine d'années, l'industrie de la microélectronique était cloisonnée : D'un côté, on trouvait les circuits logiques complémentaires, représentant l'écrasante majorité du marché, principalement pour des applications logiques comme les microprocesseurs. Le composant sur lequel reposent ces technologies, le transistor MOSFET (Metal Oxide Semiconductor Field Effect Transistor), est en effet facilement réalisable à partir d'un substrat en silicium et propice à de forts volumes de production en milieu industriel, ce qui permet de gros volumes de production à des coûts relativement faibles. Ceci explique l'incroyable essor des technologies silicium CMOS (Complementary MOS) dans la seconde moitié du XX^e siècle, prenant ainsi l'avantage sur les technologies utilisant des transistors bipolaires à homojonction (logique TTL notamment).

D'un autre côté, les circuits III-V (GaAs, GaN, InP, ou alliages plus complexes...) permettent d'adresser des applications analogiques et hyper-fréquences en utilisant des transistors bipolaires à hétérojonction, des MESFET (MEtal Semiconductor Field Effect Transistor) ou des HEMT (High Electron Mobility Transistor). Utilisés dans le domaine des télécommunications ou en environnement militaire pour la détection et le guidage, ces circuits sont produits en plus faibles quantités, et à des coûts bien supérieurs à ceux d'un circuit CMOS. Le prix élevé des technologies III-V est d'une part lié au faible volume de production, et d'autre part à la limitation de la taille du substrat ainsi qu'au coût de celui-ci, bien supérieur au coût d'un substrat silicium. Les transistors bipolaires rapides développés pour les applications de télécommunication et de détection étaient historiquement fabriqués en matériaux III-V. Il est en effet possible de modifier la structure de bandes du composant et de créer ainsi le profil favorable à l'obtention des meilleures performances dynamiques. Les fréquences obtenues avec un transistor bipolaire InP s'élèvent au dessus du THz. Ces composants III-V sont encore très utilisés pour les applications millimétriques ou bien nécessitant de fortes puissances, car les propriétés de transport des électrons sont très bonnes dans ce type de matériau, permettant d'excellentes performances.

Cependant, la diversification des applications grand public entraîne des contraintes de coût, d'encombrement et de consommation de plus en plus importantes. On a vu apparaître la nécessité de faire cohabiter des circuits logiques avec des circuits analogiques ou hyperfréquences sur une même puce, notamment pour les applications de communication sans fil.

A mesure de l'amélioration des performances dynamiques des circuits BiCMOS, les composants III-V ont donc été remplacés petit à petit par des circuits tout silicium, ce qui permet d'augmenter l'intégration et de réduire les coûts de fabrication. Un nouveau type de technologie a donc été développé pour répondre aux besoins des applications hautes fréquences à destination du grand public. Les technologies dites BiCMOS utilisent des transistors bipolaires et des transistors MOS, et allient ainsi les avantages des deux types de composants : Fortes fréquences de fonctionnement et faible bruit pour le premier, haute densité d'intégration et faible consommation pour le second.

Historiquement, les technologies BiCMOS employaient des transistors bipolaires tout silicium. Il y a une dizaine d'années, l'introduction du Germanium dans la base du composant

a permis l'ingénierie de la structure de bandes. Les performances des transistors bipolaires se sont donc rapidement améliorées grâce aux progrès réalisés sur les matériaux. Le SiGe offre à l'heure actuelle une maturité de procédé, un niveau d'intégration et des rendements proches de ce qui est couramment obtenu pour le silicium. Les technologies utilisant du SiGe commencent à remplacer les matériaux III-V pour la partie réception des systèmes de communication dans le domaine RF-basse fréquence ; la partie émission utilisant des amplificateurs de puissance restant encore l'apanage des composants III-V, ainsi que le domaine des très hautes fréquences.

C'est dans le cadre du développement de technologies BiCMOS que s'inscrit mon travail de thèse. Le but de l'étude est la réalisation et l'analyse de transistors bipolaires rapides réalisés en anticipation des technologies BiCMOS futures. Cette thèse a pour objectif de mettre en avant les limitations existant dans la structure de transistor bipolaire la plus récente, et de déboucher sur des solutions performantes et maîtrisables en milieu industriel afin d'améliorer les performances des composants. Pour cela, nous disposons des procédés de fabrication avancés disponibles à STMicroelectronics, ainsi que des techniques et méthodes d'analyse expérimentales de l'IEF. Cette thèse a également pour but de dresser un tableau de perspectives pour l'optimisation ultérieure du composant.

Dans un premier temps, nous rappelons la théorie de fonctionnement du transistor bipolaire. Nous mettons en avant les propriétés du matériau SiGe, et nous présentons les améliorations qu'apporte ce matériau sur le transistor bipolaire. Le fonctionnement théorique du TBH est ensuite décrit, en régime statique et dynamique. Nous détaillons les outils d'analyse expérimentale (caractérisation physique et électrique), et notamment l'intérêt de l'analyse du composant à basses températures.

Au chapitre II, nous détaillons les procédés de fabrication des transistors bipolaires à hétérojonction Si/SiGe. Après un bref passage en revue des structures développées par le passé, nous présentons le choix d'architecture qui s'offrait à nous pour les futures technologies BiCMOS. Les raisons ayant mené à la structure actuelle sont explicitées, et le procédé de fabrication des composants est décrit en détail. En fin de ce chapitre, nous proposons de situer les performances des transistors bipolaires de ST dans une analyse de l'état de l'art. Ainsi, les architectures et les résultats d'autres acteurs du marché de la microélectronique sont présentés. Cette analyse permet de mettre en avant les avantages et points faibles de chaque structure.

Le chapitre III traite de l'optimisation dite « conventionnelle » des transistors bipolaires qui a permis d'atteindre les performances présentées dans le chapitre précédent. Nous étudions tout d'abord des améliorations apportées par l'optimisation du profil vertical de chaque partie du composant (émetteur, base et collecteur), et présentons ensuite l'optimisation de l'extension latérale du composant. Une synthèse de cette optimisation classique est ensuite réalisée, en résumant l'évolution des performances obtenues aux cours des trois ans qu'a duré cette étude. Nous terminons le chapitre par des considérations sur l'intégration de ces composants pour réaliser des circuits rapides.

Le chapitre sur l'optimisation conventionnelle a mis l'accent sur le compromis existant entre les performances dynamiques et la tenue en tension du transistor. Le chapitre IV présente trois procédés innovants permettant de repousser ce compromis en augmentant le courant de base du transistor. Chaque procédé offre ainsi la possibilité d'améliorer la tenue en tension du composant, par une variante technologique appropriée, permettant de modifier le comportement du composant. Ainsi, nous étudierons successivement les améliorations apportées par le procédé émetteur métallique, l'insertion de Ge dans l'émetteur, et l'augmentation de la recombinaison

dans la base neutre par l'ajout de fortes doses de carbone. Chaque procédé est décrit, ainsi que ses avantages et ses inconvénients, et nous replaçons ensuite les performances obtenues par rapport à l'état de l'art.

Le dernier chapitre a pour objet l'étude du transistor bipolaire en fonction de la température. La première partie du chapitre porte sur l'auto-échauffement du composant. Le chapitre III a en effet mis en avant la possibilité d'améliorer les performances grâce à l'augmentation de la densité de courant, mais ceci a pour effet néfaste de générer beaucoup de chaleur. L'élévation de température interne, inhérente au fonctionnement du transistor apparaît donc comme une limitation des performances. Après avoir mis en place une méthode pour caractériser cet auto-échauffement, nous présentons les techniques existantes pour réduire la montée en température. Le second point porte sur l'analyse cryogénique du transistor bipolaire. En effet, les performances dynamiques sont améliorées lorsqu'on descend à des températures très basses, ce qui permet d'en tirer des informations importantes sur la capacité de la structure à évoluer vers des performances ultimes. Nous montrerons à ce titre que la principale limitation du transistor bipolaire se situe dans la partie intrinsèque.

Nous concluons cette thèse par une synthèse des résultats obtenus, et nous évoquerons des perspectives quant à l'évolution future du transistor bipolaire à hétérojonction Si/SiGe pour les technologies BiCMOS à venir.

Chapitre I

Le transistor bipolaire à hétérojonction Si / SiGeC

I.1 Introduction sur le transistor bipolaire

Le transistor bipolaire a été inventé en 1948 par Bardeen et Brattain, et sa théorie a été élaborée en 1949 par Shockley. Depuis 1951, date du premier transistor à jonction, les développements poussés, tant sur les matériaux que sur les dimensions du dispositif, ont permis d'atteindre des performances records, largement utilisées dans le domaine des télécommunications ou dans des applications nécessitant de fortes puissances.

Le transistor bipolaire est un composant électronique composé de trois zones semiconductrices successives dopées dans les deux configurations possibles NPN ou PNP, appelées successivement *émetteur*, *base* et *collecteur*. Il s'agit donc de deux jonctions PN tête-bêche ayant une région en commun comme représenté sur la figure I.1.

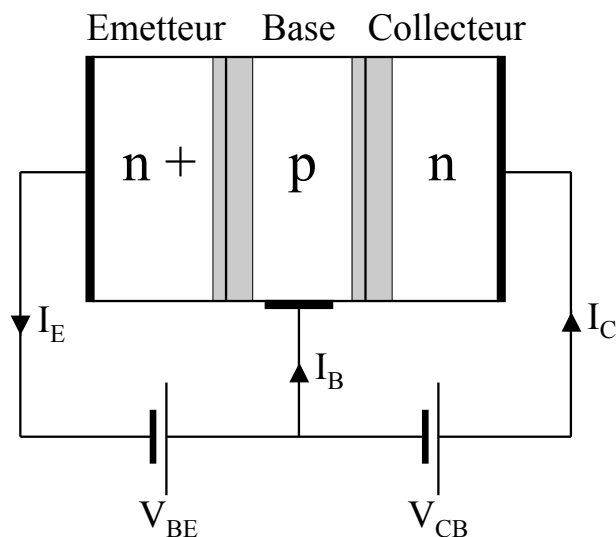


FIG. I.1 – Schéma d'un transistor bipolaire NPN. En noir : contacts métalliques ; zones grisées : zones de charge d'espace des jonctions E/B et B/C.

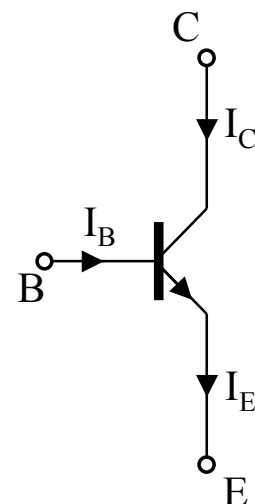


FIG. I.2 – Représentation symbolique du NPN utilisée dans les schémas électriques.

Le transistor NPN est le plus utilisé pour les applications nécessitant des circuits rapides, car la conduction repose principalement sur les électrons, qui sont plus rapides que les trous. Le transistor considéré tout au long de cette thèse est le transistor NPN, c'est celui qui a fait

l'objet des développements les plus poussés au cours des dernières années. Il faut cependant noter que les différents courants du transistor bipolaire font intervenir les deux types de porteurs (électrons *et* trous), contrairement aux composants à effet de champ, où un seul type de porteur est sollicité.

Le transistor bipolaire peut être polarisé de 4 manières différentes, qui déterminent chacune un mode de fonctionnement. Les différents modes de fonctionnement du transistor bipolaire sont :

- Mode direct : Jonction émetteur/base en direct et base/collecteur en inverse.
- Mode saturé : Jonction émetteur/base en direct et base/collecteur en direct.
- Mode inverse : Jonction émetteur/base en inverse et base/collecteur en direct.
- Mode bloqué : Jonction émetteur/base en inverse et base/collecteur en inverse.

Le mode le plus couramment utilisé dans les applications analogiques et radio-fréquences est le mode direct. C'est la proximité des deux jonctions du composant qui est à la base de l'*effet transistor* : Pour obtenir un bon fonctionnement du composant, il faut que les porteurs minoritaires injectés dans la base par l'émetteur, les électrons, parviennent jusqu'à la jonction base/collecteur. Ceci implique l'emploi de bases assez fines pour éviter la recombinaison en volume, il est impératif que la longueur de diffusion des électrons soit supérieure à l'épaisseur de la base neutre W_B .

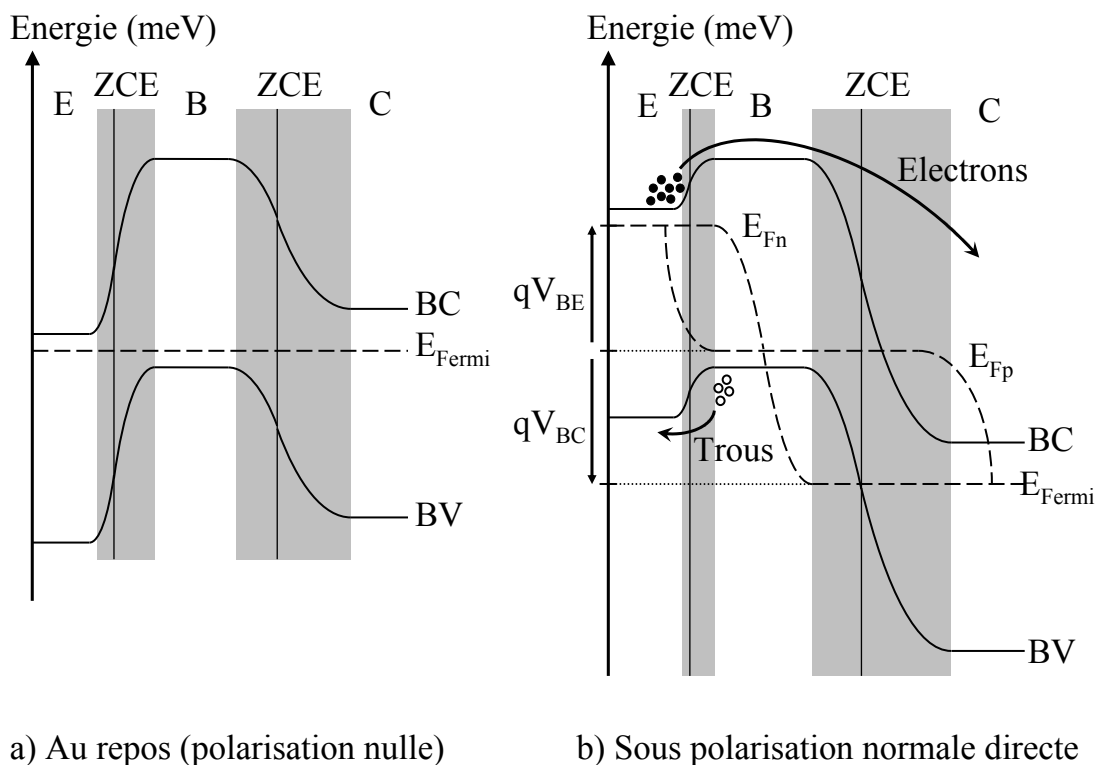


FIG. I.3 – Diagramme de bandes d'un transistor bipolaire NPN : a – au repos et b – en régime de fonctionnement normal. L'abaissement des barrières pour les électrons et pour les trous autorise le passage du courant.

Le profil de bandes d'un transistor NPN au repos est présenté sur la figure I.3–a. Lorsque la jonction émetteur/base est polarisée en direct, les barrières s'abaissent pour les trous et les électrons, autorisant ainsi le passage du courant. Les électrons diffusent de la base vers le collecteur, et sont happés par la zone de charge d'espace (ZCE) base/collecteur. La polarisation inverse de cette jonction base/collecteur sert à créer un très fort champ électrique qui accélère les électrons (figure I.3–b).

Le transistor bipolaire est un composant dit actif qui se comporte comme une source de courant commandée en tension. La jonction émetteur/base contrôle le courant principal du transistor. Dans le cas d'un transistor à homojonction, il faut que la jonction E/B soit fortement dissymétrique (i.e. dopage d'émetteur très supérieur au dopage de base) pour que la majorité des porteurs injectés soient des électrons, permettant ainsi d'obtenir une efficacité d'injection maximale. La quantité de trous injectés dans l'émetteur sera très faible en regard des électrons injectés dans la base.

Ces électrons injectés dans la base atteignent la jonction base/collecteur par un mécanisme de diffusion. Le courant collecteur dépend donc du gradient d'électrons dans la base. Dans un transistor idéal le courant ne doit pas varier lorsque la polarisation de la jonction base/collecteur varie. Afin de garantir cet effet, il ne faut pas que le gradient d'électrons dépende de la tension B/C. Ceci est possible en assurant un dopage de base très supérieur à celui du collecteur. On en déduit le profil de dopage générique d'un transistor représenté sur la figure I.4 suivante :

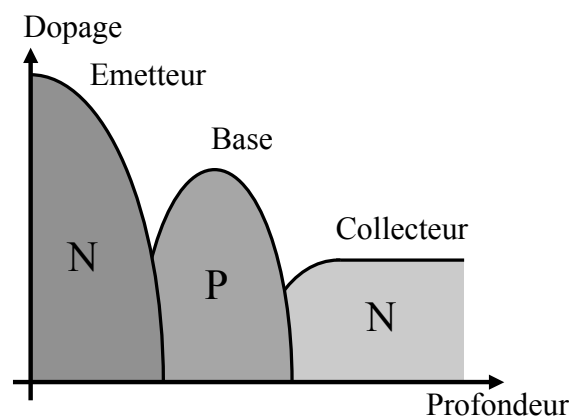


FIG. I.4 – Profil de dopage générique d'un transistor bipolaire

I.2 L'alliage Silicium-Germanium

Afin d'obtenir de meilleures performances, le profil du transistor bipolaire doit être optimisé. Ainsi, l'ingénierie de bandes rendue possible par l'introduction de Germanium (Ge) dans le composant consiste en fournir le meilleur profil de bandes d'énergie pour le passage des électrons, approche déjà utilisée en III-V pour la réalisation des transistors bipolaires à hétérojonction. Dans cette partie, nous présentons l'alliage silicium-germanium utilisé pour réaliser la base des transistors bipolaires, et les améliorations apportées par l'introduction du germanium sur la structure électrique du composant.

I.2.1 Propriétés du SiGe

I.2.1.a Paramètre de maille et caractéristiques de la couche

Le silicium et le germanium sont deux éléments de la colonne IV du tableau de Mendeleïev, cristallisant sous la structure diamant, au même titre que le carbone. Les paramètres de maille du Si et du Ge sont les suivants :

$$\begin{cases} a_{Si} = 5.431 \text{ \AA} \\ a_{Ge} = 5.657 \text{ \AA} \end{cases}$$

Le silicium et le germanium sont complètement miscibles, quelles que soient les fractions molaires de chaque espèce. L'alliage SiGe est donc stable et reproduit la structure cristallographique du Si et du Ge. Le paramètre de maille de l'alliage $Si_{1-x}Ge_x$ dépend de la proportion de Ge introduit, selon la loi de Végard décrite par l'équation I.1, où x représente le taux de Germanium.

$$a_{Si_{1-x}Ge_x} = a_{Si} + (a_{Ge} - a_{Si}) \cdot x \quad (I.1)$$

La loi donnant le paramètre de maille d'un cristal de $Si_{1-x}Ge_x$ est linéaire en fonction du taux de Germanium x . A température ambiante, la différence de maille entre le Si et le Ge est de 4.17 %. A cause de ce désaccord de maille, la croissance d'une couche de SiGe sur un substrat de silicium peut se faire selon deux modes différents exposés dans la figure I.5.

- Soit la couche est contrainte, on parle de croissance *pseudomorphique* (cas (a)). La maille de l'alliage SiGe reproduit la maille plus petite du substrat dans le plan de l'interface et se déforme élastiquement dans la direction orthogonale. La contrainte exercée sur la maille de SiGe dans les deux directions constituant le plan de l'interface est dite *biaxiale*.
- Soit la couche est relaxée (cas (b)). Le SiGe ne reproduit pas la maille du substrat et garde son paramètre de maille propre. L'interface entre le Si et le SiGe est marquée par des dislocations, liaisons manquantes ou pendantes, générant des états d'interface.

I.2.1.b Épaisseur critique

Si la couche déposée est suffisamment fine, et si le paramètre de maille diffère peu de celui du substrat, il a été démontré que la couche est intégralement contrainte [People85b]. En revanche, au delà d'une épaisseur critique h_C , l'énergie de la couche devient trop importante et la couche se relaxe. L'épaisseur critique dépend du désaccord de maille entre le SiGe et le substrat, donc du taux de Ge dans la couche, et des conditions de dépôt, notamment la température, comme indiqué dans la figure I.6. Au delà de h_C , des dislocations apparaissent, et le SiGe retrouve son paramètre de maille normal. Ces dislocations, génératrices de défauts cristallographiques peuvent créer des défauts électriques et des courts-circuits, tueurs pour le composant.

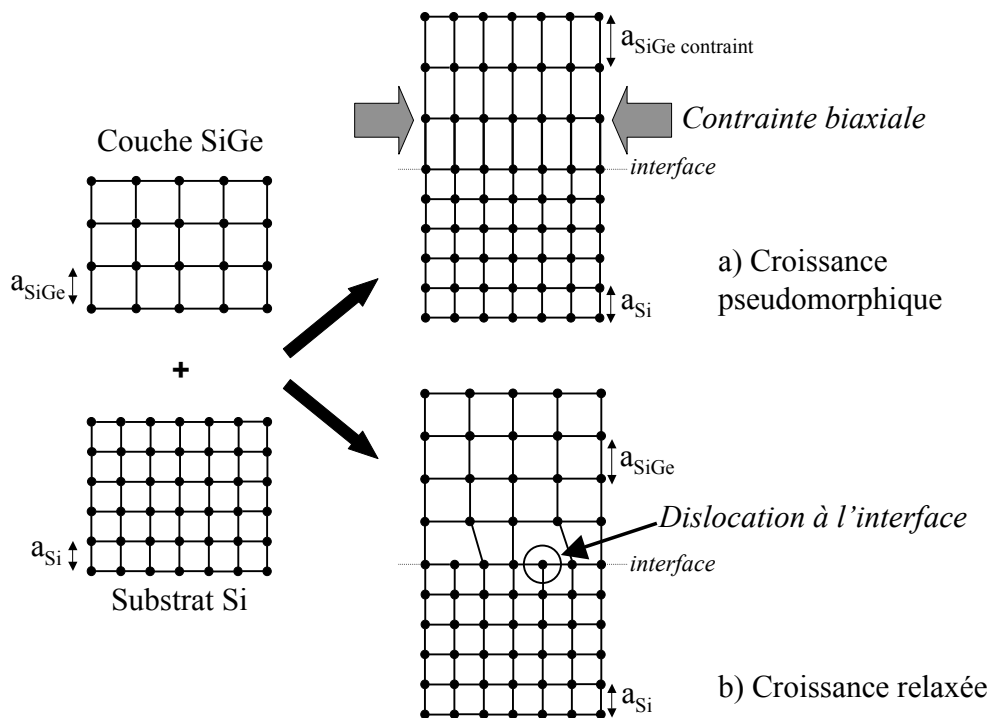


FIG. I.5 – Représentation 2D des deux types de croissance possibles : (a) Croissance pseudomorphique, avec contrainte biaxiale dans le plan de l'interface ; (b) Croissance relaxée, des dislocations apparaissent dans le plan de l'interface.

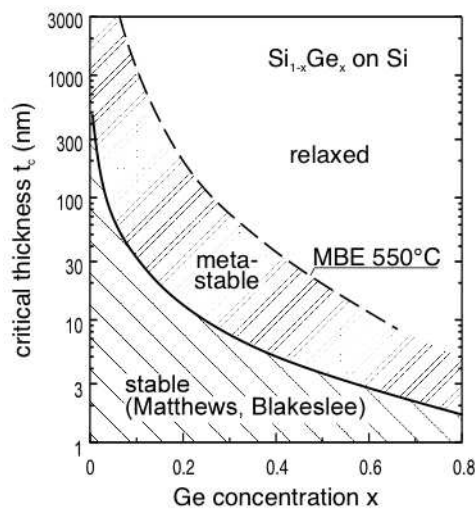


FIG. I.6 – Epaisseur critique h_C en fonction du taux de germanium x .

Entre une épaisseur faible, où la couche est stable, et une grande épaisseur, où la couche est entièrement relaxée, il existe une zone dite *métastable*, dans laquelle la couche est intégralement contrainte, mais tout apport d'énergie supplémentaire (implantation ionique, recuit...) pourrait avoir pour conséquence l'apparition de dislocations dans la couche de SiGe.

L'épaisseur critique h_C est généralement augmentée par la présence, au dessus de la couche de SiGe déposée, d'une couche de silicium qui aura donc le même paramètre de maille que le substrat. Le SiGe pris en « sandwich » entre deux couches de Si est stabilisé et risque ainsi moins de relaxer, ce qui est le cas dans la base de nos transistors bipolaires.

I.2.1.c Influence des contraintes mécaniques sur le diagramme de bandes

Le silicium Si et l'alliage silicium-germanium SiGe ont le même type de réseau cristallographique, mais les énergies de bande interdite (ou *gap*) sont différentes. A 300K, le silicium non contraint a un gap de 1.12eV, le germanium 0.66eV.

On voit dans la figure I.7 que le minimum de bande de conduction est localisé dans des directions différentes selon le matériau : direction [100] pour le Si (dégénéré 6 fois) et [111] pour le Ge (dégénéré 8 fois).

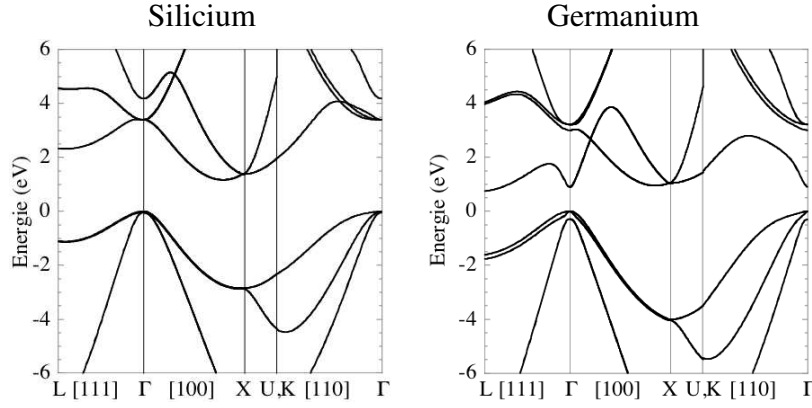


FIG. I.7 – Diagrammes de bande du Si et du Ge à 0K (données [Richard04])

L'alliage de $\text{Si}_{1-x}\text{Ge}_x$ aura un gap compris entre les gaps du Si et du Ge, qui dépend de divers paramètres :

- Le taux de Germanium x : La bande interdite de l'alliage $\text{Si}_{1-x}\text{Ge}_x$ a une valeur comprise entre celle du Si pur et du Ge pur (voir la courbe du haut de la figure I.8). La rupture qui intervient au delà de $x = 0.85$ dans le cas non contraint est due au fait que le minimum de bande de conduction se trouve dans des directions différentes de l'espace réciproque pour le Si et le Ge.
- Le fait que la couche soit contrainte ou non : Une couche de SiGe contrainte aura une bande interdite systématiquement plus faible que la couche non contrainte ayant le même taux de Ge [Lang85] (figure I.8, courbes du bas). Les deux courbes sont dues à une levée de dégénérescence en bande de valence : HH pour "Heavy Holes" (trous lourds) et LH pour "Light Holes" (trous légers).

On peut approximer la valeur de la bande interdite du $\text{Si}_{1-x}\text{Ge}_x$ contraint sur Si en fonction de x selon l'expression de People [People85a] :

$$E_{g\text{SiGe}}(x_{\text{Ge}}, T) = E_0(T) - 1.02 \cdot x_{\text{Ge}} + 0.52 \cdot x_{\text{Ge}}^2 \quad [\text{eV}] \quad (\text{I.2})$$

Où $E_0(T)$ est l'énergie de bande interdite du Si pur non contraint, et x_{Ge} la teneur en Ge.

I.2.2 Hétérojonction Si / SiGe

I.2.2.a Raccordement des bandes

L'hétérojonction Si/SiGe contraint est de type I (voir figure I.9) [People86]. Les décalages des bandes sont telles que la discontinuité est reportée principalement en bande de valence. La figure I.10 montre les décalages des bandes de conduction et de valence d'une couche de $\text{Si}_{1-x}\text{Ge}_x$ contrainte sur un substrat de $\text{Si}_{1-x_s}\text{Ge}_{x_s}$, en fonction de x et de x_s , d'après [Schäffler97].

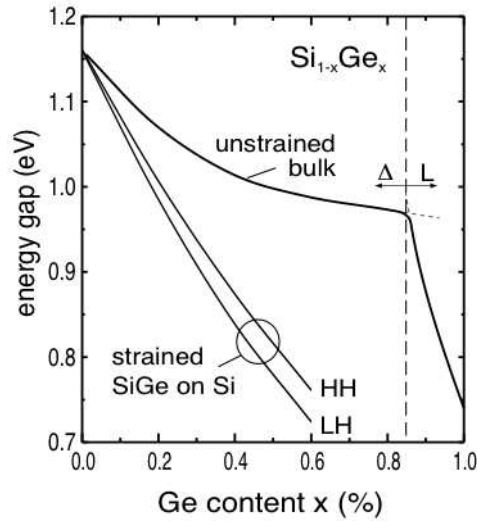


FIG. I.8 – Valeur de la bande interdite du $\text{Si}_{1-x}\text{Ge}_x$ contraint ou non sur Si en fonction du taux de Ge x (données [Lang85]).

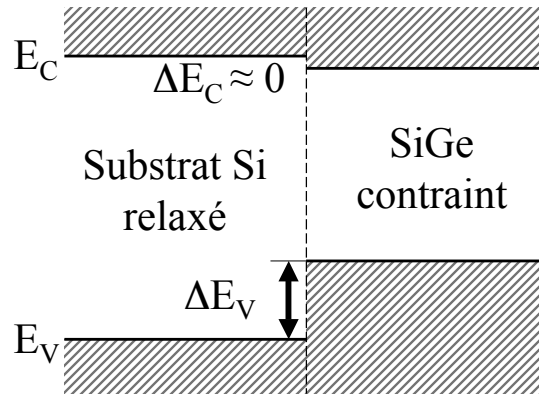


FIG. I.9 – Schéma de l'alignement de bandes entre le Si et le SiGe contraint sur Si.

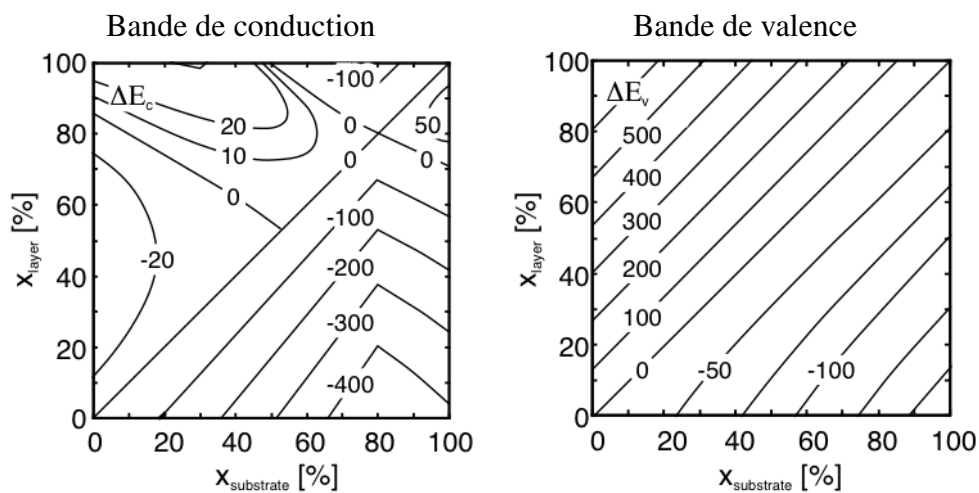


FIG. I.10 – Variations (en meV) en bande de conduction et de valence d'une couche de $\text{Si}_{1-x}\text{Ge}_x$ contrainte sur un substrat de $\text{Si}_{1-x_s}\text{Ge}_{x_s}$ relaxé en fonction des taux de germanium x et x_{substrat} (données [Schäffler97]).

On remarque que pour des concentrations de Ge comprises entre 0 et 30 % (typiques des bases des transistors bipolaires), la principale discontinuité est visible en bande de valence, et peut être évaluée simplement par l'équation I.3 suivante :

$$\Delta E_V = 0.74 \cdot x_{Ge} \text{ [eV]} \quad (\text{I.3})$$

I.2.2.b Structure électronique d'un transistor bipolaire à hétérojonction

L'effet du Germanium sur la bande interdite de la base est avantageusement utilisé dans le transistor bipolaire à hétérojonction afin d'assurer des courants collecteur plus forts et diminuer le temps de transit des porteurs dans la base, et ce grâce à deux effets conjugués :

- Le SiGe présent dans la base permet d'abaisser la barrière énergétique vue par les électrons (figure I.11), ce qui assure une plus forte injection de porteurs. Le courant collecteur est donc plus important que dans le cas d'une homojonction à même polarisation. Comme nous le verrons plus loin, un courant collecteur fort assure un f_T plus important (Chapitre I.4.3 page 41).
- La création d'un champ accélérateur est possible en utilisant un profil rétrograde de SiGe. Dans un matériau de type P, l'écart entre le niveau de Fermi et la bande de valence est fixé par le dopage. La concentration de Ge étant plus forte côté collecteur que côté émetteur, la bande interdite du SiGe de la base diminue progressivement, il en résulte un pseudo-champ électrique vu par les électrons, créé par l'abaissement de la bande de conduction entre l'émetteur et le collecteur (voir figure I.12).

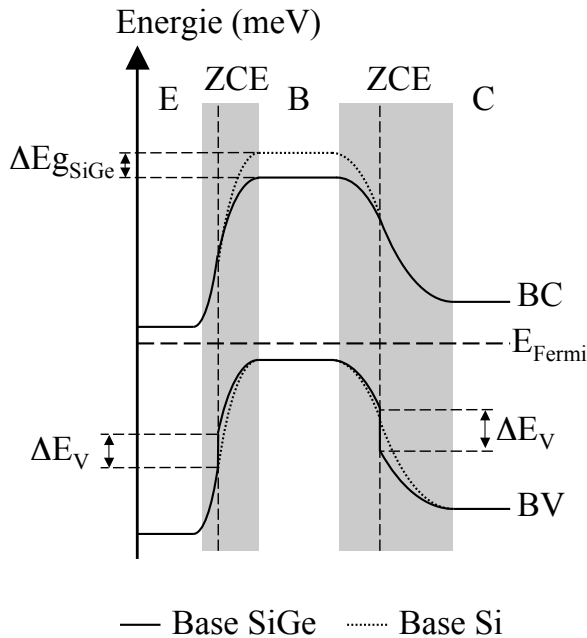


FIG. I.11 – Diagramme de bandes d'un transistor bipolaire à hétérojonction SiGe, à taux de Ge constant.

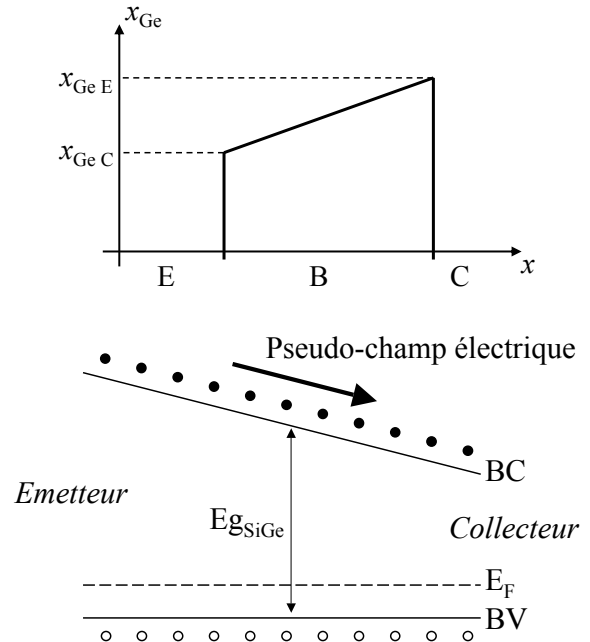


FIG. I.12 – Pseudo-champ électrique créé par l'abaissement de la bande de conduction dans le cas d'un profil rétrograde de Germanium.

I.2.2.c Introduction de carbone

I.2.2.c.i Incorporation en site interstitiel ou substitutionnel

Un autre matériau a vu le jour dans les technologies BiCMOS, il s'agit du SiGeC. Historiquement, le carbone a été introduit dans les bases SiGe des TBH pour compenser la contrainte générée par le Germanium [Lanzerotti96a]. Par la suite, il a été démontré que le carbone a la propriété de bloquer la diffusion du bore dans la base SiGe [Lanzerotti96b, Osten97], et ainsi permettre des bases plus fines grâce à un meilleur confinement des atomes dopants. En revanche, l'introduction de carbone n'est pas neutre sur le fonctionnement du transistor bipolaire : Outre la modification du profil de dopage dans la base, le carbone a également un effet direct sur le diagramme de bande lorsqu'il est incorporé en site substitutionnel. En site interstitiel, il peut créer des complexes avec les atomes du réseau, résultant en l'apparition de centres recombinant [Osten97]. Ces centres ont un effet non négligeable sur les courants du transistor. Selon leur localisation dans le composant, il peut en résulter des courants de fuite importants. L'effet du carbone sur la recombinaison de porteurs peut toutefois être tourné à notre avantage lorsque l'insertion de carbone est maîtrisée (voir chapitre IV.4 sur la recombinaison en base neutre).

I.2.2.c.ii Modification de la structure de bande

Le paramètre de maille du carbone diamant est de 3.546 Å, donc beaucoup plus petit que celui du silicium ou du germanium. En conséquence, le principal effet du carbone est de diminuer les contraintes de la couche de SiGe, et ainsi augmenter le gap du matériau. D'après [Boucaud94], 1% de carbone augmente le gap de 17 meV. Il y a donc un impact direct du carbone sur les niveaux de courant I_C et I_B .

Dans les bases des transistors bipolaires standard, la concentration de carbone est de quelques $10^{19} \text{ at}\cdot\text{cm}^{-3}$, soit moins de 0.05 %. La variation de bande interdite due au carbone dans un TBH standard est donc négligeable.

I.3 Fonctionnement statique du transistor bipolaire

I.3.1 Bilan des différents courants du transistor

Comme signalé précédemment, le transistor bipolaire est un composant qui fait intervenir les deux types de porteurs, électrons et trous. En régime de fonctionnement normal, les différents courants du composant sont répartis de la manière suivante :

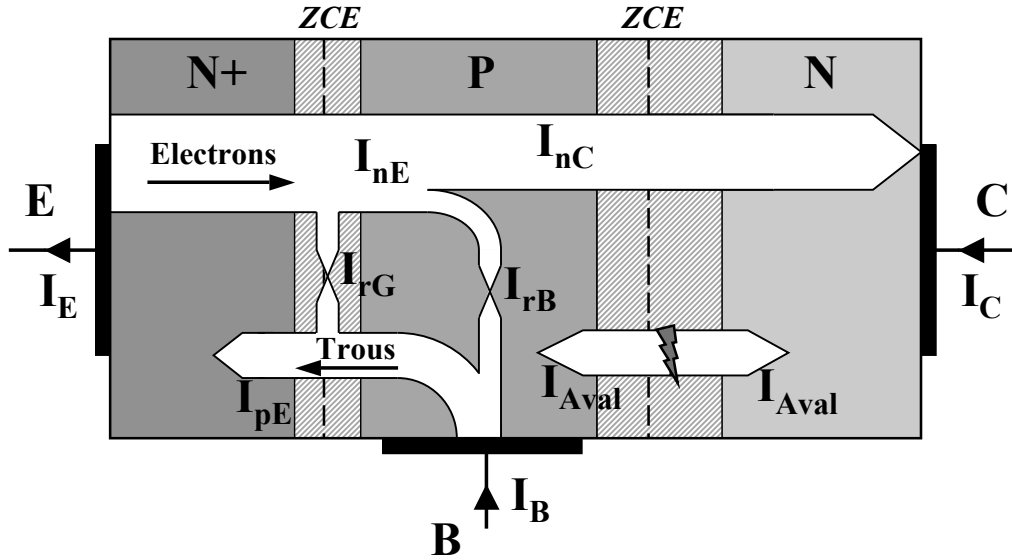


FIG. I.13 – Principaux courants du transistor bipolaire dans un mode de fonctionnement direct.

La jonction émetteur/base polarisée en direct injecte des électrons dans la base (courant I_{nE}) et des trous côté émetteur (I_{pE}). Il est possible d'avoir une recombinaison entre les trous et les électrons dans la ZCE, créant ainsi la composante I_{rG} . Les trous sont injectés dans un milieu fortement dopé N et se recombinent rapidement dans le volume de l'émetteur. Les électrons diffusent de l'émetteur vers le collecteur dans la base dopée P, il est donc également possible d'y avoir de la recombinaison en volume, créant la composante de recombinaison en base neutre I_{rB} . Les électrons qui ont traversé la base (courant I_{nC}) constituent l'intégralité du courant collecteur.

Selon la figure I.13, le courant d'émetteur a trois composantes :

- I_{nE} : Le courant d'électrons injectés de l'émetteur vers la base. C'est la composante principale du courant
- I_{pE} : Le courant de trous injectés de la base vers l'émetteur
- I_{rG} : Le courant de recombinaison dans la ZCE de la jonction émetteur/base.

On en déduit la décomposition du courant d'émetteur :

$$I_E = I_{nE} + I_{pE} + I_{rG} \quad (I.4)$$

Le courant de base est composé de :

- I_{pE} : Le courant de trous injectés de la base vers l'émetteur
- I_{rG} : Le courant de recombinaison dans la ZCE émetteur/base.
- I_{rB} : Le courant de recombinaison dans la base neutre.

Il en découle :

$$I_B = I_{pE} + I_{rB} + I_{rG} \quad (I.5)$$

Le courant collecteur est constitué uniquement du courant d'électrons ayant traversé la base, c'est à dire uniquement la composante I_{nC} .

$$I_C = I_{nC} = I_{nE} - I_{rB} \quad (I.6)$$

Les trois équations précédentes permettent bien de vérifier :

$$\begin{aligned} I_E &= \underbrace{I_{nC}} + \underbrace{I_{pE} + I_{rG} + I_{rB}} \\ I_E &= I_C + I_B \end{aligned} \quad (I.7)$$

Compte-tenu des dimensions du dispositif et des temps de transit très faibles de porteurs dans le composant, les composantes de recombinaison I_{rG} et I_{rB} sont généralement négligeables. Cependant, des problèmes lors de la fabrication du transistor peuvent déboucher sur la création d'un grand nombre de centres recombinant. S'ils sont situés dans la zone de charge d'espace émetteur/base, ils seront néfastes pour le composant, créant des courants non idéaux.

Sur la figure I.13 apparaissent également des courants d'avalanche lorsque la jonction base/collecteur est fortement polarisée en inverse. Ce courant d'avalanche a pour conséquence, lorsqu'il existe, d'augmenter le courant collecteur et de diminuer le courant de base. La figure I.14 présente les concentrations de porteurs minoritaires dans un transistor bipolaire en régime de fonctionnement normal.

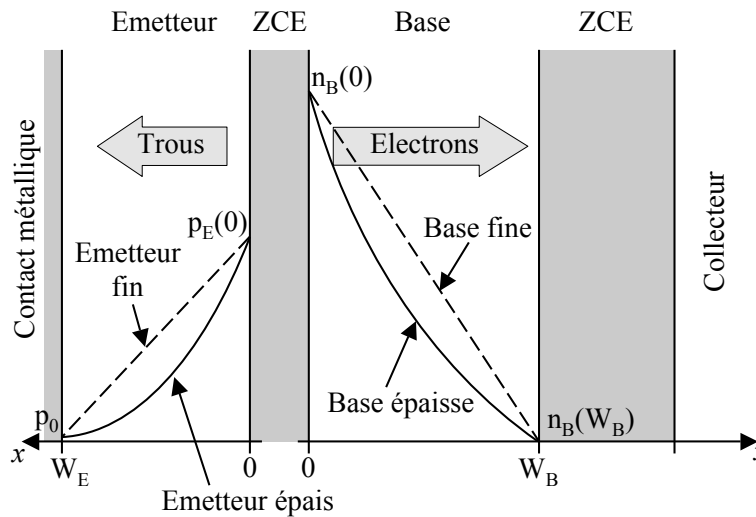


FIG. I.14 – Densités de porteurs dans un transistor bipolaire en régime de fonctionnement normal.

I.3.2 Expression des courants

Ce paragraphe vise à fournir une expression analytique des courants du TBH en régime statique, lorsque le composant est en mode de fonctionnement direct. Nous parlons uniquement dans cette partie du transistor bipolaire à hétérojonction Si/SiGe, l'apport du SiGe dans le transistor bipolaire ayant déjà été largement discuté dans différents travaux [Jouan01, Baudry01].

I.3.2.a Équations fondamentales

Nous rappelons brièvement dans cette partie les principales équations du transport dans les semiconducteurs. Loin de vouloir être exhaustif, le but est de poser les équations qui seront utiles pour les calculs des différents courants.

Dans un semiconducteur non dopé, la concentration des porteurs est égale à la concentration intrinsèque, donnée par l'équation suivante :

$$n_i^2(T) = N_C \cdot N_V \cdot \exp\left(\frac{-E_G}{kT}\right) \quad (\text{I.8})$$

Pour le silicium pur, la concentration intrinsèque dépend de la température selon l'équation suivante :

$$n_i(T) = 4.9 \times 10^{15} \left(\frac{m_{de}m_{dh}}{m_0^2}\right)^{3/4} M_C^{1/2} T^{3/2} \exp\left(\frac{-E_G}{kT}\right) \quad (\text{I.9})$$

où m_0 est la masse de l'électron libre, m_{de} et m_{dh} sont respectivement les masses effectives de densité d'état de la bande de conduction et de valence. M_C est le nombre de minima équivalents dans la bande de conduction, et vaut donc 6 dans le cas du silicium. Dans le cas d'un alliage de $\text{Si}_{1-x}\text{Ge}_x$, on peut exprimer la densité de porteurs intrinsèque en fonction de celle du Si pur :

$$(n_i^2)_{\text{SiGe}} = (n_i^2)_{\text{Si}} \frac{(N_C N_V)_{\text{SiGe}}}{(N_C N_V)_{\text{Si}}} \exp\left(\frac{\Delta E_G}{kT}\right) \quad (\text{I.10})$$

Dans l'équation précédente, la variation de bande interdite ΔE_G et le rapport $\gamma = \frac{(N_C N_V)_{\text{SiGe}}}{(N_C N_V)_{\text{Si}}}$ dépendent du taux de Germanium. En général $\gamma < 1$. Cependant, grâce à la réduction de band gap ΔE_G , le rapport $(n_i^2)_{\text{SiGe}} / (n_i^2)_{\text{Si}}$ est largement supérieur à 1.

Les constantes de diffusion des électrons et des trous sont liées à la mobilité par les relations d'Einstein (T est la température du réseau cristallin) :

$$D_n = \mu_n \frac{kT}{q} \quad (\text{I.11}) \quad \text{et} \quad D_p = \mu_p \frac{kT}{q} \quad (\text{I.12})$$

I.3.2.b Courant collecteur

Pour calculer le courant collecteur, nous reprendrons premièrement l'approche décrite par Kroemer [Kroemer85]. Le courant collecteur étant principalement un courant d'électrons ayant traversé la base, le principe est de calculer la concentration de porteurs dans la base neutre de la manière la plus générale possible. La seule hypothèse restrictive est qu'on ne considère pas de recombinaison en volume dans la base neutre, donc que la densité de courant est constante dans toute la base (à une dimension). Une autre méthode détaillée dans [Ashburn88] propose de calculer le courant J_n en chaque point de la base, afin de tenir compte de la recombinaison, mais suppose une base uniforme (dopage et taux de Germanium). Nous donnerons juste les grandes lignes de cette seconde approche.

Dans une optique de simplification des calculs, l'approche retenue pour calculer le courant I_C est une approche dérive-diffusion. On considère également que le courant de trous dans la base est négligeable comparé au courant d'électrons, et que les matériaux ne sont pas dégénérés. Bien qu'extrêmement simplificatrices, ces hypothèses offrent une bonne estimation des niveaux de courant dans le transistor.

Les abscisses sont celles définies dans la figure I.14 : $x = 0$ correspond à l'entrée dans la base neutre côté émetteur, et $x = W_B$ correspond à la sortie de la base neutre, côté collecteur.

I.3.2.b.i Cas général sans recombinaison

Dans le cas général, les courants d'électrons et de trous dépendent de leur niveaux de Fermi respectifs E_{F_n} et E_{F_p} selon les relations suivantes :

$$J_n = q n \mu_n \frac{d}{dx} \left(\frac{E_{F_n}}{q} \right) = n \mu_n \frac{dE_{F_n}}{dx} \quad (\text{I.13})$$

$$J_p = -q p \mu_p \frac{d}{dx} \left(\frac{E_{F_p}}{q} \right) = -p \mu_p \frac{dE_{F_p}}{dx} \quad (\text{I.14})$$

Compte-tenu de l'efficacité d'injection de la jonction émetteur/base et du fort gain en courant de nos transistors, nous pourrions considérer que le courant de trous dans la base est négligeable, le courant collecteur se réduisant à un courant d'électrons. Le courant de trous étant nul, on a donc $\frac{dE_{F_p}}{dx} = 0$, qui permet d'écrire :

$$J_n = n \mu_n \frac{d(E_{F_n} - E_{F_p})}{dx} \quad (\text{I.15})$$

Dans le cas d'un semiconducteur non dégénéré (niveau de Fermi suffisamment "loin" des bandes de conduction ou de valence), on peut écrire :

$$n \cdot p = n_i^2 \exp \left(\frac{E_{F_n} - E_{F_p}}{kT} \right) \quad (\text{I.16})$$

La dérivée logarithmique de I.16 donne :

$$\frac{d(E_{F_n} - E_{F_p})}{dx} = kT \cdot \frac{n_i^2}{n \cdot p} \cdot \frac{d}{dx} \left(\frac{n \cdot p}{n_i^2} \right) \quad (\text{I.17})$$

En réinjectant l'équation I.17 dans I.15 et en tenant compte de la relation d'Einstein (I.11), on peut écrire :

$$d \left(\frac{n \cdot p}{n_{iB}^2} \right) = \frac{J_n p}{q D_{nB} n_{iB}^2} dx \quad (\text{I.18})$$

On intègre l'équation précédente sur toute la largeur de la base neutre, entre $x = 0$ et $x = W_B$, ce qui donne :

$$\left. \frac{n \cdot p}{n_{iB}^2} \right|_{x=W_B} - \left. \frac{n \cdot p}{n_{iB}^2} \right|_{x=0} = \int_0^{W_B} \frac{J_n p}{q D_{nB} n_{iB}^2} dx \quad (\text{I.19})$$

En $x = W_B$, c'est à dire proche de la jonction base/collecteur, l'équation I.16 permet d'écrire (cf figure I.3(b)) :

$$\left. \frac{n \cdot p}{n_{iB}^2} \right|_{x=W_B} = \exp \frac{E_{F_n} - E_{F_p}}{kT} \Big|_{x=W_B} = \exp \frac{q V_{BC}}{kT} \quad (\text{I.20})$$

De même, en $x = 0$, c'est à dire proche de la jonction émetteur/base, on a :

$$\left. \frac{n \cdot p}{n_{iB}^2} \right|_{x=0} = \exp \frac{E_{F_n} - E_{F_p}}{kT} \Big|_{x=0} = \exp \frac{q V_{BE}}{kT} \quad (\text{I.21})$$

Compte-tenu du fait que la jonction E/B est polarisée en direct, et la jonction B/C en inverse, $\left. \frac{n \cdot p}{n_{iB}^2} \right|_{x=0}$ sera très grand devant $\left. \frac{n \cdot p}{n_{iB}^2} \right|_{x=W_B}$. Cela signifie que la concentration de porteurs est très faible à la jonction B/C, à comparer aux fortes concentrations d'électrons injectés du côté de la jonction E/B. L'équation I.19 devient ¹ :

¹Le signe «-» obtenu dans l'expression du courant collecteur est normal : il est dû à la convention de signes de départ. Dans la figure I.14, les électrons se déplacent de la gauche vers la droite, dans le sens *positif*, et le courant électrique est de sens contraire, donc *négatif*.

$$J_n = - \frac{q}{\int_0^{W_B} \frac{p(x)}{D_{nB}(x) n_{iB}^2(x)} dx} \exp \frac{q V_{BE}}{kT} \quad (I.22)$$

Afin de simplifier l'expression précédente, on introduit le *nombre de Gummel* de la base G_B , qui est un nombre dépendant uniquement des paramètres de la base du transistor.

$$G_B = \int_0^{W_B} \frac{p(x)}{D_{nB}(x) n_{iB}^2(x)} dx \quad (I.23)$$

L'expression obtenue précédemment est valable quel que soit le type de base. En effet, nous n'avons fait aucune hypothèse sur le terme n_{iB}^2 . On peut donc exprimer le courant collecteur I_C en fonction du nombre de Gummel de base et de la surface d'émetteur A_e :

$$I_C = \frac{q A_e}{G_B} \exp \frac{q V_{BE}}{kT} \quad (I.24)$$

Nous avons obtenu par le calcul une expression générale du courant collecteur prenant en compte toute la largeur de la base, en ne faisant aucune hypothèse sur le profil de dopage ou la variation du taux de Ge. Dans le cas d'une base uniforme, le nombre de Gummel devient :

$$G_B = \frac{N_{aB} W_B}{D_{nB} n_{iB}^2} \quad (I.25)$$

avec $p = N_{aB}$ dans l'hypothèse d'ionisation complète des espèces dopantes. On retrouve ainsi l'expression classique du nombre de Gummel.

I.3.2.b.ii Cas avec recombinaison dans une base uniforme

Nous avons exprimé le courant collecteur dans le cas le plus général possible, sans toutefois tenir compte des phénomènes de recombinaison en volume. Or il se peut qu'il y ait de la recombinaison dans la base neutre. Notamment si la base est épaisse, tous les électrons injectés à la jonction E/B ne parviendront pas à la jonction B/C. Nous pouvons prendre en compte cette recombinaison dans la base neutre en considérant que la base est homogène (dopage et taux de Ge). L'équation de continuité pour les électrons s'écrit :

$$\frac{\partial n}{\partial t} = G_n - U_n + \frac{1}{q} \operatorname{div} \vec{J}_n \quad (I.26)$$

En régime permanent, $\frac{\partial n}{\partial t} = 0$, de plus on considère qu'il n'y a pas de génération de porteurs, donc $G_n = 0$. Le taux de recombinaison U_n s'écrit :

$$U_n = \frac{n - n_{0B}}{\tau_n} \quad (I.27)$$

où n_{0B} est la concentration d'électrons à l'équilibre. Le courant d'électrons s'écrit :

$$\vec{J}_n = q D_n \overrightarrow{\operatorname{grad}} n + q n \mu_n \vec{E} \quad (I.28)$$

Le champ électrique $\vec{E}$ est considéré comme nul sur toute la base neutre. Compte tenu de ces hypothèses, la combinaison des trois équations I.26, I.27 et I.28 donne une équation différentielle du second ordre en $n(x)$ qui s'écrit sous la forme :

$$L_{nB}^2 \frac{d^2 n}{dx^2} - (n_B - n_{0B}) = 0 \quad (I.29)$$

dans laquelle on a posé $L_{nB} = \sqrt{D_{nB} \cdot \tau_{nB}}$. L'équation I.29 a pour solution une concentration d'électrons donnée en fonction de x :

$$n(x) = n_{0B} \left[1 + \frac{\sinh \frac{W_B - x}{L_{nB}} \left(\exp \frac{q V_{BE}}{kT} - 1 \right) - \sinh \frac{x}{L_{nB}}}{\sinh \frac{W_B}{L_{nB}}} \right] \quad (I.30)$$

La relation I.28 avec un champ électrique nul permet de calculer les densités de courant J_{nE} et J_{nC} , ainsi que la densité de courant de recombinaison J_{rB} :

$$J_{nE} = J_n(0) = \frac{-q n_{0B} D_{nB}}{L_{nB} \sinh \frac{W_B}{L_{nB}}} \left(\cosh \frac{W_B}{L_{nB}} \left(\exp \frac{q V_{BE}}{kT} - 1 \right) + 1 \right) \quad (I.31)$$

$$J_{nC} = J_n(W_B) = \frac{-q n_{0B} D_{nB}}{L_{nB} \sinh \frac{W_B}{L_{nB}}} \left(\exp \frac{q V_{BE}}{kT} - 1 + \cosh \frac{W_B}{L_{nB}} \right) \quad (I.32)$$

$$J_{rB} = J_{nE} - J_{nC} = \frac{-q n_{0B} D_{nB}}{L_{nB} \sinh \frac{W_B}{L_{nB}}} \left(\exp \frac{q V_{BE}}{kT} - 2 \right) \left(\cosh \frac{W_B}{L_{nB}} - 1 \right) \quad (I.33)$$

Compte-tenu du fait que la jonction émetteur/base est polarisée en direct on peut négliger 1 devant $\exp q V_{BE}/kT$.

I.3.2.b.iii Discussion

Dans chacune des expressions précédentes, nous pouvons constater que le facteur W_B/L_{nB} est toujours présent, nous pouvons donc faire une analyse en fonction du rapport entre la largeur de la base et de la longueur de diffusion des électrons.

Si $W_B/L_{nB} \ll 1$, cela signifie que la longueur de diffusion des électrons est très supérieure à la longueur de la base, le courant de recombinaison J_{rB} sera nul, et J_{nE} et J_{nC} seront égaux et constitueront la totalité du courant collecteur I_C . Les expressions I.31 et I.32 se simplifient fortement en tenant compte du fait que $n_{0B} = n_i^2/N_{aB}$ pour donner :

$$J_C = \frac{-q n_i^2 D_{nB}}{W_B N_{aB}} \exp \frac{q V_{BE}}{kT} \quad (I.34)$$

Le courant collecteur ne dépend plus de la longueur de diffusion des trous, mais uniquement de la largeur de la base (traits pointillés sur la figure I.14). On retrouve bien une expression identique à l'expression I.22 avec une base uniforme. Si au contraire la base est très épaisse, ou si les phénomènes de recombinaison sont prépondérants, on peut poser $W_B/L_{nB} \gg 1$, qui donne les expressions suivantes de J_{nE} et J_{nC} :

$$J_{nE} = \frac{-q n_i^2 D_{nB}}{L_{nB} N_{aB}} \exp \frac{q V_{BE}}{kT} \quad (I.35)$$

$$J_{nC} = \frac{-q n_i^2 D_{nB}}{L_{nB} N_{aB} \sinh \frac{W_B}{L_{nB}}} \exp \frac{q V_{BE}}{kT} \quad (I.36)$$

Il s'agit quasiment de la même expression que dans l'équation I.34, en ayant remplacé W_B par L_{nB} . Le courant d'émetteur I_{nE} ne dépend donc pas de la largeur de la base. Ce courant est d'autant plus faible que la base est épaisse, et correspond aux traits pleins de la figure I.14. Le courant de recombinaison se simplifie pour donner :

$$J_{rB} = \frac{-q n_{0B} D_{nB}}{L_{nB}} \exp \frac{q V_{BE}}{kT} \quad (I.37)$$

Si la largeur de la base est réellement plus grande que la longueur de diffusion des électrons, tous les électrons se recombinent en volume, et la densité d'électrons collectés tend vers 0.

Nous avons fait apparaître deux expressions possibles du courant collecteur, selon les hypothèses sur les paramètres technologiques de la base du transistor. En pratique, nous avons des bases non homogènes où le taux de Ge et le dopage varient ce qui implique le champ électrique n'est pas nul, et dans lesquelles peuvent apparaître des courants de recombinaison.

Dans toute la suite de l'étude, on considère que le courant de recombinaison est négligeable, excepté dans les cas où la recombinaison est souhaitée (voir chapitre IV.4 sur la recombinaison en base neutre).

I.3.2.c Courant de base

De même que pour le courant collecteur, nous allons exprimer le courant de base en fonction des paramètres technologiques de l'émetteur. L'émetteur est généralement réalisé en silicium (poly-silicium ré-épitaxié), mais des émetteurs en SiGe ont également été réalisés au cours de cette étude. Nous exprimerons donc le courant de base de la manière la plus générale possible.

Dans le cas classique (c'est à dire dans le cas d'une homojonction Si, ou dans le cas d'une hétérojonction graduelle Si/SiGe), l'injection des trous dans l'émetteur est principalement fixée par la polarisation de la jonction émetteur/base. Si l'hétérojonction émetteur/base est abrupte, il y aura une discontinuité en bande de valence, sans toutefois changer la hauteur de barrière totale vue par les trous, comme indiqué dans la figure I.15.

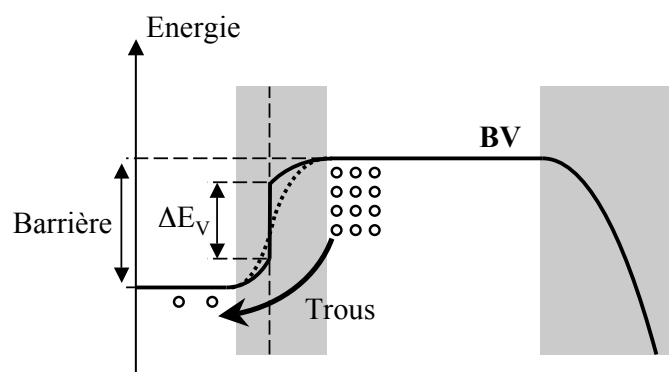


FIG. I.15 – Schéma de la discontinuité en bande de valence à la jonction émetteur/base. Malgré cette discontinuité, la hauteur de barrière totale vue par les trous reste constante.

La présence d'une discontinuité en bande de valence change fondamentalement le type d'injection des trous : d'une injection reposant sur les courants de dérive/diffusion, on passe à une injection de type thermoïonique, qui devient alors en théorie le facteur limitant le courant de trous. Le courant de base n'est alors plus contrôlé par la diffusion dans l'émetteur, mais par l'injection en elle-même.

Cependant, les expériences montrent que le courant d'émetteur est très peu dépendant du taux de Ge dans la base, ce qui signifie que l'hétérojonction émetteur/base n'est jamais infiniment abrupte. Il faut néanmoins rappeler que le modèle dérive-diffusion que nous avons retenu pour le calcul des courants suppose que les gradients de E_C et E_V sont faibles. Cette hypothèse tend à devenir caduque dans le cas des TBH modernes, présentant de fortes variations du taux de Germanium à l'entrée de la base. Les calculs présentés dans cette partie permettent toutefois d'obtenir une bonne vision des différents facteurs influençant le courant de base.

Pour calculer le courant de base, nous allons exprimer la concentration de porteurs de la manière suivante :

$$\vec{J}_p = -q D_{pE} \overrightarrow{\text{grad}} p + q p \mu_p \vec{E} \quad (\text{I.38})$$

On peut considérer que le champ électrique est nul dans l'émetteur. De même que pour les électrons, l'équation de continuité des trous s'exprime de la manière suivante :

$$\frac{\partial p}{\partial t} = G_p - U_p - \frac{1}{q} \text{div} \vec{J}_p \quad (\text{I.39})$$

On considère également qu'on se place en régime permanent ($\partial p / \partial t = 0$) et qu'il n'y a pas de génération externe de porteurs ($G_p = 0$). En tenant compte du taux de recombinaison des trous U_p qui s'exprime de la manière suivante :

$$U_{pE} = \frac{p_E - p_{0E}}{\tau_{pE}} \quad (\text{I.40})$$

où p_{0E} est la concentration de trous à l'équilibre. On en déduit l'équation différentielle vérifiée par les trous dans l'émetteur :

$$L_{pE}^2 \frac{d^2 p}{dx^2} - (p_E - p_{0E}) = 0 \quad (\text{I.41})$$

dans laquelle on a posé $L_{pE} = \sqrt{D_{pE} \cdot \tau_{pE}}$. L_{pE} est la longueur de diffusion des trous dans l'émetteur. La solution générale de l'équation I.41 est donnée par :

$$p_E(x) = p_{0E} \left[1 + \left(\exp \frac{q V_{BE}}{kT} - 1 \right) \frac{\sinh(W_E - x)/L_{pE}}{\sinh(W_E)/L_{pE}} \right] \quad (\text{I.42})$$

Le courant de base dans le cas général est donné par ² :

$$J_{pE} = -q D_{pE} \left. \frac{dp_E}{dx} \right|_{x=0} = + \frac{q D_{pE} p_{0E}}{L_{pE}} \coth \frac{W_E}{L_{pE}} \exp \frac{q V_{BE}}{kT} \quad (\text{I.43})$$

On trouve encore que le rapport W_E/L_{pE} intervient dans le calcul du courant de base. En effet, si l'épaisseur de l'émetteur est grande devant la longueur de diffusion des trous, le mécanisme de recombinaison en volume sera prépondérant. A l'inverse, si l'émetteur est très fin, la recombinaison des trous sur la surface du contact métallique contrôlera le courant de base. C'est cet effet qui est utilisé de manière volontaire pour augmenter artificiellement le courant de base (voir section IV.2 sur l'émetteur métallique).

Avec $W_E \gg L_{pE}$, on a :

$$J_{pE} = \frac{q D_{pE} p_{0E}}{L_{pE}} \exp \frac{q V_{BE}}{kT} \quad (\text{I.44})$$

Dans le cas $W_E \ll L_{pE}$, on obtient :

$$J_{pE} = \frac{q D_{pE} p_{0E}}{W_E} \exp \frac{q V_{BE}}{kT} \quad (\text{I.45})$$

De même que pour l'expression du courant collecteur, on définit un nombre de Gummel pour l'émetteur, qui s'exprime par :

$$G_E = \frac{N_{dE} L_{pE}}{D_{pE} n_i^2} \quad (\text{I.46})$$

²Là encore, attention à la convention de signes retenue : On trouve un courant se déplaçant dans le sens positif, mais le système d'axes a changé entre le calcul du courant collecteur et du courant de base (voir figure I.14)

avec $p_{0E} = n_i^2 / N_{dE}$, où N_{dE} est le dopage de l'émetteur. Dans le cas d'un émetteur très fin, on remplace L_{pE} par W_E . Cependant, compte-tenu des dimensions des dispositifs, la longueur de diffusion des trous est généralement inférieure à l'épaisseur de l'émetteur dans les transistors qui font l'objet de ce travail. On gardera donc l'expression de G_E faisant intervenir L_{pE} . De cette manière, le courant de base s'écrit :

$$I_{pE} = \frac{q A_e}{G_E} \exp \frac{q V_{BE}}{kT} \quad (\text{I.47})$$

Nous voyons apparaître les principaux paramètres technologiques influant sur le courant de base :

- Le dopage d'émetteur N_{dE} : Plus l'émetteur est dopé, moins l'injection de trous est efficace, et le courant de base diminue ;
- Le facteur n_i^2 : si on change de matériau pour réaliser l'émetteur (typiquement en remplaçant le Si par du SiGe), on change également le niveau d'injection des trous à l'entrée de l'émetteur ;
- La longueur de recombinaison des trous : Le paramètre important est la pente de la concentration de trous à l'origine, donc si les trous se recombinent plus rapidement (dans le volume ou sur une surface proche de la jonction), le courant de base sera également augmenté.

Au cours de l'optimisation du transistor bipolaire, nous disposons donc de plusieurs paramètres pour régler le courant de base à notre convenance. Les transistors bipolaires à hétérojonction SiGe ayant généralement des gains élevés, il est intéressant de dégrader volontairement le gain en augmentant I_B , et ainsi disposer d'un degré de liberté supplémentaire pour l'optimisation.

I.3.3 Différents gains en courant

Le rôle du transistor bipolaire est de se comporter comme un amplificateur de courant. Pour quantifier cette amplification, on définit le *gain en courant* du transistor, qui est tout simplement égal au courant en sortie divisé par le courant en entrée. Il existe plusieurs types de gains en courant, selon le type de connexion du transistor.

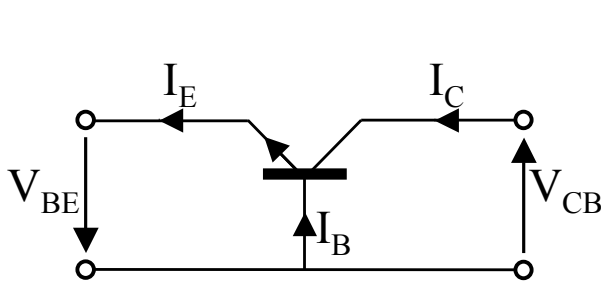


FIG. I.16 – Montage base commune.

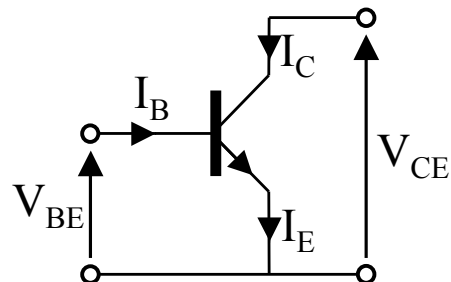


FIG. I.17 – Montage émetteur commun.

I.3.3.a Montage base commune

Le montage base commune se présente comme indiqué dans la figure I.16. Le gain en courant du composant dans cette configuration est noté α et est donc égal au courant en entrée divisé par le courant de sortie, soit :

$$\alpha = \frac{I_C}{I_E} \approx 1 \quad (\text{I.48})$$

Comme on l'a vu précédemment, le courant d'émetteur est légèrement plus grand que le courant collecteur, le gain en courant α est donc toujours inférieur à l'unité. Ce type de montage est utilisé dans les circuits lorsque le transistor doit fonctionner avec un certain type d'impédance d'entrée, ou lorsque le gain en courant n'est pas un paramètre primordial du circuit.

I.3.3.b Montage émetteur commun

Le montage dit *émetteur commun* présenté en figure I.17 est le plus courant pour le transistor bipolaire. En effet, c'est dans cette configuration que le transistor fournit le gain en courant le plus fort. Le gain en courant dans cette configuration est défini par :

$$\beta = \frac{I_C}{I_B} \quad (\text{I.49})$$

D'après les équations I.24 et I.47, le gain en courant est donc égal au rapport des nombres de Gummel de l'émetteur et de la base.

$$\beta = \frac{G_E}{G_B} = \frac{N_{dE} L_{pE} D_{nB} n_{iB}^2}{N_{aB} W_B D_{pE} n_{iE}^2} \quad (\text{I.50})$$

En tenant compte de l'équation I.10, on peut exprimer le gain de la manière suivante dans le cas d'un émetteur tout Si, avec une base en SiGe :

$$\beta = \frac{\gamma N_{dE} L_{pE} D_{nB}}{N_{aB} W_B D_{pE}} \exp \frac{\Delta E_G}{kT} \quad (\text{I.51})$$

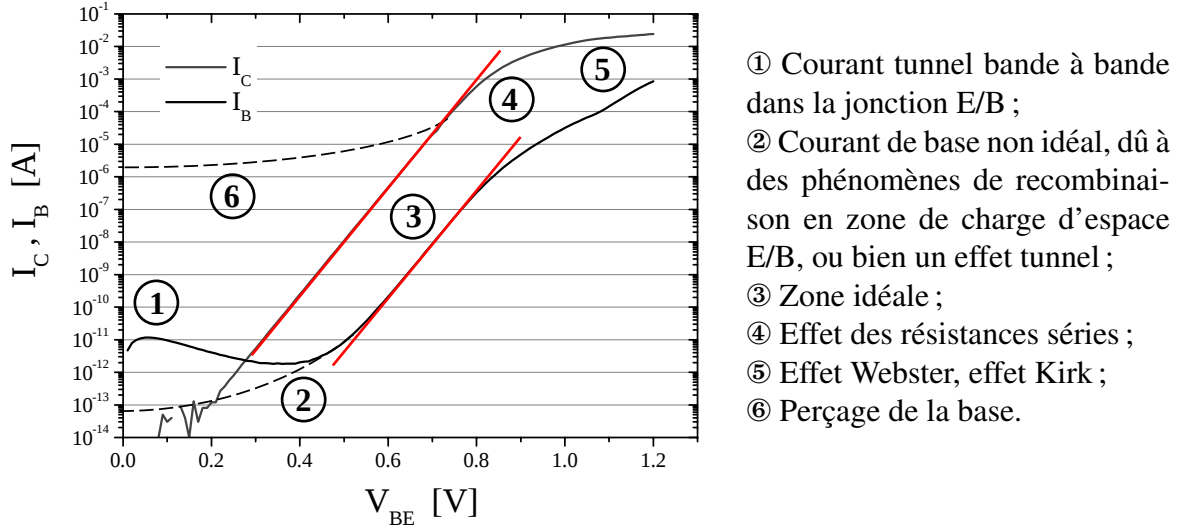
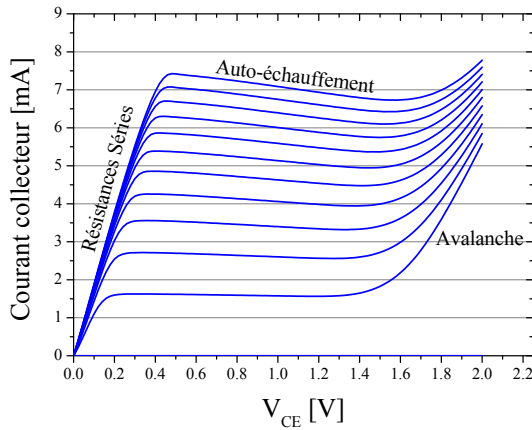
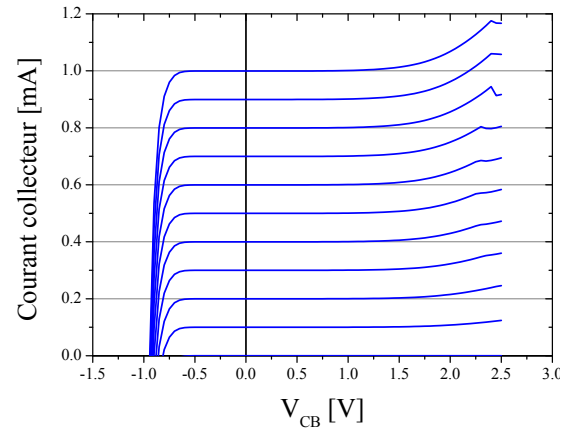
où γ est un paramètre d'ajustement proche de l'unité. Cette équation permet de faire apparaître les paramètres technologiques ayant une influence au premier ordre sur le gain en courant.

Ainsi, le rapport des dopages N_{dE}/N_{aB} et les épaisseurs respectives d'émetteur et de base ont une forte importance. Ces paramètres sont à prendre en compte dans la manière dont le transistor est réalisé. Ils ont toutefois une influence moindre que le terme $\exp \Delta E_G/kT$: Pour un taux de germanium moyen de 20%, le gain en courant sera multiplié par environ 1150 par rapport au transistor tout silicium. Nous voyons aussi que la température du composant a un fort impact sur ses caractéristiques statiques. Les transistors à base SiGe étudiés durant ce travail de thèse ont généralement des gains très forts (supérieurs à 1000). Nous verrons plus tard qu'un fort gain en courant est bénéfique pour les performances hyperfréquences, mais pénalise la tenue en tension du composant.

I.3.4 Effets du second ordre

Dans la partie précédente, nous avons exprimé les courants de base et de collecteur en fonction de la tension V_{BE} uniquement, n'indiquant aucune dépendance avec la tension V_{CB} . Or, la variation de I_C et I_B en fonction de la polarisation E/B n'est pas rigoureusement $\exp q V_{BE}/kT$, et la tension B/C a une forte influence sur le fonctionnement du composant.

De même, nous avons obtenu l'expression d'un gain en courant indépendant des tensions appliquées. En réalité, ces considérations ne sont vraies que dans une zone restreinte des caractéristiques du composant, que nous appellerons *zone idéale* (zone ③ de la figure I.18). En dehors de cette zone idéale, différents effets non linéaires se font sentir et influent sur les caractéristiques du composant.

FIG. I.18 – Exemple de courbe de Gummel. Dimensions du composant : $0.15 \times 3.6 \mu\text{m}^2$.FIG. I.19 – Caractéristiques de sortie en montage émetteur commun. I_B varie de 0 à 11 μA par pas de 1 μA . La diminution du gain aux forts courants est due à l'auto-échauffement.FIG. I.20 – Caractéristiques de sortie en montage base commune. I_E varie de 0 à 1 mA par pas de 0.1 mA

I.3.4.a Courants non-idéaux à faible polarisation

Nous pouvons voir sur la figure I.18 des courants fortement non idéaux, notamment en ce qui concerne le courant de base. Ces courants sont parfois visibles à faible polarisation, puis sont masqués, ou bien disparaissent, lorsque la tension V_{BE} augmente.

I.3.4.a.i Recombinaison en zone de charge d'espace

Dans le cas où la jonction E/B comporte des défauts, le courant de recombinaison I_{rG} peut ne plus être négligeable. Ce courant est en général exprimé de la manière suivante, où n prend une valeur comprise entre 1 et 2 [Ashburn88] :

$$I_{rG} \propto \exp \frac{q V_{BE}}{n k T} \quad (\text{I.52})$$

Le courant de recombinaison peut avoir une origine surfacique (les défauts sont répartis sur toute la surface du composant) ou bien périmétrique (défauts placés sur la périphérie de

la jonction E/B). Une analyse de ce courant en fonction de la taille du dispositif permet de remonter à la localisation du phénomène.

I.3.4.a.ii Courant tunnel bande à bande

Le phénomène à l'origine de la résistance différentielle négative visible sur le courant de base entre 0 et 0.4 V est un courant tunnel bande-à-bande et a été décrit dans [Lagarde06]. Ce courant apparaît lorsque la jonction E/B est très abrupte : Forts dopages, fort taux de Ge et épaisseur de jonction faible. Dans le cas de forts dopages, les zones de base et d'émetteur sont partiellement dégénérées, ce qui autorise le passage d'électrons directement de bande-à-bande par effet tunnel, comme expliqué sur la figure I.21.

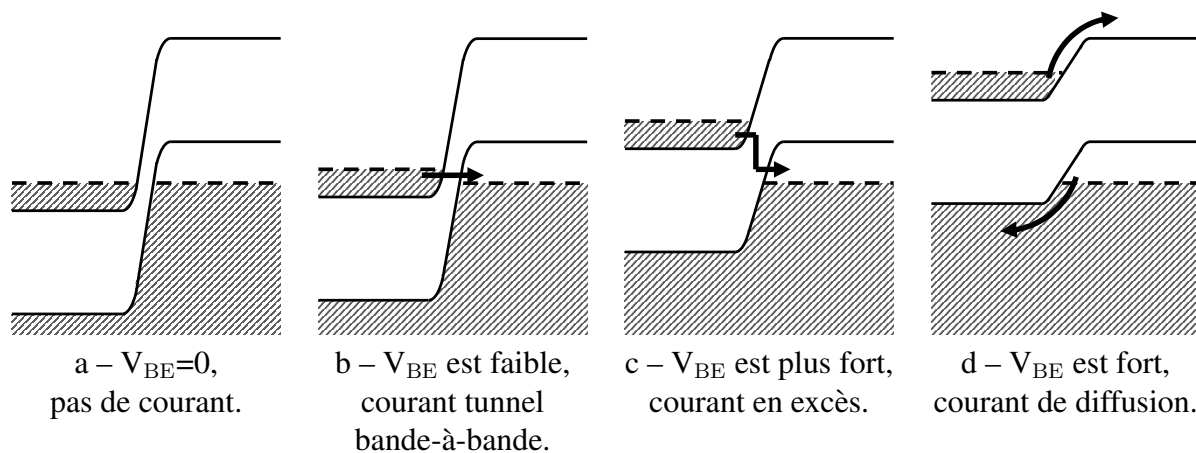


FIG. I.21 – Explication de l'effet tunnel à faible polarisation [Lagarde06].

Ce courant tunnel peut se décomposer en plusieurs phases successives :

- Lorsque V_{BE} est nul, il n'y a pas de courant (figure I.21-a).
- Un courant tunnel bande-à-bande, figure I.21-b.
- Un courant en excès (figure I.21-c). Ce courant en excès est un courant de recombinaison assisté par pièges. Les porteurs utilisent des états d'énergie localisés dans la bande interdite. Le trajet des porteurs peut être compliqué, en fonction de la position des pièges dans la zone de charge d'espace.
- Le courant thermique classique du composant (figure I.21-d). Le courant tunnel n'est plus possible, et le courant en excès est noyé sous le courant thermique classique des porteurs.

La figure I.22 résume les différents courants impliqués dans l'effet tunnel à faible polarisation, et donne également le courant résultant. Ce courant tunnel bande-à-bande est en réalité assisté par les phonons, car les extrema de bande de conduction et de valence ne sont pas situés au même moment $k = 0$. La bande de conduction du Si présente une vallée supérieure en $k = 0$, les phonons permettent aux électrons de passer dans cette vallée supérieure et ensuite transiter par effet tunnel vers la base (I.23).

I.3.4.b Effet Early

L'expression du courant collecteur en fonction de la tension V_{BE} fait intervenir en paramètre la largeur de la base neutre W_B , qui n'est pas indépendante des tensions appliquées sur le composant. La modulation de la largeur de base en fonction des tensions V_{BE} ou V_{CB} a pour conséquence une déviation du courant collecteur par rapport à l'idéalité. C'est ce qu'on appelle l'*effet Early*.

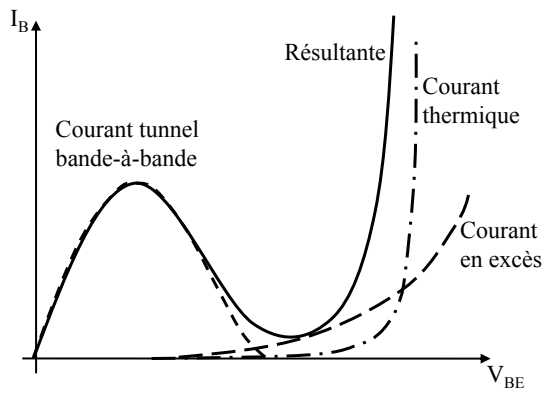


FIG. I.22 – Courants contribuant à l'effet tunnel à faible polarisation, et résultante du courant.

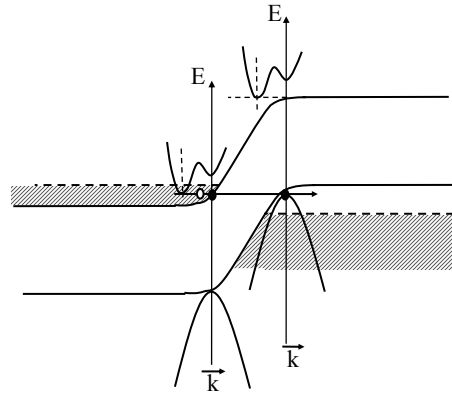


FIG. I.23 – Effet tunnel assisté par phonons : Les électrons passent dans la vallée supérieure en $k = 0$, puis franchissent la barrière par effet tunnel direct.

Il existe deux effets Early, selon que la modulation de l'épaisseur de base provient de la jonction émetteur/base ou base/collecteur.

I.3.4.b.i Effet Early direct

L'effet Early direct est dû à une variation de la largeur de base W_B avec la tension V_{CB} , comme illustré dans la figure I.24.

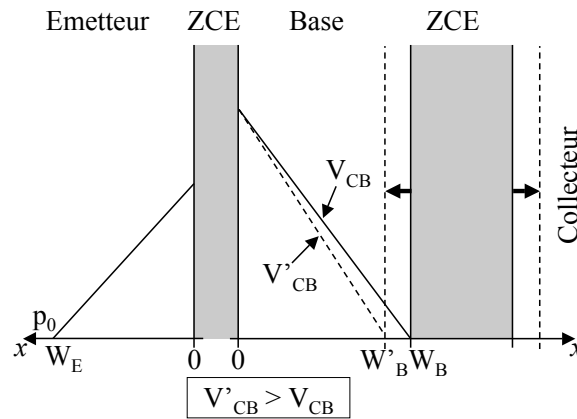


FIG. I.24 – Effet Early direct : La variation d'épaisseur de zone de charge d'espace B/C est à l'origine d'une variation de courant collecteur.

Lorsque la tension V_{CB} augmente, la zone de charge d'espace s'étend plus dans la base, et l'épaisseur de la base neutre diminue. Le gradient d'électrons étant plus fort, le courant collecteur augmente. L'effet Early direct est caractérisé par la tension du même nom, noté V_{Af} ("f" pour "forward"), dont la définition est donnée sur la figure I.25.

Le dopage de base étant généralement bien supérieur au dopage du collecteur, les tensions d'Early directes sont généralement élevées (supérieures à une centaine de Volt).

I.3.4.b.ii Effet Early inverse

L'effet Early inverse est dû à une variation de W_B avec la tension V_{BE} , comme expliqué figure I.26. Comme la jonction E/B contrôle l'injection des porteurs dans la base, cet effet se

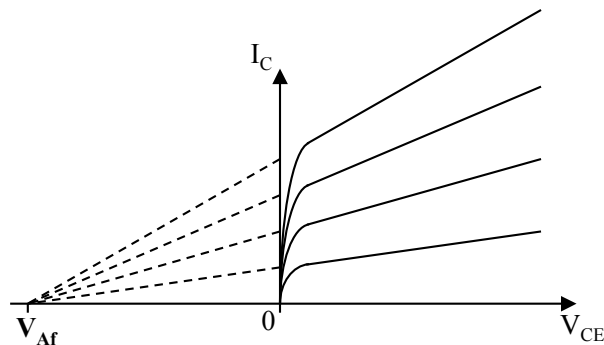


FIG. I.25 – Caractéristiques de sortie d’un transistor bipolaire mettant en évidence la tension d’Early directe.

caractérise par un écart à l’idéalité : la caractéristique de I_C en fonction de V_{BE} ne se comportera pas tout à fait comme $\exp q V_{BE}/kT$. L’effet Early inverse est surtout visible sur la caractéristique du gain en courant. Le rapport de dopage très important existant entre la base et l’émetteur joue en défaveur de l’effet Early inverse.

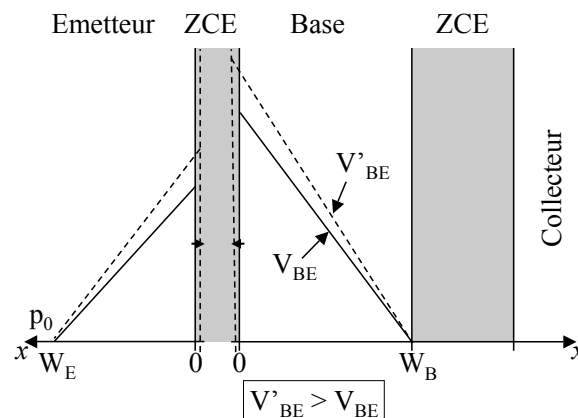


FIG. I.26 – Explication de l’effet Early inverse par une modulation de l’épaisseur de base neutre par la polarisation V_{BE} .

Cet effet Early inverse est très pénalisant pour le transistor bipolaire à hétérojonction, car il a pour effet d’augmenter la largeur de base neutre, ce qui fait chuter le gain en courant : il en découle une perte importante des performances hyperfréquences. L’effet Early inverse est lui aussi quantifié par une tension d’Early inverse, notée V_{Ar} (“r” pour “reverse”).

I.3.4.b.iii Influence du profil graduel de SiGe

L’effet Early (direct et inverse) est modifié par la présence d’un taux de Germanium graduel sur la largeur de la base : En effet, le courant collecteur est proportionnel au nombre de Gummel de la base, qui dépend des paramètres technologiques sur toute la largeur de la base neutre. Étudions donc les variations de ce nombre de Gummel avec la largeur de base, dans le cas d’un profil de Ge plat et graduel, et leur influence sur les deux tensions d’Early.

- Dans le cas de l’effet Early direct, on voit sur la figure I.27 que lorsque le profil de Ge est graduel, la dépendance de G_B avec la largeur de base est plus faible que lorsque la concentration de Ge est constante. Le profil rétrograde améliore donc la tension d’Early.
- L’effet Early inverse est en revanche dégradé par le profil graduel de Germanium (figure I.28) : En effet, on voit que la dépendance du nombre de Gummel est beaucoup plus im-

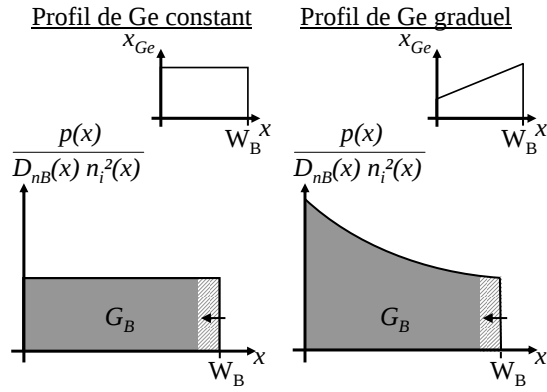


FIG. I.27 – Influence du profil Germanium sur la tension d'Early directe. La variation du nombre de Gummel avec l'épaisseur de la base est plus faible lorsque le profil de Ge est graduel, ce qui améliore la tension d'Early.

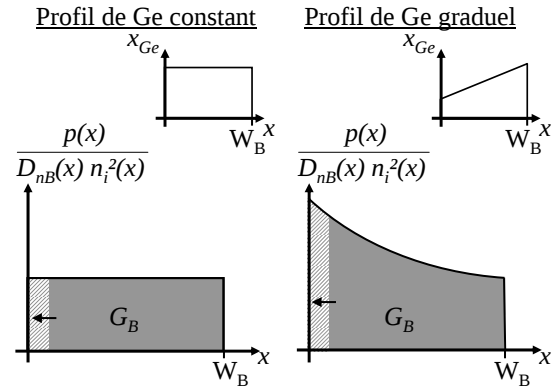


FIG. I.28 – Influence du profil graduel de Germanium sur la tension d'Early inverse. La dépendance du nombre de Gummel avec V_{BE} est plus forte dans le cas d'un profil graduel qu'un profil plat.

portante que pour un profil plat. Ceci implique d'apporter une attention toute particulière à la position de la jonction E/B. S'assurer que la limite de ZCE côté base soit dans une zone où le profil de Ge est relativement plat permet de limiter l'effet Early inverse.

Notons que lorsque le profil de Ge est plat, les tensions V_{Af} et V_{Ar} sont théoriquement les mêmes que dans le cas du transistor bipolaire tout silicium. Par exemple, dans le cas d'un transistor à base graduelle 20-30 % de Ge, les tensions d'Early directe et inverse sont respectivement de l'ordre de 200 V et 2 V, contre 100 V et 8 V respectivement pour une base à taux de Ge constant de 25 %.

I.3.4.c Tensions d'avalanche

Sur la figure I.19 apparaissent des tensions que l'on appelle tensions de claquage du composant. Il s'agit en fait de phénomènes d'avalanche qui interviennent lorsque les jonctions (principalement B/C) sont fortement polarisées.

I.3.4.c.i Définition et caractérisation de l'avalanche

L'avalanche est un phénomène de multiplication des porteurs soumis à un fort champ électrique. Lorsque la polarisation en inverse de la jonction devient très forte, un nombre important de paires électrons/trous sont créées par un phénomène d'ionisation par choc. Ces porteurs ainsi libérés dérivent grâce au fort champ électrique et peuvent, par un phénomène d'ionisation par impact générer au passage d'autres paires électrons-trous : C'est le phénomène d'avalanche. Le phénomène est décrit par le schéma présenté en figure I.29

L'avalanche est caractérisée par le facteur de multiplication M , qui mesure l'augmentation du courant dans la zone de charge d'espace pour un type de porteurs. Par exemple, pour un courant d'électrons I_n qui rentre dans la jonction, le courant de l'autre côté de la jonction sera $M I_n$. En pratique, on utilise plutôt le terme $M - 1$, plus représentatif du mécanisme d'augmentation du courant dans la jonction. Comme on le voit sur la figure I.29, pour un électron injecté côté base, M électrons seront récupérés côté collecteur. Les courants de base et collecteur varient d'une même quantité, égale à $(M - 1)I_{nC}$.

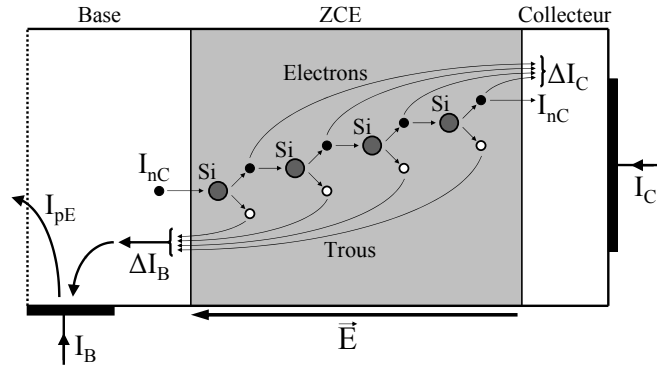


FIG. I.29 – Mécanisme d’avalanche dans la jonction base/collecteur polarisée en inverse.

I.3.4.c.ii BV_{CBO}

La tension BV_{CBO} est la tension d’avalanche de la jonction B/C. Elle est mesurée avec l’émetteur ouvert, c’est à dire non connecté. La caractéristique courant-tension de la diode B/C fait apparaître la tension de claquage BV_{CBO} , comme indiqué sur la figure I.30. La jonction B/C étant fortement dissymétrique, la tension d’avalanche BV_{CBO} dépend principalement des paramètres du collecteur :

$$BV_{CBO} = \frac{\varepsilon_0 \varepsilon_r E_{crit}^2}{2q N_{dC}} \quad (I.53)$$

Le champ critique E_{crit} dépend du matériau et du dopage. Dans le silicium dopé N, la valeur du champ critique est typiquement de l’ordre de 6.10^5 V.cm^{-1} . Lorsque la tension V_{CB} augmente, le facteur d’avalanche dans la ZCE B/C augmente également. Un calcul empirique du facteur de multiplication donne, lorsque V_{CB} n’est pas trop grand :

$$M - 1 = \exp \left(\frac{V_{CB}}{BVC} \right)^{M_f} - 1 \approx \left(\frac{V_{CB}}{BVC} \right)^{M_f} \quad (I.54)$$

où M_f est un paramètre d’ajustement, que l’on appelle facteur d’avalanche, de valeur comprise entre 3 et 6 pour les transistors faisant l’objet de cette étude. La tension BVC n’est pas la tension d’avalanche BV_{CBO} , mais un paramètre de modèle de la jonction B/C, de valeur typique 3.5 V pour nos composants.

I.3.4.c.iii BV_{EBO}

La tension BV_{EBO} est la tension d’avalanche de la jonction émetteur/base, avec collecteur ouvert. Compte-tenu des dopages plus forts de cette jonction, BV_{EBO} est généralement plus faible que la tension BV_{CBO} . Les tensions typiques obtenues sont entre 1 et 3 V. La caractéristique de la diode émetteur-base est donnée figure I.31.

I.3.4.c.iv BV_{CEO}

La tension BV_{CEO} est un paramètre important de l’optimisation du transistor bipolaire. Elle donne une mesure de l’avalanche dans la jonction base/collecteur d’un transistor bipolaire en fonctionnement, contrairement à la tension BV_{CBO} qui est une caractéristique d’une jonction PN uniquement. A cause de l’effet transistor, l’avalanche apparaît beaucoup plus tôt dans la jonction B/C lorsque le transistor est en fonctionnement : Pour les transistors bipolaires rapides, cette tension est de l’ordre de 1.5 V, et elle dépend généralement du point de fonctionnement.

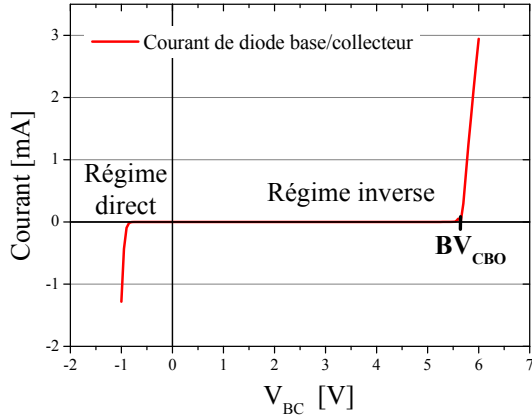


FIG. I.30 – Caractéristique courant-tension de la diode base-collecteur (émetteur ouvert). La tension de claquage BV_{CBO} est ici de 5.7 V.

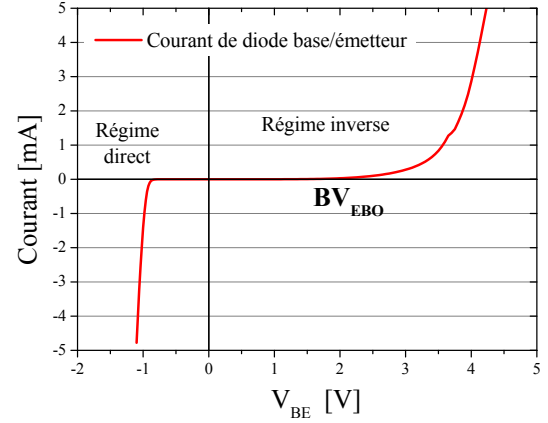


FIG. I.31 – Caractéristique courant-tension de la diode émetteur-base (collecteur ouvert). La tension de claquage BV_{EBO} est ici comprise entre 1 et 3 V.

Cette faible valeur de BV_{CEO} est un des principaux points demandant une grande vigilance lors de l'optimisation du TBH : En effet, la tension maximale applicable à un transistor connecté en montage émetteur commun (voir figure I.17) doit être inférieure à la tension de claquage collecteur-émetteur, ce qui limite fortement les tensions d'alimentation.

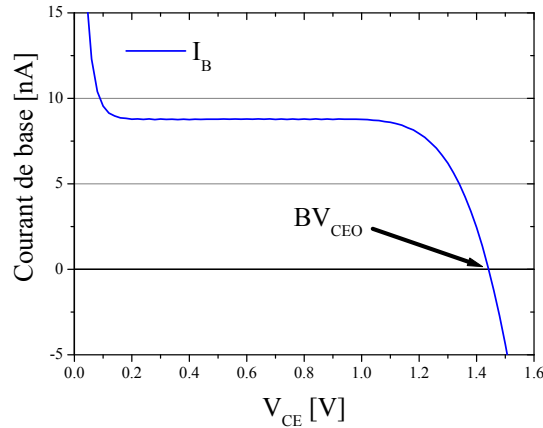


FIG. I.32 – Définition de la tension de claquage BV_{CEO} .

On voit sur la figure I.29 que le courant d'avalanche a tendance à diminuer le courant I_B . La tension BV_{CEO} est par définition la tension à laquelle le courant de base s'annule et change de signe, à cause de l'augmentation du courant d'avalanche, comme signalé sur la figure I.32. En tenant compte du fait que $I_{pE} = I_{nC}/\beta$, cette condition est remplie lorsque :

$$\frac{I_{nC}}{\beta} = (M - 1)I_{nC}$$

où β est le gain en courant du transistor. Compte-tenu de l'équation I.53, la tension de claquage BV_{CEO} s'écrit :

$$BV_{CEO} = V_{BE} + \frac{BVC}{\beta^{1/M_f}} \quad (I.55)$$

La tension de claquage BV_{CEO} diminue lorsque le gain du transistor augmente. Il y a donc un compromis à faire entre fort gain et forte tension d'avalanche. Cependant, si la tension

BV_{CEO} est un facteur de mérite simple du transistor bipolaire, la tension V_{CE} maximale disponible peut lui être bien supérieure. Selon [Rickelt01], lorsque V_{CE} est supérieur à BV_{CEO} , le courant I_B est négatif, ce qui est source d'instabilités dans le composant à cause de l'ionisation par impact dans la jonction B/C. Il est néanmoins possible, en modifiant les conditions de polarisation du composant (par exemple en fixant I_E constant ou bien V_{BE} constant), de retarder l'apparition de ces instabilités, et donc de disposer d'une plus grande plage de variation de V_{CE} .

De plus, la condition $I_B = 0$ (base ouverte) n'est pas une solution réaliste, car la base est toujours chargée par une résistance non nulle. [Kraft05] introduit la notion de BV_{CER} pour tenir compte de cet effet. BV_{CER} dépend de la valeur de la résistance de base et des conditions de polarisation, et peut être largement supérieure à BV_{CEO} . Ainsi, la donnée des valeurs de BV_{CEO} et BV_{CBO} n'est pas suffisante pour caractériser complètement les phénomènes d'avalanche dans la ZCE base/collecteur.

I.3.4.d Effet de haute injection dans la base (effet Webster)

Une des hypothèses ayant mené à l'équation I.22 est que la concentration d'électrons injectés est faible devant le dopage de base N_{aB} . Or à de très fortes polarisations, la quantité d'électrons peut devenir supérieure à ce dopage. Dans ce cas, la concentration de porteurs majoritaires augmente pour maintenir la neutralité électrique, comme indiqué sur la figure I.33.

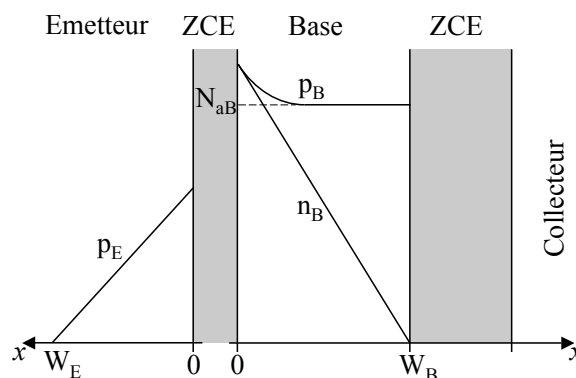


FIG. I.33 – Effet de haute injection dans la base : lorsque le courant augmente, la concentration de trous augmente pour compenser les électrons en excès.

Cet effet est équivalent à une augmentation du dopage apparent de la base. Comme le nombre de Gummel de la base est inversement proportionnel au dopage, le gain diminue lorsque la tension augmente. Un calcul rigoureux permet de démontrer que le courant collecteur est de la forme :

$$I_C \propto \exp \frac{q V_{BE}}{2kT} \quad (I.56)$$

L'effet de haute injection est visible sur les caractéristiques électriques dans la partie ⑤ de la figure I.18. Cependant, l'augmentation des dopages de base contribue à retarder l'apparition de l'effet Webster.

I.3.4.e Effet Kirk

Une autre cause de diminution du gain à forte polarisation est l'effet Kirk [Kirk62] qui est cette fois-ci lié au dopage du collecteur.

L'effet Kirk est le principal effet limitant des performances hyperfréquences du composant. Il est dû au fait que les électrons injectés dans le collecteur se déplacent selon un mécanisme de dérive. Compte-tenu du fort champ électrique qui règne à la jonction B/C, on peut considérer que les électrons ont atteint leur vitesse limite v_{sat} . Dès lors, leur temps de transit n'est plus négligeable dans la jonction B/C, et lorsque la densité de courant devient importante, la concentration d'électrons injectés atteint puis dépasse le dopage collecteur. Il en découle une extension importante de la zone de charge d'espace dans le collecteur. La base s'élargit de manière brutale, provoquant une forte diminution du gain.

On peut calculer le courant auquel apparaît l'effet Kirk selon la formule ci-dessous :

$$J_{C\ Kirk} = q N_{dC} v_{sat} \quad (I.57)$$

Nous voyons donc que le seuil de l'effet Kirk dépend du dopage collecteur. Afin d'augmenter les performances fréquentielles du composant, il faut donc augmenter le dopage collecteur, ce qui dégrade de manière non négligeable la tenue en tension du composant (voir chapitre I.3.4.c).

I.3.4.f Effet de barrière

L'effet de barrière à l'hétérojonction (HBE pour *Heterojunction Barrier Effect*) est un effet qui apparaît à forte polarisation dans les TBH lorsque le collecteur a un profil très abrupt. A faible et moyenne injection, la discontinuité de la bande de valence est masquée par la courbure des bandes, et a donc peu d'impact sur les caractéristiques électriques. Cependant, à plus forte polarisation, la diminution du champ électrique due à l'extension de la zone de charge d'espace B/C rend cette barrière apparente, ce qui bloque l'injection des trous dans le collecteur. Cette accumulation de trous a pour effet de modifier la courbure de la bande de conduction, qui va à son tour bloquer l'injection des électrons. En conséquence, on observe une augmentation importante de la charge stockée dans la base, donc une diminution très forte de f_T .

L'effet de barrière est dû à l'hétérojonction Si/SiGe, et ne peut pas être évité dans le cas du TBH. Cependant, il existe plusieurs possibilités pour réduire son effet :

- Réaliser un profil de Ge graduel de la base vers le collecteur ;
- Déplacer le plan de l'hétérojonction plus profondément dans le collecteur ;
- Augmenter le dopage collecteur, ce qui permet de retarder l'apparition de l'effet de barrière vers les forts courants.

L'effet Kirk et l'effet de barrière sont issus de deux phénomènes physiques différents. Cependant, le fait qu'ils interviennent tous deux à fort courant a induit une confusion de ces effets dans la plupart des modèles existant [Liang02]. Il est donc important de noter que, selon la position de l'hétérojonction SiGe/Si à la jonction base-collecteur, un des deux effets est généralement favorisé par rapport à l'autre. Du fait des forts dopages collecteurs utilisés dans les TBH étudiés durant ce travail, le principal facteur limitant est l'effet Kirk.

I.3.5 Effet des forts dopages

Pour les composants ayant des performances dynamiques à l'état de l'art, les niveaux de dopage, notamment d'émetteur et de base deviennent très importants. Ces forts dopages ne sont pas sans effet sur les caractéristiques du composant, et dégradent les performances électriques. Ces dégradations sont au nombre de trois : la réduction de bande interdite (ou BGN pour *Band Gap Narrowing*), la dégradation de la mobilité des porteurs et l'effet tunnel bande-à-bande décrit dans la partie I.3.4.a.ii.

I.3.5.a Réduction de bande interdite

Le dopage d'un matériau consiste en l'insertion d'un niveau énergétique dans la bande interdite, proche de la bande de conduction pour un dopage type N (niveau donneur), proche de la bande de valence pour un dopage P (niveau accepteur). Lorsque le dopage est faible, les atomes dopants sont suffisamment éloignés dans le réseau cristallin pour que leur niveau énergétique reste discret. Dans ce cas, les dopants n'ont pas d'effet sur la bande interdite du matériau, qui garde sa valeur de référence E_{G0} .

Cependant, lorsque le dopage du matériau devient très fort, les impuretés sont très proches les unes des autres, de telle sorte que les fonctions d'onde de leurs électrons se recouvrent, ce qui modifie la structure de bande du matériau en créant des niveaux d'énergie supplémentaires. Lorsque ce dopage est très important, ces états d'énergie supplémentaires viennent chevaucher la bande de valence ou de conduction, ce qui va avoir pour effet de diminuer l'énergie de bande interdite, comme expliqué dans la figure I.34.

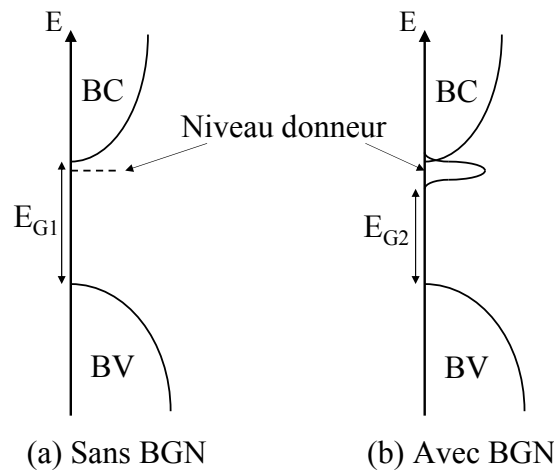


FIG. I.34 – Effet de réduction de bande interdite sous l'effet de forts dopage : a – Sans BGN, le gap du matériau est le même que le matériau intrinsèque ; b – Avec BGN, le gap apparent du matériau est réduit.

L'effet néfaste du BGN réside dans le fait que dans un TBH optimisé, l'émetteur est plus dopé que la base, donc la réduction de bande interdite sera plus importante dans l'émetteur que dans la base. Le bénéfice de l'hétérojonction Si/SiGe à la jonction E/B est donc atténué par les forts dopages.

[Klaassen92c] propose un modèle de BGN dépendant du niveau de dopage. Il permet d'exprimer la réduction de gap ΔE_g due au dopage. Notons que le BGN ne devient significatif qu'au delà de $10^{17} \text{at/cm}^{-3}$. L'expression du BGN s'écrit en fonction du dopage N :

$$\Delta E_G = 6.92 \cdot \left[\ln \left(\frac{N}{1.3 \cdot 10^{17}} \right) + \sqrt{\left(\ln \left(\frac{N}{1.3 \cdot 10^{17}} \right) \right)^2 + 0.5} \right] \quad [\text{meV}] \quad (\text{I.58})$$

I.3.5.b Modification des équations des courants

La réduction de bande interdite due au dopage a un effet sur I_C et I_B , ainsi que sur le gain en courant du transistor. On appelle ΔE_{gE} et ΔE_{gB} respectivement les réductions de bande interdite de l'émetteur et de la base. Généralement on a $\Delta E_{gE} > \Delta E_{gB}$.

La réduction de la bande interdite d'un matériau modifie la concentration intrinsèque des porteurs. Dans l'émetteur, la concentration intrinsèque devient :

$$n_{iE}^2 = n_{i0}^2 \cdot \exp \frac{\Delta E_{gE}}{kT} \quad (\text{I.59})$$

où n_{i0}^2 est la concentration intrinsèque du matériau constituant l'émetteur. De même pour la base :

$$n_{iB}^2 = n_{i0}^2 \cdot \exp \frac{\Delta E_{gB}}{kT} \quad (\text{I.60})$$

où n_{i0}^2 est la concentration intrinsèque du SiGe constituant la base du transistor. On note ΔE_g la différence entre la réduction de bande interdite de l'émetteur et de la base.

$$\Delta E_g = \Delta E_{gE} - \Delta E_{gB} \quad (\text{I.61})$$

Les équations du courant collecteur I.24 et du courant de base I.47 deviennent respectivement :

$$I_C = \frac{q A_e}{G_B} \exp \frac{\Delta E_{gB}}{kT} \exp \frac{q V_{BE}}{kT} \quad (\text{I.62})$$

$$I_B = \frac{q A_e}{G_E} \exp \frac{\Delta E_{gE}}{kT} \exp \frac{q V_{BE}}{kT} \quad (\text{I.63})$$

Dans le cas d'un profil de Germanium constant et uniformément dopé, l'expression du courant collecteur donnée par l'équation I.62 devient :

$$I_C = \frac{q n_{i0}^2 A_e D_{nB}}{N_{aB} W_B} \exp \frac{\Delta E_G + \Delta E_{gB}}{kT} \exp \frac{q V_{BE}}{kT} \quad (\text{I.64})$$

ΔE_G dans l'expression I.64 est la réduction de bande interdite due à la présence de Germanium dans la base. Le phénomène de réduction de bande interdite est favorable à une augmentation du courant de base. D'après les expressions I.63 et I.64, l'expression du gain en courant donnée par l'équation I.51 devient :

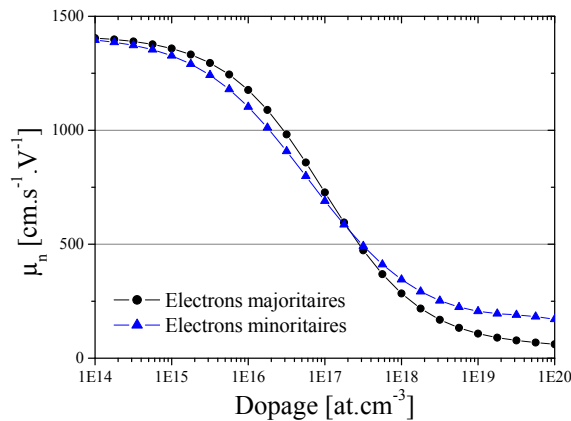
$$\beta = \frac{\gamma N_{dE} W_E D_{nB}}{N_{aB} W_B D_{pE}} \exp \frac{\Delta E_G - \Delta E_g}{kT} \quad (\text{I.65})$$

L'émetteur étant généralement plus dopé que la base, $\Delta E_g > 0$, le BGN a tendance à diminuer l'effet d'hétérojonction du Germanium. Cependant, la condition $\Delta E_G - \Delta E_g > 0$ est toujours vérifiée dans le cas d'un TBH rapide. Le gain en courant augmente donc fortement lorsque la température diminue. Ce n'est pas le cas d'un transistor bipolaire à homojonction, où $\Delta E_G = 0$. L'émetteur devant être plus fortement dopé que la base, le gain décroît lorsque la température diminue.

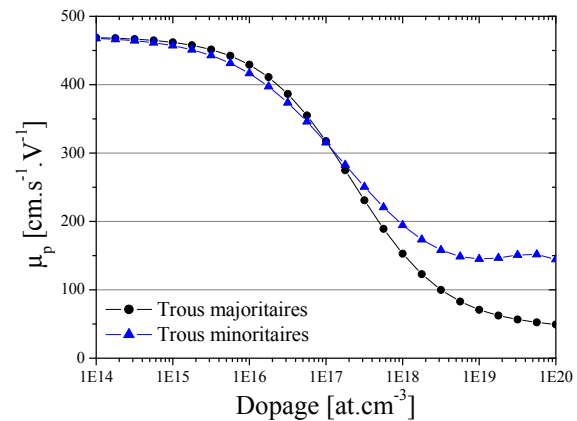
I.3.5.c Mobilité des porteurs

Il n'existe pas de modèle complet de la mobilité des porteurs dans le matériau Silicium-Germanium dopé, nous baserons notre analyse sur les modèles existants du Silicium.

Klaassen [Klaassen92a, Klaassen92b] propose un modèle de mobilité unifié qui prend en compte le type (N ou P) et le dopage du matériau. Ce modèle fait la distinction entre porteurs majoritaires et minoritaires. En considérant que le transport est avant tout diffusif, la mobilité des porteurs dépend de la concentration en impuretés dans le matériau. La figure I.35 présente l'évolution de la mobilité des porteurs à 300 K, en fonction du niveau de dopage.

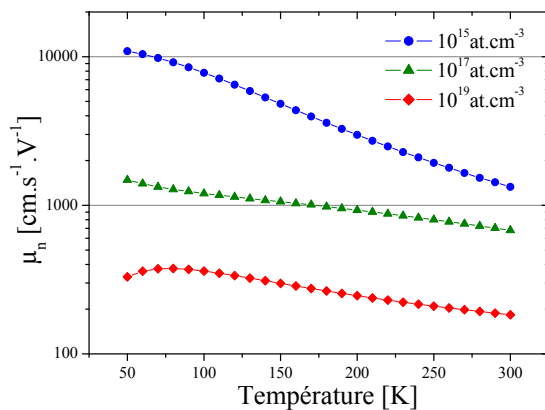


a – Mobilité des électrons

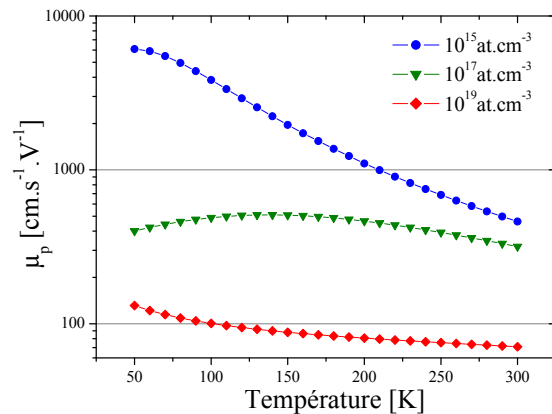


b – Mobilité des trous

FIG. I.35 – Evolution de la mobilité des porteurs dans le Silicium en fonction du dopage, à 300 K, dans le cas où les porteurs sont minoritaires ou majoritaires.



a – Mobilité des électrons en fonction de la température



b – Mobilité des trous en fonction de la température

FIG. I.36 – Évolution de la mobilité des porteurs dans le Silicium en fonction de la température pour différents niveaux de dopage.

On constate qu'avec l'augmentation du dopage, la mobilité est fortement réduite. Une mobilité réduite des électrons et des trous implique donc des niveaux de courant plus faibles, ainsi que des temps de transit plus longs.

La mobilité des porteurs dépend également de la température du matériau. Dans une gamme 50–400 K, la mobilité s'améliore lorsque la température diminue, comme indiqué sur la figure I.36. L'augmentation très forte de la mobilité des électrons à basse température est une des raisons expliquant l'amélioration des performances du composant aux températures cryogéniques (cf. chapitre V.3 sur les basses températures).

I.4 Fonctionnement dynamique du transistor bipolaire

Dans les parties précédentes, nous avons étudié le fonctionnement statique du transistor bipolaire. Or le transistor bipolaire est un composant qui présente des performances dynamiques très élevées, ce qui en fait le composant de choix pour les applications RF. Dans cette partie, nous allons détailler les différents facteurs de mérites servant à évaluer le fonctionnement dynamique du TBH.

A travers l'expression des temps de transit des porteurs et des retards dûs aux différentes capacités du TBH, nous introduisons les fréquences f_T et f_{MAX} , respectivement fréquences de transition des gains en courant et en puissance.

I.4.1 Temps de transit des porteurs

En présence de polarisation, dans un TBH fonctionnant en régime normal, on voit l'apparition de charges en excès, par rapport aux concentrations de repos n_0 et p_0 . On note Q_F la charge en excès des électrons et des trous. Pour conserver la neutralité globale de la structure, la charge négative due aux électrons en excès compense la charge positive créée par les trous.

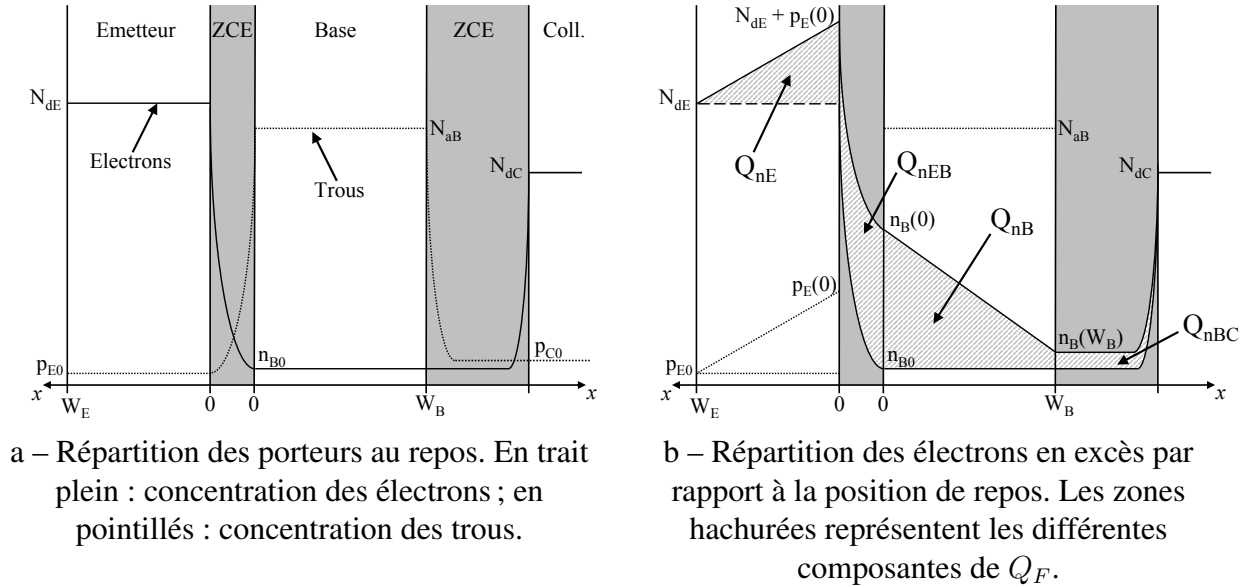


FIG. I.37 – Répartition des porteurs (a) à l'équilibre, et (b) sous polarisation normale directe.

La figure I.37 représente les concentrations d'électrons dans le transistor au repos (a) et en régime direct (b). La charge Q_F créée par les électrons en excès est représentée par les zones hachurées. Ces charges sont réparties de la manière suivante pour les électrons :

$$-Q_F = Q_{nE} + Q_{nEB} + Q_{nB} + Q_{nBC} \quad (I.66)$$

De même pour les trous :

$$+Q_F = Q_{pE} + Q_{pEB} + Q_{pB} + |Q_{nBC}| \quad (I.67)$$

Les électrons sont renouvelés par le courant collecteur (courant d'électrons minoritaires dans la base), alors que les trous sont renouvelés par le courant de base (courant de trous mi-

noritaires dans l'émetteur). De ce fait, à cause du fort gain en courant du transistor, la charge $-Q_F$ se renouvelle plus vite que $+Q_F$.

Le temps de transit des porteurs τ_F est défini comme le temps nécessaire pour renouveler la charge $-Q_F$ due aux électrons. Par définition, on a donc :

$$\tau_F = \frac{Q_F}{I_C} \quad (\text{I.68})$$

τ_F se décompose de la manière suivante :

$$\tau_F = \tau_E + \tau_{EB} + \tau_B + \tau_{BC} = \frac{|Q_{nE}|}{I_C} + \frac{|Q_{nEB}|}{I_C} + \frac{|Q_{nB}|}{I_C} + \frac{|Q_{nBC}|}{I_C} \quad (\text{I.69})$$

où τ_E , τ_{EB} , τ_B et τ_{BC} sont les retards relatifs respectivement à l'émetteur, à la jonction E/B, à la base neutre et à la jonction B/C. Pour chaque partie du transistor, on peut associer charge stockée et temps de transit. On note qu'il n'y a pas d'expression du temps de transit pour le collecteur neutre : Nous considérons qu'il s'agit d'une zone uniquement résistive, et l'influence du collecteur est rétablie dans l'expression de f_T et f_{MAX} .

I.4.1.a Temps de transit émetteur

τ_E est le temps de transit émetteur. Il représente le temps de renouvellement des charges créées par les électrons en excès dans l'émetteur. Cette charge Q_E est identique à la charge des trous injectés dans l'émetteur, pour maintenir la neutralité de la structure. D'après la figure I.37b, on peut exprimer Q_{nE} :

$$Q_{nE} = q A_e \times (\text{aire du triangle}) = q A_e \frac{1}{2} W_E p_{E0} \exp \frac{q V_{BE}}{kT} \quad (\text{I.70})$$

D'après l'expression I.24 du courant collecteur, l'expression du temps de transit émetteur devient :

$$\tau_E = \frac{W_E W_B N_{aB}}{2 D_{nB} N_{dE}} \exp \frac{-\Delta E_G}{kT} \quad (\text{I.71})$$

Or, en utilisant l'expression I.51 du gain en courant :

$$\begin{aligned} \tau_E &= \frac{W_E W_B N_{aB}}{2 D_{nB} N_{dE}} \exp \frac{-\Delta E_G}{kT} \times \frac{W_E D_{pE}}{W_E D_{pE}} \\ \tau_E &= \frac{1}{\beta} \frac{W_E^2}{2 D_{pE}} \end{aligned} \quad (\text{I.72})$$

Cette expression du temps de transit émetteur est en fait équivalente à une évacuation des charges en excès par le courant émetteur. Compte-tenu du terme en $1/\beta$, le temps de transit émetteur est très faible, voire négligeable, devant τ_B .

L'expression I.72 peut être complétée pour prendre en compte la vitesse de recombinaison S_E au contact émetteur, dans le cas d'un émetteur fin :

$$\tau_E = \frac{1}{\beta} \left(\frac{W_E^2}{2 D_{pE}} + \frac{W_E}{S_E} \right) \quad (\text{I.73})$$

On voit que l'augmentation du gain en courant a pour effet de diminuer le temps de transit émetteur τ_E . D'après [Agarwal05], l'expression du temps de transit émetteur peut se transformer en :

$$\tau_E = \frac{q n_{iE}^2 W_E}{J_C 2N_{dE}} \left(1 + \frac{1}{1 + S_E W_E / D_{pE}} \right) \quad (\text{I.74})$$

Cette expression a pour mérite de ne plus faire apparaître le gain en courant, mais le courant collecteur lui-même, mettant ainsi l'accent sur la nécessité d'augmenter la densité de courant collecteur pour réduire les temps de transit. Le courant de base n'apparaît donc plus dans l'expression I.74, ce qui est normal, car les charges en excès sont évacuées par le courant collecteur. Cette dernière expression a également pour intérêt de faire apparaître les paramètres de l'émetteur jouant au premier ordre sur le temps de transit τ_E .

I.4.1.b Temps de transit dans la jonction émetteur/base

τ_{EB} est le temps de transit des électrons dans la zone de charge d'espace E/B. L'expression de τ_{EB} est donnée par :

$$\tau_{EB} = \frac{Q_{nEB}}{I_C} \quad (\text{I.75})$$

D'après [Roulston90], la charge contenue dans la jonction E/B polarisée en direct vaut :

$$Q_{nEB} = q A_e W_{EB} \exp \frac{q V_{BE}}{2kT} \quad (\text{I.76})$$

Le courant étant proportionnel à $\exp \frac{q V_{BE}}{kT}$, la dépendance du temps de transit en fonction de V_{BE} est donc :

$$\tau_{EB} \propto \exp -\frac{q V_{BE}}{2kT} \quad (\text{I.77})$$

On remarque que lorsque V_{BE} augmente, τ_{EB} diminue très rapidement. De plus, la jonction émetteur/base étant polarisée en direct, l'extension de la ZCE est très faible, on peut généralement considérer cette composante comme négligeable au premier ordre. D'après [Kerr75], le temps de transit dans la ZCE E/B dépend du profil de dopage de l'émetteur : Un dopage plus abrupt permet de stocker moins de charges en excès, ce qui diminue le temps de transit τ_{EB} . La tendance actuelle d'augmenter les niveaux de dopage et de réduire les dimensions verticales du composant accentue encore la réduction de τ_{EB} . Cependant la charge stockée dans la jonction E/B a une contribution non nulle au temps de transit total, notamment lorsqu'il existe une couche peu dopée entre l'émetteur et la base afin d'éviter l'effet tunnel [Agarwal05].

I.4.1.c Temps de transit dans la base

τ_B est le temps de transit dans la base. Il correspond au temps de renouvellement des électrons par le courant I_C . Compte-tenu de la figure I.37b, la charge Q_{nB} vaut :

$$Q_{nB} = q A_e \times (\text{aire du triangle}) = q A_e \frac{1}{2} W_B n_{B0} \exp \frac{q V_{BE}}{kT} \quad (\text{I.78})$$

En tenant compte de l'équation I.24, on obtient l'expression suivante pour τ_B :

$$\tau_B = \frac{W_B^2}{2D_{nB}} \quad (\text{I.79})$$

Dans le cas de profils de base particuliers, notamment avec le pseudo-champ accélérateur créé par le profil de Ge graduel, le temps de transit de base est réduit par rapport à l'expression I.79. Le facteur 2 est remplacé par un facteur η supérieur à 2.

$$\tau_B = \frac{W_B^2}{\eta D_{nB}} \quad (\text{I.80})$$

Il est également possible de tenir compte de la vitesse limite des porteurs dans la base. Dans ce cas, l'expression de τ_B est changée pour devenir :

$$\tau_B = \frac{W_B^2}{\eta D_{nB}} + \frac{W_B}{V_{sat}} \quad (\text{I.81})$$

D'après I.81, plus l'épaisseur de base se réduit, plus le terme W_B/V_{sat} devient influent. Toujours à cause de la réduction des dimensions, l'effet du champ accélérateur devient de plus en plus réduit sur l'amélioration des performances. Le temps de transit de base est une des composantes principales du temps de transit τ_F des TBH SiGe. Il s'agit donc d'un des principaux termes limitant la montée en fréquence.

I.4.1.d Temps de transit dans la jonction base/collecteur

La dernière composante de τ_F est le temps de transit B/C. Son calcul est délicat, car le courant qui circule dans la jonction modifie la répartition des charges et la conformation de la zone de charge d'espace. Cette jonction étant polarisée en inverse, et le champ électrique étant très fort, on peut considérer que les porteurs sont en vitesse de saturation v_{sat} sur la majorité de la largeur de la ZCE. La densité de charges mobiles due aux électrons dans la ZCE est donc (en valeur absolue) :

$$Q_e = J_C / v_{sat} \quad (\text{I.82})$$

Cependant, il ne s'agit pas de la charge en excès dans la jonction B/C, car cette charge s'ajoute à la densité de charge négative créée par les dopants accepteurs ionisés côté base et se retranche à la densité de charge positive créée par les dopants donneurs ionisés côté collecteur. Le courant qui circule dans la jonction va avoir pour effet de modifier légèrement la conformation de la ZCE : Afin de conserver la neutralité électrique globale de la jonction, la ZCE côté base se réduit légèrement et côté collecteur elle s'étend légèrement plus.

Pour calculer la densité de charge en excès dans cette ZCE B/C, il faut considérer que le courant qui circule dans la jonction est dû à l'effet transistor, et est indépendant de la polarisation appliquée sur la jonction [Meyer87, Roulston90] (en dehors des effets d'avalanche). La charge totale stockée dans la ZCE (charges fixes + charge due aux porteurs mobiles) ne dépend que de V_{CB} .

La figure I.38 représente les densités de charge et le champ électrique dans deux cas distincts, avec courant électrique et sans courant. Lorsque le courant circule, la zone de déplétion B/C se décale légèrement vers le collecteur, la surface de la zone hachurée représentant la différence de potentiel V_{CB} est la même qu'en l'absence de courant. En considérant que la jonction est fortement dissymétrique, toute la chute de tension est répercutée sur le côté collecteur. On appelle W_{C1} et W_{C2} l'extension de la ZCE côté collecteur, respectivement au repos et en présence de courant (voir figure I.38).

$$V_{CB} \approx \frac{q N_{dC} W_{C1}^2}{2 \varepsilon_{Si}} = \frac{(q N_{dC} - Q_e) W_{C2}^2}{2 \varepsilon_{Si}} \quad (\text{I.83})$$

On en déduit :

$$\frac{W_{C2}}{W_{C1}} = \sqrt{\frac{N_{dC}}{N_{dC} - Q_e/q}} = \sqrt{\frac{1}{1 - Q_e/q N_{dC}}} \quad (\text{I.84})$$

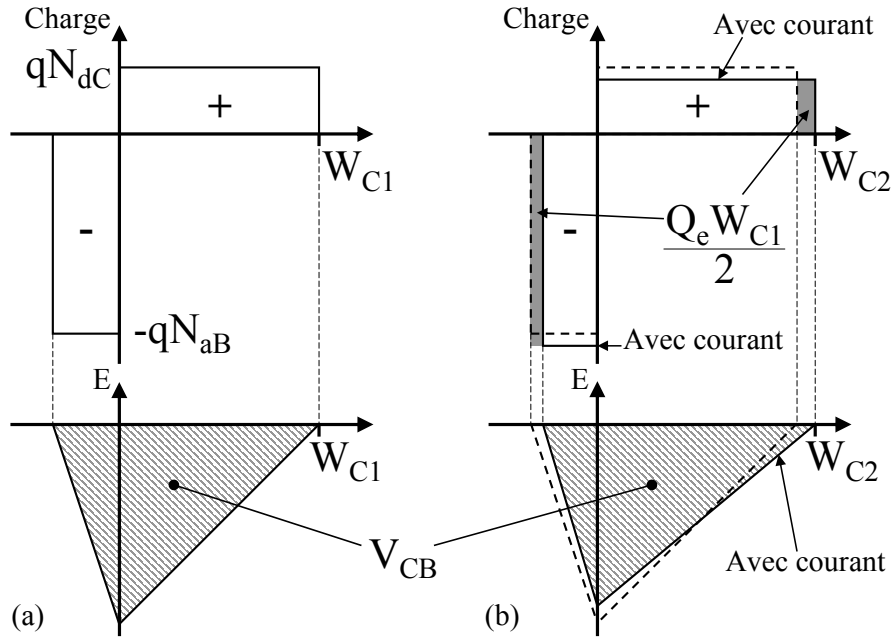


FIG. I.38 – Répartition des charges dans la jonction B/C (a) au repos, (b) avec passage de courant.

En faible injection, on considère que $Q_e \ll qN_{dC}$, un développement limité au premier ordre donne :

$$\frac{W_{C2}}{W_{C1}} = \left(1 - \frac{Q_e}{qN_{dC}}\right)^{-\frac{1}{2}} \approx 1 + \frac{Q_e}{2qN_{dC}} \quad (\text{I.85})$$

Soit :

$$W_{C2} - W_{C1} = \frac{Q_e W_{C1}}{2qN_{dC}} \quad (\text{I.86})$$

Entre une situation sans circulation de courant et une situation où le courant est établi, la charge totale due aux électrons (charges mobiles) a varié d'une quantité :

$$\Delta Q_{\text{mobile}} = -Q_e W_{C2} \approx -Q_e W_{C1}$$

La charge totale due aux atomes dopants ionisés (charges fixes) a quant à elle varié d'une quantité :

$$\Delta Q_{\text{fixe}} = qN_{dC} (W_{C2} - W_{C1}) = \frac{Q_e W_{C1}}{2}$$

La variation de charge totale de la ZCE côté collecteur vaut donc :

$$\Delta Q_{\text{total}} = \Delta Q_{\text{fixe}} + \Delta Q_{\text{mobile}} = -\frac{Q_e W_{C1}}{2} \quad (\text{I.87})$$

L'établissement du courant d'électrons a provoqué l'apparition d'une charge *négative* additionnelle $-Q_e W_{C1}$, compensée pour moitié par une variation de l'épaisseur de la ZCE côté collecteur (zone grisée sur la figure I.38). De plus, pour conserver la neutralité électrique globale de la jonction, une réduction de l'extension de la ZCE côté base a dû générer la charge *positive* opposée.

La moitié de la charge d'électrons qui n'a pas été compensée par les dopants du collecteur constitue donc la charge des électrons en excès. Cette charge vaut donc :

$$Q_{nBC} = \frac{Q_e W_{C1}}{2} = \frac{Q_e W_{BC}}{2} \quad (I.88)$$

D'après les équations I.82 et I.88, on en déduit l'expression de τ_{BC} :

$$\tau_{BC} = \frac{Q_{nBC}}{J_C} = \frac{W_{BC}}{2 v_{sat}} \quad (I.89)$$

Cette composante du temps de transit a une influence croissante dans τ_F avec la réduction des épaisseurs de base. On voit que pour réduire τ_{BC} , il faut diminuer l'extension de la ZCE B/C, ce qui peut être effectué en augmentant le dopage collecteur, mais cela va diminuer les tensions d'avalanche. Là encore, un compromis est à réaliser entre temps de transit faible et forte tension de claquage.

I.4.2 Capacités de jonction

Les temps de transit des porteurs dans la structure ne sont pas les seuls éléments limitant les performances du composant. Il existe également deux capacités de jonction, C_{BE} et C_{BC} , respectivement capacité de jonction émetteur/base et base/collecteur. Ces capacités de jonction sont dues aux charges stockées de chaque côté des zones déplétées, comme indiqué dans la figure I.39 :

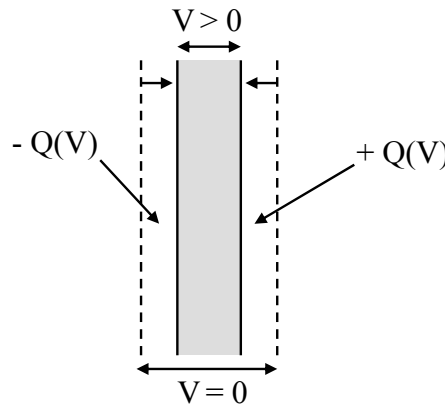


FIG. I.39 – Charges stockées dans une jonction en fonctionnement, deux charges opposées se font face de part et d'autre de la jonction.

La variation de la charge stockée avec la tension V appliquée sur la jonction permet de définir une capacité de jonction en régime petit signal :

$$C_j = \frac{dQ}{dV} \quad (I.90)$$

Dans notre cas, il y a deux jonctions, la jonction émetteur/base et base/collecteur correspondant à deux capacités dépendant des points de polarisation V_{BE} et V_{BC} .

I.4.3 Fréquences de transition f_T et f_{MAX}

I.4.3.a Définition et expression de f_T

f_T est la fréquence de transition du gain en courant petit signal. Lorsque le transistor fonctionne en mode émetteur commun, il amplifie les variations de I_B , créant des variations de I_C .

On appelle i_B et i_C ces courants petit signal, et on définit le gain en courant dynamique h_{fe} , ou h_{21} , comme :

$$h_{fe} = \left| \frac{i_C}{i_B} \right|_{V_{CE} \text{ constant}} \quad (\text{I.91})$$

Attention à ne pas confondre les gains β et h_{fe} : L'un est le rapport des courants I_C et I_B , l'autre est le gain dynamique autour d'un point de fonctionnement donné.

En régime sinusoïdal forcé, h_{fe} suit une loi du premier ordre, donnant le diagramme de Bode représenté sur la figure I.40. La fréquence f_T est par définition la fréquence pour laquelle $|h_{fe}| = 1$.

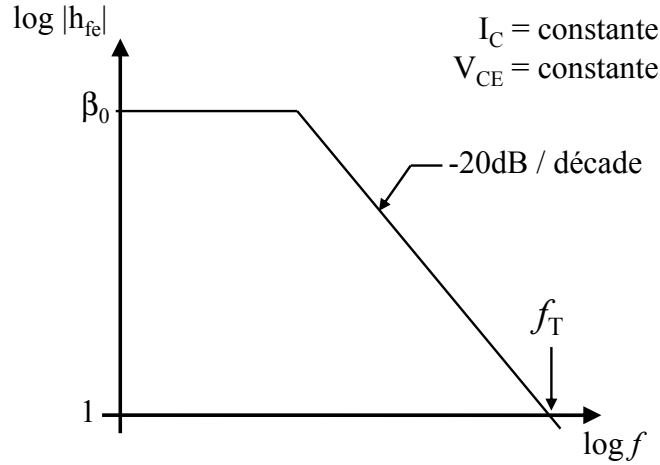


FIG. I.40 – Définition de la fréquence de transition f_T .

Pour exprimer la fréquence de transition f_T , il faut tenir compte des différents éléments induisant un retard dans le transport des électrons. En plus des deux capacités de jonction émetteur/base et base/collecteur, il faut tenir compte de la résistance d'émetteur et de collecteur, ainsi que de la capacité de diffusion C_{Diff} ajoutée pour modéliser le retard dû à la charge Q_F en excès calculée plus haut. Le temps de transit total vaut (équation I.68) :

$$\tau_F = \frac{dQ_F}{dI_C} = \frac{dQ_F}{dV_{BE}} \frac{dV_{BE}}{dI_C} \quad (\text{I.92})$$

On pose $C_{Diff} = dQ_F/dV_{BE}$. La charge en excès est contrôlée par la tension V_{BE} uniquement, elle se comporte donc comme si C_{Diff} est en parallèle sur la jonction E/B. Comme on est en régime petit signal, la capacité C_{Diff} peut être considéré comme constante, ce qui n'est pas le cas en réalité (il n'y a pas proportionnalité entre Q_F et V_{BE}).

Le terme dV_{BE}/dI_C est égal à l'inverse de la transconductance g_m , que l'on peut exprimer de la manière suivante grâce à l'équation I.24 :

$$g_m = \frac{dI_C}{dV_{BE}} = \frac{q I_C}{kT} \quad (\text{I.93})$$

En régime petit signal, la variation du courant collecteur i_C s'écrit donc $i_C = g_m \cdot v_{BE}$. Le courant i_C dépend du point de polarisation V_{BE} choisi.

On peut également exprimer C_{Diff} de la manière suivante :

$$C_{Diff} = \tau_F \cdot g_m = \tau_F \cdot \frac{q I_C}{kT} \quad (\text{I.94})$$

Compte-tenu des différents éléments pouvant induire un retard pour le fonctionnement dynamique du transistor bipolaire, on en déduit le schéma électrique équivalent petit signal présenté sur la figure I.41. Ce schéma est équivalent au modèle en π [Roulston90, Ashburn03].

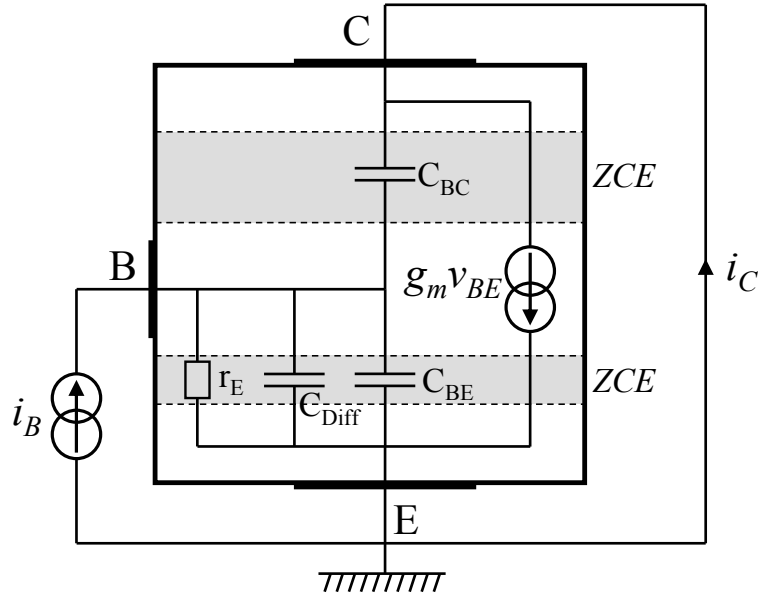


FIG. I.41 – Schéma équivalent petit signal du transistor bipolaire pour le calcul de f_T .

Le courant collecteur petit signal s'exprime de la manière suivante (régime sinusoïdal forcé) :

$$i_C = g_m \cdot v_{BE} - j\omega C_{BC} \cdot v_{BE} \quad (\text{I.95})$$

En considérant que le courant de base sert à charger les deux capacités C_{BC} et C_{BE} , plus la capacité de diffusion C_{Diff} , on en déduit :

$$i_B = \frac{v_{BE}}{r_E} + j\omega (C_{BC} + C_{BE} + C_{Diff}) \cdot v_{BE} \quad (\text{I.96})$$

On en déduit l'expression du gain en courant :

$$h_{fe} = \frac{g_m \cdot r_E - j\omega r_E C_{BC}}{1 + j\omega r_E (C_{BC} + C_{BE} + C_{Diff})} \quad (\text{I.97})$$

Le gain à $\omega = 0$ du composant vaut donc $g_m \cdot r_E$. La transconductance de nos composants est très forte, on peut donc négliger $j\omega C_{BC}$ devant g_m . Aux fortes fréquences, le module du gain en courant devient :

$$|h_{fe}| = \frac{g_m}{\omega (C_{BC} + C_{BE} + C_{Diff})} \quad (\text{I.98})$$

On peut ainsi exprimer la fréquence f_T à laquelle le gain vaut 1 :

$$f_T = \frac{\omega}{2\pi} = \frac{1}{2\pi \left(\frac{kT}{qI_C} (C_{BE} + C_{BC} + C_{Diff}) \right)} \quad (\text{I.99})$$

D'après les équations I.93 et I.94, on en déduit l'expression générale de f_T :

$$f_T = \frac{1}{2\pi \left(\tau_F + \frac{kT}{qI_C} (C_{BE} + C_{BC}) \right)} \quad (\text{I.100})$$

Néanmoins, l'expression I.100 est théorique et incomplète. Il faut en effet tenir compte des retards dû aux couples résistance d'accès / capacités. Ces retards sont au nombre de deux : $R_E C_{BC}$ et $R_C C_{BC}$. L'expression complète de f_T devient alors :

$$f_T = \frac{1}{2\pi \left(\tau_F + \frac{kT}{qI_C} (C_{BE} + C_{BC}) + (R_E + R_C) C_{BC} \right)} \quad (\text{I.101})$$

L'expression ci-dessus est valable aux faibles courants, la fréquence de transition augmente avec le courant collecteur, pour plafonner à une valeur maximale en théorie égale à $1/\tau_{FC}$, soit $1/(\tau_F + (R_E + R_C)C_{BC})$. Cependant, à cause des différents effets non idéaux détaillés au chapitre I.3.4, f_T n'atteint jamais cette valeur théorique, comme indiqué sur la figure I.42. Les différents retards capacitifs limitent la fréquence f_T , et le brutal écroulement de la fréquence de transition aux forts courants est dû aux effets de forte injection (effet Kirk principalement).

La figure I.43 présente l'évolution de $\tau_{EC} = 1/2\pi f_T$ en fonction de $1/I_C$. Le temps $\tau_{EC \text{ Min}}$ est le temps de transit minimal. Il est possible d'établir une régression linéaire sur la zone de dépendance de f_T avec I_C . La pente de la régression donne une indication sur les capacités de jonction C_{BE} et C_{BC} , l'ordonnée à l'origine permet de calculer $\tau_{EC,0}$, composé de τ_F et du terme $(R_E + R_C) \cdot C_{BC}$. Pour pouvoir remonter à la valeur de τ_F seul, il est nécessaire de réaliser des extractions plus précises de paramètres de modèle du composant (voir chapitre V.3).

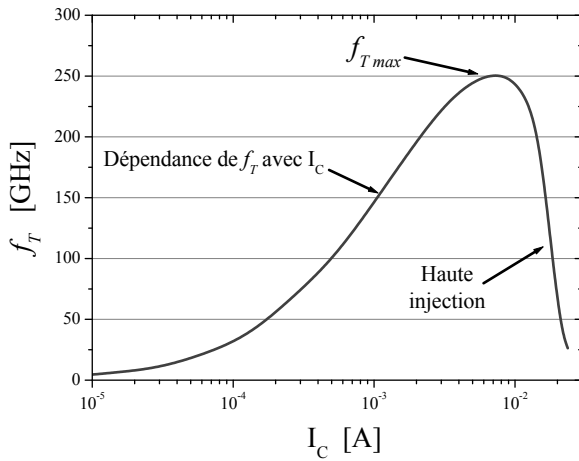


FIG. I.42 – Evolution de la fréquence de transition f_T avec le courant collecteur.

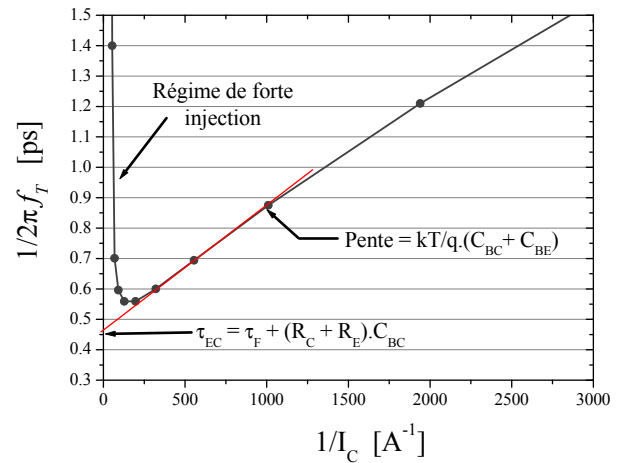


FIG. I.43 – Extraction de τ_{EC} en fonction de l'inverse du courant collecteur.

L'expression I.101 donne la fréquence de transition du gain en courant. C'est un paramètre important, car il donne une information fondamentale sur le temps de transit des porteurs dans le transistor. Cependant, il a quelques inconvénients : le fait qu'on commande en courant ne permet pas de tenir compte de la résistance de base, et la sortie court-circuitée n'est pas une situation réaliste. Dans un circuit, il y a toujours une charge d'utilisation.

I.4.3.b Expression de f_{MAX}

Il est donc intéressant de définir un deuxième facteur de mérite, f_{MAX} , basé sur le gain en puissance. La fréquence f_{MAX} est la fréquence maximale d'oscillation, et est obtenue à partir du gain en puissance du transistor, dans des conditions optimales d'adaptation et de neutrodynage.

Cependant, si le gain en courant est facile à définir, il existe plusieurs gains en puissance, selon les conditions que l'on désire. Le gain que nous retiendrons pour le calcul de la fréquence f_{MAX} est le gain de Mason, ou gain en puissance unilatéralisé, noté U. Il correspond au gain

en puissance obtenu lorsqu'il n'y a aucune transmission d'énergie de la sortie vers l'entrée. La fréquence f_{MAX} est généralement plus difficile à définir et donc à extraire que la fréquence f_T , car les effets extrinsèques ont une forte importance. Il est délicat de déterminer précisément ce qui fait partie des éléments intrinsèques et ce qui relève des aspects extrinsèques. Notamment, la résistance de base r_B du transistor joue un rôle important dans l'expression de f_{MAX} , la limite intrinsèque/extrinsèque étant, là encore, soumise à discussion.

Toutefois une expression théorique de f_{MAX} peut être obtenue [Roulston90] en ajoutant au schéma I.41 des résistances d'accès au transistor, du côté de la base (en entrée) et du collecteur (en sortie), comme indiqué sur la figure I.44. La résistance de base r_B du transistor a également été introduite et permet d'exprimer l'impédance d'entrée. Notons alors que la tension qui s'exerce sur la jonction E/B est différente de la tension externe appliquée.

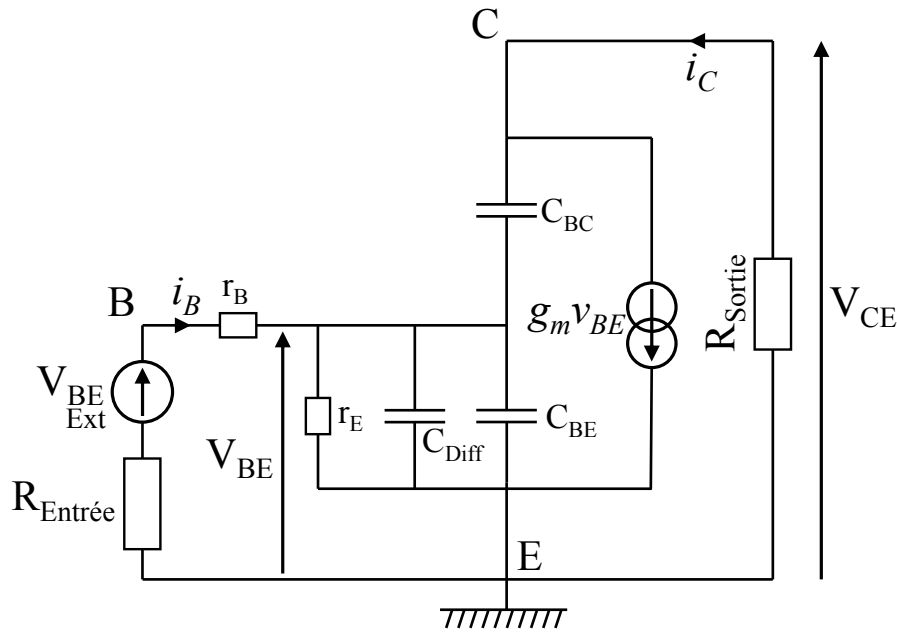


FIG. I.44 – Schéma équivalent petit signal du transistor bipolaire adapté en impédance pour le calcul de f_{MAX} .

Pour simplifier les calculs, on pose $C_\pi = C_{Diff} + C_{BE}$. L'impédance d'entrée est donc constituée de la mise en série de r_B et C_π en parallèle avec r_E , et vaut :

$$Z_e = r_B + \frac{r_E}{1 + j r_E C_\pi \omega} \approx r_B \quad (\text{I.102})$$

Cette impédance d'entrée se limite à r_B aux hautes fréquences.

L'impédance de sortie est par définition égale à V_{CE} sur I_C . D'après une analyse des courants et des tensions dans le schéma de la figure I.44, on en déduit l'expression de l'impédance de sortie, que l'on peut simplifier car la transconductance est très grande :

$$Z_s = \frac{C_\pi + C_{BC}}{C_{BC}(g_m + j C_\pi \omega)} \approx \frac{C_\pi + C_{BC}}{C_{BC} g_m} \quad (\text{I.103})$$

Dans le cas d'une adaptation simultanée, $R_{Entrée} = Z_e^*$ et $R_{Sortie} = Z_s^*$, le gain en puissance s'écrit donc :

$$G_P = \frac{Z_s^* i_C^2}{Z_e^* i_B^2} = \frac{C_\pi + C_{BC}}{r_B C_{BC} g_m} \frac{i_C^2}{i_B^2} \quad (\text{I.104})$$

D'après le schéma équivalent donné figure I.44, les courants de base et de collecteur s'expriment de la manière suivante :

$$i_C = g_m \cdot v_{BE} + j\omega C_{BC} (v_{CE} - v_{BE}) \approx g_m \cdot v_{BE} \quad (\text{I.105})$$

$$i_B = \frac{v_{BE}}{r_E} + j\omega (C_\pi + C_{BC}) v_{BE} - j\omega C_{BC} \cdot v_{CE} \quad (\text{I.106})$$

La tension de sortie v_{CE} peut s'exprimer en fonction de v_{BE} grâce aux équations I.103 et I.105 :

$$v_{CE} = -R_{Sortie} \cdot i_C = -\frac{C_\pi + C_{BC}}{C_{BC}} v_{BE} \quad (\text{I.107})$$

En injectant I.107 dans I.106, on en déduit l'expression de i_B :

$$i_B = \frac{v_{BE}}{r_E} + 2j\omega (C_\pi + C_{BC}) v_{BE} \quad (\text{I.108})$$

On peut négliger le terme $1/r_E$ lorsque ω est très grand. Le module du gain en puissance aux hautes fréquences s'écrit donc d'après I.104 :

$$G_P = \frac{g_m}{4 r_B C_{BC} \omega^2 (C_\pi + C_{BC})} \quad (\text{I.109})$$

En tenant compte du fait que $C_\pi = C_{Diff} + C_{BE}$ et de l'expression générale de f_T donnée par l'équation I.99, on obtient ($\omega = 2\pi f$) :

$$G_P = \frac{2\pi f_T}{4 r_B C_{BC} \omega^2} = \frac{f_T}{8\pi r_B C_{BC} f^2} \quad (\text{I.110})$$

La fréquence f_{MAX} à laquelle le gain G_P vaut 1 est donnée par l'équation suivante :

$$f_{MAX} = \sqrt{\frac{f_T}{8\pi R_B C_{BC}}} \quad (\text{I.111})$$

Il y a donc une relation directe entre la fréquence f_T et f_{MAX} . Lorsque le composant est correctement optimisé, les accès sont très peu résistifs et f_{MAX} est légèrement supérieure à f_T , comme indiqué dans la figure I.45.

Nous remarquons l'influence très importante de la résistance de base dans l'expression de la fréquence maximale d'oscillation. Ceci nous amène à un compromis : Une base fine permet généralement d'obtenir un faible temps de transit et une forte fréquence f_T , mais entraîne une résistance de base dégradée et une fréquence f_{MAX} faible. La solution généralement retenue pour pallier à ce problème est d'augmenter le dopage de base. Ainsi on diminue la résistance de base, mais cela implique un meilleur contrôle de la diffusion des espèces dopantes pour maintenir une épaisseur de base « électrique » suffisamment fine.

I.5 Effet des basses températures

Nous avons détaillé dans les parties précédentes le fonctionnement statique et dynamique du transistor bipolaire à température constante. Or la dépendance des courants du transistor bipolaire avec $\exp q V_{BE}/kT$ rend le composant très sensible aux variations de température [Ashburn94]. Concernant le fonctionnement dynamique, la mobilité des électrons varie énormément avec la température, les performances sont donc grandement améliorées lorsque la

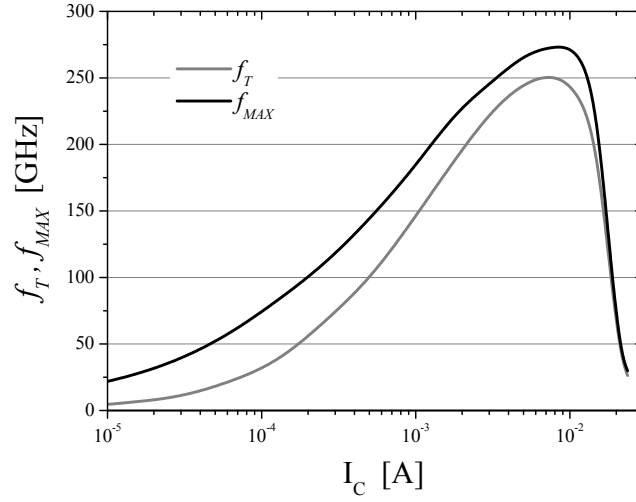


FIG. I.45 – Evolution de la fréquence maximale d’oscillation f_{MAX} avec le courant collecteur, et comparaison avec la fréquence de transition f_T .

température diminue [Banerjee03, Banerjee05] [Krithivasan06]. Les températures cryogéniques sont souvent utilisées pour analyser le transistor bipolaire car elles mettent en avant les limitations de la structure.

I.5.1 Comportement statique à basses températures

Les figures I.46 et I.47 représentent l’évolution des courbes de Gummel et de gain pour des températures comprises entre 300 K (température ambiante) et 50 K.

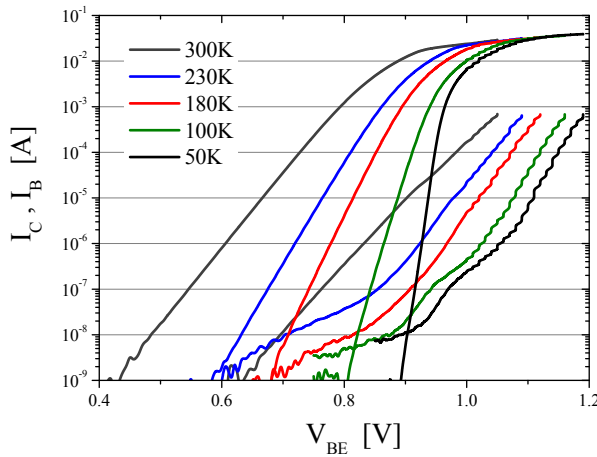


FIG. I.46 – Evolution des courbes de Gummel avec la température.

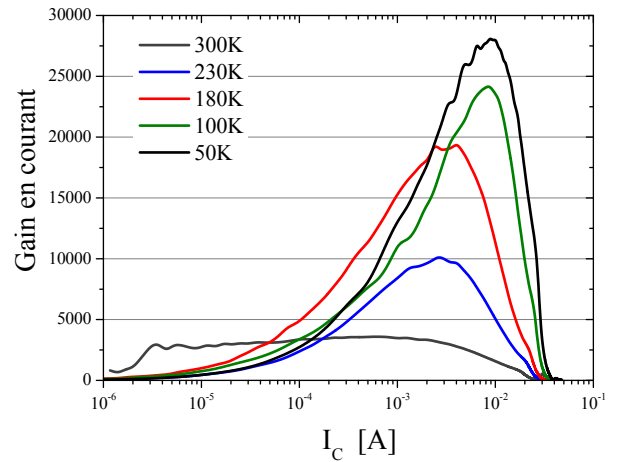


FIG. I.47 – Evolution du gain en courant avec la température.

La transconductance du transistor s’écrit (équation I.93) :

$$g_m = \frac{dI_C}{dV_{BE}} = \frac{q I_C}{n k T} \propto \frac{\exp q V_{BE} / k T}{T}$$

La diminution de température a donc pour effet d’augmenter la transconductance de manière très importante. Les courants de base et de collecteur étant tous deux proportionnels à la transconductance g_m , lorsque la température diminue, la pente du courant en fonction de V_{BE} aug-

mente, ce qu'on observe sur la figure I.46. Néanmoins, on note que pour une polarisation donnée, le courant est plus faible lorsque la température est plus basse. De plus, le courant de base perd son idéalité, ce qui est dû à des courants de fuite à basse température, comme des courants tunnels ou assistés par des pièges [Banerjee05], qui dépendent très peu de la température.

La gain en courant quant à lui est proportionnel à $\exp \frac{\Delta E_G}{kT}$, comme indiqué dans l'équation I.51. Le gain augmente fortement lorsque la température diminue : Pour un gain maximal de 4000 à température ambiante, le gain peut monter jusqu'à 30000 à 50 K (figure I.47), ce qui est dû à un ΔE_G assez fort causé par un taux de Ge important à l'entrée de la base (de l'ordre de 20 %).

I.5.2 Influence sur les performances dynamiques

La diminution de température a une nette influence sur le temps de transit du transistor bipolaire. En effet, la mobilité des porteurs est améliorée à basse température [Glicksman58], ce qui a pour effet d'augmenter la constante de diffusion des électrons et des trous (équations I.11 et I.12). Le temps de transit de base et d'émetteur est donc réduit. L'augmentation importante de la transconductance avec la diminution de la température (équation I.93) permet également d'améliorer sensiblement les performances dynamiques. Le tableau I.1 donne une évolution des valeurs de τ_{EC} et de f_T en fonction de la température.

	300 K	230 K	180 K	100 K	50 K
τ_{ECmin} [ps]	0.50	0.42	0.39	0.36	0.34
f_T [GHz]	250	291	328	365	412

TAB. I.1 – Variations du temps de transit τ_{EC} et de la fréquence f_T maximale associée avec la température (dimensions du composant : $0.15 \times 3.6 \mu\text{m}^2$).

La figure I.48 trace l'évolution de la fréquence f_T avec la température : Dans la zone de faible injection où f_T varie avec I_C , la fréquence de transition augmente avec la température, grâce à l'augmentation de g_m . La valeur maximale de f_T est aussi augmentée grâce à la diminution du temps de transit global. On note enfin un effet Kirk retardé. Ces trois phénomènes contribuent à augmenter de manière significative les performances à basse température.

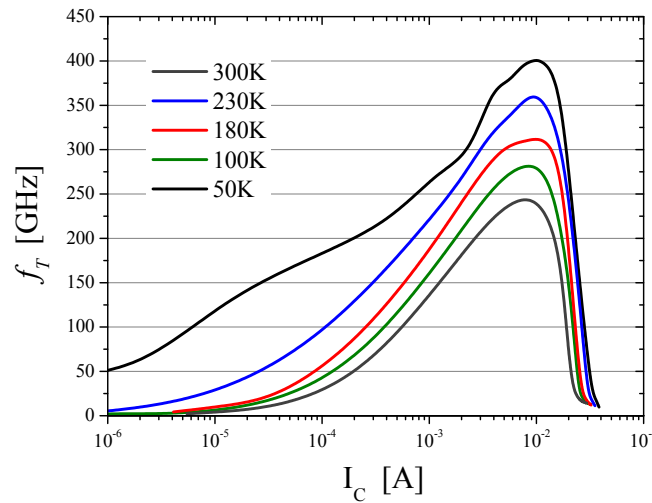


FIG. I.48 – Evolution de la fréquence de transition f_T avec la température.

I.5.3 Un outil d'anticipation

L'étude menée en faisant varier la température montre que la structure étudiée, basée sur une architecture donnée, avec un procédé de fabrication fixé, a des performances bien meilleures à basse température. Or lorsque la température diminue, ce sont les composantes intrinsèques du composant qui sont améliorées (mobilité, résistivité des couches de silicium). Les composantes extrinsèques restent *a priori* les mêmes. La conclusion est que les limitations des performances ne viennent pas de la structure en elle-même, mais bien de la partie intrinsèque du composant. Il est donc encore possible d'optimiser le profil du composant pour gagner en fréquence.

I.6 Méthodes de caractérisation

Afin de procéder à l'optimisation du composant, nous disposons de différents moyens d'investigation, nous permettant de repérer les points sur lesquels il faut s'attarder. On peut classer ces moyens de caractérisation en deux grands classes : physique et électrique. La caractérisation physique a pour but d'observer la conformation du composant (profils de dopage, contrôle des épaisseurs déposées...), la caractérisation électrique se base sur l'analyse du comportement du transistor, en régime statique ou dynamique. Chaque domaine d'analyse fait appel à des techniques et des méthodes qui lui sont propres.

I.6.1 Caractérisation physique

I.6.1.a Microscope électronique à balayage (MEB)

Le Microscope Électronique à Balayage (MEB) est un dispositif permettant d'analyser en surface des couches de matériaux, avec une précision de l'ordre de la dizaine de nm. Les couches analysées peuvent être conductrices (silicium, métal) ou isolantes (oxyde de silicium SiO_2 ou bien nitrure de silicium Si_3N_4). C'est une méthode destructive puisqu'on doit réaliser une coupe du composant pour l'analyser. On peut également utiliser le MEB en vue de dessus pour vérifier la définition des différents niveaux par les étapes de photolithographie (contrôle non destructif pendant la fabrication).

Le principe est de détecter les électrons secondaires ou rétro-diffusés pour analyser la surface de l'échantillon. Un exemple de coupe MEB est présenté sur la figure I.49.

I.6.1.b Microscope électronique à transmission (MET)

Le Microscope Electronique à Transmission (MET) permet d'analyser un matériau en étudiant la diffraction des électrons à travers son volume. Pour analyser un échantillon au MET, il est nécessaire de réaliser une coupe très fine de l'échantillon. La préparation prend donc plus de temps, mais les images sont de meilleure qualité. On a également accès à des informations non données par le MEB, comme une indication qualitative de la teneur en germanium de la base (voir figure I.50).

I.6.1.c Spectroscopie de masse par ions secondaires

La spectroscopie de masse par ions secondaires (en anglais SIMS pour *Secondary Ion Mass Spectroscopy*) sert à analyser la composition de surface d'un échantillon. Le principe est d'envoyer un flux d'ions lourds (Ar^+ , Cs^+ ou O_2^-) et très énergétiques (plusieurs keV) sur

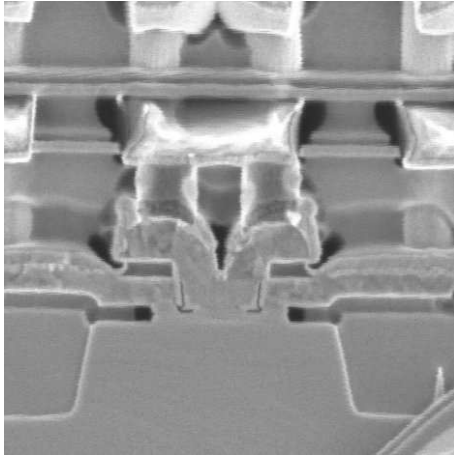


FIG. I.49 – Présentation d'une coupe MEB

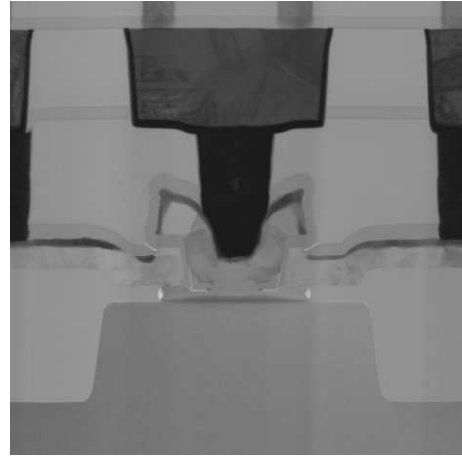


FIG. I.50 – Présentation d'une coupe MET

l'échantillon. Ce flux d'ions va dégrader la surface, ionisant les constituants de l'échantillon qui sont ensuite ré-émis et analysés par un spectromètre de masse.

Pour analyser le volume de l'échantillon, et obtenir ainsi une mesure de la concentration des espèces en fonction de la profondeur, il est nécessaire de creuser un cratère. Ce cratère va créer une rugosité en surface : La résolution du SIMS diminue avec la profondeur. Le spectromètre de masse analyse en fonction du temps les espèces recueillies, et permet de reconstituer le profil original en fonction de la profondeur. Un exemple de profil obtenu par SIMS est présenté sur la figure I.51.

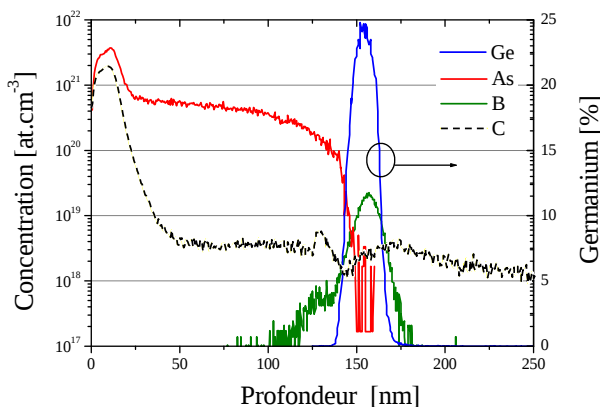


FIG. I.51 – Exemple de profil SIMS.

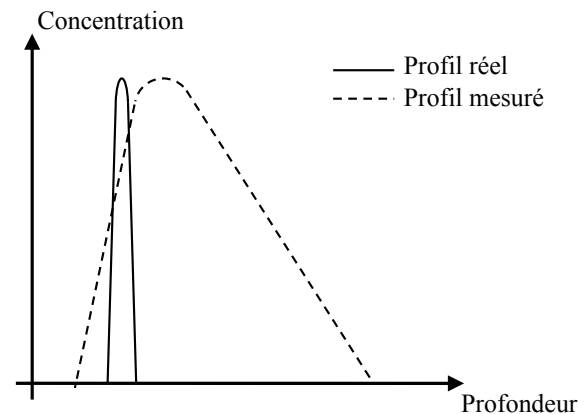


FIG. I.52 – Effet de recul observé sur un front montant abrupt.

Le profil relevé en SIMS est déformé par rapport au profil réel : la rugosité créée dans le cratère, mais aussi le recul des espèces sous le flux d'ions a pour effet d'étaler les pics de dopants, comme cela est exposé en figure I.52. Les fronts montants seront correctement détectés, mais les fronts descendants sont généralement déplacés par rapport à la réalité [Pakfar01]. En conclusion, si le SIMS apporte une information primordiale sur la répartition des espèces dopantes, il faut être prudent quant à l'interprétation des profils obtenus.

I.6.2 Mesures électriques en régime statique

L'étude du comportement électrique du composant en régime statique est la première phase de l'analyse du composant. Elle permet de repérer les éventuels défauts de fabrication (courants non idéaux, fortes résistances séries...) et permet d'extraire un grand nombre de paramètres, utiles pour la compréhension du fonctionnement du transistor.

I.6.2.a Idéalité, résistance d'émetteur

Le tracé de Gummel représente les variations des courants de base et de collecteur avec la tension V_{BE} . Afin d'être sûr que l'on est pas en régime de saturation, il peut être nécessaire de polariser la jonction B/C en inverse (typiquement $V_{CB} = 0$ V). Les courbes de Gummel donnent une indication sur l'idéalité du composant. D'après les équations I.24 et I.47, les courants de base et de collecteur varient proportionnellement avec $\exp qV_{BE}/kT$. Dans le cas de courants non idéaux, la dépendance en V_{BE} sera remplacée par une dépendance en $\exp qV_{BE}/nkT$, n étant le facteur d'idéalité du courant. Les courants de base et de collecteur ont généralement des idéalités différentes, car provenant d'effets physiques différents.

Le facteur d'idéalité est extrait dans la zone où la dépendance en V_{BE} est la plus linéaire possible, loin des effets de haute injection ou des résistances séries (figure I.53). Le principe est de comparer la pente extraite avec la pente idéale. Pour extraire la pente du courant I_C ou I_B , la méthode la plus précise repose sur une dérivée logarithmique en 3 ou 5 points.

Le tracé de Gummel permet également de faire apparaître les résistances séries du composant, notamment la résistance d'émetteur R_E (figure I.53). Cette résistance a pour effet de provoquer un écart à la droite idéale à cause d'une diminution de tension générée aux bornes de la jonction E/B, comme indiqué sur la figure I.54. L'écart à l'idéalité permet de caractériser la résistance d'émetteur. La méthode est décrite en détail dans le chapitre III.2.1.

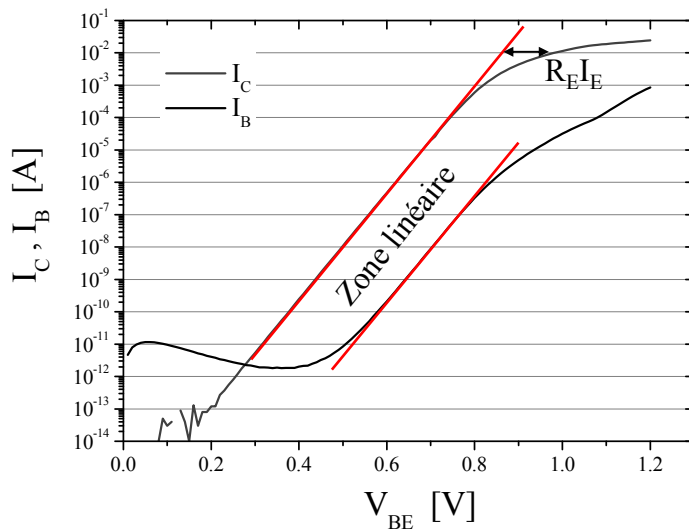


FIG. I.53 – Influence de la résistance d'émetteur sur les caractéristiques de Gummel. Le facteur d'idéalité est extrait dans la zone linéaire.

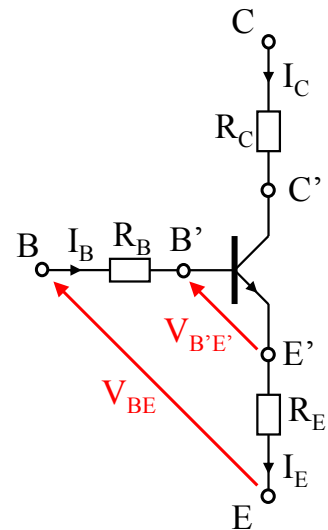


FIG. I.54 – Chute de tension dues aux résistances séries R_B et R_E .

I.6.2.b Caractérisation de l'avalanche dans la jonction B/C

Pour un transistor bipolaire en régime de fonctionnement, on peut extraire les paramètres d'avalanche en faisant varier la tension V_{CB} . Les variations du courant de base et du courant

collecteur dues à l'avalanche sont identiques, cependant la variation $\Delta I_B = \Delta I_C$ doit être extraite par rapport à I_B , le courant collecteur étant soumis à l'effet Early (voir figure I.55). ΔI_B en fonction de V_{CB} vaut donc $I_{B0} - I_B(V_{CB})$. Le facteur d'avalanche vaut :

$$M - 1 = \frac{\Delta I_B}{I_C - \Delta I_B} \quad (\text{I.112})$$

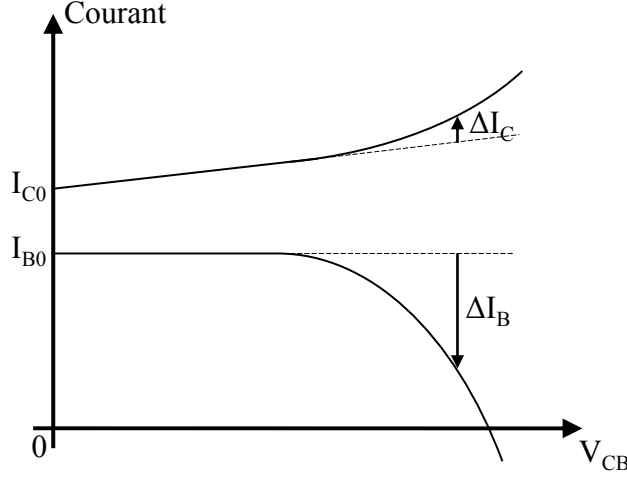


FIG. I.55 – Variation de courant de base et de collecteur à cause de l'effet d'avalanche.

D'après l'équation I.54, le logarithme de $M - 1$ vaut :

$$\ln M - 1 = M_f \ln V_{CB} - M_f \ln BVC \quad (\text{I.113})$$

La connaissance des courants de base et de collecteur en fonction de V_{CB} permet donc de tracer $\ln M - 1$ en fonction de $\ln V_{CB}$. Cette caractéristique est une droite dont la pente vaut M_f , et l'ordonnée à l'origine $-M_f \ln BVC$. On peut donc grâce à une seule mesure obtenir les caractéristiques complètes de l'avalanche dans la jonction B/C. Pour la plupart de nos dispositifs, la valeur typique de M_f est 5.7, et BVC vaut aux alentours de 3 V.

La tension de claquage BV_{CEO} est atteinte lorsque le courant de base s'annule, donc $\Delta I_B = I_{B0}$. D'après l'équation I.112, le facteur de multiplication devient :

$$M - 1 = \frac{1}{1 - I_{B0}/I_C} \approx \frac{1}{\beta} \quad (\text{I.114})$$

La tension BV_{CEO} peut donc également être définie comme la tension à laquelle le produit $(M - 1) \times \beta$ vaut 1. Cette représentation a pour avantage de repérer facilement la tension BV_{CEO} , et de comparer différents dispositifs n'ayant pas le même gain en courant ou le même facteur d'avalanche M_f , comme indiqué sur la figure I.56.

I.6.2.c Mesure des tensions d'Early

La variation du courant collecteur avec la polarisation V_{CB} et V_{BE} est due aux variations de l'épaisseur de la base neutre. On peut modifier l'expression du courant collecteur donnée d'après l'équation I.24 de la manière suivante pour tenir compte des effets Early direct et inverse :

$$I_C = \frac{qA_e}{G_B} \left(1 - \frac{V_{BE}}{V_{Ar}} + \frac{V_{CB}}{V_{Af}} \right) \exp \frac{qV_{BE}}{kT} \quad (\text{I.115})$$

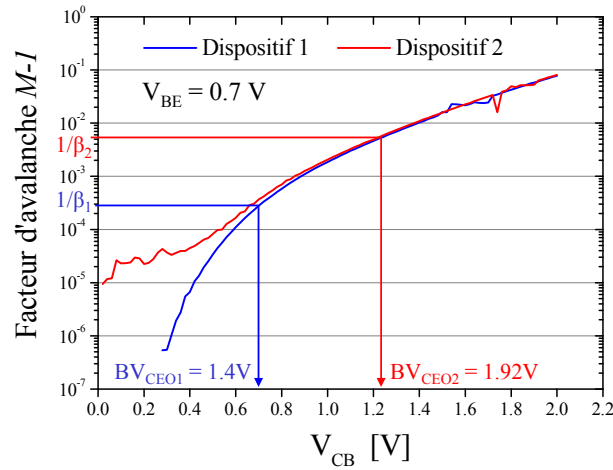


FIG. I.56 – Extraction de la tension de claquage BV_{CEO} d'après le facteur de multiplication $M - I$.

Un tracé de Gummel avec $V_{CB} = 0$ permet l'extraction de la tension V_{Ar} , en considérant que l'on a un courant collecteur idéal, ce qui est proche de la réalité. Cependant, pour extraire la tension d'Early directe, il faut tenir compte de l'effet Early inverse, car le transistor est en fonctionnement et V_{BE} ne peut être nul. Il est donc nécessaire de connaître V_{Ar} pour calculer V_{Af} . La tension d'Early directe s'extrait à partir d'une caractéristique de sortie à V_{BE} fixé (et non pas à I_B fixé, pour maintenir l'extension de la ZCE E/B). La pente de cette caractéristique permet de calculer la tension d'Early directe V_{Af} .

A noter la différence qu'il existe entre les tensions d'Early effectives, notées V_{Er} et V_{Ef} , qui dépendent d'un point de polarisation donné, et les tensions d'Early V_{Ar} et V_{Af} , qui sont des paramètres de modèle, donc indépendantes des conditions de polarisation du composant. On voit sur la figure I.57 que la tension d'Early directe effective mesurée directement avec $V_{BE} = 0.9$ V est de l'ordre de 5 V, ce qui est beaucoup plus faible que la véritable tension d'Early V_{Af} , de l'ordre de la centaine de Volt.

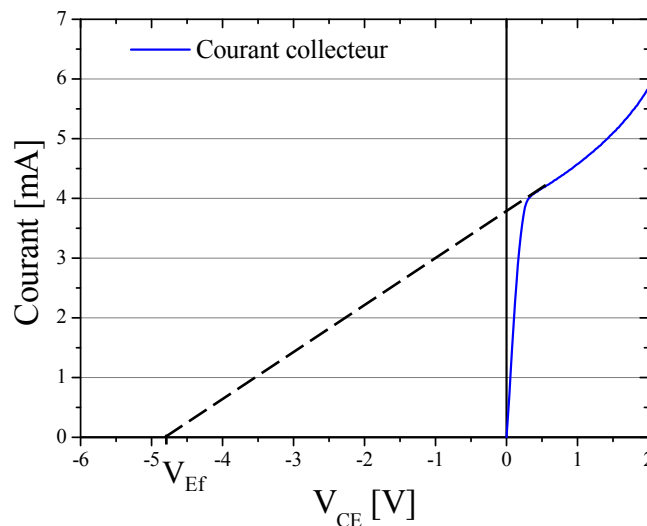


FIG. I.57 – Tension d'Early directe effective mesurée sur un TBH avec $V_{BE} = 0.9$ V. A V_{BE} plus faible, cette tension est plus élevée.

Les relations existantes entre les tensions d'Early et tensions d'Early effectives sont données par les relations :

$$V_{Ef} = V_{Af} \left(1 - \frac{V_{BE}}{V_{Ar}} \right) \quad (\text{I.116})$$

$$V_{Er} = V_{Ar} \left(1 - \frac{V_{BC}}{V_{Af}} \right) \quad (\text{I.117})$$

I.6.3 Hyperfréquences

I.6.3.a Intérêt des mesures de paramètres S

Afin d'évaluer le fonctionnement dynamique petit signal du composant, on étudie sa réponse à un signal sinusoïdal donné, à fréquence variable. Cependant, lorsqu'on atteint des fréquences de l'ordre du GHz, les déphasages induits par les lignes de transmission ne sont plus négligeables. Les longueurs d'ondes des signaux atteignent l'ordre de grandeur des dimensions caractéristiques du circuit. Il est alors inexact de parler simplement en terme de courant et de tension.

Il faut tenir compte de la propagation guidée des ondes électromagnétiques. Le composant est considéré comme un quadripôle, soumis à des ondes incidentes et réfléchies. Les relations entre ces ondes sont modélisées par une matrice 2×2, que l'on appelle matrice des paramètres S (S pour *Scattering* : dispersion). Cette matrice permet d'exprimer les ondes réfléchies b_j en fonction des ondes incidentes a_i (voir figure I.58), en tenant compte des coefficients de transmission à travers le quadripôle. Les relations entre ces différentes ondes sont données par l'équation I.118.

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (\text{I.118})$$

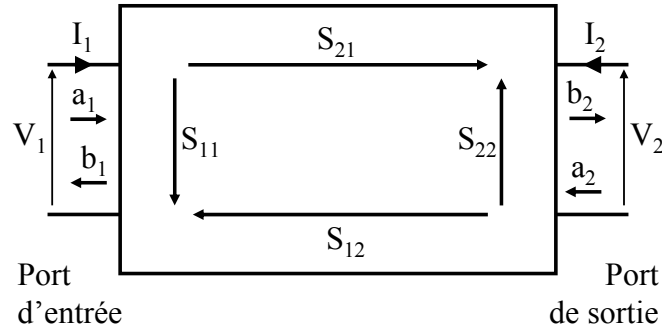


FIG. I.58 – Définition des paramètres S pour un quadripôle.

En analysant successivement les réponses b_1 et b_2 à des excitations a_1 et a_2 , et ce à différentes fréquences, on en déduit les quatre termes de la matrice $S_{i,j}$. A partir de cette matrice de paramètres S, on peut exprimer les différentes matrices de transfert H, Y ou Z, moyennant des transformations adaptées. Les différentes représentations S, H, Y et Z sont rigoureusement équivalentes : il n'y a pas de perte d'information en passant de l'une à l'autre.

I.6.3.b Méthode d'analyse

Les mesures hyperfréquences font l'objet de techniques et de protocoles particuliers [Zerounian00], qui sont rappelés brièvement dans ce chapitre.

I.6.3.b.i Adaptation

Afin de réaliser l'analyse d'un composant par mesure des paramètres S, il est nécessaire d'adapter l'impédance d'entrée et de sortie pour contrôler la propagation guidée des ondes électromagnétiques. L'adaptation permet également d'assurer la stabilité électrique du composant, car celui-ci a tendance à générer des oscillations du fait de ses grandes performances dynamiques. On utilise des accès spécifiques, à base de lignes de transmissions coaxiales et des sondes coplanaires. Le tout constitue un environnement adapté à 50 Ω jusqu'aux pointes coplanaires.

I.6.3.b.ii Calibrage

Toutefois, il est nécessaire de connaître les dispersions inhérentes à l'appareil de mesure lui-même. Les systèmes d'analyse n'étant pas parfaits, et la mesure n'étant pas réalisée au niveau du transistor, il y a des erreurs de mesure. Ces erreurs induites par le dispositif expérimental sont de 3 types :

- Erreurs aléatoires : Ces incertitudes de mesure ne peuvent être évitées, de par leur caractère totalement aléatoire. Une mesure statistique permet de réduire l'incertitude d'une mesure. La principale cause d'incertitude est la non-reproductibilité de la pose des pointes sur le dispositif à tester.
- Erreur de dérive dans le temps ou en température : En fonction du temps, les caractéristiques de l'appareillage peuvent varier. Une dérive en température, ou bien une relaxation mécanique des accès va modifier la propagation des ondes. Le système de mesure ne se comportera pas de la même manière au début et à la fin de la mesure, faussant ainsi l'observation du composant testé.
- Erreurs systématiques : Ce sont des erreurs reproductibles d'une mesure à une autre. Comme elles sont approximativement constantes dans le temps, elles sont entièrement caractérisées par un protocole de calibrage, et soustraites mathématiquement durant les mesures.

Une attention toute particulière doit donc être apportée durant toute la phase de mesure, du calibrage à la mesure du composant. Notamment, une température stable doit être assurée dans la pièce afin qu'il n'y ait aucune variation de la longueur des câbles d'accès semi-rigides, et de changement d'état des connexions.

Le calibrage a pour but de modéliser les erreurs systématiques induites par l'analyseur, aussi bien en réflexion qu'en transmission, vue de chacun des ports d'entrée ou de sortie. Il permet d'établir une correction du comportement du système de mesure jusqu'au plan des pointes hyperfréquence par une mesure d'éléments standards parfaitement connus, introduits à la place du dispositif à tester.

Il existe différentes méthodes de calibrage, dépendant des éléments standards à mesurer. Celle retenue pour l'analyse de nos composants est appelée OSTL (pour *Open, Short, Through* et *Load*), employant un substrat de calibrage externe. Ainsi, il n'est pas nécessaire de réaliser des motifs de calibrage sur le substrat des composants. L'étape d'épluchage viendra corriger le désaccord qui existe entre les deux types de substrat. Les éléments de calibrage mesurés successivement sont :

- Un circuit ouvert. En pratique, les pointes sont levées, ce qui n'est pas équivalent à poser les pointes sur un circuit ouvert, et est néanmoins pris en compte dans le modèle électrique du circuit ouvert.
- Un court-circuit : Chaque pointe hyperfréquence est posée sur un court-circuit.
- Une charge de 50 Ω .

- Une ligne de transmission adaptée $50\ \Omega$ pour étudier la transmission des ondes d'un port vers l'autre.

I.6.3.b.iii Épluchage

La connaissance des caractéristiques du système de mesure ne permet pas de mesurer directement un dispositif sur plaque de Silicium. Les plots de connexion du composant vont introduire des retards de propagation, dont il va falloir tenir compte lors de l'analyse du composant. Ces plots d'accès ajoutent au dispositif des lignes de propagation et des composantes capacitatives (voir figures I.59 et I.60). Ces éléments parasites dûs aux plots peuvent être éliminés a posteriori par le calcul en mesurant uniquement le plot d'accès sans le composant. C'est ce qu'on appelle l'épluchage (De-embedding).

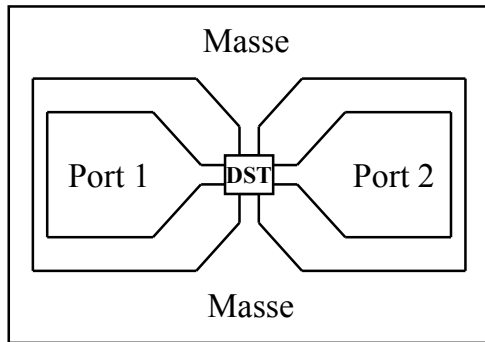


FIG. I.59 – Représentation du dispositif sous test (DST) encadré de ses plots d'accès hyperfréquence.

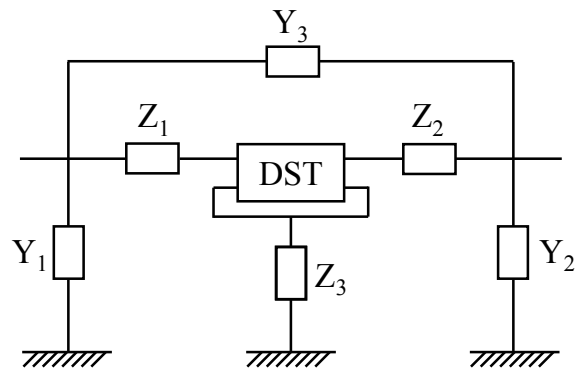
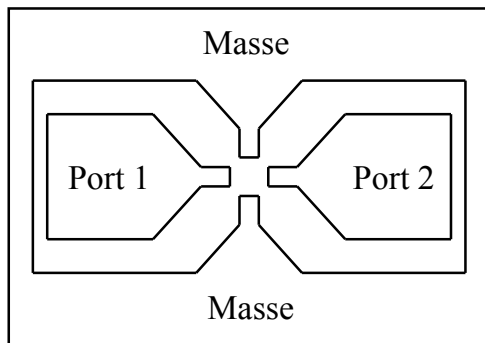
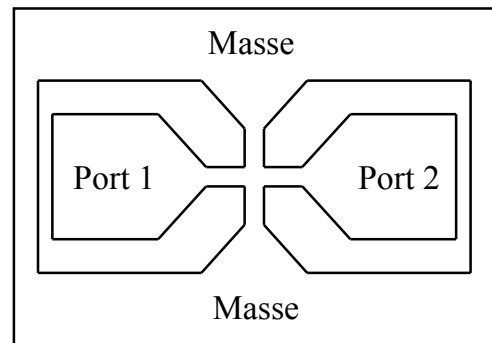


FIG. I.60 – Modélisation des composantes parasites ajoutées par les plots d'accès au composant.

Le composant testé ne sera vu qu'à travers ses accès, il faut donc éliminer ces composantes parasites si l'on veut connaître les caractéristiques du composant seul. On mesure donc les paramètres S du composant entouré de ses plots d'accès, puis on mesure les paramètres S des accès seuls. Deux motifs d'épluchage sont importants : motif « circuit ouvert », ou « Open » (figure I.61-a) pour tenir compte des admittances parallèles, et motif « court-circuit », ou « Short » (figure I.61-b) pour tenir compte des impédances séries.



a – Motif d'épluchage «open».



b – Motif d'épluchage «short».

FIG. I.61 – Motifs d'épluchage du composant.

L'épluchage est réalisé en deux fois : On élimine d'abord les admittances parasites du motif «open». Afin d'obtenir les caractéristiques du composant, il faut transformer la matrice de paramètres S mesurée en matrice admittance. On obtient la matrice intermédiaire comme indiqué par l'équation I.119.

$$[Y_{intermédiaire}] = [Y_{mesuré}] - [Y_{Open}] \quad (I.119)$$

On enlève ensuite la contribution du motif «short». Pour éliminer les effets des impédances séries, il est nécessaire de passer en matrice impédance, et retrancher la matrice des accès à la matrice intermédiaire obtenue précédemment, comme indiqué dans l'équation I.120.

$$[Z_{DST}] = [Z_{intermédiaire}] - [Z_{Short}] \quad (I.120)$$

On obtient ainsi la matrice $[Z_{DST}]$, caractéristique du composant uniquement. Pour extraire les différents paramètres hyperfréquences du composant, il est encore nécessaire de réaliser quelques transformations.

Dans la réalité, on considère que les accès sont essentiellement capacitifs, les retards des lignes de propagation étant négligeables. On peut donc se contenter uniquement de l'épluchage «open».

I.6.3.c Extraction des différents paramètres

Pour extraire les différents facteurs de mérite représentatifs des performances du composant, il est nécessaire de réaliser diverses transformations sur la matrice de paramètres S (ou Y, ou Z) obtenue. Notamment passer en matrice de transfert H permet l'extraction de la fréquence f_T , en étudiant la fréquence de transition du terme h_{21} .

L'extraction de f_{MAX} nécessite la définition du gain en puissance unilatéral U. La fréquence de transition de ce gain en puissance donne la valeur de f_{MAX} . D'après [Mason54], l'expression de U s'écrit :

$$U = \frac{|Y_{21} - Y_{12}|^2}{4 (\Re(Y_{11})\Re(Y_{22}) - \Re(Y_{12})\Re(Y_{21}))} \quad (I.121)$$

La figure I.62 représente les évolutions de h_{21} et du gain de Mason U en fonction de la fréquence, ainsi que les extractions de f_T et f_{MAX} d'après ces paramètres.

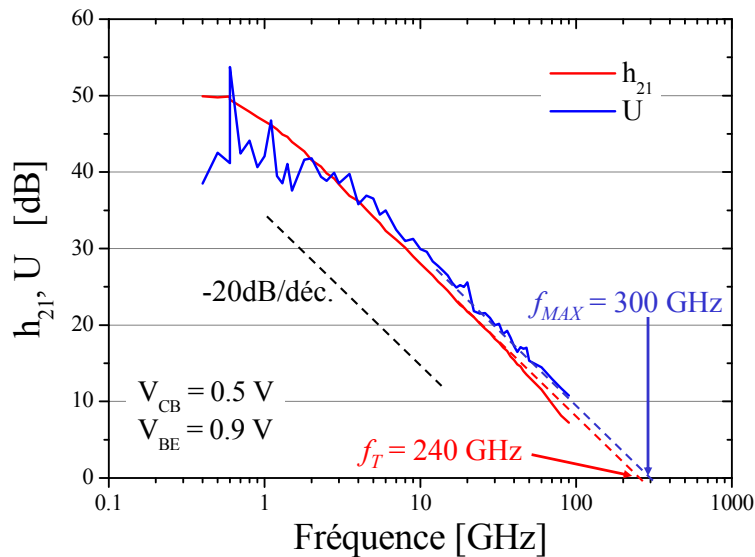


FIG. I.62 – Extraction des fréquences f_T et f_{MAX} pour un transistor de dimensions $0.12 \times 3.6 \mu\text{m}^2$ d'après les gains h_{21} et U.

L'extraction des autres paramètres significatifs du transistor bipolaire (résistances d'accès, capacités de jonction, capacités parasites...) nécessite de s'appuyer sur un schéma équivalent,

puis de définir une stratégie d'extraction dépendante du schéma retenu pour modéliser le TBH. Un exemple de stratégie est détaillée dans [Zerounian00]. C'est l'approche que nous avons retenue dans la suite de l'étude.

I.6.3.d Mesures hyperfréquences à basses températures

Afin de caractériser le composant à des températures très basses, jusque 50 K, nous devons le placer dans un cryostat, qui permet de refroidir une enceinte dans laquelle sont placés le composant, les pointes hyperfréquences et les motifs de calibration. Le vide est fait dans l'enceinte, et le support des échantillons est refroidi par circulation continue d'hélium. L'enceinte et les accès sont refroidis à l'azote liquide pour limiter l'apport de chaleur par les câbles hyperfréquences. Compte tenu de la différence de température entre l'intérieur et l'extérieur de l'enceinte, les câbles sont soumis à de forts gradients thermiques qui vont mettre du temps à se stabiliser, il faut donc en tenir compte lors de la mise en place de l'expérience et de la phase de calibration.

La phase de calibrage et l'acquisition des résultats s'effectuent de même manière à température ambiante et à basses températures. Cependant, les câbles hyperfréquences utilisés à basse température sont différents de ceux employés à température ambiante, la fréquence d'analyse est limitée (110 ou 50 GHz à température ambiante, 35 GHz dans le cryostat). Ceci peut être problématique pour la détection des fréquences de transition f_T et f_{MAX} , car les performances du composant étant améliorées à basse température, il se peut qu'à 35 GHz nous n'atteignons pas la décroissance à -20 dB/décade des gains en courant et en puissance.

Outre des précautions particulières lors des manipulations, il faut donc porter une attention accrue lors de l'extraction des paramètres S : Nous ne sommes pas à l'abri d'oscillations survenues lors de la mesure ou de divergences des algorithmes d'extraction.

I.6.4 Régime impulsionnel

En parallèle des analyses statiques (courant continu) et dynamiques (régime sinusoïdal forcé), il existe aussi le domaine des mesures impulsionnelles. Au lieu d'envoyer un signal de manière continue sur le dispositif, on peut envoyer de courtes impulsions, et analyser la réponse du composant à ces impulsions.

I.6.4.a Intérêt des mesures pulsées (isothermes)

Le principal intérêt des mesures en régime impulsionnel court est de pouvoir observer le composant avec un auto-échauffement négligeable : En régime continu, le composant s'échauffe sous l'effet des fortes densités de courant qui le parcourent (plusieurs mA par μm^2), ce qui modifie son fonctionnement (les performances du TBH se dégradent lorsque la température augmente). Cependant, lorsqu'une courte impulsion de tension est envoyée sur le transistor, celui-ci n'a pas le temps de s'échauffer, car les constantes de temps thermiques sont beaucoup plus grandes que les temps d'établissement des courants électriques. On peut donc remonter aux caractéristiques électriques du composant en dehors de tout effet d'auto-échauffement, comme indiqué sur la figure I.63.

Les mesures pulsées permettent aussi d'extraire la résistance thermique du composant (voir chapitre V.2 sur l'auto-échauffement), en comparant la caractéristique non-isotherme du transistor avec un réseau de courbes isothermes à différentes températures. On a ainsi accès à la température interne du composant, et indirectement à la valeur de sa résistance thermique R_{Th} .

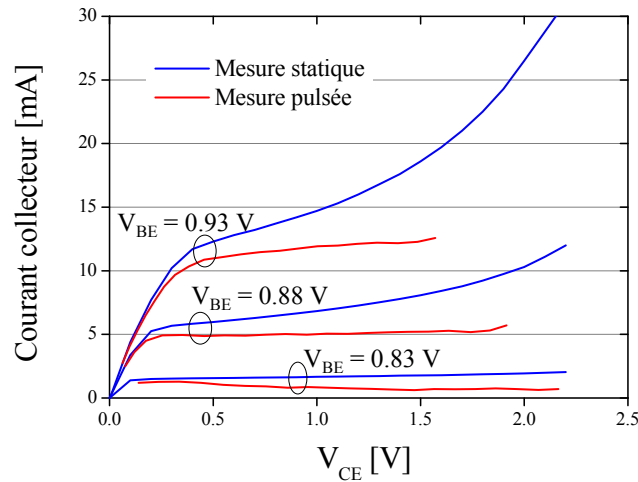


FIG. I.63 – Comparaison des caractéristiques de sortie obtenues par une mesure statique et une mesure pulsée. L’auto-échauffement visible en régime statique ne l’est plus sur les caractéristiques pulsées.

I.6.4.b Mesures pulsées dynamiques

En plus des mesures pulsées de paramètres statiques, on note que la société Amcad, en collaboration avec le laboratoire de l’université de Limoges propose des mesures de paramètres S en régime impulsionnel court. Ces mesures permettent en théorie de caractériser les composants en régime dynamique pulsé, et d’obtenir les performances hyperfréquences non influencées par l’auto-échauffement.

I.6.5 Mesures de bruit

Ce que l’on appelle bruit dans les composants électroniques est en fait des fluctuations aléatoires de courant et de tension, dues soit à des fluctuations de concentration de porteurs, soit à des fluctuations de la vitesse des porteurs. Selon l’équation I.122, la densité de courant varie si la concentration n varie (par exemple sous l’effet de génération-recombinaison des porteurs), ou bien si la vitesse $\vec{v}$ varie, sous l’effet du mouvement Brownien des particules.

$$\vec{j} = q n \vec{v} \quad (\text{I.122})$$

Le bruit est issu d’une fluctuation aléatoire de ces grandeurs. Le bruit électronique dans un semiconducteur en général peut être classifié en trois catégories :

- Le bruit thermique : Il s’agit du bruit lié à l’agitation thermique des électrons.
- Le bruit de grenaille, ou *shot noise* : Il est créé lors du passage d’une barrière de potentiel par les électrons. Il est dû à la nature quantique des porteurs de charge et a été découvert par Schottky.
- Le bruit basse fréquence (BF), ou bruit en $1/f$: Ce bruit est causé par les défauts d’états de surface et les défauts volumiques.

Le bruit est caractérisé par son spectre, la densité spectrale évoluant en fonction de la fréquence d’étude. Le bruit thermique et le bruit de grenaille sont des sources de bruit blanc, ce qui signifie que leur densité spectrale est la même quelle que soit la fréquence. Généralement, le bruit basse fréquence a un spectre en $1/f$, et peut donc être considéré comme négligeable à haute fréquence.

Cependant, le bruit BF induit par le transistor bipolaire est une principale limitation de la pureté spectrale du système, car il se convertit en bruit de phase sous l'effet des non-linéarités des différents oscillateurs. Le bruit large bande, ou bruit haute fréquence (HF) est aussi une préoccupation majeure, car il fixe le niveau minimal de bruit du système. Il est donc important de réduire aussi bien le bruit BF que le bruit HF pour les applications radiofréquences et millimétriques.

Nous allons maintenant passer en revue les différents phénomènes physiques à l'origine du bruit dans le transistor bipolaire à hétérojonction Si/SiGeC.

I.6.5.a Bruit en $1/f$

Le bruit basse fréquence dans le TBH peut être dû à deux phénomènes :

- Une fluctuation de la vitesse de recombinaison en surface dans la zone de déplétion de la jonction émetteur/base ;
- Un procédé de génération/recombinaison en volume due à la présence de pièges ayant des états d'énergie profonds et distribués dans la bande interdite du matériau. Cette origine du bruit en $1/f$ est mise en avant dans la partie consacrée à la recombinaison en base neutre (chapitre IV.4).

La figure I.64 présente un spectre de bruit basse fréquence. On observe une décroissance de la densité spectrale $S_{IB}(f)$ avec une dépendance en $1/f$ en dessous de 1 kHz. Au delà de ce point, les sources de bruit en $1/f$ sont masquées par le bruit de grenaille, de valeur constante égale à $2qI_B$.

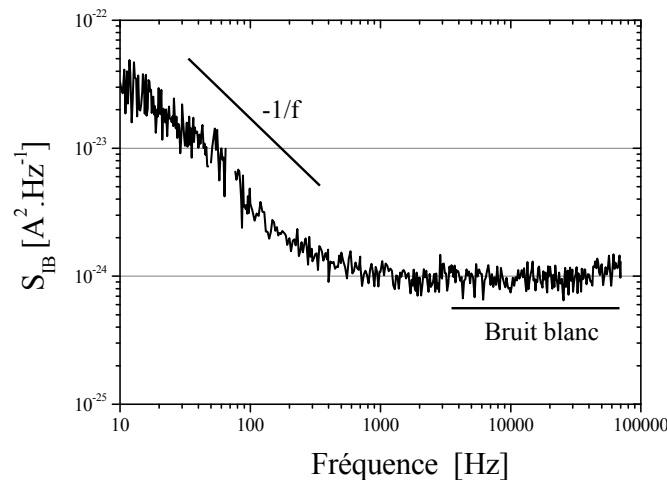


FIG. I.64 – Présentation d'un spectre de bruit basse fréquence. La décroissance en $1/f$ et le bruit de grenaille (bruit blanc) sont identifiables ($0.15 \times 3.6 \mu\text{m}^2$).

Le bruit basse fréquence est principalement un bruit du courant de base, et donc dépend largement du courant I_B . Le bruit en $1/f$ est lié à la polarisation de la jonction E/B, et à l'interface avec le polyémetteur. Les progrès réalisés dans le dépôt et le contrôle des matériaux (espaceurs en L, émetteur dopé in-situ) contribuent fortement à réduire le bruit BF. Le fonctionnement vertical du transistor bipolaire rend ce composant particulièrement compétitif en ce qui concerne le bruit, notamment en comparaison du CMOS.

I.6.5.b Bruit haute fréquence

Le bruit haute fréquence, également appelé bruit large bande, est la somme des trois composantes suivantes :

- Le bruit thermique induit par la résistance de base $R_{BB'}$;
- Le bruit de grenaille lié au courant de base ;
- Le bruit de grenaille lié au courant collecteur.

L'équation I.123 donne la valeur en dB du niveau de bruit NF du transistor bipolaire en fonction de la fréquence f , mesurée sous l'impédance R_S :

$$NF(f) = 1 + \frac{1}{R_S} \left[r_b + \frac{r_e}{2} + \left(\frac{1}{\beta} + \frac{f^2}{f_T^2} \right) \frac{r_b^2 + R_S^2}{2r_e} \right] \quad (\text{I.123})$$

Il est donc bénéfique d'augmenter le gain en courant pour diminuer le courant de base, augmenter la fréquence f_T , et ainsi diminuer le niveau de bruit. La diminution de résistance de base, nécessaire pour atteindre de fortes fréquences f_{MAX} , contribue à réduire le bruit HF. Le fait qu'il existe une corrélation entre les sources de bruit (shot noise des courants de base et de collecteur) contribue également à réduire le bruit dans le transistor bipolaire [Yau06].

I.7 Applications

I.7.1 Avantages du transistor bipolaire par rapport au MOS

Les principaux avantages du transistor bipolaire par rapport au transistor MOS résident dans le compromis entre fréquence de transition, tension de claquage et bruit. Les fréquences de transition du TBH SiGe sont généralement supérieures à celles des transistors MOS pour un nœud technologique donné, et le fonctionnement vertical du transistor bipolaire lui assure un bien meilleur bruit BF à fréquences f_T équivalentes. Le transistor bipolaire, du fait de sa meilleure tenue en tension, est particulièrement mieux adapté que le transistor MOS pour les applications nécessitant des fortes densités de puissance.

La plupart des chaînes de traitement RF et millimétrique font donc l'usage de TBH SiGe pour la réalisation de blocs tels que des amplificateurs faible bruit (LNA), des oscillateurs contrôlés en tension (VCO) ou des mélangeurs. La qualité du système RF dépend fortement des performances de ces blocs, surtout en ce qui concerne le bruit en $1/f$ et le bruit large bande. Notamment le LNA est utilisé en tant qu'étage d'entrée pour la plupart des applications radiofréquences, et bénéficie largement des avancées du transistor bipolaire SiGe. Le bruit dans un système RF étant fixé en grande partie par l'étage d'entrée, les contraintes imposées aux LNA sont très importantes. On retrouve donc des TBH dans la majorité de ces applications, car ils permettent d'obtenir de forts gains et un faible bruit.

I.7.2 Utilisation dans le domaine des hautes fréquences

Les transistors bipolaires à hétérojonction SiGeC sont utilisés pour des applications grand public nécessitant des fréquences de fonctionnement élevées. Les applications peuvent être divisées en plusieurs catégories :

- Les applications de télécommunication, comme les fibres optiques ou les réseaux de téléphonie mobile, et les réseaux informatiques WLAN (Réseau local sans fil) à 60 GHz. L'augmentation des fréquences de coupure des composants permet d'augmenter les débits d'informations transmises.
- Les applications de détection, comme les radars anti-collision à 77 GHz destinés au marché automobile. Ces applications sont très sensibles au bruit, ce qui fait du transistor bipolaire un composant de choix pour ce type de marché.

Les TBH SiGe sont aussi utilisés en commutation pour réaliser des multiplexeurs / démultiplexeurs pour les réseaux de fibres optiques. L'augmentation des débits de communication (jusqu'à 80 Gb/s) nécessite là encore la réalisation de transistors de plus en plus rapides.

Chapitre II

Technologie et fabrication des TBH SiGe

II.1 Introduction

Ce chapitre a pour objectif de présenter l'évolution des technologies de fabrication des transistors bipolaires SiGe. Les technologies BiCMOS de STMicroelectronics s'appuient dès leur origine sur une structure non auto-alignée, c'est à dire nécessitant plusieurs étapes de photolithographie pour réaliser la partie active du composant, et la base est réalisée par épitaxie non sélective. Nous allons étudier ces différentes architectures de fabrication du TBH, et nous mettrons en avant les éléments limitant chaque structure. Nous détaillerons ainsi les choix technologiques pris en faveur d'un auto-alignement complet utilisant l'épitaxie sélective de la base, et qui nous ont permis d'atteindre les performances à l'état de l'art actuel. Enfin, les performances obtenues par la structure auto-alignée sont comparées aux résultats publiés par les sociétés et laboratoires extérieurs.

II.2 Structures non auto-alignées développées dans le passé

II.2.1 *Structure simple polysilicium*

La première technologie BiCMOS utilisant un transistor bipolaire avec une base en silicium-germanium utilise le noeud CMOS 0.35 μm . Cette technologie possède des isolations par LOCOS (*LOCAl Oxidation of Silicon*), et les interconnexions métalliques sont réalisées en aluminium. La figure II.1 présente une vue en coupe d'un transistor correspondant à cette technologie.

Cette structure de transistor bipolaire est dite quasiment auto-alignée (QSA pour *Quasi Self Aligned*) car la base extrinsèque est implantée de manière auto-alignée sur le polyémetteur. Cependant, le polyémetteur n'est pas aligné par rapport à la fenêtre émetteur, il y a donc un désalignement possible entre base extrinsèque et base intrinsèque.

La fabrication de ce composant débute avec la réalisation des isolations par LOCOS, ainsi que la couche enterrée, le puits collecteur et le SIC. Un polysilicium (polybuffer) est déposé puis délimité par une gravure sèche afin qu'il n'en reste que sur l'oxyde de champ. Ce polysilicium amorphe sert à minimiser les effets de charge induits par la présence de l'oxyde. Vient ensuite le dépôt pleine plaque de la base en SiGe, qui a une croissance monocristalline sur les zones silicium et polycristalline sur l'oxyde de champ et le polybuffer. La base est recouverte d'une couche d'oxyde et de nitrure, qui est ensuite gravée pour délimiter la fenêtre émetteur. Le polysilicium d'émetteur (polyémetteur) est déposé, implanté et gravé pour libérer la base extrinsèque située en dessous. La base extrinsèque est alors implantée et délimitée par une étape de

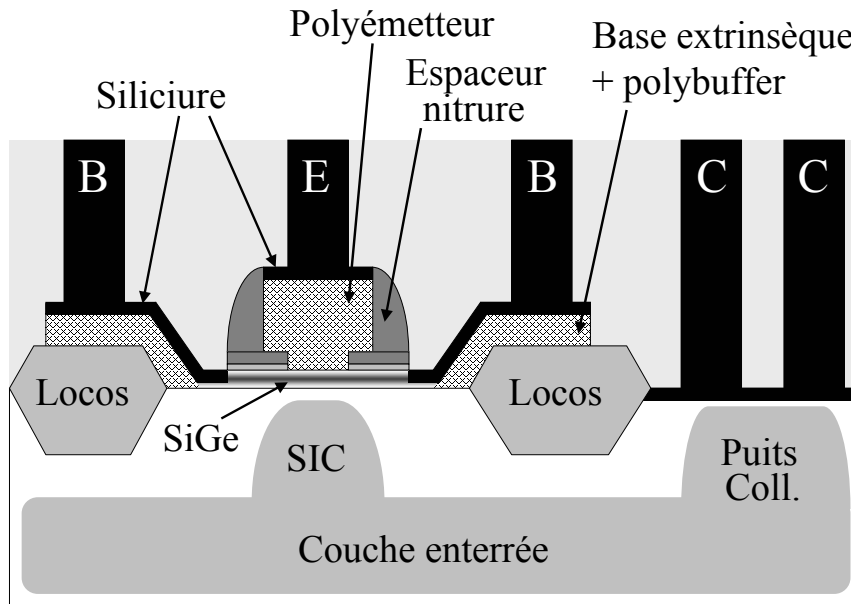


FIG. II.1 – Présentation d'une structure quasi auto-alignée (QSA) à simple polysilicium.

photolithographie, suivie d'une gravure sèche. Enfin, des espaceurs sont réalisés sur les flancs du polyémetteur, afin d'isoler électriquement la base de l'émetteur, et l'ensemble du transistor est siliciuré.

Cette structure est très simple, et demande peu d'étapes de fabrication. Les performances obtenues par cette structure sont $f_T = 45$ GHz pour une tension BV_{CEO} de 4 V et f_{MAX} de 60 GHz en technologie BiCMOS6G (0.35 μm). La largeur d'émetteur de cette technologie est $W_E = 0.4$ μm . Ce type de transistor a également été utilisé dans un noeud 0.13 μm avec de meilleures performances, la base étant réalisée en SiGeC et l'émetteur étant déposé de manière monocristalline. Cette structure est également utilisée par l'IMEC, qui obtient des fréquences f_T et f_{MAX} de 205/275 GHz (voir partie II.4.2).

Cependant le principal problème causé par ce schéma d'intégration est la création de défauts lors de l'implantation de la base extrinsèque, qui favorisent la diffusion du bore dans la base (mécanisme TED, pour Transient Enhanced Diffusion). L'introduction de carbone permet de corriger ce problème de diffusion. De plus, d'après des observations du transistor de BiCMOS6G, la siliciuration de base est moins efficace dans le SiGe que dans du silicium [Baudry01]. Il est donc préférable d'utiliser deux matériaux différents pour la base extrinsèque et intrinsèque.

II.2.2 Structure double polysilicium

Une structure à double polysilicium (d'émetteur et de base) a donc été introduite. Une vue en coupe de cette structure est présentée figure II.2. L'emploi d'un polysilicium supplémentaire pour réaliser la base extrinsèque permet de supprimer la création de défauts dans le SiGe dûs à l'implantation (le polysilicium est source de diffusion des dopants en se superposant à la couche de SiGe), et favorise la siliciuration.

Cette technologie de transistor est plus complexe qu'une structure simple polysilicium, et a été pour la première fois intégrée dans une technologie BiCMOS 0.25 μm , les interconnexions étant toujours en aluminium. Le LOCOS est remplacé par des STI (Shallow Trench Isolation) et les DTI (Deep Trench Isolation) ont été introduits pour l'isolation des composants les uns par rapport aux autres, ce qui permet aussi de réduire la capacité collecteur/substrat et l'encombre-

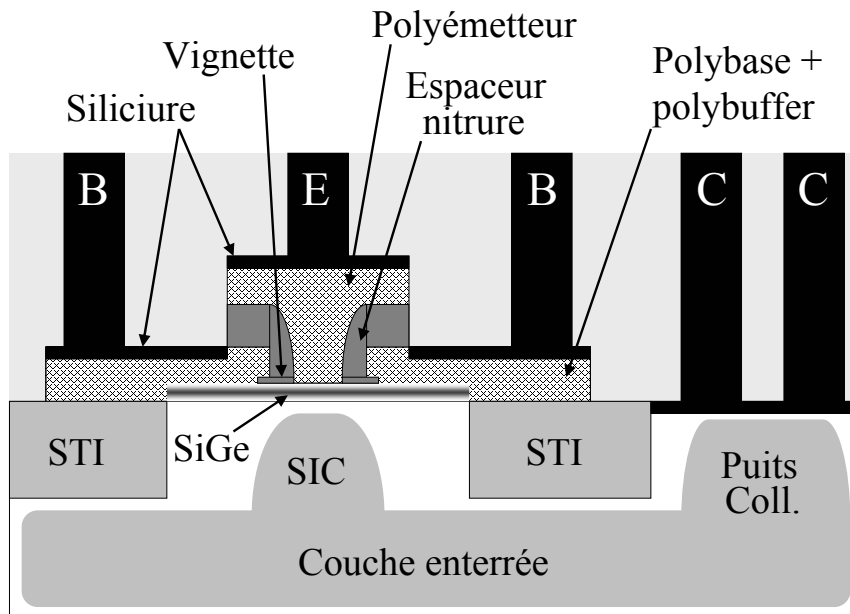


FIG. II.2 – Présentation d'une structure quasi auto-alignée (QSA) à double polysilicium.

ment sur un circuit. Les progrès sur le dépôt des matériaux ont permis également d'introduire pour la première fois des émetteurs monocristallins dopés in-situ.

Comme pour la structure simple polysilicium, la fabrication du transistor bipolaire double polysilicium commence par la réalisation de l'isolation (STI, DTI, couche enterrée), l'implantation du collecteur (puits collecteur et SIC), et le dépôt du polybuffer. La base en SiGe est déposée par épitaxie non sélective, suivie par le dépôt d'un empilement d'oxyde et de nitrure pour créer ce qu'on appelle la vignette, qui est ensuite délimitée par une étape de photo. Le polysilicium de base (polybase) est déposé et implanté, et on procède au dépôt d'une couche d'oxyde destinée à isoler les deux polysilicium. L'empilement total est délimité par une photo suivie d'une gravure, afin de libérer les contacts collecteur. Par une autre étape de photolithographie, on délimite la fenêtre émetteur d'une largeur de $0.3\ \mu\text{m}$, et on grave l'empilement oxyde - polybase. La gravure du polysilicium étant différente de celle du nitrure, la vignette n'est pas attaquée. Cette vignette, nécessitant une étape de photolithographie supplémentaire, est nécessaire pour stopper la gravure de la fenêtre émetteur sans endommager la base SiGe.

On réalise dans cette fenêtre émetteur des espaceurs en nitrure, afin de diminuer la largeur effective d'émetteur et d'isoler électriquement le polybase du polyémetteur. Au moment de la gravure du nitrure des espaceurs, la vignette est également gravée, libérant l'oxyde piédestal qui est éliminé par une gravure humide juste avant le dépôt de l'émetteur. Après dépôt du polyémetteur par RTCVD, puis gravure de ce dernier, on procède au recuit d'activation, et l'ensemble du transistor est siliciuré. Suivent ensuite le dépôt du premier niveau de diélectrique et la gravure des trous de contacts, qui sont remplis par un empilement Ti/TiN/W.

Cette structure est décrite plus précisément dans [Baudry01], et les performances obtenues avec une base SiGe atteignent $70\ \text{GHz}$ de f_T pour une valeur de f_{MAX} de $90\ \text{GHz}$. BV_{CEO} vaut $2.6\ \text{V}$. C'est avec cette structure que le carbone a été introduit pour la première fois, dans la technologie BiCMOS7RF utilisant le noeud $0.25\ \mu\text{m}$ ($W_E = 0.25\ \mu\text{m}$) [Baudry03]. Les performances obtenues sont alors de $60\ \text{GHz}$ de f_T pour $120\ \text{GHz}$ de f_{MAX} , BV_{CEO} valant $3\ \text{V}$.

La dernière technologie utilisant cette architecture est BiCMOS9, utilisant le noeud $0.13\ \mu\text{m}$ ($W_E = 0.17\ \mu\text{m}$), avec des interconnexions en cuivre. Les performances atteintes sont des fréquences f_T et f_{MAX} de l'ordre de $150\ \text{GHz}$, avec une tension BV_{CEO} de $1.7\ \text{V}$ [Laurens03].

II.3 Développement d'une structure complètement auto-alignée

II.3.1 Nécessité d'une structure complètement auto-alignée

Les avancées technologiques tant sur le dépôt des matériaux que sur la photolithographie et la gravure sèche ont permis d'atteindre d'excellentes performances avec la structure double polysilicium. Cependant, le manque d'auto-alignement est critique pour atteindre des performances ultimes. En effet, pour réduire au maximum les effets parasites, il faut réduire l'extension latérale du composant, comme indiqué au chapitre III. Or, l'existence d'un désalignement possible entre les différents masques empêche de diminuer au maximum l'encadrement entre la fenêtre émetteur et la vignette [Baudry01].

La figure II.3 présente en effet une comparaison entre un alignement correct et un fort désalignement entre fenêtre émetteur et vignette. Si la photo de la fenêtre déborde de la vignette, la base SiGe pourra être dégradée par la gravure, et le composant ne sera pas fonctionnel. La vignette doit donc être délimitée plus grande que nécessaire pour prendre en compte ce désalignement, ce qui dégrade la résistance d'accès à la base. Afin de pallier à ce problème, une structure complètement auto-alignée a été développée. Toutes les parties actives du transistor étant réalisées à partir d'un seul masque, celui de la fenêtre émetteur, il est possible de réduire l'encombrement latéral du composant, sans risque de dysfonctionnement causé par un problème d'alignement.

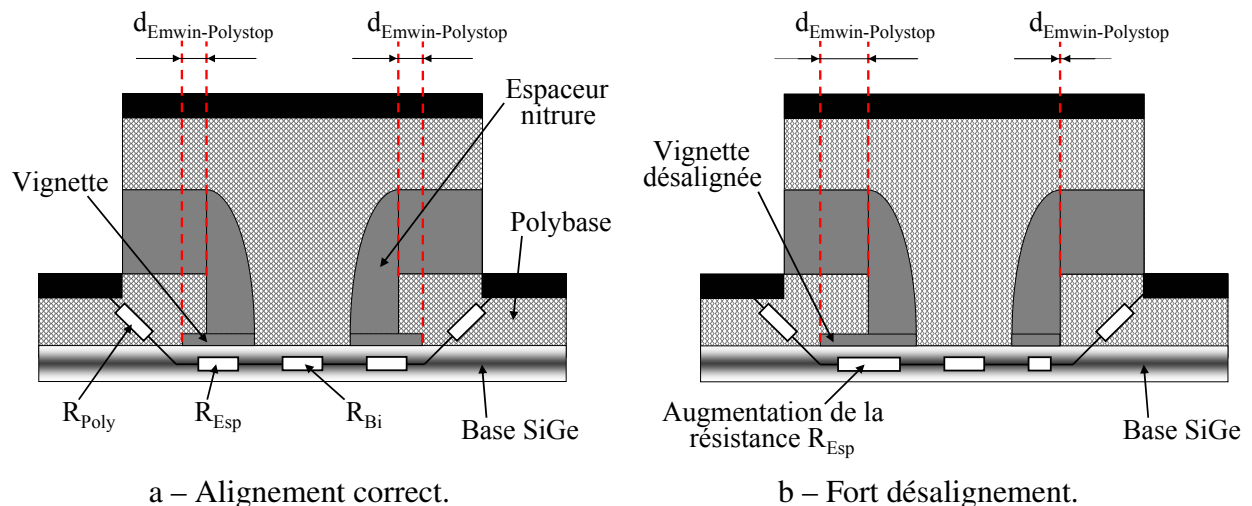


FIG. II.3 – Problématique d'alignement d'une structure QSA double polysilicium.

Cependant, il n'existe pas de schéma d'intégration simple de transistors bipolaires complètement auto-alignés. La difficulté d'intégration doit être reportée soit sur l'étude des matériaux, ou bien sur les étapes spécifiques comme le polissage mécano-chimique (CMP) ou la photolithographie. Différentes techniques permettant d'obtenir l'auto-alignement existent, nous allons détailler les deux approches qui ont été étudiées chez STMicroelectronics.

II.3.2 Structure auto-alignée par émetteur sacrificiel

II.3.2.a Présentation de la structure inverse émetteur

Une première possibilité pour réaliser un TBH complètement auto-aligné repose sur le remplacement d'un polyémetteur sacrificiel, que l'on appelle technologie damascène. La figure II.4 présente une vue en coupe de ce type de structure.

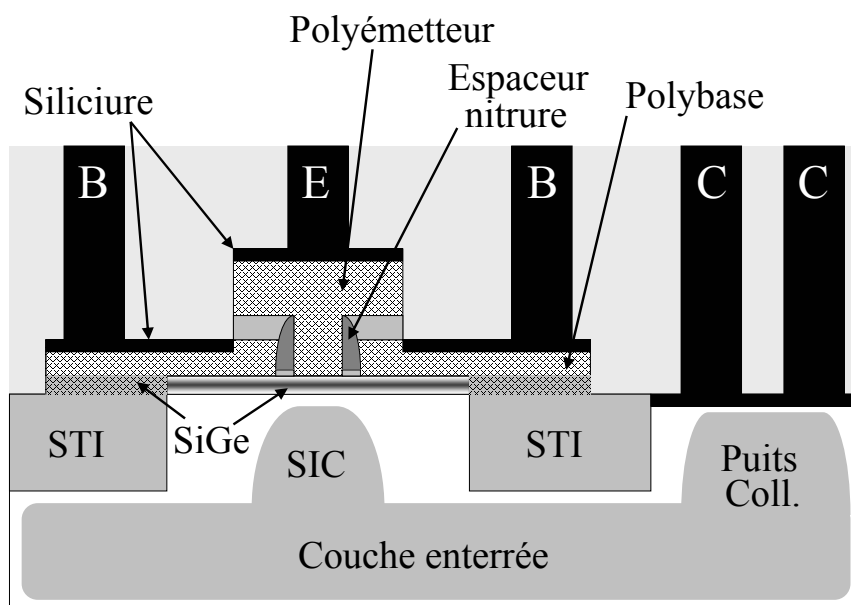


FIG. II.4 – Vue en coupe d'une structure complètement auto-alignée (FSA) par le procédé damascène.

La structure damascène utilise une épitaxie non sélective pour la base intrinsèque, et nécessite une épitaxie sélective de silicium P+ pour réaliser la base extrinsèque. L'étape critique est la libération du haut du polyémetteur sacrificiel : Elle peut se faire soit par une étape de CMP adaptée, approche retenue par IBM [Ahlgren01], soit par une étape de photolithographie utilisant un contre-masque émetteur, qui est l'approche retenue par Jazz [Racanelli01a].

Le procédé de fabrication décrit dans la suite de cette partie utilise la première méthode.

II.3.2.b Procédé de fabrication

La fabrication de transistors bipolaires par le procédé damascène commence avec la réalisation de la base intrinsèque en SiGe par épitaxie non sélective. Puis cette couche de base est recouverte d'un oxyde fin, et du polysilicium sacrificiel, lui même recouvert d'un diélectrique de protection. L'émetteur sacrificiel est délimité par une photolithographie suivie d'une gravure, puis des espaceurs externes en nitrure de silicium sont réalisés. Ensuite, on procède à la croissance de la base extrinsèque par épitaxie sélective. Celle-ci sera donc auto-alignée par rapport à l'émetteur.

Après dépôt d'un matériau diélectrique épais (empilement nitrure / oxyde), une étape de CMP vient polir le haut des plaques de manière à libérer le haut de l'émetteur. Il s'agit de l'étape la plus critique du procédé de fabrication. Une gravure sélective sert à vider l'émetteur sacrificiel, puis après élimination de l'oxyde, on dépose le polyémetteur dopé in-situ. Celui-ci est délimité par une étape de photo, et les contacts de base sont libérés. L'ensemble du transistor est siliciuré.

II.3.2.c Avantages / Inconvénients

Cette structure offre la possibilité de réaliser des émetteurs de largeur très faible, car la dimension de l'émetteur sacrificiel bénéficie des procédés de photolithographie de la grille des transistors MOS. On note que ce procédé nécessite tout de même une épitaxie sélective, mais le dépôt de Si dopé P+ sélectif est plus facile à mettre en œuvre qu'une épitaxie sélective de SiGeC dopé. L'incorporation de carbone dans la base est également plus simple dans un dépôt non sélectif que dans un dépôt sélectif, permettant un meilleur contrôle du profil de dopants.

Cependant, cette structure n'a pas connu de développements poussés, pour plusieurs raisons :

- Le procédé de CMP est très difficile à mettre en œuvre. Il faut en effet aplanir jusqu'au sommet de l'émetteur sacrificiel, sans aller trop loin car on risque de libérer le polysilicium de base. Si celui-ci est apparent, il sera également gravé lors du vidage de la cavité émetteur. Ce contrôle important à apporter au polissage est un problème majeur.
- Le dépôt de polysilicium dans une cavité étroite et haute pose des problèmes de remplissage. Des espaces vides vont apparaître, et à cause de la topologie élevée du composant, la résistance d'émetteur sera très forte (voir figure II.5).
- Une épitaxie sélective est tout de même nécessaire pour réaliser la base extrinsèque du composant.

La figure II.5 présente une observation au microscope électronique à transmission d'un TBH réalisé par le procédé damascène. L'espace vide dans le polyémetteur est bien visible (en blanc).

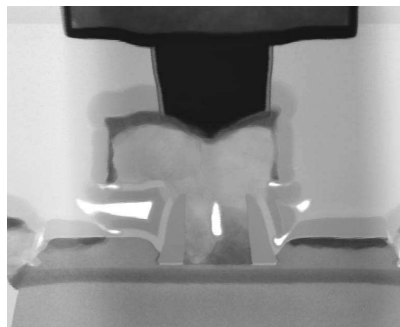


FIG. II.5 – Vue en coupe d'un transistor réalisé avec le procédé damascène, mettant en évidence un défaut de remplissage de l'émetteur.

Ainsi, les études préliminaires ont montré que ce procédé serait délicat à mettre en place pour l'intégration BiCMOS. La topologie élevée due à la présence des grilles des transistors MOS pose en effet problème lors de la CMP. En ce sens l'utilisation d'un procédé contre-masque semble plus favorable, mais l'alignement est critique. Dans le même temps, les progrès réalisés sur les matériaux, notamment en épitaxie, ont démontré qu'une structure utilisant un dépôt sélectif de la base poserait moins de problèmes d'intégration.

II.3.3 Structure auto-alignée par épitaxie sélective de la base

Les premières études sur le dépôt de SiGeC sélectif ont démontré que le blocage du bore par le carbone était légèrement moins efficace que dans un dépôt non sélectif. Cependant, les récents développements permettent d'obtenir un contrôle de la diffusion du bore aussi efficace avec un procédé sélectif que non sélectif [Brossard06].

Le choix de ST pour la réalisation de transistors bipolaires rapides s'est donc porté sur une structure auto-alignée à épitaxie sélective de la base. Cette structure offre de multiples avantages, malgré la difficulté que représente l'épitaxie sélective :

- Elle est dite « Super Self Aligned ». En effet, outre l'auto-alignement émetteur-base, le SIC peut être implanté au travers de la fenêtre émetteur avant le dépôt de la base. On a donc un système émetteur/base/collecteur complètement auto-aligné.
- La base extrinsèque est isolée du collecteur par un oxyde, ce qui permet de diminuer la capacité base/collecteur.
- La réalisation d'espaceurs internes permet de réduire fortement la largeur effective d'émetteur, sans problème de remplissage.
- La base intrinsèque est déposée très tard dans l'enchaînement des opérations, elle subit donc un budget thermique plus faible.

La suite de cette partie présente plus en détail le procédé de FSA-SEG (Fully Self Aligned - Selective Epitaxial Growth).

II.3.3.a Auto-alignement émetteur-base-collecteur

La figure II.6 présente une vue en coupe de la structure double polysilicium auto-alignée. L'auto-alignement est réalisé autour de la fenêtre émetteur, une seule étape de photolithographie est nécessaire lors de l'élaboration de la partie intrinsèque du composant. Comme il n'y a plus de risque de désalignement, les contacts de base et de collecteur peuvent être rapprochés, ce qui diminue les résistances d'accès et permet ainsi l'augmentation de f_T et f_{MAX} . La base est déposée dans une cavité spécialement ouverte à cet effet dans l'oxyde piédestal, par une épitaxie dite sélective détaillée dans la partie II.3.3.b suivante.

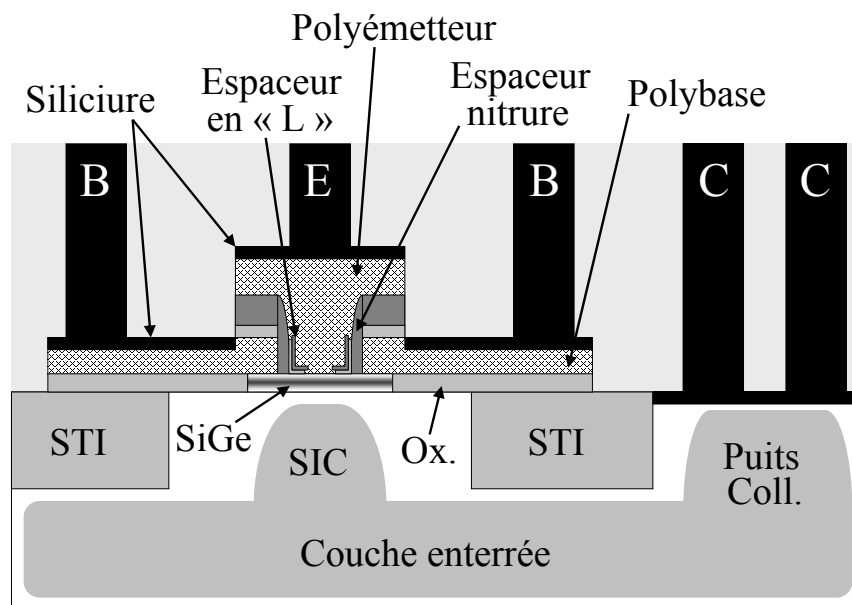


FIG. II.6 – Présentation d'une structure complètement auto-alignée (FSA) à double polysilicium.

L'avantage de l'auto-alignement est de diminuer la distance entre l'émetteur et le contact de base, comme indiqué dans la figure II.7. La distance émetteur/contact de base d est importante dans le cas d'une structure quasi auto-alignée (a – QSA) car le débordement de la vignette est important. Dans le cas d'une structure complètement auto-alignée (b – FSA), les vignettes ont disparu, donc le contact de base est fortement rapproché.

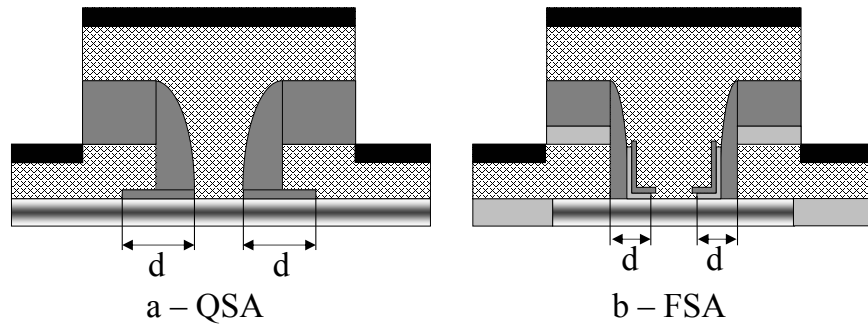


FIG. II.7 – Réduction de la distance émetteur/contact de base (d) induite par l'auto-alignement.

II.3.3.b Epitaxie sélective de la base

L'épitaxie est une technique de dépôt spécifique, utilisée entre autres pour réaliser les bases des transistors bipolaires. Une croissance épitaxiale consiste en un dépôt progressif d'atomes de manière ordonnée : Les atomes de Si et de Ge, vont donc se déposer progressivement au dessus de la plaque de silicium, en reproduisant la maille du substrat. Ce mode de croissance, plutôt lent (de quelques Å / min à quelques dizaines d'Å / min) et contrôlé permet d'obtenir des couches d'excellente qualité cristallographique.

Le dépôt de SiGe pour les bases de nos transistors se fait dans un réacteur RP-CVD (Reduced Pressure Chemical Vapor Deposition), ce qui signifie dépôt chimique en phase vapeur à pression réduite. Les épitaxies Si et SiGe se font à une pression comprise entre 1 et 100 Torr (1 Torr = 133.3 Pa) ; la température dans la chambre varie entre 650 et 850°C. [Fellous02]. Les gaz précurseurs sont introduits dans la chambre de dépôt, et la plaque de silicium est portée à haute température. La température élevée du substrat permet d'apporter l'énergie d'activation nécessaire à la réaction chimique (craquage moléculaire).

Les conditions de dépôt (température, pression, gaz précurseurs...) jouent énormément sur les propriétés de la couche de SiGe épitaxiée. L'épaisseur critique h_C , avant que des dislocations des mailles contraintes apparaissent, dépend des conditions qui règnent dans la chambre, et peut être améliorée en jouant sur la vitesse de croissance. En particulier, des conditions de dépôt judicieusement choisies assureront une croissance uniquement sur les zones de silicium, et non sur les zones diélectriques : il s'agit d'une épitaxie sélective. Cette sélectivité est obtenue en ajoutant aux gaz précurseurs des gaz additionnels qui modifient les propriétés des dépôts de matière. Notamment, l'utilisation d'une chimie chlorée empêche le dépôt de Si ou de Ge sur les zones diélectriques (nitrure ou oxyde de silicium), la croissance se fera donc uniquement sur les zones où le silicium est apparent.

Lors de l'épitaxie sélective de la base (ce qui est le cas pour tous les transistors étudiés durant cette thèse), le SiGe croît uniquement dans une cavité ouverte à cet effet, comme exposé dans la figure II.8.

II.3.3.c Enchaînement des étapes de fabrication

Les opérations successives pour réaliser le transistor bipolaire sont détaillées dans la figure II.9. Il s'agit d'une boucle courte afin de réaliser uniquement le transistor bipolaire, mais intégrant les contraintes de budget thermique du CMOS (recuit d'activation final). L'enchaînement des opérations est adapté de la structure QSA à double polysilicium de BiCMOS9.

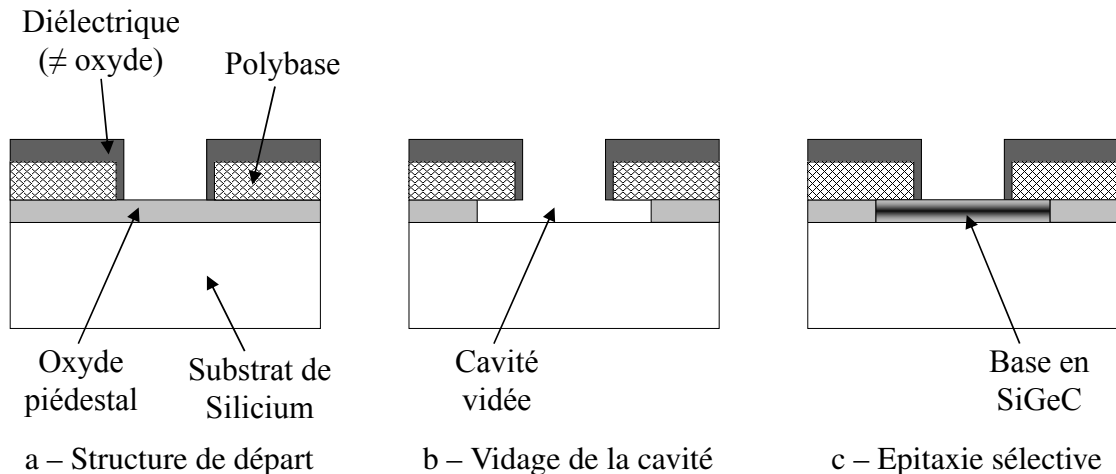


FIG. II.8 – Epitaxie sélective de la base : Une cavité est aménagée dans l'oxyde piédestal isolant le substrat du Polybase, puis le SiGe est déposée dans cette cavité, sélectivement par rapport aux diélectriques.

II.3.3.c.i Couche enterrée – Isolation

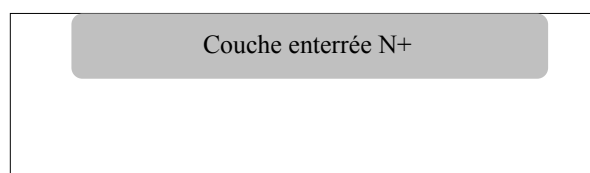
La fabrication commence par la réalisation de la couche enterrée N+ sur un substrat massif (figure II.9-a) par implantation d'arsenic, puis une épitaxie d'épaisseur 0.3 μm est réalisée. Ensuite, on passe dans le module DTI (*Deep Trench Isolation*). Les tranchées profondes sont gravées, puis remplies par de l'oxyde et du polysilicium. Le polysilicium est nécessaire pour absorber les contraintes mécaniques induites par l'oxyde des tranchées. Les tranchées ont une largeur de 1 μm , et l'oxyde a une épaisseur de 200 nm. Ensuite, on délimite les zones actives par le module STI (*Shallow Trench Isolation*). Les tranchées, d'une profondeur d'environ 0.3 μm , sont gravées puis remplies par de l'oxyde (figure II.9-b). Le puits collecteur est réalisé par implantation de phosphore.

II.3.3.c.ii Base extrinsèque – Fenêtre émetteur

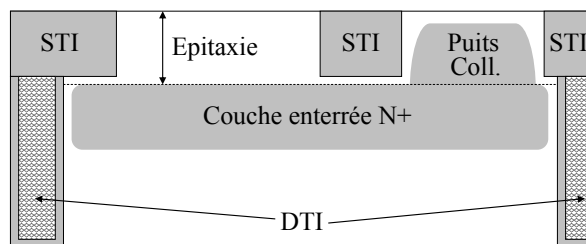
La fabrication continue par le dépôt de l'oxyde piédestal. Celui-ci est un oxyde à faible densité déposé par LPCVD (CVD basse pression), dont le précurseur est le TEOS (Tetra-EthOxySilane). Sa faible densité lui assure un dépôt rapide, ainsi que des vitesses de gravure élevées. L'épaisseur du piédestal est de 50 nm, c'est dans cette épaisseur que sera réalisée la base intrinsèque. On dépose ensuite le polysilicium de base d'épaisseur 100 nm, et le dopage P+ est assuré par implantation ionique de bore. La base est ensuite délimitée par une photo suivie d'une gravure sèche, et l'ensemble est recouvert de 20 nm de TEOS et de 60 nm de nitrure, qui servira à isoler les polysilicium de base et d'émetteur. La fenêtre émetteur est réalisée par une étape de photolithographie suivie d'une gravure sèche, s'arrêtant sur l'oxyde piédestal. La largeur de la fenêtre est de 0.27 μm . Le collecteur implanté sélectivement (*Selectively Implanted Collector* ou SIC) est implanté à travers la fenêtre émetteur, et est donc auto-aligné par rapport à cette ouverture. L'implantation du SIC sert à doper localement le collecteur, et assurer le contact avec la couche enterrée. Cette implantation peut être faite avec du Phosphore ou de l'Arsenic (figure II.9-c).

II.3.3.c.iii Base intrinsèque

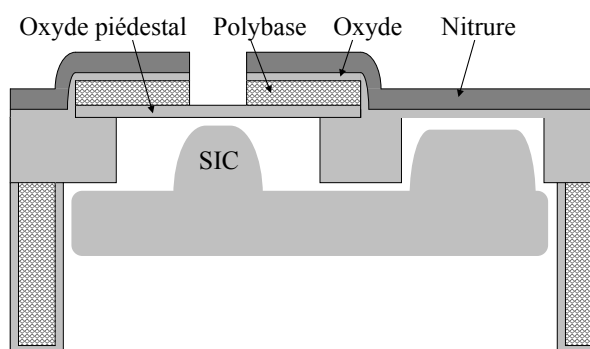
Après l'ouverture de la fenêtre émetteur, des espaceurs en nitrure sont réalisés, pour protéger les flancs du polysilicium de base lors de l'épitaxie. Pour réaliser ces espaceurs, une couche de



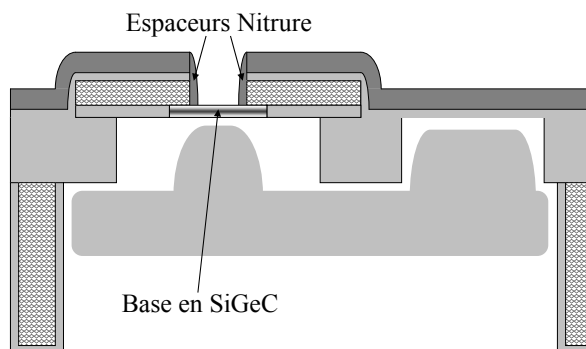
a – Implantation de la couche enterrée.



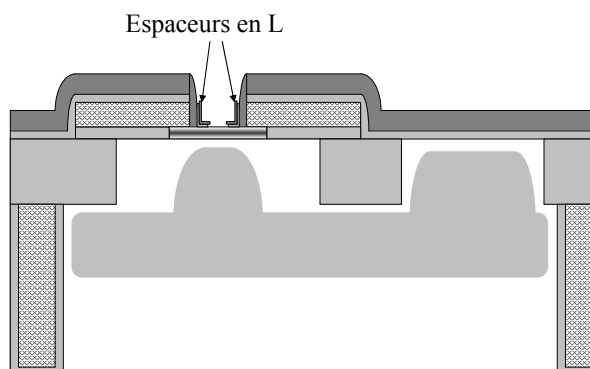
b – Reprise d'épitaxie, modules DTI et STI, implantation du puits collecteur.



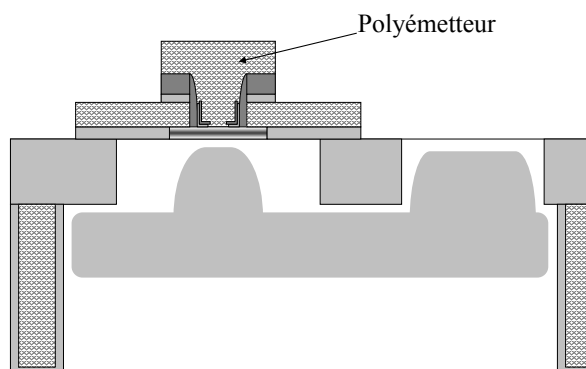
c – Dépôt/implantation de la base extrinsèque, ouverture de la fenêtre émetteur.



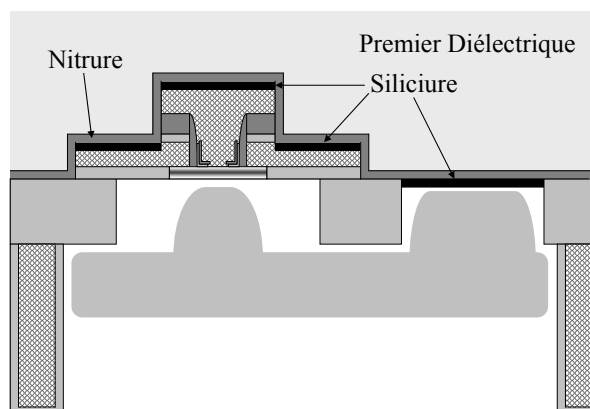
d – Espaceurs nitrure, épitaxie sélective de la base.



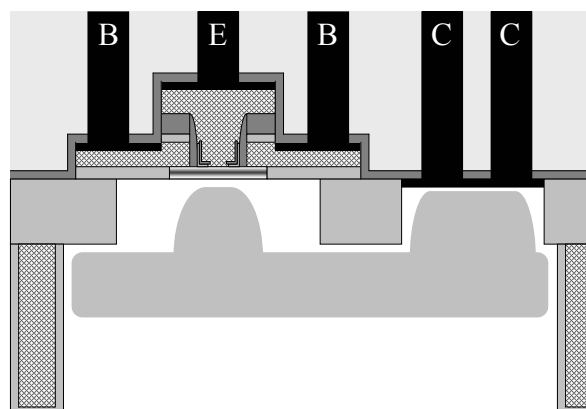
e – Réalisation des espaceurs internes en L.



f – Dépôt et gravure du polyémetteur.



g – Siliciuration, dépôt nitrure pleine plaque, dépôt du premier niveau de diélectrique.



h – Réalisation des contacts métalliques.

FIG. II.9 – Étapes de fabrication du transistor bipolaire

30 nm de nitrure est déposée par LPCVD de manière conforme, puis on procède à une gravure sèche anisotrope, de manière à graver la couche déposée. La plaque est alors entièrement recouverte d'une épaisse couche de nitrure, sauf dans le fond de la fenêtre émetteur où l'oxyde piédestal est apparent. Ensuite, on procède au dépôt de la base intrinsèque : L'oxyde piédestal est gravé par une solution d'acide fluorhydrique pour aménager la cavité de la base. La gravure progresse latéralement, consommant l'oxyde sous le polybase. La cavité de la base est auto-alignée par rapport à la fenêtre émetteur, car l'attaque chimique est isotrope.

Avant de réaliser l'épitaxie, on procède à un nettoyage du substrat par une attaque en milieu humide que l'on appelle HF-Last. Il s'agit d'une attaque non-oxydante par de l'acide fluorhydrique laissant la surface du silicium à nu. En effet, une couche d'oxyde natif à tendance à se former à la surface du silicium laissé à l'air libre, il faut donc veiller à ce que le temps entre la fin du nettoyage et l'entrée dans la chambre d'épitaxie soit le plus court possible. Juste avant le dépôt, on procède à un recuit à 900°C sous H_2 pendant quelques minutes, de manière à désorber les éventuelles traces d'oxyde ou d'impuretés présentes sur la surface du silicium. Lors de l'épitaxie, le SiGe croît sélectivement par rapport au nitrure (figure II.9-d). La croissance de la base est monocristalline sur le substrat et polycristalline sur le polybase, le lien entre base intrinsèque et extrinsèque est assuré correctement, comme le montre la figure II.10.

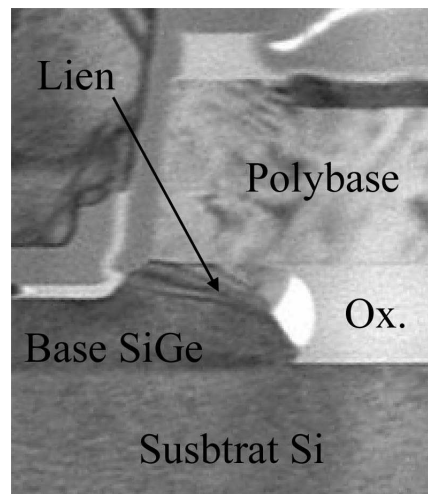


FIG. II.10 – Observation TEM du lien entre base intrinsèque et base extrinsèque. Le lien s'effectue correctement grâce à la consommation de l'oxyde piédestal sous le polybase.

II.3.3.c.iv Espaceurs – Module émetteur

Après la base en SiGe, on réalise des espaceurs en L à l'intérieur de la fenêtre émetteur (figure II.9-e). Ces espaceurs sont réalisés en déposant un empilement composé d'oxyde, de nitrure et d'oxyde, puis en gravant cet empilement par une gravure sèche. La largeur des espaceurs est donc fixée par l'épaisseur de la dernière couche d'oxyde. Ensuite, le polyémetteur dopé in-situ est déposé par RTCVD, d'épaisseur 150 nm. Il est déposé de façon polycristalline, et un réarrangement s'opère grâce au budget thermique élevé, permettant ainsi d'obtenir un monocristal dans le bas de l'émetteur. Une étape de photolithographie sert à délimiter l'extension latérale de l'émetteur, et permet de libérer le polybase (figure II.9-f).

II.3.3.c.v Recuit d'activation – Siliciuration – Back-End

La fabrication du composant se poursuit par le dépôt d'un masque dur composé d'oxyde et de nitrure, afin de protéger les zones ne devant pas être siliciurées. Le TBH étant entièrement

siliciuré, le masque dur est déposé, puis le nitrure est gravé dans la zone correspondant au transistor bipolaire. Le recuit d'activation est ensuite réalisé. Il s'agit d'un recuit très rapide à une température de 1080°C pour les transistors étudiés dans cette thèse.

La siliciuration s'opère ensuite par un procédé de dépôt de cobalt, qui réagit en diffusant dans le silicium (figure II.9-g). La diffusion est contrôlée par un recuit thermique adapté. Ceci permet d'assurer des contacts métal-silicium les moins résistifs possibles en assurant une transition douce entre les deux matériaux. Après siliciuration, on dépose une couche de nitrure pleine plaque servant à encapsuler les composants actifs pour les protéger des métaux introduits dans les interconnexions. Le cuivre en particulier est une espèce qui diffuse très rapidement dans le silicium, et induit des défauts profonds dans la bande interdite du silicium, délétères pour les composants.

Le premier niveau de diélectrique est déposé. Il s'agit d'un oxyde fluoré (FSG pour Fluoro-Silicon Glass) peu dense, déposé par PECVD (CVD assistée par plasma). L'ensemble est aplani par CMP, puis les trous de contacts sont définis par photolithographie. La gravure s'arrête dans le siliciure des contacts. Ensuite, on dépose une couche de 10 nm de titane par pulvérisation cathodique, 20 nm de TiN par CVD et le reste des trous de contact est rempli par du tungstène déposé par CVD (figure II.9-h). Le dépôt est ensuite aplani par CMP, et la fabrication continue dans le Back-end pour réaliser les 6 niveaux d'interconnexion en cuivre.

La figure II.11 présente une observation au microscope électronique à transmission de la structure FSA-SEG ainsi réalisée.

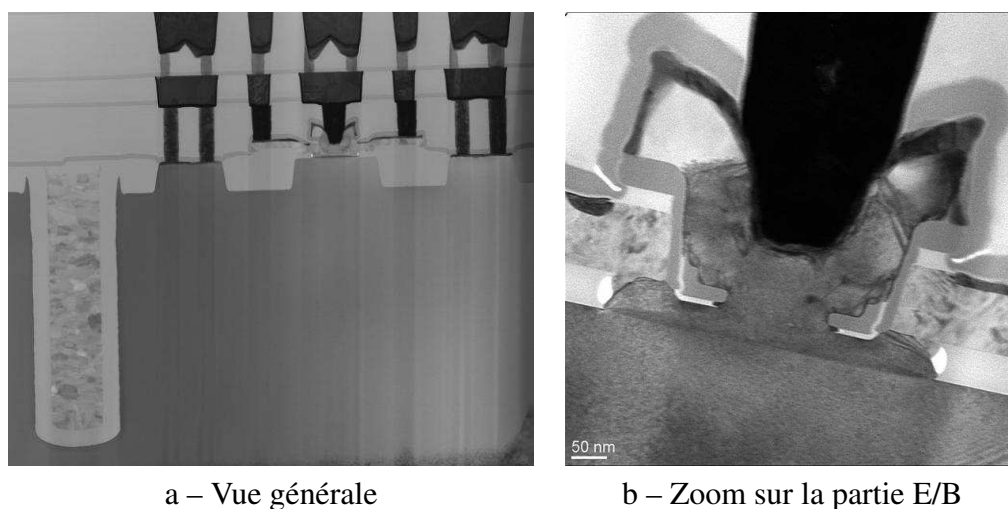


FIG. II.11 – Observation TEM de la structure FSA-SEG achevée.

II.3.4 Mise en place de la technologie BiCMOS

II.3.4.a Enjeux de l'intégration

Après avoir détaillé les opérations successives nécessaires à la réalisation du transistor avec la structure FSA, nous présentons l'intégration des procédés de fabrication dans la route complète. En effet, le but est d'intégrer ces composants dans une filière BiCMOS, c'est à dire présentant sur une même plaque des transistors MOS à effet de champ et des transistors bipolaires. Les contraintes sont importantes, car le budget thermique du MOS est élevé, et les TBH sont dégradés par ces recuits. L'épaisseur et le dopage de la couche de base sont en effet déterminants pour les performances des transistors bipolaires, et un trop fort budget thermique aurait pour conséquence des bases trop larges et peu dopées, donc peu performantes.

A l'inverse, les différents dépôts nécessaires pour la réalisation des TBH ont des températures non négligeables, qui peuvent être source de diffusion et de désactivation des espèces dopantes dans les transistors MOS, ce qui dégrade les performances des composants. Il faut donc trouver un schéma d'intégration permettant de réaliser de manière optimale les deux types de composants.

II.3.4.b Schéma d'intégration BiCMOS

Le schéma d'intégration retenu pour l'intégration des transistors bipolaires dans une technologie BiCMOS est donné dans la figure II.12 [Boissonnet06]. Les étapes spécifiques au système émetteur-base sont réalisées en une seule fois, entre le dépôt et la gravure des grilles des MOS. C'est également le schéma d'intégration retenu pour la technologie BiCMOS millimétrique, présentée en fin de chapitre III.

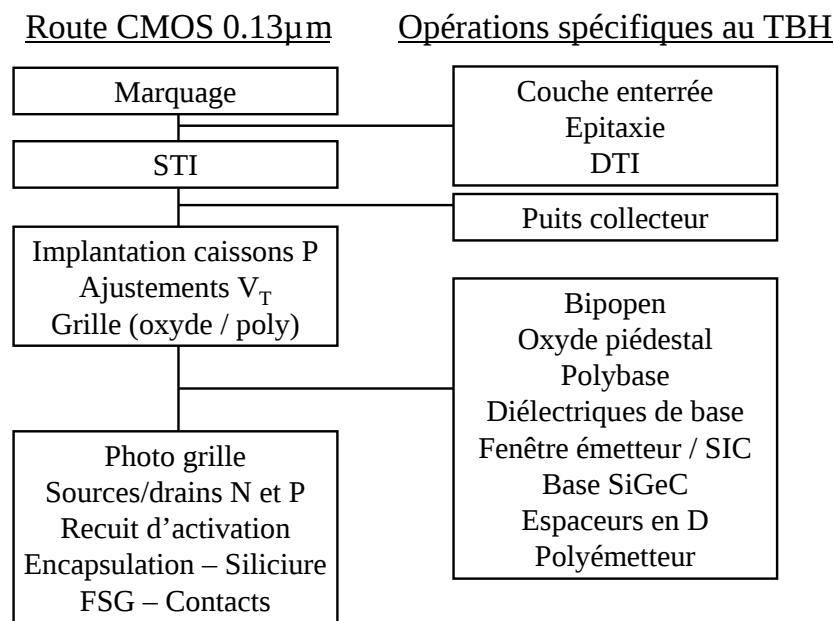


FIG. II.12 – Intégration des étapes spécifiques au TBH dans la route CMOS 0.13 µm [Boissonnet06].

II.4 Comparaison par rapport à la concurrence

La structure FSA-SEG a connu aux cours des dernières années des progrès importants, grâce notamment aux études présentées dans ce travail de thèse [Chevalier03] [Chevalier04] [Chevalier05a] [Chevalier06b]. Les technologies de réalisation des transistors bipolaires varient d'un fabricant à l'autre, et offrent une gamme de performances très vaste, adaptables en fonction de l'utilisation recherchée. Ce chapitre vise à recenser les dernières publications d'autres fabricants de semi-conducteurs dans le monde, et présenter les architectures choisies, ainsi que les performances obtenues.

II.4.1 NXP (anciennement Philips)

Le hollandais Philips s'est récemment séparé de sa branche semi-conducteur, désormais appelée NXP. La technologie BiCMOS proposée par NXP repose sur le noeud 0.25 μm , avec 5 à 6 niveaux d'interconnexion métalliques réalisées en Aluminium. L'architecture retenue pour le technologie BiCMOS QUBiC4G est une structure QSA à simple polysilicium [Deixler02]. Les performances des transistors bipolaires sont de $f_T/f_{MAX} = 70/100$ GHz pour une tension BV_{CEO} de 1.8 V. Les performances des composants ont pu être poussées jusqu'à des fréquences f_T de 157 GHz en optimisant le profil vertical [Donkers03]. Pour la technologie suivante nommée QUBiC4X, les fréquences obtenues sont 130/140 GHz en f_T/f_{MAX} , grâce notamment à l'ajout d'un polysilicium pour la base extrinsèque permettant de réduire la résistance d'accès [Deixler04]. Les tensions de claquage BV_{CEO} et BV_{CBO} atteignent 2 V et 9 V respectivement.

Dans l'article sur l'émetteur métallique [Donkers04], des siliciures de cobalt et de nickels sont comparés. La structure utilisée est une structure auto-alignée utilisant le procédé damascène. Un schéma de la structure utilisée est reportée dans la figure II.13.

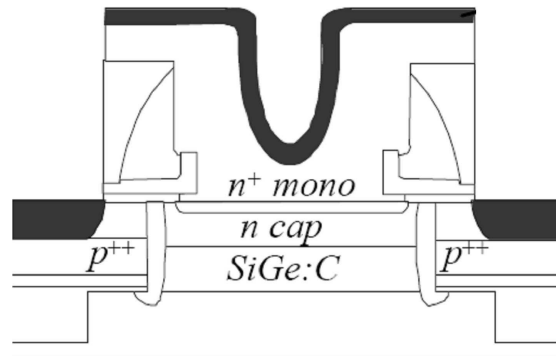


FIG. II.13 – Schéma de la structure du TBH réalisé par NXP [Donkers03].

On passe de $f_T/f_{MAX} = 190/60$ GHz avec un siliciure cobalt à $f_T/f_{MAX} = 200/150$ GHz avec le nickel. Dans le même temps, BV_{CEO} augmente de 1.49 à 1.57 V. On note que les faibles performances obtenues en f_{MAX} ne reflètent pas tous les avantages qu'apporte une structure FSA.

II.4.2 IMEC

L'IMEC est un laboratoire qui étudie également les transistors bipolaires rapides. Leur choix s'est porté sur une architecture QSA à simple polysilicium, non auto-alignée. Les performances des TBH réalisés sont néanmoins excellentes, puisqu'ils atteignent des fréquences de transition

supérieures à 200 GHz. En effet des f_T/f_{MAX} de 205/160 GHz ont été démontrées avec une tension BV_{CEO} de 1.75 V, dans le noeud CMOS 0.13 μm [Huylenbroeck04]. Il s'agit d'une technologie intégrant uniquement des TBH pour les besoins de l'étude, mais la compatibilité BiCMOS est dite maintenue. L'isolation est réalisée par des DTI et des STI.

[Piontek06] introduit des « Airgap DTI », soit des tranchées profondes d'isolations complètement vides, et démontre par la même occasion une résistance de base réduite d'un facteur 3, ce qui permet d'atteindre des valeurs de f_{MAX} de 275 GHz, en gardant la même valeur de f_T . Dans [Huylenbroeck06], BV_{CEO} vaut 1.85 V et BV_{CBO} 6.2 V. La figure II.14 propose une vue en coupe de la structure réalisée.

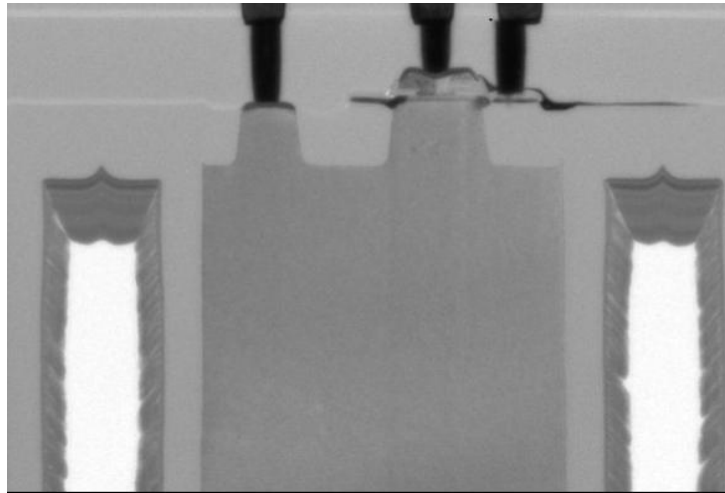


FIG. II.14 – Vue TEM du transistor bipolaire de l'IMEC [Huylenbroeck06]. Les DTI vides sont bien visibles (en blanc).

L'obtention de fréquences aussi élevées, notamment f_{MAX} , impose un contrôle très précis de l'alignement des masques les uns par rapport aux autres. Ceci est possible pour le prototypage de composants lors du développement des technologies, mais doit être sensiblement plus délicat lorsqu'il faut passer en production de volume.

II.4.3 IHP

L'IHP a fait le choix d'une structure faible coût, c'est à dire sans isolation par DTI, et avec un collecteur implanté. L'architecture du transistor bipolaire est une architecture à simple polysilicium non auto-alignée, intégrée dans un noeud CMOS 0.25 μm doté de 4 niveaux d'interconnexions métalliques. Les performances obtenues sont 200/170 GHz pour f_T/f_{MAX} , avec une tension BV_{CEO} de 2 V [Heinemann02]. La technologie complémentaire décrite dans [Heinemann03] démontre des transistors NPN à 180 GHz aussi bien pour f_T que pour f_{MAX} , en gardant BV_{CEO} à 2 V.

Ce type de transistor offre des performances améliorées lorsqu'on le place dans une technologie SOI 90 nm [Rucker04] : f_T et f_{MAX} augmentent jusqu'à 220 et 230 GHz, avec des tensions BV_{CEO} et BV_{CBO} de 2 V et 5.8 V respectivement. Le transistor bipolaire n'est pas véritablement un composant déposé sur SOI, car dans les zones de fabrication du TBH, l'oxyde enterré est gravé pour permettre le fonctionnement vertical du dispositif.

Des performances records ont été établies avec cette structure de transistor bipolaire en intégrant uniquement des transistors bipolaires sur la plaque de silicium [Heinemann04]. Le budget thermique est modifié par rapport à une technologie BiCMOS, le nombre de niveaux d'interconnexion est réduit à 2, et le dessin du collecteur est optimisé. De cette manière, des

performances de $f_T/f_{MAX} = 300/250$ GHz avec $BV_{CEO} = 1.8$ V, ou 380/190 GHz avec 1.5 V de BV_{CEO} sont possibles. Il s'agit de la plus forte valeur de f_T enregistrée pour un TBH SiGe à température ambiante. Cette forte valeur de f_T est possible grâce à une forte réduction de l'épaisseur de base, ce qui dégrade la résistance d'accès à la base, et donc f_{MAX} . Cependant, ces performances ne sont pas compatibles avec la fabrication complète des transistors MOS.

II.4.4 Jazz

Jazz propose une technologie BiCMOS 0.18 μm ou 0.13 μm . Le transistor bipolaire présenté possède une structure complètement auto-alignée grâce à un procédé à émetteur sacrificiel avec utilisation d'un contre-masque émetteur [Racanelli01b]. Les performances des TBH sont 205/210 GHz de f_T et f_{MAX} , pour une tension BV_{CEO} de 1.8 V, dans une technologie 0.13 μm avec isolation par DTI [Racanelli03]. La figure II.15 présente une observation TEM du transistor réalisé par Jazz.

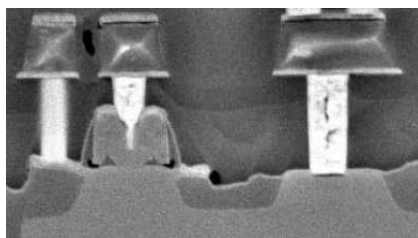


FIG. II.15 – Vue TEM du transistor bipolaire de Jazz [Racanelli01b].

II.4.5 Freescale (anciennement Motorola)

L'américain Freescale, séparé de Motorola, a également récemment effectué la conversion de sa technologie vers une structure complètement auto-alignée à épitaxie sélective [John06]. A ce titre, les performances des transistors bipolaires intégrés dans un noeud CMOS 0.18 μm présentent des f_T de 185 GHz et des f_{MAX} de 260 GHz, avec des tensions de claquage de BV_{CEO} de 2 V et BV_{CBO} de 6.2 V. L'isolation est assurée par des STI, et le contact collecteur sous le STI n'est pas réalisé par une couche enterrée, mais par un module innovant appelé SIBL (Sub-Isolation Buried Layer), les tranchées profondes d'isolation ne sont pas utilisées.

II.4.6 Hitachi

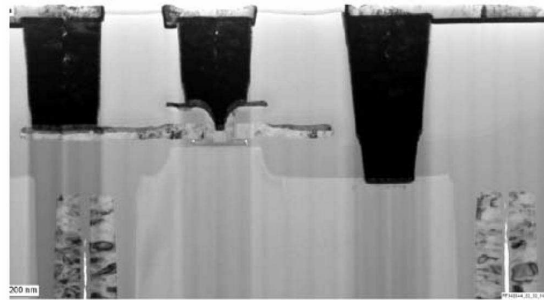
La structure complètement auto-alignée à épitaxie sélective de la base est également employée par Hitachi. Dans une technologie BiCMOS 0.18 μm , avec isolation par DTI et STI, les performances sont 140 GHz pour f_T et 183 GHz pour f_{MAX} , BV_{CEO} valant autour de 2 V [Wada02, Washio02].

Le profil vertical a par la suite été optimisé par le procédé PED (Promoting Emitter Diffusion) : Le recuit final est adapté afin de favoriser la diffusion des dopants de l'émetteur vers la base, ce qui permet d'obtenir des bases plus fines. L'épitaxie de SiGe devient moins critique, puisqu'il n'est plus nécessaire de déposer des couches très fines et bien contrôlées. On peut néanmoins s'interroger sur la reproductibilité d'un tel procédé. Avec une épaisseur de SiGe de 30 nm et un polyémetteur dopé phosphore les performances sont supérieures à celles obtenues précédemment : On démontre ainsi des couples f_T/f_{MAX} de 202/210 GHz, 220/150 GHz et

230/140 GHz [Miura04]. Le procédé PED est également employé dans [Miura06] présentant des f_T et f_{MAX} de 205 et 230 GHz.

II.4.7 Infineon

Infineon a également fait le choix d'une technologie auto-alignée utilisant l'épithaxie sélective de la base pour réaliser ses transistors bipolaires. Dans [Meister03], les performances des composants sont 206/197 GHz pour f_T/f_{MAX} , avec une tension BV_{CEO} de 1.8 V et une tension BV_{CBO} de 5.8 V. La base SiGe a une épaisseur de 40 nm, et le polyémetteur est déposé de manière monocristalline. En réduisant significativement la largeur de la fenêtre émetteur (de 0.18 à 0.14 μm), des fréquences de transition de 200/275 GHz ont pu être atteintes, la tension BV_{CEO} diminuant légèrement (1.7 V) et BV_{CBO} est maintenue constante [Böck04a]. La structure réalisée est présentée dans la figure II.16.



a – Vue générale.



b – Zoom sur la partie émetteur/base.

FIG. II.16 – Vue au microscope électronique à transmission du transistor bipolaire d'Infineon [Böck04a].

Une réduction de l'épaisseur de base intrinsèque de 40 nm à 30 nm a permis d'augmenter f_T et f_{MAX} jusqu'à 225 et 300 GHz respectivement, et BV_{CEO} jusqu'à 1.8 V, sans dégradation des résistances d'accès [Böck04b]. Enfin, [Vytlá06] présente une technologie intégrant plusieurs types de composants haute vitesse ou haute tension sur une même plaque. Le TBH rapide possède des caractéristiques légèrement dégradées : $f_T = 209$ GHz, $f_{MAX} = 237$ GHz et $BV_{CEO} = 1.8$ V, pour une tension d'avalanche BV_{CBO} de 6 V.

II.4.8 IBM

IBM possède une avance importante sur les autres acteurs du semiconducteur et a été le premier à dépasser la barre symbolique des 300 GHz. Dès 2002, des performances atteignant 207/285 GHz pour f_T/f_{MAX} ont été démontrées [Jagannathan02]. BV_{CEO} et BV_{CBO} valent respectivement 1.7 et 5.5 V. La technologie employée est une technologie bipolaire utilisant le noeud 0.18 μm , avec isolation par DTI et STI. La structure du transistor bipolaire est complètement auto-alignée, avec une épitaxie non sélective de la base, grâce au procédé inverse-emetteur réalisé par damascène. L'émetteur est dopé in-situ avec du phosphore, et la base extrinsèque est ré-épitaxiée de manière auto-alignée afin de diminuer la résistance de base, permettant les fortes valeurs de f_{MAX} ainsi obtenues. Cette structure intégrée dans une technologie BiCMOS 0.13 μm présente les mêmes performances [Orner03]. La figure II.17 présente une coupe du transistor bipolaire réalisé par IBM observée au microscope électronique à balayage.

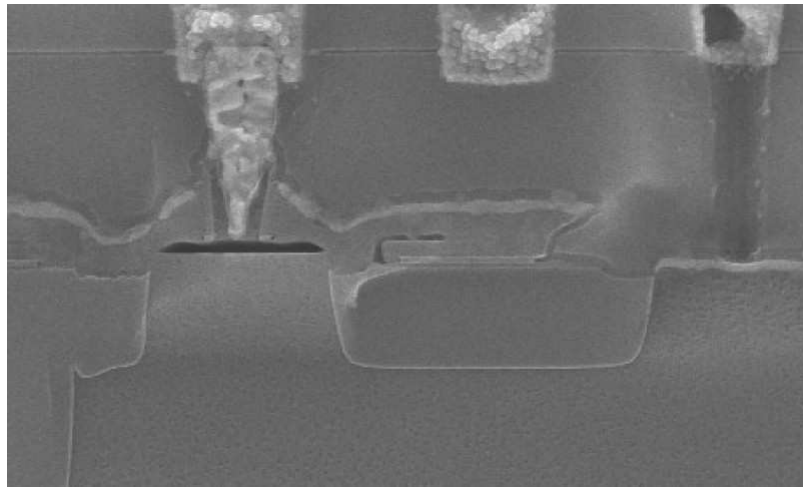


FIG. II.17 – Vue au microscope électronique à balayage du transistor bipolaire d'IBM [Orner06].

Les évolutions de cette structure sont nombreuses : Dans [Jagannathan03], la réduction de résistance de base permet d'atteindre des f_{MAX} de 338 GHz, au détriment de f_T , qui diminue jusqu'à 180 GHz. A l'inverse, une réduction de l'épaisseur de la base intrinsèque offre des performances de 350 GHz de f_T , avec néanmoins une valeur de f_{MAX} fortement dégradée de 185 GHz. Ces performances sont légèrement en retrait par rapport aux 380/190 GHz démontrés par l'IHP dans [Heinemann04].

Par la suite, f_T et f_{MAX} sont optimisées conjointement pour atteindre 302/306 GHz respectivement, BV_{CEO} valant 1.6 V et BV_{CBO} 5.5 V [Rieh04a]. Des réglages de collecteur différents permettent d'obtenir des performances diverses. Ainsi [Rieh04b] démontre des valeurs de f_T/f_{MAX} valant 310/295 GHz, avec BV_{CEO} et BV_{CBO} valant respectivement 1.6 et 5.4 V. Enfin la performance record de $f_{MAX} = 350$ GHz est démontrée dans [Khater04], pour une valeur de f_T de 300 GHz, BV_{CEO} et BV_{CBO} valant 1.7 et 5.6 V. Ces performances sont obtenues grâce à une réduction de la largeur de fenêtre émetteur jusqu'à 0.12 μm .

[Orner06] démontre l'intégration dans une technologie BiCMOS 0.13 μm de transistors bipolaires ayant comme performances 300/330 GHz de f_T/f_{MAX} pour une tension BV_{CEO} de 1.5 V, ce qui est à ce jour la meilleure performance obtenue pour une technologie BiCMOS.

II.4.9 Synthèse

Cette étude a permis de présenter les diverses architectures utilisées pour réaliser des transistors bipolaires SiGe rapides. La figure II.18 résume les valeurs de f_T et f_{MAX} obtenues, en proposant une comparaison avec les meilleures performances obtenues par ST et ayant fait l'objet de ce travail de thèse. Les résultats de ST semblent en retrait par rapport aux meilleurs points d'IBM, mais un grand nombre de publications provenant d'IBM reportent des technologies bipolaires seulement, et non des composants ayant subi tout le budget thermique d'une route BiCMOS. Les meilleures performances BiCMOS sont néanmoins à ce jour détenues par IBM (300/330 GHz).

Ainsi les technologies compatibles BiCMOS développées à ST se positionnent très favorablement aux côtés des technologies bipolaires d'IBM, d'Infineon et de l'IHP.

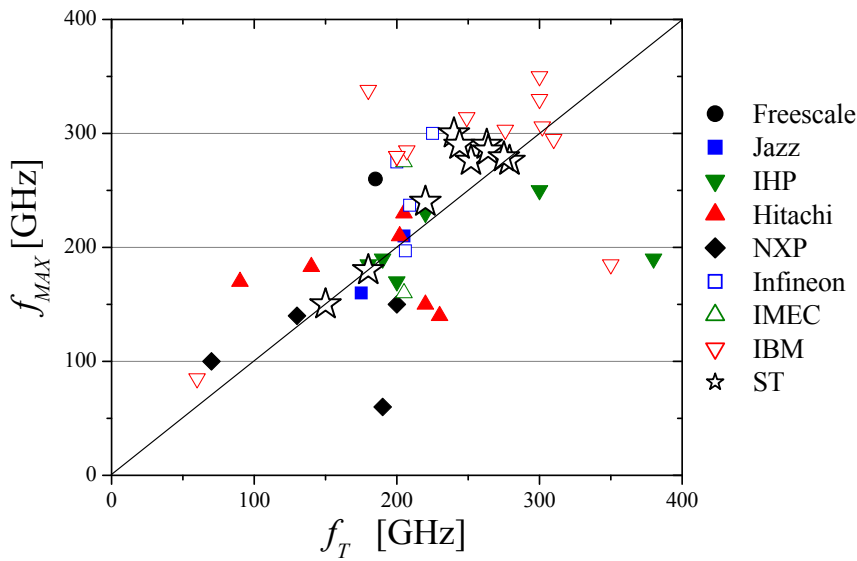


FIG. II.18 – Valeurs de f_T et f_{MAX} obtenues par les différents concurrents entre 2002 et 2006.

De cette analyse, il ressort que pour atteindre des fortes valeurs de f_{MAX} (supérieures à 250 GHz), une technologie complètement auto-alignée s'avère nécessaire. En effet, seul un auto-alignement E/B permet de réduire suffisamment les éléments extrinsèques que sont la résistance de base et la capacité B/C. Cet auto-alignement n'est pas nécessaire pour obtenir de fortes valeurs de f_T , qui dépend principalement du profil vertical émetteur/base/collecteur, comme le prouvent les résultats de l'IHP.

Ces différentes remarques sont le fruit des réflexions menées dans la suite de l'étude. Notamment, l'optimisation conventionnelle du transistor bipolaire développée au chapitre III présente successivement les améliorations réalisées sur le profil vertical et latéral du composant, afin d'optimiser conjointement f_T et f_{MAX} .

Chapitre III

Optimisation conventionnelle du TBH

III.1 But de l'optimisation verticale et latérale

L'historique des technologies réalisé au chapitre II a démontré la nécessité d'optimiser aussi bien le profil vertical du TBH que les dimensions latérales du composant, pour augmenter conjointement f_T et f_{MAX} . Ces considérations sont reprises dans [Rieh05b]. Pour obtenir une forte fréquence f_T il faut proposer une structure très compacte, permettant de minimiser les temps de transit des porteurs dans chaque partie du composant. La fréquence f_{MAX} dépendant fortement des retards liés aux capacités et résistances d'accès, il est donc nécessaire de rapprocher au maximum les contacts de base et de collecteur pour minimiser R_B et C_{BC} . La structure FSA-SEG décrite au chapitre précédent est une des architectures permettant de réduire significativement les dimensions latérales et verticales du composant.

Ce chapitre présente les optimisations réalisées sur les dimensions verticales du dispositif, en améliorant successivement les profils d'émetteur, de base, puis de collecteur. Il se poursuit par l'étude des améliorations apportées par la diminution de la fenêtre émetteur et des capacités de recouvrement. Nous réalisons ensuite la synthèse des optimisations réalisées, et l'évolution des performances des transistors bipolaires obtenues pendant ce travail de thèse est présentée. La fin du chapitre aborde les possibilités d'intégration des TBH ayant bénéficié de cette optimisation : Des performances de blocs de circuit à l'état de l'art sont présentées. Une filière nouvelle dédiée aux applications millimétriques vient d'être créée, utilisant les composants mis au point par les procédés détaillés dans ce chapitre.

III.2 Optimisation du module émetteur

L'émetteur est une partie très importante du transistor bipolaire. De son optimisation dépendent largement les paramètres statiques et dynamiques du composant. Avec un niveau de germanium modéré dans la base du côté émetteur, le courant de base dépend pour grande part des paramètres de l'émetteur, et la tenue en tension BV_{CEO} leur est également intimement liée. La résistance statique de l'émetteur provoque une chute de tension aux bornes de la jonction E/B, et pénalise également la fréquence de transition f_T . Le temps de transit émetteur dépend des trous injectés dans l'émetteur, il faut donc tenir compte de la vitesse de recombinaison des trous dans l'émetteur afin de diminuer la charge correspondante.

Le profil de dopage de l'émetteur est schématisé dans la figure III.1. Le polyémetteur fortement dopé in-situ est une source de diffusion des impuretés, qui progressent de l'émetteur vers la base sous l'influence des différents traitements thermiques. Afin de tenir compte de cet effet, une couche de silicium non dopée appelée Si-cap (Silicon Capping) est déposée au dessus de la

base SiGeC lors de l'épitaxie, et sert à espacer les matériaux constituant la base et l'émetteur.

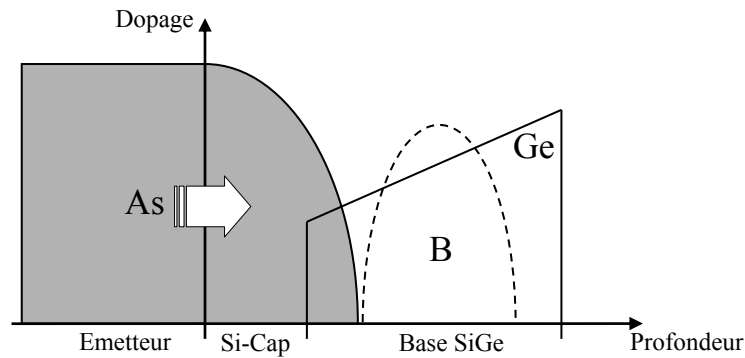


FIG. III.1 – Présentation du profil de dopage de l'émetteur.

Le principal paramètre électrique représentatif de l'émetteur est la résistance d'émetteur, R_E . Cette résistance est une résistance statique, composée par la mise en série de plusieurs contributions, détaillées dans la figure III.2. Cette décomposition de la résistance d'émetteur est prise en compte dans le modèle HiCUM (High Current Model) [Schröter00].

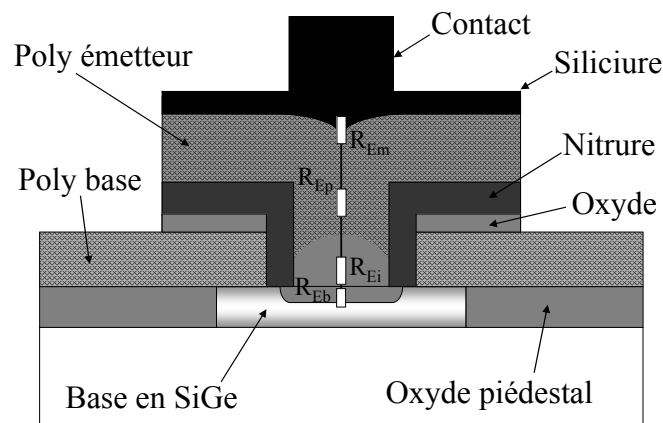


FIG. III.2 – Représentation des différentes parties de l'émetteur contribuant à R_E (d'après [Schröter00]).

Nous pouvons grâce à cette décomposition identifier les différents paramètres technologiques ayant une influence sur la résistance totale d'émetteur :

- La résistance R_{Eb} est la contribution due à la zone de diffusion des dopants de l'émetteur vers la base.
- L'interface entre l'émetteur et la base présente une résistance notée R_{Ei} . Cependant, l'amélioration des procédés de dépôt et de nettoyage a permis de diminuer sensiblement cette résistance, de fait elle sera négligée dans cette étude.
- Le polysilicium de l'émetteur contribue à la résistance R_{Ep} , qui dépend principalement du dopage et de l'épaisseur du polyémetteur.
- Les niveaux d'interconnexions métalliques. Ceux-ci comprennent le siliciure, le contact en tungstène, et les différents niveaux de cuivre composant le back-end. La contribution des niveaux de métaux à la résistance d'émetteur est notée R_{Em} .

Il est possible d'agir sur les 3 premières contributions en modifiant des conditions de dépôt de l'émetteur et de la base. L'épaisseur de polyémetteur déposée a une influence sur la contribution R_{Ep} et R_{Em} , en modifiant la forme du contact. Le dessin du contact métallique a également une influence forte sur R_{Em} .

III.2.1 Mise en place d'une méthode d'extraction de R_E

Afin d'étudier l'évolution de la résistance d'émetteur, nous avons besoin d'une méthode de mesure fiable et facile à mettre en œuvre. Compte tenu de la disponibilité des bancs de mesure, nous avons opté pour une méthode statique, ne nécessitant qu'un simple tracé de Gummel, et la connaissance d'aucun paramètre de modèle. Lorsque les courants du transistor augmentent, les chutes de tensions dans les résistances séries prennent de plus en plus d'importance. En observant l'écart à une dépendance exponentielle du courant collecteur avec la tension V_{BE} , on en déduit la valeur de R_E .

III.2.1.a Méthode basée sur des mesures statiques

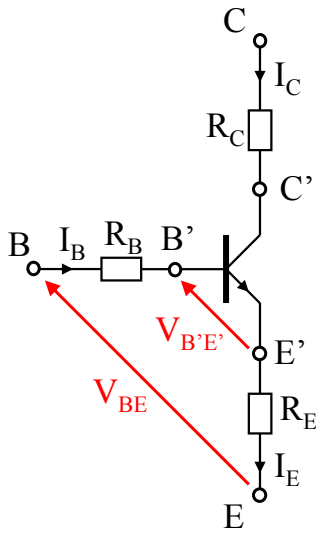


FIG. III.3 – Résistances séries du transistor bipolaire.

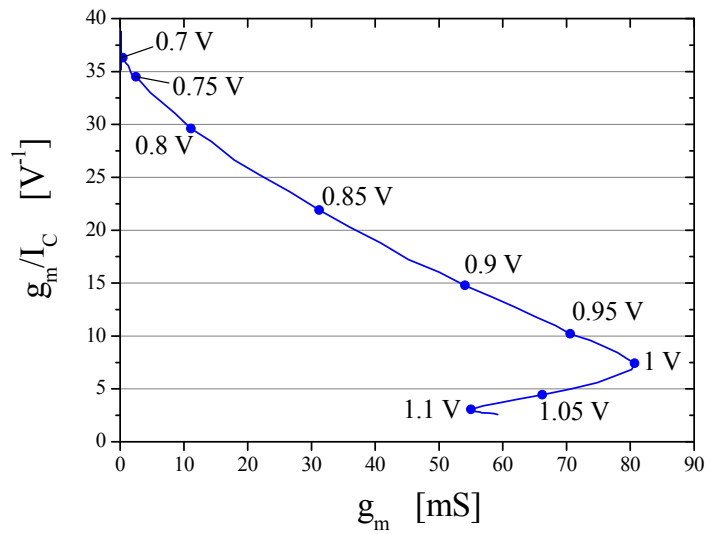


FIG. III.4 – Évolution de g_m/I_C en fonction de g_m pour l'extraction de la résistance d'émetteur. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$.

La figure III.3 représente les résistances séries ajoutées au transistor bipolaire idéal. La tension intrinsèque de la jonction E/B est notée $V_{B'E'}$. Sous l'effet des résistances séries, la tension $V_{B'E'}$ appliquée à la jonction E/B n'est pas la même que celle appliquée aux bornes du composant. D'après la figure III.3, on en déduit :

$$V_{BE} = V_{B'E'} + R_B I_B + R_E I_E \quad (\text{III.1})$$

A cause du très grand gain en courant des composants mesurés, on peut écrire $I_E \approx I_C = \beta I_B$, et on peut donc négliger $R_B I_B$ devant $R_E I_C$, la résistance de base n'étant pas trop élevée. Le courant I_C dépend uniquement de la polarisation aux bornes de la jonction émetteur/base. D'après les équations I.24 et III.1, on peut écrire :

$$I_C = \frac{q A_e}{G_B} \exp \frac{q V_{B'E'}}{kT} = \frac{q A_e}{G_B} \exp \frac{q (V_{BE} - R_E I_C)}{kT} \quad (\text{III.2})$$

La dérivée logarithmique de III.2 par rapport à V_{BE} donne, avec $g_m = dV_{BE}/dI_C$:

$$\frac{g_m}{I_C} = \frac{q}{nkT} (1 - R_E g_m) \quad (\text{III.3})$$

Dans la zone linéaire (en dessous de 0.9 V), le tracé de g_m/I_C en fonction de g_m conduit à une droite dont la pente dépend directement de la résistance d'émetteur R_E . On peut noter que l'ordonnée à l'origine vaut q/nkT . Cette méthode permet donc également d'extraire le facteur d'idéalité du courant « émetteur » n , à condition de connaître la température du transistor. Par conséquent, on cherche à réaliser l'extraction de R_E dans une zone où l'auto-échauffement du composant est limité, typiquement en dessous de $V_{BE} = 0.9$ V.

A partir d'une régression linéaire dans la gamme de tension choisie, on extrait la valeur de q/nkT , et celle de R_E . On commet une erreur en utilisant la valeur de q/kT théorique au lieu de q/nkT , car le facteur d'idéalité du composant n'est pas rigoureusement égal à 1. En fonction de la plage de tension V_{BE} choisie pour le calcul de q/nkT et R_E , on peut apprécier indépendamment les évolutions du facteur d'idéalité et de la température du transistor liés à l'auto-échauffement.

III.2.1.b Influence de la zone d'extraction

Le choix de la zone de V_{BE} sur laquelle est faite la régression linéaire influe beaucoup sur le résultat final [Barbalat04] : A faible tension, la résolution des appareils de mesure est insuffisante, et l'influence de R_E est minime. A fort courant, les effets de haute injection s'ajoutent à la modélisation retenue, et l'auto-échauffement fait varier la température du composant, agissant sur la détermination de n , et donc de R_E . La figure III.4 présente l'évolution de g_m/I_C en fonction de g_m lorsque V_{BE} varie. Les valeurs de V_{BE} correspondantes sont reportées sur la courbe. On note que cette caractéristique n'est pas réellement une droite : La pente diminue avec l'augmentation de V_{BE} signalant une variation du terme q/nkT . Les courants augmentant, l'auto-échauffement ne devient plus négligeable, donc la température varie (voir chapitre V.2 sur les effets thermiques).

En conséquence, la valeur de R_E extraite par cette méthode varie en fonction de la tension V_{BE} choisie. A titre d'exemple, le tableau III.1 donne les valeurs de résistance d'émetteur et de facteur d'idéalité n_0 en fonction de la zone de régression. On constate une diminution de la résistance extraite lorsque la tension V_{BE} augmente : Ceci est dû à l'augmentation du courant collecteur à cause de l'auto-échauffement, et qui a pour conséquence de diminuer la chute de tension apparente, donc la valeur de R_E extraite. On peut donc considérer que la valeur de R_E sera d'autant plus significative que l'extraction a été réalisée à faibles V_{BE} . Dans la suite de l'étude, la zone d'extraction retenue est 0.7 – 0.8 V. Sur la figure III.4, on peut remarquer l'apparition des effets de haute injection à partir de $V_{BE} = 1$ V.

V_{BE} [V]	0.7–0.8 V	0.75–0.85 V	0.8–0.9 V	0.85–0.95 V
R_E [Ω]	17.1	12.3	10.3	9.6
$n_0 V_{Th}$ [mV]	27.4	28.1	29.9	32.0

TAB. III.1 – Valeurs de R_E et de $n_0 V_{Th}$ extraites en fonction de la plage de V_{BE} utilisée pour la régression. La température du substrat est fixée à 300 K. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$.

III.2.1.c Accord extraction statique / dynamique

La méthode exposée ci-dessus doit être comparée à des résultats de mesures hyperfréquences [Zerounian00] pour s'assurer de sa pertinence. Le tableau III.2 présente la comparaison des valeurs de R_E extraites par la méthode statique et lors des mesures hyperfréquences, sur des dispositifs de dimension d'émetteur $0.15 \times 3.6 \mu\text{m}^2$, et présentant des émetteurs différents (dopages et épaisseurs de Si-Cap). La zone d'extraction retenue est 0.7 – 0.8 V. On note un bon accord

entre les extractions en régime statique et dynamique. L'influence des variations technologiques sur les valeurs de R_E est détaillée dans la suite du chapitre.

Débit d'As [sccm]	90	200	200	200	200	250	250	250	280	280
Si-Cap [nm]	12	12	16	20	24	16	20	24	20	24
R_E dynamique [Ω]	14.4	8.3	9.1	12.8	14.1	6.9	9.8	13.0	8.6	10.8
R_E statique [Ω]	14.1	8.4	8.9	10.7	12.2	7.4	9.5	10.5	9.0	9.7

TAB. III.2 – Comparaison des valeurs de R_E (en Ω) extraites manuellement et lors des mesures HF. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$.

En moyenne, l'écart de valeur entre les deux méthodes est inférieur à 10 %. On peut donc en conclure que la méthode d'extraction statique fournit des résultats cohérents avec ceux provenant de l'extraction analytique des éléments du schéma équivalent de TBH, sur la base de mesures de paramètres S.

III.2.2 Effet des différentes variantes technologiques

Afin d'optimiser le module émetteur, il est nécessaire de connaître l'influence des paramètres technologiques sur la résistance d'émetteur. Cette partie vise à recenser les contributions de chaque partie de la figure III.2 à la résistance d'émetteur totale. L'analyse est menée sur un lot présentant diverses variations sur le module émetteur. Dans un premier temps, on étudie l'effet du dopage du polysilicium d'émetteur. Dans un deuxième temps, nous étudions l'effet de l'épaisseur du Si-cap. Enfin, nous abordons le rôle joué par le contact métallique, situé au dessus de l'émetteur, sur la résistance d'émetteur globale. Les principaux résultats reportés ici ont été publiés dans [Chevalier05a].

III.2.2.a Effet du dopage du polysilicium

Le premier paramètre qui a été soumis à des modifications est le dopage du polysilicium de l'émetteur. Il est possible de modifier l'espèce dopante, ainsi que les débits de gaz, pour contrôler la concentration de dopants. Les débits de gaz sont mesurés en sccm (1 sccm correspond à un débit de gaz de $1 \text{ cm}^3/\text{min}$ à 0°C et à 1 atmosphère).

III.2.2.a.i Type de dopage (Phosphore ou Arsenic)

Pour réaliser un dopage de type N sur silicium, deux espèces dopantes sont facilement disponibles en microélectronique : Le phosphore (P) ou l'arsenic (As), qui peuvent tous les deux être incorporés in-situ lors de la croissance du polyémetteur. Il existe néanmoins une différence fondamentale entre ces deux dopants : Alors que l'arsenic nécessite un recuit d'activation à haute température après dépôt, le phosphore est directement actif lors de son incorporation. Généralement, à niveau de dopage identique, le dopage phosphore assure une résistance d'émetteur plus faible. Les dispositifs devront subir des budgets thermiques différents selon le type de dopage, l'émetteur dopé phosphore ne permettant pas d'importants traitements thermiques : il faut adapter l'épaisseur du Si-cap afin d'obtenir un réglage de jonction émetteur/base équivalent à celui d'un émetteur dopé avec de l'arsenic.

Le tableau III.3 compare deux transistors ayant des émetteurs dopés phosphore et arsenic, présentant les caractéristiques de jonction E/B sensiblement égales : Le niveau de courant collecteur est quasiment le même, et les deux transistors présentent des capacités E/B, des tensions de claquage BV_{EBO} et des résistances de base pincée équivalentes. Le débit de dopants est le

même dans les deux cas (200 sccm). Le Si-cap a une épaisseur de 5 nm dans le cas du phosphore et 12 nm pour l'émetteur As. Les composants ont une dimension de $0.15 \times 3.6 \mu\text{m}^2$, et une base de 30 nm d'épaisseur, de profil 20–30 % de Ge.

Dopage d'émetteur	As	P
I_C [μA] ($V_{BE} = 0.75 \text{ V}$)	244	255
C_{BE} ($V_{BE} = 0.75 \text{ V}$) [fF]	12.6	12.7
BV_{EBO} [V]	1.50	1.56
R_{bp} [$k\Omega/\square$]	2.2	1.9
R_E [Ω]	13.3	7.3

TAB. III.3 – Comparaison des caractéristiques statiques de TBH ayant des émetteur dopés As ou P. L'émetteur phosphore n'est pas recuit. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$.

On constate qu'à réglage de jonction identique, l'émetteur dopé phosphore est moins résistif. Cependant, le phosphore induit des difficultés supplémentaires lors la fabrication du composant. En effet, le phosphore diffuse beaucoup lors du recuit final de la filière BiCMOS (au dessus de 1000°C), ce qui rend son intégration difficile dans une telle filière.

III.2.2.a.ii Niveau de dopage

Il est possible lors de la croissance de l'émetteur de faire varier le débit d'arsine (AsH_3), gaz précurseur du dopant arsenic. Ainsi, différents niveaux de dopage d'émetteur ont pu être comparés. Les analyses SIMS de la figure III.5 montrent les profils d'arsenic obtenus en faisant varier le débit d'arsine de 90 sccm à 280 sccm. Le dopage maximal de l'émetteur est multiplié par 2 lorsqu'on passe d'un débit de 90 sccm à 280 sccm ($2.5 \times 10^{20} \text{ at.cm}^{-3}$ pour 90 sccm contre $5 \times 10^{20} \text{ at.cm}^{-3}$ pour un débit de 280 sccm). L'incorporation d'arsenic ne dépend pas que du débit de gaz précurseur, mais également des conditions régnant dans la chambre de dépôt (température, pression) et également de la vitesse de croissance du matériau. Les développements en cours sur les matériaux ont pour but de réaliser des couches fortement dopées, et ce dès les premiers plans atomiques déposés.

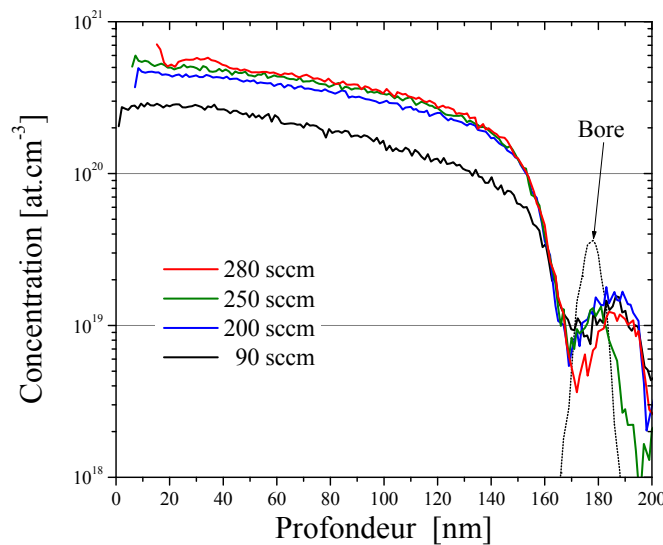


FIG. III.5 – Analyse SIMS de TBH ayant différents dopages d'émetteur (après recuit final).

On constate que la concentration d'arsenic après recuit n'est pas proportionnelle au débit de gaz précurseur durant la croissance du matériau. Il y a en effet peu de différence entre

les dopages 200, 250 et 280 sccm. L'incorporation d'As sature, et il est délicat d'obtenir des concentrations chimiques de dopants au delà de $5 \times 10^{20} \text{ at.cm}^{-3}$ (après recuit final).

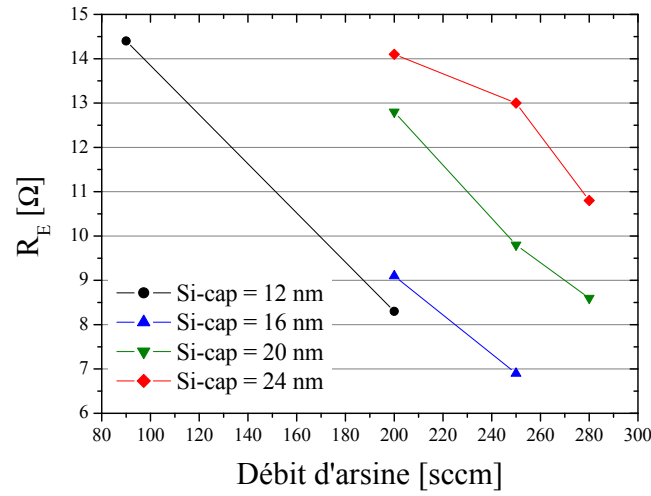


FIG. III.6 – Comparaison des résistances d'émetteur obtenues pour différents débits d'Arsine.

La variation de concentration d'As a une influence directe sur la résistance d'émetteur des transistors bipolaires. La figure III.6 représente les valeurs de R_E extraites sur des transistors de dimension $0.15 \times 3.6 \mu\text{m}^2$, ayant différentes épaisseurs de Si-cap, en fonction du débit d'arsine utilisé lors de la croissance du polyémetteur. On constate une forte décroissance de R_E lorsque le dopage du polyémetteur augmente. Ceci est dû à deux effets qui sont intimement liés :

- L'augmentation du dopage permet de diminuer la contribution R_{Ep} , résistance du matériau massif ;
- L'augmentation de la concentration de dopants du polyémetteur permet également de diminuer R_{Eb} , contribution de la zone de diffusion dans le Si-cap. Si le dopage du matériau massif augmente, le Si-cap sera plus dopé, la constante de diffusion des dopants étant a priori inchangée.

Pour étudier l'influence du dopage sur la composante R_{Ep} uniquement, il faut donc avoir un réglage de jonction émetteur/base identique, pour s'assurer que la valeur de R_{Eb} sera inchangée. Le tableau III.4 compare les valeurs de R_E obtenues sur des TBH ayant différents niveaux de dopage d'émetteur, mais des caractéristiques de jonction E/B quasiment constantes. Les transistors ont été choisis pour avoir des caractéristiques de jonction émetteur/base (capacité C_{BE} et tension de claquage BV_{EBO}) et de résistances de base pincée similaires.

Débit d'As [sccm]	90	200	250	280
Si-Cap [nm]	12	16	20	20
C_{BE} [fF]	12.1	12.6	12.0	11.1
BV_{EBO} [V]	1.63	1.50	1.96	1.77
R_{bp} [$k\Omega/\square$]	2.26	2.20	2.14	2.12
R_E [Ω]	14.4	9.1	9.8	8.6

TAB. III.4 – Évolution des valeurs de R_E à réglage de jonction équivalent. L'épaisseur du Si-cap est adaptée au niveau de dopage de l'émetteur. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$.

Lorsque les caractéristiques de la jonction E/B sont maintenues quasiment constantes, on peut observer que la résistance d'émetteur varie de manière significative avec le niveau de dopage : R_E diminue d'environ 5 Ω (50 %) lorsqu'on augmente le débit de 90 à 200 sccm. Au delà

de 200 sccm, les valeurs de R_E sont centrées autour de 9Ω . On constate en effet sur les profils SIMS de la figure III.5 que la concentration maximale d'As est quasiment identique pour les trois débits les plus forts, il n'y a donc pas d'évolution notable de R_{Ep} .

Une fois le dopage d'émetteur fixé, le réglage de la jonction E/B et l'optimisation de R_E dépendent de l'épaisseur du Si-cap, dont l'influence est décrite au paragraphe suivant.

III.2.2.b Influence du Si-Cap

La figure III.6 présente l'influence du niveau de dopage sur la résistance d'émetteur, et met également en avant l'impact très fort de l'épaisseur du Si-cap sur la résistance d'émetteur globale. On considère que les dopants de type P (bore) de la base sont fortement bloqués par le carbone et par la contrainte mécanique créée par le germanium, et donc diffusent peu en comparaison avec les dopants de l'émetteur.

L'épaisseur du Si-cap est un paramètre primordial du transistor bipolaire. En effet, comme indiqué sur la figure III.7, l'épaisseur de silicium optimale assure une diffusion de l'As jusqu'au pied du pic de bore (a). Un Si-cap trop épais aura pour conséquence une résistance d'émetteur trop forte (b). Une épaisseur trop fine aura pour effet de pincer fortement la base du composant (c), ce qui augmente la résistance de base pincée, pénalisante pour f_{MAX} , et dégrade la tenue en tension du composant.

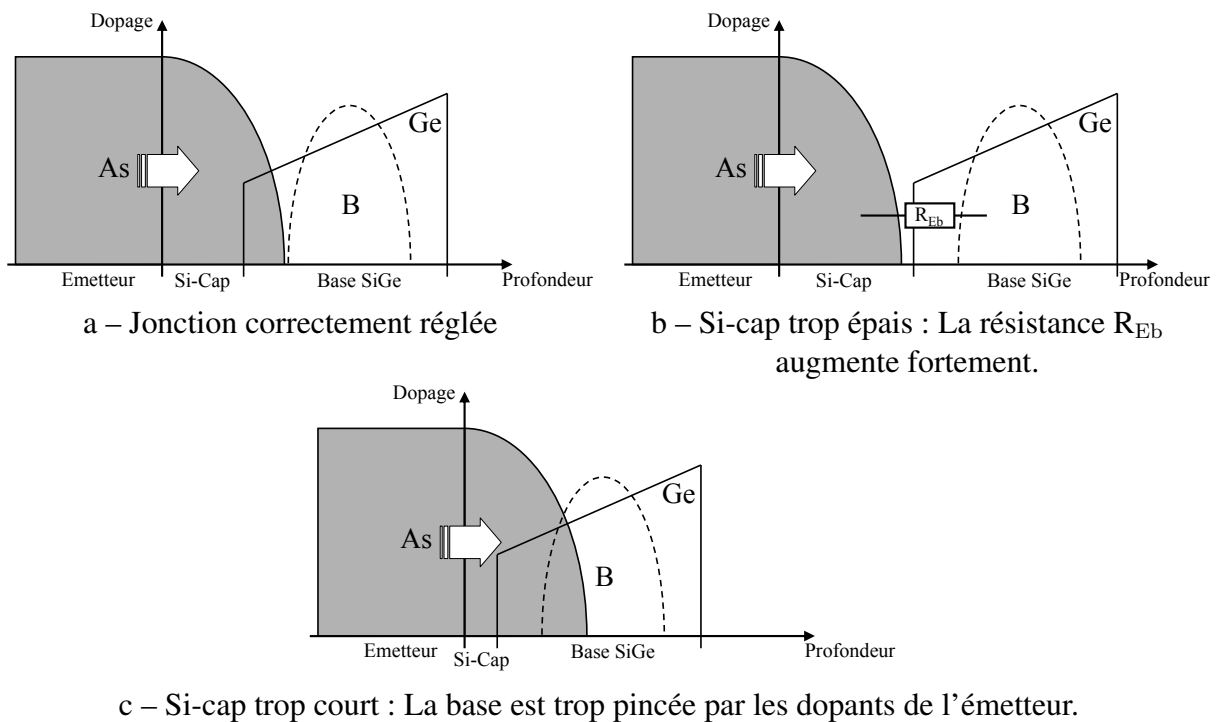


FIG. III.7 – Influence de l'épaisseur du Si-cap sur le réglage de la jonction E/B.

L'effet du Si-cap sur la jonction E/B est détaillé dans le tableau III.5, au travers de l'analyse des capacités, tensions de claquage et résistances de base pincée de transistors ayant un dopage d'émetteur de 200 sccm.

On note que lorsque l'épaisseur du Si-cap augmente, la capacité C_{BE} diminue fortement, et la tension de claquage BV_{EBO} augmente de manière significative. La résistance de base pincée diminue également, signe que la distance entre l'arsenic et le bore augmente. On note également une forte augmentation de la résistance d'émetteur lorsque l'épaisseur du Si-cap augmente, imputable à une augmentation de la composante R_{Eb} uniquement.

Débit d'As [sccm]	200	200	200	200
Si-Cap [nm]	12	16	20	24
C_{BE} [fF]	13.2	12.6	11.4	11.7
BV_{EBO} [V]	0.90	1.50	2.17	2.93
R_{bp} [$k\Omega/\square$]	2.40	2.20	2.05	2.03
R_E [Ω]	8.3	9.1	12.8	14.1

TAB. III.5 – Influence de l'épaisseur du Si-cap sur les caractéristiques de la jonction E/B et sur la résistance d'émetteur. Taille du composant : $0.15 \times 3.6 \mu m^2$.

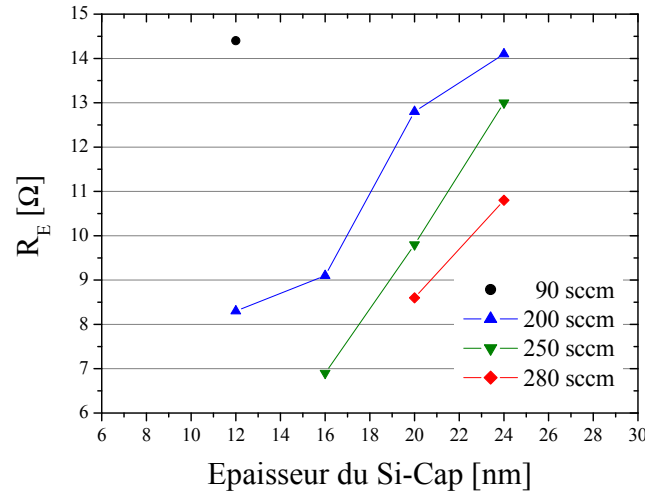


FIG. III.8 – Comparaison des résistances d'émetteur obtenues pour différents débits d'Arsine en fonction de l'épaisseur du Si-cap.

La figure III.8 présente les variations de résistance d'émetteur obtenues pour les différents débits d'arsine, en fonction de l'épaisseur du Si-cap. L'influence de l'épaisseur du cap sur la résistance d'émetteur est importante. Cela prouve que la principale contribution à la résistance d'émetteur est la composante R_{Eb} . Dans la suite de l'étude, le dopage As utilisé est le dopage maximal permis par les équipements (débit de 280 sccm), et le Si-cap aura une épaisseur de 18 voire 16 nm selon le profil de Germanium utilisé.

La réduction de l'épaisseur du Si-cap permet de réduire R_E , mais augmente le courant tunnel bande-à-bande. La figure III.9 présente les tracés de Gummel de trois transistors ayant des épaisseurs de Si-cap variables. Plus le Si-cap est fin, plus la distance entre l'arsenic et le bore diminue, ce qui favorise l'apparition d'un courant tunnel, comme l'indique l'évolution du niveau de I_B à 0.05 V. Cet effet tunnel apparaît comme une limitation à la réduction de l'épaisseur du Si-cap.

III.2.3 Effet de la géométrie du contact

Une étude a également été menée sur la forme du contact d'émetteur. Compte tenu des fortes densités de courant des transistors bipolaires rapides et de la faible résistance d'émetteur souhaitée, la surface de contact métallique doit être suffisamment importante, afin de respecter les règles d'électromigration dans les interconnexions. La taille du contact métallique étant fixée ($0.2 \times 0.2 \mu m^2$), pour augmenter la surface de contact il faut en augmenter le nombre, ce qui a débouché sur les matrices de contacts utilisées en technologie BiCMOS9 [Laurens03], et présentées dans la figure III.10–a. Cette forme de contact a l'inconvénient de rajouter une

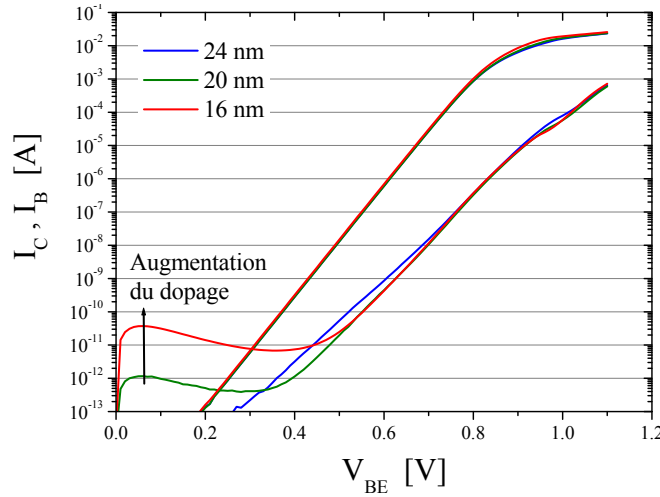


FIG. III.9 – Tracés de Gummel de TBH de dimension $0.15 \times 3.6 \mu\text{m}^2$, pour différentes épaisseurs de Si-cap.

composante à la résistance d'émetteur : En effet, pour avoir la place de poser les contacts sur le polyémetteur, sa largeur a été augmentée, il en résulte un trajet plus long pour les électrons, ce qui dégrade R_E .

Pour remédier à cet effet, un autre type de contact a été évalué, nommé contact en ruban [Chevalier05a]. Ce contact est centré sur la partie intrinsèque du TBH, et il a une longueur adaptable à la longueur de l'émetteur, comme indiqué sur la figure III.10–b. Ceci représente une violation – minimale – des règles de dessin de BiCMOS9 mais possède l'avantage de réduire la résistance d'émetteur de façon significative.

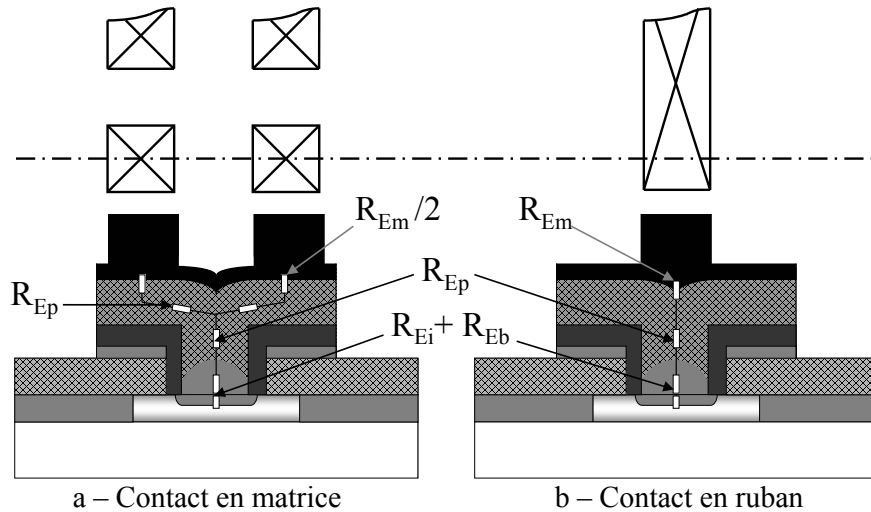


FIG. III.10 – Présentation des deux types de contacts d'émetteur étudiés [Chevalier05a].

III.2.4 Gain sur les fréquences de transition

L'optimisation complète de l'émetteur a permis de diminuer la résistance d'émetteur de 16Ω à environ 7Ω pour un TBH de surface d'émetteur $0.15 \times 3.6 \mu\text{m}^2$. Avec $5 \Omega \cdot \mu\text{m}^2$ pour un TBH doté d'un polyémetteur dopé par débit d'arsine de 90 sccm, d'un Si-cap de 12 nm et d'un contact matriciel, la résistance spécifique de l'émetteur a été réduite à $1.6 \Omega \cdot \mu\text{m}^2$ pour un

TBH avec un polyémetteur à 250 sccm, un Si-cap de 16 nm et un contact d'émetteur en ruban. La figure III.11 présente des tracés de Gummel de composants ayant l'émetteur de référence, et un émetteur optimisé.

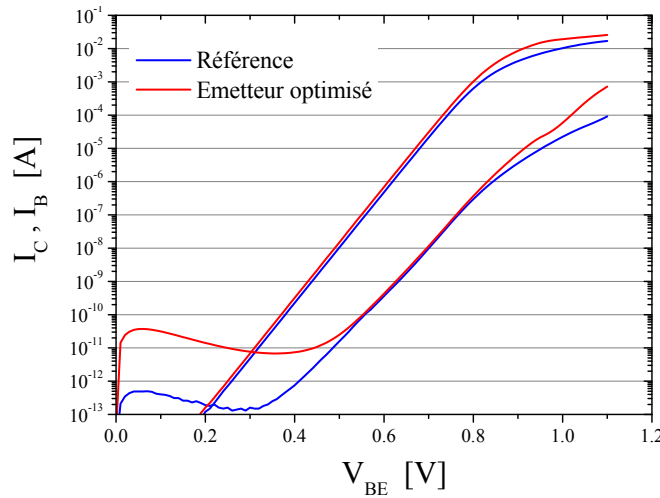


FIG. III.11 – Tracés de Gummel de TBH de dimension $0.15 \times 3.6 \mu\text{m}^2$, dans le cas d'un émetteur de référence et d'un émetteur optimisé.

Dans le cas de l'émetteur optimisé, le courant collecteur est plus important, ce qui est dû à un profil de dopage plus abrupt, comme le prouve l'augmentation du courant tunnel bande-à-bande. La réduction de R_E et les modifications apportées sur le réglage de la jonction E/B ont permis d'augmenter significativement les performances dynamiques, comme l'illustre la figure III.12, où sont comparées les valeurs de f_T de TBH ayant des émetteurs optimisés (250 sccm, Si-cap de 16 nm) à celles des TBH de référence (90 sccm, Si-cap de 12 nm), pour les deux types de contacts possibles (matrice ou ruban).

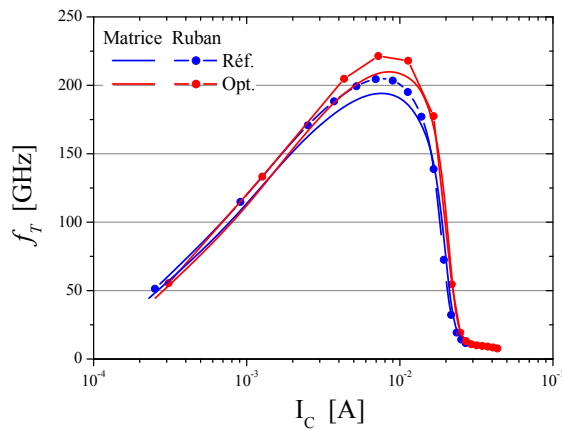


FIG. III.12 – Évolution de la fréquence f_T en fonction de I_C . *Réf.* : Émetteur de référence (90 sccm, Si-cap de 12 nm) ; *Opt.* : Émetteur optimisé (250 sccm, Si-cap de 16 nm).

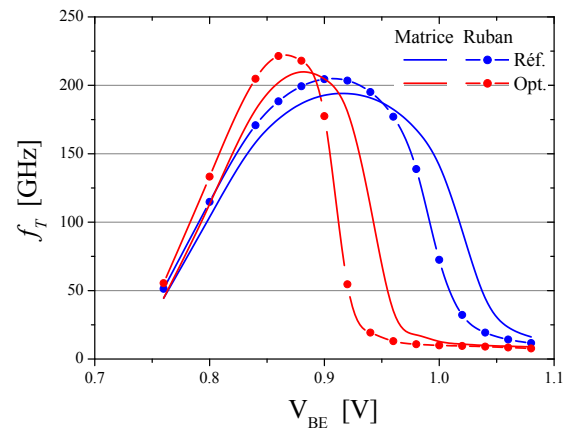


FIG. III.13 – Évolution de f_T en fonction de la tension V_{BE} . *Réf.* : Émetteur de référence (90 sccm, Si-cap de 12 nm) ; *Opt.* : Émetteur optimisé (250 sccm, Si-cap de 16 nm).

Un autre effet caractéristique de l'optimisation de R_E , est la tension V_{BE} à laquelle le pic de f_T est atteint. L'effet Kirk apparaît à partir d'un certain seuil de courant, qui est fixé par le collecteur. On note que deux composants ayant des résistances d'émetteur très différentes atteignent leur maximum de f_T pour une même valeur de I_C , comme le montre la figure III.12.

Lorsqu'on réduit la résistance d'émetteur, on réduit la chute de tension à ses bornes, donc le courant optimal est atteint à une tension V_{BE} plus faible, comme l'illustre la figure III.13 qui propose une comparaison entre deux réglages d'émetteur (débit d'arsine de 90 sccm avec un cap de 12 nm contre 250 sccm avec un cap de 16 nm), pour chacun des deux types de contacts (matrice ou ruban). On note qu'avec la diminution de la résistance d'émetteur, la fréquence f_T augmente, le pic de f_T se décale vers les V_{BE} plus faibles. Cette diminution de la tension V_{BE} optimale est bénéfique car elle s'inscrit dans la tendance actuelle de réduction des tensions d'alimentation.

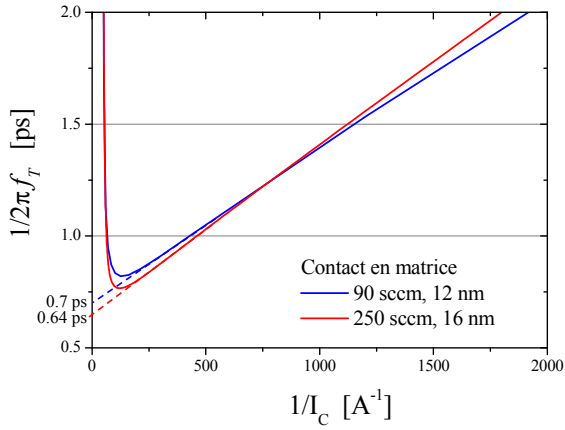


FIG. III.14 – Évolution de τ_{EC} en fonction du courant collecteur. Contact d'émetteur en matrice.

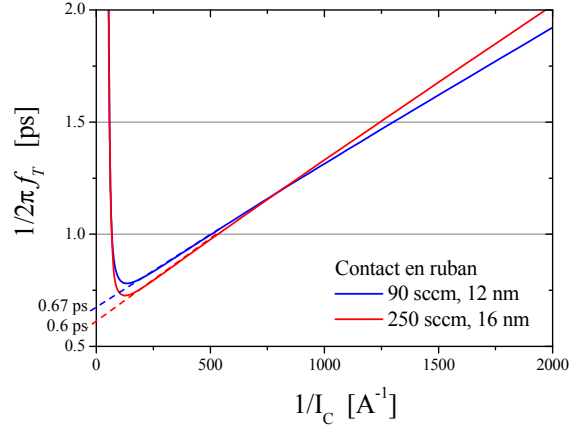


FIG. III.15 – Évolution de τ_{EC} en fonction du courant collecteur. Contact d'émetteur en ruban.

L'influence de la réduction de R_E est également visible sur le temps de transit total τ_{EC} . Les figures III.14 et III.15 présentent l'évolution de τ_{EC} en fonction de I_C dans le cas de l'émetteur de référence et de l'émetteur optimisé, pour les deux types de contacts (matrice ou ruban). La valeur de $\tau_{EC,0}$ est réduite de 0.7 ps à 0.64 ps dans le cas de contacts d'émetteur matriciés, et de 0.67 ps à 0.6 ps dans le cas de contacts d'émetteur en ruban. Cette réduction est attribuée à plusieurs effets différents :

- Le temps de transit émetteur τ_E est réduit, grâce à un profil de dopage plus abrupt.
- Le temps de transit de base τ_B est également réduit. Lorsque l'émetteur est optimisé, les dopants progressent vers la base, qui devient plus fine.
- La résistance d'émetteur étant plus faible, le retard $R_E C_{BC}$ est également réduit

L'augmentation de la pente de τ_{EC} en fonction de $1/I_C$ est due à l'augmentation de la capacité C_{BE} lorsque le profil du composant devient plus abrupt. Cet effet contribue à diminuer le temps de transit extrapolé $\tau_{EC,0}$.

Le tableau III.6 résume les principaux paramètres électriques du transistor de référence (90 sccm, Si-cap de 12 nm) et du transistor présentant la plus faible résistance d'émetteur (250 sccm, Si-cap de 16 nm), pour les deux types de contact d'émetteur, matrice et ruban.

L'optimisation de la résistance d'émetteur décrite ci-dessus a permis d'améliorer la fréquence f_T de 195 GHz à 221 GHz, soit une amélioration de l'ordre de 13 %, alors que les profils de base et de collecteur sont restés les mêmes. La fréquence f_{MAX} a bénéficié d'une amélioration de 10 %, notamment grâce à la diminution du recouvrement du polyémetteur sur le polybase. Ceci prouve que le module émetteur a une importance prépondérante sur les performances dynamiques du transistor bipolaire. En conséquence de cette étude, l'émetteur qui sera utilisé dans la suite, sauf indication contraire, est un émetteur dopé grâce à un débit d'arsine de 280 sccm. Le Si-cap optimal a une épaisseur comprise entre 16 et 18 nm.

Contact	Matrice		Ruban	
Débit d'As [sccm]	90	250	90	250
Si-cap [nm]	12	16	12	16
R_E [Ω]	16.0	8.9	14.4	6.9
C_{BE} [pF]	11.8	13.6	12.1	13.8
BV_{EBO} [V]	1.63	1.26	1.63	1.26
R_{bp} [$k\Omega/\square$]	2.3	2.2	2.3	2.2
f_T [GHz]	194	210	205	221
f_{MAX} [GHz]	219	223	230	239
V_{BE} à f_{Tmax}	0.92	0.88	0.9	0.86
$\tau_{EC,0}$ [ps]	0.70	0.64	0.67	0.60

TAB. III.6 – Principaux paramètres électriques de TBH avec émetteur de référence et émetteur optimisé.

III.3 Optimisation du profil de base

Principale partie active du transistor bipolaire, la base détermine en grande partie ses performances statiques et dynamiques. De l'optimisation de la base dépendent largement les fréquences de transition f_T et f_{MAX} . Cependant, l'optimisation du profil vertical de la base doit tenir compte des paramètres de l'émetteur et du collecteur. Les principales variations étudiées dans ce travail portent sur le profil de Germanium.

L'optimisation du profil du carbone a fait l'objet d'une étude poussée dans [Chevalier03]. Il en découle que le profil de carbone le plus favorable est un profil entourant complètement le pic de bore, avec une concentration de $3 \cdot 10^{19} \text{ at.cm}^{-3}$. Le profil de dopants optimal est un pic de bore dopé à $5 \cdot 10^{19} \text{ at.cm}^{-3}$ sur 4 nm.

Lorsque la dose d'atomes dopants augmente, ceux-ci vont avoir tendance à diffuser plus largement, entraînant une base électrique plus large et la fréquence f_T chute à cause de l'augmentation du temps de transit τ_B . Un dopage plus faible assure un gain en courant plus important, ce qui améliore f_T , mais la résistance de base pincée augmente fortement, ce qui fait chuter f_{MAX} . La concentration de bore choisie est telle que la résistance de base pincée visée se situe autour de $2 \text{ k}\Omega/\square$.

Il existe plus généralement un compromis délicat entre f_T et f_{MAX} . La présence de Germanium est un des facteurs permettant d'augmenter conjointement les deux fréquences de transition. Un fort taux de Ge dans la base assurera une valeur de f_T plus forte par l'augmentation de I_C , ce qui permet en retour d'augmenter le dopage de base pour augmenter f_{MAX} . L'analyse de l'influence des principaux paramètres technologiques sur les performances du composant permet de mettre en avant les choix à effectuer afin de réaliser des transistors ayant des performances ultimes.

III.3.1 Profils de germanium

Le profil de SiGe dans la base des TBH joue au premier ordre sur les caractéristiques électriques du composant. Le taux de Ge à l'entrée de la base a une influence sur la hauteur de l'hétérojonction, donc fait sensiblement évoluer le gain en courant du transistor. Le gradient de Ge a un impact sur le pseudo-champ électrique accélérateur, jouant sur le transport des électrons dans la couche de base.

III.3.1.a Taux de Germanium à l'entrée de la base

La concentration de Ge à l'entrée de la base joue sur le gain en courant du composant. Un taux de Ge élevé diminue la hauteur de barrière vue par les électrons, ce qui améliore le gain en courant. En revanche, une forte différence de taux de germanium entre l'entrée et la sortie de la base améliore le pseudo-champ accélérateur et favorise le transport des électrons dans la base. L'impact du taux de Ge à l'entrée de la base est évalué grâce à l'analyse de 3 bases ayant des profils différents. Les trois profils testés sont 20–30 %, 15–30 %, et 10–30 %. La couche de SiGe est scindée en marches successives variant à chaque fois de 5 %, comme indiqué sur la figure III.16.

La dose de bore est la même dans les trois configurations, et le pic d'atomes dopants est placé au milieu de la base. Cependant, en fonction du profil de Ge retenu, le pic de bore sera incorporé dans une marche avec un taux de Ge variable : 25 % pour le profil 20–30, entre 20 et 25 % pour le profil 15–30, et 20 % pour le profil 10–30. Cette différence d'incorporation et/ou d'activation a une influence sur les performances du transistor bipolaire.

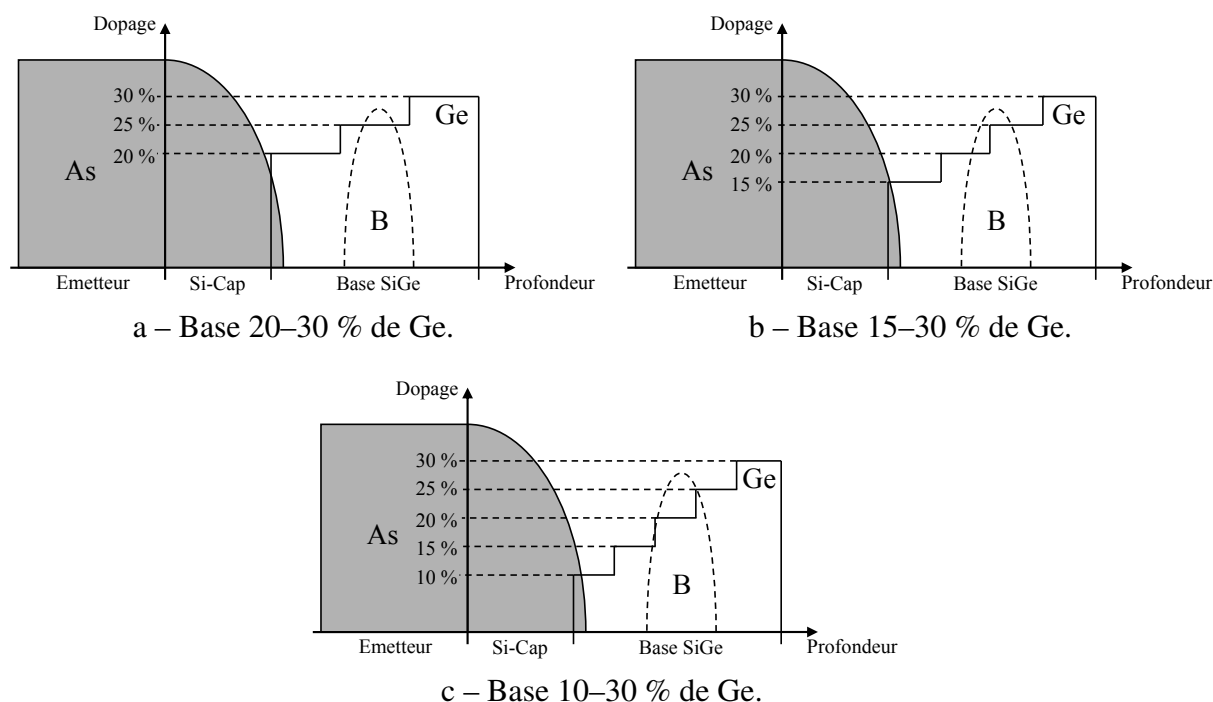


FIG. III.16 – Présentation des trois profils de base. La couche totale de SiGe a une épaisseur de 30 nm.

Le tableau III.7 donne une comparaison des principaux paramètres statiques de TBH de dimension $0.15 \times 3.6 \mu\text{m}^2$, dotés de bases de 30 nm d'épaisseur, avec un profil 20–30 %, 15–30 % et 10–30 %. L'émetteur est dopé grâce à un débit d'arsine de 90 sccm, et le Si-cap a une épaisseur de 12 nm. Le contact d'émetteur est un contact matricé standard.

Profil de Ge [%]	20–30	15–30	10–30
I_C à $V_{BE} = 0.75 \text{ V}$ [μA]	140	100	80
I_B à $V_{BE} = 0.75 \text{ V}$ [nA]	50	80	70
Gain à $V_{BE} = 0.75 \text{ V}$	2670	1330	1180
C_{BE} [pF]	10.9	11.1	9.8
R_{bp} [$\text{k}\Omega/\square$]	2.2	3.4	4.8
R_E [Ω]	16.6	19.0	21.7
V_{Af} [V]	136	118	130
V_{Ar} [V]	7.2	2.9	1.7
BV_{EBO} [V]	1.57	1.80	2.30
f_T [GHz]	200	210	207
f_{MAX} [GHz]	203	196	197

TAB. III.7 – Évolution des principaux paramètres électriques avec le taux de Ge à l'entrée de la base (épaisseur de base : 30 nm).

Comme indiqué dans la partie théorique, la tension d'Early dépend du profil de Germanium de la base. Concernant la tension d'Early directe V_{Af} , on n'observe pas d'évolution significative, alors que le gradient de Ge dans la base a évolué. Ceci est dû à la variation indirecte de la dose de bore lorsque le taux de Ge varie. En revanche, la tension d'Early inverse se dégrade de façon significative lorsque le taux de Ge côté émetteur est réduit de 20 % à 10 %. On constate une réduction de la tension V_{Ar} de 7 V à 1.7 V avec l'augmentation du gradient de germanium.

Lorsqu'on réduit le taux de Germanium à l'entrée de la base, le courant collecteur est réduit,

ce qui diminue fortement le gain en courant (de plus d'un facteur 2). Cet effet est dû au fait que le pic de bore est incorporé dans des taux de Ge différents selon les composants, et confirmé par l'évolution de la résistance de base pincée du tableau III.7 : Plus le taux de Ge est faible, plus la résistance carrée est importante. Cette évolution est expliquée par la différence d'incorporation du Bore lorsque le taux de Ge varie. La figure III.17 présente le tracé de Gummel et le gain en courant des trois transistors présentés précédemment. La figure III.18 présente l'évolution des fréquences f_T obtenues avec les trois profils de Ge.

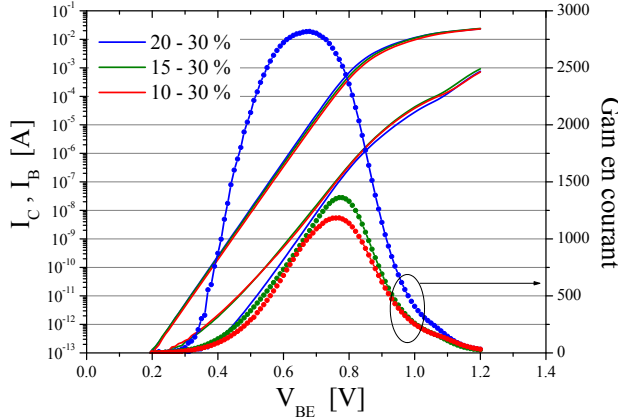


FIG. III.17 – Courbes de Gummel et de Gain pour trois profils de Ge.

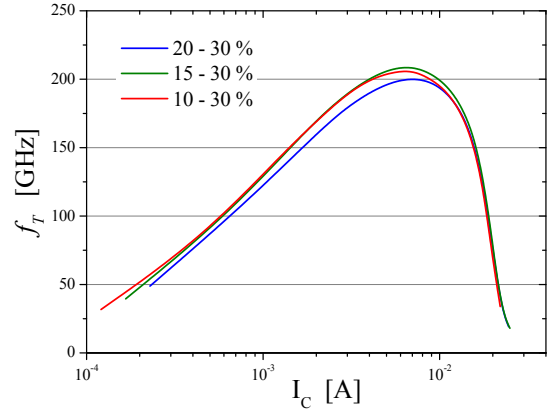


FIG. III.18 – Evolution de f_T en fonction de I_C pour trois profils de Ge.

On constate que le profil intermédiaire 15–30 % est le profil offrant les meilleures performances dynamiques ($\Delta f_T = 10$ GHz). Il existe un compromis entre un fort taux de Germanium à l'entrée de la base, un gradient important de SiGe dans l'épaisseur de la base et concentration de bore. Les temps de transit $\tau_{EC,0}$ sont du même ordre de grandeur (0.65 à 0.68 ps), comme le prouve la figure III.19.

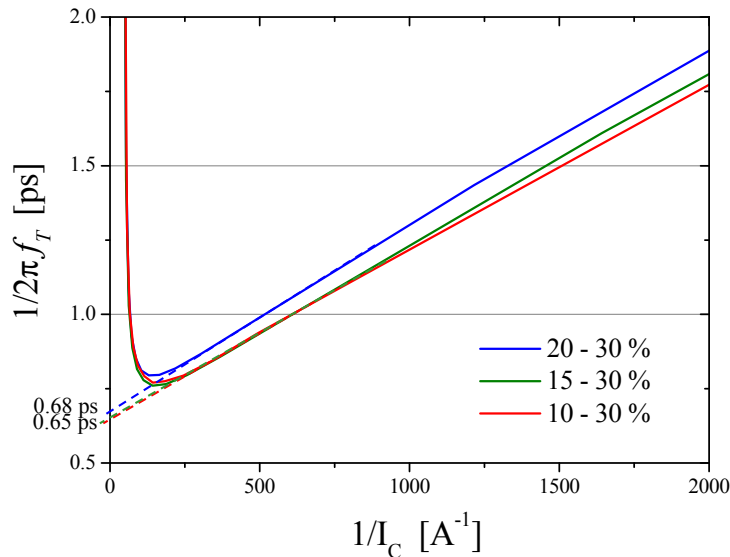


FIG. III.19 – Temps de transit τ_{EC} en fonction de I_C pour différents profils de SiGe.

Concernant le profil de Ge optimal, il est donc judicieux de choisir le plus fort taux de Ge, soit un profil 20–30 %. En effet, ce type de profil présente des fréquences f_T similaires aux autres profils, avec toutefois un f_{MAX} (+7 GHz) plus important, grâce à une résistance de base pincée plus faible. En effet, un fort taux de Ge est favorable à des résistances de couches faibles.

La résistance d'émetteur, ainsi que la tension d'Early sont améliorées par le fort taux de Ge à l'entrée de la base. Les performances statiques étant meilleures avec une base 20–30 %, c'est ce type de profil qui est retenu afin de réaliser les transistors bipolaires ayant des performances ultimes.

III.3.1.b Épaisseur de la couche de SiGe

Lorsque l'épaisseur de la base varie, le gradient d'électrons dans la base neutre varie. Ainsi, une base fine est favorable à un temps de transit τ_B réduit, car le trajet suivi par les électrons est plus court, et donc à des fréquences f_T et f_{MAX} plus fortes. Il est donc bénéfique de réduire l'épaisseur de la base neutre. Il faut cependant que les atomes de bore soient intégralement contenus dans la couche de SiGe, sinon le bénéfice de l'hétérojonction est perdu. Il n'est donc pas possible de réaliser des bases aussi fines que souhaité, sous peine de voir le bore diffuser en dehors de la couche de SiGe.

Deux épaisseurs de couche de SiGe ont été comparées : 30 nm (épaisseur de référence) et 20 nm. Le profil de base retenu est un profil 20–30 % de Ge, comme indiqué dans la figure III.20.

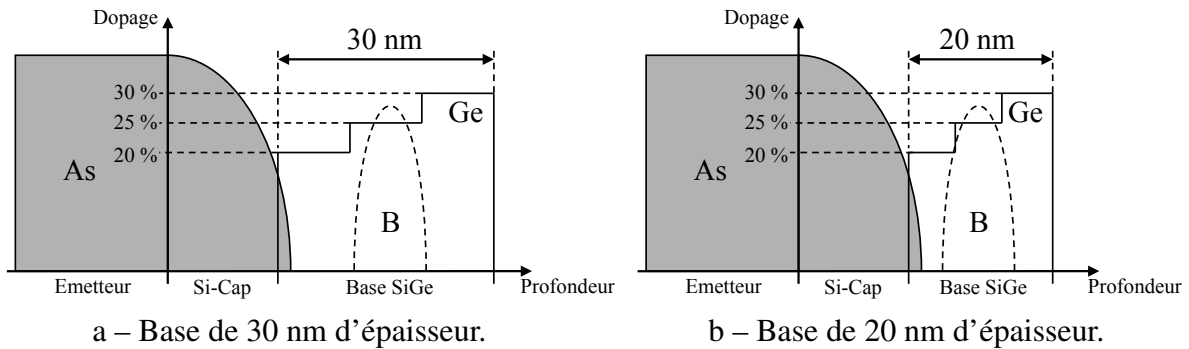


FIG. III.20 – Variation de l'épaisseur de la base. Le profil de Ge retenu est un profil 20–30 % en trois marches successives.

La variation d'épaisseur totale de Germanium est visible sur le profil SIMS de la figure III.21.

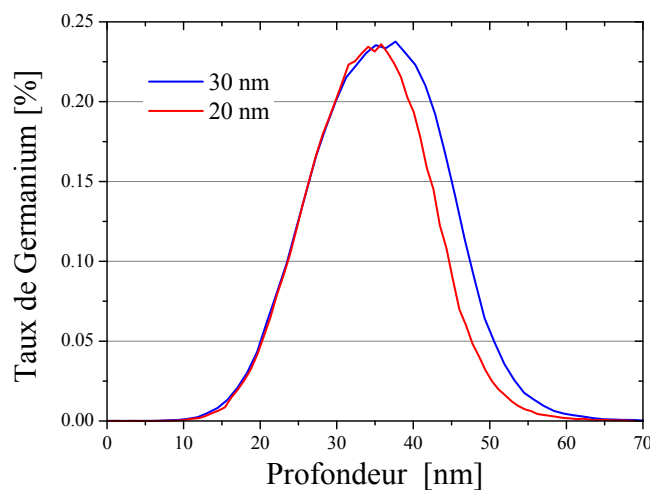


FIG. III.21 – Profils de Ge obtenus par SIMS pour différentes épaisseurs de base.

Le tableau III.8 compare les principales caractéristiques statiques et dynamiques de deux TBH ayant les profils de base présentés ci-dessus. La dimension d'émetteur est $0.15 \times 3.6 \mu\text{m}^2$,

le contact d'émetteur est un contact matricé standard, le polyémetteur est dopé à 90 sccm et le Si-cap a une épaisseur de 12 nm. La dose de bore visée est la même dans les deux cas.

Epaisseur de SiGe [%]	30 nm	20 nm
I_C à $V_{BE} = 0.75$ V [μ A]	140	115
I_B à $V_{BE} = 0.75$ V [nA]	50	55
Gain à $V_{BE} = 0.75$ V	2670	2165
C_{BE} [pF]	10.9	11.5
R_{bp} [$k\Omega/\square$]	2.2	1.9
R_E [Ω]	16.6	17.4
V_{Af} [V]	136	209
V_{Ar} [V]	7.2	4.1
BV_{EBO} [V]	1.57	1.02
f_T [GHz]	200	221
f_{MAX} [GHz]	203	233

TAB. III.8 – Évolution des principaux paramètres électriques avec l'épaisseur de la couche de SiGe (profil de SiGe : 20–30 %).

On constate dans le tableau III.8 que le transistor ayant la base la plus épaisse présente le courant collecteur et le gain les plus forts. En effet, plus la base est fine, plus la distance arsenic-bore est faible, comme le prouve l'évolution du courant tunnel bande-à-bande de la figure III.22. Les atomes de bore étant introduits au milieu de la couche de SiGe, dans le plateau à 25 % de Ge, la diminution de l'épaisseur de SiGe a pour effet d'augmenter le dopage de type P à l'entrée de la base neutre. Le dopage augmentant à l'entrée de la base, l'efficacité d'injection de la jonction E/B est diminuée, et donc le courant I_C ainsi que le gain en courant sont réduits. On note également que la résistance de base pincée diminue lorsque l'épaisseur de base diminue. Le composant ayant la base la plus fine présente les meilleures performances dynamiques. Le profil de base est plus abrupt, comme le prouve l'évolution de la tension BV_{EBO} et de la capacité C_{BE} .

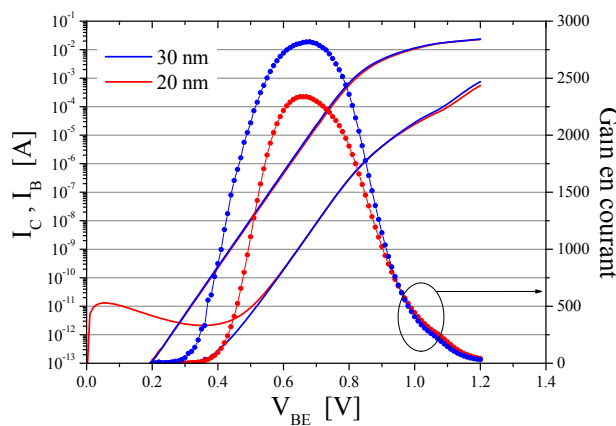


FIG. III.22 – Courbes de Gummel et de Gain en fonction de l'épaisseur de SiGe.

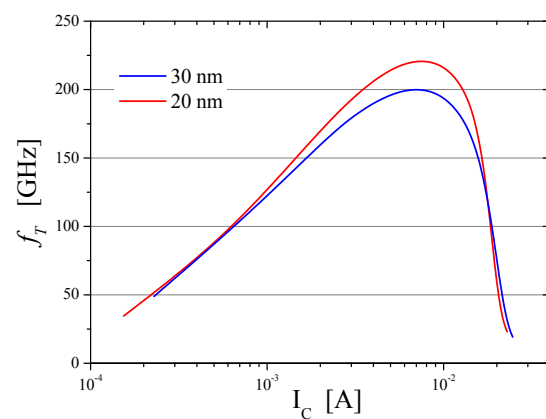


FIG. III.23 – Evolution de f_T en fonction de I_C pour différentes épaisseurs de SiGe.

La figure III.23 trace l'évolution de f_T en fonction de I_C pour les deux épaisseurs de base 30 nm et 20 nm. La base d'épaisseur 20 nm présente une amélioration de 10 % des performances par rapport à la base de 30 nm d'épaisseur. Le maximum de f_T et l'effet Kirk apparaissent aux mêmes densités de courant pour les deux épaisseurs de base. En ce qui concerne l'analyse

du temps de transit τ_{EC} , la figure III.24 montre que la base 20 nm présente également une amélioration notable des performances par rapport à la base 30 nm.

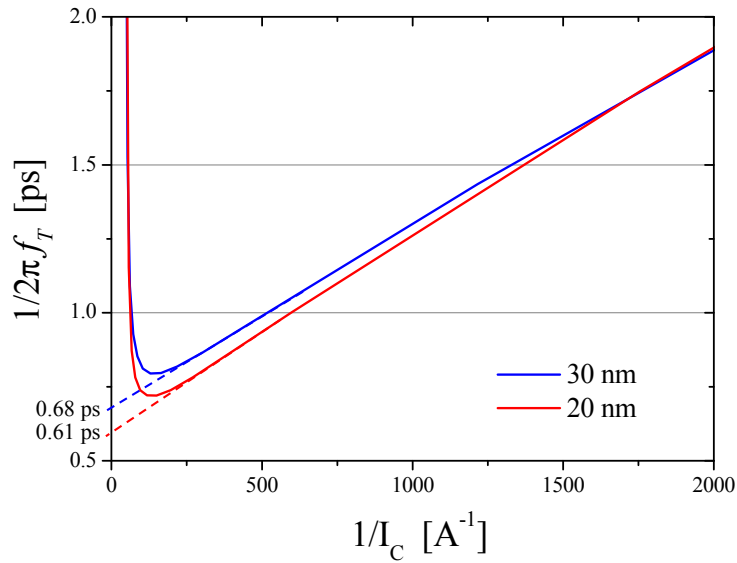


FIG. III.24 – Temps de transit τ_{EC} en fonction de I_C pour différentes épaisseurs de SiGe.

Le transistor ayant la base la plus fine présente le temps de transit le plus faible, grâce à un temps de transit de base réduit. Cependant, en l'absence d'outils de simulation performants prenant en compte le transport des électrons dans le SiGe, nous ne pouvons quantifier précisément cette diminution de τ_B . On remarque également que la pente de τ_{EC} en fonction de $1/I_C$ est plus forte dans le cas d'une base fine, ce qui est dû à une valeur légèrement plus forte de la capacité C_{BE} . En conclusion, sous réserve de bien contrôler l'incorporation du bore, une base fine est plus favorable à la montée en fréquence.

III.3.1.c Profil de Germanium graduel ou constant

Des transistors ayant des bases à Ge constant ont également été réalisés au cours de cette étude. Leurs performances sont comparées au transistor à base graduelle 20–30 % dans le tableau III.9. Les bases ont toutes une épaisseur de 20 nm. La dimension du composant est $0.15 \times 3.6 \mu\text{m}^2$, le contact d'émetteur est matricé, et l'émetteur est dopé à 90 sccm, avec un Si-cap de 12 nm. La dose de bore visée est la même pour les trois bases mesurées. La figure III.25 donne une comparaison des courbes de Gummel et de gain des trois transistors décrits précédemment.

On constate que les courants de base et de collecteurs dépendent principalement du taux de Ge à l'hétérojonction E/B : I_C et I_B sont identiques pour les profils 20–30 % et 20 %. Il faut cependant être prudent, car le dopage bore est sujet à variations lorsque le taux de Ge varie. Dans le cas du profil 25 %, l'augmentation de I_C par rapport aux deux autres profils de base est due à la réduction de la hauteur de barrière pour les électrons à l'hétérojonction E/B. Concernant les performances dynamiques, f_T est légèrement amélioré pour le profil constant à 20 % de Ge (225 GHz contre 221), et se dégrade pour le profil constant à 25 % (195 GHz). Cependant, on note une forte diminution de f_{MAX} pour le profil 20 %, due à l'augmentation de la résistance de base pincée. Le meilleur compromis entre f_T et f_{MAX} est obtenu avec la base graduelle 20–30 %.

Des profils de dopants obtenus par analyse SIMS permettent de comprendre l'augmentation des paramètres électriques en fonction du profil de base. En effet, l'incorporation du bore

Concentration de Ge	20–30 %	20 %	25 %
I_C à $V_{BE} = 0.75$ V [μ A]	115	110	138
I_B à $V_{BE} = 0.75$ V [nA]	55	55	45
Gain à $V_{BE} = 0.75$ V	2165	2080	3010
C_{BE} [pF]	11.5	11.0	10.6
R_{bp} [$k\Omega/\square$]	1.9	3.5	2.0
R_E [Ω]	17.4	19.6	15.0
V_{Af} [V]	209	45	85
V_{Ar} [V]	4.1	3.6	8.7
BV_{EBO} [V]	1.02	1.23	1.75
f_T [GHz]	221	225	195
f_{MAX} [GHz]	233	174	209

TAB. III.9 – Évolution des principaux paramètres électriques avec le profil de base (épaisseur de SiGe : 20 nm).

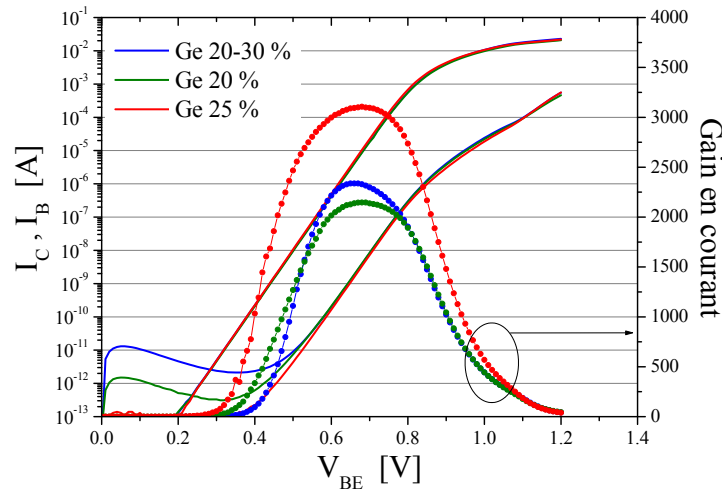


FIG. III.25 – Courbes de Gummel et de Gain en fonction du profil de SiGe.

dépend des conditions régnant dans la chambre d'épithaxie, et on observe sur la figure III.26 que la concentration de bore augmente avec le taux de Germanium. Ainsi, le profil constant à 20 % de Ge est le moins dopé, donnant ainsi une base plus fine. En conséquence, la fréquence f_T est favorisée, et f_{MAX} est dégradée.

Concernant les tensions d'Early directes, les bases à taux de Ge constant ne permettent pas de fabriquer des transistors présentant des valeurs de V_{Af} supérieures à 100 V, ce qui est pénalisant pour l'utilisation des composants dans un circuit.

III.3.1.d Synthèse de l'optimisation du profil de base

Nous avons démontré dans cette section l'amélioration des performances du transistor bipolaire en utilisant une base de 20 nm d'épaisseur, avec un taux de Ge rétrograde fixé à 20 % du côté de l'émetteur et 30 % du côté collecteur. Les profils de Ge constant n'ont pas donné de résultats satisfaisant : La dégradation des tensions d'Early et le déséquilibre entre f_T et f_{MAX} nous ont conduits à écarter cette solution. Dans la suite de l'étude, on ne considérera plus que des TBH à base graduelle.

La combinaison de ces différentes améliorations sur le profil de base, ainsi que sur la résistance d'émetteur, a permis d'augmenter significativement les performances du composant.

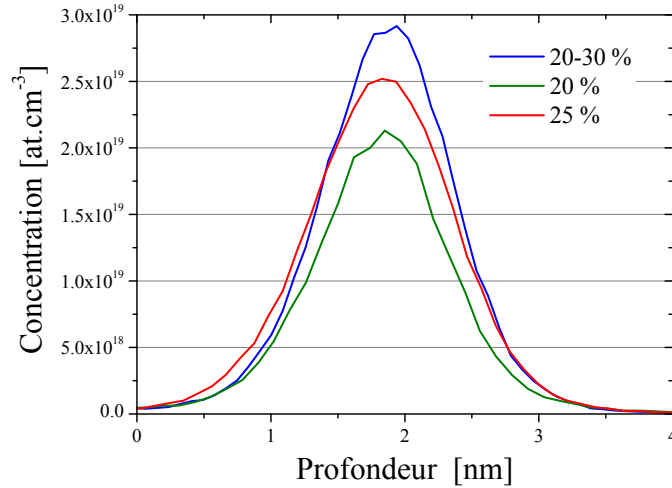


FIG. III.26 – Comparaison des profils de bore obtenus par analyse SIMS en fonction du taux de Ge dans la base.

La figure III.27 retrace l'historique des fréquences f_T en fonction de V_{BE} pour les différentes variations réalisées sur la base et l'émetteur. Cette optimisation a permis d'augmenter la fréquence f_T de 200 GHz à plus de 250 GHz. Dans le même temps, la valeur maximale de f_{MAX} est passée de 200 GHz à 280 GHz.

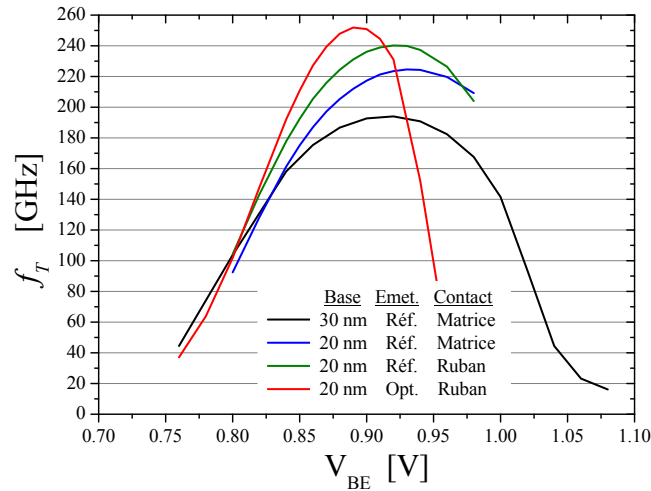


FIG. III.27 – Evolution de f_T en fonction de V_{BE} pour différentes variantes technologiques sur la base et l'émetteur.

III.3.2 Introduction d'un nouveau profil de base

Cette optimisation rendue possible par un profil de base performant s'est faite par une augmentation substantielle du courant collecteur et donc du gain en courant. Ceci pose problème pour l'intégration de ces composants dans un circuit électrique complet : les forts courants entraînent de l'électromigration dans les interconnexions métalliques, l'auto-échauffement est important. La tension de claquage BV_{CEO} est également dégradée par le fort gain du TBH. Il est donc intéressant pour une technologie BiCMOS de proposer des transistors ayant des courants collecteurs plus faibles, tout en maintenant les performances dynamiques obtenues jusqu'alors.

Afin de diminuer le courant collecteur, le taux de Germanium à l'entrée de la base doit être réduit. Le profil retenu est un profil 10–25 % de Ge, comme présenté dans la figure III.28.

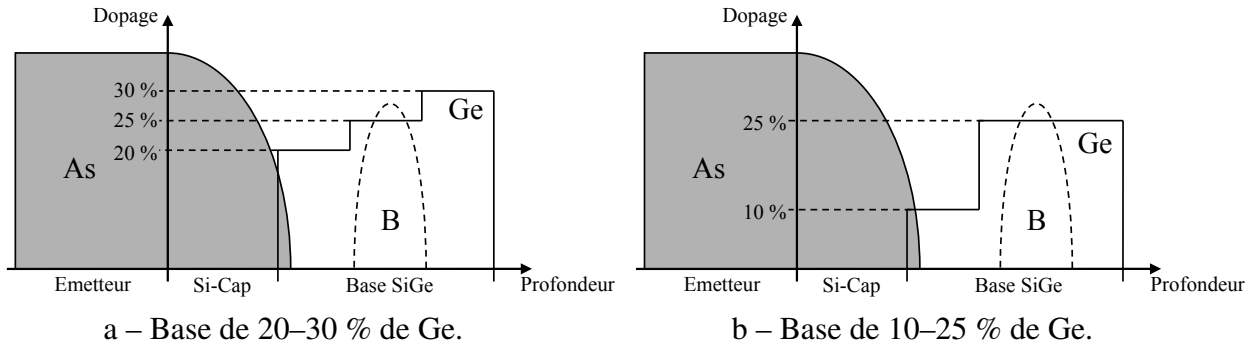


FIG. III.28 – Introduction d'un nouveau profil de Ge pour la réalisation de bases de transistors bipolaires rapides.

Les épaisseurs des différentes couches de la base sont adaptées de manière à ce que la jonction E/B ait lieu dans la marche à 10 % de Ge. Le bore est maintenu dans la partie à plus fort taux de Ge qui induit une contrainte permettant de limiter la diffusion des dopants. Il est possible de faire varier indifféremment les épaisseurs du Si-cap et de la première marche de Ge, afin de modifier la position de la jonction E/B, et permettant ainsi d'obtenir le bon réglage de base en modifiant la distance entre l'arsenic et le bore. Différents profils 10–25 % ont été comparés au profil 20–30 % de référence. Dans un premier temps, l'épaisseur du Si-cap est maintenue constante et égale à 18 nm et la première marche de Ge est réduite, comme indiqué dans la figure III.29–a. Dans un second temps, la distance As–B est maintenue constante, et on augmente l'épaisseur de la première marche de Ge, tout en réduisant l'épaisseur du Si-Cap, comme présenté dans la figure III.29–b. Dans tous les cas, $e_{25\%}$ est fixée à 14 nm.

Le tableau III.10 donne les principaux résultats statiques et dynamiques de transistors bipolaires de dimension $0.15 \times 3.6 \mu\text{m}^2$, réalisés avec les profils de base présentés en figure III.29. L'émetteur est dopé avec un débit d'arsine de 280 sccm, et le contact d'émetteur est un contact ruban. La comparaison est donnée avec la base 20–30 % de référence.

Profil de Ge	Phase 1				Phase 2	
	20–30 %	10–25 %	10–25 %	10–25 %	10–25 %	10–25 %
Si-cap [nm]	18	18	18	18	16	14
$e_{10\%}$ [nm]	6.7	10	8	6	8	10
Si-cap + $e_{10\%}$ [nm]	24.7	28	26	24	24	24
I_C à 0.75 V [μA]	56	32	34	29	58	56
I_B à 0.75 V [nA]	36	50	41	50	46	44
β à 0.75 V	1566	644	827	579	1261	1280
C_{BE} [pF]	10.1	8.9	9.3	8.8	9.5	9.5
R_{bp} [$k\Omega/\square$]	2.4	2.5	2.6	2.8	2.6	2.6
R_E [Ω]	7.4	8.1	8.1	8.8	7.6	7.1
BV_{EBO} [V]	1.28	2.17	1.94	2.13	1.61	1.58
BV_{CEO} [V]	1.48	1.64	1.61	1.66	1.55	1.54
f_T [GHz]	230	204	210	206	219	226
f_{MAX} [GHz]	270	255	258	254	265	264

TAB. III.10 – Comparaison des principaux paramètres électriques de TBH ayant le profil de base 20–30 de référence et les nouveaux profils de base 10–25 % ($0.15 \times 3.6 \mu\text{m}^2$).

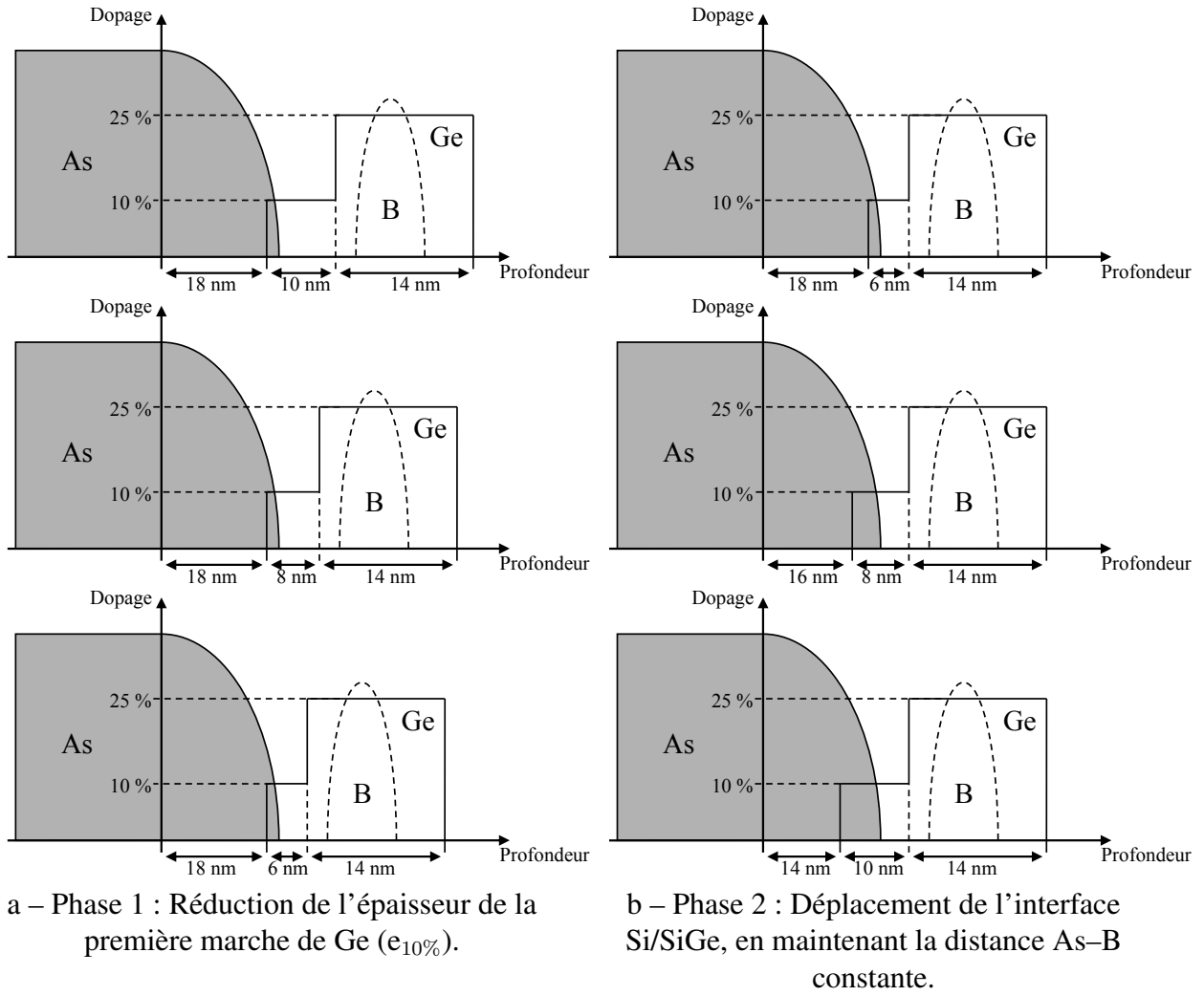


FIG. III.29 – Présentation des variations réalisées sur le profil de SiGe pour l'optimisation de la nouvelle base.

On note que la réduction du taux de Ge à l'entrée de la base permet de réduire la capacité E/B, au détriment d'une légère augmentation de la résistance de base pincée. Durant la première phase de l'optimisation (réduction de $e_{10\%}$ seule), les performances se dégradent légèrement : On constate une augmentation de R_E ayant pour conséquence une dégradation de f_T et f_{MAX} par rapport au profil 20–30 % de référence. Une jonction moins abrupte en est vraisemblablement la cause comme en témoignent les évolutions de C_{BE} et BV_{EBO} . Le courant collecteur est réduit d'un facteur 2 par rapport au profil 20–30 %, et le courant de base augmente significativement, ce qui permet de faire chuter le gain de façon assez importante (de 1566 pour la référence à 579 pour la base avec $e_{10\%} = 6$ nm). La tension de claquage BV_{CEO} augmente de 1.48 à 1.66 V grâce à ces deux effets conjugués.

Durant la deuxième phase d'optimisation, on déplace l'hétérojonction Si/SiGe vers l'émetteur en gardant la distance As–B constante. On note alors que les performances du composant s'améliorent : La résistance d'émetteur diminue, car le Ge a tendance à favoriser la diffusion de l'arsenic. Les fréquences f_T et f_{MAX} sont améliorées, par cette diminution de R_E , et l'évolution de C_{BE} et BV_{EBO} prouvent que la jonction E/B devient plus abrupte lorsque l'épaisseur du Si-cap diminue et $e_{10\%}$ augmente. Au terme de cette deuxième phase, les performances obtenues avec le profil 10–25 % sont d'un niveau similaire au profil 20–30 % de référence : f_T atteint 226 GHz et f_{MAX} 264 GHz, à comparer aux 230 / 270 GHz de la référence. Le gain est légèrement réduit, favorisant ainsi la tension de claquage BV_{CEO} . Il a donc été possible de

démontrer des performances similaires au profil 20–30 % avec une base 10–25 % de Ge.

En passant du profil de base 20-30 % à 10–25 %, on peut maintenir des valeurs de I_C et des valeurs de f_T et f_{MAX} comparables en réduisant l'épaisseur du Si-cap. La réduction de gain en courant est due à l'augmentation de I_B , car l'hétérojonction en bande de valence de la jonction E/B a diminué. La réduction du taux de Ge à l'entrée de la base a permis de diminuer le courant tunnel bande-à-bande jusqu'à un niveau négligeable.

III.4 Optimisation du profil collecteur

Dernière partie active du transistor bipolaire, le collecteur a également une influence importante sur les performances statiques et dynamiques du composant. Le but de cette section est de faire apparaître le compromis existant entre forte fréquence de transition et tension de claquage élevée.

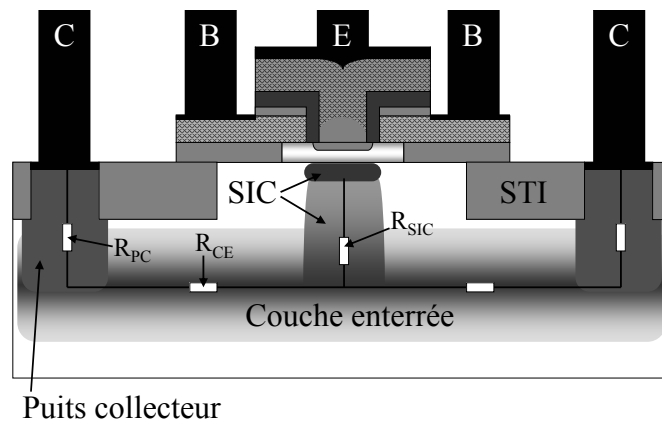


FIG. III.30 – Décomposition de la résistance collecteur en trois parties distinctes.

Le collecteur est divisé en trois parties distinctes, qui sont le SIC (*Selectively Implanted Collector*), la couche enterrée et le puits collecteur, comme indiqué dans la figure III.30. Cette dernière partie est composée d'une implantation visant à assurer le contact entre la couche enterrée et le siliciure en surface. Les différentes parties du collecteur doivent être le moins résistives possible afin de limiter la résistance de collecteur globale et permettre des fréquences de transition élevées. Cependant, la partie la plus proche de la base formée par le SIC a une influence directe sur les caractéristiques du composant, outre la résistance collecteur : Le niveau et le profil de dopage ont une forte influence sur la valeur de la capacité base/collecteur C_{BC} , sur le seuil de l'effet Kirk (cf. chapitre I) limitant la montée en fréquence, sur le temps de transit dans la jonction, et sur la tension d'avalanche de la jonction. Dès lors, il est primordial de réaliser des puits collecteurs et des couches enterrées les plus dopées possible pour diminuer R_C , et une attention particulière doit être apportée au dopage du SIC comme paramètre d'ajustement entre R_C , C_{BC} , f_T , le seuil d'effet Kirk et la tenue en tension.

III.4.1 Collecteur implanté sélectivement (SIC)

Différents profils SIC ont été réalisés afin d'évaluer l'impact du dopage sur les performances du composant. Les variations ont porté sur la dose et l'énergie d'implantation, ainsi que le type de dopant (As ou P). L'arsenic étant un atome plus lourd que le phosphore, sa diffusion est réduite, ce qui permet de diminuer la capacité C_{BC} par un profil de dopage mieux contrôlé [Chevalier04]. La réduction de diffusion latérale est également réduite comme cela est présenté dans la figure III.31 qui illustre cet effet bénéfique de l'arsenic sur la capacité base/collecteur.

Le SIC dopé phosphore est implanté en deux fois : Une première implantation à 200 keV avec une dose de $6 \cdot 10^{13} \text{ at} \cdot \text{cm}^{-2}$ est destinée à assurer le contact avec la couche enterrée. La deuxième implantation est plus proche de la surface et est réalisée avec une dose et une énergie beaucoup plus faibles (80 keV, $6 \cdot 10^{12} \text{ at} \cdot \text{cm}^{-2}$). Le phosphore étant très mobile, le collecteur a tendance à s'appauvrir en surface, cette deuxième implantation est nécessaire pour assurer un niveau de dopage suffisant à la jonction B/C.

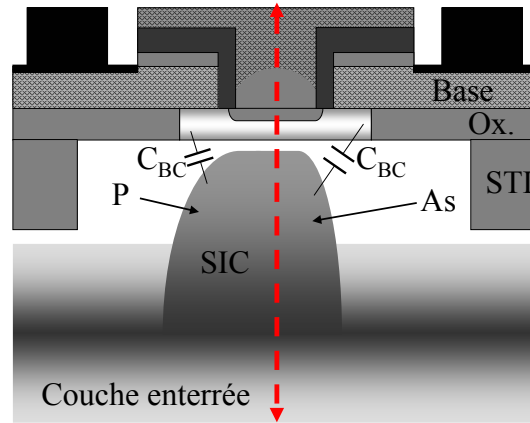


FIG. III.31 – Effet d'amélioration de la capacité C_{BC} par réduction de la diffusion latérale des dopants en passant du phosphore à l'arsenic pour le dopage du SIC.

Avec l'emploi d'un dopage arsenic, une seule implantation à forte dose et à forte énergie (dose comprise entre 4 et $8 \cdot 10^{13}$ $\text{at}\cdot\text{cm}^{-2}$, avec une énergie de 270 keV) est nécessaire. L'arsenic diffusant moins que le phosphore, le profil obtenu dans le cas de l'As sera plus abrupt.

III.4.2 Résultats électriques

Le tableau III.11 présente les caractéristiques électriques de transistors bipolaires disposant de plusieurs profils SIC. Des collecteurs dopés As à différentes doses sont comparés au collecteur dopé phosphore de référence. Un essai sans SIC a également été réalisé, et permet de mettre en évidence l'importance du collecteur sur les performances du composant. La dimension des composants est $0.15 \times 3.6 \mu\text{m}^2$, l'émetteur est dopé à 280 sccm, avec un Si-cap de 18 nm. Le contact d'émetteur est un contact en ruban. La base a un profil 20–30 % de 20 nm d'épaisseur.

Dopage	P	As	As	As	sans SIC
Dose 1 [$\text{at}\cdot\text{cm}^{-2}$]	$6 \cdot 10^{13}$	$4 \cdot 10^{13}$	$6 \cdot 10^{13}$	$8 \cdot 10^{13}$	-
Dose 2 [$\text{at}\cdot\text{cm}^{-2}$]	$6 \cdot 10^{12}$	-	-	-	-
I_C à 0.75 V [μA]	246	237	221	206	229
β à 0.75 V	3024	2910	3010	2860	2784
C_{BC} [fF]	9.2	7.8	8.4	8.7	5.9
BV_{CEO} [V]	1.45	1.52	1.47	1.46	1.90
BV_{CBO} [V]	5.6	6.0	5.6	5.4	7.8
V_{Af} [V]	103	123	91	88	129
f_T [GHz]	240	227	236	240	134
f_{MAX} [GHz]	234	258	248	249	235
$\tau_{EC,0}$ [ps]	0.56	0.58	0.57	0.56	0.88
$f_T \times BV_{CEO}$ [GHz·V]	349	344	347	349	255
$f_{MAX} \times BV_{CBO}$ [GHz·V]	1309	1543	1392	1345	1820

TAB. III.11 – Influence du dopage SIC sur les performances du composant ($0.15 \times 3.6 \mu\text{m}^2$).

On constate que le dopage collecteur a une faible influence sur les niveaux de courant à moyenne injection. En revanche, le dopage du SIC a un impact fort sur l'avalanche dans la jonction B/C : Lorsque le SIC n'est pas présent, BV_{CEO} et BV_{CBO} atteignent respectivement 1.9 V et 7.8 V, à comparer aux 1.45 V et 5.6 V obtenus avec le SIC de référence.

La capacité B/C est fortement dépendante du dopage : on constate une forte diminution de C_{BC} lorsque le dopage collecteur est réduit. A dose égale, le SIC arsenic a une capacité B/C plus faible qu'avec un SIC phosphore, ce qui est expliqué en figure III.31.

Concernant la tension d'Early directe, on note également une dégradation de ce paramètre lorsque le dopage augmente. En effet, la zone de charge d'espace B/C s'étend d'autant plus côté base que le collecteur est dopé, ce qui explique la décroissance de V_{Af} lorsque la dose d'arsenic augmente.

La figure III.32 représente les profils de dopage du collecteur pour un SIC dopé arsenic et un SIC dopé phosphore. Dans le cas de l'arsenic, on ne peut pas faire la distinction entre les atomes dopants de la couche enterrée et du SIC, la jonction entre les deux parties du collecteur s'effectue progressivement. Le dopage net, constitué de la somme des concentrations de phosphore et d'arsenic, est également représenté pour le SIC phosphore. Le SIC phosphore représenté est le SIC de référence implanté en deux fois ($6 \cdot 10^{13} \text{ at}\cdot\text{cm}^{-2}$ à 200 keV puis $6 \cdot 10^{12} \text{ at}\cdot\text{cm}^{-2}$ à 80 keV). Sur la figure III.32, le SIC As est réalisé par une seule implantation de dose $6 \cdot 10^{13} \text{ at}\cdot\text{cm}^{-2}$ à 200 keV.

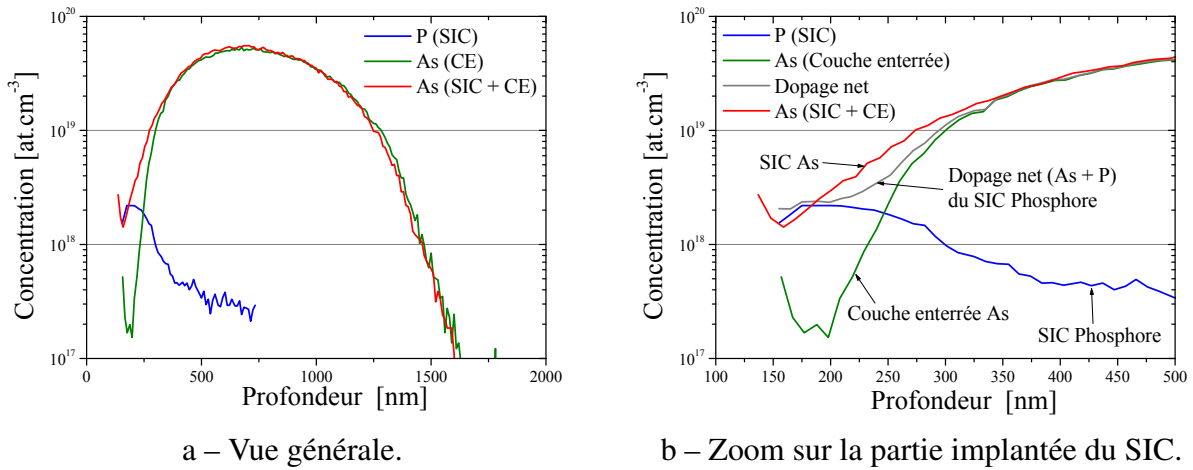


FIG. III.32 – Profils SIMS de collecteurs dopés phosphore ou arsenic. La couche enterrée est réalisée en As dans les deux cas. La dose d'As représentée est $6 \cdot 10^{13} \text{ at}\cdot\text{cm}^{-2}$. Le SIC phosphore est le SIC standard de référence.

Le profil de dopants est plus abrupt pour le SIC arsenic, en conséquence, le dopage net à la jonction B/C semble légèrement plus faible que dans le cas d'un SIC phosphore, ce qui permet de diminuer la capacité C_{BC} . En conclusion, l'arsenic offre un meilleur contrôle du profil de dopants, et permet d'atteindre rapidement le niveau de dopage de la couche enterrée, c'est donc la solution technologique retenue pour réaliser les transistors bipolaires rapides.

III.4.3 Influence du SIC sur les performances dynamiques

D'après le tableau III.11, on note que la fréquence f_T augmente lorsque le dopage collecteur augmente [Chevalier05a]. Les niveaux de courant à moyenne injection étant quasiment les mêmes quels que soient les paramètres du SIC, cette variation est due à la diminution de τ_C et au retard de l'apparition de l'effet Kirk avec l'augmentation du dopage. La figure III.33 représente les évolutions de f_T en fonction de I_C pour différents profils de SIC.

Le temps de transit $\tau_{EC,0}$ évolue légèrement avec la dose de collecteur, comme indiqué dans le tableau III.11. La valeur de $\tau_{EC,0}$ obtenue avec un SIC dopé phosphore est de 0.56 ps. Dans le cas d'un SIC arsenic, $\tau_{EC,0}$ est réduit de 0.58 à 0.56 ps lorsque la dose d'As implantée augmente de 4 à $8 \cdot 10^{13} \text{ at}\cdot\text{cm}^{-2}$. La figure III.34 montre que les 4 composants ayant un SIC

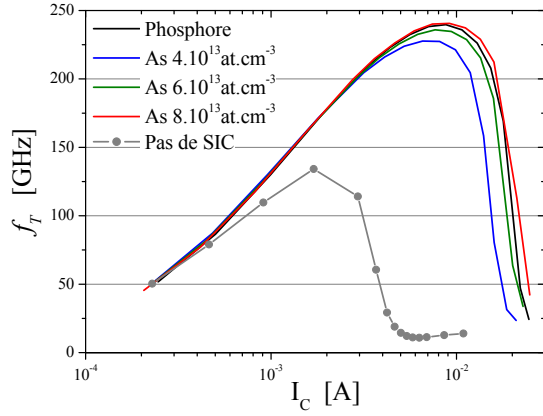


FIG. III.33 – Evolution de f_T en fonction de I_C pour différentes implantations SIC.

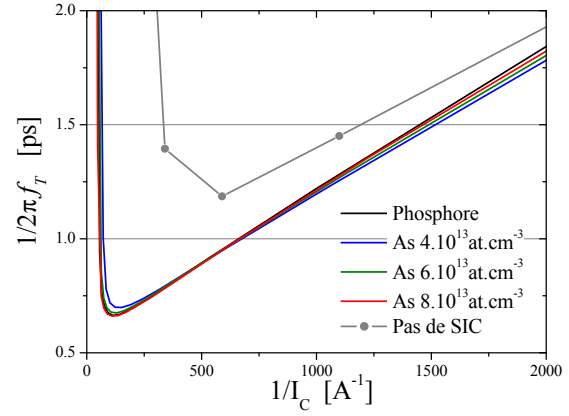


FIG. III.34 – Temps de transit τ_{EC} en fonction de $1/I_C$ pour différents profils de SIC.

ont des caractéristiques très proches. Dans le cas où le SIC n'est pas du tout réalisé, l'effet Kirk intervient très tôt, pour un courant collecteur 10 fois plus faible que lorsque le SIC est présent. Il en résulte une fréquence de transition f_T fortement réduite. Le temps de transit $\tau_{EC,0}$ atteint 0.88 ps, ce qui est due à l'augmentation de τ_C et τ_{BC} , ainsi qu'à la dégradation de la résistance collecteur.

L'évolution de f_{MAX} en fonction de la dose de collecteur indique la sensibilité de ce paramètre à la valeur de la capacité C_{BC} . En effet, lorsque le dopage diminue, f_T chute à cause de l'effet Kirk, mais f_{MAX} s'améliore. Le dopage SIC arsenic à $8 \cdot 10^{13}$ at.cm $^{-2}$ offre la même fréquence f_T que le SIC phosphore avec toutefois une amélioration de f_{MAX} de 15 GHz. Lorsque le SIC est totalement absent, f_T chute de 240 à 134 GHz, f_{MAX} restant constant, ce qui est dû à une réduction de C_{BC} de 9.2 à 5.9 fF. La configuration donnant le meilleur compromis entre f_T et f_{MAX} serait d'avoir un SIC très dopé avec un profil rétrograde, et une très faible extension latérale pour minimiser C_{BC} .

III.4.4 Compromis entre f_T et BV_{CEO}

Il existe un compromis entre la fréquence de transition f_T et la tension d'avalanche BV_{CEO} . Un fort dopage collecteur favorise f_T en diminuant τ_C et en retardant l'effet Kirk, mais la tension d'avalanche en est réduite. Afin de caractériser ce compromis, on introduit le facteur de mérite $f_T \times BV_{CEO}$, qui se mesure en GHz·V. Lorsque la dose collecteur augmente, la fréquence f_T augmente également alors que BV_{CEO} diminue, ce qui assure un produit $f_T \times BV_{CEO}$ quasiment constant, de l'ordre de 350 GHz·V, excepté pour le composant sans SIC, dont les performances sont nettement inférieures (250 GHz·V). Le fait que le composant sans SIC est en retrait par rapport aux autres est dû à la dégradation de la résistance collecteur, qui dégrade f_T et diminue le produit $f_T \times BV_{CEO}$ par rapport aux autres composants.

La figure III.35 montre l'évolution de f_T en fonction de BV_{CEO} pour les composants étudiés dans ce chapitre. Lorsque la dose du SIC diminue, on se déplace à produit $f_T \times BV_{CEO}$ quasiment constant vers les fortes tensions de claquage et les fréquences de transition plus faibles.

Cette optimisation met en avant le fait qu'il est difficile d'améliorer dans le même temps les fréquences de transition et la tenue en tension du TBH SiGe : Pour obtenir de fortes performances dynamiques, il est nécessaire de réaliser des profils de dopages très abrupts, ce qui dégrade la tenue en tension. Dès lors, afin d'augmenter le produit $f_T \times BV_{CEO}$, il faut trouver des solutions en rupture avec l'optimisation classique du composant. C'est l'approche retenue au chapitre IV, traitant de l'optimisation du produit $f_T \times BV_{CEO}$ par des architectures visant à

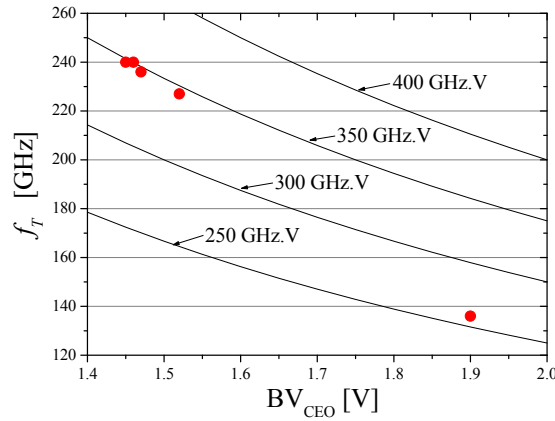


FIG. III.35 – Evolution de f_T en fonction de BV_{CEO} pour des composants ayant différentes doses de SiC.

retarder l'inversion du courant de base, et donc augmenter la tension BV_{CEO} .

III.5 Réduction des dimensions latérales du dispositif

Afin d'obtenir des performances ultimes, nous avons démontré que l'utilisation de profils de dopants très abrupts était nécessaire. Ceci permet de réaliser des bases très fines, autorisant de faibles temps de transit. Cependant, les résistances d'accès, notamment la résistance de base, sont dégradées par cette réduction de l'épaisseur du composant. Réduire les dimensions latérales du transistor apparaît donc comme une nécessité et un moyen efficace de diminuer les résistances d'accès, notamment la résistance de base.

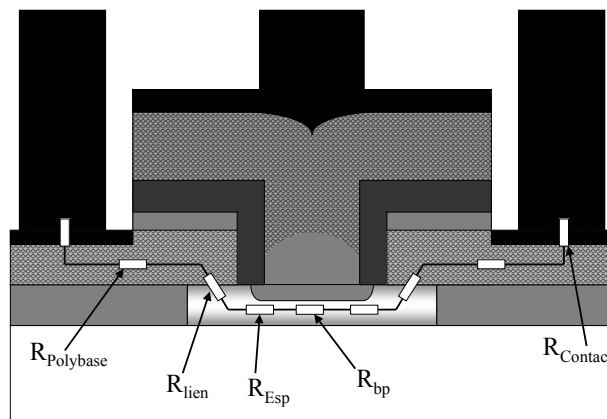


FIG. III.36 – Représentation des différentes contributions à la résistance de base [Schröter00].

La figure III.36 montre la décomposition de la résistance de base proposée dans le modèle HiCUM [Schröter00]. Les différentes contributions sont les suivantes :

- $R_{Contact}$ est la résistance du contact métallique sur le polybase. Ce contact est siliciuré afin de diminuer la résistance de l'interface entre les deux matériaux.
- $R_{Polybase}$ est la résistance du barreau de polysilicium visant à assurer le lien entre le contact et la base extrinsèque.

- R_{lien} est la résistance d'interface entre le polysilicium constituant la base extrinsèque et la base monocristalline déposée par épitaxie sélective.
- R_{Esp} est la contribution à la résistance de base située sous les espaceurs, donc dépend de la résistivité de la couche de base SiGe.
- R_{bp} est la résistance de base pincée.

La contribution R_{Polybase} dépend surtout du recouvrement du polyémetteur sur le polybase. En effet, la partie du polybase qui n'est pas située sous l'empilement de diélectriques entre la base et l'émetteur est siliciurée, donc la résistance du polybase compte peu.

La contribution la plus importante à la résistance de base est R_{Esp} , il s'agit de la partie de SiGe située sous les espaceurs internes, la base étant partiellement pincée par l'émetteur et le collecteur. Il est donc primordial de réduire la distance entre le polybase et la base intrinsèque, afin de diminuer la résistance de base, et ainsi augmenter f_{MAX} . La figure III.37 définit les dimensions latérales du transistor qui sont étudiées dans ce travail. La fenêtre émetteur est dénommée *EmWin* pour *Emitter Window*.

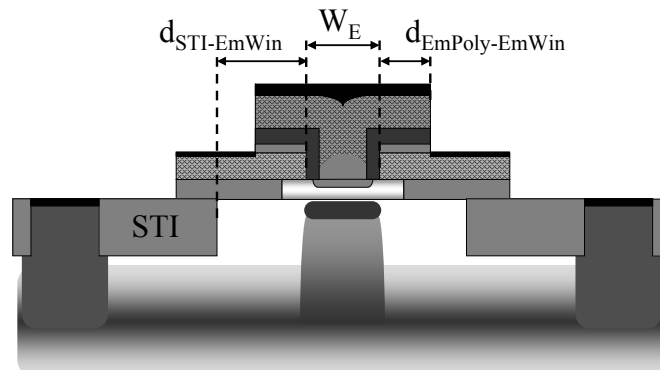


FIG. III.37 – Dimensions latérales du transistor bipolaire.

III.5.1 Largeur d'émetteur W_E

La réduction de la fenêtre émetteur permet de diminuer fortement la résistance de base et la capacité C_{BC} , qui limitent principalement la fréquence f_{MAX} . W_E peut être réduit de différentes façons :

- Soit par une réduction de la fenêtre émetteur en elle-même. Dans ce cas, les paramètres de la photolithographie sont modifiés pour avoir une fenêtre émetteur plus étroite. L'ouverture standard est de $0.30 \mu\text{m}$, elle a été réduite à 0.27 et $0.25 \mu\text{m}$.
- Soit par le biais d'espaceurs internes : En augmentant la largeur des espaceurs, on réduit la surface active de l'émetteur. Notamment la réduction de topologie du TBH a imposé le passage d'espaceurs en «L» à des espaceurs en «D», qui offrent un meilleur contrôle lors de la gravure et une largeur d'émetteur plus faible.

III.5.1.a Contact d'émetteur

Le polyémetteur est déposé de manière conforme dans la fenêtre émetteur. Ainsi, selon la largeur de celle-ci, ou bien selon la présence d'espaceurs internes plus ou moins volumineux, la conformation du contact d'émetteur varie fortement. Si la fenêtre est très large, le métal déposé au dessus de l'émetteur sera très proche de la jonction E/B, donc R_{Ep} sera faible (figure III.38–a). A l'inverse, si l'émetteur est étroit, le métal sera fortement éloigné de la zone active du composant, entraînant une forte contribution R_{Ep} (figure III.38–b).

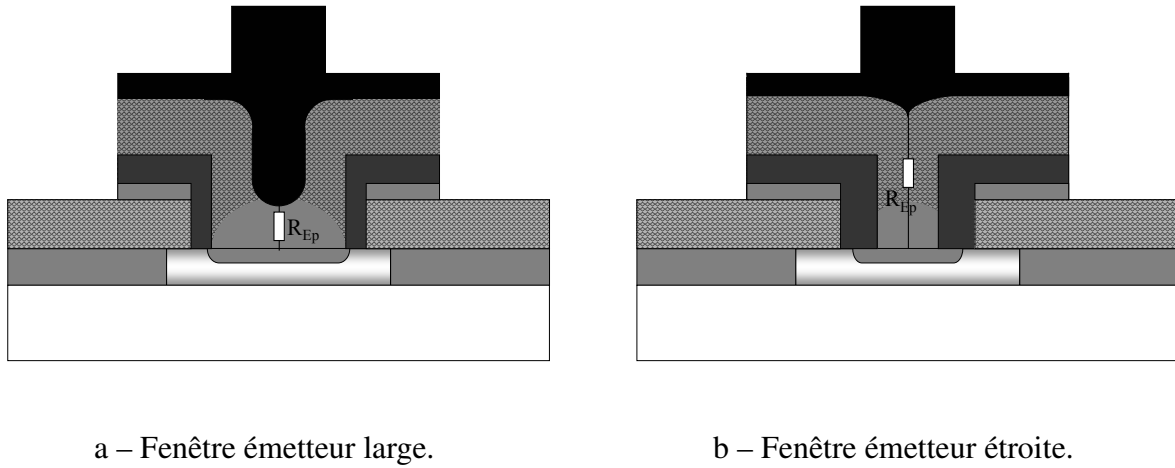


FIG. III.38 – Effet de la largeur de la fenêtre émetteur sur le contact métallique d'émetteur.

III.5.1.b Largeur de la fenêtre émetteur

La première réduction latérale étudiée est la largeur de la fenêtre émetteur. Le tableau III.12 compare les évolutions des principaux paramètres électriques de transistors bipolaires de $3.6\ \mu\text{m}$ de long en fonction de la largeur de fenêtre émetteur W_E (figure III.37). La largeur effective d'émetteur W_{Eff} est également indiquée. Elle est calculée d'après l'évolution des niveaux de courant en considérant que la densité de courant reste constante.

$W_E\ [\mu\text{m}]$	0.30	0.27	0.25
$W_{\text{Eff}}\ [\mu\text{m}]$	0.17	0.15	0.11
$I_C\ \text{à}\ 0.75\ \text{V}\ [\mu\text{A}]$	144	134	87
$J_C\ \text{à}\ 0.75\ \text{V}\ [\mu\text{A}/\mu\text{m}^2]$	235	250	220
$\beta\ \text{à}\ 0.75\ \text{V}$	2750	2850	2680
$R_E\ [\Omega]$	9.4	10.2	11.8
$R_{\text{bp}}\ [\text{k}\Omega/\square]$	2.4	2.4	2.3
$C_{\text{BE}}\ [\text{fF}]$	10.7	10.5	9.5
$C_{\text{BC}}\ [\text{fF}]$	9.2	8.9	8.4
$f_T\ [\text{GHz}]$	238	234	236
$f_{\text{MAX}}\ [\text{GHz}]$	269	293	297

TAB. III.12 – Paramètres électriques de transistors bipolaires en fonction de la largeur d'émetteur définie par la photolithographie ($L_E = 3.6\ \mu\text{m}$).

La réduction du courant collecteur avec W_E est due à la diminution de la surface active. La surface de la jonction émetteur/base étant réduite, on observe une diminution de la capacité C_{BE} . Le gain et la densité de courant restent quasiment constants pour les trois composants, signe que le réglage de la jonction E/B n'a pas été fondamentalement modifié avec la réduction de la fenêtre émetteur.

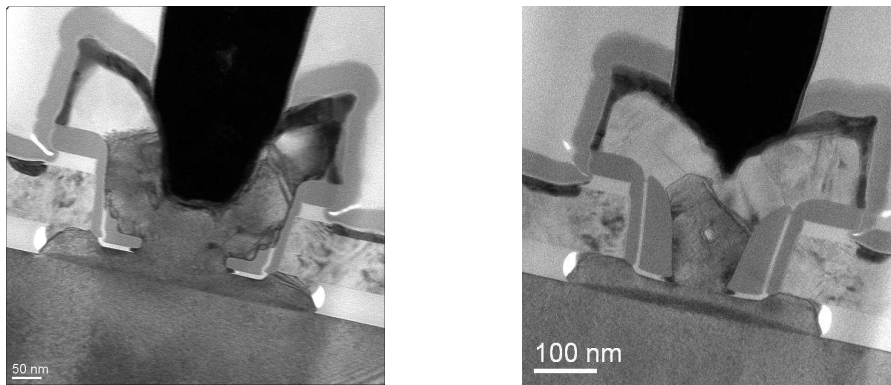
On constate que la réduction de C_{BE} avant tout due à la diminution de la surface de l'émetteur s'accompagne d'une augmentation de la résistance d'émetteur, due à une modification de la conformation du polyémetteur, comme indiqué dans la section III.5.1.a. La réduction de capacité est imputable à la réduction de la surface de jonction.

En ce qui concerne les performances dynamiques, la fréquence f_{MAX} augmente de 30 GHz pour atteindre 300 GHz dans le cas de la fenêtre émetteur la plus étroite. La fréquence f_T reste quasi constante (de 238 à 236 GHz), la réduction de capacité C_{BE} et de l'auto-échauffement

permet de compenser l'augmentation de la résistance d'émetteur. La valeur maximale de f_T est également atteinte à un courant collecteur plus faible lorsque la largeur d'émetteur est réduite, car la densité de courant reste quasiment constante.

III.5.1.c Espaceurs internes

La figure III.39 représente des photos au microscope électronique à transmission de deux transistors bipolaires utilisant deux types d'espaceurs internes. La taille de l'espaceur en L est fixée par l'épaisseur de l'oxyde, qui est ensuite éliminé par une gravure humide avant le dépôt du polyémetteur. L'espaceur en D est déterminé par le dépôt nitrure le constituant. En adaptant la durée de la gravure, on arrive à obtenir des espaceurs en D très larges, permettant de diminuer la largeur effective d'émetteur W_{Eff} jusqu'à 80 nm, obtenus en réduisant la largeur de la fenêtre émetteur.



a – Espaceurs en «L» ($W_E = 0,27 \mu\text{m}$) b – Espaceurs en «D» ($W_E = 0,25 \mu\text{m}$)

FIG. III.39 – Photos MET de deux TBH utilisant les deux types d'espaceurs internes.

Ce type d'espaceurs en D offre une capacité E/B réduite par rapport aux espaceurs en L, car la distance entre les polysilicium de base et d'émetteur est agrandie, comme indiqué dans la figure III.40.

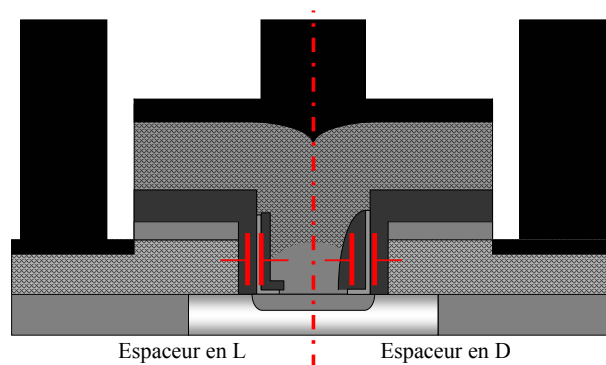


FIG. III.40 – Schéma représentant les deux types d'espaceurs (en L et en D). La capacité C_{BE} est plus faible dans le cas d'un espaceur en D.

Concernant les résultats électriques, l'espaceur en D est moins favorable que l'espaceur en L : On observe systématiquement une dégradation de la résistance d'émetteur, à cause de l'effet détaillé dans la section III.5.1.a, ce qui pénalise la fréquence f_T .

Le tableau III.13 illustre l'effet de décroissance de f_T avec l'utilisation d'espaceurs en D. La résistance d'émetteur augmente d'autant plus que W_E est faible. La réduction de la capacité C_{BE}

W_E [μm]	0.27		0.25	
Espaceur	L	D	L	D
W_{Eff} [μm]	0.13	0.13	0.10	0.10
I_C à 0.75 V [μA]	72	73	56	56
β à 0.75 V	1090	900	1040	860
R_E [Ω]	18	22	25	32
C_{BE} [fF]	10.5	9.6	9.5	8.3
f_T [GHz]	220	198	214	185

TAB. III.13 – Evolution des performances électriques en fonction du type d'espaceur pour deux largeurs de fenêtre émetteur ($L_E = 3.6 \mu\text{m}$).

n'est pas suffisante pour contrebalancer l'augmentation de R_E , ce qui dégrade les performances. Néanmoins, ce type d'espaceur est nécessaire pour une intégration dans une filière BiCMOS, la hauteur totale du composant devant être réduite pour être compatible avec des transistors MOS, les espaceurs en L ne peuvent être réalisés. La résistance d'émetteur est réduite grâce à la diminution de l'épaisseur de polyémetteur et de polybase, ce qui permet de rapprocher le métal des contacts de la jonction E/B, comme indiqué par la figure III.41.

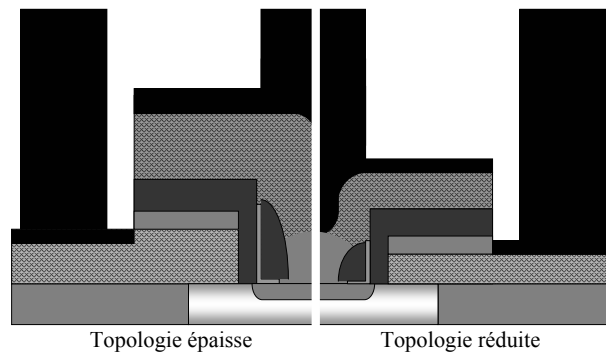


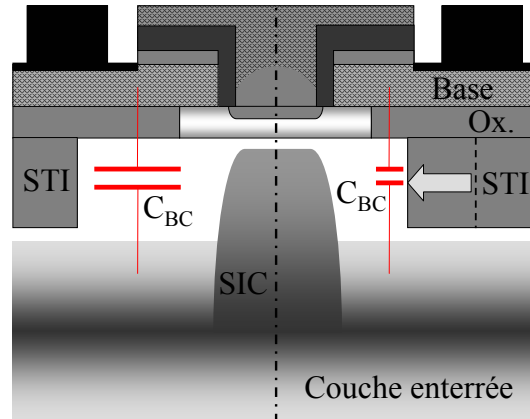
FIG. III.41 – Comparaison d'une topologie épaisse avec une topologie réduite.

Les performances obtenues récemment avec des espaceurs en D et une topologie réduite sont similaires aux performances d'un transistor doté d'espaceurs en L (260 et 280 GHz respectivement pour f_T et f_{MAX}) [Chevalier07].

III.5.2 Largeur de la zone d'active

La dimension de la zone d'active est fixée par la distance existant entre le bord de la fenêtre émetteur et le bord du STI (figure III.37). En standard, cette distance est de $0.2 \mu\text{m}$, elle a été réduite à 0.15 puis $0.1 \mu\text{m}$. La largeur de la zone d'active, dans laquelle la partie intrinsèque du transistor est réalisée, permet de jouer sur la capacité C_{BC} . Lorsque la dimension du STI augmente, le recouvrement entre la base et le collecteur diminue, ce qui diminue la capacité statique. La figure III.42 schématise cet effet.

Le tableau III.14 résume les performances obtenues en fonction de la distance STI-EmWin, pour trois composants de dimensions d'émetteur $0.15 \times 3.6 \mu\text{m}^2$. La base a une épaisseur de 20 nm , un profil 20–30 % de Ge, et le Si-cap fait 16 nm d'épaisseur. On constate une diminution de C_{BC} avec le rapprochement du STI. Les autres paramètres statiques (résistances séries et capacités) n'évoluent pas avec $d_{\text{STI-EmWin}}$. L'évolution des fréquences de transition est assez faible : f_T augmente de 5 GHz pour atteindre 271 GHz , et f_{MAX} augmente de 278 à 281 GHz .

FIG. III.42 – Influence de la largeur de la zone active sur la capacité C_{BC} .

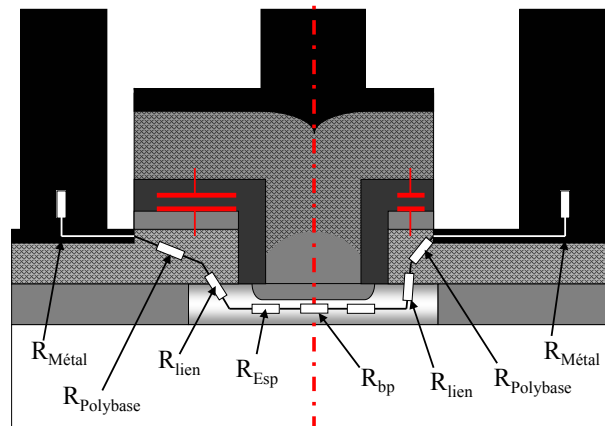
$d_{STI-EmWin}$	0.2	0.15	0.1
C_{BC} [fF]	9	8.8	8.4
f_T [GHz]	266	270	271
f_{MAX} [GHz]	278	278	281

TAB. III.14 – Performances dynamiques en fonction de la distance STI-EmWin. Dimensions des transistors : $0.15 \times 3.6 \mu m^2$.

La réduction de la capacité C_{BC} n'est pas suffisante pour améliorer significativement les performances du composant, car la valeur de C_{BC} est déjà fortement réduite par la présence de l'oxyde piédestal.

III.5.3 Largeur du polyémetteur.

Un autre paramètre ayant une influence plus sensible sur les caractéristiques du composant est la largeur du polyémetteur, déterminée par la distance entre la fenêtre émetteur et le bord du polysilicium ($d_{EmPoly-EmWin}$ sur la figure III.37). La réduction de la taille du polysilicium d'émetteur (*EmPoly* pour *E*mitter *P*olysilicon) permet d'améliorer significativement la capacité de recouvrement entre l'émetteur et la base, ainsi que la résistance de base, comme indiqué dans la figure III.43.

FIG. III.43 – Schéma expliquant l'influence de la largeur du polyémetteur sur la capacité C_{BE} et la résistance de base.

La distance $d_{\text{EmPoly-EmWin}}$ standard est $0.1 \mu\text{m}$. Elle a été réduite à 0.07 puis $0.05 \mu\text{m}$ pour les besoins de l'étude. Le tableau III.15 résume les performances obtenues en faisant varier la taille du polyémetteur.

$d_{\text{EmPoly-EmWin}}$	0.1	0.07	0.05
C_{BE} [fF]	12.4	12.3	12.2
R_{B} [Ω]	19.5	17.3	16.6
f_T [GHz]	266	270	271
f_{MAX} [GHz]	278	280	286

TAB. III.15 – Performances dynamiques en fonction de la distance STI-EmWin. Dimensions des transistors : $0.15 \times 3.6 \mu\text{m}^2$.

On constate que la réduction de $d_{\text{EmPoly-EmWin}}$ permet de diminuer sensiblement la résistance de base (de 19.5 à 16.6Ω), alors que la capacité C_{BE} diminue faiblement (de 12.4 à 12.2 fF). La composante de la résistance de base la plus impactée par la réduction de $d_{\text{EmPoly-EmWin}}$ est la résistance de polysilicium R_{Polybase} . Principalement, la réduction de R_{B} est à l'origine du gain de performances dynamiques : f_T s'améliore légèrement, alors que f_{MAX} augmente de 278 à 286 GHz .

III.5.4 Synthèse de l'optimisation latérale.

Cette étude sur l'optimisation latérale du transistor bipolaire a permis de mettre l'accent sur les principaux paramètres limitant la montée en fréquence. Avec la réduction des dimensions verticales du composant, la limitation première à la montée en fréquence est la résistance de base extrinsèque et intrinsèque. On s'aperçoit que l'effet des capacités de recouvrement est somme toute assez faible. Notamment, la réduction de la capacité C_{BC} par la variation de la distance STI-EmWin ne présente qu'une influence limitée sur les fréquences de transition.

La réduction de la largeur de la fenêtre émetteur permet de diminuer sensiblement la résistance de base, ce qui autorise une augmentation forte de f_{MAX} . La figure III.44 présente l'évolution en fonction de I_{C} de f_T et f_{MAX} pour les trois largeurs d'émetteur étudiées précédemment. On constate que l'augmentation de f_{MAX} s'accompagne d'une diminution du courant auquel le maximum de f_T est atteint. La surface effective du composant étant réduite, l'auto-échauffement est également réduit, ce qui permet d'améliorer les performances dynamiques. Le chapitre V.2 propose une étude plus poussée de l'auto-échauffement dans les dispositifs rapides.

La figure III.45 présente les évolutions de f_T et f_{MAX} en fonction de I_{C} pour les valeurs de $d_{\text{EmPoly-EmWin}}$ étudiées précédemment. La réduction du recouvrement du polyémetteur sur le polybase permet d'améliorer les deux fréquences de transition. Concernant le temps de transit total $\tau_{\text{EC Min}}$, on observe peu de différence entre les composants : Lorsqu'on réduit la largeur de fenêtre émetteur de 0.3 à $0.25 \mu\text{m}$, le temps de transit reste égal à 0.56 ps . De même, en diminuant la distance $d_{\text{EmPoly-EmWin}}$, le temps de transit varie de moins de dix fs (de 0.49 à 0.48 ps). Ainsi, les éléments extrinsèques comme les capacités de recouvrement et les résistances d'accès ne jouent pas sur le temps de transit des porteurs. Ces éléments parasites doivent donc être réduits au maximum car ils limitent dans une certaine mesure la montée en fréquence. La réduction des dimensions latérales reste donc un passage obligé de l'optimisation du transistor bipolaire.

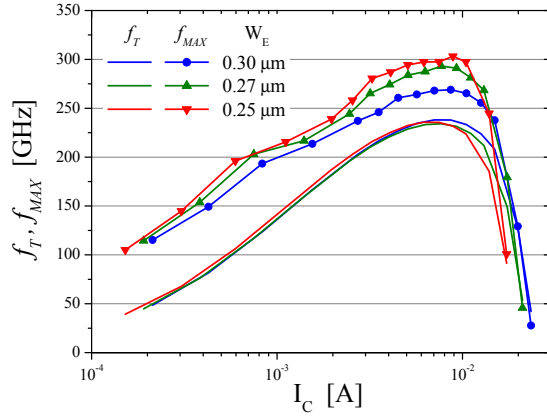


FIG. III.44 – Evolution de f_T et f_{MAX} en fonction de I_C pour différentes largeurs de fenêtre émetteur W_E . Longueur des dispositifs : 3.6 μm .

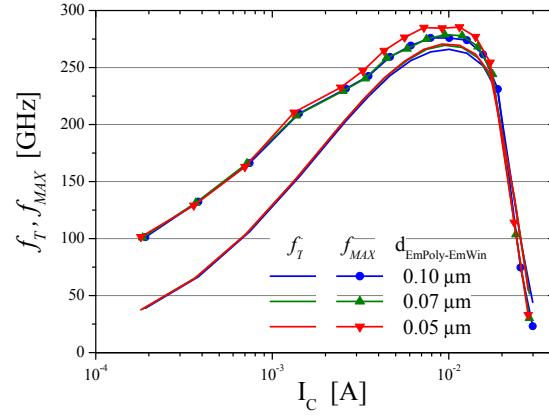


FIG. III.45 – Evolution de f_T et f_{MAX} en fonction de I_C pour différentes distances $d_{\text{EmPoly-EmWin}}$. Longueur des dispositifs : 3.6 μm .

III.6 Synthèse de l'optimisation conventionnelle

III.6.1 Performances

Ce chapitre trace l'historique de trois ans d'optimisation conventionnelle du transistor bipolaire. Les trois parties intrinsèques du composant ont été étudiées, les limitations des différentes améliorations possibles ont également été abordées. L'optimisation verticale du transistor reste le principal levier pour améliorer les fréquences de transition du TBH :

- L'émetteur limite principalement les performances du composant par sa résistance d'accès. L'emploi d'émetteurs de plus en plus dopés permet de diminuer R_E , et donc d'améliorer les performances dynamiques. Cette étude a également fait apparaître l'importance de la diffusion des dopants vers le Si-cap, la principale composante de la résistance d'émetteur étant la zone de diffusion des dopants vers la base.
- L'étude de la base fait apparaître un compromis entre les fréquences f_T et f_{MAX} . S'il est nécessaire d'avoir des bases fines pour diminuer le temps de transit des porteurs, il faut toutefois prendre garde à ne pas dégrader la résistance de base. L'optimisation du profil de germanium permet de repousser ce compromis : L'emploi de forts taux de Ge permet d'augmenter le niveau de courant collecteur en abaissant la barrière pour les électrons à la jonction E/B. En contrepartie, il est possible d'utiliser des bases plus dopées pour réduire R_B . Le Germanium a pour propriété de bloquer efficacement la diffusion du bore (au même titre que le carbone), cependant le dopage de base ne peut être augmenté trop fortement sous peine de voir la base s'élargir par diffusion à cause des forts budgets thermiques subis lors de la fabrication. Une possibilité intéressante pour obtenir des profils de composants très dopés et très abrupts est de réaliser des transistors à très faibles budgets thermiques, ou notamment le recuit final est fortement réduit. Pour l'instant, ces transistors n'ont pas démontré de performances supérieures aux TBH traditionnels.
- Le dopage du collecteur, et principalement du collecteur implanté sélectivement (SIC), joue sur les tensions d'avalanche dans la jonction B/C, ainsi que sur le seuil d'effet Kirk. Un fort dopage favorise de fortes fréquences de transition, mais la tension de claquage BV_{CEO} est dégradée. Là encore, la notion de compromis f_T - BV_{CEO} est importante : Lors de l'optimisation du transistor bipolaire, le produit $f_T \times BV_{CEO}$ est un paramètre difficile à améliorer par des méthodes conventionnelles. Le chapitre IV propose des solu-

tions technologiques en rupture avec les méthodes traditionnelles présentées dans ce chapitre, afin d'améliorer la tenue en tension du composant, sans dégrader les performances dynamiques.

Concernant la dernière partie du chapitre, l'étude de la réduction latérale a mis en avant le fait qu'il est encore possible d'améliorer les performances du composant en diminuant les résistances d'accès et les capacités de recouvrement. La réduction de la largeur de la fenêtre émetteur permet d'améliorer sensiblement les fréquences de transition f_T et f_{MAX} .

Il faut cependant noter que l'augmentation des capacités de jonction n'est pas forcément pénalisante pour le fonctionnement dynamique du TBH. En effet, une forte capacité est signe d'un profil de dopant abrupt, donc d'un temps de transit des porteurs faible, ce qui favorise f_T . D'après la figure I.43, l'augmentation des capacités C_{BE} et C_{BC} a pour effet d'augmenter la pente de la courbe de τ_{EC} en fonction de $1/I_C$. Couplée à de forts niveaux de courant, cette augmentation de capacité permet d'obtenir des temps de transits très faibles, comme indiqué sur la figure III.46. L'amélioration des performances s'est accompagnée d'une augmentation du terme $\frac{kT}{q}(C_{BE} + C_{BC})$, le temps de transit total est réduit de 0.2 ps. Ainsi, pour l'optimisation future du transistor bipolaire, il est plus favorable de réduire les résistances d'accès, qui ont uniquement une influence néfaste, que les capacités de jonction.

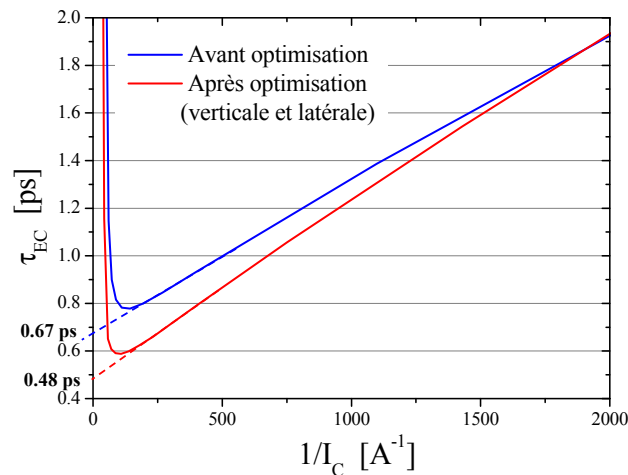


FIG. III.46 – Synthèse de l'évolution des temps de transit au cours de l'étude.

Cette optimisation verticale et latérale a permis une augmentation sensible des performances des composants : D'une fréquence de transition f_T inférieure à 200 GHz, nous avons atteint 270 GHz. Dans le même temps, sous l'impulsion de la réduction des résistances d'accès, f_{MAX} est passé d'environ 200 GHz à 300 GHz. La figure III.47 représente les améliorations de f_T et f_{MAX} par l'optimisation conventionnelle du transistor bipolaire. On observe que l'augmentation de f_T s'est toujours accompagnée d'une amélioration de f_{MAX} . La réduction des dimensions du dispositif apparaît comme étant le moyen le plus efficace d'améliorer les performances dynamiques du transistor bipolaire, notamment f_{MAX} : On note en effet que la plupart des composants sont au dessus de la droite pointillée, donc que $f_{MAX} > f_T$.

III.6.2 Développement d'une filière dédiée aux applications millimétriques

Les applications visées par un composant démontrant de telles performances sont les réseaux WLAN à 60 GHz, les radars anti-collision à 77 GHz et les communications optiques à 80 Gb/s.

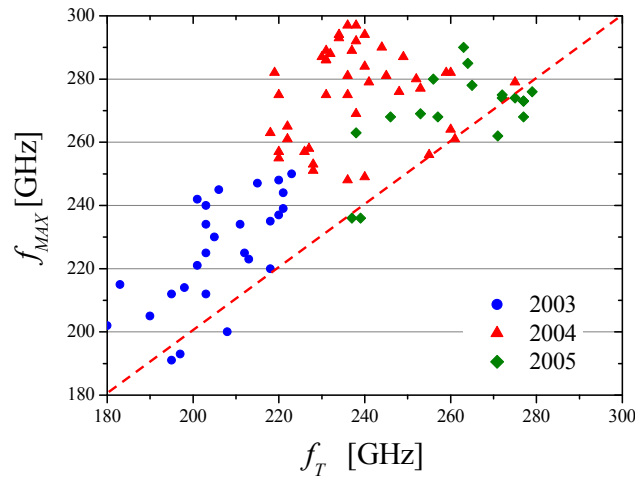


FIG. III.47 – Evolution de f_T et f_{MAX} par l'optimisation conventionnelle du TBH. La ligne pointillée représente les points vérifiant la condition $f_{MAX} = f_T$.

Il a été démontré que ces applications ne pouvaient pas être adressées par la technologie BiCMOS9, la plus récente chez ST [Chevalier07]. La principale limitation est en effet la fréquence de transition trop faible de BiCMOS9 (160 GHz), une fréquence supérieure à 200 GHz est indispensable pour satisfaire à ce type d'applications.

La structure FSA-SEG ainsi optimisée offre de remarquables performances de bruit, le facteur de bruit NF_{Min} atteint des valeurs aussi faibles que 1.2 dB à 40 GHz et 1.6 dB à 60 GHz pour un transistor ayant des f_T et f_{MAX} de 230 et 300 GHz respectivement [Chevalier06a].

Des blocs fonctionnels utilisant les transistors FSA-SEG ont par ailleurs été démontrés : [Nicolson06] démontre des blocs VCO fonctionnant au delà de 100 GHz grâce à l'emploi de transistors ayant des fréquences de 230/240 GHz ou 270/260 GHz, avec un bruit de phase inférieur à -100 dBc/GHz.

[Laskin06] démontre des blocs diviseurs de fréquence fonctionnant jusqu'à 100 GHz avec un transistor ayant une fréquence f_T de 230 GHz, ce qui en fait le candidat idéal pour la réalisation de radars anti-collisions à 77 GHz. Ces performances sont les meilleures performances obtenues avec un transistor bipolaire SiGe intégrable dans une technologie BiCMOS.

[Voiginescu06] démontre enfin la possibilité de réaliser un grand nombre de circuits hyperfréquences avec le transistor étudié au cours de ce travail, notamment pour les communications numériques jusqu'à 80 Gb/s. Le transistor bipolaire SiGe se présente donc comme un concurrent aux transistors bipolaires réalisés en InP.

Les circuits réalisés en BipX, c'est-à-dire avec une technologie bipolaire utilisant les interconnexions CMOS standards, ont démontré des performances à l'état de l'art [Laskin06, Nicolson06]. La réalisation de systèmes millimétriques complets (WLAN à 60 GHz, Radar à 77 GHz...) demande toutefois de pouvoir intégrer plus de fonctionnalités sur une même puce mais également plus de marge sur les performances. BiCMOS9MW, dernière née des filières BiCMOS de STMicroelectronics, répond à cette demande puisqu'elle intègre un transistor issu des travaux menés dans cette thèse ($f_T/f_{MAX} = 230/280$ GHz) dans un noeud CMOS 0.13 μm avec des interconnexions dédiées aux applications millimétriques offrant deux niveaux de cuivre épais et des diélectriques épais [Chevalier07].

Chapitre IV

Amélioration du produit $f_T \times BV_{CEO}$

IV.1 Présentation de l'approche

Le chapitre III sur l'optimisation conventionnelle du transistor a mis en évidence la limite d'une démarche basée sur des considérations « classiques », le produit $f_T \times BV_{CEO}$ étant difficile à améliorer. En effet, à cause des différentes contraintes d'intégration existant dans une filière BiCMOS, une augmentation de f_T s'accompagne quasi-systématiquement d'une diminution des tensions de claquage. La figure IV.1 représente l'évolution des performances des transistors bipolaires à hétérojonction SiGe dans un diagramme f_T - BV_{CEO} sur une durée de trois ans. On constate que l'augmentation des performances contribue à augmenter le produit $f_T \times BV_{CEO}$, au détriment d'une diminution de la tension de claquage.

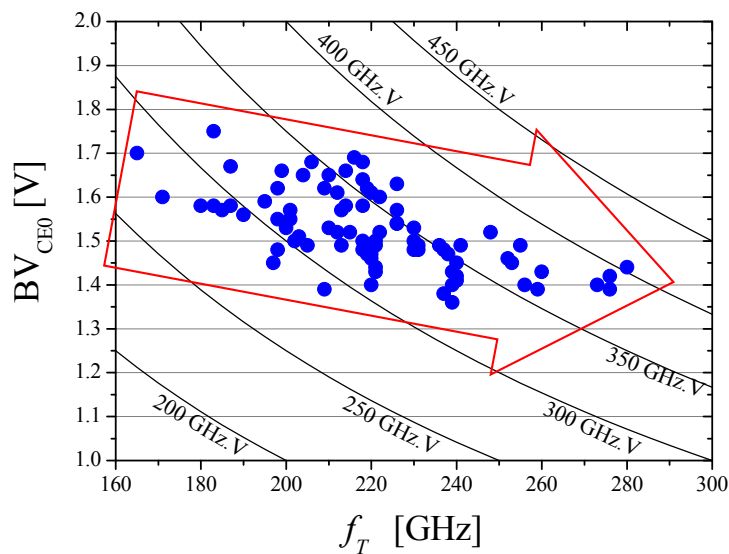


FIG. IV.1 – Evolution de BV_{CEO} en fonction de f_T lors de l'optimisation conventionnelle du TBH.

Cette diminution de BV_{CEO} est due à deux effets distincts :

- Pour atteindre des fréquences de coupure élevées, les profils de dopages doivent être de plus en plus abrupts, notamment côté collecteur. L'avalanche dans la jonction B/C intervient donc à des tensions plus basses.
- Un fort courant collecteur est nécessaire pour augmenter f_T . Le gain en courant augmente donc fortement, ce qui dégrade également la tension BV_{CEO} .

Si l'on veut améliorer le produit $f_T \times BV_{CEO}$, il faut donc trouver un moyen alternatif qui permette d'augmenter la tension de claquage sans dégrader ni le profil de dopage du collecteur,

ni le niveau de I_C . De fait, une solution existante pour améliorer BV_{CEO} est d'augmenter le courant de base. En effet, la tension de claquage BV_{CEO} est définie comme étant la tension V_{CE} à laquelle le courant de base s'inverse. Augmenter I_B permet de retarder l'apparition de l'inversion, comme indiqué dans la figure IV.2.

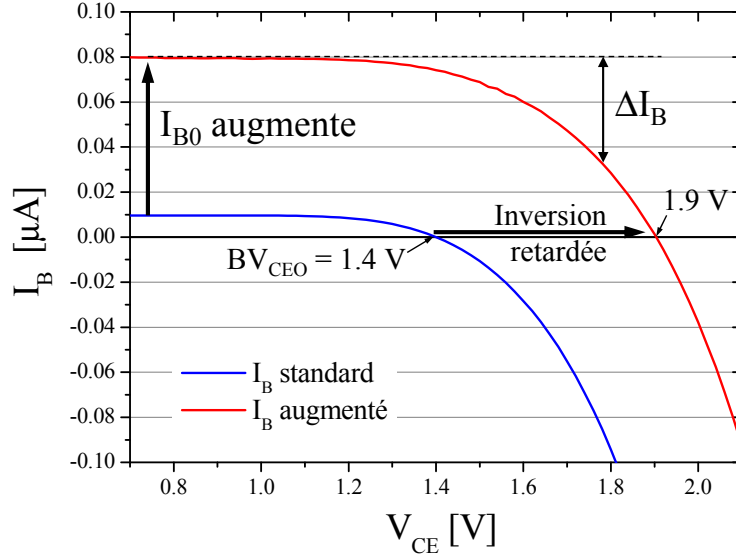


FIG. IV.2 – Variations de I_B en fonction de V_{CE} . Taille des composants : $0.15 \times 3.6 \mu\text{m}^2$. L'augmentation de I_B permet de retarder son inversion.

La variation de courant de base ΔI_B avec la tension V_{CB} due à l'avalanche dans la jonction base/collecteur, a pour expression :

$$\Delta I_B = \Delta I_C = (M - 1)I_{C0} = \left(\frac{V_{CB}}{BVC} \right)^{M_f} I_{C0} \quad (\text{IV.1})$$

ΔI_B dépend des paramètres de la jonction base/collecteur, M_f et BVC . La tension de claquage BV_{CEO} étant atteinte lorsque $\Delta I_B = I_{B0}$, et I_{B0} étant égal à I_{C0}/β , on obtient :

$$BV_{CEO} = V_{BE} + \frac{BVC}{\beta^{1/M_f}} \quad (\text{IV.2})$$

Ainsi, lorsque le profil de dopage du collecteur est fixé, M_f et BVC sont constant, et BV_{CEO} ne dépend que du gain en courant. La figure IV.3 présente l'évolution de la tension BV_{CEO} en fonction du gain pour un profil collecteur donné, calculée d'après l'équation IV.2. On note que pour atteindre des tensions de claquage de l'ordre de 2 V, le gain en courant doit être de l'ordre de 200. Des tensions de claquage plus élevées peuvent être atteintes par ce type d'approche, mais un gain trop faible est pénalisant pour l'utilisation du composant dans un circuit.

Ce chapitre vise à décrire les procédés technologiques, en rupture avec l'optimisation classique du transistor bipolaire, qui ont été mis en œuvre afin d'augmenter fortement le produit $f_T \times BV_{CEO}$ par une augmentation du courant de base uniquement. Le schéma de la figure I.13, permet d'identifier les différentes composantes du courant de base que nous pouvons modifier :

- I_{pE} : Il s'agit de la composante de diffusion des trous dans l'émetteur. Pour augmenter cette composante, il faut augmenter leur vitesse de recombinaison. En effet, lorsque le niveau de trous à l'entrée de l'émetteur p_{n0E} est maintenu constant, l'augmentation de la vitesse de recombinaison de ces trous permet d'augmenter I_{pE} . C'est l'approche retenue avec le procédé émetteur métallique exposée dans la partie IV.2 et l'émetteur SiGe détaillé dans la partie IV.3.

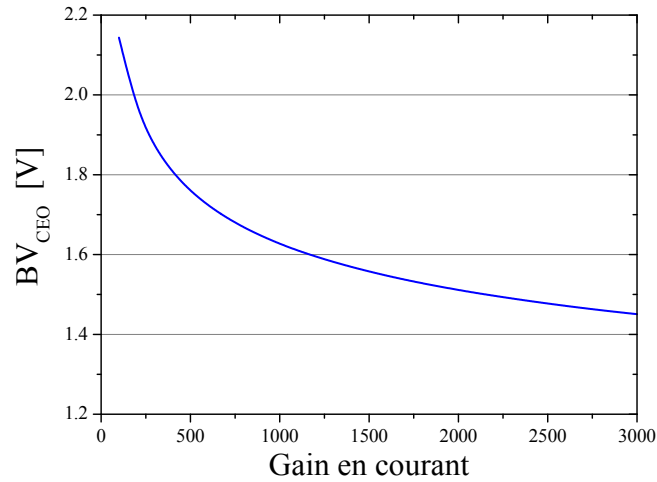


FIG. IV.3 – Estimation de la tension de claquage en fonction du gain en courant du composant pour un dopage de collecteur fixé ($V_{BE} = 0.7$ V, $M_f = 5.2$, $BVC = 3.5$ V).

- I_{rB} est la composante de recombinaison en base neutre. Généralement négligeable, ce terme peut être augmenté significativement par l'insertion de pièges dans la base neutre, à l'instar de ce qui se passe dans la base des TBH III-V grâce à l'effet Auger. Cette approche fait l'objet de la partie IV.4 de ce chapitre.

Les deux composantes I_{pE} et I_{rB} permettent toutes deux d'augmenter le courant de base de manière idéale. Nous allons donc décrire successivement les approches retenues pour augmenter chacune de ces composantes.

IV.2 Transistor bipolaire à émetteur métallique

La première approche retenue afin d'améliorer la tension de claquage BV_{CEO} des TBH est l'approche dite « émetteur métallique ». Le principe est de réaliser un émetteur très fin, ce qui permet d'accélérer la recombinaison des trous au contact métallique de l'émetteur, comme exposé dans la figure IV.4 [Chantre03].

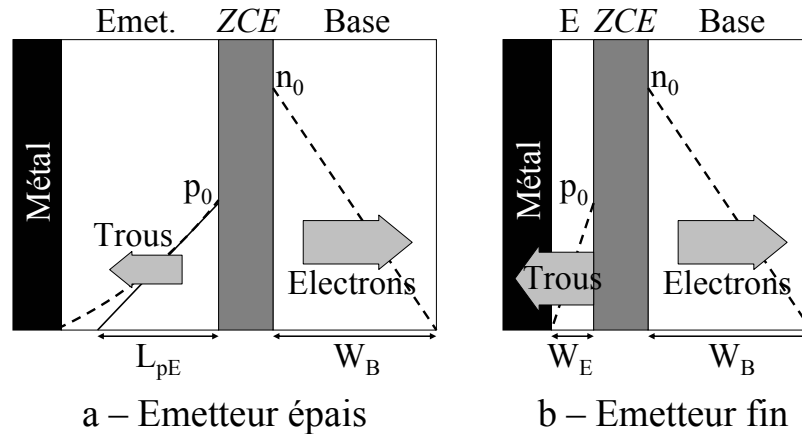


FIG. IV.4 – Schéma expliquant le principe de l'émetteur métallique.

Si l'on s'en tient à un modèle purement diffusif pour les courants du transistor bipolaire, le courant de base dépend de la longueur de recombinaison des trous dans l'émetteur. D'après [Ashburn88], cette longueur de recombinaison est égale à la longueur de diffusion des trous dans l'émetteur L_{pE} lorsque celui-ci est épais, ou à l'épaisseur d'émetteur W_E lorsque celle-ci est faible. Le but de l'émetteur métallique est de réaliser un transistor ayant un émetteur suffisamment fin pour le courant de base soit contrôlé en grande partie par l'épaisseur d'émetteur. Les trous se recombinent à l'interface polysilicium / métal, qui doit être très proche de la jonction émetteur / base. Le procédé de fabrication ainsi que les principaux résultats de ce chapitre ont été publiés dans [Barbalat06a].

IV.2.1 Réalisation technologique

IV.2.1.a Architectures possibles

Afin de réaliser un transistor à émetteur métallique, plusieurs approches existent. Des transistors à émetteur métallique ont été mentionnés pour la première fois dans [Donkers04]. Cette première approche utilise un polyémetteur épais totalement siliciuré. L'épaisseur de l'émetteur neutre est réduite en adaptant l'avancée de la siliciuration. Cette approche est assez simple à mettre en œuvre, mais elle a l'inconvénient de ne pas être compatible avec la brique CMOS 0.13 μm . Dans le cadre de cette étude, deux autres architectures permettant de réaliser des transistors bipolaires à émetteur métallique ont été mises au point :

- La première approche possible est d'utiliser un émetteur très fin, en ne modifiant que l'épaisseur de polyémetteur déposé. L'enchaînement des opérations de fabrication n'est pas modifié.
- La deuxième approche utilise le procédé PRETCH (« Poly REplacement Through Contact Hole ») décrit dans [Harrison04] pour la réalisation de transistors MOS à grille métallique. Dans cette approche, un polysilicium sacrificiel est gravé sélectivement à travers le trou de contact pour être remplacé ensuite par du métal.

L'épaisseur de polyémetteur standard est de 150 nm. Dans le cas d'un émetteur métallique, cette épaisseur est réduite à 40 voire 20 nm. La figure IV.5 présente des observations au microscope électronique à transmission de transistors bipolaires standard (a), ainsi que deux transistors à émetteur métallique (b) et (c).

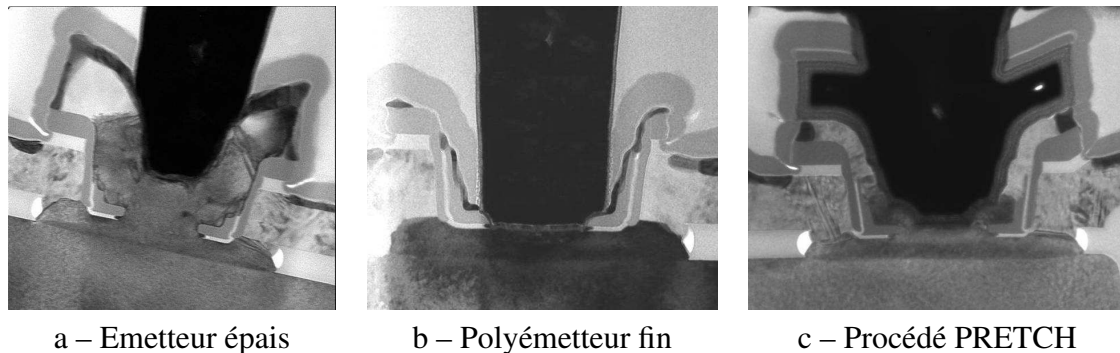


FIG. IV.5 – Comparaison d'un transistor standard avec l'architecture polyémetteur fin et l'architecture PRETCH.

L'approche utilisant un polyémetteur fin a rapidement été écartée, car ce procédé est peu contrôlable. En effet, la progression de la siliciuration dans le silicium est soumise à des variations aléatoires, il n'est donc pas possible d'obtenir à coup sûr l'épaisseur d'émetteur neutre désirée. De plus dans le cas d'émetteur très fin (20 nm) la progression du cobalt jusque dans la partie intrinsèque du transistor génère des courants non idéaux. Ainsi, nous avons retenu le procédé PRETCH, plus robuste, et offrant un excellent contrôle de l'épaisseur d'émetteur.

IV.2.1.b Détails du procédé PRETCH

Le procédé PRETCH demande un certain nombre de modifications dans l'enchaînement des étapes de fabrication. La réalisation de ce type de composant débute exactement comme pour un composant standard, jusqu'au module émetteur. À la place du polyémetteur épais déposé par RTCVD, on dépose un empilement constitué successivement :

- d'une couche de polyémetteur fin (20 à 40 nm), dopé In-situ et déposé par RTCVD, comme dans le cas d'un émetteur épais ;
- d'un oxyde interfacial (50 Å), qui sert à isoler le polyémetteur du polysilicium sacrificiel ;
- d'une couche de polysilicium amorphe sacrificiel (100 nm) ;
- d'une couche de nitrure destinée à bloquer la siliciuration au sommet de l'émetteur.

Le transistor bipolaire à émetteur métallique a la même topologie qu'un transistor standard à émetteur épais. Après gravure de cet empilement spécifique selon le masque polyémetteur standard, la fabrication se poursuit par le module siliciure, le dépôt du nitrure d'encapsulation et du premier niveau de diélectrique. La gravure des trous de contact permet d'ouvrir le nitrure situé au dessus de l'émetteur, laissant apparaître le polysilicium sacrificiel, comme indiqué dans la figure IV.6–a, en préservant le siliciure de la base et du collecteur.

Afin d'enlever le polyémetteur sacrificiel, on procède à une gravure isotrope de ce matériau par un plasma de SF_6 . La cavité émetteur est ainsi entièrement vidée, comme indiqué dans la figure IV.6–b. La gravure isotrope du polysilicium est sélective par rapport au siliciure, de telle sorte que les contacts de base et de collecteur ne sont pas modifiés, et par rapport à l'oxyde interfacial, préservant le polyémetteur situé au dessous.

Enfin, afin de réaliser l'émetteur métallique proprement dit, la couche d'oxyde interfacial est éliminée par un nettoyage non oxydant en milieu humide, similaire à un nettoyage réalisé avant une épitaxie. La surface du polyémetteur fin est donc prête pour le dépôt métallique des

contacts, constitué d'une couche de titane (Ti) de 10 nm déposée par pulvérisation cathodique, d'une couche de nitrure de titane (TiN) déposée par CVD et d'épaisseur 20 nm et de tungstène déposé par CVD afin de remplir totalement les trous de contact et la cavité émetteur, comme indiqué dans la figure IV.6–c.

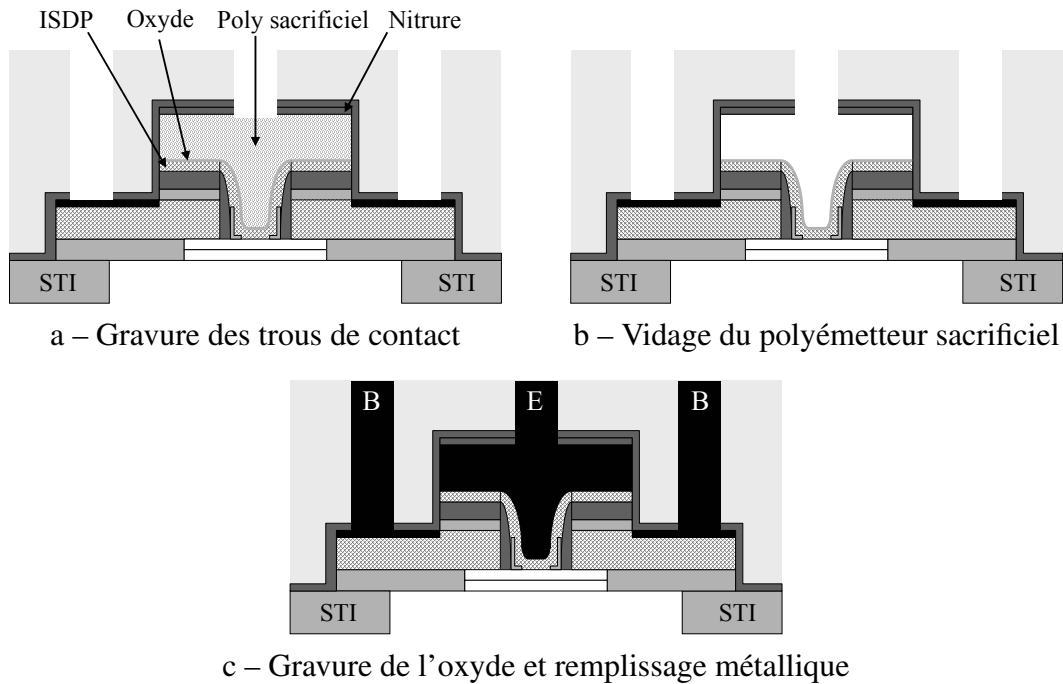


FIG. IV.6 – Procédé de fabrication PRETCH.

La figure IV.7 présente une vue en fausses couleurs d'un transistor à émetteur métallique réalisé avec le procédé PRETCH.

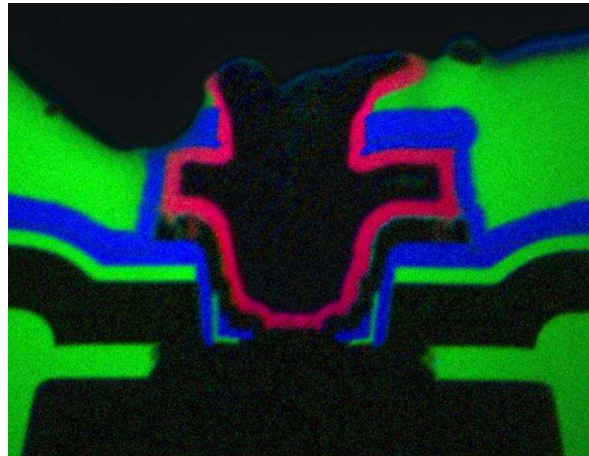


FIG. IV.7 – Photo TEM en fausses couleurs présentant un transistor à émetteur métallique réalisé avec le procédé PRETCH. Vert : Oxygène (Oxyde), Rouge : Titane (Ti, TiN), Bleu : Azote (Nitrure), Noir : Silicium et Tungstène.

IV.2.2 Résultats obtenus avec un émetteur fortement dopé

Les premiers essais pour intégrer un émetteur métallique utilisent un polyémetteur fortement dopé, c'est à dire avec le débit d'arsine maximal de 280 sccm. Ces premiers essais ont permis

de mettre au point le procédé de fabrication spécifique du PRETCH.

IV.2.2.a Résultats électriques

La figure IV.8 représente les tracés de Gummel et de gain en courant de 4 transistors de dimension $0.15 \times 3.6 \mu\text{m}^2$, avec des épaisseurs d'émetteur variant de 150 nm (épaisseur standard) à 40 nm. La base présente un profil 20-30 % de Ge sur une épaisseur de 20 nm, l'épaisseur du Si-cap est fixée à 18 nm. On note que le procédé émetteur métallique a peu d'impact sur le courant collecteur, et que le courant de base obtenu est idéal, prouvant ainsi la validité de l'approche. Le gain est diminué d'un facteur 2 (de 2350 à 1020) comme indiqué dans le tableau IV.1, grâce à l'augmentation de I_B . Cependant, cette diminution du gain a généré une augmentation minime de BV_{CEO} , qui augmente de 1.44 à 1.54 V.

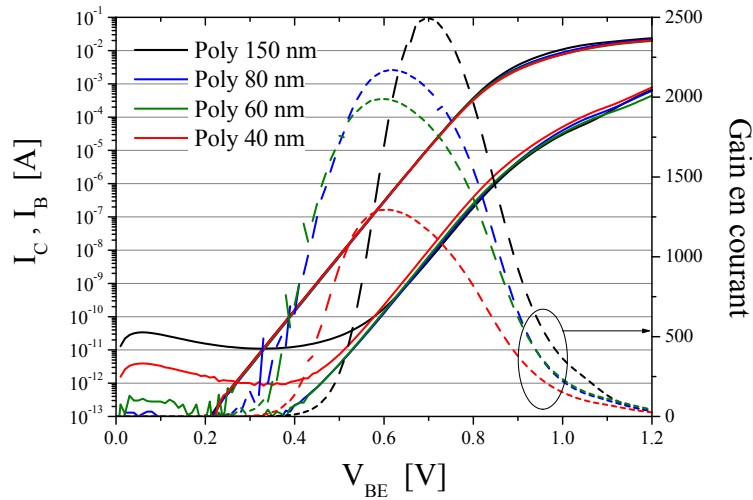


FIG. IV.8 – Courbes de Gummel et de gain de transistors bipolaires à émetteur métallique de dimension $0.15 \times 3.6 \mu\text{m}^2$, en fonction de l'épaisseur de l'émetteur.

Épaisseur d'émetteur	150 nm	80 nm	60 nm	40 nm
I_C à 0.75 V [μA]	72	68	70	66
I_B à 0.75 V [nA]	31	37	43	65
β à 0.75 V	2350	1850	1640	1020
BV_{CEO} [V]	1.44	1.46	1.48	1.54
f_T [GHz]	252	239	230	205
f_{MAX} [GHz]	274	271	262	235
R_E [Ω]	9.1	11.1	12.5	13.8
$f_T \times BV_{CEO}$ [GHz·V]	362	350	341	316

TAB. IV.1 – Principaux paramètres électriques en fonction de l'épaisseur du polyémetteur ($0.15 \times 3.6 \mu\text{m}^2$).

L'amplitude de variation du courant de base avec l'épaisseur d'émetteur est très faible, ce qui fait que l'augmentation de BV_{CEO} est minime. Concernant les performances dynamiques du composant, on note d'après le tableau IV.1 une augmentation importante de la résistance d'émetteur, qui a provoqué une chute de la fréquence f_T des transistors à émetteur métallique, lorsque l'épaisseur d'émetteur neutre est réduite. Les évolutions de f_T en fonction de I_C pour différentes épaisseurs d'émetteur sont reportées dans la figure IV.9.

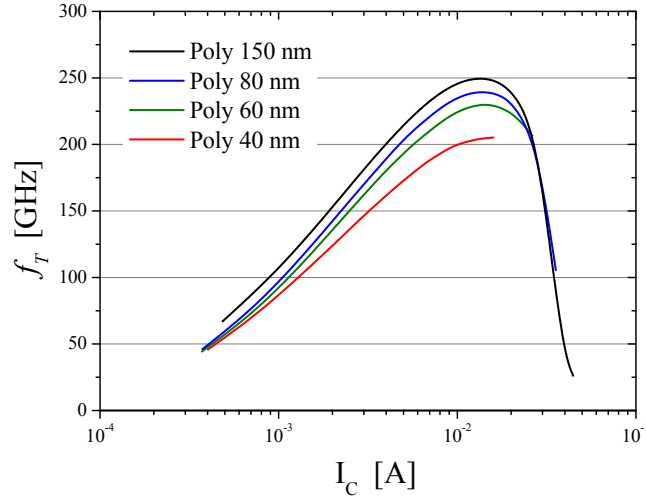


FIG. IV.9 – Évolution de f_T en fonction de I_C pour différentes épaisseurs d'émetteur. Dimension des composants : $0.15 \times 3.6 \mu\text{m}^2$.

IV.2.2.b Analyse SIMS

Afin de comprendre l'augmentation de R_E avec la réduction de l'épaisseur de polyémetteur, des analyses SIMS ont été réalisées. La figure IV.10 compare les profils de dopage de deux composants ayant des épaisseurs d'émetteur de 80 et 40 nm.

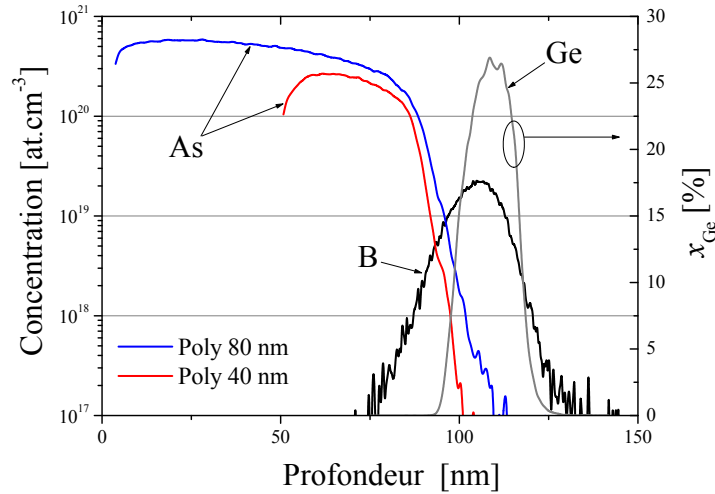


FIG. IV.10 – Profils de dopages obtenus dans le cas d'émetteurs de 40 et 80 nm d'épaisseur.

On constate d'après l'analyse SIMS que le profil de dopants arsenic n'est pas infiniment abrupt, la concentration augmente lorsque l'épaisseur de polyémetteur déposée devient plus importante. En conséquence, le réservoir de dopant, composé de toute l'épaisseur de polyémetteur, n'augmente pas proportionnellement avec l'épaisseur déposée. On constate que pour des polyémetteur de 40 nm et 80 nm d'épaisseur, les doses de dopant mesurées sur les profils SIMS sont respectivement de $7.7 \cdot 10^{14}$ et $3.8 \cdot 10^{15}$ $\text{at}\cdot\text{cm}^{-3}$. Il existe donc un facteur 5 entre les deux doses, à comparer au facteur 2 entre les épaisseurs déposées. En conséquence de cette réduction de dopage, la zone de diffusion de l'arsenic vers le Si-cap est fortement réduite. La distance arsenic-bore augmentant, la composante de la résistance d'émetteur due au Si-cap augmente fortement, comme expliqué dans la figure IV.11 : Dans le cas d'un polyémetteur épais

(a), l'épaisseur du Si-cap est adaptée à la diffusion de l'arsenic. Lorsque l'épaisseur et donc le dopage de l'émetteur sont réduits (b), la composante R_{Eb} augmente fortement.

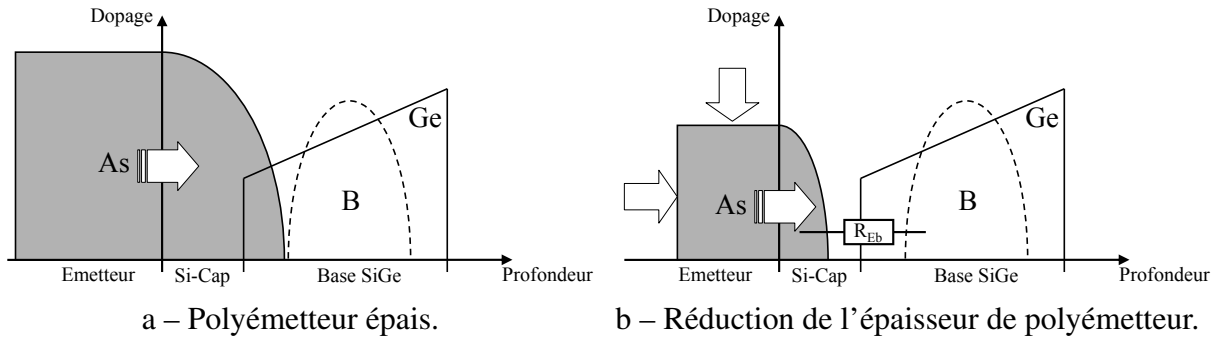


FIG. IV.11 – Effet de l'épaisseur et du dopage du polyémetteur sur le réglage de la jonction E/B.

Cette augmentation de R_E est à l'origine de la chute des performances dynamiques observées lorsqu'on réduit l'épaisseur de l'émetteur. La base étant moins pincée par les dopants de l'émetteur, l'épaisseur de la base effective augmente, ce qui entraîne également une diminution de f_T par l'augmentation du temps de transit de base. En conclusion, pour prendre en compte la diminution de la diffusion des dopants, il est nécessaire d'adapter l'épaisseur du Si-cap pour réaliser des émetteur métalliques.

IV.2.2.c Mesures d'effet Hall

Afin de comprendre la faible amplitude de variation de I_B avec l'épaisseur de l'émetteur, des mesures d'effet Hall ont été réalisées sur des échantillons de polysilicium dopés In-situ, identiques aux émetteur fortement dopés de nos TBH. On note d'après la figure IV.12 que la concentration effective des porteurs ne varie pas avec la température : Il n'y a pas de gel des porteurs, même jusqu'à une température aussi basse que 20 K. Ce comportement est typique d'un matériau dégénéré.

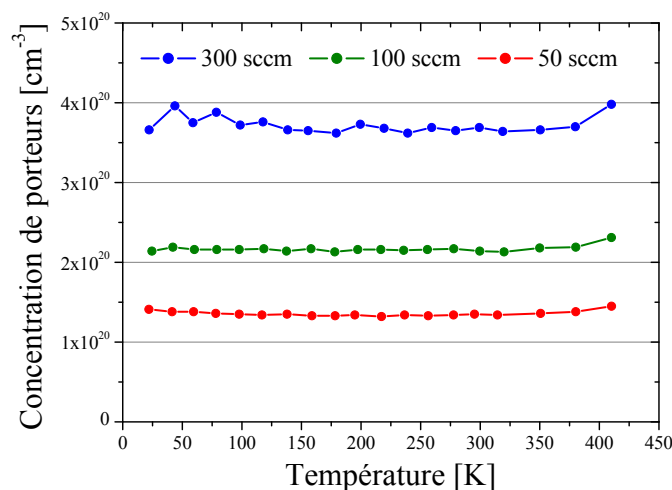


FIG. IV.12 – Évolution de la concentration de porteurs en fonction de la température pour plusieurs débit d'arsine.

Le matériau composant l'émetteur de nos transistors étant dégénéré, la longueur de diffusion des trous dans ce matériau doit être très courte. Il est donc normal que la dépendance du courant

de base avec l'épaisseur d'émetteur soit très faible, à moins que l'émetteur soit extrêmement fin, ce qui pose le problème de la réalisation technologique.

IV.2.2.d Interprétation – Suite de l'étude

L'effet très faible de l'épaisseur du polyémetteur sur le courant de base peut être interprété simplement sous l'hypothèse d'un transport des trous uniquement diffusif. Le dopage de l'émetteur étant très fort, le niveau d'injection p_0 des trous dans l'émetteur est faible. De plus, la longueur de diffusion étant très courte à cause des forts dopages, l'épaisseur de l'émetteur n'a qu'une faible influence sur le gradient de trous, comme indiqué dans la figure IV.13.

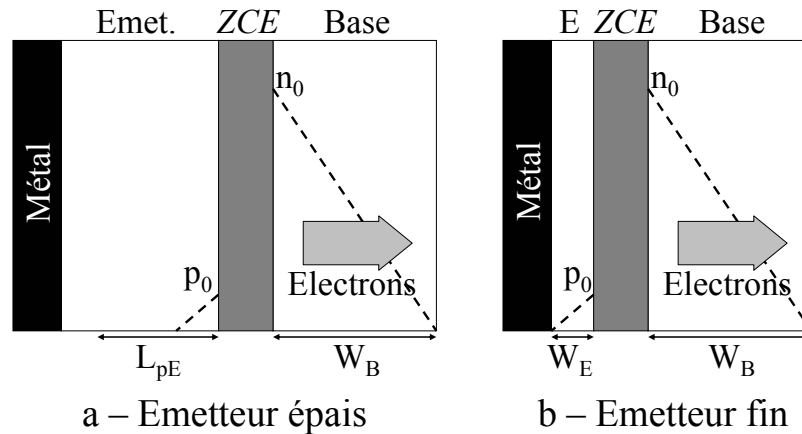


FIG. IV.13 – Interprétation des résultats dans le cas d'émetteurs fortement dopés. Lorsque l'émetteur est épais (a) ou fin (b), le gradient de trous varie peu.

Un meilleur contrôle du gradient de trous par l'épaisseur du polyémetteur est attendu avec un émetteur faiblement dopé. En effet, l'amélioration attendue est double :

- Un dopage réduit permet d'obtenir un niveau d'injection p_0 plus important, sous l'hypothèse d'un transport uniquement diffusif ;
- Le dopage étant plus faible, la longueur de diffusion des trous L_{pE} augmente et devient supérieure à la largeur d'émetteur neutre W_E .

Ces deux effets conjugués assurent que le gradient de trous à l'entrée de l'émetteur est bien contrôlé par l'épaisseur d'émetteur uniquement, comme indiqué par la figure IV.14. Plus le niveau d'injection des trous est fort, plus l'effet émetteur métallique sera important.

IV.2.3 Résultats obtenus avec un émetteur faiblement dopé

L'étude précédente a démontré la nécessité d'utiliser des émetteurs faiblement dopés pour obtenir un bon contrôle du courant de base par l'épaisseur d'émetteur. Différentes configurations d'émetteur ont été réalisées à cet effet. Le débit d'arsine a été réduit de 280 sccm à 90 sccm, réduisant le niveau de dopage d'un facteur 3. Une croissance de l'émetteur dans des conditions purement épitaxiales a également permis d'obtenir des dopages 15 fois plus faibles que le dopage standard.

Cette réduction de dopage a pour effet néfaste de diminuer encore plus la diffusion de l'arsenic vers la base. En conséquence, le Si-cap doit être adapté au niveau de dopage de l'émetteur. De 18 nm en standard, le Si-cap a été réduit à 4 nm pour les émetteurs les plus fins et les plus faiblement dopés. La base retenue pour la suite de l'étude a un profil 20-30 % de Ge, d'épaisseur 20 nm.

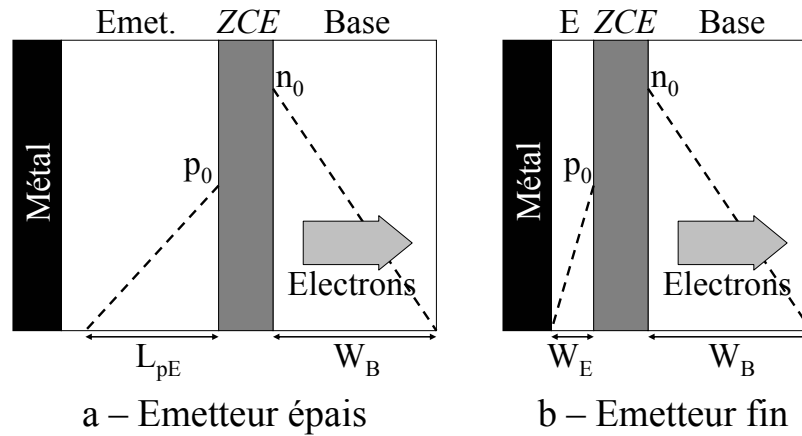


FIG. IV.14 – Effet attendu en réduisant le dopage du polyémetteur. Lorsque l'émetteur est faiblement dopé, le gradient de trous est très différent entre un émetteur épais (a) et un émetteur fortement dopé (b).

IV.2.3.a Résultats électriques

La figure IV.15 présente les courbes de Gummel et les gains en courant de transistors à émetteurs métalliques faiblement dopés. L'épaisseur du Si-cap est fixée à 6 nm. Grâce à la réduction du dopage, on observe une importante variation du courant de base avec l'épaisseur d'émetteur, validant ainsi l'interprétation des résultats de la partie précédente.

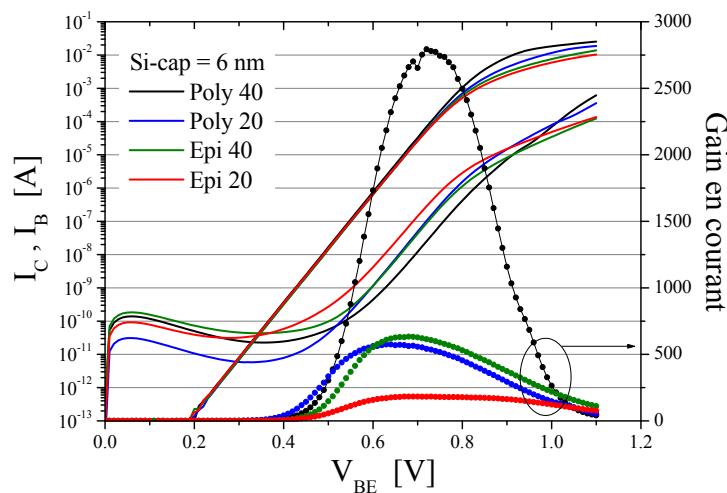


FIG. IV.15 – Courbes de Gummel et de gain en courant de TBH à émetteur métallique ayant un dopage d'émetteur réduit. Épaisseur du Si-cap = 6 nm ; dimension des composants : $0.15 \times 3.6 \mu\text{m}^2$.

Le gain en courant est réduit d'un facteur 10 grâce à l'augmentation du courant de base. Le tableau IV.2 résume l'évolution des caractéristiques statiques et dynamiques de ces composants, en fonction des variantes technologiques appliquées sur le Si-cap et sur l'émetteur. La diminution conséquente du gain a permis d'améliorer significativement la tension BV_{CEO} , qui augmente de 1.4 V à environ 2 V.

Concernant les résultats dynamiques, on observe une diminution significative des fréquences de transition lorsque l'épaisseur d'émetteur est réduite, ce qui est dû à l'augmentation des résistances séries décrite précédemment. f_T et f_{MAX} passent de 256 / 284 GHz dans le cas d'un

polyémetteur de 40 nm et Si-cap de 12 nm, à 190 / 138 GHz pour un émetteur épi de 20 nm avec un Si-cap de 6 nm. L'augmentation de la tension BV_{CEO} lorsque l'épaisseur d'émetteur diminue permet de maintenir un produit $f_T \times BV_{CEO}$ élevé (412 GHz·V dans le cas du polyémetteur de 40 nm avec Si-cap de 12 nm, 410 GHz·V pour le polyémetteur de 20 nm avec des Si-cap de 8 et 6 nm).

Type d'émetteur	Poly	Poly	Poly	Poly	Epi	Epi
Epaisseur [nm]	40	40	20	20	40	20
Dopage	90 sccm	90 sccm	90 sccm	90 sccm	Std/15	Std/15
Si-Cap [nm]	10	6	8	6	6	6
I_C à 0.75 V [μA]	136	206	140	159	149	136
I_B à 0.75 V [μA]	151	75	294	307	253	753
β à 0.75 V [μA]	900	2740	475	520	590	180
BV_{CEO} [μA]	1.61	1.59	1.75	1.74	1.69	1.92
R_E [Ω]	9.8	8.1	14.8	14.2	22.9	26.4
R_B [Ω]	28	68	33	231	184	138
f_T [GHz]	256	247	234	236	198	190
f_{MAX} [GHz]	284	195	247	179	158	138
$f_T \times BV_{CEO}$ [GHz·V]	412	392	410	410	335	363

TAB. IV.2 – Principaux paramètres électriques de transistors à émetteur métallique ayant des dopages d'émetteur réduits.

IV.2.3.b Problème d'augmentation des résistances séries

Si l'effet émetteur métallique est bien obtenu avec la réduction du dopage de l'émetteur, il subsiste toujours le problème de l'augmentation des résistances séries, notamment la résistance de base, qui devient très importante lorsque le Si-cap devient très fin. L'extraction de R_E reposant sur le fait que $R_B \cdot I_B$ est négligeable devant $R_E \cdot I_E$ (voir chapitre III.2.1), une forte résistance de base perturbe la valeur obtenue de la résistance d'émetteur.

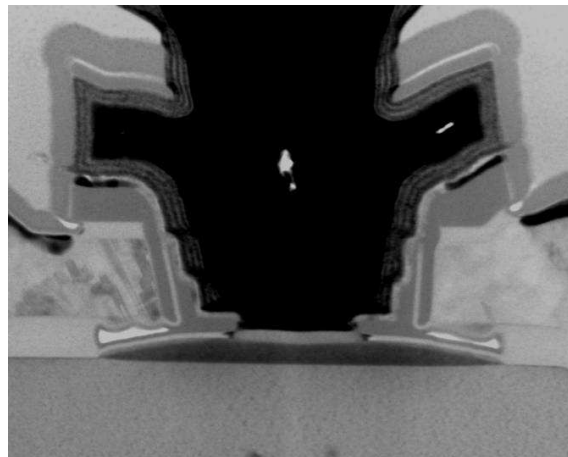


FIG. IV.16 – Observation TEM illustrant la perte du lien entre base intrinsèque et extrinsèque lorsque le Si-cap devient trop fin.

Cette augmentation de R_B a pu être identifiée grâce à une observation au microscope électronique à transmission. La figure IV.16 présente un transistor à émetteur métallique ayant un

Si-cap de 2 nm d'épaisseur. On note que la couche de base déposée par épitaxie ne remplit pas totalement la cavité qui lui est réservée dans l'oxyde piédestal. En conséquence, les couches diélectriques composant les espaceurs internes se déposent entre la base intrinsèque et la base extrinsèque, provoquant la perte de lien de base et donc l'augmentation de R_B . La réduction de l'épaisseur de l'oxyde piédestal doit donc permettre de corriger cet effet.

IV.2.4 Perspectives

L'étude a montré la possibilité d'améliorer I_B et BV_{CEO} en augmentant la vitesse de recombinaison des trous sur le contact métallique d'émetteur. Cependant, une diminution des performances dynamiques a systématiquement été observée, causée par la dégradation des résistances séries. L'approche émetteur métallique implique également systématiquement une dégradation de l'efficacité d'injection de la jonction E/B. La réduction du dopage d'émetteur notamment n'apparaît pas comme un moyen judicieux d'améliorer les performances du composant. En ce sens, l'émetteur SiGe offre la possibilité de contrôler le courant de base avec un polyémetteur épais, il n'y a donc pas de diminution du réservoir de dopants.

IV.3 Insertion de SiGe dans l'émetteur

Un autre procédé pour augmenter le courant I_{pE} est d'insérer du germanium dans l'émetteur. Le courant de trous peut ainsi être augmenté par deux effets différents :

- L'augmentation de la concentration intrinsèque des porteurs n_i dans l'émetteur. L'énergie de bande interdite du SiGe étant plus faible que celle du Si, $n_{i,SiGe} > n_{i,Si}$ donc I_B augmente par rapport à un émetteur tout silicium. Cette approche a été décrite dans [Martinet02] et [Kunz02].
- La présence de germanium a pour conséquence d'augmenter la vitesse de recombinaison par effet Auger [Ning03]. Ainsi, les trous se recombinant plus rapidement dans l'émetteur, le gradient de trous devient plus important.

Nous allons présenter dans cette partie un procédé utilisant la deuxième approche. De plus, si la quantité de Ge augmente trop fortement dans la structure, nous pouvons avoir des dislocations, auquel cas les composants ne sont plus fonctionnels. Lorsqu'un plan de Ge est inséré proche de la jonction émetteur/base, il joue le rôle de surface recombinante. Les trous injectés de la base vers l'émetteur se recombinent donc plus rapidement, et le courant de base est contrôlé par la profondeur du pic de Ge dans le cap comme indiqué dans la figure IV.17.

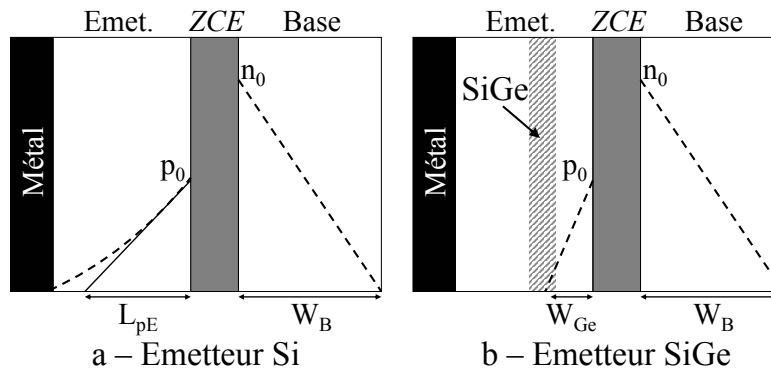


FIG. IV.17 – Schéma expliquant le principe de l'émetteur SiGe.

IV.3.1 Mise en œuvre

Afin de créer un plan recombinant dans l'émetteur, un fort pic de Ge est inséré dans le Si-cap au moment de la croissance épitaxiale de la base (voir figure IV.18). La largeur du pic de Ge est 5 nm au moment du dépôt, et le pourcentage de Ge visé est 20 ou 30 %. Le plan de Ge est situé 5 nm sous la surface du silicium.

L'insertion d'un plan de Ge dans l'émetteur a pour avantage de ne pas modifier l'enchaînement des étapes de fabrication. L'épaisseur du Si-cap doit légèrement être augmentée, car le Ge favorise la diffusion de l'arsenic. Le principe de fonctionnement est identique à celui de l'émetteur métallique (recombinaison rapide des trous sur une surface proche de la jonction E/B), et présente l'avantage de ne pas diminuer le réservoir de dopants de l'émetteur. Ainsi, on peut s'attendre à ce qu'il n'y ait pas de dégradation de la résistance d'émetteur.

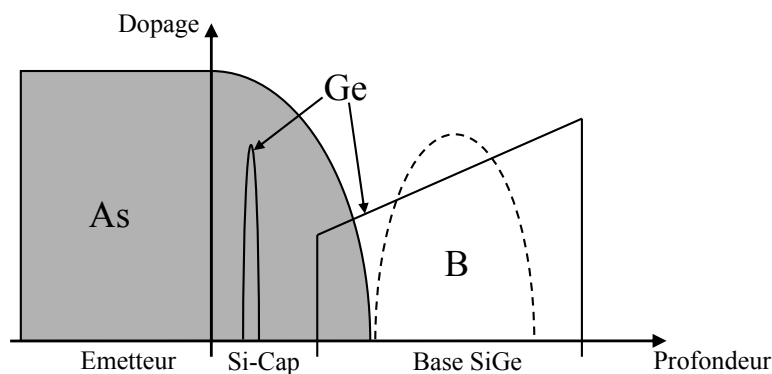


FIG. IV.18 – Profil de base réalisé. Un pic de Germanium de 5 nm de large est inséré dans le Si-cap.

IV.3.2 Résultats électriques

IV.3.2.a Résultats statiques

La figure IV.19 présente les courbes de Gummel et de gain en courant de trois transistors bipolaires de dimension $0.15 \times 3.6 \mu\text{m}^2$, présentant un Si-cap sans germanium, un pic de 20 % et un pic de 30 % de Ge. On note que la présence du germanium a un fort impact sur le courant de base, et donc sur le gain en courant du composant. Le facteur d'idéalité du courant de base augmente, ce qui peut être dû soit à des défauts présents dans la ZCE émetteur/base, soit à un phénomène de saturation des pièges identique à celui démontré dans la partie consacrée à la recombinaison en base neutre. Une caractérisation plus poussée de ces échantillons obtenus en fin de thèse est nécessaire pour identifier l'origine physique de cette non idéalité.

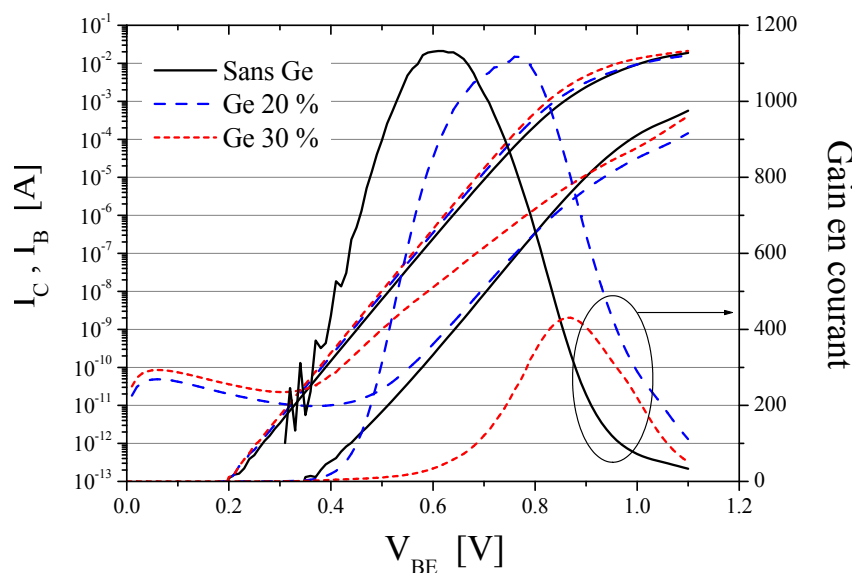


FIG. IV.19 – Courbes de Gummel et de gain en courant de transistors présentant un pic de Ge dans le Si-cap (dimensions $0.15 \times 3.6 \mu\text{m}^2$).

Le courant collecteur a également augmenté lorsqu'on ajoute du germanium. Cet effet est dû à une diffusion accrue des atomes d'arsenic venant de l'émetteur, comme le prouve l'augmentation de la capacité E/B donnée dans le tableau IV.3. Cette diffusion accrue de l'arsenic a également tendance à diminuer la résistance d'émetteur, et la résistance de base n'est pas mo-

difiée par l'insertion de Ge. On note finalement une augmentation significative de la tension BV_{CEO} , qui augmente de 1.5 V à quasiment 2 V.

Profil de Ge	Std (pas de Ge)	20 %	30 %
I_C à 0.75 V [μA]	47	76	101
I_B à 0.75 V [nA]	55	69	469
β à 0.75 V	876	1105	215
BV_{CEO} [V]	1.54	1.61	1.95
R_E [Ω]	9.9	7.2	7.0
R_B [Ω]	25.3	21.8	24.1
C_{BE} [pF]	9.78	9.8	10.11
f_T [GHz]	244	241	198
f_{MAX} [GHz]	271	273	271
$f_T \times BV_{CEO}$ [GHz·V]	376	388	386

TAB. IV.3 – Paramètres électriques statiques et dynamiques obtenus en fonction du profil de Germanium dans le Si-cap ($0.15 \times 3.6 \mu m^2$).

IV.3.2.b Performances dynamiques

Le tableau IV.3 présente les performances dynamiques des composants présentant des pics de Ge dans le Si-cap. On note que la fréquence f_T diminue fortement avec l'insertion de Ge (de 244 GHz pour le composant standard à 198 GHz dans le cas d'un pic de 30% de Ge), bien que le courant collecteur augmente et que la résistance d'émetteur diminue. La fréquence f_{MAX} reste inchangée (271 GHz), on note en effet que la résistance de base R_B reste quasiment constante lorsque la quantité de germanium de l'émetteur augmente (autour de 25 Ω).

L'augmentation de la capacité C_{BE} avec la dose de Ge (de 9.8 à 10.1 pF) est une des causes possibles de diminution de f_T . Une autre hypothèse, plus probable, est que cette chute de performances est causée par une efficacité d'injection dégradée à la jonction émetteur/base, ce qui a pour effet d'augmenter le temps de transit dans l'émetteur τ_E . Cependant, on note une légère amélioration du produit $f_T \times BV_{CEO}$ grâce à l'augmentation de la tension de claquage : De 376 GHz·V dans le cas standard, on atteint 386 GHz·V dans le cas du pic à 30 % de Ge.

IV.3.3 Conclusion sur l'insertion de Ge dans l'émetteur

L'insertion de Ge dans le Si-cap permet d'augmenter sensiblement la tension de claquage BV_{CEO} , au détriment des performances dynamiques du composant. Les effets observés sont les mêmes que dans le cas du transistor à émetteur métallique : Une forte dégradation de l'efficacité d'injection émetteur/base ayant pour conséquence d'augmenter fortement le temps de transit émetteur, ce qui dégrade f_T . Les causes de cette diminution de performances doivent encore être étudiées. Une hypothèse émise est que le Ge présent dans la ZCE E/B dégrade l'efficacité d'injection de la jonction, ce qui augmente le temps de transit d'émetteur. La présence de Ge dans la zone de déplétion émetteur/base peut aussi être à l'origine de la non-idéalité du courant de base observée sur les tracés de Gummel.

Afin de valider ces hypothèses, des études sont en cours, visant à intégrer le pic de Ge uniquement dans la zone neutre de l'émetteur, afin de ne pas modifier le réglage de la jonction E/B. Le germanium doit donc être introduit dans le polyémetteur. Ceci devrait également permettre d'obtenir des fréquences f_T supérieures à celles démontrées dans ce travail. Cependant la difficulté de réalisation de telles structures est accrue : il faut en effet s'assurer que le pic de Ge est

introduit hors de la zone de déplétion de l'émetteur, mais il doit être suffisamment proche de la jonction E/B pour permettre la recombinaison rapide des trous. Le principal facteur limitant est la largeur minimale du pic de Ge qu'il est possible de réaliser dans une couche de polysilicium fortement dopée.

Devant les difficultés de mise au point des procédés émetteur métallique et émetteur SiGe, une troisième architecture visant à augmenter le produit $f_T \times BV_{CEO}$ a été introduite. Cette dernière solution permet de ne pas modifier les réglages d'émetteur, et consiste à augmenter la recombinaison des porteurs dans la base neutre du TBH.

IV.4 Recombinaison en base neutre

Les deux parties précédentes ont montré que l'augmentation de la composante I_{pE} est un moyen efficace d'améliorer la tension de claquage BV_{CEO} , mais les performances dynamiques des transistors sont systématiquement dégradées par l'augmentation du temps de transit émetteur. Le but de cette partie est de présenter les résultats obtenus en augmentant la composante de recombinaison en base neutre I_{rB} . Afin de forcer la recombinaison des trous et des électrons dans la partie non déplétée de la base, il faut insérer une certaine quantité de pièges, en veillant à ce que ceux-ci ne soient pas présents dans les zones de charge d'espace, ce qui créerait des courants de base non idéaux.

IV.4.1 Mise en œuvre

Le moyen technologique retenu pour augmenter I_{rB} est d'insérer une forte quantité de carbone au centre de la base neutre, au moment de l'épitaxie du SiGeC. En effet, lorsqu'il est inséré en site substitutionnel, le carbone bloque la diffusion du bore, ce qui permet l'obtention de bases fines et fortement dopées. A des concentrations plus fortes, le carbone est incorporé en site interstitiel, et crée alors des centres recombinant, qui permettent d'augmenter le courant de base. L'idée de contrôler I_{B} par l'insertion de carbone a déjà été proposée dans [Saitoh04] notamment, nous présentons ici une étude plus approfondie des mécanismes de recombinaison dans la base neutre [Barbalat06c, Barbalat06b].

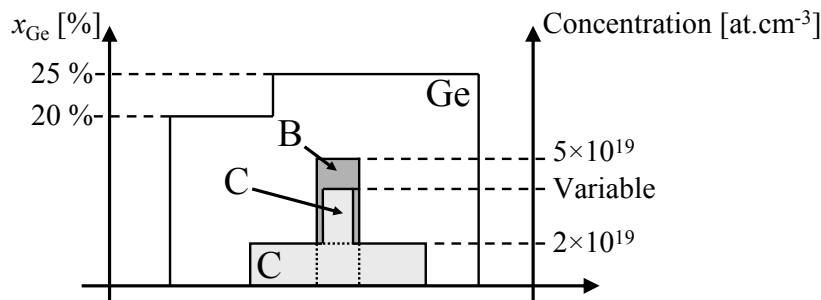


FIG. IV.20 – Profil de base déposé pour augmenter la recombinaison en base neutre.

Nous avons fait varier le taux de recombinaison dans la base neutre, en ajoutant au moment de la croissance de la base SiGeC un pic de carbone additionnel centré sur le pic de bore. Différentes doses de carbone ont été obtenues en changeant la largeur du pic, ainsi que le débit de méthylsilane CH_3SiH_3 , gaz précurseur du carbone. La figure IV.20 présente le profil de base ainsi réalisé.

Le profil de Germanium est un profil 20–25 %, qui permet d'obtenir un fort courant collecteur, afin de mettre en avant l'effet de réduction du gain par la recombinaison en base neutre. Les variations réalisées sur le profil de carbone sont résumées dans le tableau IV.4. Trois profils de carbone nommés P1, P2 et P3, correspondant à une dose croissante, sont comparés au profil standard. L'analyse SIMS (figures IV.21 et IV.22) menée sur les plaques de silicium en fin de fabrication a permis de calculer la concentration maximale ainsi que la dose par unité de surface, pour chaque profil de carbone réalisé. On note sur la figure IV.22 que la dose de bore augmente légèrement lorsque la dose de carbone est plus forte. Cet effet est expliqué par une vitesse d'incorporation du bore différente lorsque les conditions changent dans la chambre d'épitaxie.

Profil	Std	P1	P2	P3
Débit de CH_3SiH_3 [sccm]	-	7.5	15	15
Largeur du pic de C [nm]	-	4	2	4
Concentration max. [at.cm^{-3}]	$4.0 \cdot 10^{18}$	$1.7 \cdot 10^{19}$	$4.6 \cdot 10^{19}$	$1.6 \cdot 10^{20}$
Dose de C [at.cm^{-2}]	$1.3 \cdot 10^{13}$	$2.5 \cdot 10^{13}$	$5.8 \cdot 10^{13}$	$1.4 \cdot 10^{14}$

TAB. IV.4 – Résumé des caractéristiques des différents profils de carbone réalisés afin d'augmenter la recombinaison en base neutre.

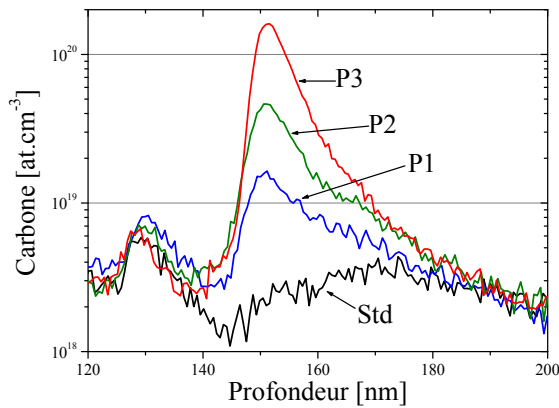


FIG. IV.21 – Profils SIMS du carbone obtenus après recuit.

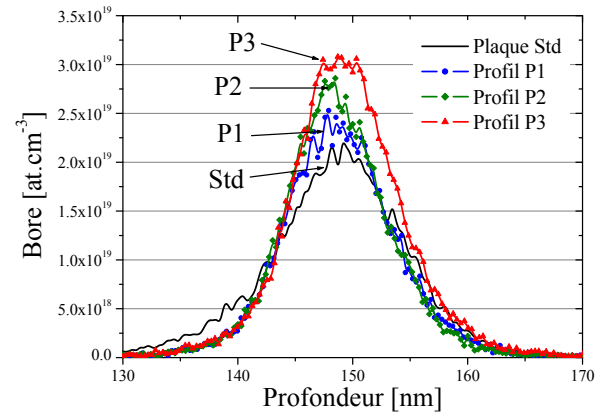


FIG. IV.22 – Profils de bore obtenus après recuit pour les trois profils P1, P2 et P3, comparés à la plaque standard.

IV.4.2 Résultats électriques

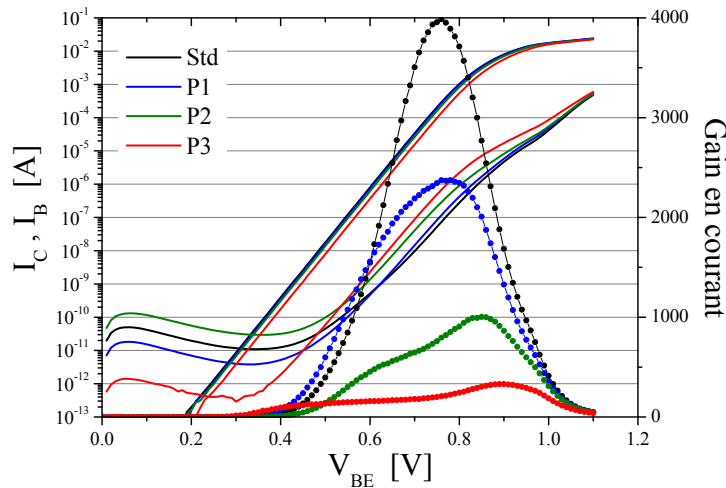


FIG. IV.23 – Courbes de Gummel et de gain en courant de TBH avec recombinaison en base neutre. Dimension des composants : $0.15 \times 3.6 \mu\text{m}^2$.

La figure IV.23 présente l'évolution des courants I_B et I_C ainsi que du gain en courant pour des TBH de dimensions $0.15 \times 3.6 \mu\text{m}^2$ dotés des profils de carbone présentés dans le tableau IV.4. Le courant collecteur décroît d'un facteur 2, ce qui est dû en partie à la variation de dose de bore entre chaque profil. On note une forte augmentation du courant de base avec la dose de carbone, tout en conservant l'idéalité. Le gain en courant décroît d'environ 4000 à 200, ce qui a

une influence notable sur la tension BV_{CEO} , qui passe de 1.4 V à 1.9 V. Le tableau IV.5 résume les principaux paramètres statiques et dynamiques obtenus.

Profil de carbone	Std	P1	P2	P3
I_C à 0.75 V [μA]	206	195	160	98
I_B à 0.75 V [nA]	52	83	211	487
β à 0.75 V	3960	2340	760	200
BV_{CEO} [V]	1.4	1.45	1.66	1.93
R_{bp} [$k\Omega/\square$]	1.7	1.5	1.5	1.5
f_T [GHz]	240	238	234	219
f_{MAX} [GHz]	294	297	294	282
$f_T \times BV_{CEO}$ [GHz·V]	336	345	388	423

TAB. IV.5 – Principaux paramètres électriques obtenus en fonction du profil de carbone ($0.15 \times 3.6 \mu m^2$).

On note que la diminution de courant collecteur avec l'augmentation de la dose de carbone a un effet néfaste sur la fréquence de transition f_T , qui chute de 20 GHz. En revanche, la résistance de base pincée étant plus faible grâce à une concentration de bore plus élevée, cette diminution de f_T a peu d'influence sur la fréquence f_{MAX} . L'augmentation conséquente de la tension BV_{CEO} a permis d'augmenter le produit $f_T \times BV_{CEO}$ de 336 à 423 GHz·V.

IV.4.3 Phénomène de saturation des pièges

Le phénomène de recombinaison en base neutre possède des caractéristiques spécifiques ayant leur influence sur le comportement du composant. L'effet le plus important est la saturation des pièges présents dans la base lorsque le courant collecteur augmente. On observe en effet d'après le tracé de Gummel que lorsque la tension augmente, le courant de base subit une transition : il passe d'un niveau élevé, fixé par la recombinaison en base neutre, à un niveau plus faible, identique à la plaque standard sans recombinaison, comme expliqué dans la figure IV.24. Cette transition du courant de base explique également la forme caractéristique du gain en courant en fonction de V_{BE} (IV.23).

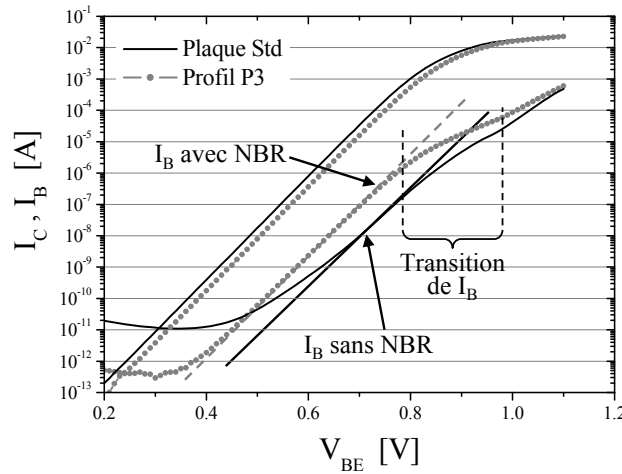


FIG. IV.24 – Mise en évidence de la transition du courant de base pour le profil P3. Au delà de 0.8 V, I_B s'écarte de l'idéalité pour rejoindre le courant standard sans recombinaison.

Ainsi à faible polarisation la recombinaison en base neutre domine, alors qu'elle devient négligeable à fort courant. Cette transition est commandée par le niveau de courant de base. En effet, l'analyse à différentes températures a révélé que le courant est toujours le même au moment de la transition. La figure IV.25 montre que la transition intervient toujours au même courant de base, quelle que soit la température. La figure IV.26 présente l'évolution du gain en courant en fonction de I_C pour différentes températures. Elle montre le contrôle du gain par la recombinaison en base neutre à bas courant, puis au delà d'une certaine valeur de I_C , la saturation des pièges fait que le gain augmente fortement.

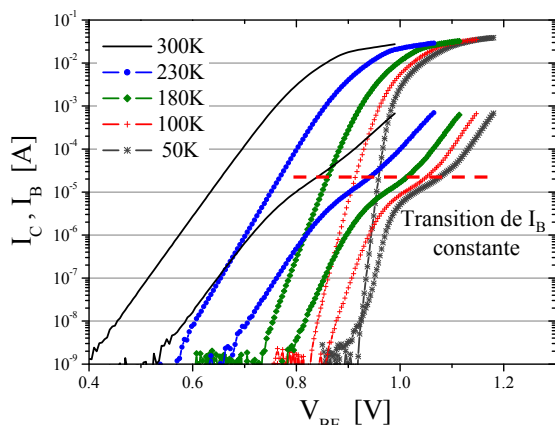


FIG. IV.25 – Tracés de Gummel entre 50 et 300 K du transistor ayant le profil de base P3. Dimensions $0.15 \times 3.6 \mu\text{m}^2$.

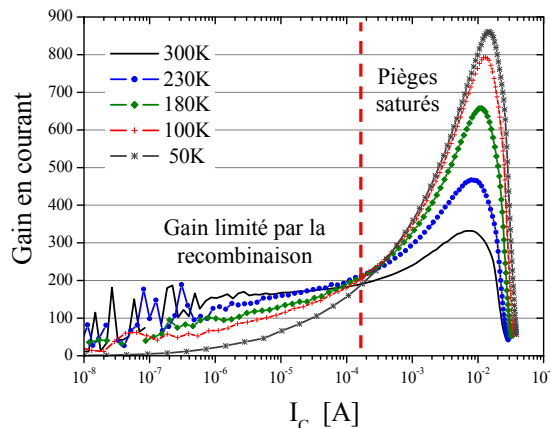


FIG. IV.26 – Évolution du gain en courant en fonction de I_C entre 50 et 300 K. Profil P3, dimensions $0.15 \times 3.6 \mu\text{m}^2$.

Ce phénomène de saturation est directement relié à la dose de carbone présente dans la base des TBH. La figure IV.27 présente les tracés de Gummel à 50 K de transistors bipolaires dotés des profils de base standard, P1, P2 et P3. La transition du courant de base, repérée par un point d'inflexion autour de 1.05 V, intervient d'autant plus tard que la dose de carbone est importante. On note qu'une légère inflexion est également visible sur le courant de base du composant standard.

La figure IV.28 trace l'évolution de la densité de courant de base à la transition en fonction de la dose de carbone. La transition du courant de base est ainsi directement proportionnelle à la quantité de carbone présente : La saturation des pièges intervient d'autant plus tard que le nombre de pièges est grand.

La forme caractéristique du gain en courant observée pour les profils de carbone P2 et P3 a pu être reproduite par des modélisations sous Synopsis DESSIS en prenant en compte la recombinaison dans la base neutre, comme indiqué dans la figure IV.29. Le type de piège considéré est un piège accepteur, ayant une énergie dans la bande interdite fixée à $E_C - 0.17 \text{ eV}$, de type $C_i - C_s$. Cependant, dans la réalité, plusieurs types de pièges cohabitent, et les simulations avec différents pièges d'énergie variable dans la bande interdite ont donné sensiblement les mêmes résultats. Lors des simulations, le niveau de courant est contrôlé uniquement par le produit $C_T \cdot \sigma$, produit de la concentration de piège et de la section efficace de capture d'un piège libre. Pour reproduire les profils P2 et P3, les produits $C_T \cdot \sigma$ ont été fixés respectivement à $2.2 \cdot 10^4$ et $8 \cdot 10^4 \text{ cm}^{-1}$.

Ces simulations permettent de calculer un taux net d'occupation des pièges. L'évolution de ce taux net d'occupation avec V_{BE} est donné figure IV.30 pour le profil P3. On note que en dessous de 0.8 V, le taux d'occupation est quasi-nul, et augmente rapidement à partir de 0.8 V. Ainsi, à partir de 0.8 V, la quantité de pièges disponibles devient très faible en comparaison du nombre de porteurs injectés. La plupart des pièges étant occupés, la majorité des porteurs

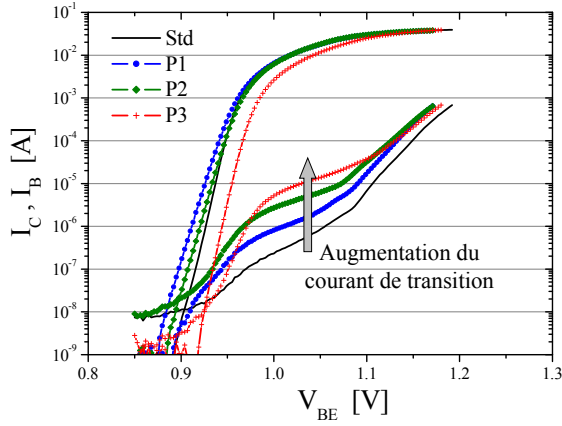


FIG. IV.27 – Tracés de Gummel à 50 K de TBH dotés des profils de base standard, P1, P2 et P3 ($0.15 \times 3.6 \mu m^2$).

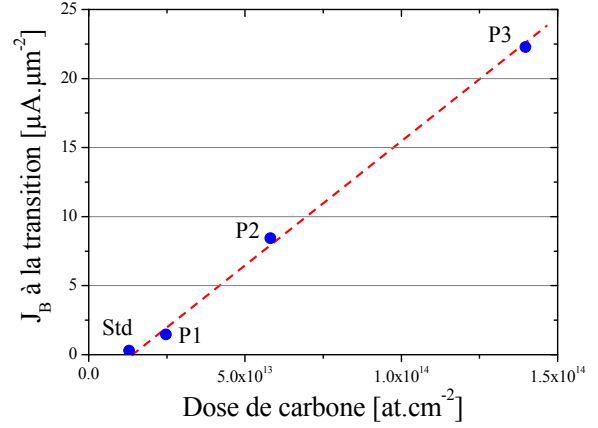


FIG. IV.28 – Densité de courant de base à la transition, en fonction de la dose de carbone ($0.15 \times 3.6 \mu m^2$).

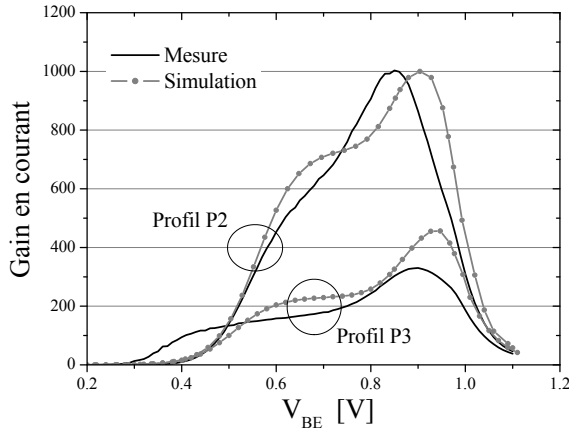


FIG. IV.29 – Évolutions du gain en courant en fonction de V_{BE} , pour les profils P2 et P3. Les simulations reproduisent la forme caractéristique du gain aux fortes tensions V_{BE} .

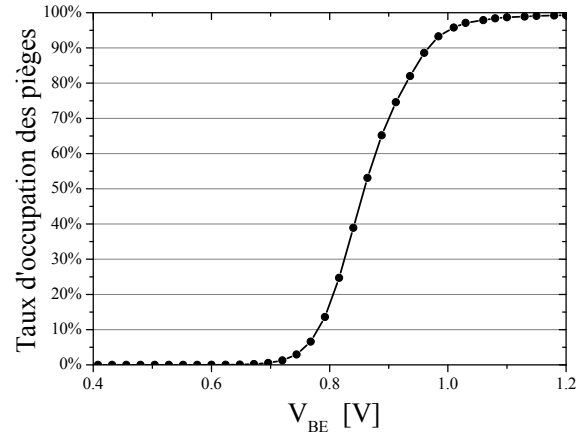


FIG. IV.30 – Taux d'occupation des pièges de la base neutre en fonction de V_{BE} (profil P3). Au delà de 0.8 V, le taux d'occupation augmente fortement.

traverse la base sans être affectée par la recombinaison. La quantité de trous piégés par unité de temps ne peut plus augmenter, ce qui explique la transition du courant de base observée.

IV.4.4 Évolution du bruit basse fréquence

La recombinaison en base neutre a un effet néfaste sur le bruit basse fréquence. Les phénomènes de génération/recombinaison provoquent des fluctuations importantes du nombre de porteurs, génératrices de bruit.

La figure IV.31 présente les spectres de bruit (densité spectrale de bruit S_{IB} en fonction de la fréquence), à 300 K, de TBH de dimensions $0.15 \times 3.6 \mu m^2$, pour chaque profil de carbone réalisé. Le composant standard présente une évolution en $1/f$ à basse fréquence, suivi d'un plateau de shot noise au niveau $2 \cdot q \cdot I_B$. Lorsque la recombinaison en base neutre est accentuée, le comportement en $1/f$ est perdu, et une (ou plusieurs) composante(s) Lorentzienne(s) associé(s) à des processus de génération/recombinaison apparaissent. Cette évolution du bruit est un problème pour les applications utilisant des blocs sensibles au bruit en $1/f$.

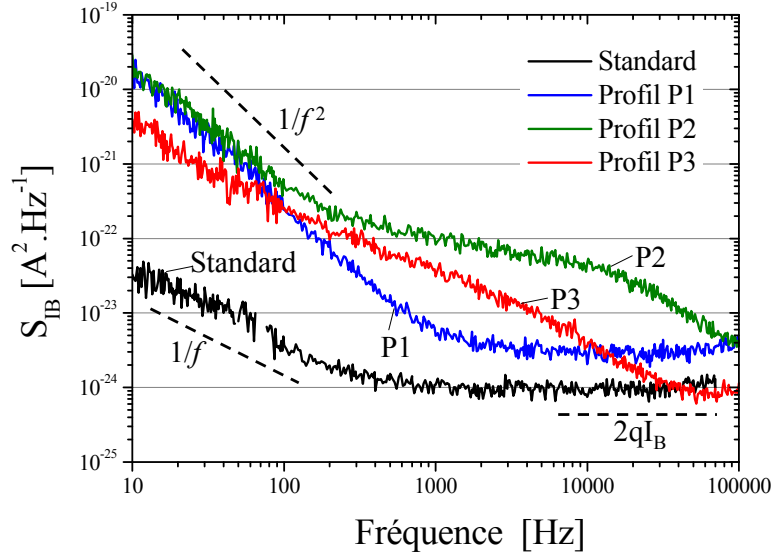


FIG. IV.31 – Spectres de bruit à $I_B = 500$ nA en fonction du profil de base. Dimension $0.15 \times 3.6 \mu\text{m}^2$.

IV.4.5 Influence du carbone sur le gap du matériau

Le carbone, en réduisant les contraintes mécaniques du SiGe, a également un impact non négligeable sur l'énergie de bande interdite de la base [Boucaud94], ainsi que sur le fonctionnement des transistors bipolaires [Osten97]. En reprenant les méthodes décrites par [Ashburn96] et [Jouan01], nous pouvons extraire par des mesures électriques à différentes températures la réduction de bande interdite due au carbone. Les différents paramètres à connaître sont :

- La résistance de base pincée. Celle-ci peut être mesurée aisément, les structures de test spécifiques étant disposées à proximité du composant.
- La mobilité des porteurs. On utilise le modèle de Klaassen [Klaassen92a, Klaassen92b] pour calculer la mobilité des électrons et des trous dans chaque composant, en fonction de la température et du dopage. Faute de modèle plus pertinent, c'est la mobilité des porteurs dans le silicium qui est prise en compte.
- La réduction de bande interdite due au dopage de base. Grâce à l'analyse SIMS, on connaît la concentration de bore, la valeur du BGN est donnée par l'équation I.58 [Klaassen92c].

La méthode repose sur le fait que l'évolution du niveau de courant d'un TBH avec la température dépend de l'énergie de bande interdite de la base. En comparant l'évolution de la densité de courant $J_C(T)$ du TBH avec celle d'un BJT théorique $J_{C0}(T)$, on peut remonter à la réduction de bande interdite totale dans la base du TBH. L'expression générique prise pour J_{C0} est donnée par :

$$J_{C0}(T, V_{BE}) = 4q \left(\frac{m_e m_h}{\hbar^2} \right)^{3/2} (kT)^4 \mu_n \mu_p R_{BP} \exp \frac{qV_{BE} - E_{G, Si}}{kT} \quad (\text{IV.3})$$

où m_e et m_h sont respectivement les masses effectives des électrons et des trous, μ_n , μ_p les mobilités des porteurs, et R_{BP} la résistance de base pincée. Ainsi le rapport des densités de courant J_C/J_{C0} s'exprime en fonction de la température :

$$\frac{J_C(T)}{J_{C0}(T)} \propto \exp \frac{\Delta E_G}{kT} \quad (\text{IV.4})$$

En calculant ce rapport à différentes températures, on a ainsi accès à la valeur de ΔE_G , qui se décompose de la manière suivante :

$$\Delta E_G = E_{Gb} + E_{G,Ge} + E_{G,C} \quad (IV.5)$$

avec ΔE_{Gb} la réduction de bande interdite due au dopage, $\Delta E_{G,Ge}$ et $\Delta E_{G,C}$ les variations de bande interdite dues respectivement au germanium et au carbone. ΔE_{Gb} peut être calculée car nous connaissons le niveau de dopage dans la base, et pour séparer les contributions du germanium et du carbone, on considère que le profil de Ge est identique pour tous les composants. Les valeurs obtenues de réduction de bande interdite sont données dans le tableau IV.6.

Profil	Std	P1	P2	P3
$\Delta E_{G,Total}$ [meV]	188.8	186.74	181.8	171.9
Dopage bore (at.cm ⁻³)	2·10 ¹⁹	2.2·10 ¹⁹	2.7·10 ¹⁹	3·10 ¹⁹
ΔE_{Gb} [meV]	72.6	73.9	76.7	78.1
$\Delta E_{G,Ge} + \Delta E_{G,C}$ [meV]	116.2	112.9	105.1	93.8
$\Delta E_{G,C}$ [meV]	-0.20	-3.5	-11.3	-22.6
C actif [%]	0.01 %	0.14 %	0.45 %	0.91 %

TAB. IV.6 – Variations de bande interdite pour les différents profils de carbone. La réduction due au dopage est prise en compte.

En considérant que $\Delta E_{G,Ge}$ est le même pour les 4 profils de base, soit 116.4 meV, ce qui correspond à un taux de Ge de 15.7 % à l'entrée de la base, on en déduit $\Delta E_{G,C}$, ainsi que le pourcentage de carbone électriquement actif. La concentration de carbone actif reste inférieure à 1 %, et augmente l'énergie de bande interdite d'environ 22.6 meV dans le cas du profil P3. Cette augmentation du gap de la base, doublée d'une quantité de bore plus importante, explique la diminution du courant collecteur lorsque la dose de carbone augmente, et la diminution de f_T .

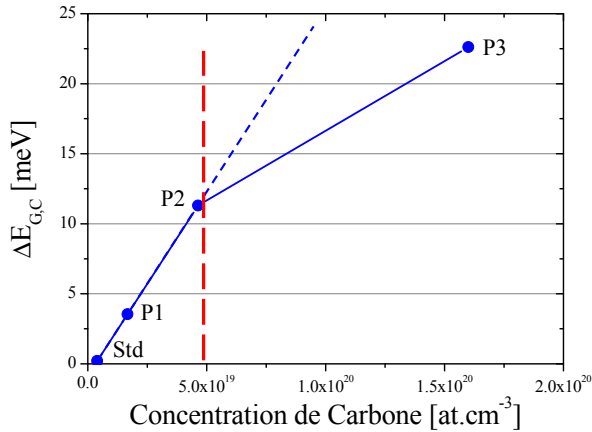


FIG. IV.32 – Variation de bande interdite $\Delta E_{G,C}$ en fonction de la concentration de carbone.

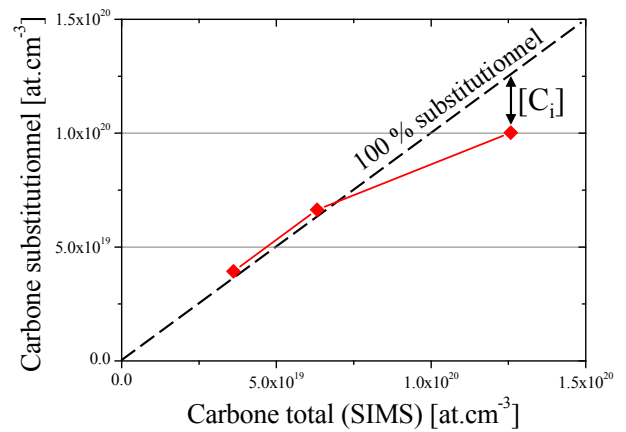


FIG. IV.33 – Concentration de carbone substitutionnel en fonction de la concentration de carbone totale.

La figure IV.32 présente l'évolution de la variation de bande interdite due au carbone $\Delta E_{G,C}$, en fonction de la concentration de carbone dans la base. On note que dans le cas de fortes concentrations de carbone, l'influence de la concentration sur $\Delta E_{G,C}$ est moindre qu'à faible

concentration. Ceci peut être interprété par le fait que seul les atomes de carbone en sites substitutionnels ont une influence sur la structure de bandes. Le carbone interstitiel crée des centres recombinant, mais ne modifie pas la bande interdite du matériau.

La figure IV.33 représente l'évolution de la concentration d'atomes de carbones substitutionnels, mesurée par photoluminescence, en fonction de la concentration totale de carbone, mesurée sur les profils SIMS. Lorsque la dose de carbone augmente une plus grande proportion d'atomes de carbone sont incorporés en site interstitiel. D'après la figure, quasiment tous les atomes sont incorporés en site substitutionnel, jusqu'à une concentration de $6.5 \cdot 10^{19} \text{ at}\cdot\text{cm}^{-3}$, légèrement supérieure à la concentration de carbone obtenue pour le profil P2. Au delà, la proportion d'atomes en sites interstitiels augmente fortement.

IV.4.6 Conclusion sur la recombinaison en base neutre.

En conclusion de cette étude, l'augmentation de la tension BV_{CEO} a bien été démontrée en insérant des pièges dans la base. Bien que le bruit basse fréquence soit fortement dégradé, cette méthode offre de plus grandes possibilités pour l'intégration de composants rapides ayant de fortes tensions de claquage. Le procédé de fabrication est simple car il n'y a pas de modification de l'enchaînement des opérations, seule la recette de base doit être adaptée. La recombinaison des trous intervenant dans la base, cela permet de ne pas dégrader l'efficacité d'injection de la jonction E/B, et ainsi de ne pas avoir de chute trop importante de f_T .

Nous avons mis en évidence un moyen d'augmenter de manière significative la tension de claquage du composant. Nous disposons donc d'une marge de manoeuvre plus importante pour l'optimisation ultérieure du transistor bipolaire : Il est possible d'augmenter le dopage collecteur afin de retarder l'apparition de l'effet Kirk, tout en maintenant une tension de claquage acceptable grâce à la recombinaison en base neutre. Par exemple, un produit $f_T \times BV_{\text{CEO}}$ de $430 \text{ GHz}\cdot\text{V}$ a pu être démontré, si par une modification du dopage collecteur, BV_{CEO} diminue pour atteindre 1.4 V, nous pouvons espérer obtenir une fréquence f_T de 307 GHz. Ce type d'approche doit être mis en place pour l'amélioration future des performances du composant.

IV.5 Synthèse de l'optimisation

IV.5.1 Comparaison des différentes méthodes

Les trois méthodes présentées dans ce chapitre (émetteur métallique, émetteur SiGe et recombinaison en base neutre) ont toutes trois permis d'augmenter sensiblement le courant de base et la tension de claquage BV_{CEO} . Le gain en courant est réduit dans les trois cas d'un facteur 10, ce qui permet de démontrer une tension d'avalanche repoussée d'environ 0.5 V. La figure IV.34 résume les gains en courant et BV_{CEO} obtenus au cours de l'étude. On note le très bon accord existant entre les points expérimentaux et la courbe théorique donnée figure IV.3.

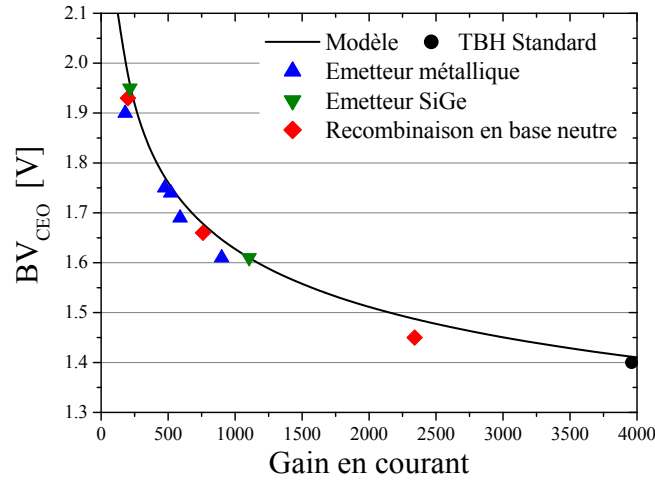


FIG. IV.34 – Evolution de la tension de claquage en fonction du gain en courant, pour l'émetteur métallique, l'émetteur SiGe et la recombinaison en base neutre.

Concernant les performances dynamiques, une diminution de f_T a systématiquement été observée lors de l'optimisation. Cependant, grâce à l'augmentation importante de BV_{CEO} , le produit $f_T \times BV_{CEO}$ a dépassé les 400 GHz·V, comme indiqué dans la figure IV.35.

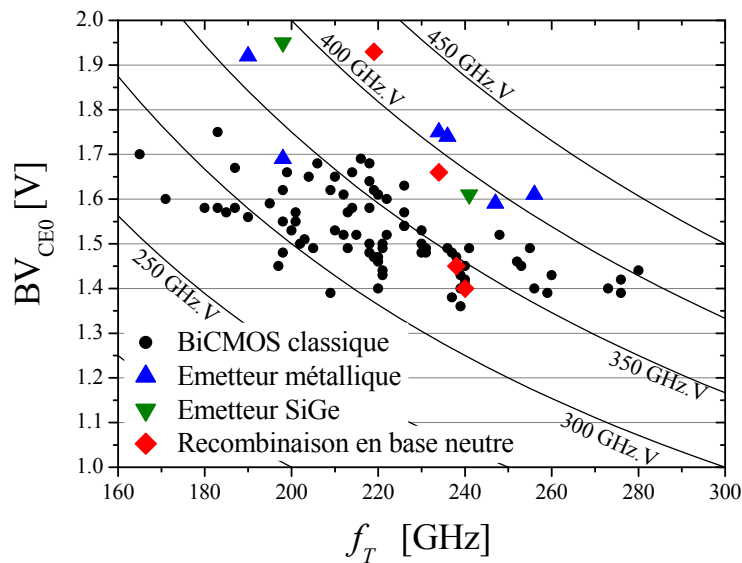


FIG. IV.35 – Positionnement des performances obtenues avec le procédé émetteur métallique, le pic de SiGe dans le Si-cap et la recombinaison en base neutre par rapport aux résultats standards.

IV.5.2 Positionnement par rapport à la concurrence

La figure IV.36 propose une comparaison des résultats obtenus par les trois méthodes avec les données de la concurrence (voir chapitre II.4). Les différentes valeurs de f_T et de BV_{CEO} sont représentées. Les performances démontrées dans ce chapitre se situent autour de 400 GHz·V, en dessous des performances démontrées par l'IHP (380 GHz pour 1.5 V de BV_{CEO}). Cependant les performances au delà de 450 GHz·V sont obtenues dans des technologies intégrant uniquement des transistors bipolaires, et non des technologies BiCMOS complètes.

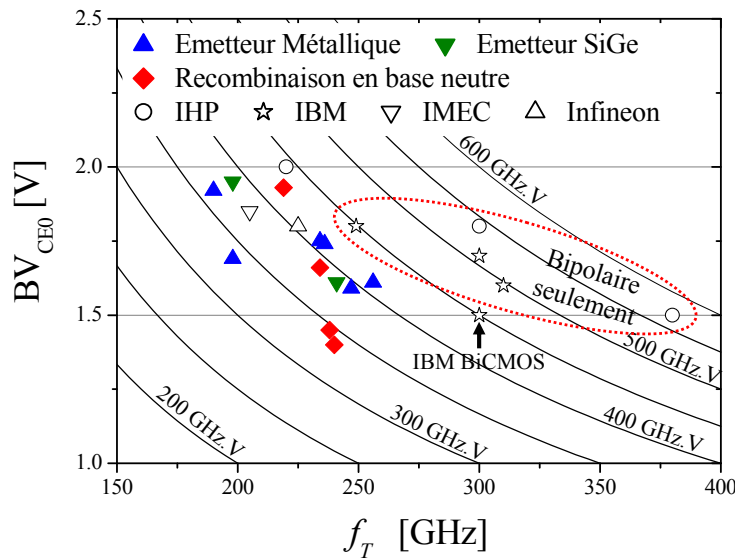


FIG. IV.36 – Positionnement des performances obtenues avec le procédé émetteur métallique, le pic de SiGe dans le Si-cap et la recombinaison en base neutre par rapport à la concurrence.

La recombinaison en base neutre et l'émetteur métallique offrent des valeurs de produit $f_T \times BV_{CEO}$ approchant 450 GHz·V, valeurs démontrées par IBM [Orner06] et par l'IHP [Rucker04] dans leurs technologies BiCMOS respectives. En ce sens, les procédés démontrés dans ce chapitre permettent de rattraper le léger retard de ST vis à vis de la concurrence.

IV.5.3 Conclusion / Perspectives

En conclusion, il est possible d'augmenter de manière importante la tension de claquage des composants par des variantes technologiques. Cependant, l'augmentation de BV_{CEO} a entraîné systématiquement une dégradation des autres paramètres électriques du composant. Il est donc indispensable d'identifier les paramètres importants pour l'utilisation du TBH dans un circuit. Selon les utilisations, et selon les contraintes fixées par les différentes applications en terme de bruit, de gain en courant, de tenue en tension ou de fréquence de fonctionnement, nous sommes en mesure de fournir une solution technologique répondant aux attentes des concepteurs. Cette étude sur le produit $f_T \times BV_{CEO}$ offre des ouvertures par rapport à l'optimisation classique du composant, permettant de sortir du compromis classique fréquence-tension de claquage.

La méthode la plus prometteuse pour l'intégration des composants dans une technologie BiCMOS est l'émetteur SiGe. En effet, comparé à l'émetteur métallique, le procédé de fabrication est beaucoup plus simple. La recombinaison en base neutre offre d'excellentes performances mais présente le défaut d'un bruit en $1/f$ dégradé. Les études visant à incorporer le pic de Ge additionnel dans l'émetteur neutre uniquement sont en cours, et devraient permettre d'obtenir des courants de base avec un facteur d'idéalité proche de 1, en évitant la présence de

Ge dans les jonctions. Les études en cours viseront également à vérifier que la présence de SiGe n'a pas d'effet sur le bruit en $1/f$.

Chapitre V

Influence de la température sur le fonctionnement du TBH

V.1 Introduction

Le fonctionnement du transistor bipolaire dépend fortement de la température, comme cela a été démontré dans le chapitre I. Le chapitre III a mis en évidence la nécessité d'augmenter les densités de courant afin d'obtenir des performances ultimes. Outre la limitation de BV_{CEO} induite par ces forts courants, le transistor bipolaire sera soumis à un auto-échauffement très important. La première partie de ce chapitre a pour but de mettre au point une méthode pour caractériser l'auto-échauffement dans les transistors bipolaires, et proposer des moyens technologiques pour réduire l'élévation de température.

La deuxième partie du chapitre propose une étude approfondie des performances des composants à très basse température. L'analyse cryogénique permet en effet d'améliorer sensiblement les performances des transistors bipolaires SiGe, ce qui donne une information sur les performances théoriquement atteignables par la structure. Dans le même temps, la décomposition des différents retards du composant permet d'identifier le principal facteur limitant la montée en fréquence.

V.2 Étude de l'auto-échauffement

V.2.1 Limitations induites par l'auto-échauffement

L'augmentation des densités de courant nécessaires pour atteindre de forts f_T introduit un auto-échauffement accru dans la partie active du transistor bipolaire. Cet auto-échauffement a une influence néfaste sur le fonctionnement du dispositif. Les performances du transistor étant largement dépendantes de la température, cet auto-échauffement pénalise la montée en fréquence. Des simulations électrothermiques prenant en compte l'auto-échauffement, réalisées sous le logiciel DESSIS de la société Synopsis, montrent que la température de fonctionnement s'élève d'environ 50 K au maximum de f_T , et de plus de 100 K au delà, lorsque le transistor est en situation de forte injection ou d'effet Kirk (voir figure V.1). La perte de performances due à l'auto-échauffement est estimée à environ 20 GHz à température ambiante.

La complexité croissante de la structure du transistor ne joue pas non plus en faveur d'une réduction de l'auto-échauffement : L'empilement de couches diélectriques autour du transistor gêne l'évacuation de la chaleur générée au sein du composant. Notamment, la présence des tranchées profondes modifie la forme du flux de chaleur vers le substrat. L'influence de ce

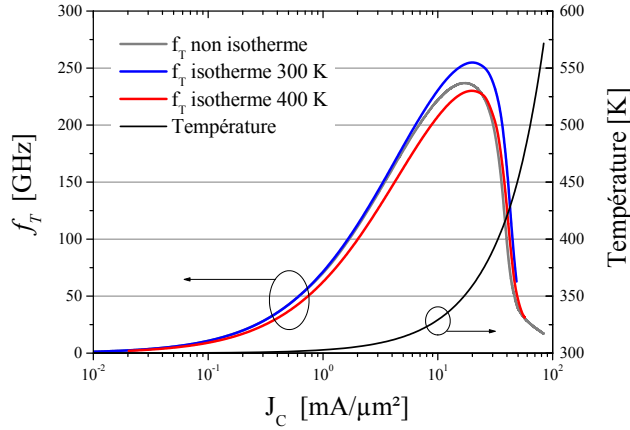


FIG. V.1 – Simulation de courbes de f_T isothermes, et non-isotherme prenant en compte l'auto-échauffement. L'augmentation de température atteint 50 K au maximum de f_T .

paramètre sur l'auto-échauffement est étudié dans la partie V.2.3.b de ce chapitre.

V.2.2 Extraction de la résistance thermique

La prise en compte de l'auto-échauffement dans les composants électroniques passe par l'élaboration d'un réseau thermique, qui vient se juxtaposer au réseau électrique du composant. On peut raisonner par analogie avec l'électrodynamique [Oettinger76] : Le courant électrique est remplacé par la puissance thermique P_{Th} dégagée par le composant, et la variation de température ΔT est assimilée à la tension électrique. Dès lors, en régime statique, la température s'élève proportionnellement à la puissance thermique dégagée. Le rapport entre les variations de ces deux grandeurs s'appelle la résistance thermique R_{Th} du composant, qui s'exprime de la manière suivante :

$$R_{Th} = \frac{\Delta T}{\Delta P_{Th}} \quad (V.1)$$

Cette équation fixe également l'unité de mesure de R_{Th} : C'est le K/W.

Lorsque l'on est en régime transitoire, les courants et tensions du transistor varient rapidement, ainsi que la puissance thermique dégagée. Cependant, compte-tenu des capacités calorifiques des différentes couches en présence, la température ne varie pas instantanément avec les variations de P_{Th} . En première approximation, le retard d'établissement de la température est modélisé par une capacité thermique C_{Th} . Le réseau thermique équivalent est donné par la figure V.2.

Un modèle plus fin de réseau thermique consiste en une succession de couples R_{Th} - C_{Th} , ayant chacun leur propre constante de temps [Mnif02].

V.2.2.a Extraction en régime continu

Pour calculer la résistance thermique du composant, il faut avoir une mesure de sa température. Les méthodes de mesures de T sont nombreuses et variées (optiques, électriques, par contact physique) [Blackburn04]. Nous nous intéressons uniquement aux mesures électriques.

La résistance thermique peut être extraite par des mesures électriques en utilisant le fait que les relations entre courants et tensions du transistor bipolaire font systématiquement intervenir au moins un terme dépendant explicitement de la température [Oettinger76]. On se sert d'un paramètre électrique du composant comme d'un thermomètre pour remonter à la température

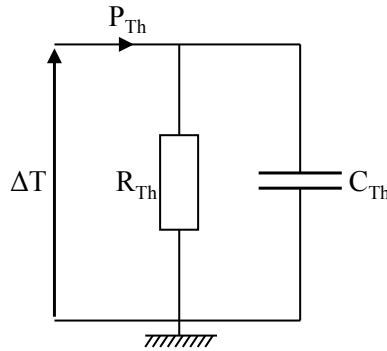


FIG. V.2 – Réseau thermique simple.

interne du transistor (généralement le gain en courant ou la tension V_{BE}). Lors de l'extraction en régime continu, on considère que les échanges thermiques sont établis avant de faire la mesure. Comme on n'a pas accès au régime transitoire, on ne peut extraire que la valeur de la résistance thermique R_{Th} , et non la valeur de C_{Th} .

L'extraction de la résistance thermique par des mesures électriques repose sur une série d'approximations [Oettinger76] : Une température de jonction uniforme sur toute la surface du composant, une résistance thermique constante quelles que soient les conditions de polarisation, et une conductivité thermique du Si et du SiGe indépendante de la température.

- Tout d'abord la température de jonction n'est pas uniforme sur toute la largeur de la base [Gao91]. La mesure électrique ne peut nous donner qu'une information sur la température moyenne du transistor. Cette température moyenne peut être bien inférieure à la température maximale dans le composant.
- R_{Th} n'est pas constante quelles que soient les conditions de polarisation du transistor. En effet, à puissance dissipée constante, la conformation de la source de chaleur peut varier (par exemple un fort courant I_C et une faible tension V_{CB} , ou bien un faible courant I_C et une forte tension V_{CB} , ne donneront pas les mêmes conformations de ZCE B/C). La répartition de la puissance dissipée sera différente, débouchant sur des valeurs de R_{Th} différentes.
- La conductivité thermique K du silicium et du SiGe varient avec la température (la figure V.3 donne l'évolution de $K(T)$ pour le silicium). Il faut donc s'arranger pour que l'extraction de R_{Th} ne mette en jeu que de faibles variations de température.

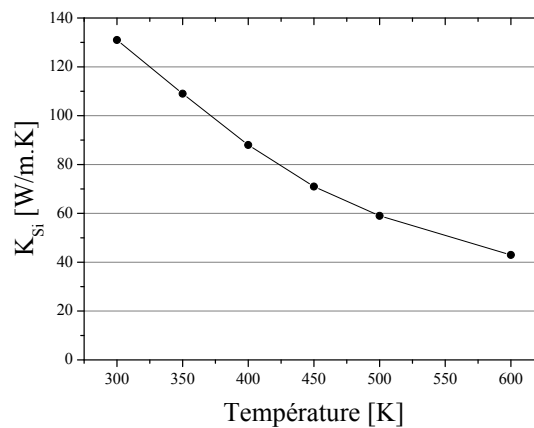


FIG. V.3 – Conductivité thermique du silicium en fonction de la température.

Malgré toutes ces approximations, il est possible d'extraire un facteur de mérite significatif

de l'auto-échauffement dans les transistors bipolaires. Plusieurs méthodes d'extraction de R_{Th} en régime statique existent dans la littérature. La méthode proposée par [Gao91] et développée dans [Ooi04] propose une extraction de R_{Th} en une seule phase de mesure (en exploitant une caractéristique de sortie $I(V)$), mais nécessite la connaissance précise de l'énergie de bande interdite du matériau et du facteur d'idéalité de la jonction émetteur/base du composant. L'énergie de bande interdite de la base étant soumise à de fortes variations au cours de nos études, cette méthode a été écartée.

Les méthodes décrites par [Waldrop92, Dawson92] utilisent la dépendance du gain en courant avec la température. Nous expliciterons dans cette partie la méthode présentée dans [Rieh01], reposant sur une mesure des variations de V_{BE} avec la température et avec la puissance thermique appliquée au composant. C'est en effet la méthode la plus facile à mettre en œuvre expérimentalement, compte-tenu des moyens disponibles, car elle ne nécessite la connaissance préalable d'aucun paramètre. La puissance thermique instantanée du composant s'exprime en fonction des courants et des tensions appliqués au transistor :

$$P_{Diss} = I_E V_{BE} + I_C V_{CB} = I_C V_{CE} + I_B V_{BE} \quad (V.2)$$

La mise en œuvre de cette extraction de R_{Th} nécessite un banc de mesure sous pointe contrôlé en température, connecté à un analyseur de paramètres Agilent 4156 pour l'acquisition des courbes. La mesure de R_{Th} est réalisée en deux phases de mesures successives :

- Tout d'abord une mesure de V_{BE} à différentes températures, pour une valeur de I_E donnée (figure V.4). On obtient une loi de V_{BE} en fonction de T .
- Ensuite une mesure de V_{BE} à différents P_{Th} , pour une température donnée, à I_E constant. La variation de P_{Th} est obtenue en faisant varier V_{CB} . La température du substrat est fixée. On obtient, grâce à l'équation V.2, une loi de V_{BE} en fonction de P_{Th} (figure V.5).

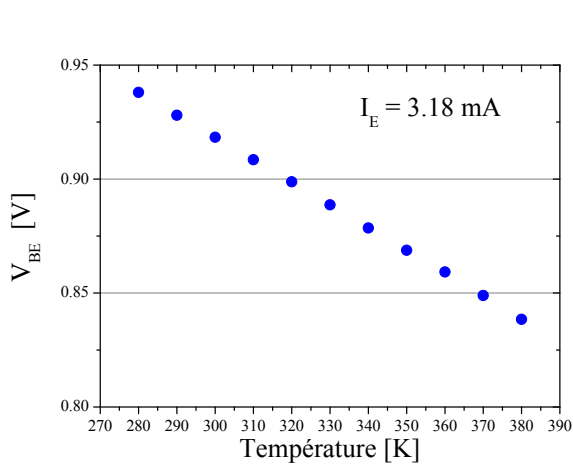


FIG. V.4 – Evolution de la tension V_{BE} avec la température du banc. $I_E = 3.18$ mA et $V_{CB} = 0$ V.

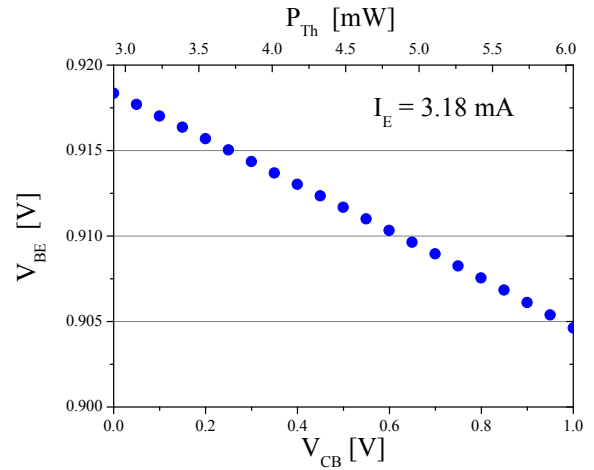


FIG. V.5 – Evolution de la tension V_{BE} avec la tension V_{CB} et la puissance thermique dissipée. $I_E = 3.18$ mA.

La valeur de I_E est la même dans les deux phases de mesure. Il faut que le courant I_E retenu corresponde à une zone de fonctionnement du composant où l'auto-échauffement est présent, typiquement à proximité du maximum de f_T .

Connaissant l'évolution de la puissance thermique P_{Th} en fonction de V_{BE} , et la loi d'évolution de V_{BE} avec la température, on en déduit la loi d'évolution de P_{Th} avec T , donc la valeur de R_{Th} . Une régression linéaire sur la première série de mesure permet d'exprimer T en fonction de V_{BE} . On en déduit la courbe traçant T en fonction de P_{Th} donnée dans la figure V.6. Cette

caractéristique est une droite, dont la pente a pour valeur la résistance thermique R_{Th} spécifique du composant. Il faut ajouter un terme correctif à la température obtenue : En effet, il existe un écart entre la température du substrat et du transistor en fonctionnement. On fixe $T = 300K$ lorsque la puissance appliquée au composant est nulle.

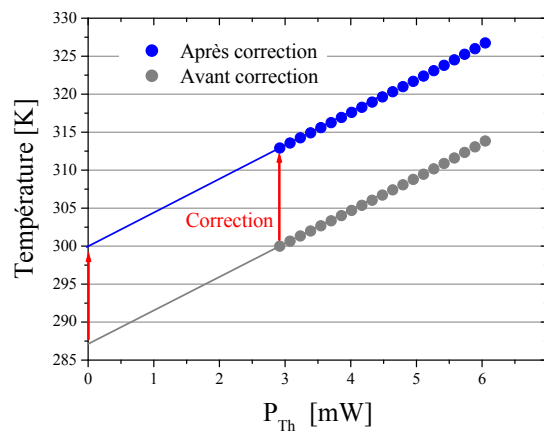


FIG. V.6 – Variation de T en fonction de la puissance thermique. Une correction est nécessaire pour tenir compte du décalage en température existant entre le substrat et le transistor en fonctionnement.

Cette méthode a le double avantage de donner par une mesure directe la valeur de R_{Th} du composant ainsi qu'une indication sur la température moyenne du transistor lorsque celui-ci est en fonctionnement.

V.2.2.b Extraction en régime pulsé

Le calcul de la résistance thermique peut également être réalisé en régime impulsionnel court. De courtes impulsions de courant collecteur sont appliquées au composant, qui n'a pas le temps d'atteindre le régime permanent au cours de l'impulsion. La société AMCAD, en partenariat avec l'université de Limoges (XLIM) propose des mesures pulsées, avec des durées d'impulsion de 400 ns. Cependant, même sur une impulsion de 400 ns, le transistor a le temps de s'échauffer. On peut considérer en première approximation, compte tenu du fait que les constantes de temps thermiques sont beaucoup plus longues que les constantes de temps électriques, que la température sera constante sur toute la durée de l'impulsion. La difficulté réside dans la mise en œuvre de mesures de durée inférieure à 400 ns. L'extraction en régime pulsé nécessite deux phases de mesure successives, et utilise l'évolution d'un paramètre sensible à la température, comme dans la méthode décrite plus haut.

- Une première série en régime pulsé : On applique des impulsions de courant I_C (correspondant à une zone dans laquelle l'auto-échauffement est présent), et on mesure la tension V_{BE} en fonction de la température à V_{CE} constant. On en déduit $\Delta V_{BE}/\Delta T$. Le transistor étant mesuré en régime pulsé, la température interne du composant est égale à la température du substrat, car le composant n'a pas le temps de s'échauffer pendant l'impulsion de courant.
- Une deuxième série en régime statique, identique à celle décrite précédemment : On fixe la température de la plaque, et on étudie les variations de V_{BE} avec la puissance thermique appliquée, en faisant varier V_{CB} . On en déduit l'expression de $\Delta V_{BE}/\Delta P_{Th}$.

La combinaison de ces deux séries de mesure permet d'extraire la loi de variation de T en fonction de P_{Th} . L'avantage de ce type de mesure est qu'elle ne nécessite pas de correction :

lors de la première phase de mesure, la température du composant est supposée connue, égale à la température appliquée.

V.2.2.c Comparaison des différentes méthodes

Les deux méthodes, en régime statique ou pulsé, reposent sur l'évolution de la tension V_{BE} avec la température. En comparant les deux méthodes, on s'aperçoit que les valeurs de R_{Th} extraites sont en bon accord, comme le prouve le tableau V.1.

Surface de composant (μm^2)	$12 \times (0.15 \times 0.49)$	$7 \times (0.15 \times 0.83)$	$5 \times (0.15 \times 1.17)$
R_{Th} statique [K/W]	820	1062	1196
R_{Th} pulsé [K/W]	773	1091	1237

TAB. V.1 – Comparaison des valeurs R_{Th} extraites en régime statique et pulsé pour différentes géométries de transistors.

V.2.3 Réduction de l'auto-échauffement par des variantes technologiques

L'auto-échauffement du composant peut être amélioré par des variantes lors de la réalisation du transistor. Les facteurs sur lesquels on peut influencer sont :

- La création de chaleur : plus la surface active est faible, moins la chaleur dégagée sera importante, donc l'auto-échauffement sera réduit.
- La répartition des sources de chaleur : Pour une même surface active, si les sources de chaleur sont espacées, la température du composant sera plus faible.
- L'évacuation de la chaleur : Elle se fait principalement par le substrat. Les isolants électriques sont généralement de mauvais conducteurs de chaleur, et même si les métaux sont de très bons conducteurs thermiques, la résistance thermique du back-end dans son ensemble reste très supérieure à celle du substrat. Il est donc possible d'améliorer R_{Th} en éliminant les éléments perturbant l'évacuation de chaleur dans le substrat, tels que les tranchées profondes d'isolation (DTI). Le type de substrat retenu influe également sur l'évacuation de la chaleur (substrat massif ou SOI).

V.2.3.a Influence de la dimension de l'émetteur

La figure V.7 trace l'évolution de la résistance thermique extraite par la méthode décrite au paragraphe V.2.2.a, pour différentes longueurs d'émetteur L_E comprises entre 0.64 et 14.92 μm . On voit que la résistance thermique diminue lorsque L_E augmente. L'effet est cependant trompeur, si l'on rapporte cette mesure de R_{Th} à la surface active générant la chaleur. L'évolution du terme $R_{Th} \times A_E$, où A_E est la surface effective de l'émetteur et donc de la source de chaleur, prouve que les composants de plus petite taille sont avantagés.

La résistance thermique est une mesure de l'évacuation de la chaleur, et ne dépend que de l'environnement du composant. Cependant, cette valeur de R_{Th} n'est pas suffisante. L'auto-échauffement du transistor ne dépend pas uniquement de l'évacuation de la chaleur, mais du rapport entre la quantité de chaleur générée et le volume de silicium disponible pour son évacuation. Les composants les plus longs ont un plus grand volume pour l'évacuation de la chaleur, mais leur densité de source de chaleur est plus importante.

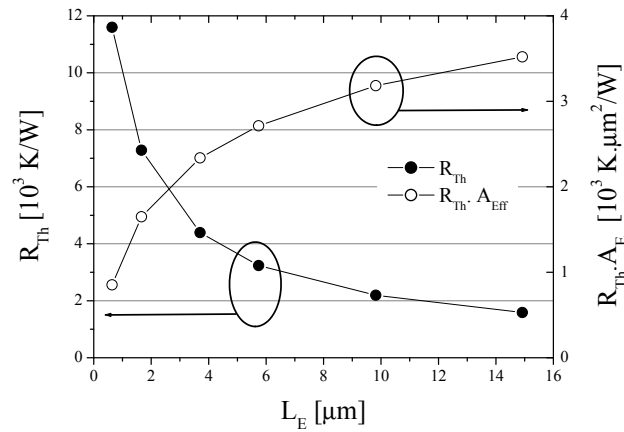


FIG. V.7 – Evolution de la résistance thermique en fonction de la longueur d'émetteur. La largeur d'émetteur est 0.14 μm .

Les longs composants sont donc d'avantage sujets à l'auto-échauffement. Ceci explique en partie la perte de performances dynamiques pour les composants de grande taille. Le tableau V.2 résume les valeurs de f_T et f_{MAX} obtenues en fonction de la longueur d'émetteur.

Longueur d'émetteur [μm]	3.7	9.82	14.92	30.22
R_E [Ω]	13.6	4.2	2.7	1.4
β à $f_{T\max}$	545	543	522	513
J_C à $f_{T\max}$ [$\text{mA}/\mu\text{m}^2$]	13.2	12.2	12.2	12.6
f_T [GHz]	228	225	221	206
f_{MAX} [GHz]	255	230	217	179

TAB. V.2 – Évolution des paramètres dynamiques en fonction de la longueur d'émetteur. Largeur du composant : 0.14 μm , contacts en matrice.

Les fréquences de transition f_T et f_{MAX} diminuent lorsque L_E augmente. La densité de courant au pic de f_T est constante quelle que soit la longueur d'émetteur, on peut donc relier la diminution de f_T à l'auto-échauffement, les résistances d'accès étant améliorées avec l'augmentation de L_E . Cependant, il n'est pas possible de réduire indéfiniment la taille du composant, car on se pénalise alors sur les résistances séries du dispositif.

Afin de réduire efficacement l'auto-échauffement du composant, il est également intéressant de diminuer la largeur d'émetteur : On ne change pas le volume disponible pour l'évacuation de la chaleur, donc R_{Th} n'évolue quasiment pas, mais la source de chaleur est plus petite. Le produit $R_{Th} \times A_E$ est donc fortement amélioré pour les dispositifs possédant une faible largeur d'émetteur. Les émetteurs actuels ont une largeur effective réduite jusqu'à 0.12 μm , pour une fenêtre émetteur de 0.27 μm de large. Le tableau III.12 présenté à la page 113 montre une amélioration de f_{MAX} lorsque la largeur d'émetteur W_E est réduite, ce qui est plutôt imputable à la diminution de la résistance de base. La réduction de l'auto-échauffement permet à la fréquence f_T de se maintenir à un niveau constant, malgré l'augmentation de R_E lorsque W_E diminue.

V.2.3.b Effet des tranchées profondes d'isolation

Sachant que la quasi-intégralité de la chaleur est évacuée par le substrat, les tranchées profondes d'isolation (DTI) perturbent l'évacuation de la chaleur. Afin d'évaluer l'influence des

tranchées profondes sur le comportement thermique du transistor, nous avons réduit leur profondeur de 5.2 μm (standard) à 3.5 et 2.5 μm . Des transistors sans module DTI ont également été réalisés.

Des simulations électrothermiques à 2 dimensions ont été réalisées sous DESSIS afin d'évaluer plus précisément l'influence des DTI sur l'auto-échauffement et les performances dynamiques. La structure simulée a une dimension de 300 μm de haut par 600 μm de large. Une température de 300 K est fixée aux bornes de la zone simulée.

Ces simulations montrent qu'une grande partie de la chaleur générée est évacuée entre les tranchées profondes (figure V.8a, b et c). En dehors de la zone délimitée par les tranchées, les isothermes retrouvent leur forme circulaire, signe d'une évacuation isotrope de la chaleur. Les tranchées jouant le rôle de «canal» d'évacuation du flux thermique, plus les tranchées sont profondes, plus la résistance thermique est élevée, comme le prouve la figure V.9.

Lorsqu'il n'y a pas d'isolation par DTI (figure V.8d), le flux de chaleur est isotrope autour de la zone active du composant, conduisant à une plus faible élévation de température, ce qui améliore fortement R_{Th} par rapport aux transistors dotés de tranchées profondes.

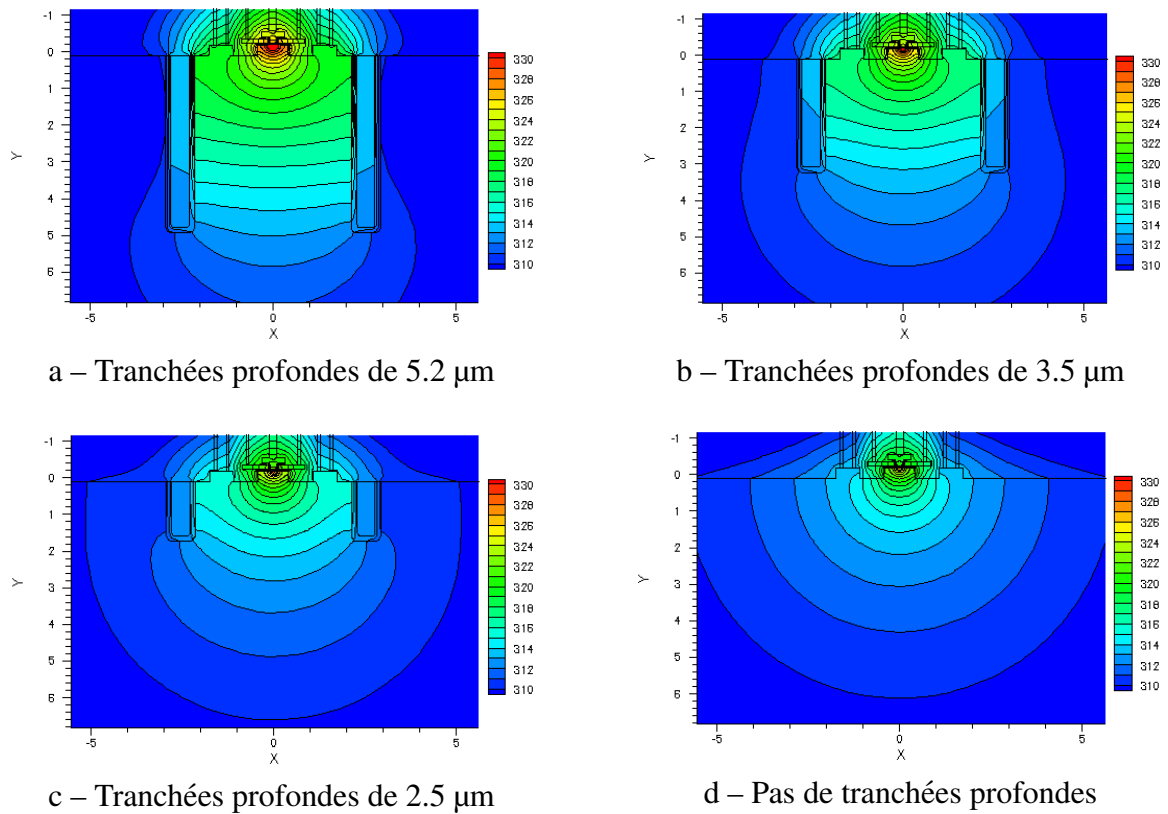


FIG. V.8 – Répartitions de température dans le substrat en fonction de la profondeur des DTI.

Le tableau V.3 présente les performances statiques et dynamiques de transistors de 3.6 μm de long, pour les 4 profondeurs de tranchées décrites précédemment. On note une diminution significative de la valeur de R_{Th} , ainsi qu'une bonne amélioration des performances dynamiques.

Cependant, l'augmentation de f_T et f_{MAX} observée n'est pas uniquement imputable à une réduction de l'auto-échauffement. En effet on constate que le courant collecteur à moyenne injection ($V_{\text{BE}} = 0.75 \text{ V}$) augmente lorsqu'on réduit la profondeur des tranchées profondes. Cette augmentation de I_C en dehors de tout effet d'auto-échauffement est probablement due à des variations des conditions de croissance de la base lors de l'épitaxie. La variation de profondeur des DTI provoque une variation locale de la température du substrat, ce qui modifie la croissance de la base. L'évolution de la capacité C_{BE} confirme bien que les jonctions E/B ne sont pas réglées

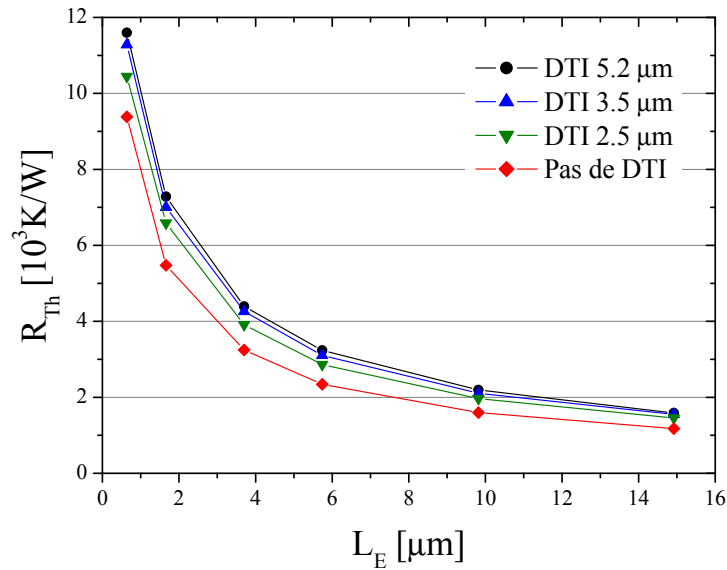


FIG. V.9 – Évolution de la résistance thermique en fonction de la profondeur des tranchées profondes, pour une longueur d'émetteur comprise entre 0.6 et 15 μm .

Profondeur de DTI [μm]	5.2	3.5	2.5	0
I_C à 0.75 V [μA]	118	123	143	152
β à $V_{BE} = 0.75$ V	2350	2380	2520	2480
C_{BE} à $V_{BE} = 0$ [fF]	11.0	11.1	12.6	12.1
C_{CS} à $V_{CS} = 1.5$ V [fF]	4.9	4.7	5.1	14.8
f_T [GHz]	252	253	263	266
f_{MAX} [GHz]	274	277	279	282
R_{Th} [K/W]	4391	4259	3910	3247

TAB. V.3 – Influence de la profondeur des tranchées sur les caractéristiques des transistors (taille de composant $0.15 \times 3.6 \mu\text{m}^2$, contact émetteur ruban).

de la même manière d'un composant à l'autre.

Les valeurs de f_T et les températures maximales obtenues par simulation sont répertoriées dans le tableau V.4. L'élévation de température maximale à la jonction E/B du composant est de l'ordre de 50 K. f_T augmente légèrement (2 GHz) avec la réduction de la profondeur des tranchées, grâce à une diminution de la température maximale d'environ 4 K. Cette variation de température est très faible, et n'induit pas d'amélioration significative des performances du composant.

Profondeur de DTI [μm]	5.2	3.5	2.5	0
f_T [GHz]	236.2	237.6	238.4	238.2
Température max. [K]	352	349	348	348

TAB. V.4 – Simulations électrothermiques 2D visant à étudier l'impact des tranchées sur les performances du composant.

Les mesures manuelles d'extraction de R_{Th} permettent d'extraire la température en fonction de la puissance appliquée au composant. Il s'agit d'une température moyenne autour de la source de chaleur. D'après la figure V.10, cette élévation de température moyenne est de l'ordre de 20 K, avec peu de différences selon les profondeurs de tranchées. En conclusion, l'isolation

par tranchées profondes réduit l'évacuation de la chaleur, donc a une forte influence sur la valeur de la résistance thermique du composant. Cependant, les densités de puissance restent faibles, ce qui fait que l'élévation de température reste raisonnable (moyenne de +20 K au pic f_T), et la différence entre les composants est mineure, sans influence majeure sur les performances dynamiques [Barbalat05].

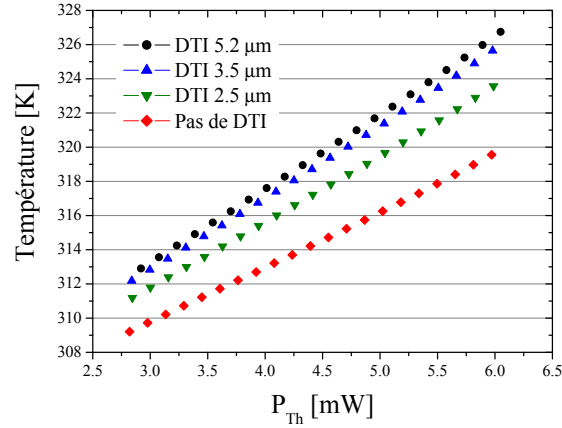


FIG. V.10 – Température moyenne en fonction de la puissance thermique appliquée, pour plusieurs profondeurs de DTI ($0.15 \times 3.6 \mu m^2$, contact émetteur ruban).

La distance du DTI par rapport à la partie intrinsèque du composant (zone active) a aussi une influence sur la valeur de R_{Th} . La distance standard est de $0.4 \mu m$. Des transistors avec des DTI placés à 0.6 et 0.8 de la zone active ont été réalisés, et sont comparés à un dessin sans DTI. Le tableau V.5 donne les valeurs de R_{Th} en fonction de la distance entre DTI et la zone active, pour des transistors de $9.7 \mu m$ de long, ainsi que les valeurs de f_T et f_{MAX} associées pour des transistors de $3.6 \mu m$ de long. La distance entre les tranchées et la zone active du composant joue peu sur les performances dynamiques. On note cependant que lorsque les DTI ne sont pas réalisés, f_T est amélioré (+3 GHz), et f_{MAX} est légèrement en retrait (-13 GHz), ce qui est dû à une capacité collecteur-substrat plus forte.

Distance DTI-Active [μm]	0.4	0.6	0.8	∞
R_{Th} [K/W]	1478	1474	1466	1352
f_T [GHz]	271	270	272	274
f_{MAX} [GHz]	280	278	279	267

TAB. V.5 – Valeurs de R_{Th} et principaux paramètres dynamiques en fonction de la distance entre les DTI et la zone active. Transistors CBECB, $0.15 \times 9.7 \mu m^2$ (R_{Th}), $0.15 \times 3.6 \mu m^2$ (f_T et f_{MAX}).

Cette étude permet de constater que le principal paramètre jouant sur la valeur de la résistance thermique est l'absence, ou bien la présence, des tranchées profondes. La distance ou la profondeur de ces tranchées jouent peu sur la valeur de R_{Th} et sur l'auto-échauffement du composant. L'élévation de température est quasi-identique avec et sans DTI.

V.2.3.c Émetteurs fractionnés ou multi-doigts

D'après la partie V.2.3.a, l'auto-échauffement du composant est amélioré lorsqu'on réduit la surface de l'émetteur par rapport au volume de silicium disponible pour l'évacuation de la chaleur. Pour obtenir de forts courants, il est donc possible de faire fonctionner plusieurs

émetteurs en parallèle, qui présenteront un auto-échauffement réduit par rapport au transistor mono-doigt de surface identique. La réduction de l'auto-échauffement dans le cas d'émetteurs multi-doigts est due à deux effets distincts :

- Le ratio périmètre/surface est plus favorable lorsque le fractionnement de l'émetteur augmente : L'évacuation de la chaleur est améliorée.
- La température maximale de jonction sera plus faible dans le cas de transistors multi-doigts que dans un transistor mono-doigt de surface identique : Les sources de chaleur sont plus étalées dans l'espace.

Pour illustrer cet effet, nous avons extrait les valeurs de R_{Th} pour plusieurs transistors de surface d'émetteur équivalentes, présentant divers fractionnements. Les fractionnements sont de deux types : Dans le premier cas, la distance entre les doigts d'émetteurs est constante, donc la surface totale du composant augmente lorsque le fractionnement augmente (figure V.11-a). Dans le deuxième cas, nous augmentons le nombre de doigts d'émetteurs, en gardant constant le volume de silicium contenu dans les DTI (figure V.11-b).

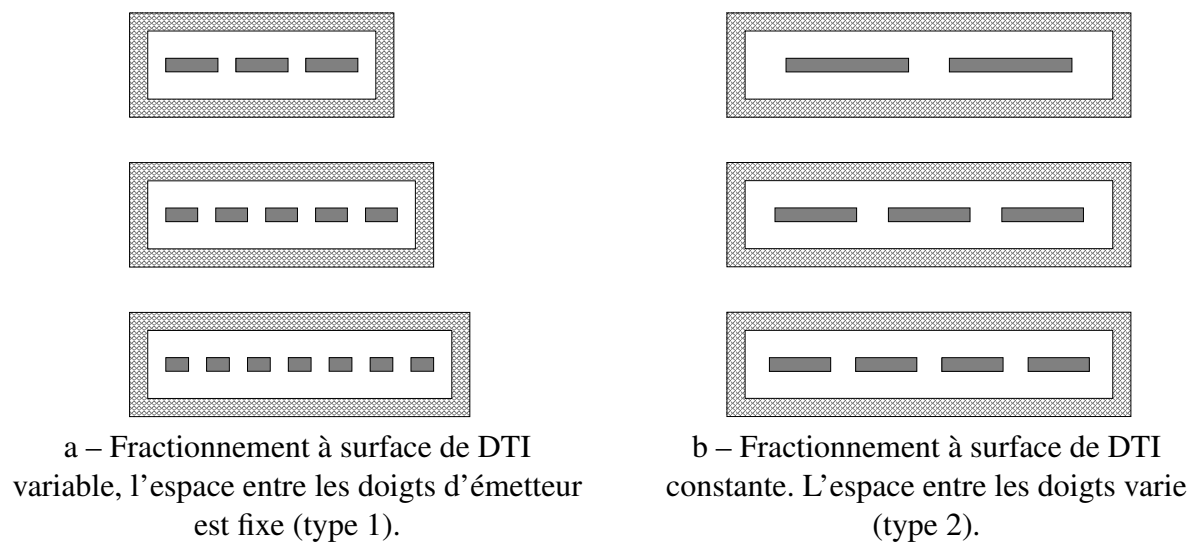


FIG. V.11 – Présentation des deux types de fractionnement possibles.

Le tableau V.6 recense les valeurs de R_{Th} extraites pour divers fractionnements d'émetteurs, avec espacement constant entre les différents doigts d'émetteur. La longueur totale d'émetteur est de $5.7 \mu m$. Chaque cellule a des contacts de base déportés, de type $C_B E^B C$ [Chevalier05b].

Surface d'émetteur	R_{Th} avec DTI	R_{Th} sans DTI
$12 \times (0.15 \times 0.49)$	1213	933
$7 \times (0.15 \times 0.83)$	1571	1208
$5 \times (0.15 \times 1.17)$	1683	1360

TAB. V.6 – Valeurs de R_{Th} pour divers fractionnements d'émetteur (type 1).

Le tableau V.6 met en évidence l'effet du fractionnement sur la résistance thermique : Lorsque le nombre de cellules augmente, R_{Th} diminue, pour une surface active d'émetteur identique.

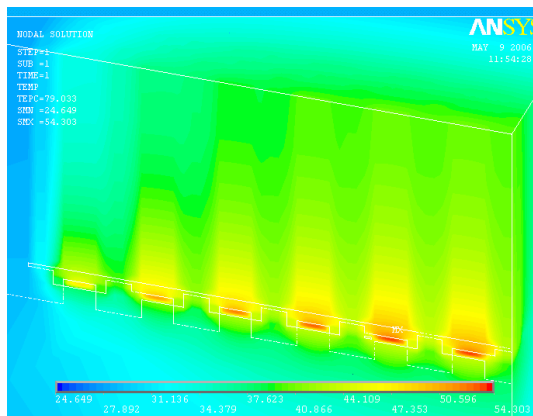
Le même type d'étude est également réalisé sur un fractionnement où le volume de silicium contenu entre les tranchées est constant (figure V.11-b). Le tableau V.7 résume les valeurs de R_{Th} obtenues, avec et sans DTI. Chaque doigt d'émetteur a une structure CBEC classique, avec isolation par STI entre les contacts collecteurs et la partie active émetteur/base.

Surface d'émetteur (μm^2)	R_{Th} avec DTI	R_{Th} sans DTI
$2 \times (0.15 \times 3)$	2226	1964
$3 \times (0.15 \times 2)$	2164	1908
$4 \times (0.15 \times 1.5)$	2130	1960
$5 \times (0.15 \times 1.2)$	2119	1967

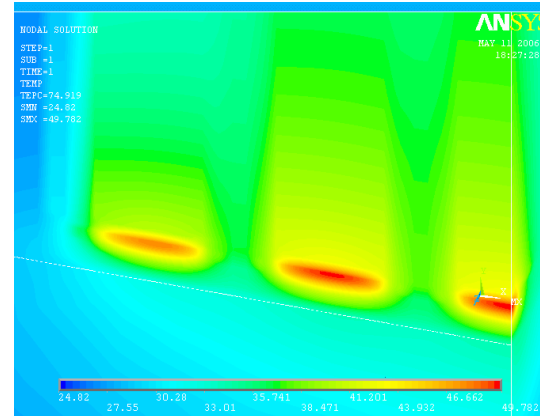
TAB. V.7 – Valeurs de R_{Th} pour différents fractionnements d'émetteur, en fonction du type de substrat (type 2).

La tendance globale est de diminuer la résistance thermique lorsque le fractionnement augmente, ce qui est surtout visible sur les transistors dotés de DTI. Dans le cas où les tranchées sont absentes, R_{Th} varie peu en fonction du fractionnement de l'émetteur.

L'auto-échauffement dans un TBH à émetteur multi-doigts est un phénomène complexe [McAllister04]. Des simulations thermiques 3D sous ANSYS ont été réalisées par l'université de Limoges et la société AMCAD, mettant en évidence l'effet de couplage thermique entre les différents doigts d'émetteur. La figure V.12 présente une simulation des répartitions de température pour deux transistors de $12 \times (0.15 \times 0.49 \mu\text{m}^2)$ et $5 \times (0.15 \times 1.17 \mu\text{m}^2)$. La structure simulée représente uniquement un quart de transistor. Les simulations font également apparaître le fait que le doigt d'émetteur situé au centre de la structure chauffe le plus, celui étant le plus éloigné du centre est plus froid de quelques degrés, comme le montre la figure V.13.



a – 12 cellules de $0.15 \times 0.49 \mu\text{m}^2$.



b – 5 cellules de $0.15 \times 1.17 \mu\text{m}^2$.

FIG. V.12 – Simulation thermiques sous Ansys simulant deux types de fractionnement d'émetteur (structure faible coût).

Les conclusions de [Rieh05a] sur l'influence du fractionnement ont été retrouvées dans notre étude : On note une très faible influence du fractionnement de l'émetteur sur les performances dynamiques des TBH. Le tableau V.8 donne les valeurs de f_T et f_{MAX} obtenues sur des transistors présentant un fractionnement de type 1. L'auto-échauffement n'est pas un facteur limitant dans ce type de structure : La réduction des capacités parasites permet d'améliorer fortement les performances lorsque le nombre de doigts d'émetteur diminue, alors que la résistance thermique augmente [Chevalier05b].

V.2.3.d Effet du type de substrat

Les transistors étudiés dans cette étude sont réalisés sur substrat massif. Des transistors bipolaires sur substrat SOI hautement résistif sont également en cours d'étude à ST, afin de pouvoir

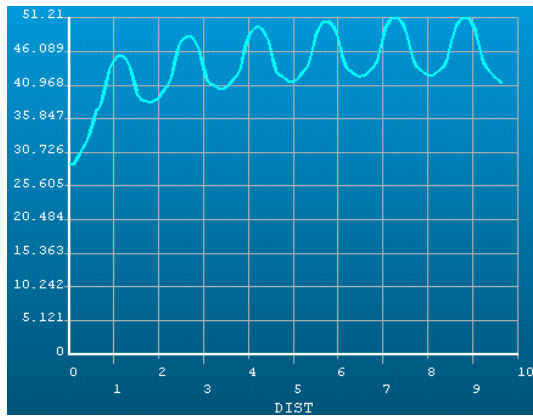
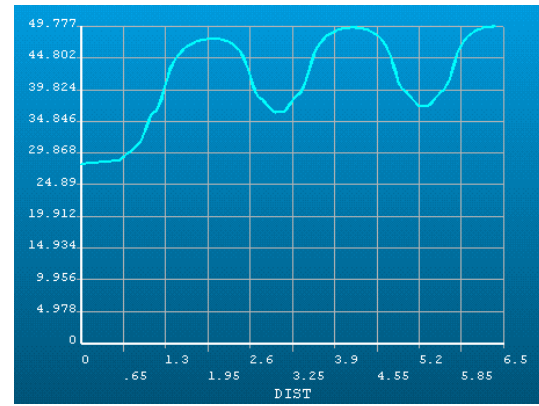
a – 12 cellules de $0.15 \times 0.49 \mu\text{m}^2$.b – 5 cellules de $0.15 \times 1.17 \mu\text{m}^2$.

FIG. V.13 – Simulations 3D sous Ansys donnant la température en fonction de la position dans le composant, pour deux fractionnements d'émetteur différents.

	$12 \times (0.15 \times 0.49)$	$7 \times (0.15 \times 0.83)$	$5 \times (0.15 \times 1.17)$
R_{Th} [K/W]	820	1062	1196
f_T [GHz]	195	219	227
f_{MAX} [GHz]	231	235	236
C_{BE} [fF]	31.4	23.9	22.6
C_{BC} [fF]	25.7	20.2	18.1
$R_{BB'}$ [Ω]	8	14.8	17.8

TAB. V.8 – Résistances thermiques et principaux paramètres statiques et dynamiques mesurés en fonction du fractionnement de l'émetteur.

bénéficier de composants passifs de haute qualité dans une technologie BiCMOS [Avenier05]. L'auto-échauffement est beaucoup plus important sur substrat SOI, du fait de la très faible conductivité thermique de l'oxyde de silicium par rapport au silicium. Le tableau V.9 compare les valeurs de R_{Th} extraites sur des transistors d'architecture faible coût, réalisés sur substrat massif ou SOI.

	$12 \times (0.15 \times 0.49)$	$7 \times (0.15 \times 0.83)$	$5 \times (0.15 \times 1.17)$
R_{Th} bulk [K/W]	820	1062	1196
R_{Th} SOI [K/W]	4434	5847	6678

TAB. V.9 – Résistances thermiques extraites sur substrat massif ou SOI

Les transistors bipolaires sur SOI ont des densités de courant plus faibles dues à un profil émetteur/base différent, donc les densités de puissance thermique sont plus faibles que leurs équivalents sur substrat massif. Leur température de fonctionnement est largement supérieure, comme le montre la figure V.14. Cette augmentation de température est une des principales causes de diminution des performances dynamiques sur substrat SOI, les effets de haute injection intervenant à plus faible polarisation. La densité de courant $J_{C,Opt}$ est donc plus faible sur SOI que sur substrat massif.

L'auto-échauffement devient donc un paramètre critique pour le fonctionnement des TBH sur substrat SOI mince. Cependant, des transistors bipolaires ayant des fréquences f_{MAX} de 200 GHz ont été démontrés sur substrat SOI [Chantre06].

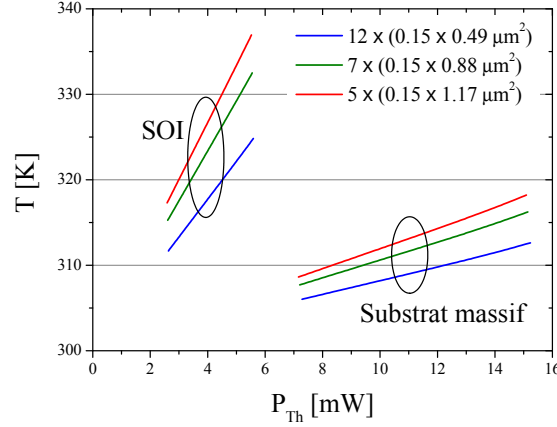


FIG. V.14 – Comparaison entre substrat massif et SOI des températures moyennes en fonction de la puissance thermique appliquée, pour plusieurs fractionnements d'émetteur.

V.2.3.e Élaboration d'un modèle simple

Afin de comprendre plus précisément la dépendance de la résistance thermique en fonction de la géométrie du transistor, un modèle analytique a pu être développé sous Mathcad. Il s'agit d'un calcul direct de R_{Th} , et non d'une méthode reposant sur la résolution de l'équation de la chaleur. De plus, nous nous intéresserons uniquement aux émetteur mono-doigts, le couplage entre les doigts ne pouvant être pris en compte simplement avec ce type d'outil.

Nous considérons que le flux de chaleur se propage uniquement vers le bas (pas de diffusion latérale). Dans ce cas, une première approximation est que le flux de chaleur est contenu uniquement dans un cône d'angle au sommet 45° [Rieh02]. La résistance thermique d'un barreau de silicium de longueur L et de surface A vaut :

$$R_{Th} = \frac{\rho L}{A} \quad (V.3)$$

où ρ est la résistivité thermique du silicium et vaut $6757 \text{ W}\cdot\mu\text{m}/\text{K}$ à 300 K . Dans notre cas, la surface A évolue en fonction de la profondeur z . R_{Th} s'écrit alors :

$$R_{Th} = \int_0^{+\infty} \frac{\rho}{A(z)} dz \quad (V.4)$$

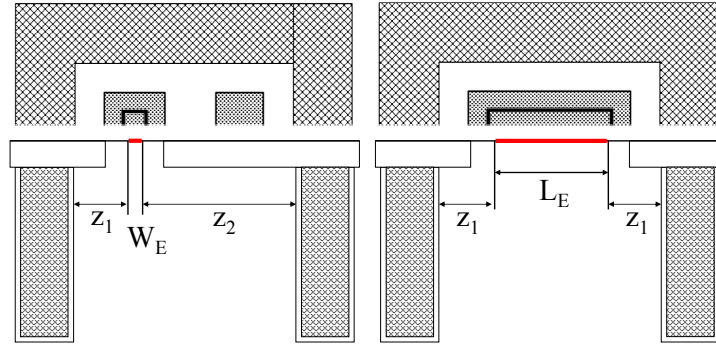
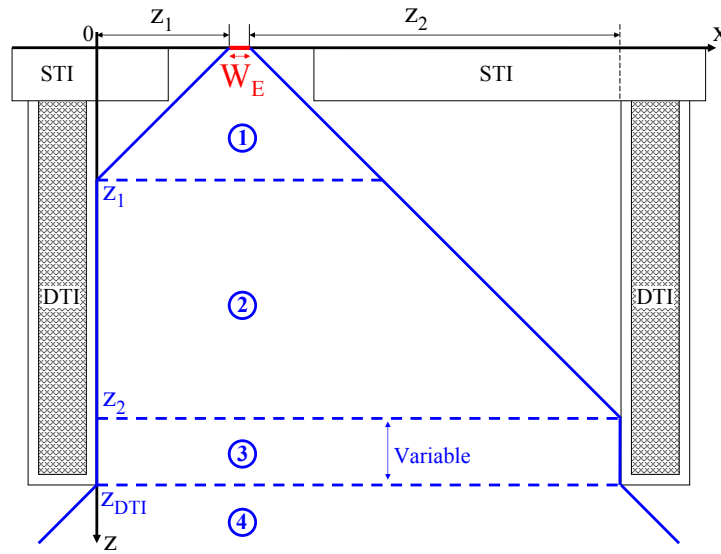
L'épaisseur du substrat est de l'ordre de $300 \mu\text{m}$, ce qui est très grand devant la dimension caractéristique du composant. On ne commet donc pas une grosse erreur en intégrant jusqu'à $+\infty$ au lieu de $300 \mu\text{m}$.

Le modèle intègre un paramétrage en fonction de la longueur et de la largeur de l'émetteur (zone de chauffe). Il faut donc déterminer $A(z, L_E, W_E)$ pour chaque profondeur z , en tenant compte du fait que les DTI se comportent comme des parois adiabatiques. En effet, la conductivité thermique de l'oxyde est très faible comparée à celle du silicium (1.4 contre $131 \text{ W}/\text{m}\cdot\text{K}$). La figure V.15 représente une vue en coupe du transistor bipolaire et des tranchées profondes. La structure modélisée a une géométrie de type BEBC.

Compte tenu des différentes distances entre la partie active et les tranchées profondes, on peut diviser la valeur de la résistance thermique en 4 contributions successives, comme indiqué sur la figure V.16.

La résistance thermique s'écrit :

$$R_{Th} = R_{Th1} + R_{Th2} + R_{Th3} + R_{Th4} \quad (V.5)$$

FIG. V.15 – Vue en coupe du transistor bipolaire. $z_1 = 0.7 \mu\text{m}$ et $z_2 = 2 \mu\text{m}$.FIG. V.16 – Répartition du flux thermique dans le substrat. Les DTI sont considérés comme des parois adiabatiques. R_{Th} est divisé en 4 parties distinctes.

R_{Th1} est une contribution conique jusqu'à la profondeur z_1 où le flux thermique rencontre le DTI. Elle ne dépend que de la longueur et la largeur d'émetteur. L'expression de R_{Th1} est :

$$R_{Th1} = \int_0^{z_1} \frac{\rho}{(W_E + 2z)(L_E + 2z)} dz \quad (\text{V.6})$$

R_{Th2} s'exprime en fonction de z_1 de la manière suivante :

$$R_{Th2} = \int_{z_1}^{z_2} \frac{\rho}{(W_E + z_1 + z)(L_E + 2z_1)} dz \quad (\text{V.7})$$

La résistance R_{Th3} est une contribution cylindrique entre z_2 et le bout des tranchées profondes. Si on appelle z_{DTI} la profondeur des tranchées, on en déduit l'expression de R_{Th3} :

$$R_{Th3} = \frac{\rho(z_{DTI} - z_2)}{(W_E + z_1 + z_2)(L_E + 2z_1)} \quad (\text{V.8})$$

La dernière contribution à la résistance thermique est la contribution du substrat sous les DTI. Cette contribution R_{Th4} s'écrit simplement :

$$R_{Th4} = \int_0^{+\infty} \frac{\rho}{(W_E + z_1 + z_2 + 2z)(L_E + 2z_1 + 2z)} dz \quad (\text{V.9})$$

Dans le cas sans DTI, R_{Th} s'exprime de la manière suivante :

$$R_{Th} = \int_0^{+\infty} \frac{\rho}{(W_E + 2z)(L_E + 2z)} dz \quad (V.10)$$

La figure V.17 compare les valeurs de R_{Th} modélisées et mesurées pour plusieurs longueurs d'émetteur. La résistance thermique globale ainsi calculée permet de modéliser correctement les valeurs de R_{Th} mesurées expérimentalement.

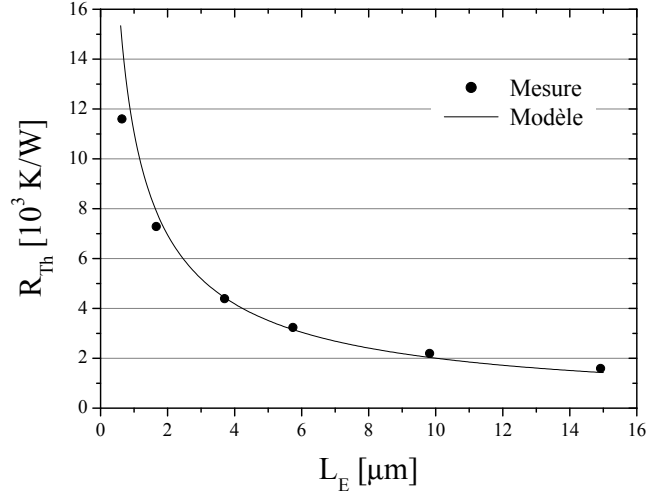


FIG. V.17 – Comparaison des valeurs de R_{Th} modélisées et mesurées en fonction de la longueur d'émetteur (Largeur d'émetteur : $0.15 \mu m$, profondeur des DTI = $5.2 \mu m$).

Ce modèle permet également de reproduire l'évolution de R_{Th} avec la profondeur des tranchées. On ne joue que sur la valeur de la contribution R_{Th3} . La figure V.18 représente les valeurs modélisées de R_{Th} en fonction de la longueur d'émetteur pour les différentes profondeurs de tranchées décrites au paragraphe V.2.3.b.

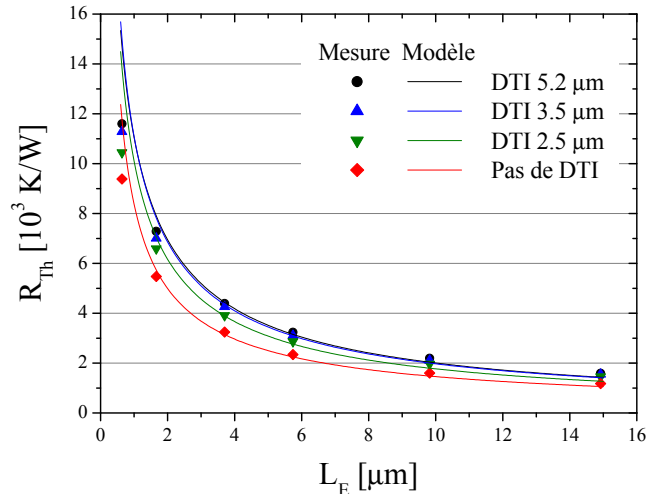


FIG. V.18 – Comparaison des valeurs de R_{Th} modélisées et mesurées en fonction de la longueur d'émetteur (Largeur d'émetteur : $0.15 \mu m$).

Le modèle permet aussi d'illustrer l'influence de la largeur effective d'émetteur sur l'auto-échauffement du composant. La figure V.19 représente les valeurs de R_{Th} et de $R_{Th} \times A_E$ simulées pour plusieurs largeurs d'émetteur W_E comprises entre $0.14 \mu m$ et $0.12 \mu m$. On voit

que la valeur de R_{Th} est quasiment indépendante de la largeur d'émetteur, cependant le produit $R_{Th} \times A_E$ est réduit de manière significative lorsque W_E diminue : L'auto-échauffement de transistors à faibles largeurs d'émetteur est réduit, non pas grâce à une amélioration de l'évacuation de la chaleur, mais car la surface générant la chaleur est réduite. Le fait que R_{Th} varie peu avec W_E est dû au fait que le volume de silicium disponible pour l'évacuation de la chaleur est identique quelle que soit la largeur de l'émetteur.

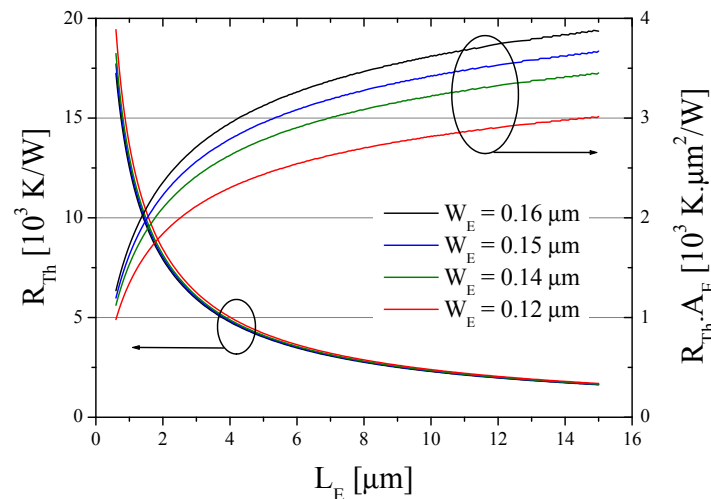


FIG. V.19 – Valeurs de R_{Th} et de $R_{Th} \times A_E$ pour plusieurs largeurs d'émetteur en fonction de L_E .

En conclusion, ce modèle simple permet grâce à un calcul direct de prédire la résistance thermique du transistor en fonction de sa géométrie. Il permet de comparer uniquement les comportements thermiques de composants différant uniquement par leur géométrie. En particulier, si les densités de courant évoluent, le modèle ne sera pas capable de prédire l'augmentation de température. Il permet juste de modéliser l'évacuation de la chaleur générée. Une des améliorations que l'on peut apporter au modèle est la prise en compte d'une évacuation de chaleur par les contacts au dessus du composant. En effet, si la chaleur est principalement évacuée par le substrat, les métaux des interconnexions sont de bons conducteurs thermiques et contribuent en partie à diminuer la valeur de R_{Th} : C'est ce qui explique que dans le cas de composants ayant de très courtes longueur d'émetteur la valeur de R_{Th} modélisée est surestimée par rapport à la valeur extraite (figure V.17). Le contact métallique au dessus de l'émetteur contribue à diminuer la résistance thermique du composant [Rieh02].

Une autre amélioration pourrait être de considérer des composants multi-doigts, en utilisant l'approche retenue dans [Rieh05a]. Cependant, l'effet de couplage thermique entre les doigts ne peut être modélisé, ce qui est une source d'erreur. L'auto-échauffement d'un doigt d'émetteur n'est en effet pas dû uniquement à son propre fonctionnement, mais à l'influence de ses voisins.

V.3 Étude du TBH à basse température

Si les performances du transistor bipolaire sont dégradées par l'auto-échauffement inhérent à son fonctionnement, ou par l'échauffement causé par son environnement, on peut exalter ses performances dynamiques en analysant les composants à basses températures. La gamme de température explorée varie entre 50 K et 300 K, grâce à un cryostat à circulation d'hélium adapté aux mesures hyperfréquences. Nous étudions dans un premier temps l'influence des températures cryogéniques sur les paramètres statiques, puis nous abordons le domaine hyperfréquence à basse température. L'analyse en température sera menée sur différents composants présentant des performances à température ambiante supérieures à 250 GHz, mais ayant des réglages différents de f_T et f_{MAX} .

Le premier composant est le composant standard du lot ayant servi à étudier la recombinaison en base neutre (chapitre IV.4) : Il s'agit d'une base de 24 nm avec un profil de Ge 20–25%, et le Si-cap a une épaisseur de 14 nm. La base étant plus épaisse, la fréquence f_{MAX} est favorisée par rapport à f_T . Nous l'appellerons composant 1.

Les trois autres composants sont les composants sur lesquels ont été faites les études sur la profondeur des tranchées d'isolation. Il s'agit de composants ayant une base 20–30% de Ge d'épaisseur 20 nm, avec un Si-cap de 16 nm. Dans ces 3 composants, la fréquence f_T est favorisée par rapport à f_{MAX} . Nous les dénommons composants 2A à 2C.

Tous les transistors étudiés sont de type CBEB, avec une surface d'émetteur active de $0.15 \times 3.6 \mu\text{m}^2$, et un contact d'émetteur en ruban pour diminuer la résistance d'émetteur. Le tableau V.10 résume l'appellation des composants dans la suite de ce chapitre :

TBH standard NBR	Composant 1
DTI = 5.2 μm	Composant 2A
DTI = 2.5 μm	Composant 2B
Pas de DTI	Composant 2C

TAB. V.10 – Dénomination des composants étudiés en hyperfréquence à des températures cryogéniques.

V.3.1 Modification des paramètres statiques à basse température

V.3.1.a Evolution des niveaux de courant

D'après I.9, l'évolution de la concentration intrinsèque des porteurs avec la température $n_i^2(T)$ s'exprime en fonction de $n_{i0}^2 = n_i^2(T_0)$ où $T_0 = 300 \text{ K}$:

$$n_i^2(T) = n_{i0}^2 \left(\frac{T}{T_0} \right)^3 \exp \frac{-2 E_G}{kT} \exp \frac{2 E_G}{kT_0} \quad (\text{V.11})$$

D'après les équations I.25 et V.11, on en déduit l'évolution du nombre de Gummel de la base avec la température :

$$G_B(T) = G_B(T_0) \frac{\mu_n(T_0)}{\mu_n(T)} \left(\frac{T_0}{T} \right)^4 \exp \frac{2 E_G}{kT} \exp \frac{-2 E_G}{kT_0} \quad (\text{V.12})$$

La mobilité des électrons dans le SiGe contraint est fortement améliorée à basse température, et est modélisée empiriquement par la formule donnée dans [Rieh00] :

$$\mu_n(T) = A.T^{-m} \text{ [cm}^2\text{/V.s]} \quad (\text{V.13})$$

où A et m sont des paramètres qui dépendent du taux de germanium. m est compris entre 0.44 et 0.68, pour des taux de Ge variant entre 20 et 30 %. On en déduit l'expression du courant collecteur en fonction de la température et de V_{BE} :

$$I_C(V_{BE}, T) = \frac{q A_e}{G_B(T_0)} \exp \frac{2 E_G}{kT_0} \left(\frac{T}{T_0} \right)^{4-m} \exp \frac{q V_{BE} - 2 E_G}{kT} \quad (\text{V.14})$$

D'après l'équation V.14, le courant collecteur diminue fortement avec la température lorsqu'on se place à V_{BE} constant. Cependant, les effets de résistance séries et de haute injection (effet Webster, effet Kirk...) sont avant tout dépendants de la densité de courant qui circule dans le transistor, et non pas de la tension V_{BE} . Pour l'analyse statique du transistor, il faut donc plutôt raisonner à courant constant. Ainsi, la transconductance du composant est améliorée lorsque la température diminue, pour un même courant I_C .

$$g_m = \frac{dI_C}{dV_{BE}} = \frac{qI_C}{kT} \quad (\text{V.15})$$

Lorsque I_C est fixé, la diminution de température de 300 K à 50 K améliore donc la transconductance d'un facteur 6. De même que pour le courant collecteur, le courant de base diminue fortement lorsque la température diminue. Ceci explique l'évolution des courbes de Gummel avec la température présentées sur la figure V.20.

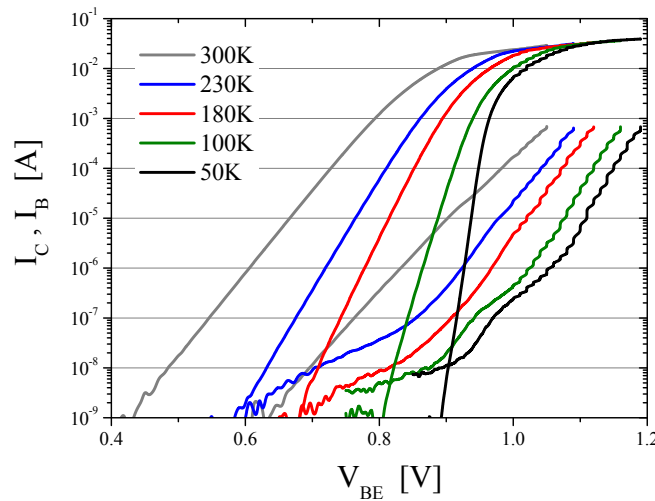


FIG. V.20 – Evolution des courbes de Gummel en fonction de la température (Composant 1).

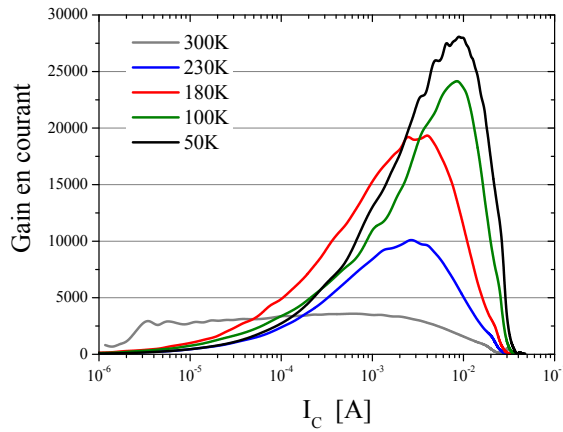
La pente des courants I_C et I_B en fonction de V_{BE} est plus forte, ce qui est dû à l'amélioration de la transconductance. On note cependant des non-idéalités de courant de base. Le courant thermique étant plus faible, les courants non idéaux correspondant à des mécanismes indépendants de la température (recombinaison, effet tunnel...) sont visibles aux températures cryogéniques, alors qu'ils ne le sont pas à température ambiante.

V.3.1.b Variation du gain en courant avec la température

Le gain dépend également de la température, selon l'équation suivante (on ne tient pas compte du BGN) [Ashburn94] :

$$\beta = \frac{\gamma N_{dE} W_E D_{nB}}{N_{aB} W_B D_{pE}} \exp \frac{\Delta E_G}{kT} \quad (\text{V.16})$$

Le gain en courant est donc fortement augmenté lorsque la température diminue. On observe en effet sur la figure V.21–a que le gain maximal varie de 3600 à 300 K à plus de 25000 à 50 K. Cette augmentation du gain en courant est importante, mais elle est limitée par les effets non idéaux qui apparaissent à basse température (un taux de Germanium moyen de 20 % donnerait une augmentation de gain d'un facteur 10^{12} entre 300 et 50 K). Compte-tenu de l'augmentation de la transconductance, le maximum de β intervient pour un V_{BE} plus important lorsque la température diminue (figure V.21–b).



a – Gain en fonction du courant collecteur

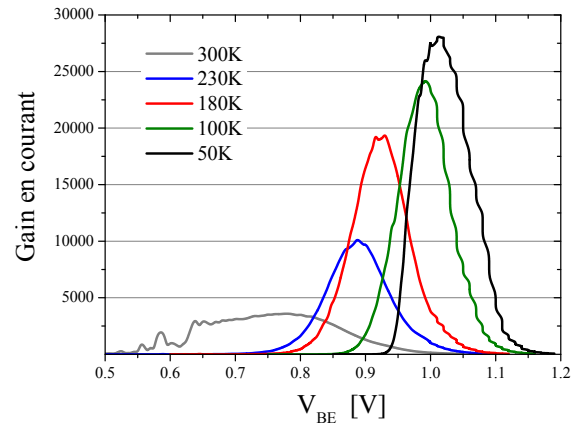
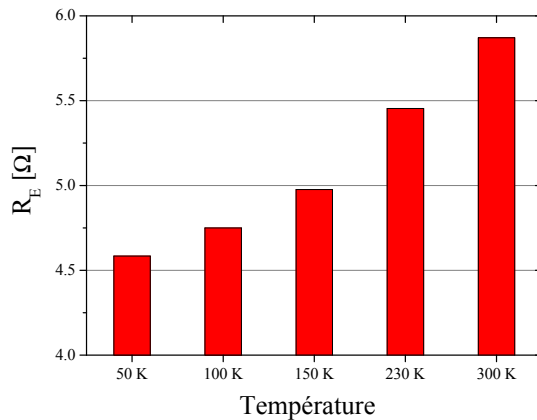
b – Gain en fonction de V_{BE}

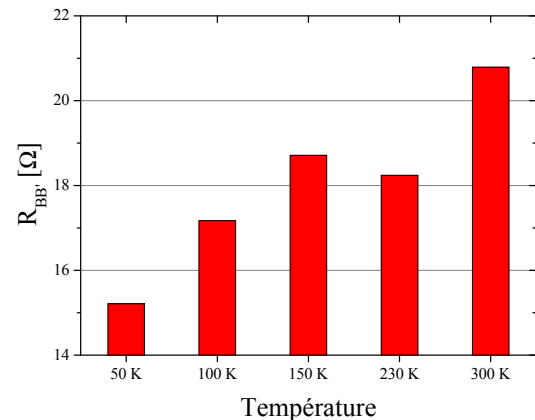
FIG. V.21 – Évolution du gain en courant avec la température (Composant 1).

V.3.1.c Évolution des autres paramètres statiques

La mobilité des porteurs étant améliorée à basse température, les résistances d'accès extrinsèques sont améliorées. Le dopage de la base extrinsèque et de l'émetteur sont très élevés, ce qui permet d'éviter la dégradation de la résistivité des matériaux par le gel des porteurs en dessous de 100 K. Les figures V.22 et V.23 montrent les évolutions des résistances d'émetteur et de base pour les 4 composants étudiés.



a – Résistance d'émetteur



b – Résistance de base

FIG. V.22 – Évolution des résistances séries du composant 1 avec la température.

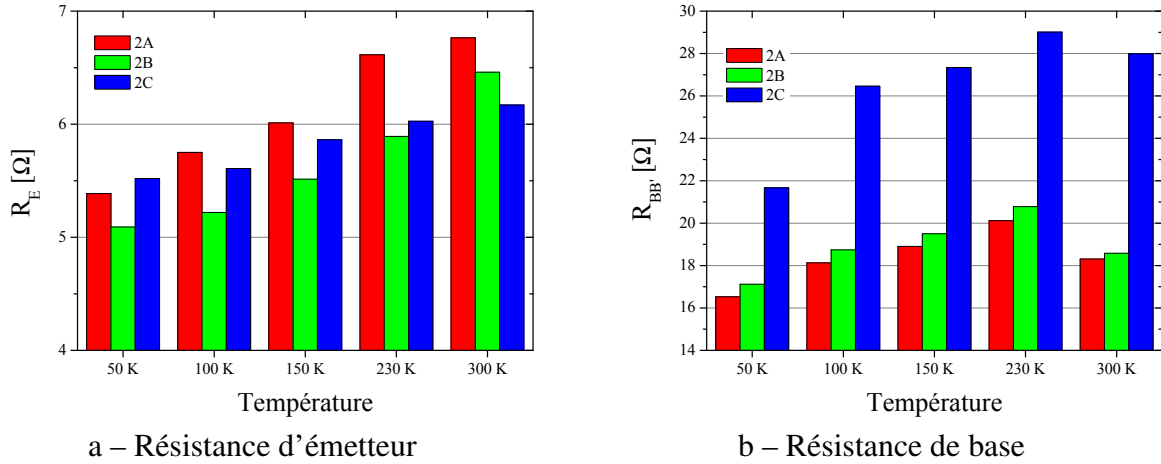


FIG. V.23 – Évolution des résistances séries des composants 2A, 2B et 2C avec la température.

Le composant 1 a une résistance d'émetteur légèrement plus faible, ce qui est dû à un Si-cap plus fin (14 nm contre 16 nm pour les composants 2). Les valeurs de résistance de base sont équivalentes, le composant 1 étant légèrement avantagé par rapport aux autres. Le composant 2C présente une résistance d'émetteur plus faible à 300 K que les composants 2A et 2B, ce qui est due à une jonction E/B plus agressive, causée par des variations des conditions de croissance lors de l'épithaxie sélective de la base, comme indiqué dans le tableau V.3. En conséquence, le composant 2C présente une résistance de base plus importante, à cause d'une base intrinsèque légèrement plus pincée. La variation de R_E avec la diminution de température est de l'ordre de - 20 % entre 300 et 50 K.

V.3.2 Performances dynamiques à basse température

V.3.2.a Fréquences f_T et f_{MAX} aux températures cryogéniques

L'amélioration à basse température des caractéristiques statiques du transistor (transconductance, résistances d'accès) et de la mobilité des porteurs permet une forte augmentation des performances dynamiques du composant. Les équations de f_T et f_{MAX} rappelées ci-dessous (équations V.17 et V.18) montrent quels sont les paramètres qui limitent les fréquences de coupure du composant :

$$f_T = \frac{1}{2\pi \left(\tau_F + \frac{1}{g_m} (C_{BC} + C_{BE}) + (R_E + R_C) C_{BC} \right)} \quad (V.17)$$

$$f_{MAX} = \sqrt{\frac{f_T}{8\pi R_B C_{BC}}} \quad (V.18)$$

Ainsi, la fréquence f_T est améliorée à basse température grâce à plusieurs effets conjugués :

- L'augmentation de la transconductance g_m permet de diminuer les temps de charge des capacités C_{BE} et C_{BC} .
- L'augmentation de la mobilité permet de réduire le temps de transit τ_F .
- La diminution des résistances séries permet de réduire les éléments parasites de type résistance-capacité, qui limitent fortement la montée en fréquence.

La fréquence f_{MAX} augmente également à basse température, sous l'impulsion de l'augmentation de f_T . Cependant, selon les valeurs de $R_B C_{BC}$, l'amplitude de variation de f_{MAX}

varie, comme nous le verrons par la suite. Une résistance de base faible favorise la fréquence f_{MAX} à température ambiante, et cet effet est encore accentué aux températures cryogéniques.

Les figures V.24 à V.27 représentent les évolutions des fréquences f_T et f_{MAX} des 4 composants testés entre 300 K et 50 K.

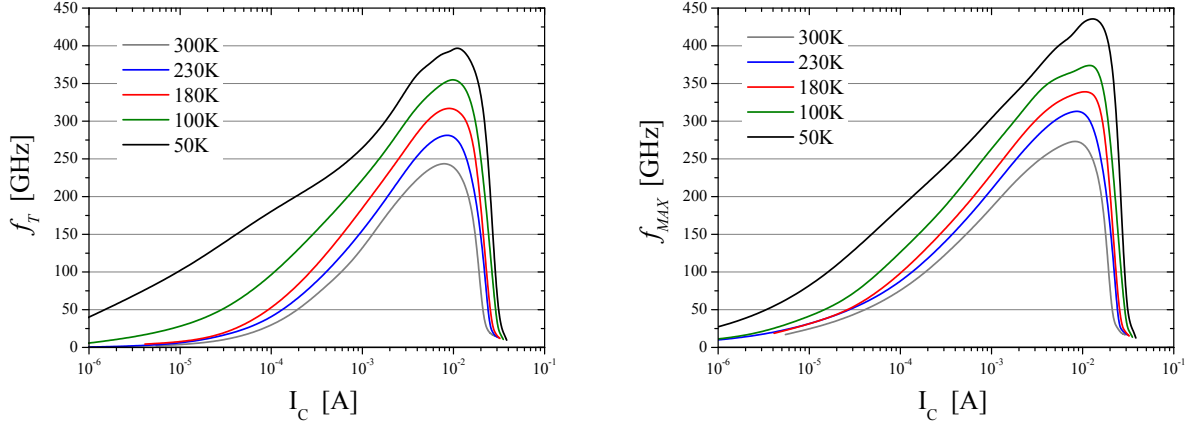


FIG. V.24 – Évolution des fréquences f_T et f_{MAX} du composant 1 avec T ($V_{CB} = 0.5$ V).

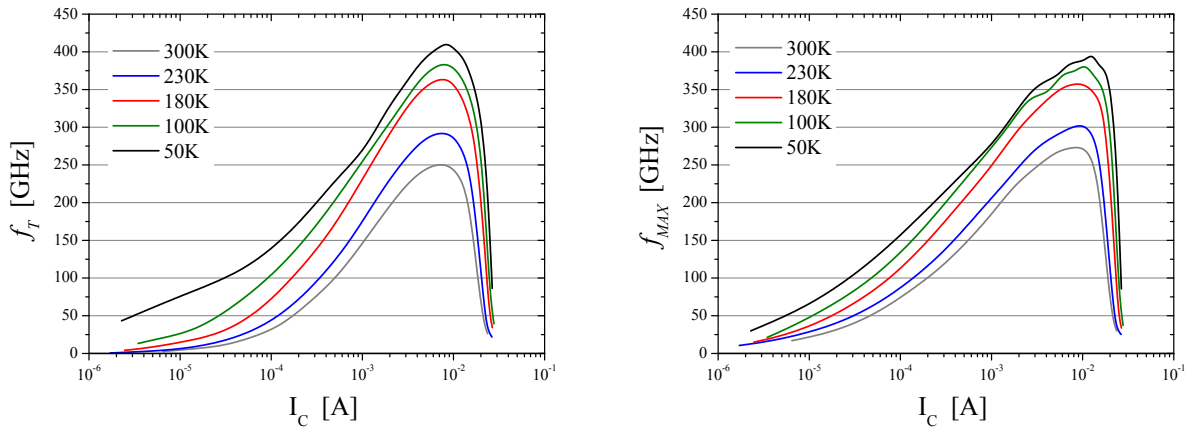


FIG. V.25 – Évolution des fréquences f_T et f_{MAX} du composant 2A avec T ($V_{CB} = 0.5$ V).

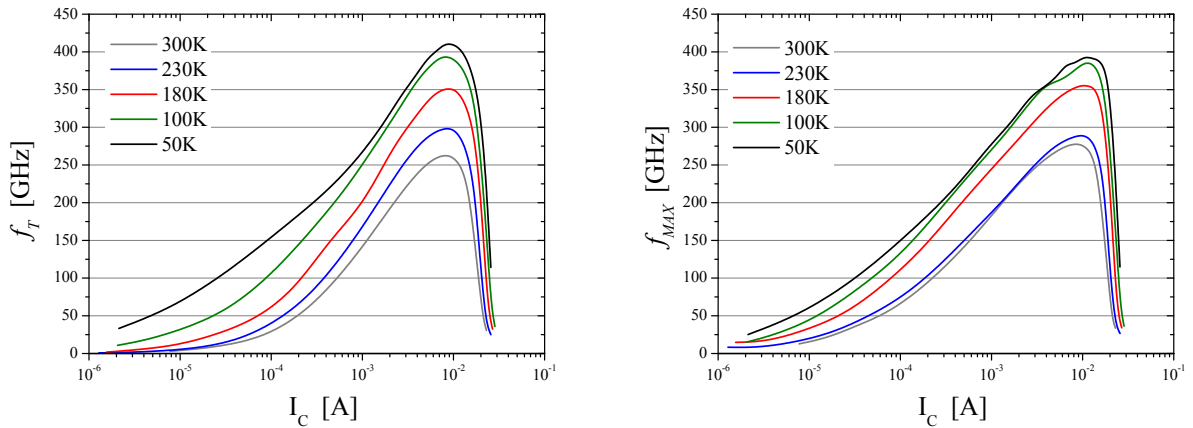
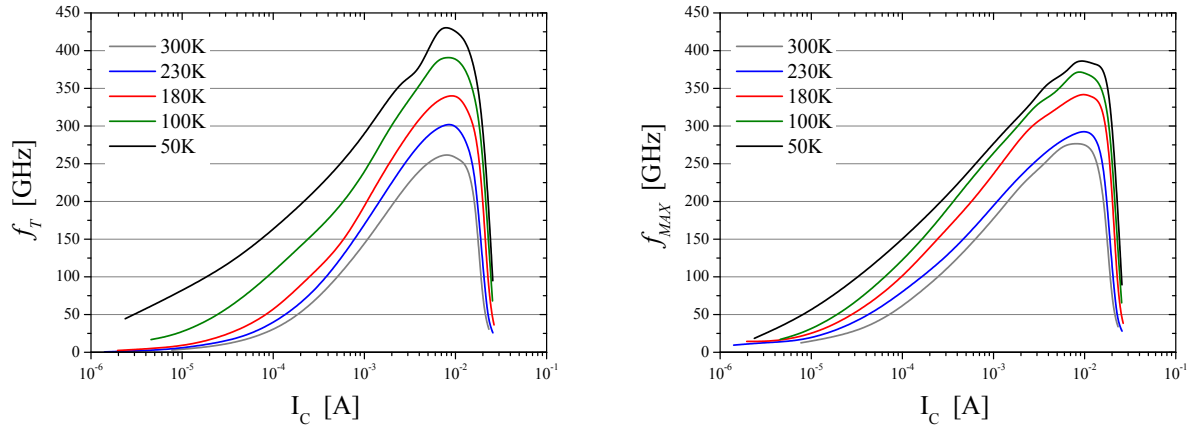


FIG. V.26 – Évolution des fréquences f_T et f_{MAX} du composant 2B avec T ($V_{CB} = 0.5$ V).

FIG. V.27 – Évolution des fréquences f_T et f_{MAX} du composant 2C avec T ($V_{CB} = 0.5$ V).

On voit que l'augmentation de f_T avec la diminution de la température est importante dans les 4 composants. Le tableau V.11 résume les fréquences maximales de f_T obtenues sur les transistors testés en fonction de la température, ainsi que l'augmentation de f_T entre 300 et 50 K. L'augmentation de f_T est identique en proportion pour 3 des 4 composants testés, le composant 2B restant légèrement en retrait.

Composant	1	2A	2B	2C
300 K	250	253	266	268
230 K	288	298	304	306
180 K	321	367	356	344
100 K	359	386	397	393
50 K	405	411	413	435
Δf_T	155	158	147	166
$\Delta f_T(\%)$	62	62	55	62

TAB. V.11 – Fréquences f_T maximales (en GHz) obtenues en fonction de la température. Δf_T est la variation de f_T entre 300 K et 50 K.

Le même type d'analyse est mené sur les fréquences f_{MAX} : Le tableau V.12 résume les valeurs maximales de f_{MAX} obtenues en fonction de la température, ainsi que les variations de f_{MAX} entre 300 et 50 K pour chaque composant testé.

Composant	1	2A	2B	2C
300 K	274	279	282	279
230 K	317	304	290	294
180 K	341	358	356	343
100 K	373	382	386	372
50 K	429	396	394	385
Δf_{MAX}	155	117	112	106
$\Delta f_{MAX}(\%)$	57	42	40	38

TAB. V.12 – Fréquences f_{MAX} maximales (en GHz) obtenues en fonction de la température. Δf_{MAX} est la variation de f_{MAX} entre 300 K et 50 K.

Excepté pour le composant 1, l'augmentation de f_{MAX} est plus modeste que celle de f_T . Les transistors démontrant les plus forts f_T ont en proportion la plus faible augmentation de

f_{MAX} . On constate également que les valeurs maximales de f_{MAX} sont inversées entre 300 et 50 K : Le composant 1, qui est celui qui présente le plus faible f_{MAX} à température ambiante, démontre le plus fort f_{MAX} aux très basses températures. Cette inversion des performances avec la diminution de température signifie que, selon le réglage des paramètres technologiques du transistor, l'une ou l'autre fréquence de coupure sera favorisée. Dans notre cas, le composant 1 dispose d'une base plus épaisse, favorable à la fréquence f_{MAX} . Les transistors 2A, 2B et 2C, dotés d'une base plus agressive, permettent d'atteindre des performances records en f_T , au détriment de l'augmentation de f_{MAX} , qui devient inférieur à f_T à 50 K.

V.3.2.b Étude des temps de transit

Nous pouvons étudier plus en détail cette variation des performances dynamiques à travers l'évolution des temps de transit des porteurs dans la structure. Le temps τ_{EC} est défini par :

$$\tau_{EC} = \frac{1}{2\pi f_T}$$

D'après l'équation V.17, on peut écrire le temps de transit sous la forme suivante :

$$\tau_{EC} = \tau_F + (R_E + R_C) C_{BC} + \frac{kT}{q I_C} (C_{BE} + C_{BC}) \quad (V.19)$$

Ainsi, le tracé de τ_{EC} en fonction de $1/I_C$ permet d'extraire différents paramètres :

- Le temps $\tau_{EC Min}$ est tout simplement égal à $1/2\pi f_{T Max}$. C'est le retard minimal du composant entouré de tous ses éléments parasites.
- Le retard $\tau_{EC,0} = \tau_F + (R_E + R_C) C_{BC}$ est obtenu en prenant l'ordonnée à l'origine de la caractéristique de $\tau_{EC}(1/I_C)$.
- La pente de $\tau_{EC}(1/I_C)$ à moyenne injection permet d'extraire le terme $kT/q(C_{BE} + C_{BC})$, représentatif des capacités de jonction. L'incertitude existe néanmoins sur le facteur d'idéalité du transistor et sur la température effective à l'intérieur du dispositif, différente de la température de la chambre.

Le temps de transit τ_F est le temps de transit de la partie idéale du TBH et peut être calculé par ailleurs lors de l'extraction analytique du transistor bipolaire [Zerounian00]. Il est donc possible de calculer la valeur de $(R_E + R_C) C_{BC}$. Néanmoins, la distinction entre τ_F et $(R_E + R_C) C_{BC}$ est soumise à caution, car dépendante du modèle choisi pour le transistor bipolaire et de l'extraction des éléments extrinsèques.

L'expression de τ_{EC} donnée par l'équation V.19 est incomplète. Il faut en effet tenir compte du retard additionnel causé par l'effet Kirk au pic de f_T , que nous nommerons $\Delta\tau$ [Malorny03]. Même si ce retard $\Delta\tau$ n'est pas très important lorsque f_T est à son maximum, l'effet Kirk est à l'origine de la chute de f_T à forts courants. La figure V.28 représente l'évolution de τ_{EC} en fonction de $1/I_C$ pour le composant 1 à 300 K, ainsi que les méthodes d'extraction des différentes contributions au temps de transit total. Le retard $\Delta\tau$ dû à l'effet Kirk est égal à la différence entre $\tau_{EC Min}$ et la courbe de τ_{EC} théorique à $I_{C Opt.}$.

Grâce à cette méthode, il est donc possible de décomposer le temps de transit τ_{EC} pour les 4 composants étudiés. La figure V.29 représente les différentes contributions du temps de transit en fonction de la température. Le temps $\tau_{EC Min}$ est subdivisé en 4 contributions, dont les proportions varient en fonction de T.

On constate que la réduction de $\tau_{EC Min}$ avec la température est due à une diminution de toutes les contributions. Le temps de transit total est réduit de 0.6 ps à température ambiante à moins de 0.4 ps à 50 K.

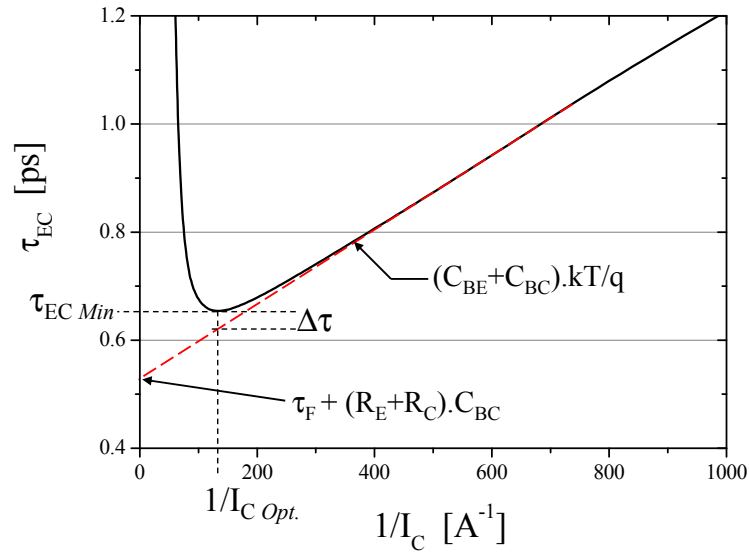
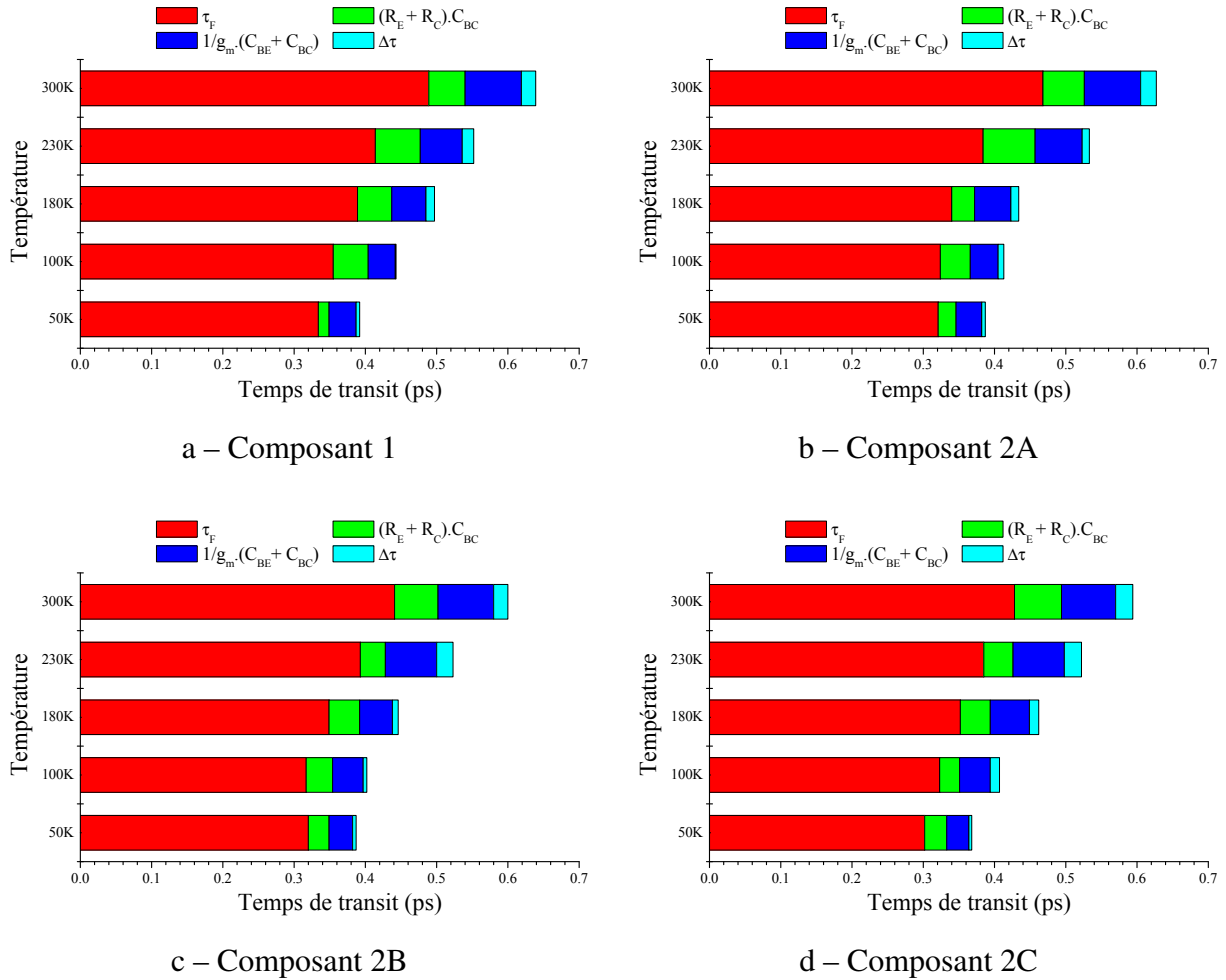
FIG. V.28 – Évolution de τ_{EC} en fonction de $1/I_C$ (composant 1, 300 K, $V_{CB} = 0.5$ V).

FIG. V.29 – Répartitions des différentes contributions au temps de transit en fonction la température.

Les effets parasites $(R_E + R_C) C_{BC}$ et $1/g_m(C_{BE} + C_{BC})$ sont également améliorés à basse température : La mobilité est augmentée à basse température, ce qui permet de diminuer

les résistances d'accès, et la transconductance est fortement augmentée lorsque la température diminue, ce qui atténue l'effet des capacités de jonction. En proportion, la contribution des effets parasites diminue à basse température. On constate en effet que le temps τ_F représente un pourcentage croissant de τ_{EC} lorsque la température diminue (75 % en moyenne à 300 K, plus de 80 % à 50 K). Le détail de l'évolution de $\tau_{EC Min}$ et τ_F est donné dans le tableau V.13.

		C1	C2A	C2B	C2C
300 K	$\tau_{EC Min}$ [ps]	0.639	0.628	0.599	0.594
	τ_F [ps]	0.489	0.468	0.441	0.428
	$\tau_F/\tau_{EC Min}$	76.6 %	74.5 %	73.6 %	71.9 %
50 K	$\tau_{EC Min}$ [ps]	0.392	0.386	0.387	0.367
	τ_F [ps]	0.334	0.321	0.320	0.302
	$\tau_F/\tau_{EC Min}$	85.1 %	83.0 %	82.6 %	82.3 %

TAB. V.13 – Comparaison des temps de transit $\tau_{EC Min}$ et τ_F à 300 K et 50 K.

A température ambiante comme aux températures cryogéniques, le temps de transit des porteurs τ_F représente la plus grosse contribution à τ_{EC} . L'amélioration du transport à basses température permet de diminuer τ_F jusqu'à 0.3 ps (composant 2C), correspondant à une fréquence de coupure de la partie idéale du transistor supérieure à 520 GHz.

On remarque grâce au tableau V.13 que les dispositifs les plus performants sont ceux pour lesquels la contribution de τ_F au temps de transit total est la plus faible. Les transistors 2A, 2B et 2C ont un profil vertical très agressif, permettant d'obtenir de très faibles temps de transit, les effets parasites ont donc plus d'importance.

V.3.3 Synthèse de l'étude cryogénique

L'analyse en température donne une vision des possibilités ultérieures de l'optimisation du transistor bipolaire. La principale limitation de la montée en fréquence est le temps de transit intrinsèque du transistor, qui compte pour plus de 75 % du temps de transit total, même à très basse température. Il est donc en théorie possible avec la structure actuelle d'obtenir des f_T de 400 GHz à 300 K en jouant sur le profil vertical du composant, les éléments parasites n'apparaissant pas comme la principale limitation. Toutefois, l'optimisation conjointe de f_T et f_{MAX} est plus délicate, car l'amélioration de τ_F nécessite une base très fine, ce qui augmente la résistance de base et pénalise f_{MAX} , comme le prouvent les évolutions des performances des composants 1 et 2C avec la température : L'une ou l'autre fréquence de coupure est favorisée.

[Krithivasan06] présente un transistor bipolaire ayant un f_T record de 510 GHz aux températures cryogéniques, pour un temps de transit τ_F de 0.27 ps, obtenu avec un profil de base très agressif (on obtient 352 GHz pour 0.36 ps à 300 K). f_{MAX} évolue peu : de 241 GHz à 300 K, on atteint 276 GHz à basse température. Il est à noter que dans ce cas, f_{MAX} ne bénéficie pas de l'augmentation de f_T avec la diminution de la température, montrant l'importance de l'optimisation globale du composant pour obtenir un bon équilibre entre f_T et f_{MAX} .

Les mesures hyperfréquences permettent d'obtenir la valeur τ_F globale, mais ne permettent pas de faire la décomposition entre τ_E , τ_{EB} , τ_B et τ_{BC} . Un outil de simulation et de modélisation est nécessaire pour arriver à faire cette décomposition de manière précise. Néanmoins, les équations I.73 et I.77 indiquent que τ_E et τ_{EB} diminuent fortement lorsque la température diminue. La tendance actuelle consistant en augmenter le gain en courant du transistor accentue cet effet.

Lorsqu'on est à très basses températures, il est donc possible de considérer la composante $\tau_E + \tau_{EB}$ comme négligeable devant τ_B et τ_{BC} , car le gain en courant est très important. Le

temps de transit mesuré à 50 K de 0.3 ps serait donc imputable uniquement à $\tau_B + \tau_{BC}$. Le partage est ensuite délicat entre les contributions de la base et du collecteur. [Zerounian03] considère que la principale amélioration de τ_F à basse température est due à la diminution des contributions de l'émetteur et de la base, mais cette diminution est difficile à quantifier. L'influence de l'émetteur sur le temps de transit n'est pas négligeable à 300 K, car le gain en courant diminue fortement lorsqu'on atteint le maximum de f_T .

La principale piste pour réduire τ_B est d'optimiser le profil vertical de la base. Il est nécessaire de réaliser des bases très fines avec un gradient de Ge important et un profil dopage abrupt. Concernant la réduction du temps de transit collecteur, les solutions sont plus difficiles à identifier. Les récentes avancées des transistors bipolaires à hétérojonction SiGe, dont une partie a été présentée dans ce travail de thèse, montrent que l'augmentation des densités de courant, couplée avec un recul du seuil d'effet Kirk, est le moyen le plus efficace pour améliorer les performances dynamiques. Cependant, une forte réduction de τ_B et τ_{BC} va entraîner une augmentation de la contribution relative de τ_E et τ_{EB} au temps de transit total. Si la base et le collecteur sont à l'heure actuelle les facteurs limitant la montée en fréquence, l'influence de l'émetteur et de l'efficacité d'injection de la jonction émetteur base ne doivent pas être négligées pour l'optimisation future du composant.

V.4 Conclusion / Perspectives

Ce chapitre fait la synthèse de l'influence de la température sur le fonctionnement du transistor bipolaire : D'un côté l'auto-échauffement apparaît comme une limitation somme toute assez faible des performances du composant. Le fait de réduire la largeur d'émetteur permet de diminuer l'élévation de température à densité de chaleur constante, la résistance thermique étant plus faible, et l'amélioration des performances dynamiques est obtenue plus par la diminution de la résistance de base que par la réduction de l'auto-échauffement. Une amélioration peut également être obtenue en réduisant la profondeur des tranchées d'isolation, mais celles-ci ne peuvent être totalement supprimées pour l'utilisation du composant dans un circuit.

D'un autre côté, l'analyse en température permet de mettre en évidence les limitations principales de la montée en fréquence : Le temps de transit intrinsèque représente encore plus de 75 % du temps de transit total. L'amélioration des performances du composant passe donc par un profil vertical plus agressif, des densités de courant accrues et une réduction des effets parasites. Il est donc nécessaire de réduire toutes les dimensions du composant.

Cependant, si les densités de courant augmentent pour atteindre de fortes valeurs de f_T , l'auto-échauffement sera dégradé. L'analyse menée au paragraphe V.2 montre que si l'auto-échauffement est une limitation de faible envergure à l'heure actuelle, son influence sera grandissante avec la montée en fréquence. Il est possible d'augmenter fortement les performances de la structure FSA-SEG actuelle, mais tôt ou tard, l'auto-échauffement apparaîtra comme une limitation majeure.

Conclusion générale

Le travail de thèse que nous venons de présenter a porté sur l'étude et la réalisation de transistors bipolaires très rapides. Ce travail, réalisé pour sa grande partie au sein de STMicroelectronics, a été mené en amont de l'assemblage d'une nouvelle filière BiCMOS. Pour mener à bien cette étude, nous nous sommes appuyés sur les moyens disponibles au sein de ST et de l'Institut d'Electronique Fondamentale de l'Université d'Orsay.

Notre travail a eu pour but d'identifier, au sein de la structure de transistor bipolaire qui a été retenue, les éléments pouvant être optimisés pour améliorer les performances des composants. Ainsi, l'étude des différentes architectures possibles, menée au chapitre II, a démontré que seule une structure complètement auto-alignée permettait de réduire suffisamment les effets parasites pour atteindre des fréquences f_{MAX} supérieures à 300 GHz. Le choix d'une épitaxie sélective de la base s'est fait notamment grâce aux progrès réalisés sur les matériaux ; l'incorporation de carbone dans un dépôt sélectif n'est en effet plus considéré comme problématique [Brossard06].

Le choix de la structure étant effectué, l'optimisation conventionnelle du composant a démontré la possibilité d'améliorer fortement f_T en proposant un profil vertical plus fin et plus abrupt. Ainsi, nous avons débouché sur un dopage d'émetteur augmenté, avec un Si-cap réduit. Cette étude a permis de réduire fortement la résistance d'émetteur de $5 \Omega \cdot \mu m^2$ à $1.6 \Omega \cdot \mu m^2$, et de diminuer le temps de transit des porteurs. Concernant la base, l'épaisseur de SiGe a été réduite de 30 à 20 nm, et le taux de Ge à l'entrée de la base doit être maintenu autour de 20 % pour obtenir les plus fortes valeurs de f_T .

Une base présentant un taux de germanium plus faible (10–25 %), et donc des niveaux de courant collecteur réduits par rapport à une base 20–30 %, a également été présentée. Ses performances fréquentielles sont plus limitées mais elle offre un meilleur compromis f_T / BV_{CEO} puisqu'une tension BV_{CEO} de 1.6 V ($BV_{CBO} = 5.5$ V) a été obtenue pour une fréquence f_T de 230 GHz. Le profil du collecteur a également été amélioré grâce notamment à un SIC implanté en Arsenic.

Dans le même temps, la réduction des dimensions latérales du composant (largeur d'émetteur, recouvrement émetteur/base et base/collecteur) a permis d'augmenter f_{MAX} de manière significative. Les fréquences maximales d'oscillation ont augmenté de 180 GHz jusqu'à 300 GHz, ce qui permet d'utiliser ce transistor bipolaire pour les applications utilisant de fortes fréquences. La technologie ainsi développée a été validée par la réalisation de blocs de circuits à l'état de l'art, fonctionnant à des fréquences de 100 GHz (VCOs, diviseurs) [Nicolson06, Laskin06], et ce travail d'optimisation conventionnel a permis d'obtenir les spécifications visées pour la filière BiCMOS9MW, qui sont 230 / 280 GHz pour f_T / f_{MAX} et $BV_{CEO} = 1.6$ V. Le travail d'intégration de cette filière, qui vise les applications millimétriques, a débuté cette année.

La partie traitant de l'optimisation classique du composant a démontré l'existence d'un produit $f_T \times BV_{CEO}$ difficile à améliorer, compte tenu des contraintes d'intégration existant dans une filière BiCMOS. Parallèlement à cette étude, nous avons cherché à augmenter la tension de claquage BV_{CEO} par trois procédés technologiques différents : L'émetteur métallique, l'in-

sersion de Ge dans l'émetteur et la recombinaison en base neutre ont permis d'atteindre des tensions de claquage de 2 V.

Ainsi un procédé robuste de réalisation de transistors bipolaires à émetteur métallique a été mis au point. La principale limitation de cette approche est le compromis entre l'augmentation de la résistance d'émetteur et l'effet d'augmentation du courant de base par la réduction de l'épaisseur d'émetteur. Par ailleurs le bruit BF n'est pas dégradé par le procédé émetteur métallique. Cette étude nous a apporté une bien meilleure connaissance du matériau composant l'émetteur, et une meilleure compréhension de la jonction E/B de nos composants.

La méthode par recombinaison en base neutre est prometteuse, notamment pour les applications qui ne sont pas sensibles au bruit BF. Elle ouvre notamment des perspectives intéressantes pour les applications d'amplificateur de puissance, car le gain est stable dans une large gamme de température.

Enfin, l'émetteur SiGe est une technique offrant de larges perspectives d'améliorations, et les études sur ce procédé continuent au sein de ST.

Le dernier chapitre de cette thèse a pour objet l'étude du TBH en fonction de la température. L'auto-échauffement apparaît comme une limitation des performances du transistor, et réduire la surface active du transistor est le meilleur moyen pour diminuer le rapport entre la densité de chaleur générée et la dissipation de cette chaleur. Une méthode d'extraction simple de la résistance thermique R_{Th} a été développée, ainsi qu'un modèle mathématique reproduisant l'évolution de la valeur de R_{Th} en fonction de la géométrie de transistors bipolaires à un seul doigt d'émetteur. La tendance actuelle d'augmenter les densités de courant pour atteindre de fortes fréquences de transition ne joue pas en faveur d'une réduction de l'auto-échauffement.

L'étude du fonctionnement hyperfréquence à des températures cryogéniques a révélé une forte amélioration des performances lorsque la température diminue, grâce à une amélioration du transport. La structure FSA-SEG présentant peu d'effets parasites, des fréquences de 500 GHz ont récemment été atteintes à 50 K sur les composants étudiés dans cette thèse. Il est donc en théorie possible d'optimiser encore plus la structure de transistor actuelle pour atteindre un tel niveau de performances à température ambiante. De fait, le principal retard existant est le temps de transit τ_F , les retards dus aux accès représentent seulement entre 20 et 30 % du temps de transit total $\tau_{EC,0}$.

Cette structure ainsi optimisée présente d'excellentes performances, propices à un emploi du transistor bipolaire pour des applications autour de 100 GHz. L'état de l'art des performances des TBH SiGeC se situe autour de 350 GHz à température ambiante [Khater04, Heinemann04], et au delà de 500 GHz à des températures cryogéniques [Krithivasan06]. Les transistors bipolaires intégrés avec les contraintes d'une technologie BiCMOS atteignent 330/300 GHz [Orner06], soit un tiers de THz. Qu'en est-il des deux tiers restant avant de franchir le cap symbolique du THz ?

Des simulations de profils de transistors extrêmement agressifs [Shi04, Shi05] ont permis de démontrer des performances théoriques supérieures à 1000 GHz pour les fréquences f_T et f_{MAX} . L'approche retenue repose sur une base très fine de quelques nm, avec un profil de Ge graduel pour minimiser τ_B . D'après ces simulations, lorsqu'on réduit le temps de transit dans la base, les autres temps de transit prennent des proportions plus importantes. Les temps de transit τ_E et τ_{EB} deviennent de sévères limitations des fréquences de coupure, effet amélioré par le fort courant collecteur.

Pour l'optimisation future des transistors bipolaire à hétérojonction, la réduction des dimensions verticales permet d'améliorer significativement τ_B . Un profil de base abrupt, avec un gradient de Ge important, permet d'obtenir de forts courants collecteur et de forts gains en cou-

rant, ce qui améliore les temps τ_E et τ_{EB} , avec au final un effet bénéfique sur f_T . Concernant le collecteur, un dopage trop abrupt aura pour effet de provoquer un brutal écroulement de f_T à l'apparition de l'effet Kirk. Il faut donc augmenter le dopage collecteur, sans toutefois négliger sa forme rétrograde, car elle permet de maintenir une capacité C_{BC} à un niveau relativement faible, et d'améliorer la tenue en tension.

La réduction de l'épaisseur du composant nécessite l'emploi de forts dopages de base afin d'éviter le perçage. L'augmentation du dopage de base est également nécessaire à l'amélioration de la fréquence f_{MAX} , car la réduction de l'épaisseur de base provoque une augmentation de la résistance de base. Il y a donc encore des avancées à réaliser dans le domaine des matériaux afin d'arriver à déposer des bases très fines et très fortement dopées.

La réduction de l'extension latérale du composant est un passage obligé de la montée en fréquence : Réduire la taille du dispositif permet de réduire les capacités de recouvrement et la résistance de base extrinsèque. La résistance de base intrinsèque est réduite en diminuant la largeur de la fenêtre émetteur, ce qui permet aussi de diminuer l'auto-échauffement du composant. La voie du THz semble tracée, mais nul doute qu'atteindre cette fréquence demandera d'une part d'améliorer encore notre connaissance du TBH, et d'autre part de poursuivre les innovations technologiques.

Publications de l'auteur

- [Barbalat05] B. Barbalat, T. Schwartzmann, P. Chevalier, T. Jagueneau, B. Vandelle, L. Rubaldo, F. Saguin, N. Zerounian, F. Aniel et A. Chantre, «Deep Trench Isolation Effect on Self-Heating and RF Performances of SiGeC HBTs», *European Solid-State Device Research Conference*, 2005, pp. 129–132.
- [Barbalat06a] B. Barbalat, F. Judong, L. Rubaldo, P. Chevalier, M. Proust, C. Richard, G. Borot, B. Vandelle, F. Saguin, D. Dutartre, N. Zerounian, F. Aniel et A. Chantre, «Experimental Study of Metallic Emitter SiGeC HBTs», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 251–254.
- [Barbalat06b] B. Barbalat, T. Schwartzmann, P. Chevalier, B. Vandelle, L. Rubaldo, A. Lachatter, F. Saguin, N. Zerounian, F. Aniel et A. Chantre, «Carbon effet on neutral base recombination in high speed SiGeC heterojunction bipolar transistors», *Semiconductor Science and Technology* **22**, 2006, pp. 99–102.
- [Barbalat06c] B. Barbalat, T. Schwartzmann, P. Chevalier, B. Vandelle, L. Rubaldo, F. Saguin, N. Zerounian, F. Aniel et A. Chantre, «Carbon effet on neutral base recombination in high speed SiGeC HBTs», *IEEE International SiGe Technology and Device Meeting*, 2006, pp. 238–239.

Références bibliographiques

- [Agarwal05] P. Agarwal, G. A. M. Hurkx, J. J. T. M. Donkers et J. W. Slotboom, «On the Delay Times in Vertically Scaled SiGe HBTs», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2005, pp. 264–267.
- [Ahlgren01] D. Ahlgren, B. Jagannathan, M. Meghelli, K. Chan, J.-S. Rieh, K. Schonenberg, S. Subbanna et G. Freeman, «U.S. patent application, numéro de série : 6,492,238 B1, enregistré le 22 juin 2001», 2001.
- [Ashburn88] P. Ashburn, *Design and Realization of Bipolar Transistors*, John Wiley and Sons, 1988.
- [Ashburn94] P. Ashburn, A. Nouailhat, M. D. R. Hashim, G. J. Parker, M. Mouis et D. J. Robbins, «Temperature dependence of the current gain of Si/Si_{1-x}Ge_x heterojunction and Si homojunction bipolar transistors», *IEEE European Solid-State Device Research Conference*, 1994, pp. 477–480.
- [Ashburn96] P. Ashburn, H. Boussetta, M. D. R. Hashim, A. Chantre, M. Mouis, G. J. Parker et G. Vincent, «Electrical Determination of Bandgap Narrowing in Bipolar Transistor with Epitaxial Si, Epitaxial Si_{1-x}Ge_x and Ion Implanted Bases», *IEEE Transactions on Electron Devices* **43**(5), Mai 1996, pp. 774–783.
- [Ashburn03] P. Ashburn, *SiGe Heterojunction Bipolar Transistors*, John Wiley and Sons, 2003.
- [Avenier05] G. Avenier, P. Chevalier, B. Vandelle, D. Lenoble, F. Saguin, S. Frégonèse, T. Zimmer et A. Chantre, «Investigation of fully- and partially-depleted self-aligned SiGeC HBTs on thin film SOI», *European Solid-State Device Research Conference*, 2005, pp. 133–136.
- [Banerjee03] B. Banerjee, S. Venkataraman, Y. Lu, S. Nuttinck, D. Heo, Y.-J. Emery-Chen, J. D. Cressler, J. Laskar, G. Freeman et D. C. Ahlgren, «Cryogenic Performance of a 200 GHz SiGe HBT Technology», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003, pp. 171–173.
- [Banerjee05] B. Banerjee, S. Venkataraman, Y. Lu, Q. Liang, C.-H. Lee, S. Nuttinck, D. Heo, Y.-J. Emery-Chen, J. D. Cressler, J. Laskar, G. Freeman et D. C. Ahlgren, «Cryogenic Operation of Third Generation, 200-GHz Peak- f_T , Silicon-Germanium Bipolar Transistors», *IEEE Transactions on Electron Devices* **52**(4), Avril 2005, pp. 585–593.
- [Barbalat04] B. Barbalat, *Réduction de la résistance d'émetteur de transistors bipolaires rapides*, Rapport de DEA, Université Joseph Fourier, Juin 2004.
- [Barbalat05] B. Barbalat, T. Schwartzmann, P. Chevalier, T. Jagueneau, B. Vandelle, L. Rubaldo, F. Saguin, N. Zerounian, F. Aniel et A. Chantre, «Deep Trench Isolation Effect on Self-Heating and RF Performances of SiGeC HBTs», *European Solid-State Device Research Conference*, 2005, pp. 129–132.

- [Barbalat06a] B. Barbalat, F. Judong, L. Rubaldo, P. Chevalier, M. Proust, C. Richard, G. Borot, B. Vandelle, F. Saguin, D. Dutartre, N. Zerounian, F. Aniel et A. Chantre, «Experimental Study of Metallic Emitter SiGeC HBTs», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 251–254.
- [Barbalat06b] B. Barbalat, T. Schwartzmann, P. Chevalier, B. Vandelle, L. Rubaldo, A. Lachater, F. Saguin, N. Zerounian, F. Aniel et A. Chantre, «Carbon effet on neutral base recombination in high speed SiGeC heterojunction bipolar transistors», *Semiconductor Science and Technology* **22**, 2006, pp. 99–102.
- [Barbalat06c] B. Barbalat, T. Schwartzmann, P. Chevalier, B. Vandelle, L. Rubaldo, F. Saguin, N. Zerounian, F. Aniel et A. Chantre, «Carbon effet on neutral base recombination in high speed SiGeC HBTs», *IEEE International SiGe Technology and Device Meeting*, 2006, pp. 238–239.
- [Baudry01] H. Baudry, *Développement et étude de transistors bipolaires hautes performances à base silicium-germanium*, Thèse de doctorat, Université de Grenoble I – Joseph Fourier, 2001.
- [Baudry03] H. Baudry, B. Szelag, F. Deléglise, M. Laurens, J. Mourier, F. Saguin, G. Troillard, A. Chantre et A. Monroy, «BiCMOS7RF : A Highly Manufacturable 0.25 μm BiCMOS RF-Applications-dedicated Technology using Non-Selective SiGeC Epitaxy», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003, pp. 207–210.
- [Blackburn04] D. L. Blackburn, «Temperature Measurements of Semiconductor Devices – A Review», *20th Annual IEEE Semiconductor Thermal Measurement and Management Symposium*, Mars 2004, pp. 70–80.
- [Boissonnet06] L. Boissonnet, F. Judong, B. Vandelle, L. Rubaldo, P. Bouillon, D. Dutartre, A. Perrotin, G. Avenier, P. Chevalier, A. Chantre et B. Rauber, «A 0.13 μm thin SOI CMOS technology with low-cost SiGeC HBTs and complementary high-voltage LDMOS», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 53–56.
- [Boucaud94] P. Boucaud, C. Francis, F. H. Julien, J.-M. Lourtioz, D. Bouchier, S. Bodnar, B. Lambert et J. L. Regolini, «Band-edge and deep level photoluminescence of pseudomorphic $\text{Si}_{1-x-y}\text{Ge}_x\text{C}_y$ alloys», *Applied Physics Letters* **64**(7), Février 1994, pp. 875–877.
- [Brossard06] F. Brossard, B. Vandelle, P. Chevalier et D. Dutartre, «Improved Carbon Incorporation in Selective Epitaxial Growth of SiGeC for HBT applications», *IEEE International SiGe Technology and Device Meeting*, 2006, pp. 240–241.
- [Böck04a] J. Böck, H. Schäfer, K. Aufinger, R. Stengl, S. Boguth, R. Schreiter, M. Rest, H. Knapp, M. Wurzer, W. Perndl, T. Böttner et T. F. Meister, «SiGe Bipolar Technology for Automotive Radar Applications», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2004, pp. 84–87.
- [Böck04b] J. Böck, H. Schäfer, H. Knapp, K. Aufinger, M. Wurzer, S. Boguth, T. Böttner, R. Stengl, W. Perndl et T. Meister, «3.3 ps Bipolar Technology», *IEEE International Electron Device Meeting*, 2004, pp. 255–258.
- [Chantre03] A. Chantre, M. Marty, B. Martinet et P. Chevalier, «European Patent Office, numéro de série EP 1 517 377 A1, enregistré le 17 septembre 2003», 2003.

- [Chantre06] A. Chantre, G. Avenier, P. Chevalier, B. Vandelle, F. Saguin, C. Maneux, D. Dutartre et T. Zimmer, «SiGe HBT design for CMOS compatible SOI», *IEEE International SiGe Technology and Device Meeting*, 2006, pp. 268–269.
- [Chevalier03] P. Chevalier, C. Fellous, B. Martinet, F. Leverd, F. Saguin, D. Dutartre et A. Chantre, «180 GHz f_T and f_{MAX} self-aligned SiGeC HBT using selective epitaxial growth of the base», *European Solid-State Device Research Conference*, 2003, pp. 299–302.
- [Chevalier04] P. Chevalier, C. Fellous, L. Rubaldo, D. Dutartre, M. Laurens, T. Jagueneau, F. Leverd, S. Bord, C. Richard, D. Lenoble, J. Bonnouvrier, M. Marty, A. Perrotin, D. Gloria, F. Saguin, B. Barbalat, R. Beerkens, N. Zerounian, F. Aniel et A. Chantre, «230 GHz self-aligned SiGeC HBT for 90 nm BiCMOS technology», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2004, pp. 225–228.
- [Chevalier05a] P. Chevalier, B. Barbalat, L. Rubaldo, B. Vandelle, D. Dutartre, P. Bouillon, T. Jagueneau, C. Richard, F. Saguin, A. Margain et A. Chantre, «300 GHz f_{MAX} self-aligned SiGeC HBT optimized towards CMOS compatibility», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2005, pp. 120–123.
- [Chevalier05b] P. Chevalier, D. Lagarde, G. Avenier, T. Schwartzmann, B. Barbalat, D. Lenoble, J. Bustos, F. Pourchon, F. Saguin, B. Vandelle, L. Rubaldo et A. Chantre, «Low-Cost Self-Aligned SiGeC HBT Module for High-Performance Bulk and SOI RFCMOS Platforms», *IEEE International Electron Device Meeting*, 2005, pp. 963–966.
- [Chevalier06a] P. Chevalier, D. Gloria, P. Scheer, S. Pruvost, F. Ganesello, F. Pourchon, P. Garcia, J.-C. Vildeuil, A. Chantre, C. Garnier, O. Noblanc, S. P. Voinigescu, T. O. Dickson, E. Laskin, S. T. Nicolson, T. Chalvatsis et K. H. K. Yau, «Advanced SiGe BiCMOS and CMOS platforms for Optical and Millimeter-Wave Integrated Circuits», *Compound Semiconductor IC Symposium Conference (CSICS) Digest*, 2006, pp. 12–15.
- [Chevalier06b] P. Chevalier, C. Raya, B. Geynet, F. Pourchon, F. Judong, F. Saguin, T. Schwartzmann, R. Pantel, B. Vandelle, L. Rubaldo, G. Avenier, B. Barbalat et A. Chantre, «250 GHz self-aligned Si/SiGeC HBT featuring an all-implanted collector», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 243–246.
- [Chevalier07] P. Chevalier, B. Barbalat, M. Laurens, B. Vandelle, L. Rubaldo, B. Geynet, S. Voinigescu, T. Dickson, N. Zerounian, S. Chouteau, D. Dutartre, A. Monroy, F. Aniel, G. Dambrine et A. Chantre, «High-Speed SiGe BiCMOS Technologies : 120-nm Status and End-of-Roadmap Challenges», *Digest of the 7th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, 2007, pp. 18–23.
- [Dawson92] D. E. Dawson, A. K. Gupta et M. L. Salib, «CW Measurements of HBT Thermal Resistance», *IEEE Transactions on Electron Devices* **39**(10), Octobre 1992, pp. 2235–2239.
- [Deixler02] P. Deixler, R. Colcaser, D. Bower, N. Bell, W. D. Boer, D. Szmyd, S. Bardly, W. Wilbanks, P. Barre, M. V. Houdt, J. C. J. Paasschens, H. Veenstra, E. V. D. Heijden, J. J. T. M. Donkers et J. W. Slotboom, «QUBiC4G :

- A $f_T/f_{MAX}=70/100$ GHz 0.25 μm low power SiGe-BiCMOS production technology with high quality passives for 12.5 Gb/s optical networking and emerging wireless applications up to 20 GHz», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2002, pp. 201–204.
- [Deixler04] P. Deixler, A. Rodriguez, W. D. Boer, H. Sun, R. Colclaser, D. Bower, N. Bell, A. Yao, R. Brock, Y. Bouttement, G. A. M. Hurkx, L. F. Tiemeijer, J. C. J. Paasschens, H. G. A. Huizing, D. M. H. Hartskeerl, P. Agarwal, P. H. C. Magnée, E. Aksen et J. W. Slotboom, «QUBiC4X : An $f_T/f_{MAX}=130/140$ GHz SiGe :C-BiCMOS Manufacturing Technology with Elite Passives for Emerging Microwave Applications», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2004, pp. 233–236.
- [Donkers03] J. J. T. M. Donkers, P. H. C. Magnée, H. G. A. Huizing, P. Agarwal, E. Aksen, P. Meunier-Beillard, F. Neuilly, R. J. Havens et T. Vanhoucke, «Vertical Profile Optimization of a Self-Aligned SiGeC HBT Process with an n-Cap Emitter», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003, pp. 111–114.
- [Donkers04] J. J. T. M. Donkers, T. Vanhoucke, P. Agarwal, R. J. E. Hueting, P. Meunier-Beillard, M. N. Vijayaraghavan, P. H. C. Magnée, M. A. Verheijen, R. de Kort et J. W. Slotboom, «Metal Emitter SiGe :C HBTs», *IEEE International Electron Device Meeting*, 2004, pp. 243–246.
- [Fellous02] C. Fellous, *Epitaxies d'alliages Silicium-Germanium en milieu industriel pour les application BiCMOS*, Thèse de doctorat, Université de Grenoble I – Joseph Fourier, 2002.
- [Gao91] G. Gao, M. S. Ünlü, H. Morkoç et D. L. Blackburn, «Emitter Ballasting Resistor Design for, and Current Handling Capability of AlGaAs/GaAs Power Heterojunction Bipolar Transistors», *IEEE Transactions on Electron Devices* **38**(2), Février 1991, pp. 185–196.
- [Glicksman58] M. Glicksman, «Mobility of Electrons in Germanium-Silicon Alloys», *Phys. Rev.* **111**(1), Juillet 1958, pp. 125–128.
- [Harrison04] S. Harrison, P. Coronel, A. Cros, R. Cerutti, F. Leverd, A. Beverina, R. Wacquez, J. Bustos, D. Delille, B. Tavel, D. Barge, J. Bienacel, M. P. Samson, F. Martin, S. Maitrejean, D. Munteanu, J. L. Autran et T. Skotnicki, «Polygate REplacement Through Contact Hole (PRETCH) : A new method for High K/Metal gate and multi-oxide implementation on chip», *IEEE International Electron Device Meeting*, 2004, pp. 291–294.
- [Heinemann02] B. Heinemann, H. Rucker, R. Barth, J. Bauer, D. Bolze, E. Bugiel, J. Drews, K.-E. Ehwald, T. Grabolla, U. Haak, W. Hoppner, D. Knoll, D. Kruger, B. Kuck, R. Kurps, M. Marschmeyer, H. H. Richter, P. Schley, D. Schmidt, R. Scholz, B. Tillack, W. Winkler, D. Wolnsky, H.-E. Wulf, Y. Yamamoto et P. Zaumseil, «Novel Collector Design for High-Speed SiGeC HBTs», *IEEE International Electron Device Meeting*, 2002, pp. 775–778.
- [Heinemann03] B. Heinemann, R. Barth, D. Bolze, J. Drews, P. Formanek, O. Fursenko, M. Glante, K. Glowatzki, A. Gregor, U. Haak, W. Hoppner, D. Knoll, R. Kurps, S. Marschmeyer, S. Orłowski, H. Rucker, P. Schley, D. Schmidt, R. Scholz, W. Winkler et Y. Yamamoto, «A Complementary BiCMOS Technology with High-Speed npn and pnp SiGeC HBTs», *IEEE International Electron Device Meeting*, 2003, pp. 117–120.

- [Heinemann04] B. Heinemann, R. Barth, D. Bolze, J. Drews, P. Formanek, T. Grabolla, U. Haak, W. Höppner, D. Knoll, B. Kuck, R. Kurps, K. Köpke, S. Marschmeyer, H. Richter, H. Rücker, P. Schley, D. Schmidt, W. Winkler, D. Wolanski, H. E. Wulf et Y. Yamamoto, «A Low-Parasitic Collector Construction for High-Speed SiGeC HBTs», *IEEE International Electron Device Meeting*, 2004, pp. 251–254.
- [Huylenbroeck04] S. van Huylenbroeck, A. Sibaja, A. Piontek, L. J. Choi, M. W. Xu, N. Ouassif, F. Vieugels, K. van Wichelen, L. Witters, E. Kunnen, P. Leray, K. Devriendt, X. Shi, R. Loo et S. Decoutere, «Lateral and Vertical Scaling of a QSA HBT for 0.13 μm 200 GHz SiGeC BiCMOS Technology», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2004, pp. 247–250.
- [Huylenbroeck06] S. van Huylenbroeck, L. J. Choi, A. Sibaja-Hernandez, A. Piontek, D. Lintén, M. Dehan, O. Dupuis, G. Carchon, F. Vleugels, E. Kunnen, P. Leray, K. Devriendt, X. P. Shi, R. Loo, E. Hijzen et S. Decoutere, «A 205/275 f_T/f_{MAX} Airgap Isolated 0.13 μm BiCMOS Technology featuring on-chip High Quality Passives», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 57–60.
- [Jagannathan02] B. Jagannathan, M. Kather, F. Pagette, J. S. Rieh, D. Angell, H. Chen, J. Florkey, F. Golan, D. R. Greenberg, R. Groves, S. J. Jeng, J. Johnson, E. Mengistu, K. T. Schonenberg, C. M. Schnabel, P. Smith, A. Stricker, D. C. Ahlgren, G. Freeman, K. Stein et S. Subbanna, «Self Aligned SiGe NPN Transistors with 285 GHz f_{MAX} and 207 GHz f_T in a Manufacturable Technology», *IEEE Electron Device Letters* **23**(5), Mai 2002, pp. 258–260.
- [Jagannathan03] B. Jagannathan, M. Meghelli, K. Chan, J.-S. Rieh, K. Schonenberg, D. Ahlgren, S. Subbanna et G. Freeman, «3.9 ps SiGe HBT ECL Ring Oscillator and Transistor Design for Minimum Gate Delay», *IEEE Electron Device Letters* **24**(5), Mai 2003, pp. 324–326.
- [John06] J. P. John, J. Kirshgessner, M. Menner, H. Rueda, F. Chai, D. Morgan, J. Hildreth, M. Dawdy, R. Reuter et H. Li, «Development of a Cost-Effective, Selective-Epi, SiGeC HBT Module for 77 GHz Automotive Radar», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 247–250.
- [Jouan01] S. Jouan, *Développement et caractérisation de transistors bipolaires à hétérojonctions Si/SiGe pour les circuits radiofréquences*, Thèse de doctorat, Université de Grenoble I – Joseph Fourier, 2001.
- [Kerr75] J. A. Kerr et F. Berz, «The Effect of Emitter Doping Gradient on f_T in Microwave Bipolar Transistors», *IEEE Transactions on Electron Devices* **22**(1), Janvier 1975, pp. 15–20.
- [Khater04] M. Khater, J. S. Rieh, T. Adam, A. Chinthakindi, J. Johnson, R. Krishnasamy, M. Meghelli, F. Pagette, D. Sanderson, C. Schnabel, K. T. Schonenberg, P. Smith, K. Stein, A. Stricker, S. J. Jeng, D. Ahlgren et G. Freeman, «SiGe HBT Technology with $f_{MAX}/f_T = 350/300$ GHz and Gate Delay Below 3.3 ps», *IEEE International Electron Device Meeting*, 2004, pp. 247–250.
- [Kirk62] C. T. Kirk, «A Theory of Transistor CutOff Frequency (f_T) Falloff at High Current Densities», *IRE Transactions on Electron Devices* **9**, Mars 1962, pp. 164 – 174.

- [Klaassen92a] D. B. M. Klaassen, «A Unified Mobility Model for Device Simulation - I. Model Equations and Concentration Dependence», *Solid-State Electronics* **35**(7), juillet 1992, pp. 953–959.
- [Klaassen92b] D. B. M. Klaassen, «A Unified Mobility Model for Device Simulation - II. Temperature Dependence of Carrier Mobility and Lifetime», *Solid-State Electronics* **35**(7), juillet 1992, pp. 961–967.
- [Klaassen92c] D. B. M. Klaassen, J. W. Slotboom et H. C. DeGraaff, «Unified apparent bandgap narrowing in n- and p-type silicon», *Solid-State Electronics* **35**(2), février 1992, pp. 125–129.
- [Kraft05] J. Kraft, D. Kraft, B. Löffler, H. Jauk et E. Wachman, «Usage of HBTs beyond BV_{CEO} », *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2005, pp. 33–36.
- [Krithivasan06] R. Krithivasan, Y. Lu, J. D. Cressler, J. S. Rieh, M. Khater, D. C. Ahlgren et G. Freeman, «Half-Terahertz Operation of SiGe HBTs», *IEEE Transactions on Electron Devices* **27**(7), Juillet 2006, pp. 567–569.
- [Kroemer85] H. Kroemer, «Two integral relations pertaining to the electron transport through a bipolar transistor with a non-uniform energy gap in the base region», *Solid-State Electronics* **28**, 1985, pp. 1101–1103.
- [Kunz02] V. D. Kunz, C. H. de Groot, S. Hall, I. M. Anteney, A. I. Abdul-Rahim et P. Ashburn, «Application of Polycrystalline SiGe for Gain Control in SiGe Heterojunction Bipolar Transistors», *European Solid-State Device Research Conference*, 2002, pp. 171–174.
- [Lagarde06] D. Lagarde, P. Chevalier, T. Schwartzmann et A. Chantre, «Band-to-Band Tunneling in Vertically Scaled SiGeC HBTs», *IEEE Transactions on Electron Devices* **27**(4), Avril 2006, pp. 275–277.
- [Lang85] D. V. Lang, R. People, J. C. Bean et A. Sergent, «Measurement of the band gap of Ge_xSi_{1-x}/Si strained-layer heterostructures», *Applied Physics Letter* **47**(12), Décembre 1985, pp. 1333–1335.
- [Lanzerotti96a] L. D. Lanzerotti, A. S. Amour, C. Liu, J. C. Sturm, J. K. Watanabe et N. D. Theodore, «Si/Si_{1-x-y}Ge_xC_y/Si Heterojunction Bipolar Transistor», *IEEE Electron Device Letters* **17**(7), juillet 1996, pp. 334–337.
- [Lanzerotti96b] L. D. Lanzerotti, J. C. Sturm, E. Stach, R. Hull, T. Buyuklimanli, et C. Magee, «Suppression of boron outdiffusion in SiGe HBT's by carbon incorporation», *IEEE International Electron Device Meeting*, 1996, pp. 249–252.
- [Laskin06] E. Laskin, S. T. Nicolson, P. Chevalier, A. Chantre, B. Sautreuil et S. P. Voiginescu, «Low-Power, Low-Phase Noise SiGe HBT Static Frequency Divide Topologies up to 100 GHz», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 235–238.
- [Laurens03] M. Laurens, B. Martinet, O. Kermarrec, Y. Campidelli, F. Deleglise, D. Dutartre, G. Troillard, D. Gloria, J. Bonnouvrier, R. Beerkens, V. Rousset, F. Leverd, A. Chantre et A. Monroy, «A 150 GHz f_T/f_{MAX} 0.13 μm SiGe :C Technology», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003, pp. 199–202.
- [Liang02] Q. Liang, J. D. Cressler, G. Niu, R. M. Malladi, K. Newton et D. L. Hareme, «A Physics-Based High-Injection Transit-Time Mode Applied to Barrier Effect in Si_{1-x}Ge_x HBTs», *IEEE Transactions on Electron Devices* **49**(10), Octobre 2002, pp. 1807–1813.

- [Malorny03] M. Malorny, M. Schröter, D. Céli et D. Berger, «An improved method for determining the transit time of Si/SiGe bipolar transistors», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003, pp. 229–232.
- [Martinet02] B. Martinet, F. Romagna, O. Kermarrec, Y. Campidelli, F. Saguin, H. Baudry, M. Marty, D. Dutartre et A. Chantre, «An investigation of the static and dynamic characteristics of high speed SiGe :C HBTs using a poly-SiGe emitter», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2002, pp. 147–150.
- [Mason54] S. J. Mason, «Power Gain in Feedback Amplifier», *IRE Transaction on Circuit Theory* **1**(2), Juin 1954, pp. 20–25.
- [McAllister04] S. P. McAllister, W. R. McKinnon, S. J. Kovacic et H. Lafontaine, «Self-heating in multi-emitter SiGe HBTs», *Solid-State Electronics* **48**, Juin 2004, pp. 2001–2006.
- [Meister03] T. F. Meister, H. Schäfer, K. Aufinger, R. Stengl, S. Boguth, R. Schreiter, M. Rest, H. Knapp, M. Wurzer, A. Mitchell, T. Bottner et J. Bock, «SiGe Bipolar Technology with 3.9 ps gate delay», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003, pp. 103–106.
- [Meyer87] R. G. Meyer et R. S. Muller, «Charge-Control Analysis of the Collector-Base Space-Charge Region Contribution to Bipolar Transistor Time Constant τ_T », *IEEE Transactions on Electron Devices* **34**(2), Février 1987, pp. 450–452.
- [Miura04] M. Miura, H. Shimamoto, R. Hayami, A. Kodama, T. Tominari, T. Hashimoto et K. Washio, «Optimization of Vertical Profiles of SiGe HBT/BiCMOS by Promoting Emitter Diffusion Process», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2004, pp. 92–95.
- [Miura06] M. Miura, H. Shimamoto, R. Hayami, A. Kodama, T. Tominari, T. Hashimoto et K. Washio, «SiGe BiCMOS Technologies for Improving Sensitivity and High-Speed Characteristics of the Communication LSIs», *IEEE International SiGe Technology and Device Meeting*, 2006, pp. 228–229.
- [Mnif02] H. Mnif, T. Zimmer, J.-L. Battaglia et B. Ardouin, «Modeling the self-heating effect in SiGe HBTs», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2002, pp. 96–99.
- [Nicolson06] S. T. Nicolson, K. Yau, K. Tang, P. Chevalier, A. Chantre, B. Sautreuil et S. P. Voiginescu, «Design and Scaling of SiGe BiCMOS VCOs Above 100 GHz», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 142–145.
- [Ning03] T. Ning, «Polysilicon-Emitter SiGe-Base Bipolar Transistors — What Happens When Ge Gets Into the Emitter ?», *IEEE Transaction on Electron Devices* **50**(5), Mai 2003, pp. 1346–1352.
- [Oettinger76] F. F. Oettinger, D. L. Blackburn et S. Rubin, «Thermal Characterization of Power Transistors», *IEEE Transactions on Electron Devices* **23**(8), Août 1976, pp. 831–838.
- [Ooi04] B. L. Ooi, B. Chen, F. Lin, P. S. Kooi et C. S. Hui, «A Fast and Practical Approach to the Determination of Junction Temperature and Thermal Resistance for BJT/HBT Devices», *IEEE International Conference on Communications Systems*, 2004, pp. 588–591.

- [Orner03] B. A. Orner, Q. Z. Liu, R. Rainey, A. Stricker, P. Geiss, P. Gray, M. Zierack, M. Gordon, D. Collins, V. Ramachandran, W. Hodge, C. Willets, A. Joseph, J. Dunn, J.-S. Rieh, J. S. Jeng, E. Eld, G. Freeman et D. Ahlgren, «A 0.13 μm BiCMOS Technology Featuring a 200/280 GHz (f_T/f_{MAX}) SiGe HBT», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2003, pp. 203–206.
- [Orner06] B. A. Orner, M. Dahlström, A. Pothiwala, R. M. Rassel, Q. Liu, H. Ding, M. Khater, D. Ahlgren, A. Joseph et J. Dunn, «BiCMOS Technology Featuring a 300/330 GHz (f_T/f_{MAX}) SiGe HBT for Millimeter Wave Applications», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 49–52.
- [Osten97] H. J. Osten, G. Lippert, P. Gaworzewski et R. Sorge, «Impact of low carbon concentrations on the electrical properties of highly boron doped SiGe layers», *Applied Physics Letters* **71**(11), Septembre 1997, pp. 1522–1524.
- [Pakfar01] A. Pakfar, «How to extract accurate dopant profile from SIMS measurement ?», Rapport technique. TM01_084, STMicroelectronics, Mai 2001.
- [People85a] R. People, «Indirect band gap of coherently strained $\text{Ge}_x\text{Si}_{1-x}$ bulk alloys on $\langle 001 \rangle$ silicon substrates», *Physical Review B* **32**(2), Juillet 1985, pp. 1405–1408.
- [People85b] R. People et J. C. Bean, «Calculation of critical layer thickness versus lattice mismatch for $\text{Ge}_x\text{Si}_{1-x}$ strained-layer heterostructures», *Applied Physics Letters* **47**(3), Août 1985, pp. 322–324.
- [People86] R. People et J. C. Bean, «Band alignments of coherently strained $\text{Ge}_x\text{Si}_{1-x}/\text{Si}$ heterostructures on $\langle 001 \rangle$ $\text{Ge}_y\text{Si}_{1-y}$ substrates», *Applied Physics Letters* **48**(8), Février 1986, pp. 538–540.
- [Piontek06] A. Piontek, T. Vanhoucke, S. van Huylbroeck, L. J. Choi, G. A. M. Hurkx, E. Hijzen et S. Decoutere, «A 205/275 f_T/f_{MAX} Airgap Isolated 0.13 μm BiCMOS Technology featuring on-chip High Quality Passives», *IEEE International SiGe Technology and Device Meeting*, 2006, pp. 242–243.
- [Racanelli01a] M. Racanelli, «International patent application, numéro de série : WO 02/43132 A1, enregistré le 19 novembre 2001», 2001.
- [Racanelli01b] M. Racanelli, K. Schuegraf, A. Kalburge, A. Kar-Roy, B. Shen, C. Hu, D. Chapek, D. Howard, D. Quon, F. Wang, G. U'ren, L. Lao, H. Tu, J. Zheng, J. Zhang, K. Bell, K. Yin, P. Joshi, S. Akhtar, S. Vo, T. Lee, W. Shi et P. Kempf, «Ultra High Speed SiGe NPN for Advanced BiCMOS Technology», *IEEE International Electron Devices Meeting*, 2001, pp. 336–339.
- [Racanelli03] M. Racanelli et P. Kempf, «SiGe BiCMOS technology for communication products», *IEEE Custom Integrated Circuits Conference*, 2003, pp. 331–334.
- [Richard04] S. Richard, *Modélisation physique de la structure électronique, du transport et de l'ionisation par choc dans les matériaux IV-IV massifs, contraintes et dans les puits quantiques*, Thèse de doctorat, Université Paris-Sud XI, 2004.
- [Rickelt01] M. Rickelt, H. M. Rein et E. Rose, «Influence of Impact-Ionization-Induced Instabilities on the Maximum Usable Output Voltage of Si-Bipolar Transistors», *IEEE Transaction on Electron Devices* **48**(4), Avril 2001, pp. 774–783.

- [Rieh00] J. S. Rieh, P. K. Bhattacharya et E. T. Croke, «Temperature Dependent Minority Electron Mobilities in Strained $\text{Si}_{1-x}\text{Ge}_x$ ($0.2 \leq x \leq 0.4$) Layers», *IEEE Transactions on Electron Devices* **47**(4), Avril 2000, pp. 883–890.
- [Rieh01] J. S. Rieh, D. Greenberg, B. Jagannathan, G. Freeman et S. Subbanna, «Measurements and Modeling of Thermal Resistance of High Speed SiGe Heterojunction Bipolar Transistor», *Topical Meeting on Silicon Integrated Circuits in RF Systems*, 2001, pp. 110–113.
- [Rieh02] J. S. Rieh, J. Johnson, S. Furkay, D. Greenberg, G. Freeman et S. Subbanna, «Structural Dependence of the Thermal Resistance of Trench Isolated Bipolar Transistors», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2002, pp. 100–103.
- [Rieh04a] J. S. Rieh, D. Greenberg, M. Khater, K. T. Schonnenberg, S.-J. Jeng, F. Pagette, T. Adam, A. Chinthakindi, J. Florkey, B. Jagannathan, J. Johnson, R. Krishnasamy, D. Sanderson, C. Schnabel, P. Smith, A. Stricker, S. Sweeney, K. Vaed, T. Yanagisawa, D. Ahlgren, K. Stein et G. Freeman, «SiGe HBTs for Millimeter-Wave Applications with Simultaneously Optimized f_T and f_{MAX} of 300 GHz», *IEEE Radio Frequency Integrated Circuits Symposium*, 2004, pp. 395–398.
- [Rieh04b] J. S. Rieh, M. Khater, K. T. Schonnenberg, F. Pagette, P. Smith, T. N. Adam, K. Stein, D. Ahlgren et G. Freeman, «Collector Vertical Scaling and Performance Tradeoffs in 300 GHz SiGe HBTs», *Device Research Conference*, 2004, pp. 235–236.
- [Rieh05a] J. S. Rieh, D. Greenberg, Q. Liu, A. J. Joseph, G. Freeman et D. C. Ahlgren, «Structure Optimization of Trench-Isolated SiGe HBTs for Simultaneous Improvements in Thermal and Electrical Performances», *IEEE Transactions on Electron Devices* **52**(12), Décembre 2005, pp. 2744–2752.
- [Rieh05b] J. S. Rieh, D. Greenberg, A. Stricker et G. Freeman, «Scaling of SiGe Heterojunction Bipolar Transistors», *Proceedings of the IEEE* **93**(9), Septembre 2005, pp. 1522–1538.
- [Roulston90] D. J. Roulston, *Bipolar Semiconductor Devices*, Electronic Engineering Series, McGraw-Hill International Editions, 1990.
- [Rucker04] H. Rucker, B. Heinemann, R. Barth, D. Bolze, J. Drews, O. Fursenko, T. Grabolla, U. Haak, W. Hoppner, D. Knoll, S. Marschmeyer, N. Mohapatra, H. H. Richter, P. Schley, D. Schmidt, B. Tillack, G. Weidner, D. Wolansky, H. E. Wulf et Y. Yamamoto, «Integration of High-Performance SiGeC HBTs with Thin-Film SOI CMOS», *IEEE International Electron Device Meeting*, 2004, pp. 239–242.
- [Saitoh04] T. Saitoh, T. Kawashima, Y. Kanzawa, J. Sato-Iwanaga, K. Idota, T. Takagi, T. Ohnishi, K. Yuki, T. Sano et S. Sawada, «Base Current Control in Low- V_{BE} -Operated SiGeC Heterojunction Bipolar Transistors Using SiGe-Cap Structure and High-Carbon-Content Base», *Japanese Journal of Applied Physics* **43**(4B), Avril 2004, pp. 2250–2254.
- [Schröter00] M. Schröter, «HICUM - A scalable physics-based compact bipolar transistor model - Description of model version 2.1», www.iee.et.tu-dresden.de/~schroter/Models/hicman.pdf, Décembre 2000.

- [Schäffler97] F. Schäffler, «High-mobility Si and Ge structures», *Semiconductor Science and Technology* **12**, 1997, pp. 1515–1549.
- [Shi04] Y. Shi et G. Niu, «Vertical Profile Design and Transit Time Analysis of Nano-Scale SiGe HBTs for Terahertz f_T », *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2004, pp. 213–216.
- [Shi05] Y. Shi et G. Niu, «2-D Analysis of Device Parasitics for 800/1000 GHz f_T/f_{MAX} SiGe HBT», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2005, pp. 252–255.
- [Voiginescu06] S. P. Voiginescu, T. Chalvatsis, K. H. K. Yau, A. Hazneci, A. Garg, S. Shahramian, T. Yao, M. Gordon, T. O. Dickinson, E. Laskin, S. T. Nicolson, A. C. Carusone, L. Tchoketch-Kebir, O. Yuryevich, G. Ng, B. Lai et P. Liu, «SiGe BiCMOS for Analog, High-Speed Digital and Millimeter-Wave Applications beyond 50 GHz», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 223–230.
- [Vytla06] R. K. Vytla, T. F. Meister, K. Aufinger, D. Lukashevich, S. Boguth, H. Knapp, J. Böck, H. Schäfer et R. Lachner, «Simultaneous Integration of SiGe High Speed Transistor and High Voltage Transistors», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2006, pp. 61–64.
- [Wada02] S. Wada, Y. Nonaka, T. Saito, T. Tominari, K. Koyu, K. Ikeda, K. Sakai, K. Sasahara, K. Watanabe, H. Fujiwara, F. Murata, E. Ohue, Y. Kiyota, H. Shimamoto, K. Washio, R. Takeyari, H. Hosoe et T. Hashimoto, «A Manufacturable 0.18 μm SiGe BiCMOS Technology for 40-Gb/s Optical Communication LSIs», *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, 2002, pp. 84–87.
- [Waldrop92] J. R. Waldrop, K. C. Wang et P. M. Asbeck, «Determination of the Junction Temperature in AlGaAs/GaAs Heterojunction Bipolar Transistors by Electrical Measurements», *IEEE Transactions on Electron Devices* **39**(5), Mai 1992, pp. 1248–1250.
- [Washio02] K. Washio, E. Ohue, R. Hayami, A. Kodama, H. Shimamoto, M. Miura, K. Oda, I. Suzumura, T. Tominari et T. Hashimoto, «Ultra-High-Speed Scaled-down Self-Aligned SEG SiGe HBTs», *IEEE International Electron Device Meeting*, 2002, pp. 767–770.
- [Yau06] K. H. K. Yau et S. P. Voiginescu, «Modeling and Extraction of SiGe HBT Noise Parameters from Mesured Y-Parameters and Accounting for Noise Correlation», *Digest of the 6th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, 2006, pp. 226–229.
- [Zerounian00] N. Zerounian, *Etude expérimentale et modélisation de composants micro-ondes bipolaires et à effet de champ SiGe*, Thèse de doctorat, Université Paris-Sud XI, 2000.
- [Zerounian03] N. Zerounian, «Transit times of SiGeC HBTs using non selective base epitaxy», *International Semiconductor Device Research Symposium*, Décembre 2003, pp. 166–167.

Liste des figures

I.1	Schéma d'un transistor bipolaire NPN. En noir : contacts métalliques ; zones grisées : zones de charge d'espace des jonctions E/B et B/C.	5
I.2	Représentation symbolique du NPN utilisée dans les schémas électriques. . . .	5
I.3	Diagramme de bandes d'un transistor bipolaire NPN : a – au repos et b – en régime de fonctionnement normal. L'abaissement des barrières pour les électrons et pour les trous autorise le passage du courant.	6
I.4	Profil de dopage générique d'un transistor bipolaire	7
I.5	Représentation 2D des deux types de croissance possibles : (a) Croissance pseudomorphique, avec contrainte biaxiale dans le plan de l'interface ; (b) Croissance relaxée, des dislocations apparaissent dans le plan de l'interface.	9
I.6	Epaisseur critique h_C en fonction du taux de germanium x	9
I.7	Diagrammes de bande du Si et du Ge à 0K (données [Richard04])	10
I.8	Valeur de la bande interdite du $Si_{1-x}Ge_x$ contraint ou non sur Si en fonction du taux de Ge x (données [Lang85]).	11
I.9	Schéma de l'alignement de bandes entre le Si et le SiGe contraint sur Si.	11
I.10	Variations (en meV) en bande de conduction et de valence d'une couche de $Si_{1-x}Ge_x$ contrainte sur un substrat de $Si_{1-x_s}Ge_{x_s}$ relaxé en fonction des taux de germanium x et $x_{substrat}$ (données [Schäffler97]).	11
I.11	Diagramme de bandes d'un transistor bipolaire a hétérojonction SiGe, à taux de Ge constant.	12
I.12	Pseudo-champ électrique créé par l'abaissement de la bande de conduction dans le cas d'un profil rétrograde de Germanium.	12
I.13	Principaux courants du transistor bipolaire dans un mode de fonctionnement direct.	14
I.14	Densités de porteurs dans un transistor bipolaire en régime de fonctionnement normal.	15
I.15	Schéma de la discontinuité en bande de valence à la jonction émetteur/base. Malgré cette discontinuité, la hauteur de barrière totale vue par les trous reste constante.	20
I.16	Montage base commune.	22
I.17	Montage émetteur commun.	22
I.18	Exemple de courbe de Gummel. Dimensions du composant : $0.15 \times 3.6 \mu m^2$	24
I.19	Caractéristiques de sortie en montage émetteur commun. I_B varie de 0 à 11 μA par pas de 1 μA . La diminution du gain aux forts courants est due à l'auto-échauffement.	24
I.20	Caractéristiques de sortie en montage base commune. I_E varie de 0 à 1 mA par pas de 0.1 mA	24
I.21	Explication de l'effet tunnel à faible polarisation [Lagarde06].	25
I.22	Courants contribuant à l'effet tunnel à faible polarisation, et résultante du courant. .	26

I.23	Effet tunnel assisté par phonons : Les électrons passent dans la vallée supérieure en $k = 0$, puis franchissent la barrière par effet tunnel direct.	26
I.24	Effet Early direct : La variation d'épaisseur de zone de charge d'espace B/C est à l'origine d'une variation de courant collecteur.	26
I.25	Caractéristiques de sortie d'un transistor bipolaire mettant en évidence la tension d'Early directe.	27
I.26	Explication de l'effet Early inverse par une modulation de l'épaisseur de base neutre par la polarisation V_{BE}	27
I.27	Influence du profil Germanium sur la tension d'Early directe. La variation du nombre de Gummel avec l'épaisseur de la base est plus faible lorsque le profil de Ge est graduel, ce qui améliore la tension d'Early.	28
I.28	Influence du profil graduel de Germanium sur la tension d'Early inverse. La dépendance du nombre de Gummel avec V_{BE} est plus forte dans le cas d'un profil graduel qu'un profil plat.	28
I.29	Mécanisme d'avalanche dans la jonction base/collecteur polarisée en inverse. .	29
I.30	Caractéristique courant-tension de la diode base-collecteur (émetteur ouvert). La tension de claquage BV_{CBO} est ici de 5.7 V.	30
I.31	Caractéristique courant-tension de la diode émetteur-base (collecteur ouvert). La tension de claquage BV_{EBO} est ici comprise entre 1 et 3 V.	30
I.32	Définition de la tension de claquage BV_{CEO}	30
I.33	Effet de haute injection dans la base : lorsque le courant augmente, la concentration de trous augmente pour compenser les électrons en excès.	31
I.34	Effet de réduction de bande interdite sous l'effet de forts dopage : a – Sans BGN, le gap du matériau est le même que le matériau intrinsèque ; b – Avec BGN, le gap apparent du matériau est réduit.	33
I.35	Evolution de la mobilité des porteurs dans le Silicium en fonction du dopage, à 300 K, dans le cas où les porteurs sont minoritaires ou majoritaires.	35
I.36	Évolution de la mobilité des porteurs dans le Silicium en fonction de la température pour différents niveaux de dopage.	35
I.37	Répartition des porteurs (a) à l'équilibre, et (b) sous polarisation normale directe. .	36
I.38	Répartition des charges dans la jonction B/C (a) au repos, (b) avec passage de courant.	40
I.39	Charges stockées dans une jonction en fonctionnement, deux charges opposées se font face de part et d'autre de la jonction.	41
I.40	Définition de la fréquence de transition f_T	42
I.41	Schéma équivalent petit signal du transistor bipolaire pour le calcul de f_T	43
I.42	Evolution de la fréquence de transition f_T avec le courant collecteur.	44
I.43	Extraction de τ_{EC} en fonction de l'inverse du courant collecteur.	44
I.44	Schéma équivalent petit signal du transistor bipolaire adapté en impédance pour le calcul de f_{MAX}	45
I.45	Evolution de la fréquence maximale d'oscillation f_{MAX} avec le courant collecteur, et comparaison avec la fréquence de transition f_T	47
I.46	Evolution des courbes de Gummel avec la température.	47
I.47	Evolution du gain en courant avec la température.	47
I.48	Evolution de la fréquence de transition f_T avec la température.	48
I.49	Présentation d'une coupe MEB	50
I.50	Présentation d'une coupe MET	50
I.51	Exemple de profil SIMS.	50
I.52	Effet de recul observé sur un front montant abrupt.	50

I.53	Influence de la résistance d'émetteur sur les caractéristiques de Gummel. Le facteur d'idéalité est extrait dans la zone linéaire.	51
I.54	Chute de tension dues aux résistances séries R_B et R_E	51
I.55	Variation de courant de base et de collecteur à cause de l'effet d'avalanche. . .	52
I.56	Extraction de la tension de claquage BV_{CEO} d'après le facteur de multiplication $M - 1$	53
I.57	Tension d'Early directe effective mesurée sur un TBH avec $V_{BE} = 0.9$ V. A V_{BE} plus faible, cette tension est plus élevée.	53
I.58	Définition des paramètres S pour un quadripôle.	54
I.59	Représentation du dispositif sous test (DST) encadré de ses plots d'accès hyperfréquence.	56
I.60	Modélisation des composantes parasites ajoutées par les plots d'accès au composant.	56
I.61	Motifs d'épluchage du composant.	56
I.62	Extraction des fréquences f_T et f_{MAX} pour un transistor de dimensions $0.12 \times 3.6 \mu m^2$ d'après les gains h_{21} et U.	57
I.63	Comparaison des caractéristiques de sortie obtenues par une mesure statique et une mesure pulsée. L'auto-échauffement visible en régime statique ne l'est plus sur les caractéristiques pulsées.	59
I.64	Présentation d'un spectre de bruit basse fréquence. La décroissance en $1/f$ et le bruit de grenaille (bruit blanc) sont identifiables ($0.15 \times 3.6 \mu m^2$).	60
II.1	Présentation d'une structure quasi auto-alignée (QSA) à simple polysilicium. .	64
II.2	Présentation d'une structure quasi auto-alignée (QSA) à double polysilicium. .	65
II.3	Problématique d'alignement d'une structure QSA double polysilicium.	66
II.4	Vue en coupe d'une structure complètement auto-alignée (FSA) par le procédé damascène.	67
II.5	Vue en coupe d'un transistor réalisé avec le procédé damascène, mettant en évidence un défaut de remplissage de l'émetteur.	68
II.6	Présentation d'une structure complètement auto-alignée (FSA) à double polysilicium.	69
II.7	Réduction de la distance émetteur/contact de base (d) induite par l'auto-alignement.	70
II.8	Epitaxie sélective de la base : Une cavité est aménagée dans l'oxyde piédestal isolant le substrat du Polybase, puis le SiGe est déposée dans cette cavité, sélectivement par rapport aux diélectriques.	71
II.9	Étapes de fabrication du transistor bipolaire	72
II.10	Observation TEM du lien entre base intrinsèque et base extrinsèque. Le lien s'effectue correctement grâce à la consommation de l'oxyde piédestal sous le polybase.	73
II.11	Observation TEM de la structure FSA-SEG achevée.	74
II.12	Intégration des étapes spécifiques au TBH dans la route CMOS $0.13 \mu m$ [Boissonnet06].	75
II.13	Schéma de la structure du TBH réalisé par NXP [Donkers03].	76
II.14	Vue TEM du transistor bipolaire de l'IMEC [Huylenbroeck06]. Les DTI vides sont bien visibles (en blanc).	77
II.15	Vue TEM du transistor bipolaire de Jazz [Racanelli01b].	78
II.16	Vue au microscope électronique à transmission du transistor bipolaire d'Infineon [Böck04a].	79
II.17	Vue au microscope électronique à balayage du transistor bipolaire d'IBM [Orner06].	80
II.18	Valeurs de f_T et f_{MAX} obtenues par les différents concurrents entre 2002 et 2006.	81

III.1	Présentation du profil de dopage de l'émetteur.	84
III.2	Représentation des différentes parties de l'émetteur contribuant à R_E (d'après [Schröter00]).	84
III.3	Résistances séries du transistor bipolaire.	85
III.4	Évolution de g_m/I_C en fonction de g_m pour l'extraction de la résistance d'émetteur. Taille du composant : $0.15 \times 3.6 \mu m^2$	85
III.5	Analyse SIMS de TBH ayant différents dopages d'émetteur (après recuit final).	88
III.6	Comparaison des résistances d'émetteur obtenues pour différents débits d'Arsine.	89
III.7	Influence de l'épaisseur du Si-cap sur le réglage de la jonction E/B.	90
III.8	Comparaison des résistances d'émetteur obtenues pour différents débits d'Arsine en fonction de l'épaisseur du Si-cap.	91
III.9	Tracés de Gummel de TBH de dimension $0.15 \times 3.6 \mu m^2$, pour différentes épaisseurs de Si-cap.	92
III.10	Présentation des deux types de contacts d'émetteur étudiés [Chevalier05a].	92
III.11	Tracés de Gummel de TBH de dimension $0.15 \times 3.6 \mu m^2$, dans le cas d'un émetteur de référence et d'un émetteur optimisé.	93
III.12	Évolution de la fréquence f_T en fonction de I_C . <i>Réf.</i> : Émetteur de référence (90 sccm, Si-cap de 12 nm); <i>Opt.</i> : Émetteur optimisé (250 sccm, Si-cap de 16 nm).	93
III.13	Évolution de f_T en fonction de la tension V_{BE} . <i>Réf.</i> : Émetteur de référence (90 sccm, Si-cap de 12 nm); <i>Opt.</i> : Émetteur optimisé (250 sccm, Si-cap de 16 nm).	93
III.14	Évolution de τ_{EC} en fonction du courant collecteur. Contact d'émetteur en matrice.	94
III.15	Évolution de τ_{EC} en fonction du courant collecteur. Contact d'émetteur en ruban.	94
III.16	Présentation des trois profils de base. La couche totale de SiGe a une épaisseur de 30 nm.	97
III.17	Courbes de Gummel et de Gain pour trois profils de Ge.	98
III.18	Evolution de f_T en fonction de I_C pour trois profils de Ge.	98
III.19	Temps de transit τ_{EC} en fonction de I_C pour différents profils de SiGe.	98
III.20	Variation de l'épaisseur de la base. Le profil de Ge retenu est un profil 20–30 % en trois marches successives.	99
III.21	Profils de Ge obtenus par SIMS pour différentes épaisseurs de base.	99
III.22	Courbes de Gummel et de Gain en fonction de l'épaisseur de SiGe.	100
III.23	Evolution de f_T en fonction de I_C pour différentes épaisseurs de SiGe.	100
III.24	Temps de transit τ_{EC} en fonction de I_C pour différentes épaisseurs de SiGe.	101
III.25	Courbes de Gummel et de Gain en fonction du profil de SiGe.	102
III.26	Comparaison des profils de bore obtenus par analyse SIMS en fonction du taux de Ge dans la base.	103
III.27	Evolution de f_T en fonction de V_{BE} pour différentes variantes technologiques sur la base et l'émetteur.	103
III.28	Introduction d'un nouveau profil de Ge pour la réalisation de bases de transistors bipolaires rapides.	104
III.29	Présentation des variations réalisées sur le profil de SiGe pour l'optimisation de la nouvelle base.	105
III.30	Décomposition de la résistance collecteur en trois parties distinctes.	107
III.31	Effet d'amélioration de la capacité C_{BC} par réduction de la diffusion latérale des dopants en passant du phosphore à l'arsenic pour le dopage du SIC.	108

III.32	Profils SIMS de collecteurs dopés phosphore ou arsenic. La couche enterrée est réalisée en As dans les deux cas. La dose d'As représentée est $6 \cdot 10^{13} \text{ at} \cdot \text{cm}^{-2}$. Le SIC phosphore est le SIC standard de référence.	109
III.33	Evolution de f_T en fonction de I_C pour différentes implantations SIC.	110
III.34	Temps de transit τ_{EC} en fonction de $1/I_C$ pour différents profils de SIC.	110
III.35	Evolution de f_T en fonction de BV_{CEO} pour des composants ayant différentes doses de SIC.	111
III.36	Représentation des différentes contributions à la résistance de base [Schröter00].	111
III.37	Dimensions latérales du transistor bipolaire.	112
III.38	Effet de la largeur de la fenêtre émetteur sur le contact métallique d'émetteur. .	113
III.39	Photos MET de deux TBH utilisant les deux types d'espaceurs internes.	114
III.40	Schéma représentant les deux types d'espaceurs (en L et en D). La capacité C_{BE} est plus faible dans le cas d'un espaceur en D.	114
III.41	Comparaison d'une topologie épaisse avec une topologie réduite.	115
III.42	Influence de la largeur de la zone active sur la capacité C_{BC}	116
III.43	Schéma expliquant l'influence de la largeur du polyémetteur sur la capacité C_{BE} et la résistance de base.	116
III.44	Evolution de f_T et f_{MAX} en fonction de I_C pour différentes largeurs de fenêtre émetteur W_E . Longueur des dispositifs : $3.6 \mu\text{m}$	118
III.45	Evolution de f_T et f_{MAX} en fonction de I_C pour différentes distances $d_{EmPoly-EmWin}$. Longueur des dispositifs : $3.6 \mu\text{m}$	118
III.46	Synthèse de l'évolution des temps de transit au cours de l'étude.	119
III.47	Evolution de f_T et f_{MAX} par l'optimisation conventionnelle du TBH. La ligne pointillée représente les points vérifiant la condition $f_{MAX} = f_T$	120
IV.1	Evolution de BV_{CEO} en fonction de f_T lors de l'optimisation conventionnelle du TBH.	121
IV.2	Variations de I_B en fonction de V_{CE} . Taille des composants : $0.15 \times 3.6 \mu\text{m}^2$. L'augmentation de I_B permet de retarder son inversion.	122
IV.3	Estimation de la tension de claquage en fonction du gain en courant du composant pour un dopage de collecteur fixé ($V_{BE} = 0.7 \text{ V}$, $M_f = 5.2$, $BVC = 3.5 \text{ V}$). .	123
IV.4	Schéma expliquant le principe de l'émetteur métallique.	124
IV.5	Comparaison d'un transistor standard avec l'architecture polyémetteur fin et l'architecture PRETCH.	125
IV.6	Procédé de fabrication PRETCH.	126
IV.7	Photo TEM en fausses couleurs présentant un transistor à émetteur métallique réalisé avec le procédé PRETCH. Vert : Oxygène (Oxyde), Rouge : Titane (Ti, TiN), Bleu : Azote (Nitrure), Noir : Silicium et Tungstène.	126
IV.8	Courbes de Gummel et de gain de transistors bipolaires à émetteur métallique de dimension $0.15 \times 3.6 \mu\text{m}^2$, en fonction de l'épaisseur de l'émetteur.	127
IV.9	Évolution de f_T en fonction de I_C pour différentes épaisseurs d'émetteur. Dimension des composants : $0.15 \times 3.6 \mu\text{m}^2$	128
IV.10	Profils de dopages obtenus dans le cas d'émetteurs de 40 et 80 nm d'épaisseur. .	128
IV.11	Effet de l'épaisseur et du dopage du polyémetteur sur le réglage de la jonction E/B.	129
IV.12	Évolution de la concentration de porteurs en fonction de la température pour plusieurs débit d'arsine.	129
IV.13	Interprétation des résultats dans le cas d'émetteurs fortement dopés. Lorsque l'émetteur est épais (a) ou fin (b), le gradient de trous varie peu.	130

IV.14 Effet attendu en réduisant le dopage du polyémetteur. Lorsque l'émetteur est faiblement dopé, le gradient de trous est très différent entre un émetteur épais (a) et un émetteur fortement dopé (b).	131
IV.15 Courbes de Gummel et de gain en courant de TBH à émetteur métallique ayant un dopage d'émetteur réduit. Épaisseur du Si-cap = 6 nm ; dimension des composants : $0.15 \times 3.6 \mu\text{m}^2$	131
IV.16 Observation TEM illustrant la perte du lien entre base intrinsèque et extrinsèque lorsque le Si-cap devient trop fin.	132
IV.17 Schéma expliquant le principe de l'émetteur SiGe.	134
IV.18 Profil de base réalisé. Un pic de Germanium de 5 nm de large est inséré dans le Si-cap.	135
IV.19 Courbes de Gummel et de gain en courant de transistors présentant un pic de Ge dans le Si-cap (dimensions $0.15 \times 3.6 \mu\text{m}^2$).	135
IV.20 Profil de base déposé pour augmenter la recombinaison en base neutre.	138
IV.21 Profils SIMS du carbone obtenus après recuit.	139
IV.22 Profils de bore obtenus après recuit pour les trois profils P1, P2 et P3, comparés à la plaque standard.	139
IV.23 Courbes de Gummel et de gain en courant de TBH avec recombinaison en base neutre. Dimension des composants : $0.15 \times 3.6 \mu\text{m}^2$	139
IV.24 Mise en évidence de la transition du courant de base pour le profil P3. Au delà de 0.8 V, I_B s'écarte de l'idéalité pour rejoindre le courant standard sans recombinaison.	140
IV.25 Tracés de Gummel entre 50 et 300 K du transistor ayant le profil de base P3. Dimensions $0.15 \times 3.6 \mu\text{m}^2$	141
IV.26 Évolution du gain en courant en fonction de I_C entre 50 et 300 K. Profil P3, dimensions $0.15 \times 3.6 \mu\text{m}^2$	141
IV.27 Tracés de Gummel à 50 K de TBH dotés des profils de base standard, P1, P2 et P3 ($0.15 \times 3.6 \mu\text{m}^2$).	142
IV.28 Densité de courant de base à la transition, en fonction de la dose de carbone ($0.15 \times 3.6 \mu\text{m}^2$).	142
IV.29 Évolutions du gain en courant en fonction de V_{BE} , pour les profils P2 et P3. Les simulations reproduisent la forme caractéristique du gain aux fortes tensions V_{BE}	142
IV.30 Taux d'occupation des pièges de la base neutre en fonction de V_{BE} (profil P3). Au delà de 0.8 V, le taux d'occupation augmente fortement.	142
IV.31 Spectres de bruit à $I_B = 500 \text{ nA}$ en fonction du profil de base. Dimension $0.15 \times 3.6 \mu\text{m}^2$	143
IV.32 Variation de bande interdite $\Delta E_{G,C}$ en fonction de la concentration de carbone.	144
IV.33 Concentration de carbone substitutionnel en fonction de la concentration de carbone totale.	144
IV.34 Evolution de la tension de claquage en fonction du gain en courant, pour l'émetteur métallique, l'émetteur SiGe et la recombinaison en base neutre.	146
IV.35 Positionnement des performances obtenues avec le procédé émetteur métallique, le pic de SiGe dans le Si-cap et la recombinaison en base neutre par rapport aux résultats standards.	146
IV.36 Positionnement des performances obtenues avec le procédé émetteur métallique, le pic de SiGe dans le Si-cap et la recombinaison en base neutre par rapport à la concurrence.	147

V.1	Simulation de courbes de f_T isothermes, et non-isotherme prenant en compte l'auto-échauffement. L'augmentation de température atteint 50 K au maximum de f_T	150
V.2	Réseau thermique simple.	151
V.3	Conductivité thermique du silicium en fonction de la température.	151
V.4	Evolution de la tension V_{BE} avec la température du banc. $I_E = 3.18$ mA et $V_{CB} = 0$ V.	152
V.5	Evolution de la tension V_{BE} avec la tension V_{CB} et la puissance thermique dissipée. $I_E = 3.18$ mA.	152
V.6	Variation de T en fonction de la puissance thermique. Une correction est nécessaire pour tenir compte du décalage en température existant entre le substrat et le transistor en fonctionnement.	153
V.7	Evolution de la résistance thermique en fonction de la longueur d'émetteur. La largeur d'émetteur est $0.14 \mu\text{m}$	155
V.8	Répartitions de température dans le substrat en fonction de la profondeur des DTI.	156
V.9	Évolution de la résistance thermique en fonction de la profondeur des tranchées profondes, pour une longueur d'émetteur comprise entre 0.6 et $15 \mu\text{m}$	157
V.10	Température moyenne en fonction de la puissance thermique appliquée, pour plusieurs profondeurs de DTI ($0.15 \times 3.6 \mu\text{m}^2$, contact émetteur ruban).	158
V.11	Présentation des deux types de fractionnement possibles.	159
V.12	Simulation thermiques sous Ansys simulant deux types de fractionnement d'émetteur (structure faible coût).	160
V.13	Simulations 3D sous Ansys donnant la température en fonction de la position dans le composant, pour deux fractionnements d'émetteur différents.	161
V.14	Comparaison entre substrat massif et SOI des températures moyennes en fonction de la puissance thermique appliquée, pour plusieurs fractionnements d'émetteur.	162
V.15	Vue en coupe du transistor bipolaire. $z_1 = 0.7 \mu\text{m}$ et $z_2 = 2 \mu\text{m}$	163
V.16	Répartition du flux thermique dans le substrat. Les DTI sont considérés comme des parois adiabatiques. R_{Th} est divisé en 4 parties distinctes.	163
V.17	Comparaison des valeurs de R_{Th} modélisées et mesurées en fonction de la longueur d'émetteur (Largeur d'émetteur : $0.15 \mu\text{m}$, profondeur des DTI = $5.2 \mu\text{m}$).	164
V.18	Comparaison des valeurs de R_{Th} modélisées et mesurées en fonction de la longueur d'émetteur (Largeur d'émetteur : $0.15 \mu\text{m}$).	164
V.19	Valeurs de R_{Th} et de $R_{Th} \times A_E$ pour plusieurs largeurs d'émetteur en fonction de L_E	165
V.20	Evolution des courbes de Gummel en fonction de la température (Composant 1).	167
V.21	Évolution du gain en courant avec la température (Composant 1).	168
V.22	Évolution des résistances séries du composant 1 avec la température.	168
V.23	Évolution des résistances séries des composants 2A, 2B et 2C avec la température.	169
V.24	Évolution des fréquences f_T et f_{MAX} du composant 1 avec T ($V_{CB} = 0.5$ V).	170
V.25	Évolution des fréquences f_T et f_{MAX} du composant 2A avec T ($V_{CB} = 0.5$ V).	170
V.26	Évolution des fréquences f_T et f_{MAX} du composant 2B avec T ($V_{CB} = 0.5$ V).	170
V.27	Évolution des fréquences f_T et f_{MAX} du composant 2C avec T ($V_{CB} = 0.5$ V).	171
V.28	Évolution de τ_{EC} en fonction de $1/I_C$ (composant 1, 300 K, $V_{CB} = 0.5$ V).	173
V.29	Répartitions des différentes contributions au temps de transit en fonction la température.	173

Liste des tableaux

I.1	Variations du temps de transit τ_{EC} et de la fréquence f_T maximale associée avec la température (dimensions du composant : $0.15 \times 3.6 \mu\text{m}^2$).	48
III.1	Valeurs de R_E et de $n_0 V_{Th}$ extraites en fonction de la plage de V_{BE} utilisée pour la régression. La température du substrat est fixée à 300 K. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$	86
III.2	Comparaison des valeurs de R_E (en Ω) extraites manuellement et lors des mesures HF. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$	87
III.3	Comparaison des caractéristiques statiques de TBH ayant des émetteur dopés As ou P. L'émetteur phosphore n'est pas recuit. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$	88
III.4	Évolution des valeurs de R_E à réglage de jonction équivalent. L'épaisseur du Si-cap est adaptée au niveau de dopage de l'émetteur. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$	89
III.5	Influence de l'épaisseur du Si-cap sur les caractéristiques de la jonction E/B et sur la résistance d'émetteur. Taille du composant : $0.15 \times 3.6 \mu\text{m}^2$	91
III.6	Principaux paramètres électriques de TBH avec émetteur de référence et émetteur optimisé.	95
III.7	Évolution des principaux paramètres électriques avec le taux de Ge à l'entrée de la base (épaisseur de base : 30 nm).	97
III.8	Évolution des principaux paramètres électriques avec l'épaisseur de la couche de SiGe (profil de SiGe : 20–30 %).	100
III.9	Évolution des principaux paramètres électriques avec le profil de base (épaisseur de SiGe : 20 nm).	102
III.10	Comparaison des principaux paramètres électriques de TBH ayant le profil de base 20–30 de référence et les nouveaux profils de base 10–25 % ($0.15 \times 3.6 \mu\text{m}^2$).	104
III.11	Influence du dopage SIC sur les performances du composant ($0.15 \times 3.6 \mu\text{m}^2$).	108
III.12	Paramètres électriques de transistors bipolaires en fonction de la largeur d'émetteur définie par la photolithographie ($L_E = 3.6 \mu\text{m}$).	113
III.13	Evolution des performances électriques en fonction du type d'espaceur pour deux largeurs de fenêtre émetteur ($L_E = 3.6 \mu\text{m}$).	115
III.14	Performances dynamiques en fonction de la distance STI-EmWin. Dimensions des transistors : $0.15 \times 3.6 \mu\text{m}^2$	116
III.15	Performances dynamiques en fonction de la distance STI-EmWin. Dimensions des transistors : $0.15 \times 3.6 \mu\text{m}^2$	117
IV.1	Principaux paramètres électriques en fonction de l'épaisseur du polyémetteur ($0.15 \times 3.6 \mu\text{m}^2$).	127
IV.2	Principaux paramètres électriques de transistors à émetteur métallique ayant des dopages d'émetteur réduits.	132

IV.3	Paramètres électriques statiques et dynamiques obtenus en fonction du profil de Germanium dans le Si-cap ($0.15 \times 3.6 \mu\text{m}^2$).	136
IV.4	Résumé des caractéristiques des différents profils de carbone réalisés afin d'augmenter la recombinaison en base neutre.	139
IV.5	Principaux paramètres électriques obtenus en fonction du profil de carbone ($0.15 \times 3.6 \mu\text{m}^2$).	140
IV.6	Variations de bande interdite pour les différents profils de carbone. La réduction due au dopage est prise en compte.	144
V.1	Comparaison des valeurs R_{Th} extraites en régime statique et pulsé pour différentes géométries de transistors.	154
V.2	Évolution des paramètres dynamiques en fonction de la longueur d'émetteur. Largeur du composant : $0.14 \mu\text{m}$, contacts en matrice.	155
V.3	Influence de la profondeur des tranchées sur les caractéristiques des transistors (taille de composant $0.15 \times 3.6 \mu\text{m}^2$, contact émetteur ruban).	157
V.4	Simulations électrothermiques 2D visant à étudier l'impact des tranchées sur les performances du composant.	157
V.5	Valeurs de R_{Th} et principaux paramètres dynamiques en fonction de la distance entre les DTI et la zone active. Transistors CBEBC, $0.15 \times 9.7 \mu\text{m}^2$ (R_{Th}), $0.15 \times 3.6 \mu\text{m}^2$ (f_T et f_{MAX}).	158
V.6	Valeurs de R_{Th} pour divers fractionnements d'émetteur (type 1).	159
V.7	Valeurs de R_{Th} pour différents fractionnements d'émetteur, en fonction du type de substrat (type 2).	160
V.8	Résistances thermiques et principaux paramètres statiques et dynamiques mesurés en fonction du fractionnement de l'émetteur.	161
V.9	Résistances thermiques extraites sur substrat massif ou SOI	161
V.10	Dénomination des composants étudiés en hyperfréquence à des températures cryogéniques.	166
V.11	Fréquences f_T maximales (en GHz) obtenues en fonction de la température. Δf_T est la variation de f_T entre 300 K et 50 K.	171
V.12	Fréquences f_{MAX} maximales (en GHz) obtenues en fonction de la température. Δf_{MAX} est la variation de f_{MAX} entre 300 K et 50 K.	171
V.13	Comparaison des temps de transit $\tau_{EC Min}$ et τ_F à 300 K et 50 K.	174

Résumé :

Cette thèse a pour objet l'étude et la réalisation de transistors bipolaires à hétérojonction Si/SiGeC à très hautes performances pour les applications de télécommunications et de détection radar (> 50 GHz). Le premier chapitre rappelle la théorie du fonctionnement du transistor bipolaire, aussi bien en régime statique que dynamique. Nous présentons ensuite dans un second chapitre un historique des technologies de fabrication des TBH SiGe, spécifiquement dans un cadre de compatibilité BiCMOS, en expliquant les choix pris en faveur d'un auto-alignement complet et d'une épitaxie sélective de la base. Les résultats obtenus dans cette thèse sont comparés aux performances de l'état de l'art. Nous étudions ensuite l'optimisation classique du TBH, et nous démontrons comment, par différentes optimisations du profil vertical et des dimensions latérales du composant nous avons pu augmenter les fréquences f_T et f_{MAX} de 200 GHz au début de la thèse jusqu'à environ 300 GHz. La possibilité d'optimiser la tenue en tension collecteur est démontrée par des procédés technologiques en rupture avec l'optimisation classique du transistor, tels que l'émetteur métallique, l'insertion de Ge dans l'émetteur et la recombinaison en base neutre. Des améliorations significatives de BV_{CEO} ont pu être démontrées grâce à ces procédés, ce qui permet d'obtenir des produits $f_T \times BV_{CEO}$ à l'état de l'art (> 400 GHz·V). La dernière partie de la thèse vise à étudier le comportement du TBH en fonction de la température. Nous décrivons tout d'abord l'auto-échauffement du transistor, et son impact sur les performances dynamiques. Des variantes technologiques permettant de réduire l'auto-échauffement sont étudiées (profondeur des DTI et fragmentation de l'émetteur), et un modèle simple de calcul de R_{Th} est développé. Nous terminons par l'étude du TBH aux températures cryogéniques. f_T et f_{MAX} étant fortement améliorées à basse température (plus de 400 GHz), nous en tirons, grâce à l'extraction des différents temps de transit, des perspectives intéressantes pour l'amélioration ultérieure du composant, avec pour objectif d'atteindre des fréquences de transition de l'ordre du THz.

Abstract :

The purpose of this thesis is the study and optimisation of high performance Si/SiGeC heterojunction bipolar transistors for telecommunication and radar detection applications (> 50 GHz). The first chapter is a reminder of the bipolar transistor theory, both in its static and dynamic operations. We present in the second chapter an history of the SiGe HBTs manufacturing technologies, explaining the choices leading to a fully self-aligned structure with selective epitaxial growth of the base. Results obtained in this thesis are compared with the state-of-the-art performances published in literature. Next, we study the classical optimisation of the HBT and we demonstrate how, by several optimisations made on the vertical profile and the lateral dimensions of the device, we were able to push up the f_T and f_{MAX} frequencies from 200 GHz at the beginning of the thesis up to 300 GHz. The possibility to increase the collector to emitter breakdown voltage BV_{CEO} is demonstrated using innovative process features, such as metallic emitter, Ge insertion in the emitter, and neutral base recombination. Significant improvements of BV_{CEO} are demonstrated, which enables state-of-the-art $f_T \times BV_{CEO}$ products (> 400 GHz·V). The study of the temperature behaviour of the HBT is presented in the last part of the thesis. First, we describe the self-heating of the transistor, and its impact on dynamic performances. Technological variations enable the reduction of the self-heating, such as DTI depth diminution and emitter fragmentation, and a simple model for R_{Th} calculation is developed. Finally, the study of the HBT at cryogenic temperatures completes this thesis. Since f_T and f_{MAX} are strongly improved at low temperature (more than 400 GHz), we can deduct from the extraction of the different transit times interesting perspectives for further optimisation of the device, with the goal to reach transition frequencies around 1 THz.