



HAL
open science

**Techniques de conception des circuits intégrés
analogiques pour des applications en haute température,
en technologies de conception des circuits intégrés
analogiques pour des applications en haute température,
en technologies sur substrat de silicium**

R.A. Bianchi

► **To cite this version:**

R.A. Bianchi. Techniques de conception des circuits intégrés analogiques pour des applications en haute température, en technologies de conception des circuits intégrés analogiques pour des applications en haute température, en technologies sur substrat de silicium. Micro et nanotechnologies/Microélectronique. Institut National Polytechnique de Grenoble - INPG, 1999. Français. NNT : . tel-00011575

HAL Id: tel-00011575

<https://theses.hal.science/tel-00011575>

Submitted on 9 Feb 2006

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.

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

THÈSE

pour obtenir le grade de

DOCTEUR de l'Institut National Polytechnique de Grenoble

Discipline: Microélectronique

***Techniques de conception des circuits intégrés analogiques pour
des applications en haute température, en technologies sur
substrat de silicium***

Présentée et soutenue publiquement par

Raúl-Andrés BIANCHI

le 21 octobre 1999

Composition du jury:

Messieurs: Pierre GENTIL

Yves DANTO

Jean-Pierre CHANTE

Bernard COURTOIS

Jean-Michel KARAM

Jean-Michel FOURNIER

Eric DONZIER

Président

Rapporteur

Rapporteur

Directeur

Examineur

Examineur

Examineur

Thèse préparée au Laboratoire TIMA-INPG

Remerciements

Je tiens à exprimer ma reconnaissance à,

Monsieur Pierre Gentil, Directeur du Collège Doctoral de l'INPG, pour l'honneur qu'il m'a fait en acceptant de présider le jury de cette thèse et pour m'avoir donné, lors de mon arrivée en France, la possibilité de me spécialiser en microélectronique, à Grenoble.

Messieurs Jean-Pierre Chante, Directeur du Laboratoire CEGELY - INSA-Lyon, et Yves Danto, Chercheur au laboratoire IXL - Université Bordeaux 1, pour avoir accepté d'être rapporteurs de cette thèse.

Messieurs Bernard Courtois, Directeur du laboratoire TIMA, et Jean Michel Karam, Président de la société MEMSCAP S.A. et Responsable du groupe Microsystèmes (MCS) jusqu'en 1998, pour m'avoir donné la possibilité de faire ce travail.

Monsieur Eric Donzier, Section Manager de la société SCHLUMBERGER, Messieurs Patrick Perciot et Pierre-Marie Petit et Mlle Sandrine Lelong, Ingénieurs de la société SCHLUMBERGER pour tout leur soutien pendant le déroulement du projet MIPEG.

Monsieur Francis Presseccq, Ingénieur du Centre National d'Etudes Spatiales pour sa précieuse collaboration dans le test des capteurs de température intégrés développés pendant cette thèse.

Membres et ex-membres du groupe MCS-TIMA, notamment Jérôme Goy, Zein Juneidi et Filipe Vinci dos Santos avec lesquels j'ai eu des échanges fructueux.

Madame Danuta Metz et Mlle Delphine Vuillerme pour leur patience à réviser le français de cette thèse.

Au personnel du Laboratoire TIMA et du Laboratoire CIME, notamment Monsieur Alexandre Chagoya, pour leur disponibilité et leur support technique au cours de ces années d'étude.

Messieurs Oscar Sartori, ex-Doyen de la faculté de génie de l'Universidad Católica de Córdoba, Aldo Brunetti, Chercheur au CONICET et ex-Doyen du FaMAF, Ernesto Perea, Group Manager à STMicroelectronics et Carlos Marques, Ingénieur et Enseignant à l'UCC, pour leur soutien.

Et finalement, je suis spécialement reconnaissant envers ma famille et mes amis, qui ont toujours été présents.

Résumé

Cette thèse se situe dans le domaine de la microélectronique à haute température, et concerne particulièrement les applications industrielles à bas coût et grand volume de production. Actuellement les technologies de circuits intégrés plus poussées, en termes de densité d'intégration, coût, et fiabilité, utilisent encore le silicium comme matériau de base. Ces technologies sont conçues pour permettre une large durée de vie, dans une plage de température donnée, typiquement de 0 à 100°C. D'autres technologies se développent aussi, notamment le *SiC* (de plus large bande interdite) et le SOI couche mince (où la surface des jonctions parasites sont fortement réduites), pour satisfaire les besoins des applications à des très hautes températures. Ayant actuellement des performances inférieures, en termes de coût et de densité d'intégration, elles restent beaucoup moins compétitives que les technologies standards sur substrat de silicium. De plus, il est prévu que, dans les prochains dix ans, plus de 70% des applications en haute température correspondront encore à des applications automobiles et pétrolières ayant des températures d'opération intermédiaires, inférieures à 200°C.

A partir de l'étude de la physique des composants semiconducteurs et des matériaux pour la microélectronique, cette thèse élargit jusqu'à environ 250°C la plage de température d'utilisation des technologies CMOS et BiCMOS standards, sur substrat de silicium, à travers des techniques de conception de circuits intégrés, sans toutefois modifier les procédés de fabrication. Les études et les tests ont été expérimentés sur une technologie CMOS et une technologie BiCMOS commerciales. Ces conclusions sont transférables à toute technologie semblable actuelle. De plus, les performances en température tendent à s'améliorer dans le futur, du fait que l'augmentation de la densité d'intégration entraîne une augmentation des concentrations des dopants et une réduction de la surface des jonctions d'isolation.

Deux applications industrielles, représentatives du marché potentiel des applications des circuits intégrés à haute température, ont permis de vérifier dans la pratique les résultats théoriques obtenus.

Abstract

This Ph.D. report deals with high temperature microelectronics, for low-cost and high-volume applications. Present day most advanced microelectronic technologies, in terms of density, cost and reliability, still use silicon substrates. These technologies are designed to obtain long MTF (Mean Time to Failure) at normal temperatures: 0 to 100°C, typically. Other technologies, such as Wide Bandgap Semiconductors (SiC, Diamond, etc.) and thin film SOI, are under development to provide good performance at much higher temperatures (>250°C). But, since they presently have lower performances in terms of density and cost, they are less competitive than present day silicon-substrate based technologies. In addition, it is expected for the next ten years that more than 70% of high temperature applications will still correspond to automotive and oil prospecting applications, having intermediate operating temperatures, lower than 200°C.

Based on semiconductor devices and microelectronics materials physics, this work will increase, up to 250°C, the operation temperature range of silicon CMOS and BiCMOS standard technologies, by means of integrated circuit design techniques (no process modifications are added, to keep the low-cost feature). The experiments were done using a commercial CMOS and BiCMOS technology; the obtained results can be applied to any similar technology. In addition, the high temperature performance seems to ameliorate in future technologies, since the density increase involves a doping concentration increase and the reduction of the isolation junctions' dimensions.

Two industrial applications, which are good examples of the potential market for high temperature applications, helped to experimentally verify the obtained theoretical results.

Table des matières

1. Introduction	7
2. Physique des matériaux en Haute Température	11
2.1. Silicium	11
2.1.1. Structure cristalline du Si	11
2.1.2. Bandes d'énergie du Si	12
2.1.3. Fonction de Fermi	14
2.1.4. Concentration intrinsèque	15
2.1.5. Semiconducteur extrinsèque	17
2.1.6. Température intrinsèque	18
2.1.7. Courants dans le semiconducteur	19
2.1.7.1. Courant de conduction	20
2.1.7.2. Courant de diffusion	22
2.1.7.3. Courant total	22
2.1.8. Génération et recombinaison des porteurs	22
2.1.8.1. Semiconducteur de type n	23
2.1.8.2. Zone désertée	23
2.1.9. Conductivité thermique du silicium	24
2.1.9.1. Modèles de dissipation de la chaleur	24
2.2. Couches conductrices et contacts	25
2.2.1. Métallisations	26
2.2.1.1. Electromigration	26
2.2.1.1.a. Aluminium	28
2.2.1.1.b. Alliages Al-Cu	28
2.2.1.2. Corrosion	28
2.2.2. Silicium poly-cristallin	30
2.3. Isolation	31
2.3.1. Dioxyde de Silicium	31
2.3.1.1. Injection d'électrons chauds	32
2.3.1.2. Transport de charge dans les diélectriques	33
2.3.1.3. Fiabilité des diélectriques	33
2.3.2. Jonctions pn	34
2.3.2.1. La jonction pn à l'équilibre	34
2.3.2.2. La jonction pn en dehors de l'équilibre	36
2.3.2.3. Densités de porteurs dans les zones neutres et densité idéale de courant.	37
2.3.2.4. Génération et recombinaison dans la ZCE	38
2.3.2.5. Courants de fuite	38
2.3.2.6. Tension de rupture	40
2.3.2.6.a. Effet Zener et effet d'avalanche	41
2.3.2.7. Capacité parasite	42
2.3.2.8. Latchup	44
2.4. Conclusions	44
3. Physique des dispositifs en Haute Température	46
3.1. Résistances	46
3.1.1. Résistances en silicium monocristallin	46
3.1.2. Résistances en silicium polycristallin	48
3.2. Capacités	49
3.2.1. Capacités de jonction pn	49
3.2.2. Capacités MOS	49

3.2.2.1. Capacité MOS en accumulation	52
3.2.3. Capacités faites en Double Poly-Si	54
3.3. Diodes	54
3.4. Transistors MOS	55
3.4.1. Tension de seuil	57
3.4.2. Courant de conduction en inversion forte	59
3.4.2.1. Régime linéaire	61
3.4.2.2. Régime de saturation	61
3.4.2.2.a. Conductance drain-source des transistors MOS en régime de saturation	62
3.4.3. Mobilité	63
3.4.3.1. Mobilité effective dans la couche d'inversion	65
3.4.3.2. Saturation de la vitesse de dérive	66
3.4.4. Courant drain-source en fonction de la température	67
3.4.5. Régime de faible inversion des transistors MOS en haute température.	68
3.4.6. Courants de fuite et conductances dans le drain et la source.	69
3.4.6.1. Courants de fuite et conductance dans les transistors NMOS	70
3.4.6.2. Courants de fuite et conductance dans les transistors PMOS	71
3.4.7. Courants de fuite et conductance dans le caisson des PMOS	72
3.5. Transistors bipolaires	73
3.5.1. Transistors bipolaires verticaux	73
3.5.1.1. Les courants dans la Base et le Collecteur	73
3.5.1.2. Gain de courant	76
3.5.1.3. Transconductance	77
3.5.1.4. Modulation de la largeur effective de la base	77
3.5.1.5. Courant de fuite et conductance dans la jonction collecteur-substrat	78
3.5.2. Transistors bipolaires latéraux CLBT	80
3.5.2.1. Modèle du courant du collecteur latéral	80
3.5.2.1.a. Fonction analytique approximative de f_g	85
3.5.2.2. Courant du collecteur vertical	85
3.5.2.3. Gain de courant et conductance des collecteurs latéral et vertical	85
3.5.2.4. Courants de fuite dans le collecteur latéral	87
3.6. Conclusions	87
4. Techniques d'endurcissement pour la haute température	88
4.1. Techniques de conception	88
4.1.1. Tolérances des paramètres des dispositifs	88
4.1.2. Dégradation des composants en fonction de la température	89
4.1.3. Architectures robustes	91
4.1.3.1. Compensations en boucle ouverte	92
4.1.3.1.a. Sources proportionnelles à la température avec un coefficient positif	92
4.1.3.1.b. Sources proportionnelles à la température avec un coefficient négatif	93
4.1.3.1.c. Source de courant et de tension avec un coefficient thermique nul	95
4.1.3.2. Compensations par boucles de contrôle	96
4.1.3.3. Sous-circuits analogiques robustes	96
4.1.4. Modèles de simulation	97
4.1.4.1. Transistors MOS	98
4.1.4.1.a. Modèles fournis par le fabricant	98
4.1.4.1.b. Caractérisation et extraction des paramètres	100
4.1.4.2. Transistors bipolaires	101
4.1.4.2.a. Modèles fournis par le fabricant	101
4.1.4.2.b. Caractérisation et extraction des paramètres	101
4.2. Techniques d'implémentation	102
4.2.1.1. Maximisation de la précision relative	102
4.2.1.2. Minimisation des courants de fuite	103
4.2.1.3. Robustesse aux <i>latchups</i>	104

4.3. Fiabilité	105
5. Applications	106
5.1. High Temperature Oil Well Application : Miniature Permanent Gauge (MIPEG) project	107
5.1.1. The ASIC_MIPEG	108
5.1.1.1. The sensors interface	110
5.1.1.1.a. Oscillators structure	110
5.1.1.1.b. Implemented circuits	116
5.1.1.1.c. Mixer structure	127
5.1.1.1.d. Filters and Buffers	128
5.1.1.1.e. Operational Amplifier	129
5.1.1.1.f. Fabricated circuit and measured results	130
5.2. Automotive application	131
5.2.1. CMOS Compatible Temperature Sensors	132
5.2.1.1. The possible sensitive elements	132
5.2.1.1.a. Resistors.	133
5.2.1.1.b. Thermal diffusion constant of silicon.	133
5.2.1.1.c. MOS transistors.	133
5.2.1.1.d. Bipolar transistors and diodes.	133
5.2.1.2. Temperature sensor with an analog output signal	134
5.2.1.2.a. The sensing principle	134
5.2.1.2.b. Implemented Circuit	137
5.2.1.2.c. Measurement results	138
5.2.1.3. Sensor interfaces and digital interfaces	140
5.2.1.3.a. Sensor interfaces	140
5.2.1.3.b. Digital interfaces	141
5.2.1.3.c. Suitable AD Converters	141
5.2.1.4. Temperature sensor with a digital output signal	142
5.2.1.4.a. The sensing principle	142
5.2.1.4.b. The Vbe signal	143
5.2.1.4.c. Intrinsically referenced temperature signal $I_t(T)$	144
5.2.1.4.d. The reference signal $I_r(T)$	144
5.2.1.4.e. The output digital signal DS	145
5.2.1.4.f. The implemented circuit	146
5.2.1.4.g. Measurement results	148
6. Conclusions	151
7. Références	152

1. Introduction

Cette thèse se situe dans le domaine de la microélectronique à haute température, notamment celui des applications dont les températures d'opération dépassent les 125°C (température maximale de la plupart des composants commerciaux). Le marché mondial de l'électronique et de la microélectronique à haute température est en pleine croissance, représentant un pourcentage de 1.5% du marché potentiel mondial en 1994 et estimé à 6% du marché potentiel pour 2005 [67]. Les applications les plus importantes proviennent de l'industrie de l'automobile et, en second lieu, des industries pétrolière et aérospatiale (voir la Figure 1).

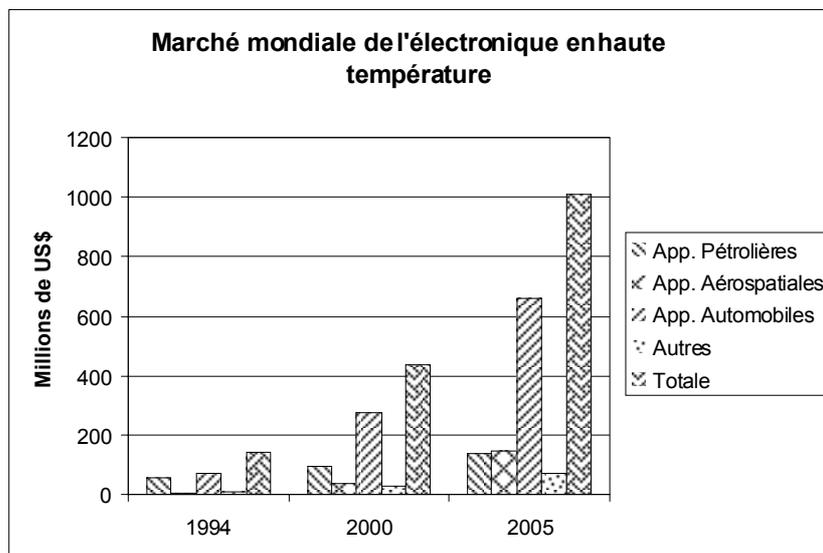


Figure 1. Marché mondial de l'électronique à haute température [67].

a. Industrie pétrolière

L'industrie de l'exploitation pétrolière a été la force qui a supporté et financé le développement de l'électronique à haute température pendant les vingt dernières années. Différents types de capteurs et de systèmes d'acquisition des données sont utilisés pour mesurer les caractéristiques de l'environnement lors du perçage et l'exploitation des puits pétroliers. Actuellement, 82% des applications pétrolières ont des températures d'opération inférieures à 200°C, tandis que les 18% restant nécessitent une électronique résistante à 260°C. La génération actuelle d'équipements pour l'exploration pétrolière utilise des technologies de semiconducteurs sur silicium ou SOI exclusivement. Les composants en diamant, *SiC* (*Carbure de Silicium*) et d'autres *WBS* (*Wide Bandgap Semiconductors*) seront nécessaires pour l'exploitation des puits géothermiques (350 à 400°C).

b. Industrie aérospatiale

En ce qui concerne l'industrie aérospatiale, l'électronique à haute température pourrait améliorer sensiblement la qualité (performance, poids, et encombrement), le coût, et la fiabilité des systèmes de contrôle déjà existants. L'environnement le plus contraignant est celui des moteurs d'avions (jusqu'à

250°C), où l'électronique de contrôle fonctionne à des températures inférieures à 105°C, grâce aux systèmes de refroidissement et aux dissipateurs de chaleur. L'implémentation des capteurs, et d'une électronique résistant aux hautes températures permettrait d'éliminer, ou de réduire, ces lourds et encombrants systèmes de refroidissement. Les technologies des semiconducteurs sur silicium et SOI seraient utilisées pour l'implémentation des fonctions les plus complexes (i.e. circuits intégrés analogiques, mémoires, microcontrôleurs, etc.) à des températures respectivement inférieures à 250°C et 300°C. Les capteurs et actuateurs, qui fonctionnent à plus de 300°C, utiliseront des technologies *WBS*.

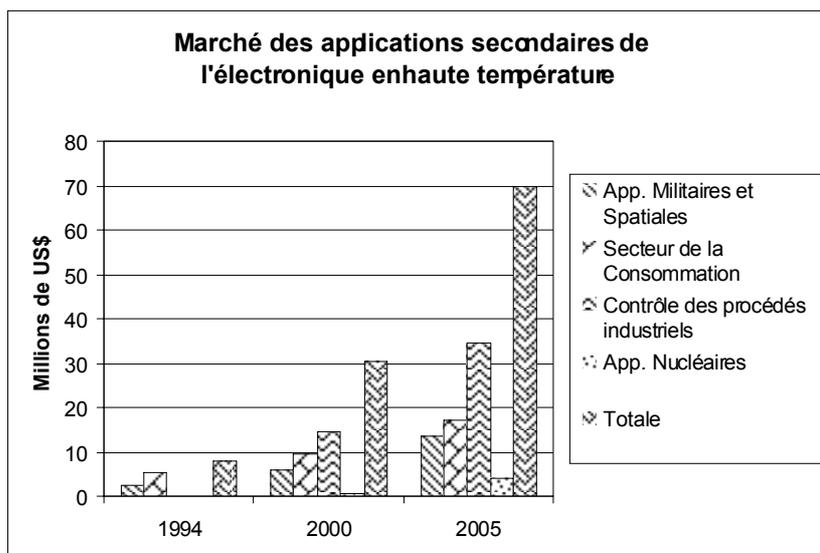


Figure 2. Marché des applications secondaires à haute température [67].

c. Industrie automobile

L'industrie automobile constitue le marché potentiel le plus large, en ce qui concerne l'électronique à haute température. Cependant le marché des applications automobiles pourrait être le plus difficile à pénétrer : La durée des cycles de conception des circuits et microsystèmes intégrés est incompatible avec l'évolution des spécifications de plus en plus exigeantes imposées par l'industrie automobile, en termes de coût, performance et fiabilité à long terme.

Les composants pour l'industrie automobile ont normalement des températures de fonctionnement inférieures à 125°C. Les systèmes de contrôle du moteur, exception faite de l'*ABS (Antilock Braking Systems)* et de la direction assistée, nécessitent une électronique résistant à 180°C. Mais, il devient de plus en plus difficile de maintenir la température sous le capot à des valeurs inférieures à 125°C, car l'aérodynamique impose des contraintes de réduction d'espace sous le capot, et les matériaux métalliques lourds sont remplacés par des matériaux composites plus légers, mais moins conducteurs de chaleur.

Les cinq applications automobiles les plus importantes pour l'électronique à haute température sont le contrôle du moteur, celui de la transmission, les *ABS*, les suspensions actives, et les capteurs pour les systèmes de contrôle de la vitesse et de la pression dans les pneus. La plupart des applications automobiles seront implémentées avec des technologies capables de fonctionner à 250°C maximum. L'insertion des technologies *WBS* sera très lente (3% vers le 2005) du fait des fortes contraintes de coût et de fiabilité à long terme, exigences difficiles à remplir par ces technologies émergentes, mais pas encore assez matures.

d. Autres applications

Un bon nombre d'applications dans d'autres secteurs de l'industrie, pouvant bénéficier des technologies des semiconducteurs en haute température, n'ont pas encore des programmes actifs de développement, ou correspondent à des petites niches du marché demandant de faibles volumes de production (voire la Figure 2). Le secteur potentiellement le plus représentatif est celui des procédés industriels (chimiques, métallurgiques, etc.), où l'électronique à haute température pourrait servir à améliorer la sécurité, la fiabilité et le contrôle des procédés de fabrication. Dans le secteur du nucléaire, le marché est très réduit, mais les bénéfices potentiels de l'utilisation des technologies résistantes aux hautes températures seraient pourtant très importants.

Il en résulte que la majorité des applications actuelles de l'électronique et la microélectronique à haute température correspondent à des températures inférieures à 200°C (voir la Figure 3) et à des technologies de circuits intégrés sur silicium et SOI (voir la Figure 4). D'après les prévisions, cette affirmation continuera à être valable dans les dix prochains années, puisque l'insertion des nouvelles technologies des semiconducteurs dans les différents marchés de la haute température est aujourd'hui très lente, ce qui, semble-t-il, ne devrait pas changer dans un avenir proche.

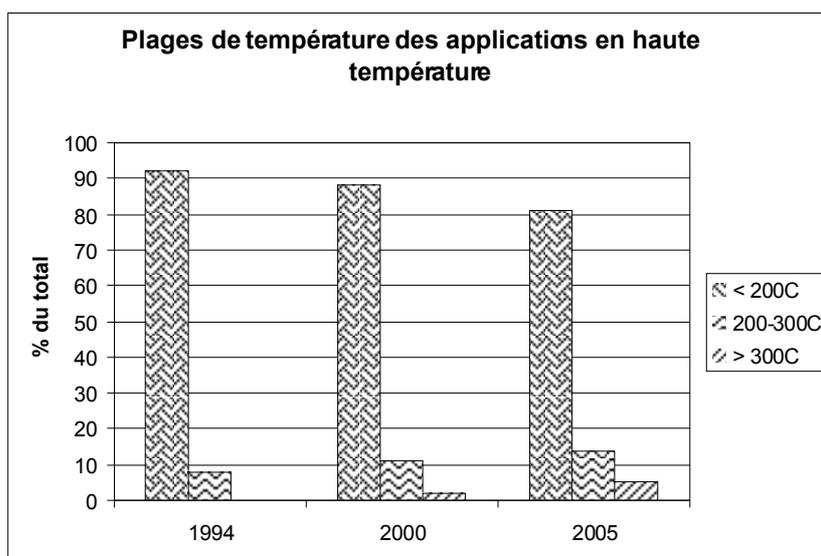


Figure 3. Plages de température des applications en haute température [67].

Les technologies CMOS et BiCMOS sur substrat de silicium sont encore les technologies de circuits intégrés les plus compétitives, en termes de densité d'intégration, coût, et fiabilité. Ces technologies sont conçues pour permettre une large durée de vie, dans une plage de température donnée, généralement située entre 0 et 100°C. D'autres technologies se développent aussi, notamment le SiC (de plus large bande interdite) et le SOI couche mince (où la surface des jonctions parasites sont fortement réduites), pour satisfaire les besoins des applications à très hautes températures (>300°C). Ayant actuellement des performances inférieures, en termes de coût et de densité d'intégration, elles restent beaucoup moins compétitives que les technologies standards sur substrat de silicium. De plus, il est prévu que, dans les prochains dix ans, plus de 70% des applications en haute température correspondront encore à des applications automobiles et pétrolières ayant des températures d'opération intermédiaires inférieures à 200°C.

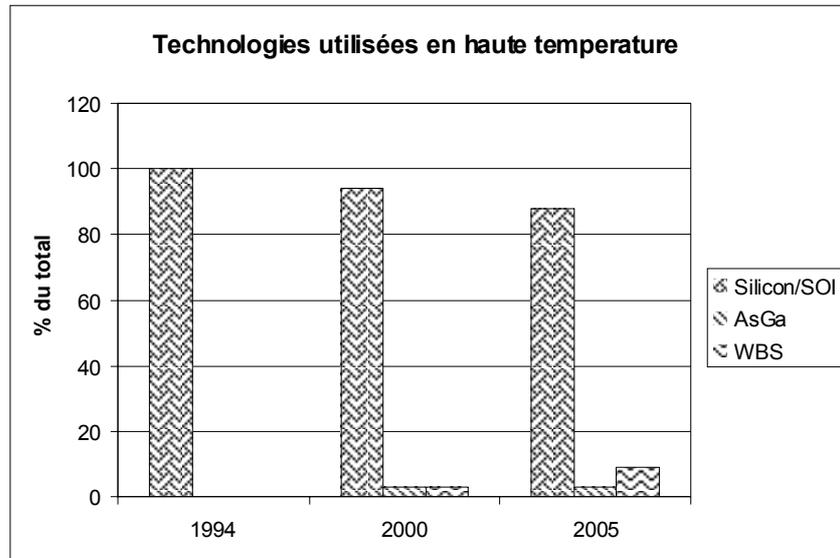


Figure 4. Technologies utilisées dans les applications en haute température [67].

A partir de l'étude de la physique des composants semiconducteurs (chapitre 3) et des matériaux pour la microélectronique (chapitre 2), cette thèse élargit jusqu'à environ 250°C la plage de température d'utilisation des technologies CMOS et BiCMOS standards, sur substrat de silicium, à travers des techniques de conception de circuits intégrés (chapitre 4), sans toutefois modifier les procédés de fabrication. Les études et les tests ont été expérimentés sur une technologie CMOS et une technologie BiCMOS commerciales. Ces conclusions sont transférables à toute technologie actuelle semblable. De plus, les performances en température tendent à s'améliorer dans le futur, en terme général, car l'augmentation de la densité d'intégration entraîne une augmentation des concentrations des dopants et une réduction de la surface des jonctions d'isolation.

Deux applications industrielles (chapitre 5), représentatives du marché potentiel des applications des circuits intégrés à haute température, ont permis de vérifier dans la pratique les résultats théoriques obtenus.

2. Physique des matériaux en Haute Température

Le comportement en haute température des matériaux conducteurs, semiconducteurs et isolants, utilisés pour la fabrication des circuits intégrés, est le premier facteur déterminant les caractéristiques en haute température des composants et des circuits réalisés avec une technologie donnée. Les propriétés de ces matériaux de base doivent être étudiées, afin de pouvoir les exploiter et ainsi améliorer la tenue en température des composants et des circuits intégrés.

2.1. Silicium

Actuellement les technologies de circuits intégrés utilisent encore le silicium comme matériau semiconducteur de base. Le silicium étant un matériau semiconducteur de largeur de bande interdite moyenne s'adapte bien aux contraintes imposées par les applications à des températures normales, soit de 0 à 100°C. A de plus hautes températures, la performance des technologies sur silicium se dégrade progressivement.

2.1.1. Structure cristalline du Si

La matière condensée peut prendre deux états solides différents, suivant les conditions dans lesquelles elle s'est formée : un état dit *amorphe*, dans lequel la disposition des atomes est aléatoire, ou un état *cristallisé*, caractérisé par un arrangement périodique des atomes. Les solides *amorphes* sont généralement appelés des verres et considérés comme des liquides de haute viscosité. Les solides *cristallisés* peuvent former des *monocristaux*, si la périodicité est ininterrompue dans tout le volume, ou des *poly-cristaux*, s'il existe une périodicité locale et que le solide est composé d'un ensemble de grains cristallins de taille à peu près uniforme. Dans la microphotographie de la Figure 5, prise par Transmission Electron Microscopy (TEM), une vue de coupe du réseau cristallin du silicium permet d'observer sa périodicité.

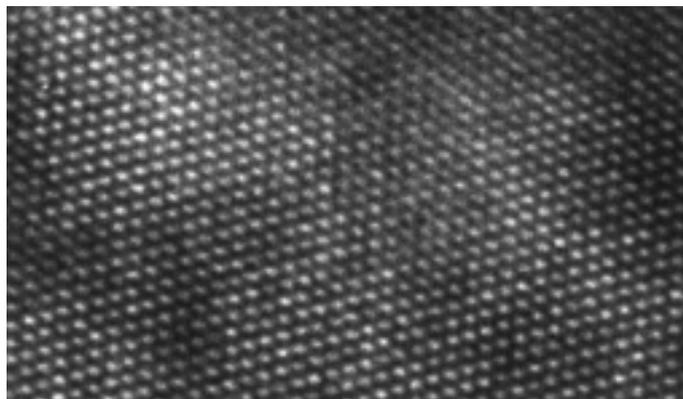


Figure 5. Microphotographie du réseau cristallin du silicium prise par TEM

Le résultat est un ensemble ordonné, ou localement ordonné, des noyaux d'atomes et d'électrons liés entre eux par des forces principalement coulombiennes. Ces forces sont plus ou moins importantes

suivant l'électronégativité des éléments intervenant dans les liaisons, ce qui confère aussi des propriétés mécaniques et électriques au solide cristallisé résultant. On distingue quatre familles de solides cristallins : les cristaux ioniques, les cristaux covalents, les métaux et les cristaux moléculaires [2].

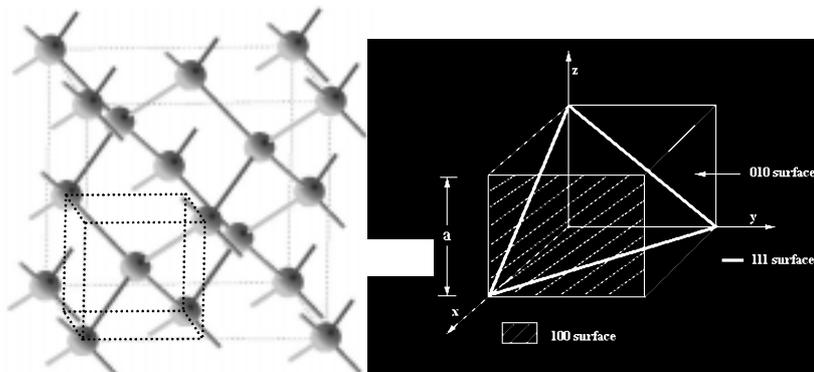


Figure 6. Réseau du silicium et sa cellule élémentaire (en pointillé noir), à gauche. Plans réticulaires principaux et ses indices de Miller, à droite.



Figure 7. Plans réticulaires principaux du silicium et leurs indices de Miller.

Le silicium, ainsi que les autres éléments de la colonne IV du tableau périodique (C, Ge, Sn) forment des cristaux *covalents*. Ces éléments génèrent des liaisons covalentes, avec leurs quatre atomes voisins, en mettant en commun leurs quatre électrons de valence. Les électrons de valence ont une énergie de liaison importante dans le cas du carbone diamant, ce qui en fait un isolant (ou plutôt un semiconducteur de large bande interdite), et une énergie nulle dans le cas de l'étain, ce qui en fait un bon conducteur. Dans le cas du germanium et du silicium cette énergie a une valeur intermédiaire à des températures normales, faisant de ces matériaux des semiconducteurs aux applications intéressants.

Le réseau cristallin du silicium est celui du diamant (voir Figure 6). Même si le silicium est un matériau monoatomique, la cellule élémentaire (en pointillé noir dans la Figure 6) contient deux atomes de silicium, l'atome central plus un quart de chacun de ses quatre voisins. Chaque atome a une coordination tétraédrique et a des liaisons covalentes avec les atomes voisins. Les plans réticulaires principaux du silicium sont montrés dans la Figure 6 et la Figure 7. Les plans 100 et 111 sont les plus utilisés par l'industrie de semiconducteurs sur silicium monocristallin.

2.1.2. Bandes d'énergie du Si

Bien que les électrons d'un atome isolé (i.e. un gaz) aient des niveaux d'énergie permis bien définis, le comportement de ces électrons dans un réseau cristallin périodique est différent. Si la distance entre les atomes est suffisamment petite, comme dans le cas du réseau cristallin du silicium, la présence des atomes voisins génère une sorte d'interférence entre les niveaux permis de chaque atome. Cela entraîne une multitude quasi continue d'états autour de chaque niveau permis de l'atome individuel. Cet ensemble d'états, très proches entre eux, est plutôt considéré alors comme une bande d'énergie permise, avec une densité d'états associée.

Les niveaux d'énergie des électrons de valence et le niveau d'énergie d'ionisation génèrent ainsi deux bandes d'énergie permises particulièrement intéressantes, la *bande de valence* et la *bande de conduction*, séparées par une bande dépourvue de niveaux permis, nommée *bande interdite* (voir Figure 8). La particularité de ces bandes permises tient au fait que, s'agissant d'électrons partagés par les atomes du réseau cristallin (Les électrons dans la bande de valence sont partagés par les atomes pour générer des liaisons covalentes. Dans la bande de conduction, les électrons ne sont liés à plus aucun atome et servent à la neutralité de la charge globale), ils peuvent circuler librement dans toute la maille cristalline et ainsi générer des courants de conduction électrique.

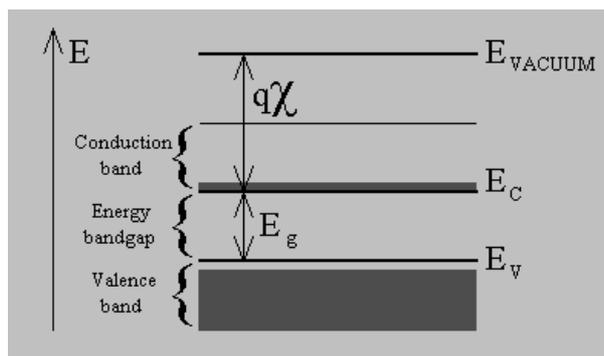


Figure 8. Diagramme des bandes d'énergie des semiconducteurs.

A des températures très basses (quelques K), la bande de conduction se trouve vide d'électrons et la bande de valence est complètement occupée, car l'énergie thermique n'est pas suffisamment importante pour ioniser les atomes du réseau cristallin. A des températures normales un bon nombre des atomes ionisés ont déjà cédé leurs électrons à la bande de conduction, dans laquelle ils participent à la conduction électrique avec une charge négative q . La bande de valence, même quasiment pleine, présente un nombre de niveaux inoccupés (trous) égal au nombre d'électrons dans la bande de conduction (dans le cas d'un semiconducteur intrinsèque); ces niveaux inoccupés (désormais nommés *trous*) peuvent être occupés par un électron de valence d'un atome voisin et ainsi se déplacer dans le réseau cristallin, et participer à la conduction électrique comme s'ils étaient des particules chargées positivement.

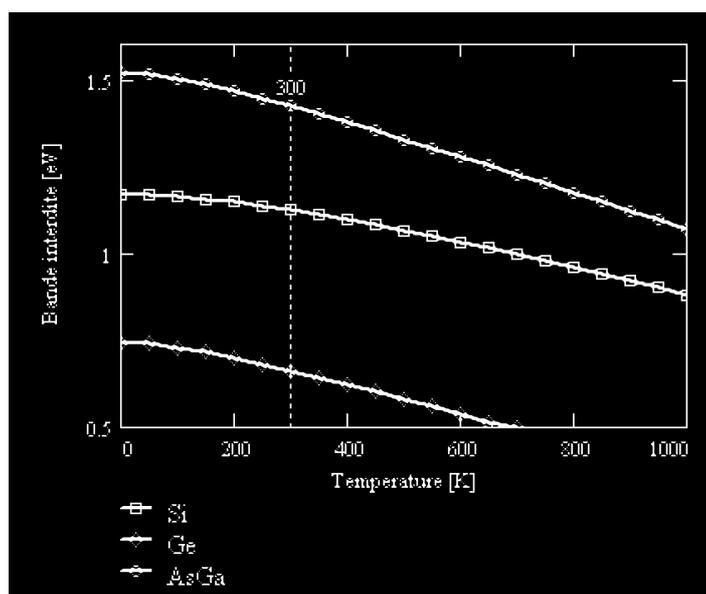


Figure 9. Valeur de la bande interdite du Si, Ge et AsGa en fonction de la température.

La taille E_g de la *bande interdite* donne des caractéristiques électriques importantes à chaque semiconducteur. Des résultats expérimentaux montrent que la largeur de la *bande interdite* décroît avec la température (voir Figure 9). Ces variations peuvent être décrites approximativement par la fonction universelle suivante [1],

$$(1) \quad E_g(T) = E_g(0) - \frac{aT^2}{(T+b)}$$

Semiconducteur	$E_g(0)$ [eV]	a [eV/K]	b [K]
AsGa	1.519	5.405	204
Si	1.170	4.73	636
Ge	0.7437	4.774	235

Tableau 1. Paramètres de la variation de la bande interdite en fonction de la température.

Dans le cas du silicium $E_g = 1.12 \text{ eV}$ à température ambiante.

Au voisinage de chacune des limites E_c et E_v , respectivement de la *bande de conduction* et de la *bande de valence*, la densité d'états permis $N(E)$ peut être estimée avec l'approximation parabolique suivante [2],

$$(2) \quad N_c(E) = K_c (E - E_c)^{\frac{1}{2}} \quad \text{pour la bande de conduction.}$$

$$(3) \quad N_v(E) = K_v (E_v - E)^{\frac{1}{2}} \quad \text{pour la bande de valence.}$$

K_c et K_v étant des constantes pratiquement indépendantes de la température.

2.1.3. Fonction de Fermi

Les électrons ne pouvant qu'occuper des niveaux d'énergie permis, leur distribution dans l'espace d'énergie dépendra de la distribution de ces niveaux dans le dit espace, et de la probabilité que les niveaux d'énergie mentionnés soient inoccupés. Cette probabilité d'occupation d'un niveau d'énergie permis E par un électron, à la température T , est donnée par la *fonction de Fermi* $f(E)$,

$$(4) \quad f(E) = \frac{1}{1 + e^{\frac{(E-E_f)}{kT}}}$$

Le niveau de Fermi E_f correspond au niveau d'énergie où la probabilité d'occupation est de 50% pour toute température (voir Figure 10). Pour des niveaux d'énergie inférieurs à E_f , la probabilité d'occupation va vers 100%. Pour des niveaux supérieurs à E_f , cette probabilité devient quasiment nulle après quelques dixièmes de eV. Cette transition devient de plus en plus abrupte au fur et à mesure qu'on abaisse la température (voir Figure 10).

La fonction de Fermi décrit des systèmes denses, dans lesquels le nombre de particules est comparable au nombre d'états permis, et elle découle de l'application du principe d'exclusion de Pauli. Pour des niveaux d'énergie supérieurs à E_f de quelques kT , la quantité d'électrons devient faible par rapport au nombre d'états permis; par conséquent, la fonction de Fermi (4) approche la *distribution de Boltzmann* (voir Figure 11) suivante,

$$(5) \quad f_B(E) = e^{\frac{-(E-E_f)}{kT}}$$

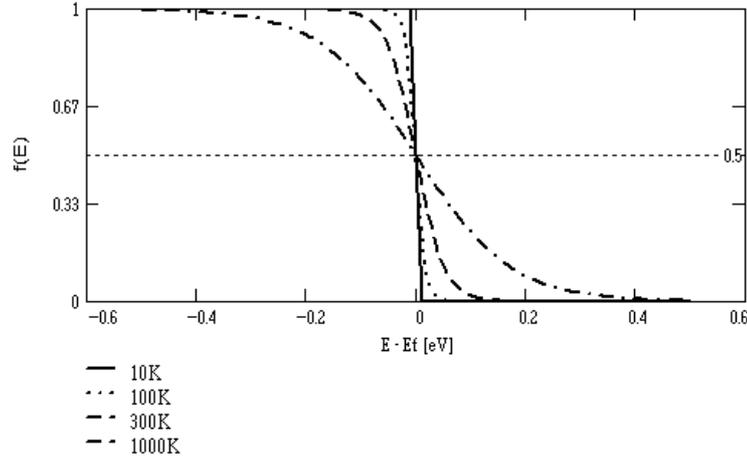


Figure 10. Fonction de Fermi $f(E)$ selon la température.

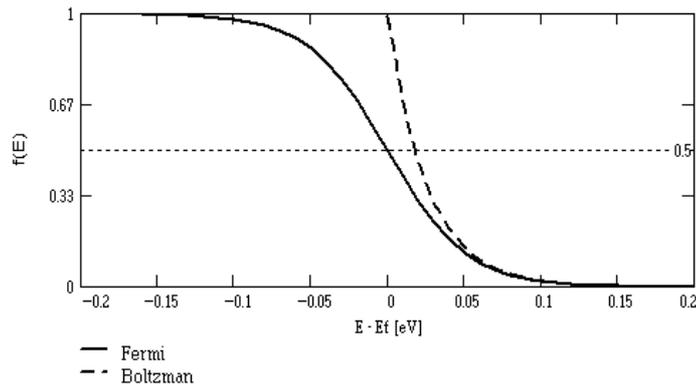


Figure 11. Fonction de Fermi $f(E)$ à 300K estimée par une distribution de Boltzmann.

2.1.4. Concentration intrinsèque

Le nombre d'électrons ou de trous d'énergie E dans un semiconducteur est donné par le produit de la densité d'états $N(E)$ par la fonction de distribution $f(E)$ (voir Figure 12). Ainsi, dans un semiconducteur non dégénéré (le niveau de Fermi E_f tombe dans la bande interdite, donc $f(E) \cong f_B(E)$ près de E_c ou E_v) à l'équilibre thermique, les densités totales d'électrons dans la bande de conduction et des trous dans la bande de valence sont données, à partir de (2), (3) et (5), par les expressions suivantes [2],

$$(6) \quad n = \int_{E_c}^{\infty} N_c(E) e^{\frac{-(E-E_f)}{kT}} dE = K n_i T^2 e^{\frac{-(E_c-E_f)}{kT}}$$

$$(7) \quad p = \int_{-\infty}^{E_v} N_v(E) e^{-\frac{(E-E_f)}{kT}} dE = K_p T^{\frac{3}{2}} e^{-\frac{(E_v-E_f)}{kT}}$$

K_n et K_p étant des constantes essentiellement indépendantes de la température. Les surfaces hachurées de la Figure 12 représentent les intégrales mentionnées.

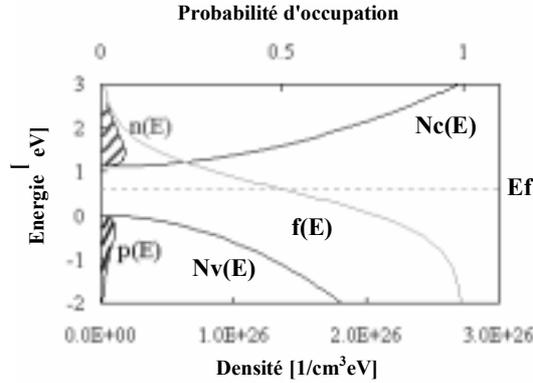


Figure 12. Densités de porteurs dans les bandes permises.

Un semiconducteur intrinsèque est un matériau dépourvu de toute impureté susceptible de modifier la densité des porteurs. Les électrons de la bande de conduction ne pouvant résulter que de l'excitation thermique des électrons de la bande de valence, les trous et les électrons existent nécessairement par paires ($n=p=ni$), donc à partir de (6) et (7) on obtient les expressions suivantes, pour la concentration intrinsèque des porteurs et le niveau de Fermi intrinsèque,

$$(8) \quad ni = \sqrt{n \cdot p} = \sqrt{K_n \cdot K_p} T^{\frac{3}{2}} e^{-\frac{E_g}{2kT}}$$

$$(9) \quad E_i = \frac{E_c + E_v}{2} + \frac{kT}{2} \ln \frac{K_v}{K_c} \sqrt{\frac{E_c + E_v}{2}}$$

Comme normalement $K_c \approx K_v$, il en résulte que le niveau de Fermi intrinsèque E_i est toujours très proche du centre de la bande interdite.

La concentration intrinsèque $ni(T)$ est fortement dépendante de la température. Comme cela est évident dans la Figure 13, elle varie quasi exponentiellement avec la température. A des températures supérieures à 800°K, la courbe s'éloigne de la droite parce que le terme $T^{3/2}$ commence à prendre de l'importance et parce que E_g décroît avec la température.

On peut réécrire (6) et (7) à partir de (8) et (9) de la façon simplifiée suivante, toujours dans le cas d'un semiconducteur non dégénéré et à l'équilibre thermique ($p=p_0$, $n=n_0$, $p_0 \cdot n_0 = ni^2$),

$$(10) \quad n_0 = ni e^{-\frac{(E_i-E_f)}{kT}}$$

$$(11) \quad p_0 = ni e^{\frac{(E_i-E_f)}{kT}}$$

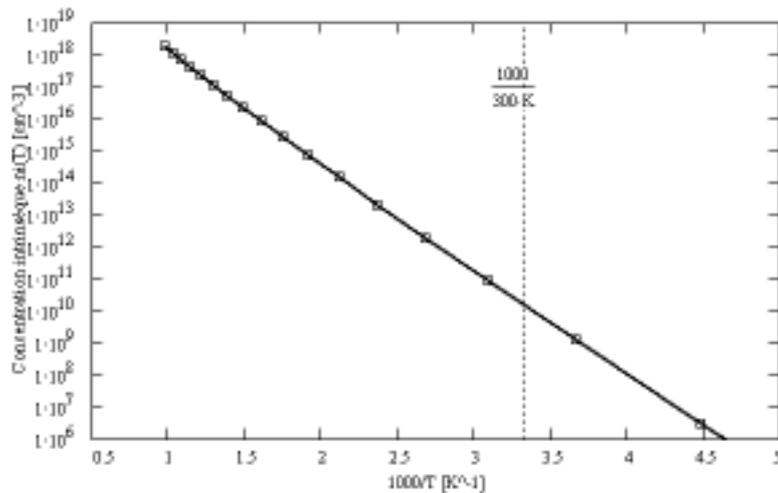


Figure 13. Concentration intrinsèque des porteurs dans le silicium.

2.1.5. Semiconducteur extrinsèque

Les propriétés électriques d'un semiconducteur de la colonne IV, comme le silicium, sont modifiées considérablement quand on le dope de manière contrôlée avec des atomes des colonnes III et V du tableau périodique (voir Figure 14). Considérons le cas du silicium monocristallin dans lequel on introduit un atome de la colonne V tel que l'arsenic. Après un traitement thermique l'atome d'arsenic remplace un atome de silicium dans le réseau cristallin et établit des liaisons avec ses quatre voisins. L'atome d'arsenic aura 9 électrons autour de lui, dont 8 saturent les orbitales liantes du cristal. Le 9ème occupe alors une orbitale beaucoup plus délocalisée autour de l'ion positif As⁺, et il a une faible énergie de liaison $E_D = E_c - E_d$. A la température ambiante, cet électron est donc libéré dans le réseau et occupe un niveau dans la bande de conduction. L'atome d'arsenic ionisé As⁺ reste comme une charge positive fixe dans le réseau cristallin. On dit que l'arsenic dans le silicium est un donneur, puisqu'il donne un électron de conduction.

La densité des donneurs ionisés N_d^+ s'écrit [2],

$$(12) \quad N_d^+ = \frac{N_d}{1 + 2 \gamma e^{\frac{E_f - E_d}{kT}}}$$

N_d étant la densité des donneurs et E_d le niveau d'énergie donneur.

IIIB	IVB	VB
B Boron 2-3	C Carbon 2-4	N Nitrogen 2-4
Al Aluminium 2-4-3	Si Silicon 2-4-4	P Phosphorus 2-4-6
Ga Gallium 2-4-10-3	Ge Germanium 2-4-10-4	As Arsenic 2-4-10-6

Figure 14. Extrait du Tableau périodique, autour du silicium.

Si un atome de silicium est remplacé par un atome de la colonne III, tel que le bore, il apparaît alors un déficit d'un électron dans son environnement. L'électron manquant est alors facilement remplacé par un électron provenant d'une liaison voisine, ce qui entraîne d'une part la création d'un trou dans la bande de valence, et d'autre part l'apparition d'une charge négative fixe autour de l'atome de bore. L'atome de bore devient un ion négatif B⁻ et est appelé accepteur. L'énergie nécessaire pour transférer un électron de la bande de valence et supprimer le déficit électronique autour de l'atome accepteur est appelée énergie de liaison $E_A = E_a - E_v$ du trou sur l'atome accepteur.

La densité d'accepteurs N_a s'écrit [2],

$$(13) \quad N_a^- = \frac{N_a}{1 + \frac{1}{4} e^{\frac{E_a - E_f}{kT}}}$$

N_a étant la densité des accepteurs et E_a le niveau d'énergie accepteur.

Le matériau semiconducteur étant globalement neutre, l'ensemble des charges positives est égal à l'ensemble des charges négatives. L'équation de la neutralité électrique s'écrit,

$$(14) \quad n + N_a^- = p + N_d^+$$

Si on remplace dans (14) les équations (8), (12) et (13) et si l'on résout le système résultant on obtient les courbes de la Figure 15, pour le silicium, avec $N_d = 10^{16} \text{ cm}^{-3}$, $N_a = 10^{14} \text{ cm}^{-3}$ et $E_c - E_d = E_a - E_v = 50 \text{ meV}$.

Trois régimes différents peuvent être identifiés dans la Figure 15. Le régime intrinsèque (c) : à très haute température la génération thermique des paires électron-trou est dominante, alors que les électrons générés à partir des dopants ionisés sont négligeables, le semiconducteur devient donc intrinsèque ($E_f = E_g/2$). Le régime de gel (a) : à très basse température, l'agitation thermique devient insuffisante pour ioniser les atomes donneurs, donc les électrons sont piégés par les atomes donneurs et leur concentration chute au fur et à mesure que la température diminue. Le régime d'épuisement (b) : à des températures normales, les atomes donneurs sont tous ionisés et la concentration d'électrons générée en conséquence est dominante par rapport à la génération thermique. Dans ce régime intermédiaire la concentration d'électrons est pratiquement égale à la concentration des impuretés donatrices.

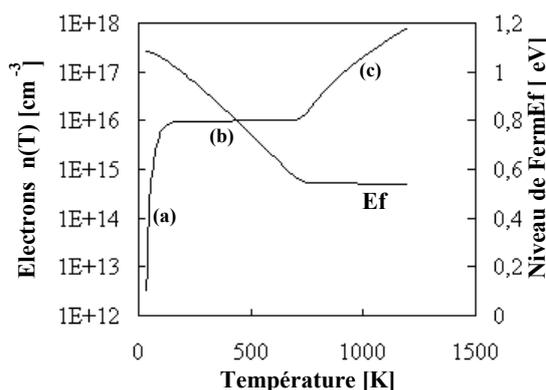


Figure 15. Densité des électrons et niveau de Fermi en fonction de la température.

2.1.6. Température intrinsèque

La transition entre le régime d'épuisement et le régime intrinsèque marque la température limite pour un dispositif semiconducteur, dite température intrinsèque. D'où l'importance, dans les cas d'applications en haute température, de repousser au maximum cette transition.

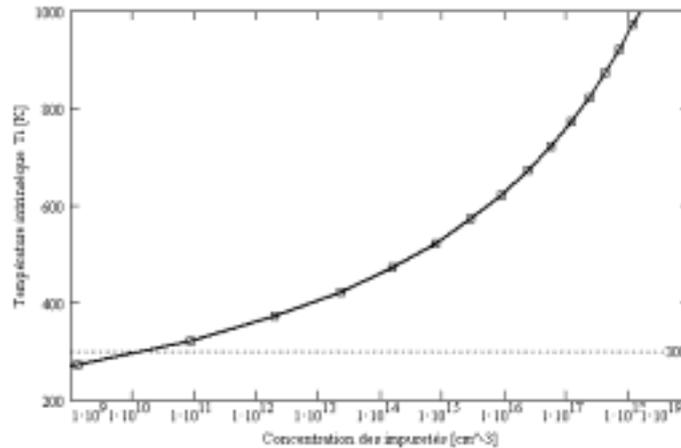


Figure 16. Température intrinsèque du silicium en fonction du dopage.

Si on définit la température intrinsèque T_i comme la température où la concentration intrinsèque $n_i(T)$ devient comparable ou égale à la concentration des impuretés dopantes nette $|N_d - N_a|$,

$$(15) \quad n_i(T_i) = |N_d - N_a|$$

on peut tracer la courbe de la Figure 16 à l'aide des équations (8) et (15). Pour élargir la plage de température d'utilisation des dispositifs semiconducteurs sur silicium, on peut jouer sur le dopage en augmentant le plus possible la concentration des impuretés. Le silicium moyennement dopé avec une concentration d'impuretés de 10^{16} cm^{-3} devient intrinsèque autour de 350°C .

2.1.7. Courants dans le semiconducteur

Les courants dans le semiconducteur résultent du déplacement des porteurs de charge, électrons et trous, sous l'action de différentes forces, l'origine de ces forces étant un champ électrique (*courant de conduction*) ou un gradient de concentration (*courant de diffusion*). La théorie des solides établie qu'un réseau parfaitement périodique ne dispersera pas les porteurs libres qui le traversent. Autrement dit, les porteurs libres n'échangeront pas d'énergie avec un réseau périodique statique. En revanche, à des températures au-delà du zéro absolu, les vibrations du réseau affecteront sa périodicité, permettant un échange d'énergie (des chocs), qui caractérisera le déplacement des porteurs dans le réseau semiconducteur.

Dans un semiconducteur où le champ électrique et le gradient de concentration des porteurs sont nuls, les porteurs libres sont uniquement animés d'un mouvement brownien, similaire à celui des molécules dans un gaz confiné. Entre deux collisions le mouvement est rectiligne uniforme, et les porteurs de masse m^* ont une vitesse caractéristique appelée *vitesse thermique du porteur*, donnée par [2],

$$(16) \quad V_{th} = \sqrt{\frac{3kT}{m^*}}$$

Où m^* correspond à la *masse effective de conductivité* des trous m_t ou des électrons m_e ; dans le silicium $m_t \approx 0.5m_e$, et $V_{th} \approx 10^7 \text{ cm/s}$ à la température ambiante.

Le *libre parcours moyen* l_m des porteurs est défini par [2],

$$(17) \quad l_m = V_{th} \tau_c$$

Où t_c est le temps entre deux collisions. A l'issue d'un choc toutes les directions du vecteur vitesse sont équiprobables, il en résulte que la vitesse moyenne des porteurs est nulle en l'absence d'un champ électrique ou d'une différence de concentration des porteurs.

2.1.7.1. Courant de conduction

Si un champ électrique E est appliqué, un porteur de charge q est soumis à une force $F=qE$ qui génère un courant de dérive des porteurs superposé à l'agitation thermique (voir Figure 17). La composante de vitesse instantanée $v_i(t)$ du porteur, correspondant à ce mouvement de dérive dans la direction du champ électrique, est alors donnée par [2],

$$(18) \quad F = m^* \frac{dv_i(t)}{dt} = qE$$

si on intègre $dv_i(t)$ de 0 à t ,

$$(19) \quad v_i(t) = \frac{qE}{m^*} t$$

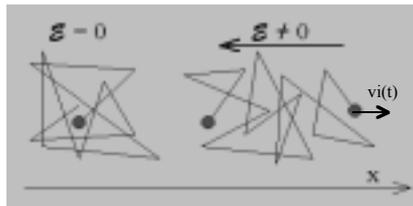


Figure 17. Agitation thermique superposée au mouvement de dérive.

Cette composante de vitesse augmente linéairement suivant le temps entre deux chocs et elle s'annule statistiquement à chaque choc. Donc, la valeur moyenne de la vitesse de dérive est donnée par [2],

$$(20) \quad v = \frac{1}{t_c} \int_0^{t_c} v_i(t) dt = \frac{qE t_c}{2m^*} = \frac{q \tau_r}{m^*} E = \mu E$$

donc,

$$(21) \quad \mu = \frac{q \tau_r}{m^*}$$

Où τ_r est appelé *temps de relaxation* et μ est appelé *mobilité des porteurs*. La mobilité détermine le rapport entre la vitesse de dérive et le champ électrique appliqué pour des champs électriques faibles. Comme le temps de relaxation des électrons est beaucoup plus important que celui des trous, il en résulte que la mobilité des électrons est plus importante que celle des trous. Pour le silicium [2], à température ambiante, $\mu_n \approx 1350 \text{ cm}^2 \text{V}^{-1} \text{sec}^{-1}$ et $\mu_p \approx 480 \text{ cm}^2 \text{V}^{-1} \text{sec}^{-1}$. La mobilité varie fortement selon la température et la concentration des impuretés (voir Figure 18).

La dispersion des porteurs, créée par l'interaction avec les vibrations du réseau, est plus efficace en haute température. La mobilité due à l'interaction avec des phonons acoustiques μ_{ph} est proportionnelle à [1],

$$(22) \quad \mu_{ph} \propto m^{*\frac{5}{2}} T^{-\frac{3}{2}}$$

Elle décroît avec la température et la valeur de la masse effective.

Les impuretés dopantes génèrent aussi des distorsions dans la périodicité du réseau qui dispersent les porteurs libres. Cette interaction, par collision coulombienne, est moins importante en haute température parce que les électrons se déplacent plus vite et qu'ils restent pendant un temps plus court dans la région influencée par une impureté. La mobilité due à l'interaction avec des impuretés ionisées μ_c est proportionnelle à [1],

$$(23) \quad \mu_c \propto m^* \frac{1}{2} N_i^{-1} T^{\frac{3}{2}}$$

Où N_i est la densité des impuretés ionisées. La mobilité μ_c augmente avec la température et elle décroît selon N_i et la valeur de la masse effective.

Dans le cas du silicium, la dispersion causée par les vibrations du réseau et la présence des impuretés, sont les deux phénomènes dominants qui déterminent la mobilité. La mobilité totale μ_{tot} est une combinaison des deux phénomènes [1],

$$(24) \quad \frac{1}{\mu_{tot}} = \frac{1}{\mu_{ph}} + \frac{1}{\mu_c}$$

Le modèle le plus couramment utilisé par les simulateurs du type SPICE pour la mobilité est le suivant,

$$(25) \quad \mu_{tot, p}(T) = \mu_{tot, p}(T_{amb}) \sqrt{\frac{T}{T_{amb}}}^{TC_{n,p}}$$

Ce modèle est valable pour la plage de température de 150 à 500°K, environ.

Dans le cas du silicium pur à des températures normales $TC_n = -2.42$ et $TC_p = -2.20$ [1].

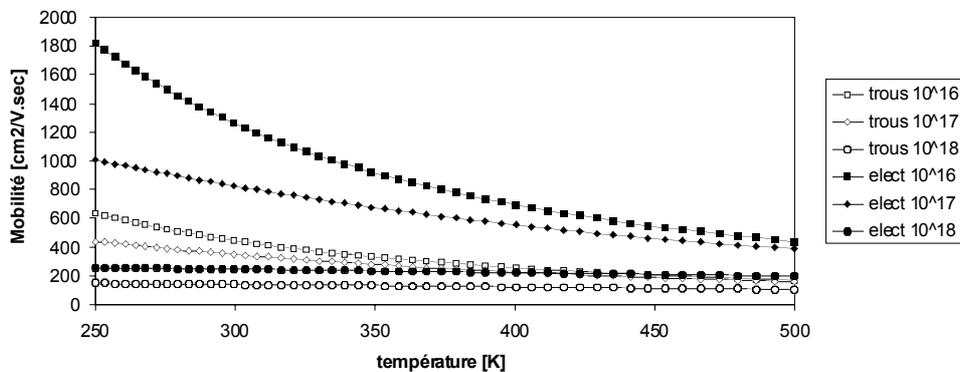


Figure 18. Mobilité dans le silicium en fonction de la température et de la densité des impuretés.

La conductivité électrique σ est due à la contribution, à la densité de courant totale j_t , des deux types de porteurs (leurs concentrations, leurs charges et leurs vitesses de dérive), en réponse à un faible champ électrique E .

$$(26) \quad j_t = n q v_n + p q v_p = q (n \mu_n + p \mu_p) E = \sigma E$$

ainsi, la conductivité $\sigma(T)$ et la résistivité $\rho(T)$, sont données par,

$$(27) \quad \sigma(T) = q [n(T) \mu_n(T) + p(T) \mu_p(T)]$$

$$(28) \quad \rho(T) = \frac{1}{q [n(T) \mu_n(T) + p(T) \mu_p(T)]}$$

2.1.7.2. Courant de diffusion

Si dans le semiconducteur il y a des variations dans la distribution spatiale des porteurs, un processus de diffusion se produit. Le mouvement des porteurs libres s'effectuera de façon à uniformiser leur concentration dans tout le volume du semiconducteur. Conformément à la première loi de Fick, le flux de porteurs sera proportionnel à leur gradient de concentration, donc la densité de courant de diffusion sera donnée par,

$$(29) \quad j_{d_n} = q D_n \frac{dn}{dx}$$

$$(30) \quad j_{d_p} = -q D_p \frac{dp}{dx}$$

D_n et D_p sont les coefficients de diffusion des porteurs correspondant respectivement aux électrons et aux trous. La relation de Einstein, pour les semiconducteurs non dégénérés, établit le rapport entre le coefficient de diffusion et la mobilité suivant,

$$(31) \quad D_{n,p} = \frac{kT}{q} \mu_{n,p}$$

2.1.7.3. Courant total

La densité totale de courant pour les électrons et les trous est donnée, à partir de (26), (29), (30) et (31), en explicitant la relation de dépendance avec la température, par,

$$(32) \quad j_n = q \mu_n(T) \frac{dn(T)}{dx} + \frac{kT}{q} \frac{dn(T)}{dx}$$

$$(33) \quad j_p = q \mu_p(T) \frac{dp(T)}{dx} - \frac{kT}{q} \frac{dp(T)}{dx}$$

2.1.8. Génération et recombinaison des porteurs

Chaque fois que la condition d'équilibre thermique ne se vérifie pas quelque part dans le volume du semiconducteur ($p \neq p_0$, $n \neq n_0$, $p \cdot n \neq n_i^2$), des processus s'activent afin de restaurer cet équilibre ($p = p_0$, $n = n_0$, $p \cdot n = n_i^2$). Dans la Figure 19 les trois processus basiques de recombinaison sont représentés. Dans le cas d'une *recombinaison électron-trou directe* (a) ou d'une *recombinaison Auger* (c), un électron passe de la bande de conduction directement à la bande de valence en émettant un photon (a) ou en transférant l'énergie à un autre électron ou à un autre trou (c). Le processus (a) est l'inverse des transitions optiques directes, et le processus (c) est l'inverse de l'ionisation par impact; ces processus concernant plutôt les semiconducteurs III-V à gap direct [1]. Dans la *recombinaison assistée par*

centres de recombinaison (b) on vérifie que la présence des impuretés incontrôlées joue un rôle important dans le retour à l'équilibre thermique de semiconducteurs. Un centre de niveau d'énergie E_t est un *centre de recombinaison* si, quand il capture un électron, la probabilité de capturer en suite un trou (et de recombiner ainsi la paire électron-trou) est plus grande que celle de remettre l'électron dans la bande de conduction.

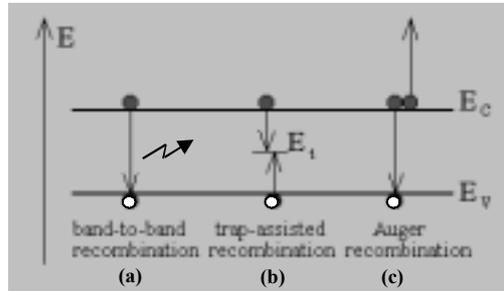


Figure 19. Processus de recombinaison des porteurs.

Le taux de génération et de recombinaison U des porteurs représente la variation du nombre de porteurs par unité de volume et de temps, résultant des phénomènes de *recombinaison assistée par centres de recombinaison* et de *génération thermique spontanée*. La théorie de *Shockley-Read* établie [1] [2],

$$(34) \quad U = N_t \frac{C_n C_p (p n - n_i^2)}{C_n n + n_i e^{\frac{E_t - E_i}{kT}} + C_p p + n_i e^{\frac{E_i - E_t}{kT}}}$$

N_t est la densité des centres de recombinaison d'énergie E_t , C_n et C_p sont respectivement les coefficients de capture d'électrons et de trous. Dans la pratique, les centres les plus efficaces, concernant le processus de recombinaison, introduisent dans la bande interdite des niveaux d'énergie proches de E_i et ont des coefficients de capture tels que $C_n \approx C_p = C$. On peut donc écrire (34) sous la forme simplifiée suivante,

$$(35) \quad U = N_t \frac{C (p n - n_i^2)}{2 n_i + p + n}$$

2.1.8.1. Semiconducteur de type n

Dans les cas d'un semiconducteur type n sous faible excitation, on constate que $n = n_0 + dn \approx n_0 \gg n_i \gg p_0$ et $p = p_0 + dp \ll n_0$. L'équation (35) s'écrit alors,

$$(36) \quad U = N_t \frac{C (p n - n_i^2)}{n} \quad U(T) = N_t C (p(T) - p_0(T))$$

2.1.8.2. Zone désertée

Dans le cas d'une zone désertée, comme à l'interface d'une jonction $p-n$, on peut négliger p et n devant n_i , alors (35) s'écrit,

$$(37) \quad U(T) = -N_t C \frac{n_i(T)}{2}$$

2.1.9. Conductivité thermique du silicium

Une bonne conductivité thermique est nécessaire dans le substrat semiconducteur, d'une part pour réduire la température du substrat par dissipation dans le cas d'un auto-échauffement généré par une forte injection locale de chaleur, et d'autre part pour assurer une température uniforme dans tout le substrat, afin d'améliorer la similitude ("*matching*") entre le comportement des dispositifs de nature identique mais placés dans différentes positions dans le substrat semiconducteur.

Le problème du management thermique des systèmes électroniques était une des difficultés qui a limité l'évolution des technologies de tubes à vide. L'électronique d'état solide a pris le relais avec succès, grâce à sa plus haute densité d'intégration et à sa plus faible consommation d'énergie, ce qui ne rendait plus nécessaire aucune considération d'ordre thermique. Mais les technologies des circuits intégrés sur silicium ont fortement évolué en densité d'intégration, et densité de puissance dissipée, tout en conservant le même type de substrat semiconducteur (i.e. les mêmes caractéristiques d'évacuation de la chaleur générée), produisant une augmentation de la température, et des gradients de température, dans le substrat. En revanche, de considérables améliorations de la conductivité thermique des boîtiers ("*packages*") et de la dissipation passive ou dynamique de la chaleur dans l'environnement ont été apportées. De même, des nouveaux matériaux semiconducteurs de plus large bande interdite et de plus importante conductivité thermique, tels que le *SiC* et le *Diamant*, sont étudiés pour des applications à plus de 300°C. En ce qui concerne la gestion de la chaleur dans les circuits intégrés analogiques et mixtes, quelques techniques d'implémentation peuvent être appliquées afin de minimiser les gradients et les transitoires thermiques autour des structures les plus sensibles (voir la section 4.2.1.1).

2.1.9.1. Modèles de dissipation de la chaleur

La quantité de chaleur Q (en Joules) dans un matériau est proportionnelle à sa température absolue T ,

$$(38) \quad Q = C_T \cdot T = M \cdot c \cdot T$$

Où M est la masse du matériau et c sa capacité thermique par unité de masse ou capacité thermique spécifique (en $J/Kg \cdot ^\circ K$) celle-ci étant une propriété fondamentale du matériau. La capacité thermique spécifique est une mesure de la propriété d'un matériau d'accumuler de l'énergie thermique, en analogie avec la capacité d'un condensateur électrique qui accumule de l'énergie électrostatique.

Une différence de température entre deux points d'un matériau se comporte de la même manière qu'une différence de tension, car elle génère un flux de chaleur du point le plus chaud vers le point le plus froid. Le débit dQ/dt du flux de chaleur à travers le matériau dépend du gradient de température dT/dx ,

$$(39) \quad P_Q = \frac{dQ}{dt} = -\kappa \cdot S \cdot \frac{dT}{dx}$$

Où P_Q (en Watts) est la puissance thermique dissipée, S la surface de la section du matériau à travers lequel la chaleur passe, et κ (en W/mK) la conductivité thermique du matériau. Dans le cas du silicium la conductivité κ est de $1.56 W/cmK$; elle est supérieure à celle du *GaAs* ($0.44 W/cmK$), *GaN* ($1.3 W/cmK$), *GaP* ($0.8 W/cmK$), *AlN* ($0.3 W/cmK$), et inférieure à celle du *SiC* ($4.9 W/cmK$) et du *Diamant* ($10 W/cmK$) [3][5].

La résistance thermique R_T d'un corps, définie comme la réticence au passage de la chaleur à travers le corps (en analogie avec la résistance électrique au passage d'un courant), dépend de la conductivité du matériau et de sa géométrie (la surface de sa section S et la longueur L),

$$(40) \quad R_T = \frac{L}{\kappa \cdot S}$$

La température T et le flux thermique P_Q sont les variables fondamentales (comme la tension et le courant dans les systèmes électriques) tandis que R_T et C_T décrivent les propriétés basiques du système thermique. Dans le Tableau 2, l'analogie entre les systèmes thermiques et les systèmes électriques, est établie. Mais, il n'y a pas d'équivalent thermique pour l'inductance électrique. De plus, le comportement réel des systèmes thermiques n'est pas linéaire car les valeurs des résistances et des capacités thermiques dépendent normalement de la température (Le coefficient de diffusion thermique du silicium : $Dth_{si} = \kappa_{si}/c_{si}$ à une dépendance de la température de l'ordre de $-0.57\%/^{\circ}C$ [32] [33]).

Système	Variables	Eléments	Loi statique	Loi dynamique
Electrique	$I - V$	$R - C - L$	$q = C/V$	$V = I.R$
Thermique	$P_Q - T$	$R_T - C_T - \dots$	$Q = C_T/T$	$T = P_Q.R$

Tableau 2. Analogie entre les systèmes thermiques et électriques

Cependant, un modèle linéaire du système thermique (i.e. des fonctions exponentielles pour l'évolution de la température selon le temps) représente dans la plupart des cas une bonne approximation du comportement réel. Dans la Figure 20, un modèle simple pour la dissipation de la chaleur d'un circuit intégré encapsulé est décrit. Où S = Substrat, Csb = Contact entre le substrat et le boîtier, B = Boîtier, P = pattes du boîtier, Cbd = Contact entre le boîtier et le dissipateur, D = Dissipateur, C = Convection spontanée ou forcée, et Amb = Ambiante.

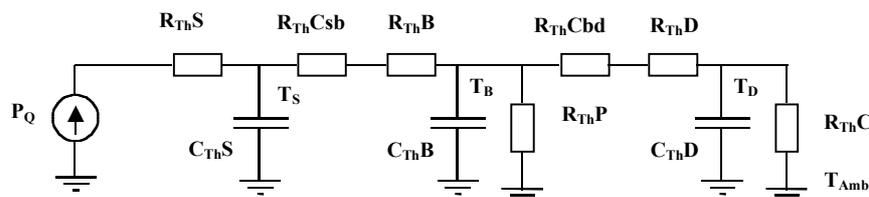


Figure 20. Modèle simple pour la dissipation de la chaleur d'un circuit intégré encapsulé.

2.2. Couches conductrices et contacts

Dans les technologies des circuits intégrés planaires, la définition de l'interaction électrique entre les différents dispositifs intégrés (i.e. l'interconnexion) se fait à travers les contacts et les couches conductrices. La qualité de ces interconnexions implique un compromis entre la taille des conducteurs, leur résistivité, les propriétés mécaniques de la couche conductrice déposée, et la qualité des interfaces générées. Il en résulte que différents matériaux (des métaux, du silicium polycristallin fortement dopé) sont utilisés selon le type d'interconnexion (i.e. grilles des transistors MOS, plaques des capacités intégrées, lignes d'interconnexion, etc.). Aussi, des nouveaux matériaux et alliages sont recherchés afin de pouvoir réduire la résistivité des conducteurs, augmenter la densité d'intégration et la fiabilité des interconnexions. Des couches conductrices sont aussi utilisées pour la génération des dispositifs passifs intégrés tels les inducteurs, capacitances et résistances. En ce qui concerne l'intégration des résistances, les propriétés recherchées sont plutôt une haute résistivité et un faible coefficient thermique, tout en restant le plus proche possible du comportement décrit par la loi de Ohm (i.e. faible non-linéarité). La dégradation de la qualité des interconnexions et des dispositifs passifs, à long terme et au fur et à mesure que la température augmente, a une grande importance pour la détermination de la tenue en température d'une technologie des circuits intégrés.

2.2.1. Métallisations

Les facteurs les plus importants qui limitent la durée de vie des conducteurs et des contacts métalliques en haute température sont l'électromigration et la corrosion [5]. Les propriétés et l'origine de ces phénomènes doivent être prises en compte par les concepteurs des procédés de fabrication (pour le choix des matériaux) et par le concepteur des circuits intégrés (pour le dimensionnement des conducteurs) afin d'éviter des conséquences irréversibles.

2.2.1.1. Electromigration

Les effets de l'électromigration (EM) dans les couches minces conductrices ont été étudiés à la fin des années 60, du fait de l'utilisation généralisée des couches minces métalliques pour les interconnexions dans les circuits intégrés sur silicium. Il a été observé que cette migration des atomes déforme significativement les conducteurs et peut être une cause importante de défaillance. Ce problème, qui paraissait complètement résolu dans les années 80, est réapparu récemment et son domaine d'action s'est étendu aux structures des contacts. En effet, la réduction constante des dimensions des dispositifs entraîne aussi une forte réduction de la section des lignes conductrices et de la surface des contacts, se fabriquant sur des jonctions de moins en moins profondes. Il en résulte que pour un facteur de réduction d'échelle K , la densité de courant augmente avec un facteur K^2 pour les interconnexions et avec un facteur K^3 pour les contacts [4].

Le facteur essentiel de l'électromigration est la densité de courant électrique. Même s'il y a d'autres causes significatives d'électromigration, telles que les gradients de température, les gradients de tension, les gradients de concentration, etc., il n'y a pas d'électromigration sans courant électrique. En général les forces qui agissent sur les atomes dans l'électromigration résultent de la contribution de deux forces: l'une due à l'action directe du champ électrostatique sur les ions, l'autre due à l'action du mouvement des porteurs libres sur le mouvement d'ions (force du *vent électronique*). Dans le cas des métaux pour la microélectronique, tel que l'aluminium, où les électrons sont essentiellement les porteurs de charge, une diffusion des atomes vers l'électrode positive a été clairement observée. En revanche dans le cas du tungstène, où les porteurs de charge sont les trous plutôt que les électrons, le mouvement de matériau se fait en sens inverse. Ce qui indique que la force du *vent électronique*, due à la densité de courant électrique, est dominante [4].

Le flux atomique est caractérisé par un coefficient de diffusion D , qui a une énergie d'activation E_a ,

$$(41) \quad D = D_0 ? e^{\frac{-E_a}{kT}}$$

A des températures proches de celle de la fusion du matériau, le coefficient de diffusion et l'énergie d'activation, pour la migration des atomes dans le volume, sont des paramètres bien connus. Mais l'analyse de l'électromigration en couches minces est beaucoup plus complexe, surtout à des températures inférieures à 300°C, où la diffusion des atomes se fait principalement le long des joints des grains et dans les surfaces du film, si elles n'ont pas été passivées.

La surface d'un film représente le plus large des défauts. Si les atomes en surface ne sont pas immobilisés, leur migration dominera l'électromigration du film. Une couche de passivation bien collée au film conducteur permet de réduire les dégradations par électromigration du conducteur d'une manière directe et d'une manière indirecte. La réduction directe des endommagements provient du fait que les fortes liaisons entre les atomes de la surface du film conducteur et ceux de la couche de passivation permettent d'immobiliser les atomes en surface, en réduisant énormément l'électromigration en surface. La réduction indirecte résulte de l'élimination de la source principale des lacunes : la surface du film. Pour créer des cavités dans le film conducteur, par accumulation des

lacunes, il faut un très grand apport de lacunes neuves qui viennent principalement de grands défauts, telle une surface non passivée du film. Tout en éliminant la source des lacunes, on réduit énormément la création des cavités.

Il en résulte que le mouvement des atomes, dans un film conducteur avec ses surfaces passivées, varie plutôt en fonction de la distribution de la taille des grains et en fonction de l'orientation des joints entre les grains. En plus, d'autres mécanismes de diffusion sont présents, chacun avec sa propre énergie d'activation (voir Tableau 3).

Mécanisme de diffusion	Energie d'activation (eV)
EM en volume	1.4
EM dans les joints des grains	0.4 - 0.5
Passage d'un joint au volume	0.62
Passage d'un défaut au volume	> 0.62
EM en surface	0.28

Tableau 3. Energies d'activation correspondantes aux couches minces d'aluminium [4].

Le flux total de matériau étant composé de plusieurs facteurs, l'énergie d'activation effective résultante dépend de la plage de température considérée, de la taille et de l'orientation des grains, du type et de la quantité des défauts, etc., ce qui explique la dispersion des valeurs d'énergie d'activation effective obtenues par différents auteurs [5].

Mais la migration d'atomes ne peut pas seule générer des défauts dans une couche métallique, s'il n'y a pas une importante divergence dans le flux atomique. Les cavités se produisent là où le mouvement atomique génère une déplétion massique, et elles peuvent augmenter jusqu'à ouvrir une ligne de connexion. Les protubérances se forment là où il y a une accumulation de masse, et peuvent court-circuiter deux lignes d'interconnexion ou même percer les couches de passivation, s'exposer ainsi à la corrosion.

Plusieurs facteurs, comme par exemple des gradients de température, la taille et l'orientation des grains, etc., peuvent être responsables de la divergence dans le flux. C'est en particulier la structure non uniforme des grains qui est à l'origine de l'électromigration dans les couches minces. Une non homogénéité bien connue se trouve dans le joint de trois grains, dit *point triple*, où la diffusion d'atomes change de direction. Une autre source de divergence se trouve dans un changement local de la taille des grains, où les atomes passent d'un milieu à forte densité de joints (des voies de passage rapide) à un milieu où la densité de joints est plus faible, et vice versa, en générant respectivement une accumulation ou une déplétion de masse dans l'interface entre les deux milieux. Pour conclure, la distribution uniforme de la taille des grains et l'augmentation de la taille moyenne des grains, par rapport à la largeur des conducteurs, permettent d'augmenter sensiblement la résistance à l'électromigration.

Les techniques expérimentales les plus utilisées pour quantifier l'effet d'électromigration cherchent à déterminer le temps moyen de vie, ou temps moyen avant défaillance (*Mean Time to Failure MTF*), des conducteurs. Les mesures se font à des températures et des densités de courant bien supérieures aux conditions normales de fonctionnement. Le modèle pour la prédiction du *MTF*, proposé par Black [4] est le suivant,

$$(42) \quad MTF^{-1} = A j^n e^{-\frac{E_{a,eff}}{kT}}$$

Où n prend une valeur dans la plage de 1 à 7, en fonction de la densité du courant, de la température, et de la géométrie du film [4]. Il est important de remarquer la forte dépendance de *MTF* par rapport à la densité de courant j et en particulier par rapport à la température, ce qui indique une forte accélération du processus de dégradation des conducteurs, au fur et à mesure que la température augmente. Normalement, les fabricants de circuits intégrés spécifient les résultats des tests de fiabilité, concernant l'électromigration, en termes de densité de courant maximal JI (par unité de

largeur du conducteur) admissible à une température de référence $T1$. A partir de l'équation (42) on peut écrire la relation, qui détermine approximativement la densité de courant maximal à respecter pour une température $T2$ et un MTF équivalent,

$$(43) \quad J2 = J1 \cdot e^{\frac{-Ea_{eff}}{nk} \cdot \frac{T2-T1}{T2T1}}$$

En estimant les valeurs de Ea_{eff} et n , et à l'aide de l'équation (43) et des spécifications techniques fournies par le fabricant des circuits intégrés, le concepteur de circuits intégrés en haute température peut prendre une juste marge de sécurité afin d'éviter le raccourcissement de la durée de vie du circuit, causée par le phénomène d'électromigration.

2.2.1.1.a. Aluminium

L'aluminium (Al) est le matériau le plus utilisé pour les interconnexions métalliques dans les circuits intégrés, à cause de sa haute conductivité électrique et thermique, de sa bonne adhérence à l'oxyde de silicium (SiO_2), de sa résistance à la corrosion, de la facilité à le déposer et le graver, et de son coût peu élevé. L'électromigration génère dans les conducteurs en aluminium un transport de masse, comme précédemment décrit, caractérisé par une énergie d'activation comprise entre 0.4 et 1.3 eV [4]. Les méthodes typiquement utilisées pour augmenter la robustesse des conducteurs Al vis à vis de l'électromigration, sont fondées sur:

- Ø L'optimisation de la structure granulaire, comme précédemment définie.
- Ø L'utilisation de couches de passivation d'oxyde de silicium, pour inhiber la formation des protubérances et retarder la formation des cavités.
- Ø La déposition des couches d'aluminium avec un faible pourcentage de cuivre (Cu) ou d'autres matériaux.
- Ø La formation des structures style sandwich, alternant des couches d' Al avec un matériau réfractaire.

2.2.1.1.b. Alliages $Al-Cu$

Il a été vérifié lors des tests de résistance à l'électromigration que l'addition d'un deuxième élément au film conducteur augmente sensiblement le MTF . En particulier l'alliage $Al-Cu$ a été largement étudié et utilisé. Pour de faibles concentrations de Cu (inférieures à 3 ou 4 %), le MTF augmente de façon régulière, principalement à cause de l'absorption des atomes de cuivre dans les joints entre les grains d'aluminium. En fait, le nombre de défauts dans les joints se réduit considérablement avec l'intrusion des atomes de cuivre plus fortement liés. La diffusion se faisant essentiellement le long des joints des grains; si les positions disponibles sont occupées par les atomes de cuivre, il faut un apport supplémentaire d'énergie pour permettre la diffusion des atomes d'aluminium. Il en résulte que l'énergie d'activation pour l'électromigration augmente.

Ainsi le cuivre est aussi considéré comme un possible remplaçant du $Al-Cu$, grâce à sa meilleure conductivité électrique et à son plus haut point de fusion, ce qui impliquerait une meilleure résistance à l'électromigration.

2.2.1.2. Corrosion

La corrosion des couches minces métalliques peut être décrite comme une perte involontaire du matériau métallique, soit par dissolution dans un électrolyte, soit par transformation du matériau

métallique en produits non-conducteurs par réaction chimique, en réponse à des contraintes externes de nature principalement chimique ou électrochimique. Ces contraintes externes résultent de la combinaison des éléments suivants,

- ∅ Les impuretés externes au film (i.e. poussière, des gaz corrosifs, etc.).
- ∅ La température.
- ∅ La tension appliquée.
- ∅ Les contraintes mécaniques, principalement liées à la déposition des couches protectrices, etc.

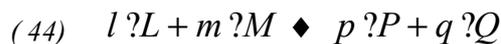
Cette interaction entre les métaux et l'environnement peut conduire à la destruction des connexions électriques au niveau des couches d'interconnexion métalliques, au niveau des fils de connexion entre le circuit intégré et le boîtier et au niveau des soudures dans le boîtier et les plots d'entrée/sortie du circuit intégré.

Normalement les mécanismes de corrosion correspondent à l'un des deux types suivants :

- ∅ Corrosion sèche : C'est, par exemple, le cas de l'oxydation de l'aluminium au contact de l'air.
- ∅ Corrosion humide : Où la réaction nécessite la présence d'un électrolyte, d'un environnement pollué et d'une force électromotrice.

La corrosion sèche est généralement moins importante du point de vue de la fiabilité des dispositifs semiconducteurs. Par exemple, l'oxydation de l'aluminium au contact de l'air est considérée comme une réaction auto-passivée, qui forme des oxydes d'une épaisseur inférieure à 50 Å. Dans un environnement sec, tel l'intérieur d'un boîtier hermétiquement scellé sous une atmosphère contrôlée (azote, gaz inerte), la corrosion de l'aluminium est quasiment impossible [4].

Pour déclencher la corrosion humide, la formation d'une cellule électrolytique est une condition nécessaire. Cette cellule est composée d'une anode et d'une cathode métalliques, connectées entre elles, à l'extérieur de la cellule par un circuit qui permet la circulation des électrons, et à l'intérieur de la cellule par un milieu électrolytique qui permet la diffusion des ions. La dynamique de réaction d'une corrosion électrochimique de type,



est décrite par la relation de Nerst suivante [4],

$$(45) \quad A_P^p \text{ ?}A_Q^q = A_M^m \text{ ?}A_L^l \text{ ?}e^{-\frac{K}{T}(E-E_0)}$$

Où,

E = Force électromotrice de la cellule électrolytique.

E_0 = Potentiel d'oxydation standard.

T = Température absolue.

K = Constante qui ne dépend pas de la température.

l, m, p, q = moles correspondant aux espèces chimiques L, M, P, Q.

A = Activités chimiques.

L'intensité de la réaction, d'après l'équation (45), dépend des éléments intervenant (par exemple, le l et le K donnent des valeurs de E_0 positives et très importantes) et dépend exponentiellement de la température, ce qui indique une forte accélération du processus de dégradation des conducteurs au fur et à mesure que la température augmente. Le besoin d'une bonne tenue à la corrosion, dans le cas des applications des circuits intégrés en haute température, exige leur encapsulation étanche sous atmosphère contrôlée (azote, gaz inerte), afin d'éviter toute corrosion humide, dans des boîtiers (céramiques ou autres) résistants à la température maximale considérée.

2.2.2. Silicium poly-cristallin

Comme précédemment mentionné, les solides *cristallisés* peuvent former des *monocristaux*, si la périodicité est ininterrompue dans tout le volume, ou des *poly-cristaux*, s'il existe une périodicité locale et que le solide est plutôt composé d'un ensemble de grains cristallins, de taille à peu près uniforme, séparés par des joints de grains. La résistivité des films de silicium poly-cristallin est une fonction complexe de la taille des grains cristallins, de la concentration des dopants et de la température [6] [27] Robert A. Pease, *Troubleshooting Analog Circuits*, Butterworth-Heinemann, Boston, 1991. ISBN: 0-7506-9184-0

[28]. Pour de faibles concentrations des dopants, la conduction électrique est déterminée par les caractéristiques des joints de grains et par le piégeage des porteurs dans des centres profonds induits par les joints des grains. Les charges mobiles piégées dans les joints établissent une barrière de potentiel. Dans ces conditions, la résistivité est décrite par,

$$(46) \quad \sigma = \sigma_0(T) \cdot e^{-\frac{E_B}{kT}}$$

Où E_B est une énergie d'activation qui prend en compte la hauteur de la barrière de potentiel. Le comportement exponentiel de la résistivité, décrit par (46), donne des coefficients thermiques négatifs aux résistances sur silicium poly-cristallin peu dopé. Au fur et à mesure que la concentration des dopants augmente, la hauteur de la barrière de potentiel augmente initialement jusqu'à ce que tous les centres profonds soient remplis. Puis, la largeur et la hauteur de la barrière commencent à diminuer et pour de très fortes concentrations de dopants, ou des tailles de grains très larges, elles deviennent négligeables et la résistivité est plutôt déterminée par la résistivité dans le volume des grains : le comportement du polycristal vis à vis de la résistivité correspond à celui du monocristal, pour des concentrations de dopants équivalentes (i.e. le coefficient thermique de la résistivité devient positif).

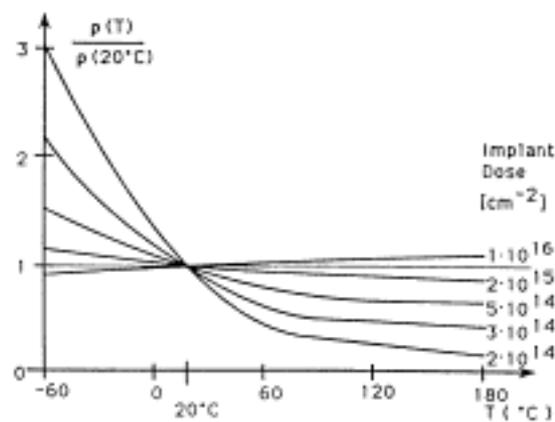


Figure 21. Variations, selon la température, de la résistivité normalisée du polysilicium dopé Bore.

La résistivité du poly-silicium peut être contrôlée en modifiant le dopage du film (voir la Figure 21). Pour de très faibles concentrations de dopants, la résistivité est très haute et elle décroît exponentiellement avec la température. A de fortes concentrations, le régime de conduction en volume donne une résistivité plus faible, qui augmente légèrement avec la température. Ce dernier cas est celui des couches de silicium polycristallin normalement utilisées dans les technologies CMOS et BiCMOS (~ 25 ohms/carré) pour la formation de la grille des transistors MOS, les capacitances dites "double Poly-Si", et les résistances en polysilicium.

2.3. Isolation

Les techniques d'isolation entre les différents dispositifs intégrés, ainsi que les techniques d'interconnexion, permettent l'intégration de milliers de dispositifs sur un substrat unique, et jouent un rôle important dans l'augmentation de la densité d'intégration dans les technologies planaires. Les éléments utilisés pour isoler les différents dispositifs actifs, passifs et les lignes d'interconnexion dans les technologies CMOS et BiCMOS sur substrat de silicium, sont les diélectriques et les jonctions *pn* inversement polarisées. Leur mission est de réduire au minimum toute interaction électrique non-désirée entre les dispositifs. La qualité de l'isolement au fur et à mesure que la température augmente (i.e. transfert de charge à travers les diélectriques, dégradation des diélectriques, courant de fuite, capacité et tension de claquage des jonctions *pn* polarisées en inverse, etc.) a une importance clé pour la détermination de la tenue en température d'une technologie des circuits intégrés.

2.3.1. Dioxyde de Silicium

Les propriétés physiques de l'oxyde de silicium (SiO_2) ont joué un rôle très important dans le développement des technologies planaires sur silicium. Le SiO_2 est un matériau isolant presque idéal, largement utilisé pour la passivation de la surface du silicium ainsi que pour l'isolation des couches d'interconnexion, du fait de sa bonne stabilité chimique. De plus, comme le SiO_2 est très imperméable à la diffusion des impuretés dopantes et qu'il peut être gravé très sélectivement, il est aussi utilisé comme masque pour le dopage sélectif du silicium. Avec l'arrivée des technologies MOS, le SiO_2 prend une part active dans le fonctionnement des dispositifs semiconducteurs. Cette fois ci, une autre propriété clé du SiO_2 est concernée: la très basse concentration des états de surface de l'interface $Si-SiO_2$, laquelle après quelques traitements reste très largement inférieure à 1% de la concentration des électrons dans un canal en inversion forte.

Le SiO_2 est utilisé dans les technologies CMOS es BiCMOS principalement comme:

- ⊖ Diélectrique entre le 1^{er} niveau de Poly-Si et la surface du silicium, pour la formation de la grille des transistors MOS.
- ⊖ Diélectrique entre les deux niveaux de Poly-Si, pour la formation des capacités double Poly-Si.
- ⊖ Isolant entre les différents niveaux d'interconnexion.
- ⊖ Oxyde de champ et couche de passivation, pour la protection de la surface du silicium.

La qualité des diélectriques et isolants est principalement affectée par deux facteurs qui dégradent d'une manière irréversible leurs propriétés : le passage d'un courant électrique et la pollution induite pendant les procédés de fabrication ou due à l'environnement et aux conditions d'opération. La conduction à travers le SiO_2 est principalement générée par l'injection d'électrons chauds ou par effet tunnel sous l'influence de champs électriques forts. Dans les cas de l'injection d'électrons chauds, les électrons, ayant une énergie supérieure à celle de la barrière de potentiel à l'interface $Si-SiO_2$, peuvent pénétrer la bande de conduction de l'oxyde après avoir subi, par collision, une réorientation de leur trajectoire. Concernant l'effet tunnel, l'application des forts champs électriques ($>10MV/cm$) augmente considérablement la probabilité de transmission à travers la barrière de potentiel à l'interface $Si-SiO_2$. Du premier cas il résulte un piégeage des électrons dans l'oxyde, tandis que l'effet tunnel génère un phénomène de dégradation plus complexe qui entraîne la formation de charges positives et une forte augmentation de la densité d'états à l'interface $Si-SiO_2$.

Les phénomènes de conduction à travers l'oxyde sont, dans la plupart des cas, des effets indésirables qui sont normalement contrôlés en limitant la valeur maximale du champ électrique entre l'oxyde et le semiconducteur. Pendant les vingt dernières années, l'épaisseur des oxydes de grille a subi une constante réduction, permettant une augmentation de la densité d'intégration, qui n'a pas été accompagnée d'une réduction équivalente des niveaux de tension des signaux. Par conséquent la

valeur du champ électrique à l'interface Si-SiO₂ n'a pas cessé d'augmenter, ce qui a fait de l'injection des porteurs chauds un des problèmes de fiabilité les plus importants des technologies MOS et BiCMOS.

2.3.1.1. Injection d'électrons chauds

Le premier phénomène physique important qui s'est manifesté dans les transistors MOS sous l'influence des larges tensions de drain est l'effet d'ionisation par impact dans le canal, côté drain. Ce phénomène, défini comme le passage d'un électron de la bande de valence à la bande de conduction, nécessite le transfert d'une quantité d'énergie au moins équivalente à la valeur de la bande interdite. La création des paires électron-trou entraîne une augmentation involontaire du courant de substrat, qui peut générer ensuite une variation de la tension du substrat et de la tension du seuil des transistors MOS. Dans le pire des cas, un courant de substrat assez fort pourrait polariser en direct la jonction source-substrat et verrouiller dans son état passant le thyristor parasite (latch-up), présent dans les technologies CMOS et BiCMOS.

L'injection d'électrons chauds dans l'oxyde se manifeste quand le champ électrique longitudinal atteint une valeur telle, qu'un bon nombre des électrons peuvent gagner une énergie supérieure à la barrière de potentiel à l'interface Si-SiO₂. Dans ce cas, les électrons franchissent la barrière de potentiel en générant un courant détectable dans la grille. Comme la hauteur de la barrière est approximativement égale à $3eV$, le courant de grille sera inférieur au courant de substrat de quelques ordres de magnitude. Non seulement des électrons du canal peuvent être injectés dans l'oxyde, mais les électrons générés thermiquement dans les substrats peuvent aussi être injectés, pourvu que le substrat soit assez négativement polarisé pour permettre le franchissement de la barrière de potentiel.

À des températures normales, l'influence des électrons chauds générés dans le substrat est moins importante pour la fiabilité des transistors MOS que celle des électrons chauds du canal, à cause de la faible quantité de paires électron-trou thermiquement générée. Mais il faut bien prendre en compte que l'effet des électrons chauds générés dans le substrat augmente avec la température, puisqu'ils sont générés thermiquement. Au contraire, l'effet des électrons chauds du canal diminue au fur et à mesure que la température augmente à cause d'une réduction de la mobilité de électrons.

L'injection d'électrons chauds dans l'oxyde a été modélisée selon différentes interprétations du phénomène, arrivant cependant à des résultats comparables. Un des ces modèles, nommé modèle de la *température équivalente*, considère que la fonction de distribution des électrons est maxwellienne à la température équivalente $Te(x)$, où x représente la position le long du canal du transistor. La densité de courant injectée Jg est donc donnée par l'équation d'émission thermoionique de Richardson [30] [4],

$$(47) \quad Jg = q \cdot ns \cdot \frac{-k \cdot Te(x)}{2 \cdot \pi \cdot m^*} \sqrt{\frac{1}{2}} \cdot e^{\left(\frac{q\Phi_b}{kTe(x)}\right)}$$

$$(48) \quad Te(x) = T + (Te(0) - T) \cdot e^{-\frac{x}{\frac{5}{3} \cdot \lambda_s \cdot \tau_e}} + \frac{2 \cdot q \cdot x}{5 \cdot k_0} \cdot E(x - x') \cdot e^{-\frac{x'}{\frac{5}{3} \cdot \lambda_s \cdot \tau_e}} \cdot f(x')$$

où T est la température absolue, ns est la densité des électrons à l'interface Si-SiO₂, vs est la vitesse des électrons et te le temps de relaxation (21). Si $E(x)$ est connu (par résolution numérique des équations du semiconducteur appliquées au transistor MOS) l'équation (48) permet d'obtenir la valeur de la température équivalente $Te(x)$ en fonction de la position x dans le canal. La courant passant par la grille d'un transistor MOS de largeur W et longueur L est obtenu en intégrant Jg le long du canal,

$$(49) \quad I_g = W \int_0^L J_g(x) dx$$

L'analyse et les conclusions concernant le comportement des électrons chauds dans les transistors à canal n , s'appliquent aussi au cas de trous chauds dans les dispositifs MOS de canal p . Mais, les injections de porteurs mentionnées affectent plus fortement les transistors NMOS que les transistors PMOS, car la mobilité des électrons, dans le silicium et dans l'oxyde de silicium, est plus importante que celle des trous.

2.3.1.2. Transport de charge dans les diélectriques

Les expressions pour la densité de courant J , due aux différents phénomènes de transport de charge dans les diélectriques, sont résumées dans le Tableau 4 afin d'analyser leur dépendance par rapport à la température absolue T et la tension appliquée V . Dans tous les types de conduction mentionnés, la densité de courant augmente selon la tension appliquée (même très fortement quand l'effet tunnel est présent) et dans plusieurs cas elle augmente très sensiblement avec la température.

Tableau 4. Procès de conduction basiques dans les diélectriques [1].

Type de conduction.	Densité de courant.	Dépendance de la tension V et la température T .
Emission Schottky	$J = A^* T^2 \exp\left(\frac{-q}{kT} \left(\Phi_b - \sqrt{\frac{q\xi}{4\pi\epsilon_i}}\right)\right)$	$\propto T^2 \exp\left(\frac{a\sqrt{V}}{T} - \frac{q\Phi_b}{kT}\right)$
Emission Frenkel-Poole	$J \propto \xi \exp\left(\frac{-q}{kT} \left(\Phi_b - \sqrt{\frac{q\xi}{\pi\epsilon_i}}\right)\right)$	$\propto V \exp\left(\frac{2a\sqrt{V}}{T} - \frac{q\Phi_b}{kT}\right)$
Emission tunnel	$J \propto \xi^2 \exp\left(\frac{-4}{3q\hbar\xi} \sqrt{2m^*} (q\Phi_q)\right)^3$	$\propto V^2 \exp\left(\frac{-b}{V}\right)$
Limitation par charge d'espace	$J = \frac{8\epsilon_i \mu V^2}{9d^3}$	$\propto V^2$
Ohmique	$J \propto \xi \exp\left(\frac{-\Delta E_{ae}}{kT}\right)$	$\propto V \exp\left(\frac{-c}{T}\right)$
Conduction Ionique	$J \propto \frac{\xi}{T} \exp\left(\frac{-\Delta E_{ai}}{kT}\right)$	$\propto \frac{V}{T} \exp\left(\frac{-p}{T}\right)$

A^* = constante effective de Richarson, Φ_b = hauteur de la barrière, ξ = Champ électrique, ϵ_i = Permittivité dynamique du diélectrique, m^* = masse effective, d = épaisseur du diélectrique, ΔE_{ae} = Energie d'activation des électrons, ΔE_{ai} = Energie d'activation des ions, $a = (q/4\pi\epsilon_i)^{1/2}$, $V = \xi d$. Et b, c, p sont des constantes positives qui ne dépendent ni de la température ni de la tension.

2.3.1.3. Fiabilité des diélectriques

La variation du temps de claquage des diélectriques avec la température et la valeur du champ électrique ont été largement étudiées, du fait de l'importance du phénomène de dégradation des diélectriques pour la détermination de la durée de vie des dispositifs MOS. La distribution du temps de défaillance tf la plus souvent acceptée est logarithmique,

$$(50) \quad tf = A \exp\left(\frac{-Eg}{kT} - B E_{ox}\right)$$

où A est une constante, Ea l'énergie d'activation, Eox le champ électrique appliquée et B le paramètre d'accélération de champ. La relation (50) montre une forte accélération des processus de dégradation des diélectriques au fur et à mesure que la température et le champ électrique appliqué augmentent.

2.3.2. Jonctions pn

Une jonction pn est formée par la juxtaposition de deux régions d'un même monocristal semiconducteur, dopées respectivement avec des impuretés du type accepteur et donneur. Chacune des régions est un semiconducteur extrinsèque, dont la concentration effective des dopants donneurs et accepteurs est respectivement nommée Nd et Na , voir Figure 22.

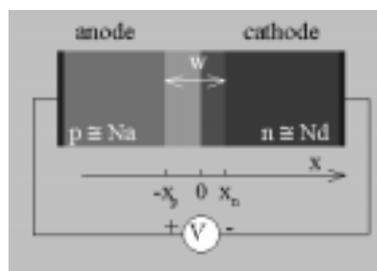


Figure 22. Structure et polarisation d'une jonction pn.

Les densités de porteurs libres dans chacune des régions sont alors données par: $n_n = Nd$, $n_p = ni^2/Na$, $p_p = Nd$, $p_n = ni^2/Nd$. La différence des densités de donneurs et accepteurs $Nd-Na$ passe d'une valeur négative dans la région p (l'anode) à une valeur positive dans la région n (la cathode). La variation de la densité des dopants près de la jonction peut être plus ou moins abrupte, linéaire, exponentielle, etc., selon la méthode de fabrication.

2.3.2.1. La jonction pn à l'équilibre

Théoriquement, si on connecte les deux régions du silicium, dopées p et n respectivement, initialement les deux niveaux de Fermi ne sont pas alignés. Les trous, majoritaires dans la région de type p , diffusent vers la région de type n où ils se recombinent avec les électrons. Et les électrons, majoritaires dans la région de type n , diffusent vers la région de type p où ils se recombinent avec les trous (voir Figure 23).

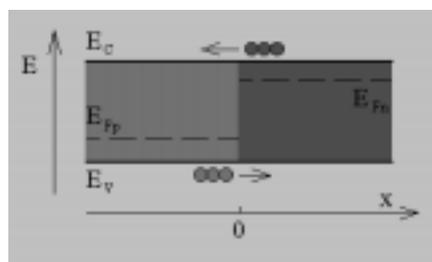


Figure 23. Formation d'une charge d'espace dans la jonction pn.

La recombinaison des porteurs libres des deux cotés de la jonction fait apparaître une charge d'espace résultant de la présence des donneurs et accepteurs ionisés, dont les charges ne sont plus intégralement compensées par celles des porteurs libres. Il s'établit alors, au voisinage de la jonction

métallurgique, un champ électrique qui s'oppose à la diffusion des porteurs majoritaires. L'équilibre thermodynamique est établi quand la force résultante de l'apparition du champ électrique équilibre la force de diffusion associée aux gradients de concentration de porteurs libres, limitant ainsi la taille de la région de charge d'espace. La présence d'une charge d'espace entraîne l'existence d'un champ électrique et d'une variation de potentiel, qui se traduit dans la courbure des bandes de la Figure 24. La différence de potentiel entre la région n et la région p est appelée *barrière de diffusion* V_{bi} , parce que c'est la barrière qui équilibre les forces de diffusion.

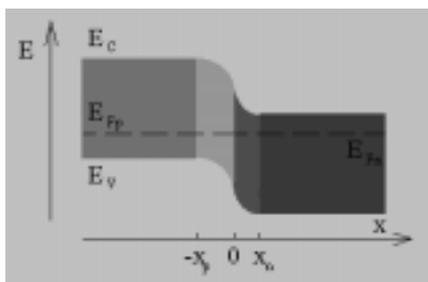


Figure 24. Bandes d'énergie de la jonction pn à l'équilibre.

A l'équilibre thermodynamique, les courants sont nuls et le niveau de Fermi est horizontal dans toute la structure. Le niveau de Fermi étant constant dans toute la structure, les concentrations des électrons dans chacune des régions s'écrivent, à partir de (10),

$$(51) \quad n_n = n_i \exp\left(\frac{-(E_{in} - E_f)}{kT}\right)$$

$$(52) \quad n_p = n_i \exp\left(\frac{-(E_{ip} - E_f)}{kT}\right)$$

ce qui permet d'écrire,

$$(53) \quad V_{bi} = \frac{E_{ip} - E_{in}}{q} = \frac{kT}{q} \ln\left(\frac{N_d N_a}{n_i^2}\right)$$

Dans la Figure 25 on voit l'évolution de la valeur de la barrière de potentiel d'une jonction pn sur silicium, en fonction de la température.

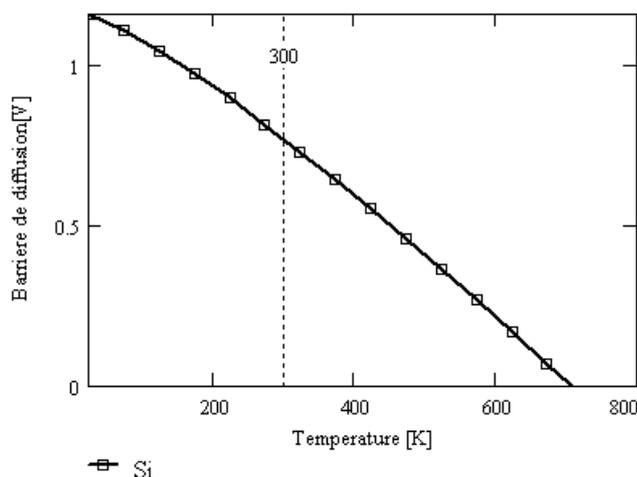


Figure 25. Hauteur de la barrière de diffusion en fonction de la température.

2.3.2.2. La jonction pn en dehors de l'équilibre

Les densités des porteurs dans le silicium en dehors de l'équilibre ne sont plus données par les équations (10) et (11) et $n.p_{ni}$. On peut cependant conserver une écriture semblable en introduisant les *quasi-niveaux* de Fermi E_{fn} et E_{fp} . On définit,

$$(54) \quad n = n_i ? e^{\frac{-(E_i - E_{fn})}{kT}}$$

$$(55) \quad p = n_i ? e^{\frac{(E_i - E_{fp})}{kT}}$$

On peut alors réécrire le produit pn à partir de (54) et (55),

$$(56) \quad p ? n = n_i^2 ? e^{\frac{E_{fn} - E_{fp}}{kT}}$$

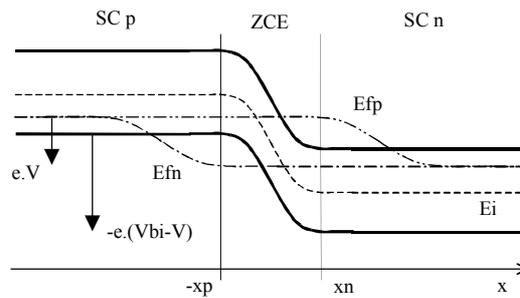


Figure 26. Bandes d'énergie de la jonction pn en dehors de l'équilibre.

A partir des équations (32), et (54) on peut aussi réécrire l'équation de la densité de courant j_n ,

$$(57) \quad j_n = \mu_n ? n ? \frac{f E_{fn}}{f x}$$

Si on néglige la recombinaison dans la zone de charge d'espace (ZCE) d'une jonction pn polarisée avec une tension V (voir la Figure 22), la densité de courant en régime permanent est constante dans la ZCE, c'est à dire: $j_n(-xp) = j_n(xn)$. Donc,

$$(58) \quad \frac{\frac{f E_{fn}}{f x} \Big|_{-xp}}{\frac{f E_{fn}}{f x} \Big|_{xn}} = \frac{n(xn)}{n(-xp)}$$

Le quasi-niveau de Fermi E_{fn} varie là où les électrons sont minoritaires, c'est à dire, du côté p de la ZCE (voir Figure 26). De même pour les trous, dont E_{fp} varie du côté n de la ZCE. Par conséquent, on peut considérer E_{fn} et E_{fp} constants dans la ZCE et $E_{fn} - E_{fp} = q.V$.

On peut alors réécrire le produit pn dans la ZCE, à partir de (56),

$$(59) \quad p ? n = n_i^2 ? e^{\frac{qV}{kT}}$$

A partir de (54), (55), (10) et (11) on obtient la concentration des porteurs minoritaires aux limites de la ZCE,

$$(60) \quad n(-xp) = n_{p0} e^{\frac{qV}{kT}}$$

$$(61) \quad p(xn) = p_{n0} e^{\frac{qV}{kT}}$$

2.3.2.3. Densités de porteurs dans les zones neutres et densité idéale de courant.

Dans un régime d'injection faible, la distribution des porteurs majoritaires n'est pas affectée. Le champ électrique dans les régions neutres est faible, alors que le gradient de la concentration des porteurs minoritaires est important à cause du phénomène d'injection. Le courant est essentiellement un courant de diffusion limitée par le phénomène de recombinaison dans le volume. La distribution des porteurs minoritaires en régime stationnaire est alors donnée par les équations de continuité suivantes,

$$(62) \quad \frac{fn}{ft} = Dn \frac{f^2 n}{fx^2} - \frac{n - n_{p0}}{\tau_n} = 0$$

$$(63) \quad \frac{fp}{ft} = Dp \frac{f^2 p}{fx^2} - \frac{p - p_{n0}}{\tau_p} = 0$$

Le deuxième terme prend en compte la recombinaison des porteurs minoritaires excédents (voir équation (36)) dans les zones neutres fortement dopées. Considérant les conditions aux limites suivantes : $n(-) = n_{p0}$, $p(-) = p_{n0}$, (60) et (61) on obtient,

$$(64) \quad n = n_{p0} + n_{p0} e^{\frac{qV}{kT}} - 1 \sqrt{\frac{(x-xp)}{Ln}} \quad \text{avec } x > 0 \text{ et } |x| > |xp|$$

$$(65) \quad p = p_{n0} + p_{n0} e^{\frac{qV}{kT}} - 1 \sqrt{\frac{(xn-x)}{Lp}} \quad \text{avec } x > 0 \text{ et } x > xn$$

Où $Lp = (Dp \cdot \tau_p)^{1/2}$ et $Ln = (Dn \cdot \tau_n)^{1/2}$ sont respectivement les longueurs de diffusion des trous et des électrons. Les courants d'électrons dans la région de type p et de trous dans la région de type n sont des courants de diffusion, donnés à partir de (29) et (30) par,

$$(66) \quad jn(x) = \frac{q \cdot ni^2 Dn}{Na \cdot Ln} e^{\frac{qV}{kT}} - 1 \sqrt{\frac{(x-xp)}{Ln}}$$

$$(67) \quad jp(x) = \frac{q \cdot ni^2 Dp}{Nd \cdot Lp} e^{\frac{qV}{kT}} - 1 \sqrt{\frac{(xn-x)}{Lp}}$$

Le courant total traversant la jonction est la somme des courants d'électrons et de trous en un même point. Si on suppose qu'il n'y a pas de recombinaison dans la ZCE, les courants j_n et j_p sont constants dans cette zone et alors $j_n(-xp) = j_n(xn)$ et $j_p(-xp) = j_p(xn)$. Le courant total J est ainsi donné par la somme des équations (66) et (67) évaluées respectivement à xp et xn ,

$$(68) \quad J = q \left[n_i(T)^2 \left(\frac{D_p}{N_d} \frac{1}{L_p} + \frac{D_n}{N_a} \frac{1}{L_n} \right) e^{-\frac{qV}{kT}} - 1 \right] = J_s \left(e^{-\frac{qV}{kT}} - 1 \right)$$

2.3.2.4. Génération et recombinaison dans la ZCE

Jusqu'à maintenant les phénomènes de générations thermiques et recombinaisons dans la ZCE n'ont pas été considérés. La densité de porteurs dans la ZCE étant faible, le taux est donné par les équations (35) et (59),

$$(69) \quad U = \frac{N_t \left(n_i^2 e^{-\frac{qV}{kT}} - 1 \right)}{2(n_i + p + n)}$$

En polarisation directe la tension V est positive, l'exponentielle est supérieure à 1, donc il y a recombinaison de porteurs dans la ZCE. En polarisation inverse la tension V est négative, l'exponentielle est inférieure à 1, alors il y a génération thermique de porteurs dans la ZCE. L'équation du courant s'écrit alors,

$$(70) \quad J = J_s \left(e^{-\frac{qV}{kT}} - 1 \right) + q U \left(\frac{xp}{xn} \right)$$

2.3.2.5. Courants de fuite

Dans le cas d'une polarisation inverse suffisamment forte ($|V| \ll kT/q$) l'expression du taux de recombinaison (69) se simplifie dans l'équation (37), d'une part parce que le terme exponentiel devient négligeable devant 1, et d'autre part parce que les concentrations des porteurs n et p deviennent inférieures à n_i . En explicitant le deuxième terme de l'équation (70), le courant de génération s'écrit,

$$(71) \quad J_g = -q N_t \left(\frac{n_i}{2} \right) W$$

Où W est la largeur de la ZCE. On considère le cas d'une jonction abrupte, entre deux régions semi-conductrices uniformément dopées (respectivement, N_a et N_d). Si on applique le théorème de Gauss dans le volume formé par la ZCE, la largeur W de la ZCE de la jonction pn en fonction de la tension appliquée V (voir Figure 22), s'écrit,

$$(72) \quad W = \sqrt{\frac{2 \epsilon_s \epsilon_0}{q} \left(\frac{N_a + N_d}{N_a N_d} \right) \left(V_{bi} - 2 \frac{kT}{q} - V \right)}$$

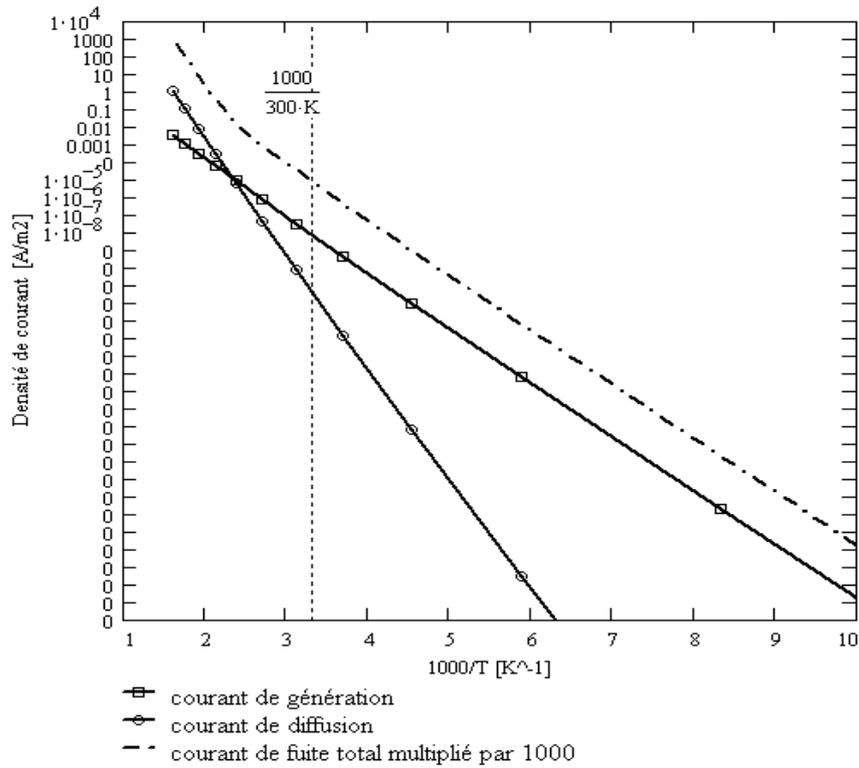


Figure 27. Densité du courant de fuite d'une jonction pn polarisée à -5 volts, d'après (73).

La densité du courant de fuite d'une jonction polarisée inverse est donnée par la somme de J_s et J_g , donc à partir des équations (68), (71) et (72) on obtient,

$$(73) \quad J_f = -q \left[n_i(T)^2 \left(\frac{D_p}{N_d L_p} + \frac{D_n}{N_a L_n} \right) + N_t C \right] n_i(T) \sqrt{\frac{\epsilon_s q}{2} \left(\frac{1}{N_d} + \frac{1}{N_a} \right) \sqrt{V_{bi}(T) - 2} \frac{k T}{q} - V}$$

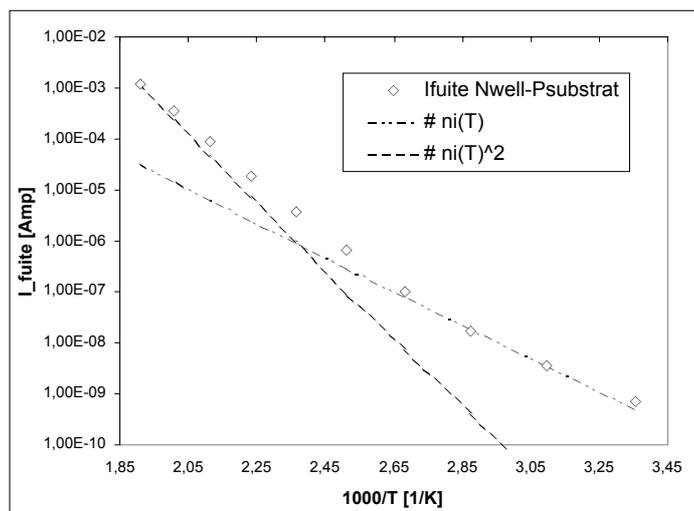


Figure 28. Courant de fuite d'une jonction n-p-, polarisée à -5 volts, en fonction de la température (de 25°C à 250°C). La température de transition, entre J_g et J_s , est de 150°C environ.

Les deux termes dans (73) dépendent fortement de la température, essentiellement à cause de la concentration intrinsèque $n_i(T)$. Dans la Figure 27 les densités du courant de diffusion J_s (proportionnel à $n_i^2(T)$) et de génération J_g (proportionnel à $n_i(T)$), d'une jonction polarisée inversement à $V=-5$ volts, ont été calculées afin d'apprécier leur évolution exponentielle selon la température. La densité du courant de fuite totale J_f (multipliée par 1000 dans la Figure 27 pour pouvoir mieux la visualiser) suit J_g à basse température jusqu'à un point d'inflexion à partir duquel elle suit l'évolution de la densité du courant de diffusion J_s , doublant sa pente exponentielle selon la température. La température de cette transition entre J_g et J_s varie selon les caractéristiques des jonctions. Cela peut être observé dans la Figure 28 et la Figure 29, où les variations des courants de fuite avec la température sont comparées à l'évolution de $n_i(T)$ et de $n_i^2(T)$.

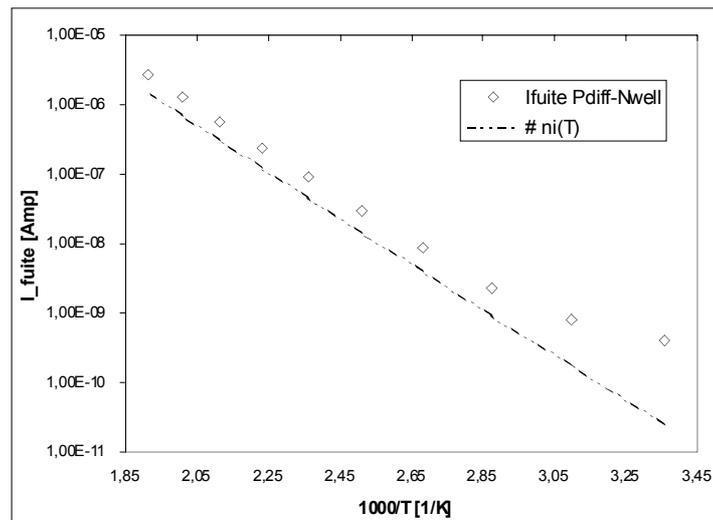


Figure 29. Courant de fuite d'une jonction p+n-, polarisée à -5 volts, en fonction de la température (de 25°C à 250°C). La température de transition, entre J_g et J_s , est supérieure à 250°C.

Une telle évolution, de la dégradation de l'isolation selon la température, limite fortement la performance des circuits intégrés en haute température. Le côté le plus faiblement dopé est celui qui détermine, d'après (73), la valeur de J_f . Il en résulte que, pour minimiser les courants de fuite, les concentrations des dopants, aux deux côtés de la jonction, doivent être maximisées et doivent être comparables (i.e. $N_a \approx N_d$).

2.3.2.6. Tension de rupture

Les courants de fuite d'une jonction pn abrupte augmentent en fonction de la racine carrée de la tension de polarisation inverse, comme précédemment mentionné, quand le phénomène de génération des porteurs dans la ZCE est dominant. Mais on observe qu'à partir d'une certaine valeur V_r de tension inverse, le courant tend vers l'infini, celui-ci étant uniquement limité par la résistance en série avec la jonction. On est alors en présence d'une rupture complète de l'isolation, par claquage de la jonction pn polarisée en inverse.

2.3.2.6.a. Effet Zener et effet d'avalanche

La largeur de la ZCE d'une jonction pn abrupte augmente selon la racine carrée de la tension de polarisation inverse, alors que la tension aux limites de la ZCE augmente linéairement. Il en résulte que le champ électrique à l'intérieur de cette région augmente aussi. Le champ électrique dans la ZCE est maximum à la jonction, et varie selon la largeur de la ZCE du côté n (x_n) ou du côté p (x_p) de la jonction. Dans le cas d'une jonction abrupte, entre deux régions semiconductrices uniformément dopées, le champ électrique maximum est donné par,

$$(74) \quad \xi_m = \frac{qNd}{\epsilon_{sc}} x_n = \frac{qNd}{\epsilon_{sc}} \sqrt{\frac{2\epsilon_{sc}}{q} \frac{Na}{Nd} \frac{1}{Na+Nd} (V_{bi} - 2) \frac{kT}{q} - V}$$

Ce qui donne, pour une polarisation inverse telle que $-V \gg V_{bi}$,

$$(75) \quad \xi_m \propto \sqrt{\frac{2q}{\epsilon_{sc}} \frac{NaNd}{Na+Nd} (-V)}$$

La tension aux bornes de la ZCE ne peut pas augmenter indéfiniment parce qu'il existe une limite à la valeur de ξ_m , due à l'augmentation de la force électrique $F = -q \cdot \xi_m$ appliquée aux électrons liés. Quand cette force dépasse la force de liaison des électrons de valence, ces électrons sont libérés, le matériau devient conducteur et, en conséquence, la tension aux bornes de la ZCE n'augmente plus. La paire électron-trou créée par l'ionisation est immédiatement évacuée par le champ électrique, provoquant un courant inverse uniquement limité par la résistance en série avec la jonction. Dans le silicium, ce champ maximum est de l'ordre de 10^6 V/cm. D'après (75), la tension limite correspondante, dite *tension Zener*, est une fonction de la concentration des dopages, dans la jonction pn .

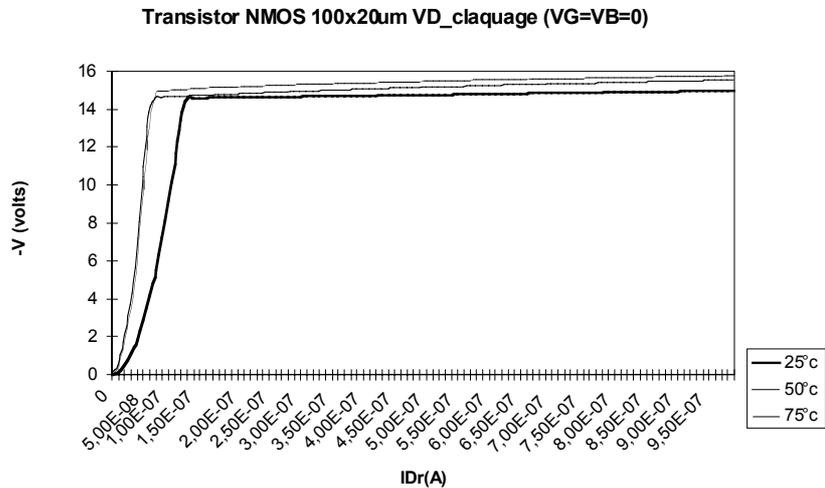


Figure 30. Courbe caractéristique en polarisation inverse de la jonction $p+n$ du drain d'un transistor NMOS, en fonction de la température.

Dans la pratique cet effet n'est observable que dans les jonctions fortement dopées, où la ZCE est très étroite (500Å environ). Autrement, c'est l'effet d'avalanche qui entraîne le claquage de la jonction. Dès que le champ électrique est de l'ordre de 10^5 V/cm, l'accélération acquise par les porteurs, qui transportent le courant inverse, est suffisante pour leur permettre de générer des paires électron-trou par ionisation par choc des atomes du réseau cristallin. Ces paires sont à leur tour accélérées et peuvent créer d'autres paires, etc. Le processus d'avalanche nécessite non seulement un champ élevé mais aussi une distance suffisante permettant l'accélération des porteurs. C'est pour cette raison que,

dans les jonctions fortement dopées, l'effet Zener précède l'effet d'avalanche, car la ZCE n'a pas la largeur nécessaire. Mais dans la plupart des jonctions pn , l'effet d'avalanche précède celui de Zener, parce que son champ de seuil est plus faible.

Les deux effets peuvent être distingués expérimentalement par la variation avec la température de la tension de claquage V_r . Dans le cas de l'effet Zener, cette tension présente un coefficient de température négatif, car il dépend de la hauteur de la bande interdite E_g [1], laquelle décroît avec la température. Au contraire, lorsque le claquage résulte de l'effet d'avalanche, le coefficient de température de V_r est positif. Les porteurs chauds qui traversent la ZCE perdent une partie de leur énergie à cause de leur interaction avec le réseau cristallin, caractérisée par un libre parcours moyen λ qui diminue avec la température du fait de l'augmentation des vibrations atomiques. Il en résulte que, à une plus haute température et sous un champ électrique constant, les porteurs cèdent plus d'énergie par distance parcourue au réseau cristallin. C'est pourquoi il faut une tension plus élevée pour qu'ils puissent accumuler l'énergie nécessaire pour la génération des paires électron-trou.

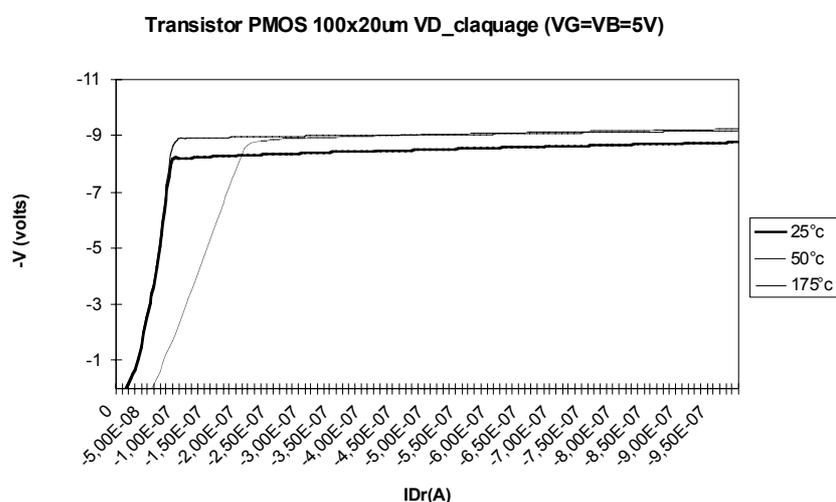


Figure 31. Courbe caractéristique en polarisation inverse de la jonction $n+p$ du drain d'un transistor PMOS, en fonction de la température.

Généralement, les tensions de claquage inférieures à 5 volts résultent de l'effet Zener, et les tensions de claquage supérieures à 7 volts résultent de l'effet d'avalanche. Dans les technologies des circuits intégrés $CMOS$ et $BiCMOS$ à 5 volts, la tension de claquage des jonctions pn , utilisées pour l'isolation des dispositifs (jonction caisson-substrat, diffusion-caisson, diffusion-substrat, etc.), est toujours supérieure à 7 volts, afin de pouvoir générer une isolation effective pour des différences de potentiel inférieures ou égal à 5 volt. C'est la raison pour laquelle la tension de claquage de ces jonctions augmente selon la température, ne représentant pas une dégradation ou un risque de dégradation de l'isolation en haute température. Les courbes caractéristiques en polarisation inverse des jonctions $n+p$ - et $p+n$ - correspondant respectivement aux diffusions des drains des transistors PMOS et NMOS des technologies étudiées, ont été mesurées à différentes températures (voir Figure 30 et Figure 31) pour évaluer la tenue en température de la tension de claquage de ces types de jonction.

2.3.2.7. Capacité parasite

Si on applique le théorème de Gauss dans le volume formé par la partie de la ZCE correspondant au semiconducteur de type n , on obtient l'expression suivante, pour la charge Q_n dans le volume mentionné,

$$(76) \quad Q_n = \epsilon_{sc} \gamma A \gamma \xi_m$$

Où A est la surface de la jonction et ξ_m est le champ électrique maximum dans la ZCE, donné par (74). La dernière relation, dans le cas d'une jonction abrupte pas trop dissymétrique et à dopages constants, s'écrit maintenant,

$$(77) \quad Q_n = A \gamma \epsilon_{sc} \gamma \sqrt{\frac{2 \gamma q}{\epsilon_{sc}} \gamma \frac{Na \gamma Nd}{Na + Nd} \gamma \sqrt{V_{bi} - 2 \gamma \frac{k \gamma T}{q} - V} \gamma}$$

Une variation de dV de la tension de polarisation entraîne une variation dQ_n de la charge dans la ZCE. La jonction pn polarisée en inverse se comporte donc comme un condensateur de capacité dynamique C_t ,

$$(78) \quad C_t = \frac{fQ}{fV} = \frac{\epsilon_{sc} \gamma A}{\sqrt{\frac{2 \gamma \epsilon_{sc}}{q} \gamma \frac{Na + Nd}{Na \gamma Nd} \gamma \sqrt{V_{bi}(T) - 2 \gamma \frac{k \gamma T}{q} - V} \gamma}} = \frac{\epsilon_{sc} \gamma A}{W}$$

Où W est la largeur totale de la ZCE. C'est donc le côté le moins dopé de la jonction qui déterminera la valeur de la capacité. Cette capacité différentielle est appelée *capacité de transition* C_t de la jonction; elle varie selon la tension inverse appliquée et selon la température. Dans la Figure 32 on peut apprécier l'évolution de la *capacité de transition* C_t en fonction de la tension de polarisation inverse et de la température. Même si les valeurs des *capacités de transition* augmentent selon la température, leurs variations ne sont pas dramatiques et, par conséquent, ne représentent pas une véritable limitation de la performance des circuits intégrées en haute température.

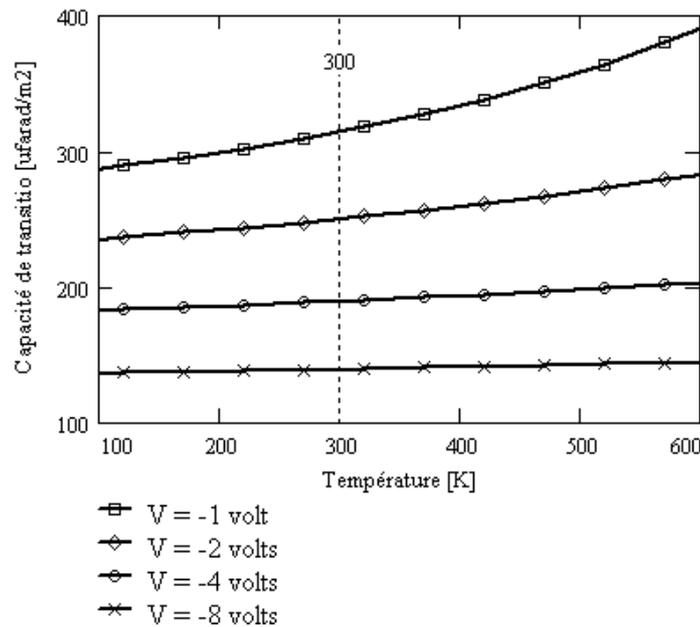


Figure 32. Capacité de transition d'une jonction pn en fonction de la tension et la température.

2.3.2.8. Latchup

L'existence, dans un circuit CMOS ou BiCMOS, de structures du type *Diode de Shockley* (ou diodes *pnpn*) entre les sources d'alimentation (voir Figure 33), entraîne le risque d'un basculement involontaire (dit *latchup*) dans l'état direct-passant de la diode *pnpn*, avec des conséquences destructrices pour le circuit intégré. Les valeurs des courants de déclenchement et de maintien sont les paramètres les plus importants décrivant la robustesse aux *latchups* d'un circuit. Ces courants dépendent fortement des paramètres technologiques, des facteurs géométriques et de la température. Par exemple, le courant de maintien I_M , dans le modèle de la Figure 33, est approximativement donné par [61],

$$(79) \quad I_M(T) = \frac{V_{be_{PNP}}(T)}{R_n} + \frac{V_{be_{NPN}}(T)}{R_p} - |I_{f_{c-s}}(T)| \quad \text{pour } \beta_{npn} \cdot \beta_{pnp} \gg 1$$

$I_{f_{c-s}}$ étant le courant de fuite dans la jonction caisson-substrat, donné par l'équation (73) multiplié par la surface de la jonction mentionnée.

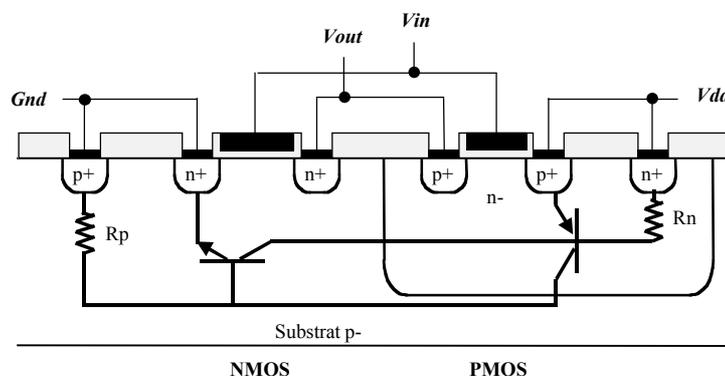


Figure 33. Transistors bipolaires parasites responsables des latchups dans un inverseur CMOS.

Le courant de déclenchement I_D est normalement décrit par la relation empirique suivante [61],

$$(80) \quad I_D(T) = I_D(T_0) \sqrt{\frac{T}{T_0}}^{-TCD} \quad \text{avec typiquement } 3 < TCD < 4.5$$

Au fur et à mesure que la température augmente, les courants de déclenchement et de maintien diminuent progressivement, augmentant les risques du *latchup* [13]. Cependant, les structures peuvent être modifiées selon quelques techniques d'implémentation (voir section 4.2.1.3), afin d'augmenter sensiblement (jusqu'à 300°C selon [3]) la robustesse aux *latchups*. Dans le cas des technologies CMOS et BiCMOS sur substrat epitaxial (substrat de base fortement dopé plus une couche epitaxiale mince faiblement dopée), la plus basse résistivité du substrat de base fait que les courants injectés sont plutôt absorbés par le substrat conducteur qu'injectés dans la base du transistor *nnp* latéral de la Figure 33, rendant encore plus effective toute technique d'implémentation. Il en résulte que la valeur du courant de maintien augmente sensiblement, dans les technologies sur substrat epitaxial [3].

2.4. Conclusions

Les matériaux semiconducteurs, monocristallins ou polycristallins, sont affectés par la température de différentes manières. D'une part, une concentration des dopants minimale doit être assurée pour contrôler le comportement électrique du matériau à une température donnée. D'autre part, la mobilité et le coefficient de diffusion des porteurs diminuent progressivement en fonction de la température.

Tout cela concerne plutôt la robustesse des dispositifs et des circuits à haute température, sans présenter de problèmes de fiabilité majeurs.

Mais, en ce qui concerne les métallisations et les isolants diélectriques minces, les problèmes de fiabilité, activés respectivement par une densité de courant (électromigration) et un champ électrique (vieillissement par les phénomènes de conduction dans les oxydes minces) sont de plus en plus présents en haute température. La réduction de la densité de courant et la multiplication des contacts et vias, à une température donnée, permet d'améliorer assez efficacement la robustesse et la fiabilité des conducteurs métalliques intégrés. La corrosion est évitée par le biais des encapsulations étanches sous atmosphère neutre. En diminuant la surface de la grille des transistors MOS, on peut réduire les conséquences de l'injection dans l'oxyde de grille de porteurs chauds, générés thermiquement dans le substrat, tout en améliorant la robustesse des circuits à haute température. L'utilisation de transistors PMOS dans les parties critiques d'un circuit est un moyen intéressant pour augmenter la fiabilité des circuits intégrés en haute température.

La variation exponentielle des courants de fuite des jonctions pn selon la température, semble être le problème majeur, du point de vue de la robustesse des circuits, parce que ces jonctions font partie de la structure de la majorité des dispositifs intégrés actifs et passifs. Ces courants de fuite peuvent aussi avoir un impact considérable sur la fiabilité, à travers le phénomène du latchup.

3. Physique des dispositifs en Haute Température

Le comportement en haute température des composants intégrés passifs et actifs dépend des caractéristiques en haute température des matériaux conducteurs, semiconducteurs et isolants avec lesquels ils sont fabriqués, et dépend aussi de leur structure propre et de la manière dont ils sont polarisés. Les caractéristiques et les propriétés en haute température de ces dispositifs doivent être étudiées, afin de pouvoir les exploiter et ainsi concevoir des circuits intégrés durcis, fonctionnels à des températures supérieures à la température maximale spécifiée par le fabricant.

3.1. Résistances

Les résistances, ainsi que d'autres composants passifs, ont conservé leur importance dans le domaine des circuits analogiques. Les propriétés les plus importantes de ces dispositifs, en haute température, sont la linéarité (i.e. leur fidélité à la loi de Ohm), la stabilité thermique (i.e. ces coefficients thermiques et la reproductibilité de leur caractéristiques après un nombre de cycles thermiques), la capacité parasite et les courants de fuite associés au type d'isolation utilisé.

3.1.1. Résistances en silicium monocristallin

Les résistances en silicium monocristallin intégrées sont normalement formées par le dopage plus ou moins fort d'une région du substrat, qui était déjà dopé avec des impuretés du type opposé (voir Figure 34).

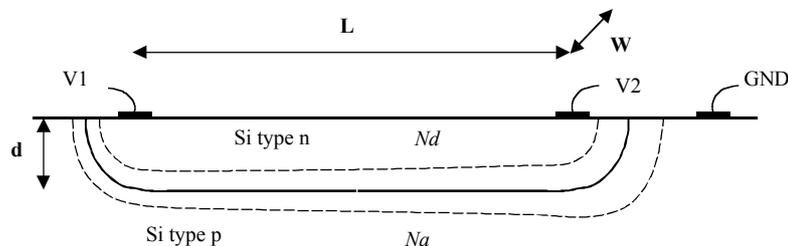


Figure 34. Résistance en silicium monocristallin intégrée ($V2 > V1 > \text{GND}$).

Les dimensions de la résistance W , L et d sont définies, par la jonction pn ainsi formée, en surface par le masquage pendant l'implantation et la diffusion des impuretés pendant les traitements thermiques, et en épaisseur par la profondeur de pénétration des impuretés implantées. Les dimensions W et d effectives de ces résistances dépendent aussi de la largeur $X_{n,p}$ de la partie de la ZCE de la jonction pn correspondante à la résistance implantée. Comme $X_{n,p}$ est une fonction de la polarisation et du dopage (voir x_n dans l'équation (74)), il en résulte que les résistances ainsi générées se comportent d'une manière très non-linéaire, quand la valeur de $X_{n,p}$ est comparable à la profondeur de l'implantation d ou à la largeur W de la résistance implantée. La valeur de la résistance est donnée par,

$$(81) \quad R(T) = \rho_{n,p}(T) \frac{L}{[W - 2 \cdot X_{n,p}(V1, V2)] \cdot [d - X_{n,p}(V1, V2)]} = \frac{\rho_{n,p} \text{carré}(T, V1, V2) \cdot L}{[W - 2 \cdot X_{n,p}(V1, V2)]}$$

L'isolation de ces résistances par polarisation inverse de la jonction pn , leur donne une importante capacité parasite associée et est la cause de leurs non-négligeables courants de fuite (voir Tableau 5 et

Tableau 6). Ces courants de fuite, comme il a été déjà mentionné dans la section 2.3.2.5, augmentent très rapidement avec la température, rendant non recommandable l'utilisation de ce type de résistances dans les applications en haute température.

Leur résistivité correspond à la résistivité du silicium extrinsèque en régime d'épuisement (région (b) dans la Figure 15) où la concentration d'électrons ou de trous est pratiquement égale, respectivement, à la concentration des donneurs Nd ou des accepteurs Na . A partir de l'équation (28), on obtient les expressions suivantes pour la résistivité,

$$(82) \quad \rho_n(T) = \frac{1}{q Nd \mu_n(T)}$$

$$(83) \quad \rho_p(T) = \frac{1}{q Na \mu_p(T)}$$

La résistivité augmente sensiblement selon la température principalement à cause de la réduction de la mobilité (voir section 2.1.7.1 et voir Tableau 5 et Tableau 6).

Implantation	Résistivité [Ohms/carré]	Coefficient Thermique α [$10^{-3}/K$]	Courant de fuite par unité de surface (-5v, Tamb) [$fA/\mu m^2$]	Courant de fuite des parois (-5v, Tamb) [$fA/\mu m$]	Capacité par unité de surface (0v, Tamb) [$fF/\mu m^2$]	Capacité des parois (0v, Tamb) [$fF/\mu m$]	Tension de claquage [V]
Caisson n-	1000 à 1400	~ +6.5	0.02	0.02	0.07	0.07	45
Diffusion n+	20 à 35	~ +1.8	0.01	0.12	0.29	0.23	17
Diffusion p+	25 à 60	~ +1.7	0.04	0.5	0.49	0.21	15

Tableau 5. Caractéristiques des résistances en Si monocristallin de la technologie CMOS étudiée.

La concentration des dopants Na ou Nd , affecte la valeur de résistivité, l'ordre de grandeur des courants de fuite (voir section 2.3.2.5), et la taille $X_{n,p}$ de la ZCE à la jonction (i.e. le comportement non-linéaire de la résistance) des résistances intégrées en silicium monocristallin.

Implantation	Résistivité [Ohms/carré]	Coefficient Thermique α [$10^{-3}/K$]	Courant de fuite par unité de surface (-5v, Tamb) [$fA/\mu m^2$]	Courant de fuite des parois (-5v, Tamb) [$fA/\mu m$]	Capacité par unité de surface (0v, Tamb) [$fF/\mu m^2$]	Capacité des parois (0v, Tamb) [$fF/\mu m$]	Tension de claquage [V]
Caisson n-	3000 à 4500	~ +6.6	0.02	0.06	0.14	0.01	55
Diffusion n+	19 à 28	~ +2.1	0.01	0.05	0.36	0.25	18
Couche enterré n+	~ 24	~ +1.2	0.05	0.10	0.095	0.95	33
Diffusion p+	35 à 51	~ +1.8	0.02	0.06	0.44	0.22	13
Base p-	3800 à 5000	~ +3.2	0.03	0.07	0.44 à 0.50	0.32	18

Tableau 6. Caractéristiques des résistances en Si monocristallin de la technologie BiCMOS étudiée

3.1.2. Résistances en silicium polycristallin

Comme il a été déjà mentionné dans la section 2.2.2, la résistivité du polysilicium peut être contrôlée en modifiant le dopage du film (voir Figure 21). Pour de très faibles concentrations de dopants, la résistivité est très haute et décroît exponentiellement avec la température. A plus fortes concentrations de dopants, le régime de conduction en volume devient dominant et fournit une résistivité plus faible, laquelle augmente légèrement avec la température. Ce dernier cas est celui des couches de silicium polycristallin normalement utilisées dans les technologies CMOS et BiCMOS (voir Tableau 7 et Tableau 8) pour la formation de résistances en polysilicium. Ces résistances sont modélisées, autour d'une température de référence T_0 , avec la relation suivante,

$$(84) \quad R(T) = \sigma_{carré}(T_0) \frac{L}{W} \left(1 + \alpha (T - T_0) + \beta (T - T_0)^2 + \gamma (T - T_0)^3 + \dots \right)$$

Où $\sigma_{carré}$ correspond à la résistance d'un carré de polysilicium dont l'épaisseur est prédéfinie, L est la longueur et W la largeur correspondant à une résistance de forme rectangulaire.

Couche de Poly-Si	Résistivité [Ohms/carré]	Coefficient Thermique α [$10^{-3}/K$]	Epaisseur [μm]	Erreur de Largeur ΔW [μm]	Capacité de la surface [fF/ μm^2]	Capacité du périmètre [fF/ μm]
Poly-Si_1	20 à 28	~ +0.9	0.39 à 0.45	-0.1 à 0.12	0.061 à 0.072	0.047 à 0.050
Poly-Si_2	21 à 33	~ +0.8	0.25 à 0.29	-0.4 à 0.2	0.061 à 0.072	0.047 à 0.050

Tableau 7. Caractéristiques des couches de polysilicium de la technologie CMOS étudiée.

La reproductibilité de la valeur des résistances sur Poly-Si, en fonction de la température, est normalement très bonne à des températures inférieures à 250°C. Au-delà de cette température des variations de la valeur de la résistance peuvent être observées après quelques cycles thermiques. Comme la linéarité des résistances sur Poly-Si fortement dopé est normalement assez bonne, dans la plupart des applications, les coefficients thermiques α (voir Tableau 7 et Tableau 8) et β permettent de suffisamment bien décrire l'évolution de leur résistance selon la température.

Couche de Poly-Si	Résistivité [Ohms/carré]	Coefficient Thermique α [$10^{-3}/K$]	Epaisseur [μm]	Erreur de Largeur ΔW [μm]	Capacité de la surface [fF/ μm^2]	Capacité du périmètre [fF/ μm]
Poly-Si_1 (région n+)	18 à 26	~ +1.0	0.38 à 0.44	-0.14 à 0.1	0.059 à 0.068	0.046 à 0.049
Poly-Si_1 (région p+)	~ 40	~ +0.7	0.38 à 0.44	-	0.059 à 0.068	0.046 à 0.049
Poly-Si_2	55 à 80	~ -0.3	0.30 à 0.36	-0.25 à 0.1	0.059 à 0.068	0.046 à 0.049
HR-Poly-Si	900 à 1500	~ -2.5	-	-0.25 à 0.15	0.059 à 0.068	0.046 à 0.049

Tableau 8. Caractéristiques des couches de polysilicium de la technologie BiCMOS étudiée.

Une propriété très importante des résistances en Poly-Si, dans la plupart des technologies CMOS et BiCMOS, est leur très bonne isolation. Les couches de Poly-Si sont normalement isolées avec des diélectriques épais, ce qui rend les courants de fuite en haute température pratiquement négligeables. En contrepartie, la capacité parasite associée aux diélectriques isolants limite la valeur maximale des résistances réalisables dans certaines applications (la capacité entre les couches Poly-Si sur l'oxyde de champ est donnée dans le Tableau 7 et le Tableau 8, étant approximativement un ordre de grandeur inférieure à celle des jonctions $p+n+$).

Pour les applications qui nécessitent des valeurs de résistance très élevées avec une faible capacité parasite et des courants de fuite négligeables, une couche supplémentaire de Poly-Si faiblement dopée (i.e. de plus haute résistivité) est souvent disponible dans quelques technologies. La résistivité diminuant de façon plutôt exponentielle avec la température, des coefficients thermiques d'ordres supérieurs (i.e. β , γ , etc., dans l'équation (84)) sont nécessaires pour bien décrire sa valeur suivant la température.

3.2. Capacités

Les capacités sont les dispositifs passifs les plus importants en ce qui concerne la réponse temporelle des circuits intégrés en général, et dans le cas des circuits intégrés analogiques en particulier. Les propriétés les plus importantes de ces dispositifs sont la linéarité (i.e. l'indépendance entre la valeur de la capacité et la tension appliquée), la stabilité thermique (i.e. ses coefficients thermiques et la reproductibilité de ses caractéristiques après un nombre de cycles thermiques) la résistance parasite en série et les courants de fuite associés au type de diélectrique utilisé.

3.2.1. Capacités de jonction pn

Les jonctions *pn* polarisées en inverse ont un comportement capacitif non-linéaire, comme cela a déjà été démontré dans la section 2.3.2.7, propriété qui est mise à profit dans les varactors. Même si la valeur de la capacité a une stabilité en température relativement acceptable, ses courants de fuite non-négligeables (voir la section 2.3.2.5), augmentent très rapidement avec la température et ne rendent pas recommandable l'utilisation de ce type de capacités dans les application en haute température.

3.2.2. Capacités MOS

Une capacité MOS est formée par la superposition d'une couche conductrice (Métallique, historiquement), d'une couche de matériau diélectrique (Oxyde de silicium) et d'un matériau Semiconducteur. Dans les technologies des circuits intégrés sur silicium, les capacités MOS sont souvent réalisées avec : polysilicium dégénéré de type *p* ou *n*, comme couche conductrice; oxyde de silicium, comme couche diélectrique; et silicium monocristallin dopé *n* ou *p*, comme matériau semiconducteur.

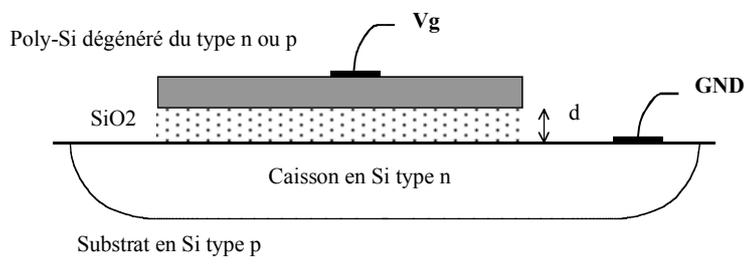


Figure 35. Structure d'une capacité MOS.

Les Figure 35 et Figure 36 montrent respectivement la structure et le diagramme de bandes correspondant à une capacité MOS polarisée avec un potentiel V_g .

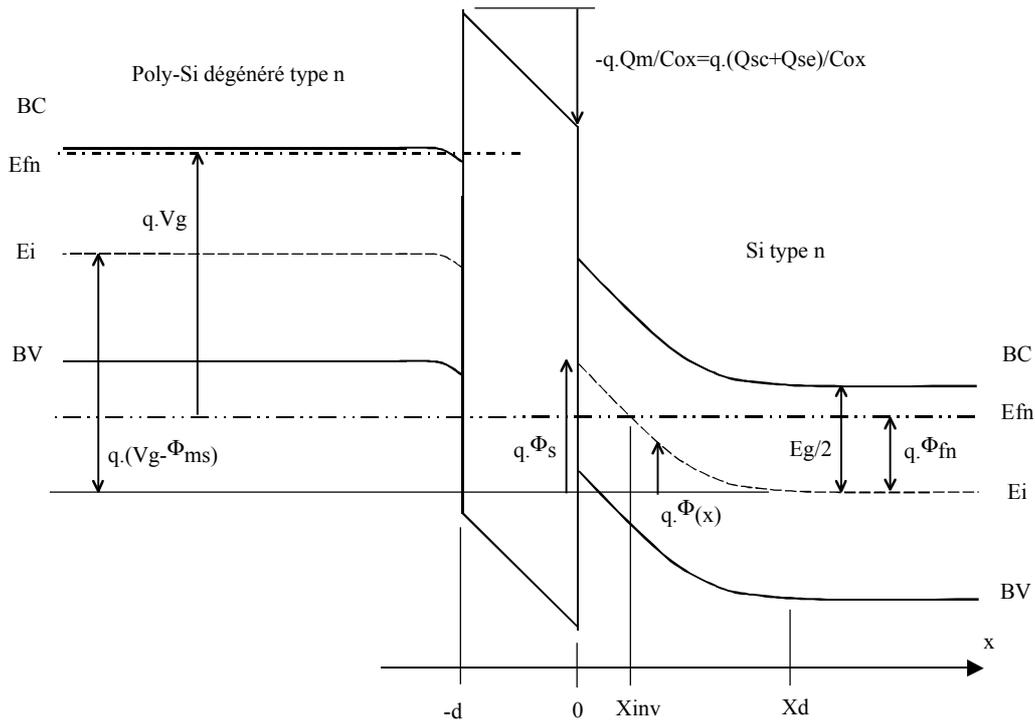


Figure 36. Diagramme des bandes d'énergie d'une capacité MOS polarisée en faible inversion.

$$(85) \quad Vg = \Phi_{ms} + \Phi_s - \frac{Q_{sc}}{C_{ox}} - \frac{Q_{se}}{C_{ox}} = V_{fb} + \Phi_s - \frac{Q_{sc}}{C_{ox}}$$

Où Φ_{ms} est le travail de sortie entre le métal (Poly-Si dégénéré) et le semiconducteur, Φ_s est le potentiel à l'interface $Si-SiO_2$, $C_{ox} = \epsilon_{ox}/d$ est la capacité de la couche d'oxyde de silicium par unité de surface, Q_{sc} est la charge dans le semiconducteur par unité de surface, Q_{se} est la charge à l'interface $Si-SiO_2$ par unité de surface et V_{fb} le potentiel de bandes plates, défini par la valeur de Vg correspondante à Φ_s et Q_{sc} nuls,

$$(86) \quad V_{fbn,p} = \frac{Eg}{2} - \Phi_{fn,p} - \frac{Q_{se}}{C_{ox}} \quad \text{pour le Poly-Si dégénéré de type } p.$$

$$(87) \quad V_{fbn,p} = -\frac{Eg}{2} - \Phi_{fn,p} - \frac{Q_{se}}{C_{ox}} \quad \text{pour le Poly-Si dégénéré de type } n.$$

Où $\Phi_{fn,p}$ est le potentiel de Fermi correspondant au matériau semiconducteur non-dégénéré de type n et p , respectivement,

$$(88) \quad \Phi_{fp} = Vt \ln \frac{Na}{ni} > 0 \quad \text{pour le Si monocristallin non-dégénéré de type } p.$$

$$(89) \quad \Phi_{fn} = -Vt \ln \frac{Nd}{ni} < 0 \quad \text{pour le Si monocristallin non-dégénéré de type } n.$$

Où $Vt = k.T/q$. A partir des équations (10) et (11) les concentrations des porteurs, dans la Figure 36, s'écrivent,

$$(90) \quad n(x) = ni \cdot e^{\frac{\Phi(x) - \Phi_{fn}}{Vt}} = n_{n0} \cdot e^{\frac{\Phi(x)}{Vt}}$$

$$(91) \quad p(x) = n_i \cdot e^{\frac{-(\Phi(x)-\Phi_{fn})}{V_t}} = p_{n0} \cdot e^{\frac{-\Phi(x)}{V_t}}$$

La concentration totale de charge dans le semiconducteur, par unité de volume, dans la direction x , est donnée par,

$$(92) \quad \rho(x) = q(N_d - N_a + p(x) - n(x)) = q \left(p_{n0} \cdot e^{\frac{-\Phi(x)}{V_t}} - 1 \right) - n_{n0} \cdot e^{\frac{\Phi(x)}{V_t}} - 1$$

Si on intègre l'équation de Poisson, à partir de l'expression (92), depuis la région neutre du semiconducteur vers la région de charge d'espace, on obtient la densité de charge Q_{sc} par unité de surface dans le semiconducteur [1] [2]:

$$(93) \quad Q_{sc} = -\frac{\Phi_s}{|\Phi_s|} \cdot \frac{\epsilon_{sc}}{L_d} \cdot V_t \cdot \sqrt{\frac{p_{n0}}{n_{n0}} \cdot e^{\frac{-\Phi_s}{V_t}} - 1} + e^{\frac{-\Phi_s - 2\Phi_{fn}}{V_t}} + e^{\frac{\Phi_s}{V_t}} - \frac{\Phi_s}{V_t} - 1$$

Où L_d est la longueur de Debye des électrons,

$$(94) \quad L_d = \sqrt{\frac{\epsilon_{sc} \cdot V_t}{2 \cdot q \cdot n_{n0}}}$$

La Figure 37 montre Q_{sc} en fonction de Φ_s , d'après les équations (85) et (93). Quatre différents régimes sont identifiés, selon la valeur de Φ_s : forte inversion ($\Phi_s < 2\Phi_{fn}$), faible inversion ($2\Phi_{fn} < \Phi_s < \Phi_{fn}$), déplétion ($\Phi_{fn} < \Phi_s < 0$) et accumulation ($0 < \Phi_s$).

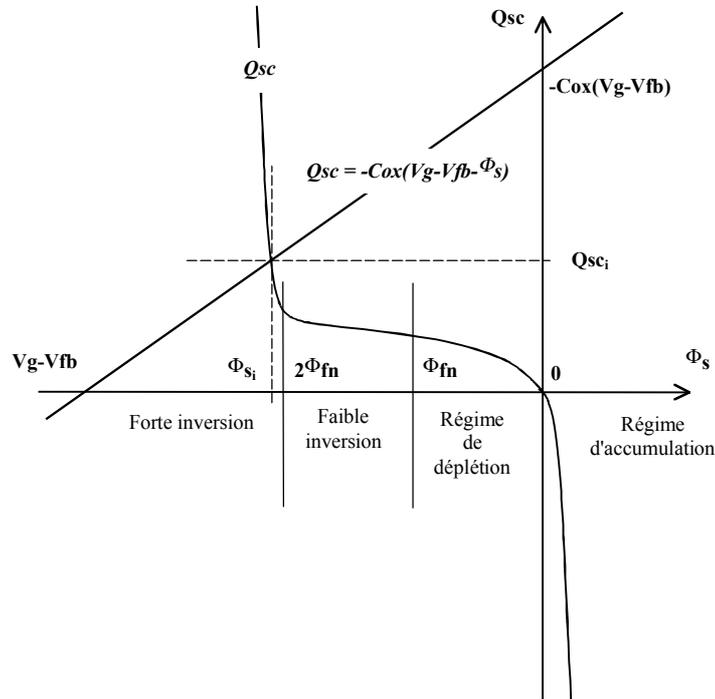


Figure 37. Densité de charges par unité de surface Q_{sc} dans le semiconducteur de type n .

La capacité de la structure de la Figure 35 est $C = dQ_g/dV_g = -dQ_{sc}/dV_g$. Prenant en compte le fait que $C_{ox} = dQ_g/d(V_g - \Phi_s)$ et $C_{sc} = -dQ_{sc}/d\Phi_s$, la capacité MOS devient,

$$(95) \quad \frac{1}{C} = \frac{1}{C_{ox}} + \frac{1}{C_{sc}}$$

On obtient la capacité C_{sc} associée au semiconducteur en dérivant l'expression (93) par rapport à Φ_s , ce qui donne une expression relativement complexe, car la capacité de la structure MOS se comporte très différemment selon le régime de fonctionnement. Dans le régime de forte inversion, la capacité MOS résultante dépend de la fréquence, et dans les régimes de déplétion et de faible inversion la capacité MOS dépend fortement du potentiel Φ_s à l'interface $Si-SiO_2$ [1][2]. Les non-linéarités de ces capacités les rendent peu intéressantes pour les applications classiques. Seul, le régime d'accumulation génère des capacités suffisamment linéaires et des valeurs de capacité par unité de surface relativement importantes, dans la plupart des technologies CMOS et BiCMOS.

3.2.2.1. Capacité MOS en accumulation

Le régime d'accumulation de la capacité MOS de la Figure 35 correspond à une valeur positive de Φ_s , produite par une polarisation tel que $V_g - V_{fb} > 0$ (voir Figure 37). Il en résulte que, lorsque Φ_s devient de l'ordre de quelques V_t , l'équation (93) devient,

$$(96) \quad Q_{sc} = -\frac{\epsilon_{sc}}{Ld} \eta V_t \eta e^{\sqrt{\frac{\Phi_s}{V_t}}}$$

La capacité dynamique C_{sc} associée au semiconducteur résulte donc de la modulation de la charge d'accumulation et est donnée par le module de la dérivée de l'expression précédente,

$$(97) \quad C_{sc} = \frac{\epsilon_{sc}}{2 \eta Ld} \eta e^{\frac{\Phi_s}{2V_t}}$$

Dans la Figure 37 on peut voir que la valeur Φ_s du potentiel à l'interface $Si-SiO_2$ dépend de la tension de polarisation V_g . A partir des équations (85) et (96) on écrit,

$$(98) \quad -C_{ox} \eta (V_g - V_{fb} - \Phi_s) = -\frac{\epsilon_{sc}}{Ld} \eta V_t \eta e^{\frac{\Phi_s}{2V_t}}$$

En regroupant les termes en Φ_s ,

$$(99) \quad C_{ox} \eta (V_g - V_{fb}) = \frac{\epsilon_{sc}}{Ld} \eta V_t \eta e^{\frac{\Phi_s}{2V_t}} \eta \left[1 + \frac{C_{ox}}{\epsilon_{sc}/Ld} \eta \frac{\Phi_s}{V_t} \eta e^{-\frac{\Phi_s}{2V_t}} \right] \sqrt{\frac{\epsilon_{sc}}{Ld} \eta V_t \eta e^{\frac{\Phi_s}{2V_t}}}$$

Compte tenu du fait que normalement $C_{ox} \ll \epsilon_{sc}/Ld$ [2] et qu'en régime d'accumulation Φ_s est de l'ordre de quelques V_t . On obtient ainsi l'expression suivante pour le potentiel Φ_s à l'interface, en régime d'accumulation,

$$(100) \quad \Phi_s = 2 \eta V_t \eta \ln \left[\frac{C_{ox}}{\epsilon_{sc}/Ld} \eta \frac{(V_g - V_{fb}) \eta}{V_t} \eta \sqrt{\frac{\epsilon_{sc}}{Ld} \eta V_t \eta} \right]$$

En explicitant Φ_s dans l'équation (97), la capacité dynamique C_{sc} associée au semiconducteur s'écrit,

$$(101) \quad C_{sc} = C_{ox} \sqrt{\frac{V_g - V_{fb}}{2 \eta t}}$$

En remplaçant C_{sc} dans l'équation (95), on obtient l'expression suivante de la capacité MOS en régime d'accumulation C_{acc} ,

$$(102) \quad C_{acc} = \frac{C_{ox}}{1 + \frac{2 \eta t}{V_g - V_{fb}}}$$

Dans les technologies CMOS et BiCMOS les structures MOS sont construites de façon à être polarisées en faible et forte inversion et ainsi servir à la réalisation des transistors NMOS et PMOS. C'est seulement quand on dispose d'un caisson semiconducteur isolé du substrat qu'on peut modifier la polarisation du matériau semiconducteur et ainsi utiliser le régime d'accumulation d'une structure MOS, sans altérer le fonctionnement du reste du circuit intégré.

Les technologies CMOS et BiCMOS étudiées ont un substrat de départ du type p , sur lequel les structures NMOS sont construites, et dans lequel des caissons du type n sont formés afin de permettre la réalisation des structures PMOS. En polarisant une structure du type PMOS comme décrit dans la Figure 35 on peut obtenir des capacités MOS en accumulation, lesquelles, dans le cas des technologies étudiées, ont la valeur la plus importante par unité de surface (voir le Tableau 9). Dans la Figure 38 on peut apprécier les variations de la valeur de cette capacité en fonction de la tension de polarisation V_g et de la température, calculée à partir de l'équation (102).

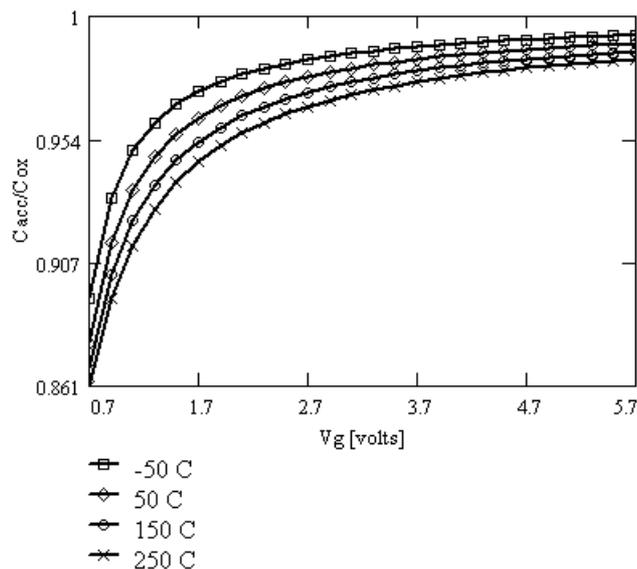


Figure 38. Capacité MOS en accumulation en fonction de la polarisation et de la température.

Technologie des circuits intégrés	Capacité C_{ox} [fF/ μm^2]	Épaisseur d'oxyde [nm]	Tension de claquage [V]
CMOS	2.03 à 2.30	15 à 17	12 à 14
BiCMOS	2.03 à 2.30	15 à 17	12 à 18

Tableau 9. Caractéristiques des capacités MOS des technologies étudiées.

Les courants de fuite des capacités MOS en accumulation sont pratiquement négligeables, même à haute température, dans le nœud du côté du Poly-Si, du fait de l'isolation par diélectrique. Par contre,

les courants de fuite peuvent être très importants dans le caisson (voir section 2.3.2.5), selon sa polarisation par rapport au substrat. Il en résulte que les applications en haute température pour les capacités MOS se limitent à celles où le caisson peut être connecté en permanence à un nœud de très basse impédance, comme par exemple des alimentations *GND*, *VDD*, etc.

Les non-linearités des capacités MOS en accumulation peuvent être mises à profit en les utilisant comme des *varactors* de capacité maximale C_{ox} , si on les polarise avec une tension V_g supérieure à V_{fb} de quelque dizaines de V_t . Pour d'autres types d'applications, des capacités MOS fortement polarisées peuvent être aussi utilisées. Dans les deux cas, l'impact des non-linéarités et des variations selon la température, des capacités MOS en accumulation, doit être soigneusement étudié (voir un exemple dans la section 5.2.1.4.f.c).

3.2.3. Capacités faites en Double Poly-Si

Les capacités Double-Poly sont composées de deux couches superposées de polysilicium séparées par une couche d'oxyde de silicium de haute qualité et d'épaisseur contrôlée. Dans les technologies étudiées, les couches de Poly-Si sont isolées du reste du circuit intégré par le diélectrique qui sert à l'isolation des couches d'interconnexion, et par l'oxyde de champ. Par conséquent, la non-linéarité et les courants de fuite de ces capacités sont pratiquement négligeables, même en haute température. Cependant il faut prendre en compte l'existence d'une capacité parasite non négligeable entre la couche inférieure de polysilicium et le substrat, du fait de la présence d'un diélectrique : l'oxyde de champ. L'influence de cette capacité parasite, dans le comportement d'un circuit, peut être éliminée en connectant la couche inférieure du Poly-Si, de la capacité Double-Poly, à un nœud très capacitif du circuit, tel *VDD*, *GND*, etc. Les caractéristiques des capacités faites en Double Polysilicium sont résumées dans le Tableau 10.

Technologie	Capacité Double-Poly par unité de surface [fF/ μm^2]	Capacité Double-Poly par unité de périmètre [fF/ μm]	Capacité entre le Poly-Si et le substrat par unité de surface [fF/ μm^2]	Tension de Claquage [volts]
CMOS	1.65 à 1.92	0.19 à 0.21	0.061 à 0.072	20 à 33
BiCMOS	1.65 à 1.92	0.19 à 0.22	0.059 à 0.068	14 à 20

Tableau 10. Caractéristiques des capacités Double Poly-Si des technologies CMOS et BiCMOS étudiées

La performance des capacités Double-Poly, dans la plage de température considérée, est pratiquement idéale.

3.3. Diodes

Les caractéristiques des diodes, en ce qui concerne leurs courbes caractéristiques de conduction (voir l'équation (70)), leurs courants de fuite (voir l'équation (73)), leurs capacités parasites (voir l'équation (78)), et leurs tensions de claquage (voir la Figure 30 et la Figure 31), sont décrites dans la section 2.3.2. Les caractéristiques des jonctions implémentables dans les technologies CMOS et BiCMOS étudiées sont résumées respectivement dans le Tableau 5 et le Tableau 6. Dans les deux technologies mentionnées, on a aussi la possibilité d'implémenter des diodes Zener (voir le Tableau 11), mettant à profit les propriétés du claquage du type Zener des jonctions $p+n+$ polarisées inversement.

Diode	Tension de Claquage [volts]	Courant de Fuite [nA] (-1 volt, Tamb)	Pente du courant de Fuite [log(Amp)/volt]
Diode Zener $p+n+$	3.4 à 5.0	< 300	1.1

Tableau 11. Caractéristiques de la Diode Zener des technologies CMOS et BiCMOS étudiées

3.4. Transistors MOS

Les structures de base des transistors MOS des technologies étudiées, sont représentées sur la Figure 39. Ces structures sont constituées essentiellement d'un substrat du type p , pour les transistors NMOS, et d'un caisson du type n , pour les transistors PMOS, dans lequel deux diffusions du type contraire, respectivement $n+$ et $p+$, constituent les électrodes de la source S et le drain D . Une capacité MOS est réalisée sur le substrat et sur le caisson, entre la source et le drain. L'électrode de commande de la capacité MOS constitue la grille G du transistor. Le semiconducteur sous la grille peut être dans les états d'accumulation, déplétion, faible inversion ou forte inversion, selon la polarisation de la grille. Le substrat $p-$ est connecté à la terre et le caisson $n-$ est normalement polarisé avec $V_{sb}=5\text{volts}$. Dans les conditions normales d'opération, les tensions de polarisation de la grille V_g , de la source V_s et du drain V_d , sont positives dans le cas des NMOS et négatives dans le cas des PMOS.

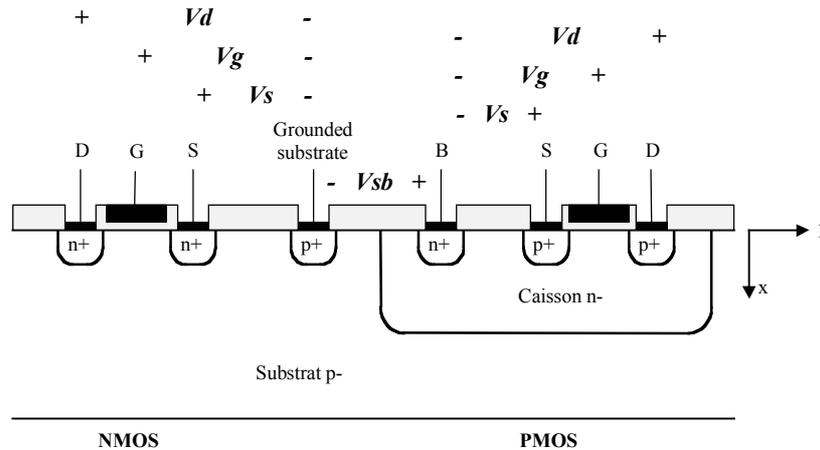


Figure 39. Structure des transistors MOS dans les technologies étudiées.

Quand la capacité MOS se trouve dans le régime d'inversion faible ou forte, la couche d'inversion créée à la surface du semiconducteur, dite canal d'inversion, génère une conductivité entre le drain et la source, laquelle est modulée par la tension V_g de polarisation de la grille. Le passage d'un courant de porteurs minoritaires entre le drain et la source modifie le diagramme des bandes d'énergie de la capacité MOS (voir la Figure 40) mentionné dans la section 3.2.2. Les quasi-niveaux de Fermi dans le semiconducteur E_{fn} et E_{fp} se séparent dans la région inversée suivant la valeur de la tension $V_c(y)$ du canal ($|V_s| < |V_c(y)| < |V_d|$). Le quasi-niveau de Fermi E_{fp} , correspondant aux porteurs minoritaires, varie selon la direction y . Dans la Figure 40, on observe le diagramme des bandes d'énergie d'un transistor PMOS en régime d'inversion, où on peut remarquer l'élargissement de la zone désertée X_d sous la grille, par rapport à celle de la capacité MOS, dû à la polarisation $V_c(y)$ du canal.

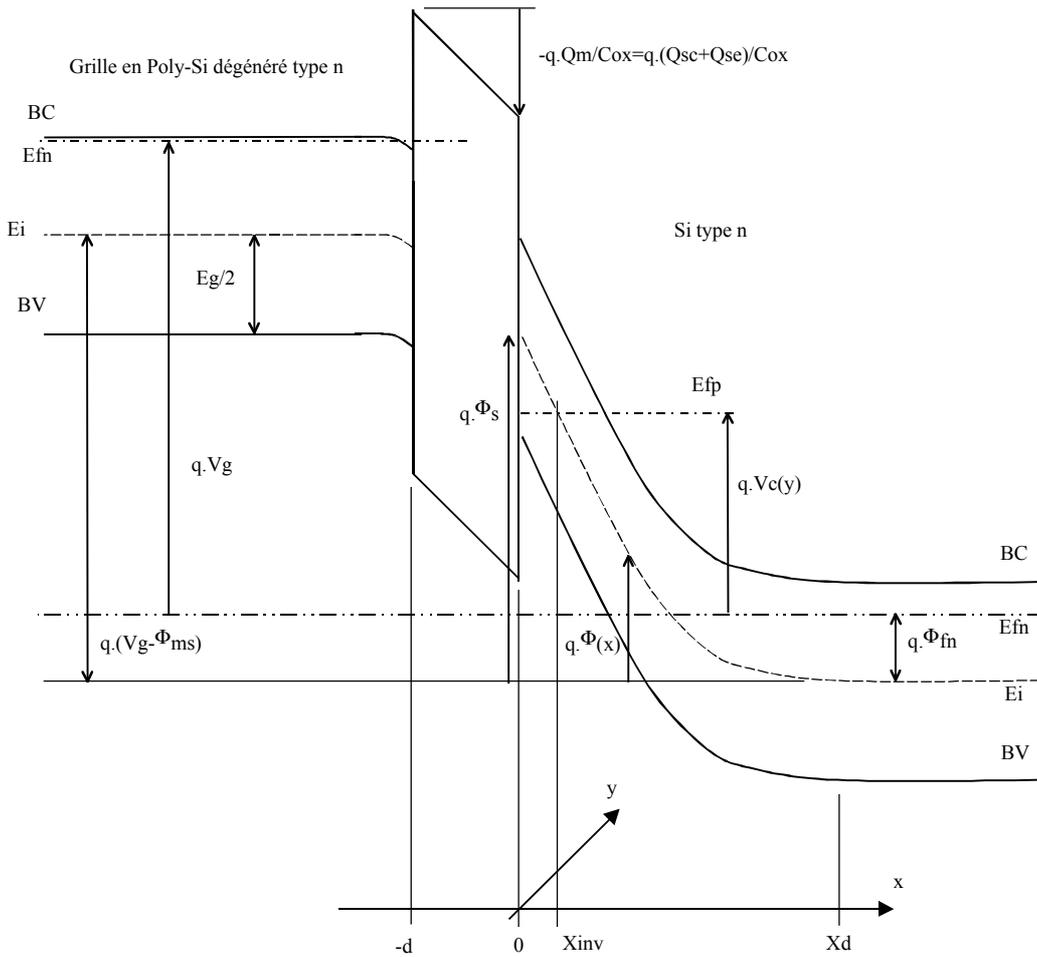


Figure 40. Diagramme des bandes d'énergie d'un transistor PMOS polarisé en inversion.

Les équations (85), (86), (87), (88) et (89) des capacités MOS restent toujours valables pour les transistors MOS. A partir des équations (10) et (11) les concentrations des porteurs, dans la Figure 40, s'écrivent,

$$(103) \quad n(x) = n_i \cdot e^{\frac{\Phi(x) - \Phi_{fn}}{V_t}} = n_{n0} \cdot e^{\frac{\Phi(x)}{V_t}}$$

$$(104) \quad p(x) = n_i \cdot e^{\frac{-(\Phi(x) - V_c(y) - \Phi_{fn})}{V_t}} = p_{n0} \cdot e^{\frac{-[\Phi(x) - V_c(y)]}{V_t}}$$

La concentration totale de charge dans le semiconducteur, par unité de volume, dans la direction x , est donc donnée par,

$$(105) \quad \rho(x) = q \cdot (Nd - Na + p(x) - n(x)) = q \cdot \left[p_{n0} \cdot e^{\frac{-[\Phi(x) - V_c(y)]}{V_t}} - 1 \right] - n_{n0} \cdot e^{\frac{\Phi(x)}{V_t}} - 1$$

Si on intègre l'équation de Poisson, à partir de l'expression (105), depuis la région neutre du semiconducteur vers la région de charge d'espace, on obtient la densité de charge Q_{sc} suivante, par unité de surface dans le semiconducteur [1] [2],

$$(106) \quad Q_{sc} = -\frac{\Phi_s}{|\Phi_s|} \frac{\epsilon_{sc}}{Ld} \sqrt{\frac{p_{n0}}{n_{n0}}} \frac{\Phi_s}{Vt} - e^{\frac{Vc(y)}{Vt}} \sqrt{+e^{-\frac{-(\Phi_s - Vc(y) - 2\Phi_{fn})}{Vt}} + e^{\frac{\Phi_s}{Vt}} - \frac{\Phi_s}{Vt} - 1}$$

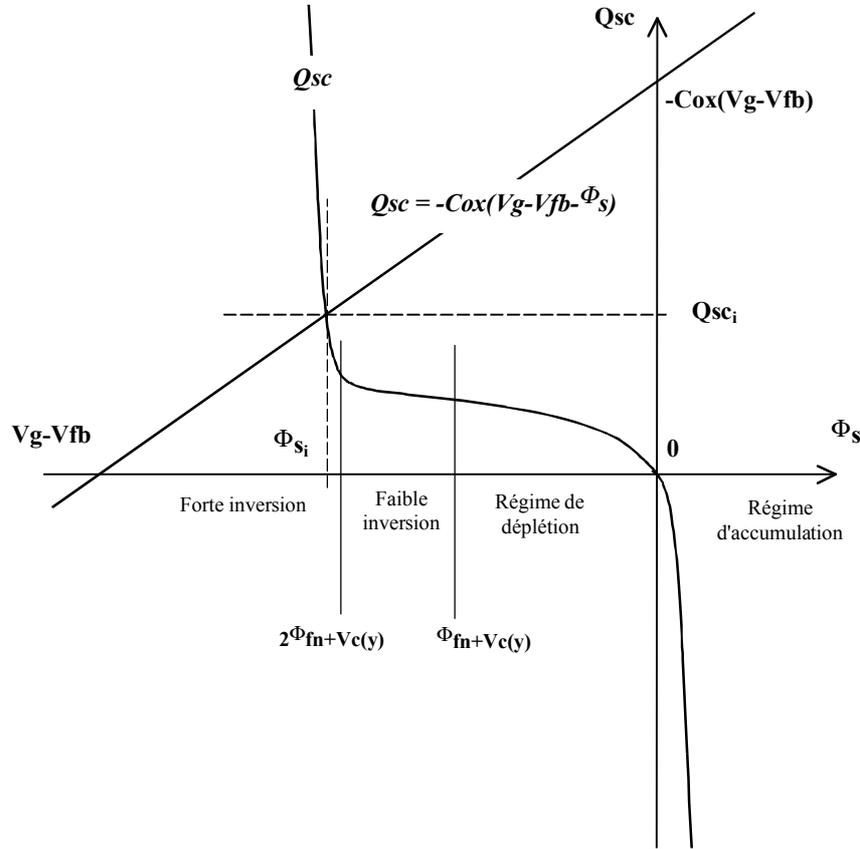


Figure 41. Densité de charges par unité de surface Q_{sc} dans le caisson n- du transistor PMOS.

La Figure 41 montre Q_{sc} en fonction de Φ_s , d'après les équations (85) et (106). Quatre différents régimes sont identifiés, selon la valeur de Φ_s : forte inversion ($\Phi_s < 2\Phi_{fn} + Vc(y)$), faible inversion ($2\Phi_{fn} + Vc(y) < \Phi_s < \Phi_{fn} + Vc(y)$), déplétion ($\Phi_{fn} + Vc(y) < \Phi_s < 0$) et accumulation ($0 < \Phi_s$).

3.4.1. Tension de seuil

A partir de la Figure 41 et des équations (85) et (106), on peut déterminer la valeur de la tension dans la grille Vg correspondant au début du régime d'inversion forte, dite *tension de seuil* V_{thP} et V_{thN} , respectivement pour les transistors PMOS et NMOS,

$$(107) \quad V_{thP} = Vg|_{PMOS:inv.forte} = V_{fbn} + 2\Phi_{fn} + Vc(y) - \gamma \sqrt{|2\Phi_{fn} + Vc(y)|} < 0$$

$$(108) \quad V_{thN} = Vg|_{NMOS:inv.forte} = V_{fbp} + 2\Phi_{fp} + Vc(y) + \gamma \sqrt{|2\Phi_{fp} + Vc(y)|} > 0$$

$$(109) \quad \gamma_p = \frac{\sqrt{2 \cdot \epsilon_{sc} \cdot q \cdot Nd}}{C_{ox}}$$

$$(110) \quad \gamma_n = \frac{\sqrt{2 \cdot \epsilon_{sc} \cdot q \cdot Na}}{C_{ox}}$$

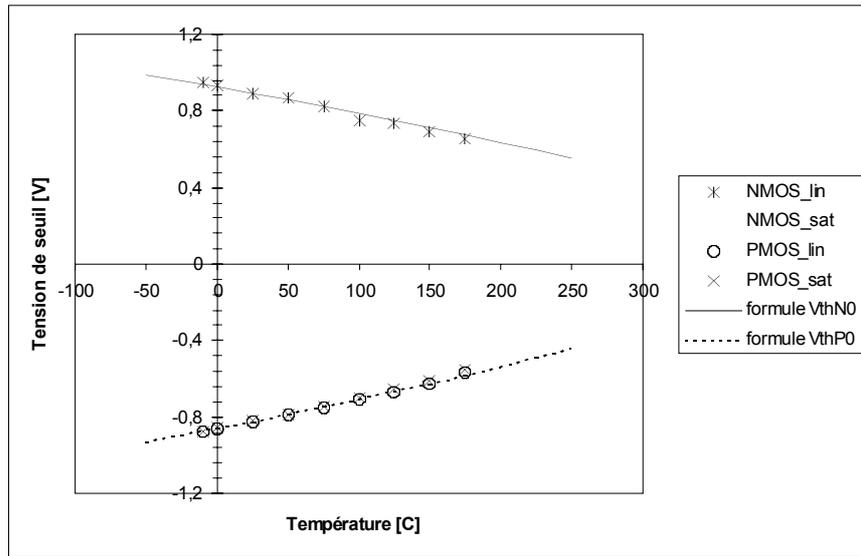


Figure 42. Tension de seuil V_{th0} des transistors NMOS et PMOS avec $W/L=20/20 \mu\text{m}$ des technologies étudiées en fonction de la température.

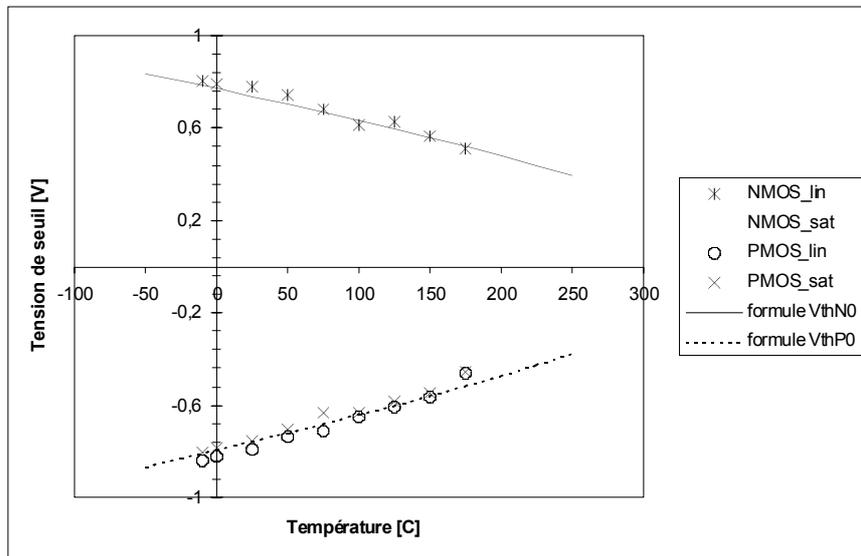


Figure 43. Tension de seuil V_{th0} des transistors NMOS et PMOS avec $W/L=20/0.8 \mu\text{m}$ des technologies étudiées en fonction de la température.

Pour $V_c(\text{source})=V_s$, et en considérant le cas d'une grille construite avec du silicium polycristallin dégénéré du type n (voir la Figure 40), les tensions des seuils s'écrivent, à partir de l'expression (87) de la manière suivante,

$$(111) \quad V_{thP}(T) = -\frac{E_g(T)}{2q} - \frac{Q_{seP}}{C_{ox}} + \Phi_{fn}(T) + V_s - \gamma_p \sqrt{2\Phi_{fn}(T) + V_s} < 0$$

$$(112) \quad V_{thN}(T) = -\frac{E_g(T)}{2q} - \frac{Q_{seN}}{C_{ox}} + \Phi_{fp}(T) + V_s + \gamma_n \sqrt{2\Phi_{fp}(T) + V_s} > 0$$

Les valeurs des densités de charges par unité de surface à l'interface $Si-SiO_2$ des transistors PMOS et NMOS, Q_{seP} et Q_{seN} respectivement, ont été estimées à partir des équations (111) et (112) prenant en compte que $V_{th0P} = V_{thP}(V_s=0) = -0.818$ volts et $V_{th0N} = V_{thN}(V_s=0) = 0.892$ volts à température ambiante (mesurés à partir des transistors NMOS et PMOS de $W/L=20/20$ [μm]), et que $V_{th0P} = -0.755$ volts et $V_{th0N} = 0.736$ volts à température ambiante (mesurés à partir des transistors NMOS et PMOS de $W/L=20/0.8$ [μm]).

Dans la Figure 42 et la Figure 43, on peut observer l'évolution des valeurs de V_{thP} et V_{thN} en fonction de la température, mesurées (en régime linéaire et saturé, voir la section 3.4.2) et calculées à partir des équations précédentes, pour une polarisation de la source $V_s=0$. Les variations selon la température de $|V_{thN}|$ et $|V_{thP}|$ sont légèrement différentes, d'une part à cause des variations de $E_g(T)$ qui s'additionnent de la même manière aux deux tensions de seuil de signes opposés, et d'autre part à cause des différences entre les variations de $\Phi_{fp}(T)$ et $\Phi_{fn}(T)$, et entre les coefficients γ_p et γ_n .

La réduction de la tension de seuil des transistors MOS selon la température, doit être prise en compte pour la conception des circuits intégrés en haute température ($<300^\circ C$), mais ne représente pas une forte limitation pour l'utilisation des transistors MOS en haute température.

3.4.2. Courant de conduction en inversion forte

Dans le cas des transistors longs et larges, le courant drain-source I_{ds} est essentiellement donné par la conductivité de la couche d'inversion dans le canal. On considère que la conductivité dans chaque point du canal est exclusivement due à la présence des porteurs minoritaires, donc, pour le cas des transistors PMOS, on obtient à partir de l'équation (27),

$$(113) \quad \sigma(x) = q p(x) \mu_p(x)$$

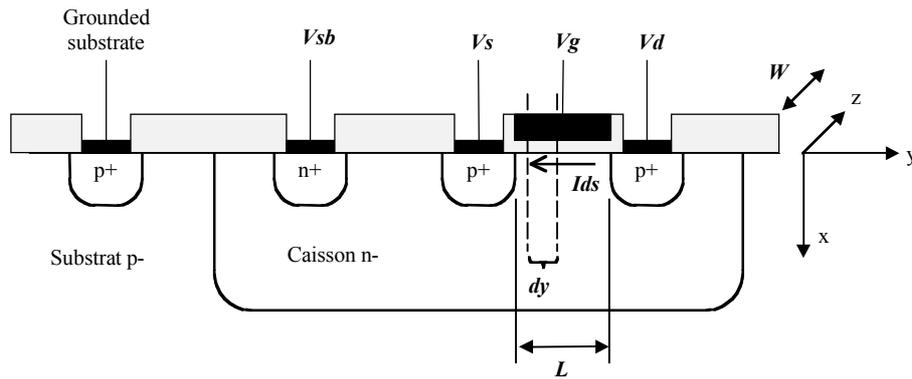


Figure 44. Structure d'un transistor PMOS de largeur W et longueur L .

En supposant que la mobilité μ_p ne varie pas en fonction de la profondeur x , la conductance dg correspondante à la tranche de canal dy dans la Figure 44 est alors donnée par,

$$(114) \quad dg = \frac{W}{fy} \int_0^{x_{inv}} \sigma(x) dx = \frac{W}{fy} \int_0^{x_{inv}} q p(x) \mu_p dx = \frac{W}{fy} \mu_p \int_0^{x_{inv}} |Q_p| dx$$

Où Q_p est la densité par unité de surface de la charge dans le canal due aux porteurs minoritaires. La variation de tension dV_c aux extrêmes de la tranche du canal dy est une fonction du courant drain-source I_{ds} est de la conductance dg ,

$$(115) \quad fV_c = \frac{I_{ds}}{fg} = \frac{I_{ds} \cdot fy}{W \cdot \mu_p \cdot |Q_p|}$$

La densité de charge par unité de surface Q_p , correspond à la concentration des porteurs minoritaires dans un point donné du canal. La charge dans le semiconducteur Q_{sc} est composée des charges fixes dans la région désertée et des porteurs libres dans la région d'inversion.

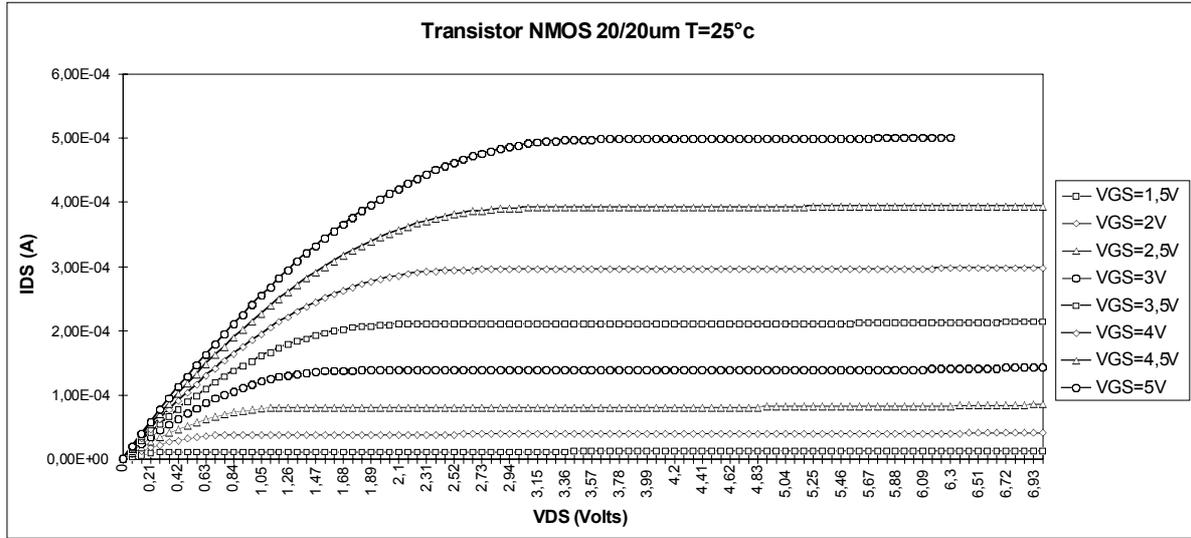


Figure 45. Caractéristique $I_{ds}(V_{ds}, V_g)$ d'un transistor NMOS long des technologies étudiées.

A partir de la Figure 41 et de l'équation (106), on peut considérer, d'une part qu'en régime de déplétion et d'inversion faible, la concentration des porteurs minoritaires dans le canal est négligeable, et d'autre part qu'en régime d'inversion forte, la zone désertée n'augmente plus. Il en résulte que les charges générées dans le semiconducteur pour $|V_g| > |V_{th}|$ sont essentiellement des porteurs minoritaires, donc dans le cas des transistors MOS en inversion forte,

$$(116) \quad |Q_p| = |V_g - V_{thP}(V_c(y))| \cdot C_{ox}$$

En combinant les deux dernières équations,

$$(117) \quad I_{ds} \cdot \int_0^L dy = \mu_p \cdot W \cdot C_{ox} \cdot \int_0^L \frac{V_d}{V_s} \cdot |V_g - V_{thP}(V_c)| \cdot fV_c$$

Et en intégrant le long du canal, on obtient l'expression suivante, du courant de drain-source I_{ds} des transistors NMOS et PMOS,

$$(118) \quad I_{ds} = \mu_{n,p} \cdot C_{ox} \cdot \frac{W}{L} \cdot \int_0^L \frac{V_d}{V_s} \cdot |V_g - V_{th_{N,P}}(V_c)| \cdot fV_c$$

Où $V_{th_{N,P}}(V_c)$ est donnée par les équations (107) et (108), respectivement, pour les transistors PMOS et NMOS. Le produit $\mu_{n,p} \cdot C_{ox}$ est normalement appelé Facteur de Gain $KP_{N,P}$ des transistors MOS. Dans la Figure 45 du courant I_{ds} d'un transistor NMOS en fonction de V_g , V_d , pour $V_s = 0$, on

observe deux régimes: un régime linéaire pour de faibles valeurs de V_d , et un régime de saturation pour de plus fortes valeurs de V_d , où le courant I_{ds} ne dépend plus de V_d .

3.4.2.1. Régime linéaire

Prenons le cas du transistor NMOS. Si on fait une approximation de premier ordre, l'équation (108) s'écrit,

$$(119) \quad V_{thN}(V_c) \cup V_{th0N} + \lambda_N \cdot V_c > 0$$

Où λ_n est un paramètre qui ne varie pratiquement pas selon la température.

Si on remplace (119) dans (118) et intègre de V_s à V_d , on obtient,

$$(120) \quad I_{ds} = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \frac{1}{2 \cdot \lambda_N} \cdot \left[(V_g - V_{th0N} - \lambda_N \cdot V_c) \right]_s^d$$

$$(121) \quad I_{ds} = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \overline{V_g - V_{th0N} - \frac{\lambda_N}{2} (V_d + V_s)} \cdot (V_d - V_s)$$

Dans le régime linéaire $V_d \ll V_g - V_{th0N}$, on obtient donc, pour $V_s = 0$, la réponse linéaire suivante,

$$(122) \quad I_{ds_{LIN}} \cup \mu_{n,p} \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_g - V_{th0_{N,P}}) \cdot V_d \quad \text{pour des NMOS ou PMOS.}$$

La transconductance g_m , pour des petits signaux, des transistors MOS en régime linéaire est donnée par,

$$(123) \quad g_{m_{LIN}} = \frac{f I_{ds_{LIN}}}{f V_g} = \mu_{n,p} \cdot C_{ox} \cdot \frac{W}{L} \cdot V_d$$

La conductance drain-source $1/r_d$ (résistance de sortie r_d), pour de petits signaux, des transistors MOS en régime linéaire s'écrit,

$$(124) \quad \frac{1}{r_{d_{LIN}}} = \frac{f I_{ds_{LIN}}}{f V_{ds}} \cup \mu_{n,p} \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_g - V_{th0_{N,P}})$$

3.4.2.2. Régime de saturation

Dans le régime de saturation, le courant drain-source I_{ds} devient indépendant de la tension de drain V_d . On constate dans l'équation (120) que, pour toute valeur de V_s , ce régime est atteint quand,

$$(125) \quad V_d = V_{d_{SAT}} = \frac{V_g - V_{th0N}}{\lambda_N}$$

Dans le régime de saturation, I_{ds} est donc donné par,

$$(126) \quad I_{ds_{SAT}} = \mu_{n,p} \cdot C_{ox} \cdot \frac{W}{L} \cdot \frac{(V_g - V_{th_{N,P}}(V_s))^2}{2 \cdot \lambda_{N,P}} \quad \text{pour des NMOS ou PMOS.}$$

La transconductance g_m , pour de petits signaux, en régime de saturation est donnée par,

$$(127) \quad gm_{SAT} = \frac{f Ids_{SAT}}{f Vg} = \sqrt{2 \mu_{n,p} Cox \frac{W}{L} Ids_{SAT}}$$

3.4.2.2.a. Conductance drain-source des transistors MOS en régime de saturation

Quand $Vc(y) \geq Vd_{SAT}$, sous la grille du transistor MOS, la couche d'inversion disparaît et le semiconducteur se trouve en régime de faible inversion ou de désertion. Les porteurs traversent cette zone désertée entraînés par le champ électrique ξ_y . Dans la Figure 46, les épaisseurs de la couche d'inversion et de la région désertée le long du canal sont représentées par les courbes X_{inv} et X_d . On observe que le canal s'étrangle à partir d'un point où $Vc(y) = Vd_{SAT}$.

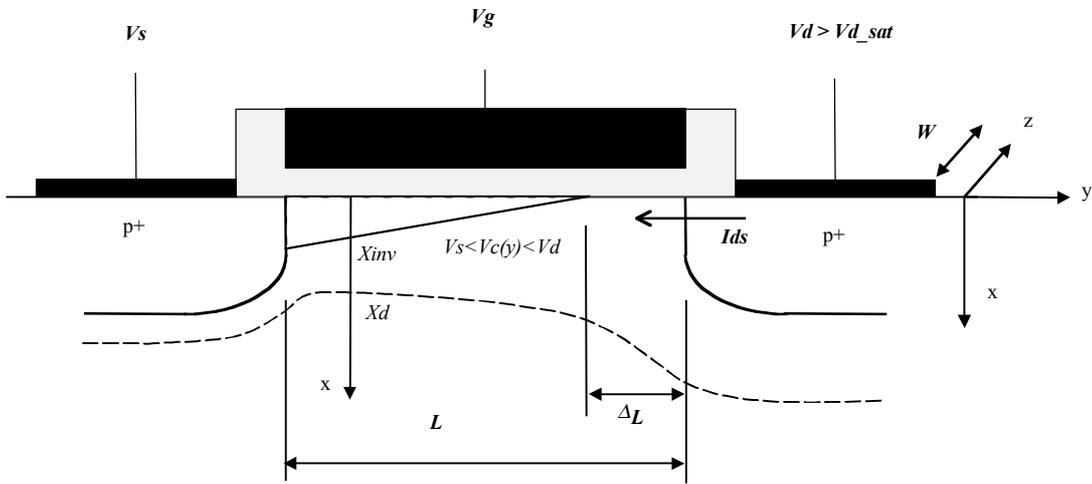


Figure 46. Canal du transistor MOS en régime de saturation.

On assiste donc à un raccourcissement de la longueur effective du canal, de la valeur de ΔL , laquelle est donnée, après avoir intégré l'équation de Poisson dans la zone désertée mentionnée, par,

$$(128) \quad \Delta L = \sqrt{\frac{2 \epsilon_{sc}}{q Nd} (Vd - Vd_{SAT})}$$

Le courant drain-source en régime de saturation, pour $Vd \geq Vd_{SAT}$, s'écrit donc,

$$(129) \quad Ids_{SAT}(Vd > Vd_{SAT}) = \mu_{n,p} Cox \frac{W}{L} \frac{(Vg - Vth_{N,P}(Vs))^2}{2 \lambda_{N,P}} \frac{1}{1 - \frac{\Delta L}{L}} = Ids_{SAT}(Vd_{SAT}) \sqrt{1 + \frac{\Delta L}{L}}$$

La conductance drain-source $1/rd$ (résistance de sortie rd) des transistors MOS, pour des petits signaux, est obtenue en dérivant l'expression précédente par rapport à Vd ,

$$(130) \quad \frac{1}{rd_{SAT}} = \frac{f Ids_{SAT}}{f Vd} = Ids_{SAT}(Vd_{SAT}) \frac{1}{L} \sqrt{\frac{\epsilon_{sc}}{2 q Nd (Vd - Vd_{SAT})}}$$

Le paramètre rd caractérise l'impédance de sortie des étapes d'amplification de source commune. La tension de Early Va , est définie (comme pour les transistors bipolaires) par,

$$(131) \quad V_a = I_{ds} \frac{fVd}{fI_{ds}} \cup L \sqrt{\frac{2 \mu_n \mu_p N_d}{\epsilon_{sc}} Vd - \frac{Vg - V_{th0_{N,P}}}{\lambda_{N,P}}}$$

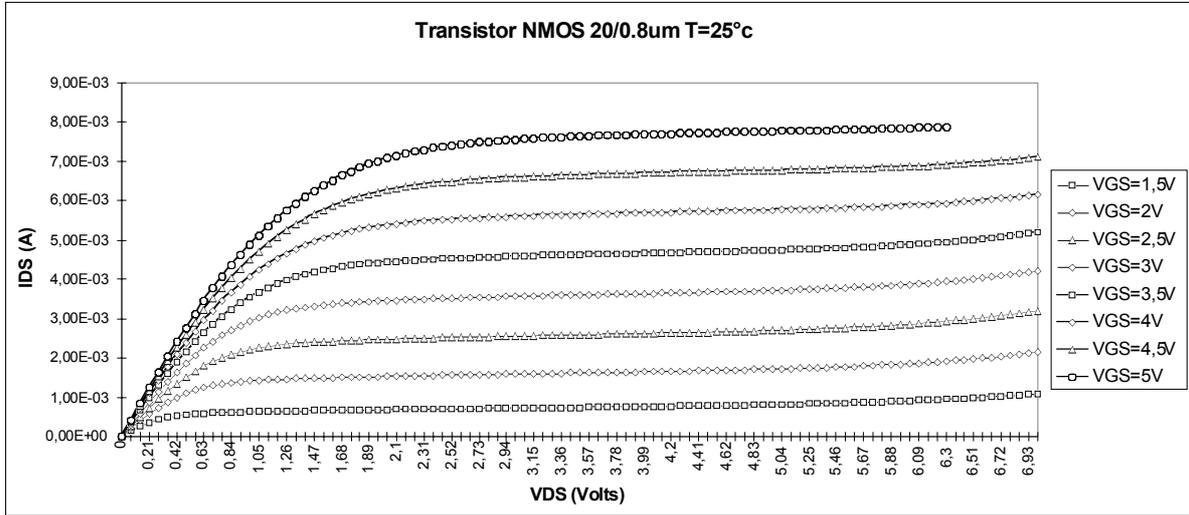


Figure 47. Caractéristique $I_{ds}(V_{ds}, V_g)$ d'un transistor NMOS court des technologies étudiées.

La tension de *Early* dépend fortement de la longueur physique du canal L (comparer les pentes en régime de saturation dans la Figure 47 et la Figure 45) et des conditions de polarisation V_d et V_g . En revanche, elle dépend faiblement de la température à travers la racine carrée de V_{th0} , ce qui ne représente pas une limitation importante pour l'utilisation des transistors MOS en haute température.

3.4.3. Mobilité

Dans la section 3.4.2 nous avons mis en évidence l'importance de la mobilité dans les caractéristiques des transistors MOS concernant le courant I_{ds} , la conductivité drain-source et la transconductance en inversion forte (voir les équations (122), (123), (124), (126) et (127)). Les variations de la mobilité modulent directement les caractéristiques des transistors MOS.

La mobilité dans le volume (voir la section 2.1.7.1) est caractérisée par les interactions des porteurs avec les phonons acoustiques et les impuretés dopantes. Comme dans un transistor MOS le flux de porteurs minoritaires est confiné dans la couche d'inversion, leur mobilité et vitesse de dérive sont influencées par la rugosité de l'interface $Si-SiO_2$. La mobilité en surface μ_s (faiblement dépendante de la température), due à la présence de l'interface $Si-SiO_2$, s'ajoute aussi à l'expression de la mobilité totale μ_0 ,

$$(132) \quad \frac{1}{\mu_0} = \frac{1}{\mu_s} + \frac{1}{\mu_{ph}} + \frac{1}{\mu_c}$$

Le modèle le plus couramment utilisé par les simulateurs du type SPICE pour la mobilité est le suivant,

$$(133) \quad \mu_{0n,p}(T) = \mu_{0n,p}(T_{amb}) \sqrt{\frac{T}{T_{amb}}}^{TC_{n,p}}$$

Ce modèle est valable pour la plage de température de 150 à 500K, environ.

La mobilité correspondant aux électrons et trous dans la couche d'inversion à été mesurée, respectivement avec des transistors NMOS ($W/L=20/20\mu\text{m}$) et PMOS ($W/L=20/20\mu\text{m}$), en régime linéaire. Les résultats des mesures, plus leurs approximations faites à partir de la formule (133), sont montrés dans la Figure 48.

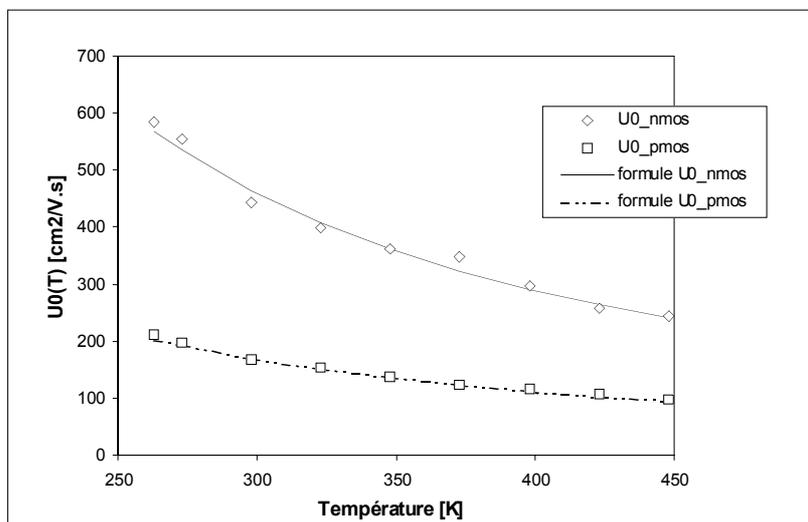


Figure 48. Mobilités des électrons(NMOS) et trous (PMOS), dans la couche d'inversion des transistors MOS.

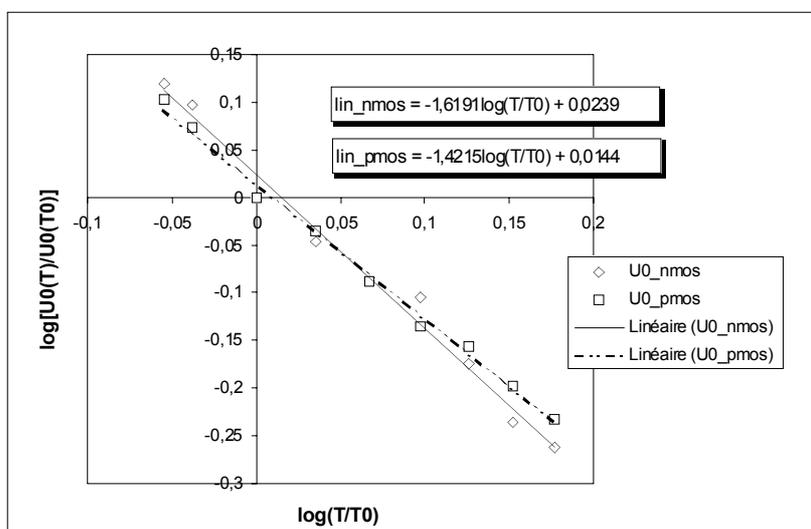


Figure 49. Extraction du coefficient thermique de la mobilité des électrons(NMOS) et trous (PMOS), dans la couche d'inversion des transistors MOS.

Les valeurs de la mobilité dans la couche d'inversion, mesurées à température ambiante, $\mu_{0n}=443 \text{ cm}^2\text{V}^{-1}\text{sec}^{-1}$ et $\mu_{0p}=167 \text{ cm}^2\text{V}^{-1}\text{sec}^{-1}$ sont bien inférieures à celles de la mobilité dans le volume (voir la section 2.1.7.1). Les coefficients thermiques $TCn=-1.62$ et $TCp=-1.42$, utilisés dans les approximations de la Figure 48, ont été obtenus à partir de la pente logarithmique dans la Figure 49.

La réduction de la mobilité selon la température affecte d'une manière générale les caractéristiques électriques des transistors MOS. Elle doit être prise en compte pour la conception des circuits intégrés en haute température, mais elle ne représente pas une limitation incontournable pour l'utilisation des transistors MOS en haute température.

3.4.3.1. Mobilité effective dans la couche d'inversion

Si un faible champ électrique ξ_y est appliqué (dans la direction y de la Figure 44) dans le canal d'un transistor MOS (faible V_{ds} et/ou transistor long), la vitesse de dérive varie linéairement selon ξ_y , la mobilité μ_{eff} étant le facteur de proportionnalité (voir l'équation (20)). Dans un transistor MOS, le flux de porteurs minoritaires étant confiné dans la couche d'inversion, la mobilité de ceux-ci dépend de la rugosité de l'interface $Si-SiO_2$ et du champ électrique ξ_x transversal à l'interface. La mobilité effective est donc une fonction du champ électrique transversal ξ_x , autrement dit, une fonction de la polarisation de la capacité MOS. Pour des valeurs de champ électrique modérées, les variations de la mobilité sont décrites par la relation empirique suivante,

$$(134) \quad \mu_{eff}(V_g, T) = \frac{\mu_0(T)}{(1 + PhiG(T))(V_g - V_{th0}(T))}$$

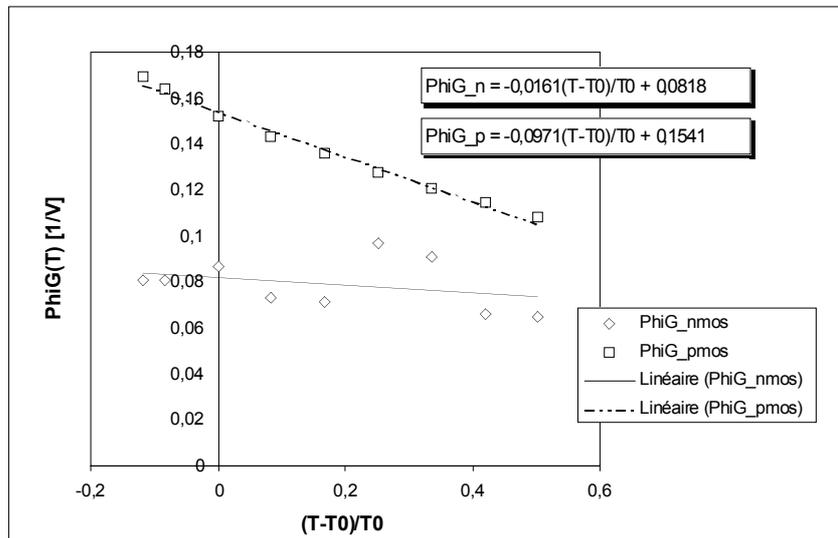


Figure 50. Extraction du coefficient $\Phi_iG(T)$ de la mobilité des électrons (NMOS) et trous (PMOS), dans la couche d'inversion des transistors MOS.

A partir de l'équation de I_{ds} en régime linéaire (122) et de l'expression précédente, les paramètres $\mu_{0N,P}(T)$ et $\Phi_iG(T)$ peuvent être extraits à l'aide de,

$$(135) \quad C_{ox} \frac{W}{L} \frac{(V_g - V_{th0_{N,P}}) V_{ds}}{I_{ds_{LIN}}} = \frac{1}{\mu_{0N,P}} + \frac{\Phi_iG_{N,P}}{\mu_{0N,P}} (V_g - V_{th0_{N,P}}) = F_{\mu} [V_g - V_{th0_{N,P}}]$$

Les variables et les paramètres à gauche dans l'équation (135) sont connus, à partir de la mesure de $I_{ds}(V_g, V_{ds})$ en régime linéaire. A chaque température, on extrait $\mu_{0N,P}$ et $\Phi_iG_{N,P}$, respectivement, à partir de la valeur à l'origine et de la pente de $F_{\mu}[V_g - V_{th0_{N,P}}]$. Les valeurs ainsi extraites, correspondant à $\mu_{0N,P}$ et $\Phi_iG_{N,P}$ en fonction de la température, sont montrées respectivement dans la Figure 48 et la Figure 50. On constate que l'influence du champ électrique transversal ξ_x dans la mobilité se réduit au fur et à mesure que la température augmente.

3.4.3.2. Saturation de la vitesse de dérive

Pour de faibles valeurs du champ électrique ξ_y la vitesse de dérive varie linéairement selon ξ_y ($v = \mu_{eff} \cdot \xi_y$), mais à partir d'une certaine valeur la vitesse de dérive sature et n'augmente plus proportionnellement à ξ_y ($v = v_{s}(\xi_y)$). Cette saturation de la vitesse des porteurs a deux effets: d'une part le courant des transistors MOS en régime de saturation est fortement réduit, particulièrement dans le cas des transistors courts polarisés avec de larges valeurs de V_{ds} , et d'autre part, le courant de saturation devient une fonction linéaire de la tension V_{gs} , au lieu de répondre à la fonction quadratique (126). Ainsi, le courant des transistors MOS en régime de saturation devient,

$$(136) \quad I_{ds_SAT} = C_{ox} \cdot W \cdot (V_g - V_{th_{N,P}}(V_s)) \cdot v_s$$

La transconductance g_m , pour des petits signaux, en régime de saturation s'écrit donc,

$$(137) \quad g_{m_SAT} = \frac{f I_{ds_SAT}}{f V_g} = W \cdot C_{ox} \cdot v_s$$

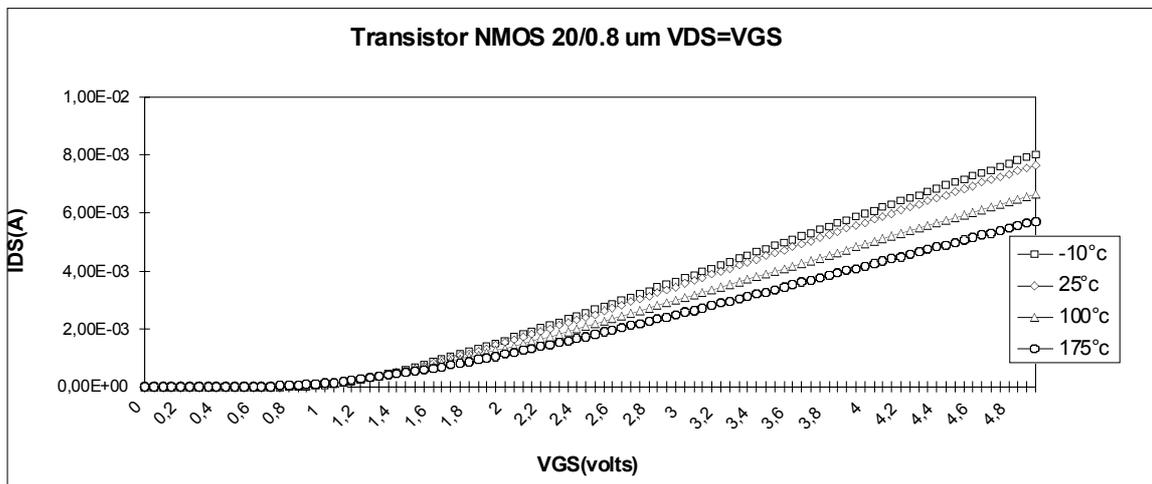


Figure 51. Courant drain-source I_{ds} d'un transistor NMOS court en régime de saturation.

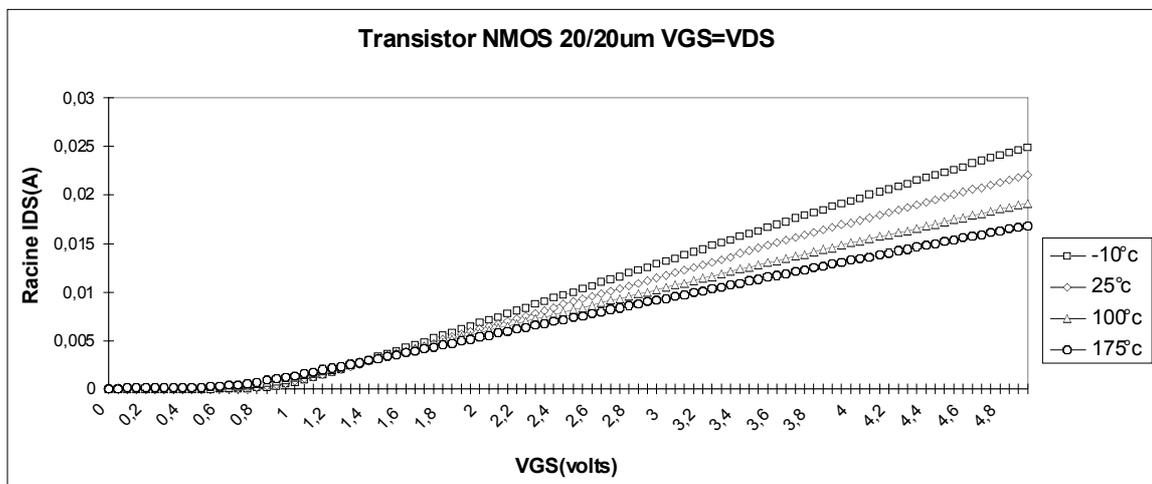


Figure 52. Racine carré de I_{ds} d'un transistor NMOS long en régime de saturation.

Dans la Figure 51 on observe la réponse linéaire de la courbe $I_{ds}(V_{gs})$, pour de fortes valeurs de V_{ds} ($V_{ds}=V_{gs}$) d'un transistor NMOS court en régime de saturation. Dans la Figure 52, la réponse linéaire correspond à la *racine carrée de $I_{ds}(V_{gs})$* d'un transistor NMOS long en régime de saturation.

3.4.4. Courant drain-source en fonction de la température

Comme il a été mentionné dans les sections précédentes, la tension seuil et la mobilité diminuent avec la température. L'équation (118) permet de voir que les influences de la diminution de ces deux parametres dans le courant I_{ds} s'opposent. Il en résulte que pour des faibles valeurs de V_{gs} , le courant augmente avec la température, et pour de fortes valeurs de V_{gs} , le courant I_{ds} diminue au fur et à mesure que la température augmente (voir la Figure 53 et la Figure 54). La valeur de V_{gs} pour laquelle le courant I_{ds} ne varie pas en fonction de la température, dans une plage de température donnée, est normalement nommée $V_{gs_{ZTC}}$ (Zero Temperature Coefficient).

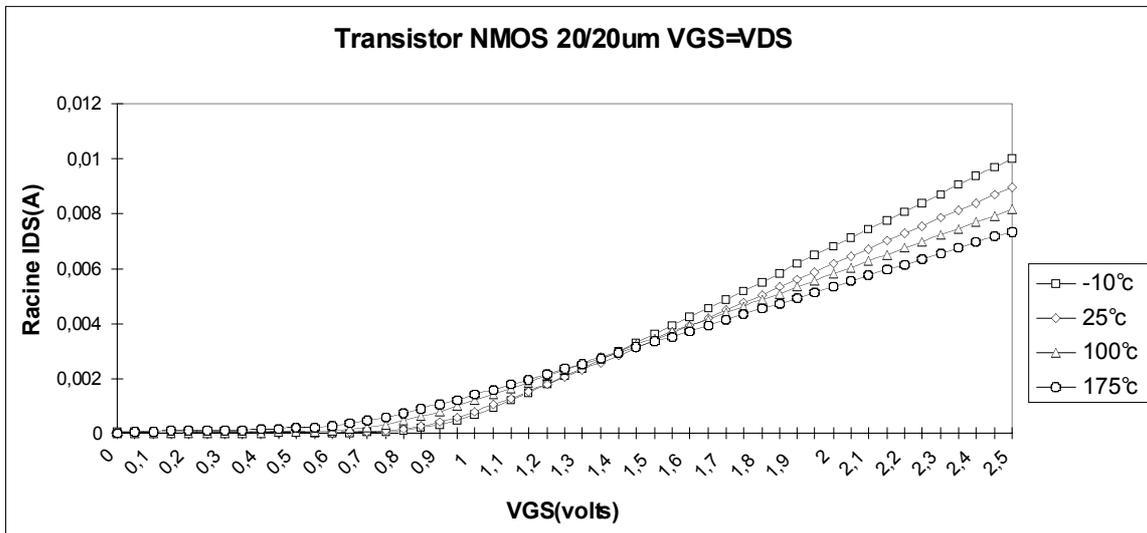


Figure 53. Racine carrée de I_{ds} d'un NMOS en régime de saturation, en fonction de la température.

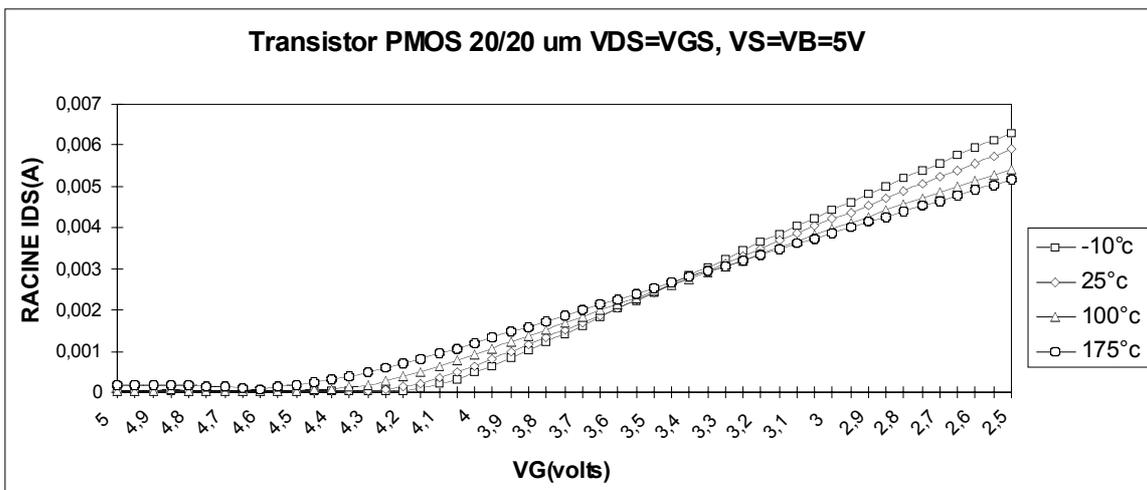


Figure 54. Racine carrée de I_{ds} d'un PMOS en régime de saturation en fonction de la température.

On linéarise l'évolution de la tension de seuil, avec la fonction suivante,

$$(138) \quad |V_{th_{N,P}}(T)| = |V_{th_{N,P}}(T_{amb})| - |TCv_{N,P}|(T - T_{amb})$$

et on considère que l'exposant thermique de la mobilité est approximativement $TC_{n,p} \approx -1.5$ pour les deux types de porteurs. A partir de la dérivée de l'équation (126) par rapport à la température, on obtient donc l'expression suivante, pour le calcul du $V_{gs_{ZTC}}$ des transistors MOS en régime de saturation,

$$(139) \quad |V_{gs_{ZTC}}| = |V_{th_{N,P}}(T_{amb})| + |TCv_{N,P}| T_{amb} + \frac{T_{max} + T_{min}}{6} \sqrt{\quad}$$

Les valeurs de $V_{gs_{ZTC}}$ calculées, à partir des données de la Figure 42 et l'équation précédente, pour les cas des NMOS et PMOS sont très proches des valeurs observés: $V_{gs_{ZTC_N}} \approx 1.45$ volts et $V_{gs_{ZTC_P}} \approx -1.65$ volts respectivement dans la Figure 53 et de la Figure 54.

3.4.5. Régime de faible inversion des transistors MOS en haute température.

En inversion faible, le courant drain-source I_{ds} est déterminé par le phénomène de diffusion des porteurs minoritaires à travers le canal, comme dans le cas des transistors bipolaires [1]. Le courant I_{ds} résultant varie exponentiellement selon la tension de grille V_g ,

$$(140) \quad I_{ds_{F_INV}} = I_{ds0} e^{-\frac{C_{ox}}{C_{ox} + Cd_0 + C_{ss}} \frac{qV_g}{kT}}$$

$$(141) \quad g_{m_{F_INV}} = \frac{fI_{ds_{F_INV}}}{fV_g} = \frac{C_{ox}}{C_{ox} + Cd_0 + C_{ss}} \frac{q}{kT} I_{ds_{F_INV}}$$

Où $Cd_0 = \epsilon_{sc}/Xd$ est la capacité par unité de surface de la zone désertée, laquelle varie selon la racine carrée de la densité des dopants dans le canal, et C_{ss} est la capacité associée aux états de surface, laquelle dépend fortement de la densité d'états à l'interface $Si-SiO_2$. Il en résulte que la caractéristique $I_{ds}(V_g)$ en inversion faible est beaucoup plus sensible aux paramètres de fabrication, et à leurs tolérances, qu'en inversion forte.

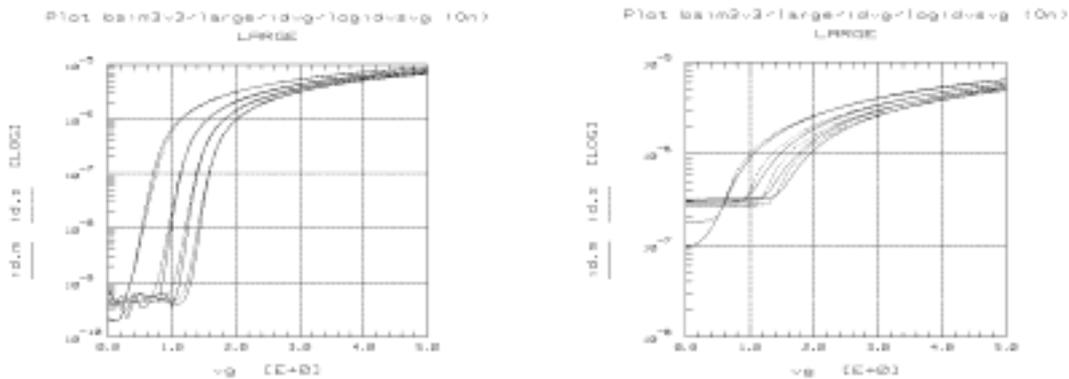


Figure 55. Caractéristiques des transistors NMOS en faible inversion à 150°C (gauche) et 250°C (droite), pour différentes valeurs de la tension du substrat.

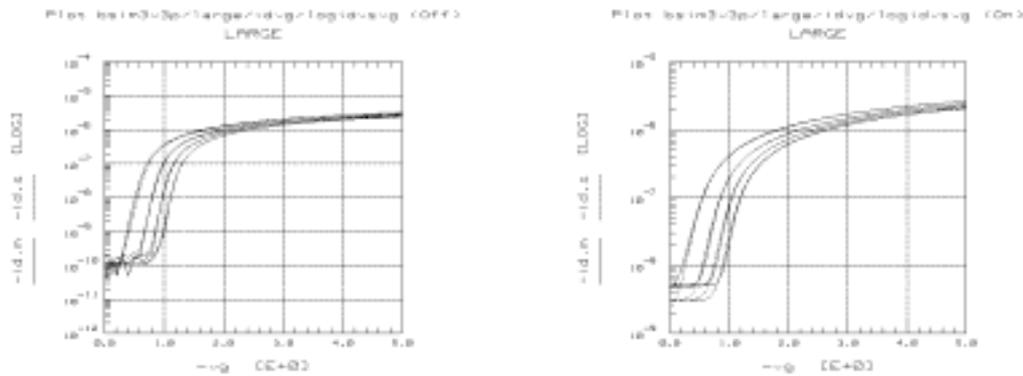


Figure 56. Caractéristiques des transistors PMOS en faible inversion à $150^\circ C$ (gauche) et $250^\circ C$ (droite), pour différentes valeurs de la tension du caisson.

A température ambiante les équations (140) et (141) s'appliquent à 5 ou 6 ordres de grandeur du courant drain-source. En haute température les courants de fuite réduisent fortement la plage des courants en faible inversion (voir Figure 55 et Figure 56), ce qui rend inintéressante l'utilisation de la réponse exponentielle des transistors MOS en inversion faible, dans les applications en haute température. Mais, il est intéressant de remarquer que dans les courbes de la Figure 55 et de la Figure 56, le courant en régime *ouvert* (i.e. $V_g=0$) ne correspond pas à la conduction en régime de faible inversion précédemment décrite, parce que le phénomène du courant de fuite dans une jonction inversement polarisée est dominant. Cependant, dans les technologies où les tensions de seuil sont réduites (à cause de la réduction des tensions d'alimentation) le courant en régime *ouvert* pourrait plutôt répondre au phénomène de conduction dans un canal en régime de faible inversion.

3.4.6. Courants de fuite et conductances dans le drain et la source.

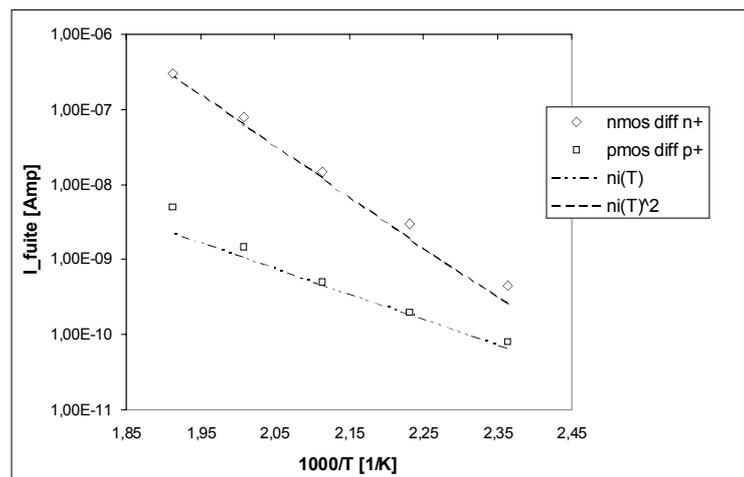


Figure 57. Courants de fuite dans les diffusions de drain ou de la source des transistors MOS, en fonction de la température (150 à $250^\circ C$).

Dans les transistors MOS des technologies sur substrat de silicium étudiées, la surface de la jonction d'isolation entre les diffusions (fortement dopées) du drain ou de la source, et le substrat p - ou le caisson n - (faiblement dopées), est très large par rapport à la surface de la section transversale du

canal. Autrement dit, ces diffusions ont une petite partie active, qui sert à l'injection et à la collection des porteurs minoritaires circulant par le canal, et une grande partie électriquement passive, qui doit être isolée du substrat ou du caisson. Cette isolation est faite, dans le cas des technologies CMOS et BiCMOS étudiées, à travers la polarisation en inverse de la jonction. Les courants de fuite dans la source I_{fs} et dans le drain I_{fd} ainsi générés (voir Figure 57), entrent en concurrence avec le courant de conduction drain-source au fur et à mesure que la température augmente.

3.4.6.1. Courants de fuite et conductance dans les transistors NMOS

Les courants de fuite dans la source I_{fs} et dans le drain I_{fd} des transistors NMOS (voir Figure 39), sont donnés par l'équation (73) multipliée par la surface A_s ou A_d de la jonction correspondante. Les courants de fuite dans ces jonctions augmentent fortement avec la température (voir la section 2.3.2.5 et la Figure 57) représentant ainsi une forte limitation à l'utilisation en haute température des transistors MOS. (voir dans le Tableau 5 et le Tableau 6 les valeurs des courants de fuite à température ambiante dans les diffusions n^+ , des technologies étudiées)

Prenons le cas du drain d'un transistor NMOS, dans la Figure 58. Dans la plage de température où la génération de porteurs dans la ZCE (voir l'équation (71)) est dominante dans l'expression (73) de la densité du courant de fuite, la dépendance de ce courant I_{fd_g} par rapport à la tension de polarisation V_d permet de définir la conductance G_{fd_g} , pour des petits signaux autour d'une valeur V_{dq} , suivante,

$$(142) \quad G_{fd_g} = \left. \frac{fI_{fd_g}}{fV_d} \right|_{V_{dq}} = q \cdot A_d \cdot N_t \cdot C \cdot \frac{ni}{2} \cdot \left. \frac{fW(V_d)}{fV_d} \right|_{V_{dq}} = A_d \cdot N_t \cdot C \cdot \frac{ni}{2} \cdot \frac{\epsilon_{sc} \cdot \frac{Na + Nd}{Na \cdot Nd}}{W(V_{dq})}$$

Dans la plage de température où la diffusion de porteurs (voir l'équation (68)) est dominante dans l'expression (73) de la densité du courant de fuite, la conductance pour les petits signaux ne résulte plus de la modulation de la densité du courant de fuite, mais elle est générée par une modulation de la taille de la surface à la limite de la ZCE, du côté le moins dopé de la jonction [34].

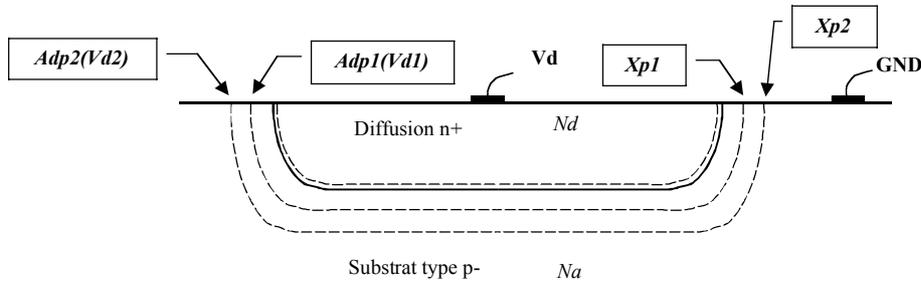


Figure 58. Variation de la surface Ad_p à la limite de la ZCE en fonction de la polarisation.

Dans une jonction dissymétrique, comme celle de la Figure 58, le courant de fuite de diffusion I_{fd_d} est dominé par la diffusion des électrons du substrat vers le collecteur, décrite par l'équation (66) évalué dans la limite x_p de la ZCE, i.e. du côté le moins dopée de la jonction,

$$(143) \quad I_{fd_d} \simeq \frac{q \cdot Ad_p \cdot ni^2 \cdot D_n}{Na \cdot Ln} \cdot \left. e^{-\frac{qV_d}{kT}} \cdot \sqrt{e^{\frac{(x-x_p)}{Ln}}} \right|_{x=x_p} = \frac{q \cdot Ad_p(V_d) \cdot ni^2 \cdot D_n}{Na \cdot Ln}$$

Dans une jonction pn tridimensionnelle, comme celle de la Figure 58, la surface Ad_p est une fonction de la tension de polarisation V_d . Il en résulte que le courant de fuite de diffusion I_{fd_d} est

modulé par V_d à travers $Adp(V_d)$. On définit la conductance Gfd_d , pour des petits signaux autour d'une valeur Vdq , de la façon suivante,

$$(144) \quad Gfd_d = \frac{fIfd_d}{fVd} = \frac{q \cdot ni^2 \cdot Dn}{Na \cdot Ln} \cdot \frac{fAdp}{fxp} \cdot \frac{fxp}{fVd} = \frac{ni^2 \cdot Dn}{Na^2 \cdot Ln} \cdot \frac{esc \cdot Nd}{(Na + Nd)} \cdot \frac{1}{xp(Vdq)} \cdot \frac{fAdp}{fxp}$$

Où $dAdp/dxp$ est un facteur qui dépend uniquement de la géométrie de la jonction.

Dans les cas d'une jonction dissymétrique la conductance totale Gfd due aux courants de fuite, est donnée par,

$$(145) \quad Gfd = \frac{esc}{Na \cdot xp(Vdq)} \cdot \overline{ni(T)} \cdot \frac{Ad \cdot Nt \cdot C}{2} + \frac{ni(T)^2 \cdot Dn}{Na \cdot Ln} \cdot \frac{fAdp}{fxp}$$

Gfd augmente fortement avec la température, d'une manière très similaire au courant de fuite qui la génère. Dans le cas des transistors NMOS courants (i.e. la diffusion du drain à une taille minimale), cette conductance Gfd reste normalement inférieure à la conductance drain-source $1/r_d$ mentionnée dans la section 3.4.2.2.a, même à haute température (<250°C [34]). Dans les courbes $idvd$ de la Figure 73, on observe que la pente en régime de saturation des NMOS ne change pas beaucoup entre 150 et 250°C.

3.4.6.2. Courants de fuite et conductance dans les transistors PMOS

Les courants de fuite dans la source I_{fs} et dans le drain I_{fd} des transistors PMOS sont affectés par la présence d'un transistor bipolaire vertical parasite pnp [35] de section A_{ev} , polarisé en configuration de base commune (voir la Figure 59). Considérons le cas des courants de fuite dans le drain du transistor de la Figure 59, où la source est connectée au caisson. Le courant d'émetteur (drain), correspondant au modèle d'Ebers-Molls unidimensionnel [1] du transistor bipolaire ainsi défini, est donné par,

$$(146) \quad \frac{I_e}{A_{ev}} = J_{s_{BE}} \cdot e^{-\frac{qV_{BE}}{kT}} - 1 \cdot \alpha_F \cdot J_{s_{BE}} \cdot e^{-\frac{qV_{BC}}{kT}} - 1 \cdot J_{g_{BE}}$$

Où la densité de courant de diffusion $J_{s_{BE}}$ et la densité de courant de génération $J_{g_{BE}}$ sont données, respectivement, par les équations (68) et (71).

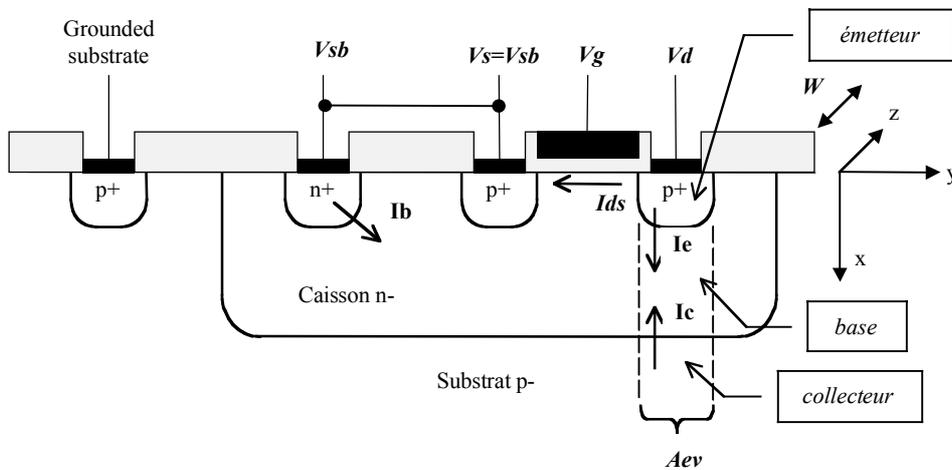


Figure 59. Courants dus au transistor bipolaire vertical parasite des transistors PMOS.

Prenant en compte que $\alpha_F = \beta_F/(\beta_F+1)$ et les conditions normales de fonctionnement du transistor PMOS : $V_{be} \ll 0$ et $V_{bc} \ll 0$, l'expression précédente s'écrit,

$$(147) \quad \frac{I_e}{A_{ev}} = Jfd = -J_{s_{BE}} \frac{1}{\beta_F + 1} + J_{g_{BE}}$$

La présence du transistor bipolaire entraîne une atténuation du courant de diffusion proportionnel au gain de courant β_F . Cette atténuation est moins importante dans le cas de la technologie BiCMOS étudiée, du fait de la présence d'une couche enterrée n^+ au fond du caisson qui réduit fortement la valeur de β_F du transistor bipolaire vertical parasite (voir dans le Tableau 13 les valeurs du gain de courant β_F du transistor vertical parasite, dans les technologies étudiées).

Il en résulte que le courant de diffusion, proportionnel à $ni^2(T)$, deviendra dominant par rapport au courant de génération (proportionnel à $ni(T)$), à des températures encore plus hautes, en fonction de la valeur de β_F . Dans la Figure 57, on observe que le courant de fuite des PMOS continue à être proportionnel à $ni(T)$ jusqu'à 250°C environ, tandis que dans le cas des NMOS le courant de fuite est déjà proportionnel à $ni^2(T)$ à 150°C.

A partir des équations (73) et (147) on obtient l'expression suivante pour la densité du courant de fuite dans le drain ou la source $Jfd_{,s}$ d'un transistor PMOS construit sur un caisson n^- ,

$$(148) \quad Jfd_{,s} = -q \cdot ni(T)^2 \cdot \left[\frac{Dp}{Nd \cdot Lp} + \frac{Dn}{Na \cdot Ln} \cdot \frac{1}{\beta_F + 1} \right] + Ni \cdot C \cdot ni(T) \cdot \left[\frac{esc}{2 \cdot q} \cdot \frac{1}{Nd} + \frac{1}{Na} \cdot \sqrt{V_{bi}(T) - 2} \cdot \frac{k \cdot T}{q} - V_{d,s} \right]$$

Dans le cas normal où les diffusions du drain et de la source sont plus fortement dopées que le caisson, la conductance du drain ou de la source due aux courants de fuite est obtenue d'une manière analogue à l'expression (145),

$$(149) \quad Gf_{d,s} = \frac{esc}{Nd \cdot xn(V_{d,s})} \cdot ni(T) \cdot \frac{A_{d,s} \cdot Ni \cdot C}{2} + \frac{ni(T)^2 \cdot Dp}{Nd \cdot Lp} \cdot \frac{f_{A_{d,s}} \cdot n}{fxn} \cdot \frac{1}{\beta_F + 1}$$

Même si sensiblement inférieurs, ces courants de fuite et conductances, dans le drain et la source des PMOS, augmentent fortement aussi avec la température, ayant pratiquement le même type de conséquences sur la performance des PMOS que dans le cas des transistors NMOS.

3.4.7. Courants de fuite et conductance dans le caisson des PMOS

Normalement, le caisson n^- est polarisé positivement par rapport au substrat, par une des sources d'alimentation ou un nœud de très basse impédance du circuit. Dans ces cas, les courants de fuite en haute température s'additionnent simplement à la consommation du circuit, sans affecter son fonctionnement. Mais dans les configurations des transistors où il est nécessaire d'éviter l'effet *body*, dans la valeur de la tension de seuil, en polarisant le caisson avec la tension de la source V_s , les courants de fuite générés dans la comparativement très large jonction caisson-substrat s'additionnent au courant de source. Le courant de fuite et sa conductance associée répondent aux équations développées dans les sections 3.4.6.1.

Dans le cas de la technologie BiCMOS étudiée, la couche enterrée n^+ fortement dopée, peut aider à la réduction de ces courants de fuite. Cependant, comme dans une jonction dissymétrique la région la moins dopée domine dans l'expression (73), si le dopage du substrat reste constant la réduction du courant de fuite caisson-substrat, due à la couche enterrée, sera en pratique très limitée.

3.5. Transistors bipolaires

Grâce à certains avantages, tels que la possibilité d'implémenter des circuits analogiques et numériques dans la même puce, la faible consommation et les faibles dimensions des transistors, etc., la technologie CMOS domine encore la conception VLSI. Même s'il est possible de réaliser la plupart des fonctions analogiques avec la technologie CMOS, quelques-unes ne peuvent être réalisées qu'avec des transistors bipolaires, ceci étant dû à la plus large transconductance, la plus large capacité de commander des charges capacitives, le plus faible bruit $1/f$, et le meilleur degré de similitude ("matching") entre transistors. C'est la raison de l'apparition des technologies BiCMOS, qui combinent les deux types de transistors dans la même puce, et de l'implémentation dans les technologies CMOS des transistors bipolaires latéraux CLBT.

3.5.1. Transistors bipolaires verticaux

La structure d'un transistor bipolaire *npn* vertical, pour des faibles puissances, est décrite dans la Figure 60. On suppose que la région active du transistor est limitée au prisme vertical, de section effective A_e (qui peut être modulée par plusieurs phénomènes, tels que les courants transversaux dans la base, etc. lesquels ne seront pas abordés dans cette analyse), que la base, l'émetteur et le collecteur sont uniformément et assez fortement dopés (chutes de tension négligeables), que le champ électrique est nul en dehors des ZCE, et qu'on est toujours en régime de faible injection.

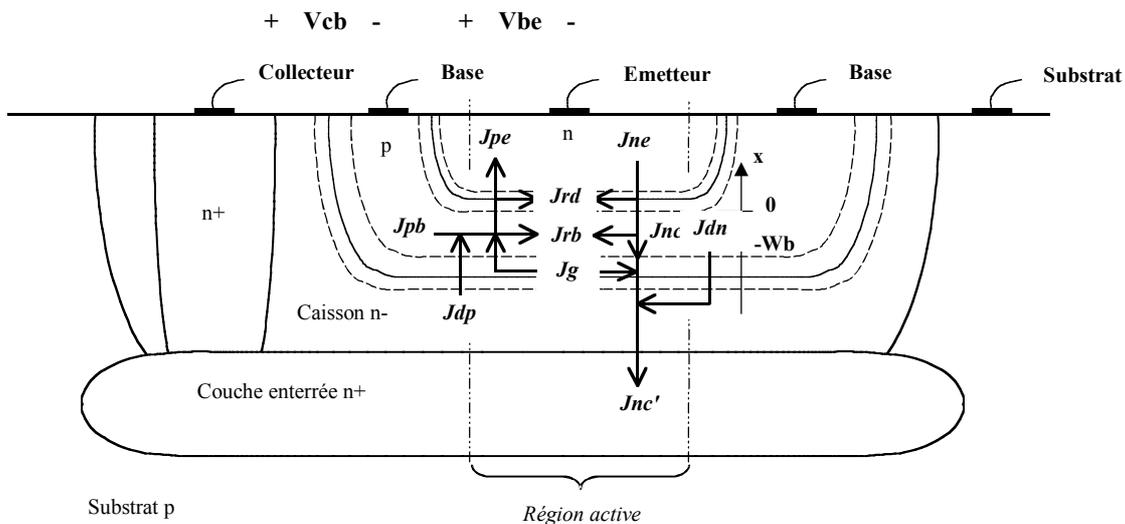


Figure 60. Structure du transistor bipolaire NPN vertical.

Les densités de courants des électrons et trous, dans les trois régions, ainsi que leur nomenclature sont décrites dans la Figure 60. L'effet des courants de fuite, de la jonction base-collecteur polarisée en inverse, sur le gain de courant β et la conductance du collecteur des transistors bipolaires, seront étudiés en particulier.

3.5.1.1. Les courants dans la Base et le Collecteur

La densité de courant à la base J_b est égal à l'injection des trous dans l'émetteur J_{pe} , plus les courants de recombinaison dans la base J_{rb} et dans la ZCE de la jonction base-émetteur J_{rd} , moins les

densités de courant (des trous) de génération Jg et de diffusion Jdp injectée dans la base, moins la densité de courant (des électrons) de diffusion Jdn (voir Figure 60). La densité de courant dans le collecteur Jc est égale à la densité de courant dans l'émetteur Jne , moins Jrd , moins Jrb , plus les densités du courant de génération Jg et de diffusion Jdn et Jdp .

$$(150) \quad Jb = Jpb - Jdn = Jpe + Jrd + Jrb - Jdp - Jg - Jdn = Jpe + Jrd + Jrb - Jf$$

$$(151) \quad Jc = Jnc' + Jdp = Jnc + Jg + Jdn + Jdp = Jnc + Jf$$

A travers Jf on met en évidence l'influence de la densité du courant de fuite dans la jonction base-collecteur, correspondant au courant de fuite d'une jonction inversement polarisée mentionné dans la section 2.3.2.5. Si l'épaisseur de l'émetteur est plus grande que la longueur de diffusion des trous dans l'émetteur Lpe , la densité de courant des trous Jpe injectée dans l'émetteur est donnée par l'équation (67) évaluée au bord de la ZCE de la jonction base-émetteur,

$$(152) \quad Jpe = jp(xn) \cup \frac{q \ni ni^2 Dpe}{Nde \ni Lpe} \ni e^{\frac{qVbe}{kT}} \quad \text{pour } Vbe \gg kT/q.$$

La densité de courant de recombinaison Jrd dans la ZCE de la jonction base-émetteur est donnée par l'intégration dans la ZCE du taux de recombinaison maximal des porteurs Ube (voir l'équation (69)), autrement dit, pour $n=p=ni \cdot \exp(qVbe/2kT)$,

$$(153) \quad Jrd = q \int_{xp}^{xn} Ube \ni fx$$

et,

$$(154) \quad Ube = \frac{Nt \ni C}{2} \ni ni \ni e^{\frac{qVbe}{2kT}} \quad \text{pour } Vbe \gg kT/q.$$

La densité de courant de recombinaison Jrd s'écrit donc,

$$(155) \quad Jrd = \frac{q \ni Nt \ni C}{2} \ni ni \ni W_{ZCEbe} \ni e^{\frac{qVbe}{2kT}} \quad \text{pour } Vbe \gg kT/q.$$

Où W_{ZCEbe} est la largeur de la ZCE de la jonction base-émetteur, donnée par l'équation (72). Dans un régime d'injection faible, la distribution des porteurs majoritaires dans la base n'est pas affectée. Le champ électrique est nul dans la base, alors que le gradient de concentration des porteurs minoritaires est important, à cause du phénomène d'injection. Le courant est essentiellement un courant de diffusion limité par le phénomène de recombinaison dans le volume. En régime stationnaire, la distribution des porteurs minoritaires dans la base est alors donnée par l'équation de continuité (62), avec les conditions suivantes,

$$(156) \quad n_b(0) = n_{bo} \ni e^{\frac{qVbe}{kT}} \quad \text{du côté de l'emetteur}$$

$$(157) \quad n_b(-Wb) = 0 \quad \text{du côté du collecteur}$$

La densité des porteurs minoritaires dans la base s'écrit alors,

$$(158) \quad n_b(x) \cup n_{bo} = \frac{\text{Sh} \frac{-Wb+x}{Lnb} \sqrt{\frac{qV_{be}}{kT}}}{\text{Sh} \frac{-Wb}{Lnb} \sqrt{\frac{qV_{be}}{kT}}} \quad \text{pour } V_{be} \gg kT/q.$$

Les densités de courant des électrons dans la base du côté de l'émetteur et du côté du collecteur, sont respectivement données par,

$$(159) \quad J_{ne} = q D_{nb} \left. \frac{fn_b}{fx} \right|_{x=0} = \frac{q D_{nb} ni^2}{Lnb Nab} \coth \frac{-Wb}{Lnb} \sqrt{\frac{qV_{be}}{kT}} \quad \text{pour } V_{be} \gg kT/q.$$

$$(160) \quad J_{nc} = q D_{nb} \left. \frac{fn_b}{fx} \right|_{x=-Wb} = \frac{q D_{nb} ni^2}{Lnb Nab \text{Sh} \frac{-Wb}{Lnb} \sqrt{\frac{qV_{be}}{kT}}} e^{\frac{qV_{be}}{kT}} \quad \text{pour } V_{be} \gg kT/q.$$

La densité de courant de recombinaison J_{rb} dans la base, correspond à la différence entre les deux équations précédentes,

$$(161) \quad J_{rb} = \frac{q D_{nb} ni^2}{Lnb Nab} \frac{\text{Ch} \frac{-Wb}{Lnb} \sqrt{\frac{qV_{be}}{kT}} - 1}{\text{Sh} \frac{-Wb}{Lnb} \sqrt{\frac{qV_{be}}{kT}}} \quad \text{pour } V_{be} \gg kT/q$$

La densité de courant de génération dans la ZCE de la jonction base-collecteur est donnée, à partir de l'équation (71), par,

$$(162) \quad J_g = q N_t C \frac{ni}{2} W_{ZCEbc}(V_{cb})$$

Où $W_{ZCEbc}(V_{cb})$ est la largeur de la ZCE de la jonction base-collecteur, donnée par l'équation (72). A partir des équations (66) et (67) évaluées aux bords de la ZCE, les densités des courants de diffusion des électrons et des trous J_{dn} et J_{dp} , dans la jonction base-collecteur polarisée en inverse (avec $V_{cb} \gg kT/q$) s'écrivent, respectivement,

$$(163) \quad J_{dn} = -jn(xp) = \frac{q ni^2 D_{nb}}{Nab Lnb} \quad \text{pour } V_{cb} \gg kT/q.$$

$$(164) \quad J_{dp} = -jp(xn) = \frac{q ni^2 D_{pc}}{Ndc Lpc} \quad \text{pour } V_{cb} \gg kT/q$$

La densité du courant de fuite $J_g = J_{dp} + J_{dn} + J_g$ s'écrit donc,

$$(165) \quad J_f = q ni^2 \left(\frac{D_{pc}}{Ndc Lpc} + \frac{D_{nb}}{Nab Lnb} \right) + \frac{N_t C ni}{2} W_{ZCEbc}$$

3.5.1.2. Gain de courant

Le gain de courant β d'un transistor bipolaire est un des paramètres les plus importants des transistors bipolaires. Il est défini comme le rapport entre le courant de collecteur et le courant de base,

$$(166) \quad \frac{1}{\beta} = \frac{J_b}{J_c} = \frac{1}{1 + \frac{J_f}{J_{nc}}} \frac{J_{pe} + J_{rb} + J_{rd} - J_f}{J_{nc}} \cong \frac{J_f}{J_{nc}} \frac{J_{pe} + J_{rb} + J_{rd} - J_f}{J_{nc}}$$

L'injection des porteurs dans le collecteur, due au courant de fuite J_f , est très faible par rapport aux forts courants de collecteur J_{nc} , dès que V_{be} est supérieur à quelques kT/q . En revanche, comme le courant de la base est normalement très faible par rapport au courant du collecteur, l'injection du courant J_f dans la base n'est pas forcément négligeable. Si on suppose que $W_b \ll L_{nb}$, le gain de courant β s'écrit,

$$(167) \quad \frac{1}{\beta} = \frac{\mu_{pe}(T) N_{ab} W_b}{\mu_{nb}(T) N_{de} L_{pe}(T)} + \frac{1}{2} \frac{W_b}{L_{nb}(T)} + \frac{N_t \tau_c N_{ab} W_b W_{ZCEbe}}{2 D_{nb}(T) n_i(T)} e^{-\frac{qV_{be}}{2kT}} - \frac{J_f}{J_{nc}} \frac{1}{1 - \frac{J_f}{J_{nc}}}$$

Avec,

$$(168) \quad \frac{J_f}{J_{nc}} = \frac{N_t \tau_c N_{ab} W_b W_{ZCEbe}}{2 D_{nb}(T) n_i(T)} + \frac{\mu_{pc}(T) N_{ab} W_b}{\mu_{nb}(T) N_{dc} L_{pc}(T)} + \frac{W_b}{L_{nb}(T)} e^{-\frac{qV_{be}}{kT}}$$

Le premier terme dans (167) correspond au phénomène de transport des porteurs minoritaires dans la base (très étroite, plus faiblement dopée et de plus forte mobilité, par rapport à l'émetteur). Il est le terme dominant dans les conditions normales d'opération, et il varie faiblement avec la température en fonction de la différence entre les coefficients thermiques des mobilités et en fonction de $L_{pe}(T)$ (Dans le cas d'un émetteur très fin, $L_{pe}(T)$ est remplacé par l'épaisseur de l'émetteur W_e). Le deuxième terme prend en compte le phénomène de recombinaison des porteurs dans la base. Il est normalement négligeable dans le cas des transistors de base très étroite, et il augmente faiblement aussi avec la température. Le troisième terme introduit l'effet de recombinaison des porteurs dans la ZCE de la jonction base-émetteur. Il est responsable de la réduction de β dans le cas des faibles valeurs de I_c (i.e. faible V_{be}). L'influence sur β de ce troisième terme diminue très fortement au fur et à mesure que V_{be} ou la température ($n_i(T)$ varie fortement selon la température) augmentent. Il est pratiquement négligeable quand $V_{be} \gg 2kT/q$.

Transistor Bipolaire	Gain du courant β latéral ($A_e=3*0.8$)	Coefficient thermique XTB de β	Courant de collecteur pour $\beta/2$ [mA/ μ m]	Early Voltage [volts]
Vertical npn	70 à 160	1.8	0.3 à 0.8	20 à 32

Tableau 12. Caractéristiques du transistor bipolaire vertical de la technologie BiCMOS étudiée.

Le quatrième terme J_f/J_{nc} dans (167) correspond à la réduction du courant de base par injections d'un flux de porteurs provenant du collecteur (courant de fuite J_f). Il a sur β un effet contraire aux phénomènes de recombinaison précédemment mentionnés, mais on voit clairement dans (168) que le terme J_f/J_{nc} devient pratiquement négligeable quand $V_{be} \gg kT/q$. Le premier terme dans l'équation (168) prend en compte l'injection des trous dans la base, générés dans la ZCE de la jonction base-collecteur. L'importance de ce terme diminue fortement au fur et à mesure que la température augmente ($n_i(T)$ varie fortement selon la température). Le deuxième terme prend en compte les courants de fuite par diffusion à travers la jonction base-collecteur inversement polarisée. Il ressemble

au premier terme dans (167), mais il a sur β un effet contraire. Dans les conditions normales d'opération (i.e. $V_{be} \gg kT/q$, $V_{cb} \gg kT/q$, faible injection), le courant de fuite J_f n'a pas un fort impact sur le gain de courant β . Le premier terme dans (167) reste le terme dominant, dans ces conditions et même en haute température. Il donnera à β un coefficient thermique raisonnablement faible (voir Tableau 12), lequel est normalement modélisé dans SPICE par,

$$(169) \quad \beta(T) = \beta(T_0) \exp\left(\frac{-T}{T_0} \right)^{XTB}$$

De plus, la dispersion de la valeur de β étant normalement assez large (généralement de l'ordre de $\pm 30\%$), les variations de β selon la température n'affecteront pas beaucoup la performance des circuits à transistors bipolaires.

3.5.1.3. Transconductance

La transconductance gm est définie comme la variation du courant de collecteur I_c par rapport à la tension base-émetteur V_{be} ,

$$(170) \quad gm = \frac{fI_c}{fV_{be}} \cup Ae \frac{fJ_{nc}}{fV_{be}} = \frac{q I_c}{k T}$$

3.5.1.4. Modulation de la largeur effective de la base

La largeur effective de la base W_b , est une fonction de la largeur physique de la base W_{b0} et de la taille X_{pe} et X_{pc} des ZCE dans la base, correspondant respectivement aux jonctions base-émetteur et base-collecteur,

$$(171) \quad W_b = W_{b0} - [X_{pe}(V_{be}) + X_{pc}(V_{cb})] \cup W_{b0} - X_{pc}(V_{cb})$$

Dans les conditions normales d'opération $X_{pe} \ll X_{pc}$, parce que la jonction base-émetteur est polarisée en direct et la jonction base-collecteur en inverse, et car le collecteur et l'émetteur sont plus fortement dopés que la base. Considérant le cas des jonctions abruptes, X_{pc} est donné par,

$$(172) \quad X_{pc} = \sqrt{\frac{2 \epsilon_{sc} \frac{N_{dc}}{q} \frac{1}{N_{ab} + N_{dc}} \sqrt{V_{bi} - 2 \frac{k T}{q}} + V_{cb}}{q}}$$

Cette modulation de la largeur effective de la base, dit effet Early, génère une variation du courant de collecteur en fonction de la valeur de V_{cb} . En négligeant les courants de fuite dans le collecteur et en supposant que $W_b \ll L_{nb}$, la densité du courant du collecteur J_c s'écrit,

$$(173) \quad J_c = \frac{q D_{nb} n_i^2}{W_b(V_{cb}) N_{ab}} \exp\left(\frac{q V_{be}}{k T}\right)$$

J_c dépend explicitement de la tension de polarisation V_{be} , et implicitement de V_{cb} , à travers le terme W_b . Les variations de J_c en fonction de V_{cb} , avec V_{be} constante, s'écrivent,

$$(174) \quad \frac{fJ_c}{fV_{cb}} = - \frac{J_c}{W_b} \frac{fW_b}{fV_{cb}} = \frac{J_c}{W_b} \frac{fX_{pc}}{fV_{cb}} = \frac{J_c}{W_b} \frac{\epsilon_{sc} \frac{N_{dc}}{q} \frac{1}{N_{ab} + N_{dc}}}{X_{pc}}$$

L'équation (174) décrit la conductance du collecteur pour des petits signaux, paramètre qui caractérise l'impédance de sortie des étapes d'amplification d'émetteur commun.

La tension de Early V_a est définie de la manière suivante,

$$(175) \quad V_a = J_c \frac{fV_{cb}}{fJ_c} = \frac{(Wb_0 - X_{pc}) \gamma X_{pc}}{\frac{\epsilon_{sc}}{q} \gamma \frac{N_{dc}}{N_{ab}} \gamma \frac{1}{N_{ab} + N_{dc}}}$$

Considérant le cas théorique où $N_{dc} \gg N_{ab}$ et $Wb_0 \gg X_{pc}$, la tension de Early V_a pourrait s'écrire de la manière simplifiée suivante,

$$(176) \quad V_a \approx \frac{q \gamma N_{ab} Wb_0}{\epsilon_{sc}} \gamma X_{pc}(V_{ce}, T)$$

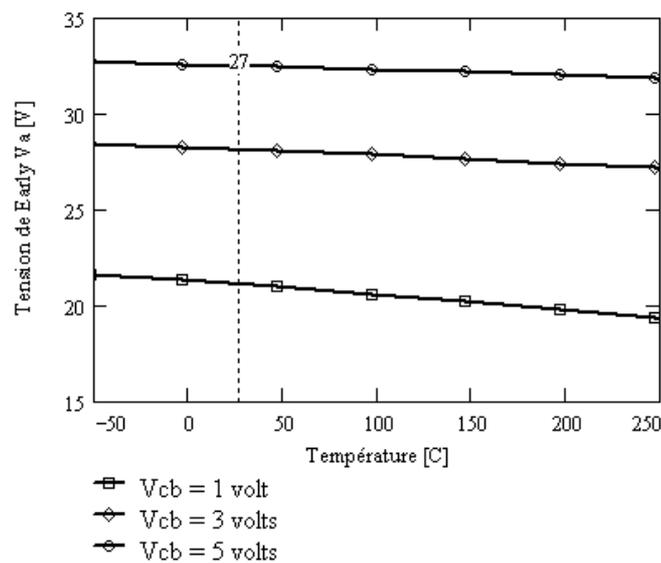


Figure 61. Tension de Early en fonction de V_{cb} et la température.

Dans la Figure 61 on peut apprécier l'évolution de la tension de Early selon l'équation plus générale (175) en fonction de la température et la tension de polarisation V_{cb} , dans le cas théorique d'un transistor bipolaire où $N_{dc} \approx 100 * N_{ab} \approx 700 * 10^{22} m^{-3}$ et $Wb_0 \approx 1.25 \mu m$. Les variations de V_a selon la température indiquent une dégradation relativement faible de l'impédance de collecteur du transistor bipolaire (dV_{cb}/dI_c) en fonction de la température, ce qui ne représente pourtant pas une limitation importante pour l'utilisation des ces transistors en haute température.

3.5.1.5. Courant de fuite et conductance dans la jonction collecteur-substrat

La jonction d'isolation entre le collecteur et le substrat (voir Figure 60) est comparativement très large. Le courant de fuite I_{fcs} , donné par l'équation (73) multipliée par la surface A_{cs} de la jonction mentionnée, entre en concurrence avec le courant du collecteur de l'expression (151) au fur et à mesure que la température augmente. La couche enterrée fortement dopée, construite pour réduire la résistance du collecteur, peut aussi aider à la réduction du courant de fuite. Cependant, comme dans une jonction dissymétrique la région la moins dopée domine dans l'expression (73), si le dopage du substrat reste constant la réduction du courant de fuite due à la couche enterrée sera très limitée (voir dans le Tableau 6 les valeurs des courants de fuite à température ambiante dans le caisson n - et la

couche enterrée $n+$). Le courant de fuite dans cette jonction représente ainsi une forte limitation pour l'utilisation en haute température des transistors bipolaires verticaux.

Dans la plage de température où la génération de porteurs (voir l'équation (71)) dans la ZCE est dominante dans l'expression (73) de la densité du courant de fuite, la dépendance de ce courant $Ifcs_g$ par rapport à la tension de polarisation Vcs permet de définir la conductance $Gfcs_g$ suivante, pour des petits signaux autour d'une valeur $Vcsq$,

$$(177) \quad Gfcs_g = \left. \frac{fIfcs_g}{fVcs} \right|_{Vcsq} = q \cdot Acs \cdot Nt \cdot C \cdot \frac{ni}{2} \cdot \left. \frac{fW(Vcs)}{fVcs} \right|_{Vcsq} = Acs \cdot Nt \cdot C \cdot \frac{ni}{2} \cdot \frac{\epsilon sc \cdot \frac{Na + Nd}{Na \cdot Nd}}{W(Vcsq)}$$

Dans la plage de température où la diffusion de porteurs (voir l'équation (68)) est dominante dans l'expression (73) de la densité du courant de fuite, la conductance pour les petits signaux ne résulte plus de la modulation de la densité du courant de fuite, mais elle est générée par une modulation de la taille de la surface à la limite de la ZCE, du côté le moins dopé de la jonction [34].

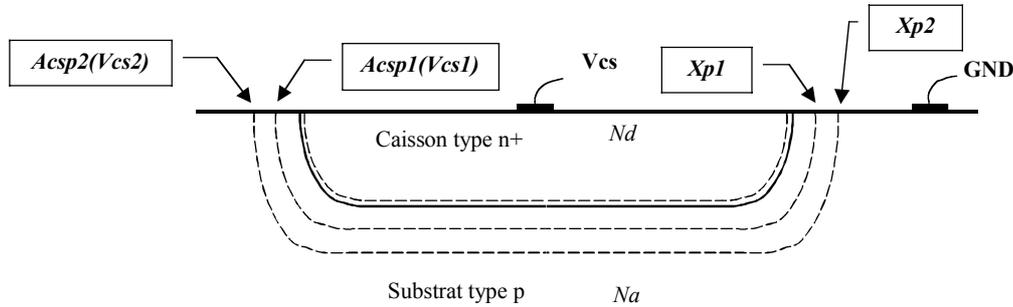


Figure 62. Variation de la surface $Acsp$ à la limite de la ZCE en fonction de la polarisation.

Dans une jonction dissymétrique, comme celle de la Figure 62, le courant de fuite de diffusion est dominé par la diffusion des électrons du substrat vers le collecteur, décrite par l'équation (66) évaluée dans la limite xp (expression similaire à l'équation (172)) de la ZCE, i.e. du côté le moins dopé de la jonction (le substrat),

$$(178) \quad Ifcs_d \approx - \frac{q \cdot Acs \cdot ni^2 \cdot Dn}{Na \cdot Ln} \cdot \left. e^{-\frac{qVcs}{kT}} \cdot \left(1 - \sqrt{e^{\frac{(x-xp)}{Ln}}} \right) \right|_{x=xp} = \frac{q \cdot Acsp(Vcs) \cdot ni^2 \cdot Dn}{Na \cdot Ln}$$

Dans une jonction pn tridimensionnelle, comme celles de la Figure 62 et de la Figure 60, la surface $Acsp$ est une fonction de la tension de polarisation Vcs . Il en résulte que le courant de fuite de diffusion $Ifcs_d$ est modulé par Vcs à travers $Acsp(Vcs)$. On définit la conductance $Gfcs_d$, pour des petits signaux autour d'une valeur $Vcsq$, suivante,

$$(179) \quad Gfcs_d = \frac{fIfcs_d}{fVcs} = \frac{q \cdot ni^2 \cdot Dn}{Na \cdot Ln} \cdot \frac{fAcsp}{fVcs} \cdot \frac{fxp}{fVcs} = \frac{ni^2 \cdot Dn}{Na^2 \cdot Ln} \cdot \frac{\epsilon sc \cdot Nd}{(Na + Nd)} \cdot \frac{1}{xp(Vcsq)} \cdot \frac{fAcsp}{fxp}$$

Où $dAcsp/dxp$ est un facteur qui dépend uniquement de la géométrie de la jonction.

Dans le cas d'une jonction dissymétrique, la conductance totale due aux courants de fuite dans la jonction collecteur-substrat est donnée par,

$$(180) \quad Gfcs = \frac{\epsilon sc}{Na \cdot xp(Vcsq)} \cdot ni(T) \cdot \frac{Acs \cdot Nt \cdot C}{2} + \frac{ni(T)^2 \cdot Dn}{Na \cdot Ln} \cdot \frac{fAcsp}{fxp}$$

Elle augmente fortement avec la température, d'une manière très similaire au courant de fuite qui la génère. Cette conductance pourrait dégrader considérablement la performance des transistors bipolaires, en haute température.

3.5.2. Transistors bipolaires latéraux CLBT

Il y a au moins, deux manières de réaliser des transistors bipolaires dans les technologies CMOS. La première possibilité est le Transistor Bipolaire Vertical CMOS (CVBT) où le substrat est utilisé comme collecteur, un caisson séparé formant la base, et une diffusion dopée de même type que le substrat formant l'émetteur. Mais les applications qui utilisent des CVBTs seront limitées à des configurations en collecteur commun, puisque les CVBTs ont tous comme collecteur commun le substrat. La deuxième possibilité est celle de faire une seconde diffusion dans le caisson et de l'utiliser comme collecteur. De cette façon, le CLBT, combiné avec le CVBT, est défini par la structure de double collecteur de la Figure 63 [15] [64]. La grille (G) du CLBT type PNP, dans la Figure 63, doit être polarisée positivement par rapport au potentiel du caisson de type n , de façon à accumuler des porteurs majoritaires à l'interface $Si-SiO_2$, éviter la conduction au niveau de la surface et ainsi obtenir une conduction bipolaire dans le volume. Quand le potentiel entre la grille et le caisson de type n devient positif et supérieur au potentiel de bandes plates V_{fb} , une accumulation d'électrons (porteurs majoritaires) a lieu à l'interface $Si-SiO_2$ du CLBT. Si la jonction entre la source (émetteur) et le caisson est polarisée en direct, des trous (porteurs minoritaires) seront injectés dans le caisson de type n (base), où ils vont diffuser vers le collecteur latéral et aussi vers le collecteur vertical.

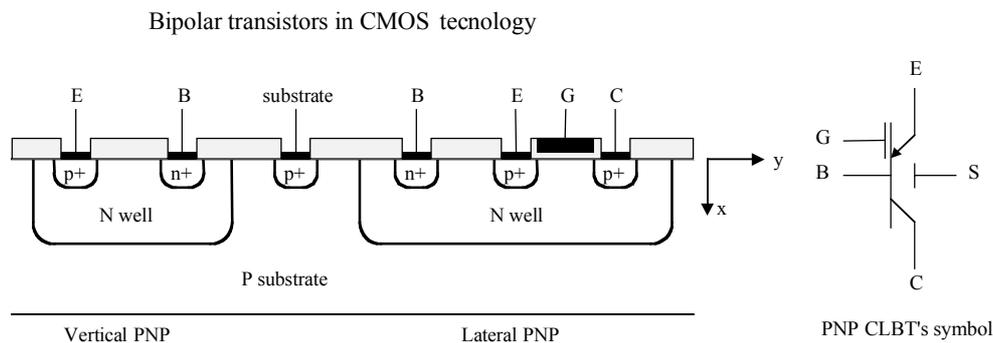


Figure 63. Structure des transistors bipolaires dans les technologies CMOS.

Quand le potentiel entre la grille et le caisson est assez important (quelques volts), la relation $I_d(V_{source_caisson})$ où $I_c(V_{be})$ devient indépendante de la polarisation de grille V_g , car les porteurs sont poussés vers le volume, loin de la surface, générant ainsi un courant purement bipolaire. Comme le courant est poussé loin de la surface, les effets de piégeage de porteurs dans les centres profonds vont être réduits et on obtiendra donc une plus faible composante du bruit $1/f$ et de meilleures caractéristiques de conduction, par rapport aux dispositifs bipolaires latéraux normaux.

La structure du CLBT étant similaire à celle des transistors bipolaires courants, à l'exception de la présence de la grille, l'influence de la polarisation de cette grille, dans le comportement du CLBT en haute température, doit être étudiée.

3.5.2.1. Modèle du courant du collecteur latéral

Un modèle analytique précis du CLBT n'est pas facile à développer car le transport de charges s'effectue dans deux régions différentes, sous la grille et sous l'émetteur, ceci suivant un chemin tridimensionnel. De plus, la distance entre les deux diffusions P+ (largeur de base), la concentration de

dopage dans la base, et la mobilité, ne sont pas constantes mais elles varient avec la profondeur. Les modèles du courant vertical et latéral du CLBT utilisés, ont été développés par [15] à partir d'une structure idéalisée du transistor, afin d'obtenir un modèle analytique unidimensionnel simplifié, mais suffisamment précis, du CLBT.

Le courant du collecteur latéral du CLBT type PNP est donné par,

$$(181) \quad I_c = I_{cfb} \cdot fg$$

Où I_{cfb} est le courant du collecteur quand le potentiel dans le caisson se trouve dans la condition de bandes plates $\Phi_s=0$. Autrement dit, I_{cfb} est le courant de collecteur d'un transistor bipolaire normal (voir l'équation (173)) où la grille n'existe pas et la surface base-émetteur $A_{be}=\omega_o \cdot X_j$ est définie par la profondeur X_j de la jonction entre la diffusion p^+ et le caisson n^- et le périmètre (dans le cas des structures d'émetteur carrées ou circulaires) de l'émetteur ω_o ,

$$(182) \quad I_{cfb} = q \cdot ni^2 \cdot \frac{\bar{D}_{pb} \cdot \omega_o \cdot X_j}{W_b \cdot N_{db}} \cdot e^{\frac{V_{be}}{V_t}}$$

Où $V_t=kT/q$. Le facteur fg est la fonction de grille, laquelle dépend uniquement des paramètres technologiques et de la polarisation de la grille V_g . fg décrit l'influence de la grille sur le courant de collecteur [15].

$$(183) \quad fg = \frac{1}{X_j} \int_0^{X_j} e^{-\frac{\Phi(x)}{V_t}} dx$$

Où $\Phi(x)$ est le potentiel du caisson, défini dans la Figure 36. La fonction de grille fg dépend de la température à travers V_t et $\Phi(x)$. Les valeurs de la fonction $\Phi(x)$ ont été déterminées par résolution numérique de l'équation différentielle suivante, qui résulte de l'application de l'équation de Poisson à la structure MOS de la Figure 36, en régime d'accumulation,

$$(184) \quad \frac{d}{dx} \left(\frac{d\Phi(x)}{dx} \right) = \frac{1}{L_d} \left[e^{\frac{\Phi(x)}{V_t}} - \frac{\ddot{\Phi}(x)}{V_t} - 1 \right] \quad \text{avec } \Phi(0)=\Phi_s \text{ comme condition initiale.}$$

Où L_d est longueur de Debye des électrons.

Le potentiel de surface Φ_s est déterminé, à partir des équations (85) et (93) dans le régime d'accumulation, par la relation non-linéaire suivante,

$$(185) \quad \frac{V_g - V_{fb}}{V_t} = \frac{\Phi_s}{V_t} + \frac{\mathcal{P}}{\sqrt{V_t}} \sqrt{e^{\frac{\Phi_s}{V_t}} - \frac{\ddot{\Phi}_s}{V_t} - 1}$$

où V_g est le potentiel entre la grille et le caisson.

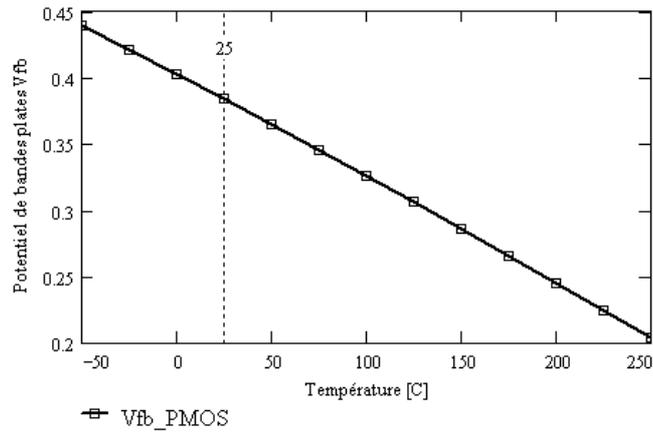


Figure 64. Potentiel de bandes plates d'une capacité PMOS (Poly-Si type n - Si-SiO₂ - SC type n).

Le potentiel de bandes plates V_{fb} dépend aussi de la température. Dans le cas du CLBT type PNP ce potentiel est celui de la capacité PMOS, nommé donc V_{fbP} . La valeur de V_{fbP} est donnée, à partir de l'expression (87) et dans le cas d'un semiconducteur du type n , par l'équation suivante,

$$(186) \quad V_{fbP} = -\frac{E_g}{2 \cdot q} - \Phi_{fn} - \frac{Q_{se}}{C_{ox}}$$

Dans la Figure 64 on peut apprécier l'évolution selon la température, du potentiel de bandes plates V_{fbP} du CLBT type PNP.

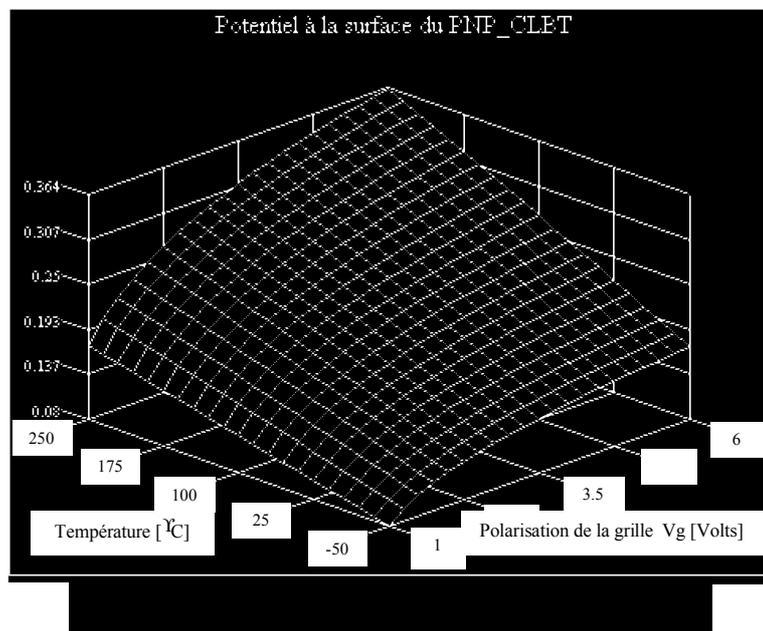


Figure 65. Potentiel Φ_s à l'interface Si-SiO₂ de la capacité PMOS du PNP_CLBT.

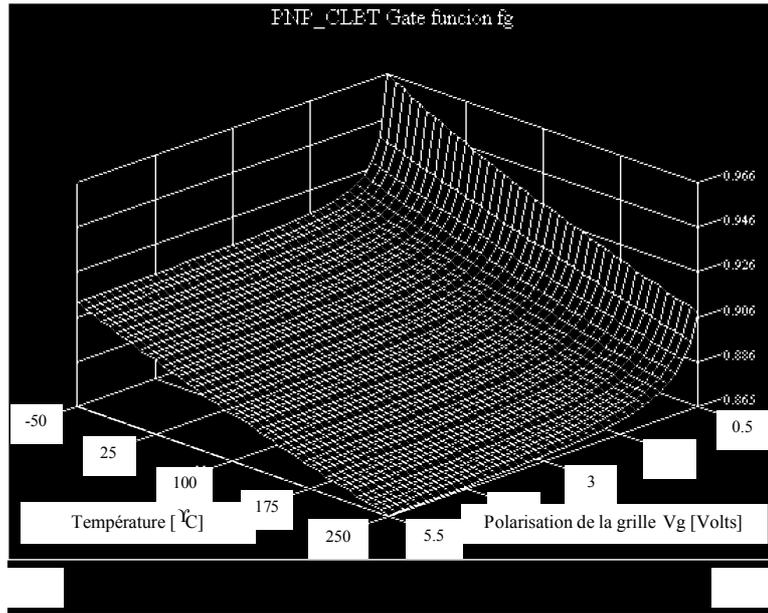


Figure 66. Fonction de grille fg du PNP_CLBT en fonction de Vg et la température.

La concentration des charges Q_{se} à l'interface a été approximativement déterminée à partir de la valeur de la tension de seuil $V_{thP}(Ta) = -0.755$ volts des transistors PMOS à température ambiante (mesurée à partir d'un transistor PMOS avec $W/L=20/0.8$ [μm]), et la relation suivante, qui découle de l'expression de la tension de seuil (111) des transistors PMOS avec $V_s=0$,

$$(187) \quad V_{fbP}(Ta) = V_{thP}(Ta) - 2\phi_{fn}(Ta) + \gamma \sqrt{2\phi_{fn}(Ta)} = -\frac{E_g(Ta)}{2q} - \phi_{fn}(Ta) - \frac{Q_{se}}{C_{ox}}$$

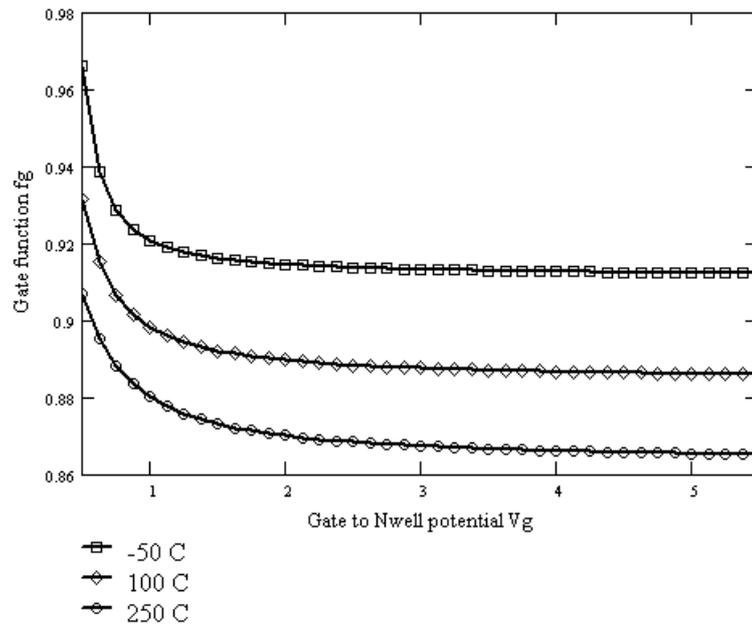


Figure 67. Fonction de grille fg du PNP_CLBT en fonction de Vg.

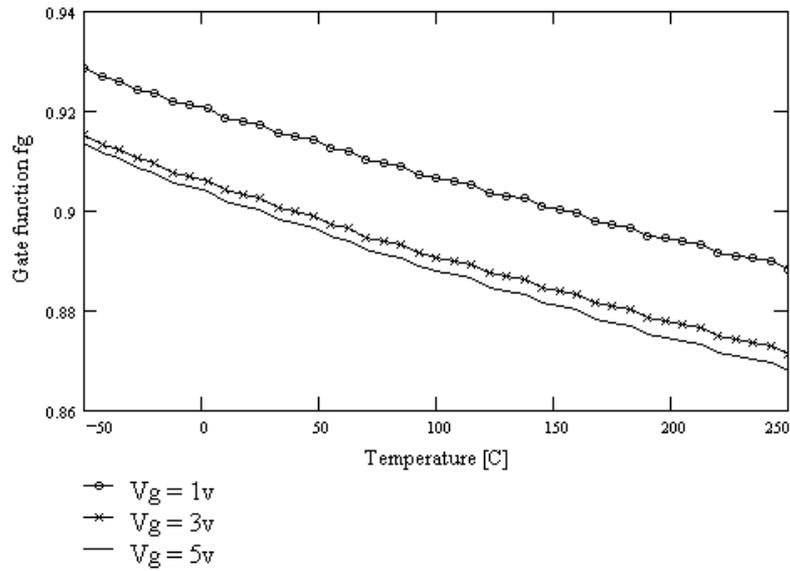


Figure 68. Fonction de grille fg du PNP_CLBT en fonction de la température.

A partir des équations (185), (186) et (187) le potentiel de surface Φ_s de la capacité PMOS du PNP_CLBT a été calculé (voir Figure 65) pour des variations de la tension de polarisation V_g de 1 à 6 volts, et de la température de -50 à 250°C. A partir des données de la Figure 65 et des équations (183) et (184), les valeurs de la fonction de grille fg ont été obtenues pour des variations de la tension de polarisation V_g de 0.5 à 5.5 volts, et de la température de -50 à 250°C. (voir la Figure 66). Dans toutes les températures considérées, la valeur de fg varie beaucoup initialement, selon V_g (voir la Figure 67), et se stabilise à partir de $V_g=1$ volt environ. Les variations de fg selon la température sont assez linéaires pour toutes les valeurs de V_g considérées (voir la Figure 68).

Dans la Figure 69, les variations maximales (i.e. dans toute la plage de température) de fg selon la température pour chaque valeur de V_g sont mises en évidence. A partir de la Figure 67 et la Figure 69, on peut déterminer la polarisation ($V_g \geq 0.75$ volt) de la capacité PMOS du PNP_CLBT, qui rend moins sensible la fonction de grille fg , par rapport à la température et à la tension de polarisation V_g , et par conséquent, qui rend également moins sensible le courant du collecteur latéral du transistor.

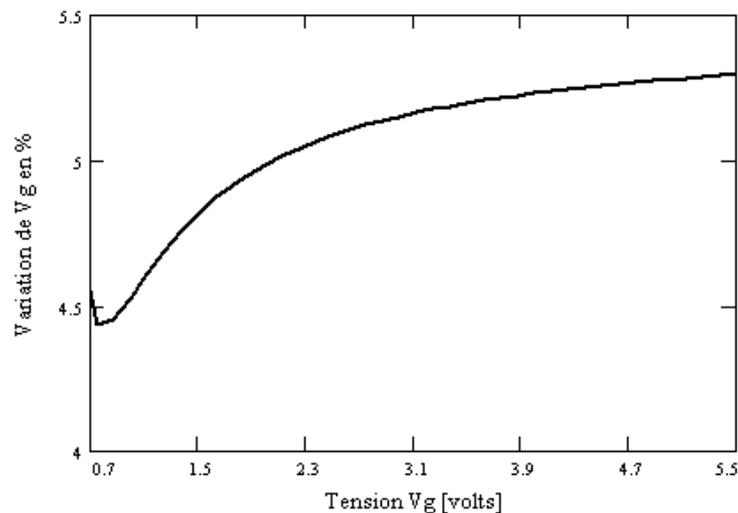


Figure 69. Variation maximale de fg entre -50 et 250°C, en fonction de Vg.

3.5.2.1.a. Fonction analytique approximative de fg

Comme une solution analytique à l'intégrale dans l'équation (183) n'a pas été trouvée, la fonction analytique approximative de fg suivante, a été proposée par [15],

$$(188) \quad fg = 1 - \frac{Ld}{fp \cdot Xj} \cdot \left(1 - \frac{1}{1 + \frac{Vg - Vfb}{\gamma \sqrt{2} \cdot Vt}} \right)^{-2\delta p}$$

Où δp est un paramètre d'ajustement. Il est intéressant de remarquer que fg dépend directement du rapport Ld/Xj , donc la fonction de grille s'approche de l'unité au fur et à mesure que la densité des dopants Nd et la profondeur de jonction Xj augmentent. Il est aussi possible de constater que l'influence de $Vfb(T)$ et $Vt(T)$ devient très faible quand Vg s'approche de la valeur de Vfb, ce qui a pour effet de réduire la sensibilité de fg par rapport à la température. La fonction approximative de fg a été évaluée, en utilisant les paramètres des technologies étudiées, et elle a été comparée avec les valeurs de fg de la Figure 66. Ainsi, la valeur du paramètre d'ajustement $\delta p = 0.6043$ a été obtenue, donnant une erreur rms du 0.2% entre la fonction approximative (188) et la résolution numérique de l'équation (183), dans toute la plage de température et de tension de polarisation de la grille considérée.

3.5.2.2. Courant du collecteur vertical

Le courant du collecteur vertical peut être calculé en appliquant directement la théorie usuelle des transistors bipolaires verticaux (voir la section 3.5.1), plus l'addition des courants de fuite de la jonction entre le caisson n- et le substrat p-,

$$(189) \quad I_{cv} = q \cdot n_i^2 \cdot \frac{D_{pb}}{W_{bv}} \cdot \frac{A_{ev}}{N_{db}} \cdot e^{\frac{V_{be}}{V_t}} + A_{N_{well-substratP}} \cdot J_{N_{well-substratP}}$$

Où A_{ev} est la surface verticale de la jonction base-émetteur (défini dans la Figure 59), et W_{bv} est la largeur de la base verticale, définie par la différence entre la profondeur de jonction du caisson et celle de la jonction base-émetteur. La jonction entre la base et le substrat, dans la Figure 63, est normalement très large et faiblement dopée. Il en résulte que le courant de fuite ($J_{N_{w-sP}}$ est donnée par l'équation (73)) peut devenir une composante importante du courant du collecteur vertical, en haute température.

La largeur de la base verticale est $W_{bv} \approx 3 \text{ à } 4 \mu m$, dans le cas des technologies étudiées, et la largeur de la base latérale peut être aussi petite que $0.8 \mu m$, ce qui a pour effet d'augmenter le rapport entre le courant latéral et le courant vertical. De plus, dans la technologie BiCMOS, le caisson "Nwell" a une couche enterrée très fortement dopée, située dans la base du transistor vertical parasite. Ainsi, la différence entre le dopage de la base Nd du transistor latéral et du transistor vertical mentionné, contribue fortement à augmenter le rapport entre le courant du collecteur latéral et le courant du collecteur vertical (voir les gains en courant des transistors verticaux dans le Tableau 13).

3.5.2.3. Gain de courant et conductance des collecteurs latéral et vertical

Le gain en courant vertical β_v est obtenu de la même manière que dans la section 3.5.1, les densités de courant dans le collecteur et la base étant approximativement données par,

$$(190) \quad J_{cv} \cup q \cdot n_i^2 \cdot \frac{\bar{D}_{pb}}{W_{bv} \cdot N_{db}} \cdot e^{\frac{V_{be}}{V_T}} + \frac{A_{Nwell-substratP}}{A_e} \cdot J_{f_{Nwell-substratP}}$$

$$(191) \quad J_b \cup J_{pe} + J_{rd} + J_{rb} - \frac{A_{Nwell-substratP}}{A_e} \cdot J_{f_{Nwell-substratP}}$$

On néglige les courants de fuite dans la jonction entre le collecteur latéral et la base, parce que sa surface est très inférieure à celle de la jonction entre le collecteur vertical (i.e. le substrat) et la base. Il en résulte que l'influence du quatrième terme dans l'équation (167) est amplifiée, approximativement, par le rapport des surfaces $A_{Nwell-substratP}/A_e$.

Transistor Bipolaire	Gain en courant β latéral ($A_{el}=3,6 \cdot 4 \cdot X_j$, $I_e=1\mu A$)	Gain en courant β latéral ($A_{el}=3,6 \cdot 4 \cdot X_j$, $I_e=10\mu A$)	Early Voltage de I_{cl} [volts]	Early Voltage de I_{cv} [volts]	Gain de courant β vertical parasite ($A_{ev}=3,6 \cdot 3,6$, $I_e=1\mu A$)
Technologie CMOS	100 à 360	50 à 190	~ 10	~ 150	~ 35
Technologie BiCMOS	110 à 310	70 à 190	~ 10	-	< 2

Tableau 13. Caractéristiques des transistors bipolaires latéraux des technologies étudiées.

En ce qui concerne le gain en courant latéral β_l , la densité du courant dans le collecteur est approximativement donnée par,

$$(192) \quad J_{cl} = \frac{I_{c_{fb}} \cdot f_g}{A_e} + \frac{A_{Nwell-Coll_lat}}{A_e} \cdot J_{f_{Nwell-Coll_lat}}$$

Le gain en courant latéral β_l sera donc affecté par la fonction de grille f_g ,

$$(193) \quad \beta_l \cong \beta_{fb} \cdot f_g$$

Où β_{fb} est le gain en courant du transistor dans le régime des bandes plates. Comme le courant de la base reste celui de l'expression (191), il en résulte que l'influence du quatrième terme dans l'équation (167), associé maintenant à β_{fb} , est aussi amplifiée, approximativement, par le rapport des surfaces $A_{Nwell-substratP}/A_e$. A travers la Figure 66 on peut apprécier comment la fonction de grille f_g module la valeur du gain en courant latéral, selon la température et la polarisation de la grille.

La transconductance g_m est définie comme la variation du courant du collecteur latéral I_{cl} par rapport à la tension base-émetteur V_{be} ,

$$(194) \quad g_m = \frac{f_{I_{cl}}}{f_{V_{be}}} \cup \frac{q \cdot I_{cl}}{k \cdot T}$$

Les modulations de la largeur effective de la base latérale et de la base verticale (effet *Early*) répondent au même type d'équations que celles développées dans la section 3.5.1.4. Les différences entre la largeur physique de la base verticale et de la base latérale expliquent les valeurs de tension de *Early* très différentes montrées dans le Tableau 13.

3.5.2.4. Courants de fuite dans le collecteur latéral

Une des différences les plus intéressantes, entre la structure des transistors bipolaires verticaux, dans la Figure 60, et latéraux, dans la Figure 63, est l'absence d'isolation par jonction inversement polarisée entre le collecteur latéral du CLBT et le substrat. C'est pourquoi, les larges courants de fuite, ainsi que leurs conductances associées, ne font pas partie du courant du collecteur latéral, comme c'est le cas du courant du collecteur des transistors verticaux (voir la section 3.5.1.5). Il en résulte que les limitations à l'utilisation des transistors bipolaires verticaux en haute température mentionnées dans la section 3.5.1.5, ne s'appliquent pas aux transistors latéraux. Une meilleure tenue en haute température est ainsi attendue.

3.6. Conclusions

Les résistances sur silicium monocristallin, dans les technologies étudiées, souffrent de plusieurs désavantages : les importants courants de fuite en haute température, les non-linéarités et capacités parasites considérables, et les variations de la résistivité selon la température, ce qui les rend peu adaptées aux applications en haute température. En revanche, les résistances sur silicium polycristallin offrent de meilleures propriétés pour la haute température, principalement à cause de leurs courants de fuite et de leurs non-linéarités négligeables, jusqu'à 250°C environ.

Les capacités Double-Poly ont aussi une très bonne performance à haute température, du fait de leurs très faibles courants de fuite et non-linéarités. Des valeurs de capacité par unité de surface plus élevées peuvent être obtenues avec les capacités PMOS en accumulation. Mais l'influence des ces non-linéarités, des variations de la capacité selon la température et des courants de fuite dans le caisson doit être étudiée en détail pour chaque application.

La tension de seuil V_t , le facteur de gain KP et la tension d'Early V_a des transistors MOS diminuent progressivement avec la température. En revanche, les courants de fuite dans le drain et la source, étant sensiblement plus faibles dans le cas des PMOS que dans celui des NMOS, augmentent de façon plus importante avec la température. Tout cela concerne plutôt la robustesse des circuits à la haute température et ne pose pas de problèmes majeurs de fiabilité. Cependant, l'injection de porteurs chauds dans l'oxyde de grille, mentionnée dans le chapitre 2, est responsable d'un vieillissement, thermiquement accéléré, des transistors MOS. L'utilisation des transistors PMOS (moins sensibles aux porteurs chauds), dans les étages plus critiques des circuits intégrés, reste un bon moyen d'améliorer la fiabilité des circuits intégrés en haute température (voir section 5.1.1.1.a.b).

De même, les paramètres des transistors bipolaires (tels que le gain de courant β , la tension d'Early V_a et la fonction de grille fg des transistors latéraux) varient progressivement selon la température, à l'exception des courants de fuite collecteur-substrat des transistors verticaux. En revanche, le transistor bipolaire latéral présente une structure beaucoup moins sensible aux courants de fuite à haute température.

L'optimisation de la structure des dispositifs et des circuits intégrés, selon les critères abordés dans les deux chapitres précédents, rend possible la conception de circuits intégrés analogiques en haute température avec une robustesse et une fiabilité sensiblement améliorées, permettant ainsi l'élargissement de la plage de température d'utilisation des technologies CMOS et BiCMOS standards.

4. Techniques d'endurcissement pour la haute température

Puisque la température agit partout dans un circuit intégré et qu'elle affecte une grande partie des paramètres des composants actifs et passifs intégrés, ainsi que leur isolation, la robustesse aux hautes températures d'un circuit intégré doit être étudiée à tous les niveaux, à partir du niveau système jusqu'au niveau d'implémentation. Cela concerne, le choix de la technologie, l'architecture du circuit, la conception et l'implémentation de chaque bloc fonctionnel du système, et les outils de simulation.

Les technologies CMOS et BiCMOS actuelles présentent quelques caractéristiques intéressantes, vis à vis de la tenue en température des circuits (voir le chapitre 2), tels que des courants de fuite pratiquement négligeables, même en haute température, à travers les diélectriques qui forment la grille des transistors MOS et qui isolent les résistances en polysilicium et les capacités Double-Poly. Cependant, les courants de fuite dans les diffusions du drain et de la source des transistors MOS, dans le collecteur des transistors bipolaires verticaux, et dans d'autres jonctions polarisées en inverse, ne sont pas négligeables en haute température. Ces courants de fuite, étant le critère le plus important pour le choix de la technologie, dépendent de la taille des jonctions, de la concentration des dopants du côté le moins dopé de chaque jonction, et éventuellement d'autres paramètres additionnels tels que la tension de seuil des transistors MOS, etc. Ainsi, les performances en température de ces technologies tendent à s'améliorer dans le futur, en terme général, car l'augmentation de la densité d'intégration entraîne une augmentation des concentrations des dopants et une réduction de la surface des jonctions d'isolation.

4.1. Techniques de conception

La conception d'un circuit intégré analogique est fondée sur les spécifications électriques de l'application, les caractéristiques et les paramètres de la technologie choisie et les contraintes imposées par l'environnement. Ces trois conditions, étant fortement liés à chaque application en particulier, varient selon les cas à étudier, ne permettant pas de faire une standardisation et une automatisation des procédures de conception. Les techniques de conception concernent donc les solutions à trouver, en termes d'architecture du circuit, de définition et d'implémentation de chaque bloc fonctionnel, pour remplir les conditions précédemment mentionnées.

4.1.1. Tolérances des paramètres des dispositifs

Les procédés de fabrication des circuits intégrés dans les technologies planaires se caractérisent par l'accumulation d'une série de traitements (oxydations, déposition des couches minces, gravures, implantations, diffusions, traitements thermiques, etc.) appliqués à la surface du substrat semiconducteur, de manière plus ou moins homogène. La reproductibilité de certains paramètres contrôlés (i.e. la capacité par unité de surface, la résistivité des couches minces, le facteur de gain des transistors MOS, etc.) en différents points de la surface d'une ou différentes tranches, étant une propriété importante des technologies des circuits intégrés, dépend de l'homogénéité des procédés de fabrication. Cette reproductibilité est exprimée en termes de tolérances acceptées des paramètres contrôlés d'une technologie de circuits intégrés (voir Tableau 14). Ces tolérances, parfois très larges, conditionnent la conception des circuits intégrés.

Il existe aussi, dans une technologie de circuits intégrés, des paramètres non-contrôlés; c'est à dire, les paramètres dont le fabricant ne spécifie pas de valeurs minimale et maximale. De plus, la plupart des paramètres varient aussi avec la température (voir les chapitres 2 et 3), rendant encore plus larges

les "tolérances effectives" des paramètres des dispositifs, pour une plage de température donnée. Il en résulte que la performance ou la précision d'un circuit analogique ne peut pas être fondée sur la précision absolue des paramètres des procédés de fabrication.

Paramètres Contrôlés		Min	Max	Tolérances [%]
Résistances [Ohms/carré]	Caisson n-	3000	4500	40
	Diffusion n+	19	28	38
	Diffusion p+	35	51	37
	Base p-	3800	5000	27
	Poly-Si 1 (région n+)	18	26	36
	Poly-Si 2	55	80	37
	HR-Poly-Si	900	1500	50
Capacités [fF/μm ²]	Cox	2.03	2.3	13
	Double-Poly	1.65	1.92	15
Transistors MOS	Tension de seuil des NMOS [Volts]	0.62	0.87	34
	Tension de seuil des PMOS [Volts]	-0.68	-0.9	28
	Facteur γ des NMOS [Volts ^{1/2}]	0.65	0.83	24
	Facteur γ des PMOS [Volts ^{1/2}]	0.4	0.5	22
	Facteur de Gain KPN [μ A/V ²]	85	115	30
	Facteur de Gain KPP [μ A/V ²]	30	40	29
Trans. bipolaires verticaux	Courant de saturation I_s [aA/μm]	0.4	2.2	138
	Gain en courant β	70	160	78
	Early voltage [Volts]	20	32	46
Trans. bipolaires latéraux	Gain en courant β	110	310	95

Tableau 14. Tolérances de quelques paramètres contrôlés de la technologie BiCMOS étudiée.

En revanche, les variations locales (i.e. distances typiquement inférieures à 100 μm, sur la surface du substrat) des paramètres contrôlés sont beaucoup plus faibles. Les gradients locaux des paramètres étant pratiquement constants, des degrés de similitude ("*matching*") entre composants de structure identique, meilleurs que 1% peuvent être obtenus à travers des techniques d'implémentation (voir section 4.2.1.1).

4.1.2. Dégradation des composants en fonction de la température

La dégradation des propriétés et des paramètres, à haute température, des matériaux et composants correspondant aux technologies CMOS et BiCMOS, a été étudiée (voir les chapitres 2 et 3) afin de pouvoir prédire et réduire son impact sur la performance des circuits analogiques en haute température. Certains paramètres et propriétés se dégradent progressivement en fonction de la température, rendant plus larges les tolérances à considérer, lors de la conception d'un circuit intégré analogique. D'autres paramètres se dégradent plus dramatiquement, limitant la plage de température d'utilisation de certains composants et de certaines manières d'implémenter certains dispositifs.

- Lignes d'alimentation et d'interconnexion intégrées : La réduction de la durée de vie (*Mean Time to Failure MTF*) des conducteurs, due à l'électromigration (donnée par l'équation (42)), dépend de la température et de la densité maximale de courant dans le conducteur. L'équation (

43) permet de calculer la densité du courant équivalente à respecter, à la température maximale de la plage considérée, afin d'éviter le raccourcissement de la durée de vie du circuit dû au phénomène d'électromigration. Le besoin d'une bonne tenue à la corrosion, dans le cas des applications des circuits intégrés en haute température, exige leur encapsulation étanche sous atmosphère contrôlée (azote, gaz inerte), afin d'éviter toute corrosion humide, dans des boîtiers (céramiques ou autres) résistants à la température maximale considérée.

- ∅ Isolation par diélectrique : L'injection des porteurs chauds dans l'oxyde de grille d'un transistor MOS, générés thermiquement dans le substrat, augmente avec la température. Cette injection peut être réduite en diminuant la surface de la grille, en réduisant la longueur L du canal pour un rapport W/L donné. L'injection des porteurs chauds provenant du canal peut être réduite en diminuant la valeur du champ électrique longitudinale, autrement dit en augmentant la longueur du canal L . Du fait que ce dernier type d'injection des porteurs dans la grille diminue avec la température, le compromis à trouver entre les deux critères mentionnés doit plutôt favoriser la réduction de l'injection de porteurs thermiquement générés. Les injections des porteurs mentionnées affectent plus fortement les transistors NMOS que les transistors PMOS, car la mobilité des électrons, dans le silicium et dans l'oxyde de silicium, est plus importante que celle des trous.
- ∅ Isolation par jonction inversement polarisée : la taille de ces jonctions doit toujours être réduite au minimum possible. Dans le cas des jonctions faiblement dopées (i.e. de haute résistivité), telles que celles dans le caisson ou le substrat, de nombreux contacts (i.e. prises de caisson, de substrat) doivent être mis, afin de réduire (dans le caisson ou le substrat) les variations de potentiel dues aux courants de fuite en haute température.
- ∅ Latchup : Au fur et à mesure que la température augmente, les courants de déclenchement et de maintien diminuent progressivement, augmentant les risques de *latchup*. Cependant, quelques techniques d'implémentation permettent d'augmenter sensiblement (jusqu'à 300°C selon [3]) la robustesse aux *latchups* (voir la section 4.2.1.3). Ces techniques sont plus efficaces encore dans le cas des technologies sur substrat épitaxial, car le substrat de base a une plus faible résistivité.
- ∅ Résistances: Les résistances sur silicium monocristallin souffrent de plusieurs désavantages, tels que les importants courants de fuite en haute température, les considérables non-linéarités et capacités parasites (selon la concentration des dopants) et les variations de la résistivité avec la température, qui les rendent peu adaptées aux applications en haute température. En revanche, les résistances sur silicium polycristallin offrent de très bonnes propriétés pour la haute température, principalement du fait de leurs courants de fuite et de leurs non-linéarités négligeables, dans la plage de température considérée (<250°C). Il faut cependant prendre en compte leurs coefficients thermiques, plus ou moins importants selon la concentration des dopants dans le film polycristallin.
- ∅ Capacités: Les capacités Double-Poly ont aussi une performance pratiquement idéale en haute température, du fait de leurs courants de fuite et de leurs non-linéarités négligeables, dans la plage de température considérée (<250°C). Par contre, des valeurs de capacité par unité de surface plus élevées peuvent être normalement obtenues avec les capacités PMOS en accumulation. Cependant l'influence des non-linéarités et des variations selon la température des capacités PMOS en accumulation doit être étudiée en détail dans chaque application (voir la section 5.2.1.4.f.c). Les courants de fuite dans le caisson des capacités PMOS limite leur utilisation aux applications où le caisson peut être connecté à un nœud de très basse impédance.
- ∅ Transistors MOS : La tension de seuil et le facteur de gain $KP = \mu_{n,p} \cdot Cox$ (plus important dans le cas des transistors NMOS que dans le cas des PMOS) diminuent avec la température de façon

comparable dans le cas des NMOS et des PMOS. La tension de *Early* dépend fortement de la longueur physique du canal L et de la polarisation. En revanche, elle dépend faiblement de la température. La saturation de la vitesse de dérive des porteurs dans le canal peut être évitée en diminuant la valeur du champ électrique longitudinal, en augmentant la longueur du canal L . L'utilisation de la réponse exponentielle des transistors MOS en inversion faible, dans les applications en haute température, n'est pas recommandable à cause des forts courants de fuite qui réduisent fortement, en haute température, la plage des courants d'inversion faible. Les courants de fuite dans le drain et la source augmentent exponentiellement avec la température, mais de façon différente pour les NMOS et les PMOS. En haute température, ces courants de fuite sont beaucoup plus importants dans le cas des NMOS que dans le cas des PMOS. Le courant de fuite de la comparativement très large jonction caisson-substrat des PMOS doit être pris en compte, si le caisson n'est pas connecté à un nœud de très basse impédance.

- ∅ Transistors bipolaires NPN verticaux : Le gain en courant β augmente avec la température, la transconductance gm diminue de façon inversement proportionnelle à la température, et la tension d'Early diminue légèrement avec la température. Le courant de fuite, de la comparativement très large jonction collecteur-substrat, et sa conductance associée, doivent être soigneusement considérés.
- ∅ Transistors bipolaires PNP latéraux : La grille doit être polarisée approximativement à $Vg \geq 0.7$ volts (Vg étant la différence de tension entre le caisson et la grille) pour réduire la sensibilité de la *fonction de grille* fg par rapport à Vg et à la température. Le courant du collecteur latéral, le gain en courant βl et la transconductance, ont un comportement similaire à celui de transistors verticaux, mais sont modulés par la fonction de grille fg . Les transistors latéraux des technologies étudiées ont un gain en courant supérieur et une tension de Early inférieure à ceux des transistors verticaux. Du fait de l'absence d'une jonction inversement polarisée entre le collecteur latéral et le substrat, les larges courants de fuite présents dans le collecteur des transistors verticaux ne font pas partie du courant du collecteur latéral.

4.1.3. Architectures robustes

Une bonne méthode, pour définir au niveau système des architectures robustes aux hautes températures, est celui d'interconnecter les différents blocs fonctionnels à travers des éléments de très faibles courants de fuite, tels que la grille des transistors MOS, les résistances en polysilicium, les capacités Double-Poly, etc., de façon à assurer l'indépendance du comportement de chaque bloc, même à la plus haute température considérée. Dans chacun des blocs fonctionnels, différentes techniques de compensation pourront être implementées afin de réduire l'impact de la dégradation des composants en haute température sur la performance de chaque circuit en particulier.

Les variations de certains paramètres selon la température peuvent être totalement ou partiellement compensées par des techniques de polarisation. En principe, un paramètre doit avoir un degré important de non-linéarité pour qu'il puisse, de cette manière, être compensé en température. Quelques exemples de ce type de paramètre sont:

- ∅ La capacité d'une jonction pn polarisée en inverse (équation (78)), la capacité MOS en accumulation (équation (102)), la tension de seuil des transistors MOS (équations (111) et (112)), la transconductance des transistors MOS en régime linéaire (équation (123)), la conductance drain-source des transistors MOS en régime linéaire (équation (124)), etc. Les variations selon la température de ces paramètres peuvent être réduites avec une polarisation en tension appropriée.

- ∅ la transconductance des transistors bipolaires (équation (170)), la transconductance des transistors MOS en régime de saturation (équation (127)) et en inversion faible (équation (141)). Les variations selon la température de ces derniers paramètres peuvent être réduites avec une polarisation en courant appropriée.

Les tensions et les courants de polarisation peuvent être appliqués de deux manières différentes:

- ∅ En boucle ouverte: des sources de courant et de tension dépendantes de la température sont utilisées pour compenser les variations d'un paramètre, dont la réponse en température est connue.
- ∅ En boucle fermée: un système de contrôle asservi permet de fixer la valeur du paramètre, à partir d'une référence donnée.

4.1.3.1. Compensations en boucle ouverte

Si la réponse en température du paramètre à compenser est connue, ses variations selon la température du paramètre peuvent être diminuées en appliquant au dispositif une polarisation, à travers des sources de tension ou de courant qui soient une fonction appropriée de la température. De même, si un paramètre d'un dispositif a une dépendance de la température qui varie, selon les conditions de polarisation, il peut être polarisé d'une manière stable autour du point le moins sensible, avec des sources de tension ou de courant invariantes par rapport à la température. Des exemples de sources de tension et de courant dépendantes de la température sont montrés dans la Figure 70 et décrits dans les trois prochaines sections.

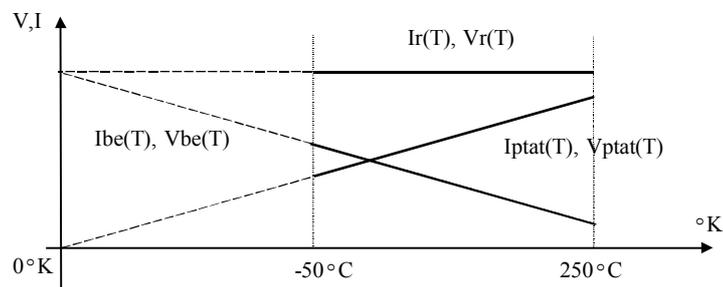


Figure 70. Sources de tension et de courant dépendantes de la température.

D'autres types de compensation en boucle ouverte sont aussi possibles. Deux exemples sont abordés dans la section 5.2.1, où la non-linéarité du signal de sortie d'un capteur de température est compensée en température, par le coefficient thermique d'une résistance en polysilicium et par une source de courant faiblement dépendante de la température.

4.1.3.1.a. Sources proportionnelles à la température avec un coefficient positif

Le courant du collecteur latéral du PNP CLBT (voir section 3.5.2.1) est donné par,

$$(195) \quad I_c = J_s \cdot fg \cdot A_e \cdot e^{\frac{V_{be}}{kT}}$$

Où fg dépend de la polarisation de la grille et dépend faiblement de la température. Mais la fonction de grille fg a pratiquement la même valeur quand deux transistors identiques sont polarisés de la même manière. Ainsi, la différence entre les tensions base-emetteur V_{be} de deux transistors bipolaires, qui

ont des densités de courant de collecteur différentes, est proportionnelle à la température absolue ($PTAT=Proportional\ To\ Absolute\ Temperature$) et elle ne dépend pas des paramètres technologiques,

$$(196) \quad V_{ptat} = V_{be3} - V_{be4} = \frac{kT}{q} \ln \frac{I_{c3}}{J_s \gamma_{fg} A_{e3}} - \ln \frac{I_{c4}}{J_s \gamma_{fg} A_{e4}} = \frac{kT}{q} \ln \frac{I_{c3} \gamma_{fg} A_{e4}}{I_{c4} \gamma_{fg} A_{e3}}$$

La Figure 71 montre le circuit utilisé pour l'implémentation de la source de courant et la source de tension, proportionnelles à la température avec un coefficient thermique positif. Ce circuit est une version améliorée de celle présentée dans [51]. La résistance en polysilicium $R1$ est placée entre les bases des transistors $T3$ et $T4$. Le courant ainsi généré est donné par,

$$(197) \quad I_{ptat} = \frac{V_{be3} - V_{be4}}{R1}$$

Dans la mesure où les transistors latéraux ont une faible tension d'Early, une boucle de contrôle ($M7 + OPAMP$) a été ajoutée afin d'imposer aux transistors $T6$ et $T5$ la même tension collecteur-émetteur. De plus, l'impédance de sortie de la source de courant est ainsi améliorée.

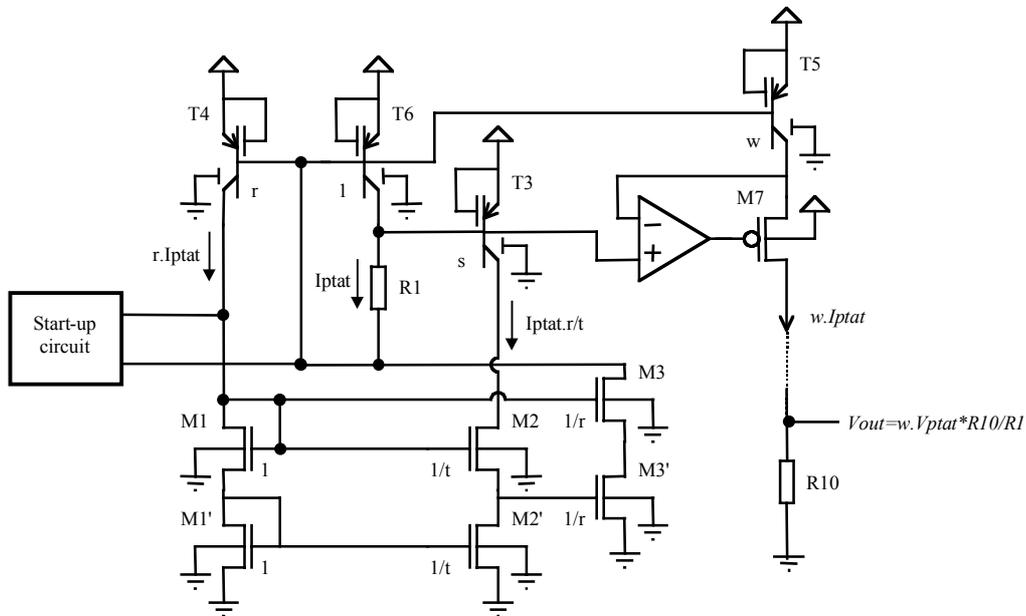


Figure 71. Circuit de la source de courant $I_{ptat}(T)$ et de tension $V_{ptat}(T)$

Dû aux bons résultats obtenus, le circuit de la Figure 71 a été utilisé dans l'implémentation du capteur de température, décrit dans la section 5.2.1, et a été aussi utilisé pour la compensation en température de la transconductance (voir l'équation (235)) de l'étape de gain des oscillateurs à cristal, et pour la compensation en température de l'erreur d'amplitude (voir l'équation (250)), dans la boucle de contrôle automatique de gain des oscillateurs à cristal, décrites dans la section 5.1.1.1.b.

4.1.3.1.b. Sources proportionnelles à la température avec un coefficient négatif

A partir de l'expression du courant de collecteur latéral I_c , donnée par (181) on obtient la valeur de la tension de base-émetteur V_{be} ,

$$(198) \quad V_{be}(T) = \frac{kT}{q} \ln \frac{I_c}{\bar{\mu}_{pb}(T) n_i(T)^2 f_g \frac{kT A_e}{W_b N_{db}}}$$

Les faibles variations de f_g avec la température sont négligeables par rapport aux variations de la mobilité et de la concentration intrinsèque. En introduisant dans l'expression précédente les équations correspondant à la mobilité et à la concentration intrinsèque, respectivement (25) et (8), on obtient,

$$(199) \quad V_{be}(T) = V_g(T) - \frac{kT}{q} \ln \frac{-T^\gamma CTE A_e}{I_c W_b N_{db}}$$

$\gamma = 4 - TC_p$, $V_g(T) = E_g(T)/q$ (voir l'équation (1)) et CTE est une constante qui ne dépend pratiquement pas de la température. Si le transistor est polarisé avec une source de courant de type $I_c = I_c(T_0) [T/T_0]^\alpha$, $V_{be}(T)$ s'écrit donc,

$$(200) \quad V_{be}(T) = V_g(T) - \frac{kT}{q} \ln \frac{-T_0^\gamma CTE A_e}{I_c(T_0) W_b N_{db}} - (\gamma - \alpha) \frac{kT}{q} \ln \frac{T}{T_0}$$

$V_{be0} = V_{be}(T_0)$ est donnée par,

$$(201) \quad V_{be0} = V_{g0} - VT_0 \ln \frac{-T_0^\gamma CTE A_e}{I_c(T_0) W_b N_{db}}$$

Où V_{g0} est le potentiel de bande interdite à T_0 et $VT_0 = kT_0/q$. A partir des deux relations précédentes, $V_{be}(T)$ s'écrit,

$$(202) \quad V_{be}(T) = V_g(T) - (V_{g0} - V_{be0}) \frac{T}{T_0} - (\gamma - \alpha) \frac{kT}{q} \ln \frac{T}{T_0}$$

On obtient, par le développement en série de Taylor du premier et troisième terme de l'équation (202), l'approximation de deuxième ordre suivante,

$$(203) \quad V_{be}(T) = V_{orig} - ST T - QT \frac{T - T_0}{T_0}$$

avec,

$$(204) \quad V_{orig} = V_{g0} + \frac{a b T_0^2}{(T_0 + b)^2} + (\gamma - \alpha) VT_0$$

$$(205) \quad ST = V_{g0} - V_{be0} + a T_0 \left[1 - \frac{b^2}{(T_0 + b)^2} \right] + (\gamma - \alpha) VT_0 \frac{1}{T_0}$$

$$(206) \quad QT = \frac{a b^2 T_0^2}{(T_0 + b)^3} + \frac{(\gamma - \alpha) VT_0}{2}$$

V_{be0} est une variable qui dépend fortement des paramètres technologiques. Cela affectera seulement la valeur de la pente ST , mais pas la constante V_{orig} , ni le coefficient QT du terme quadratique. Il en résulte que la linéarité ne sera pas affectée par les variations des paramètres technologiques. La valeur de V_{orig} est principalement déterminée par des constantes (V_{g0} , a , b , k , q). Elle est indépendante de la géométrie des transistors, et pratiquement indépendante des paramètres technologiques. La valeur de la pente ST peut être réglée, à travers V_{be0} , en modifiant la taille du transistor bipolaire et le courant de polarisation (cette propriété a été exploitée pour le contrôle de la

pente en température d'une référence de courant, et pour l'amélioration de la linéarité du signal de sortie du capteur de température décrit dans la section 5.2.1.4). Le coefficient QT est responsable de la faible non-linéarité de $V_{be}(T)$, laquelle peut être annulée si l'exposant thermique α a la valeur suivante,

$$(207) \quad \alpha_o := \gamma + \frac{2 \cdot a \cdot b^2 \cdot q \cdot T_0}{k \cdot (T_0 + b)^3} \quad \alpha_o = 3.426 \text{ pour } \gamma = 1.8$$

La Figure 72 montre le circuit utilisé pour l'implémentation de la source de courant et la source de tension, proportionnelles à la température avec un coefficient thermique négatif. C'est un circuit classique, utilisé normalement pour la génération des courants dérivés de la tension base-emetteur.

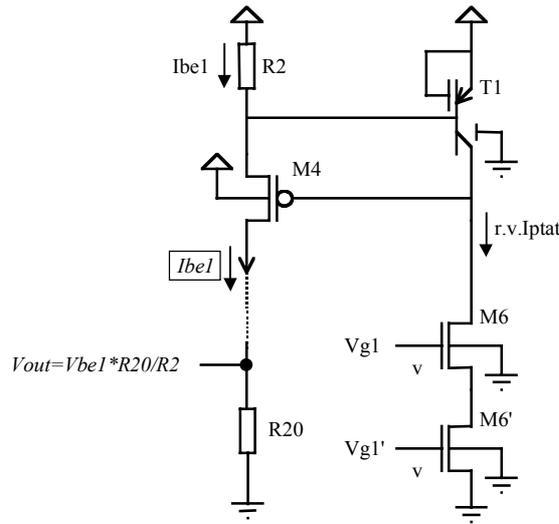


Figure 72. Circuit de la source de courant $I_{be}(T)$ et de la source de tension $V_{be}(T)$.

Le courant $I_{be}(T)$ est généré avec la résistance $R2$, en polysilicium, et la tension base-emetteur du transistor $T1$. La grille du CLBT est connectée à l'émetteur, donnant une polarisation grille-caisson égale à la tension V_{be} (~ 0.7 volts). Le courant du collecteur du CLBT est un courant PTAT généré par le circuit de la Figure 71, donc " $\alpha=1$ ". Le transistor $M4$ a été ajouté afin d'améliorer l'impédance de sortie de la source de courant. Le circuit de la Figure 72 a été utilisé dans l'implémentation du capteur de température décrit dans la section 5.2.1.

4.1.3.1.c. Source de courant et de tension avec un coefficient thermique nul

Les deux types de sources précédentes peuvent être additionnées (voir Figure 70) respectivement en mode tension ou en mode courant, afin d'obtenir des sources de tension et de courant avec un coefficient thermique nul. L'addition peut facilement être faite en mode courant, à partir des circuits de la Figure 71 et de la Figure 72, et le signal $I_r(T)$ ainsi obtenu reconverti en mode tension $V_r(T)$ avec une résistance ($R30$) qui a le même comportement en température que $R1$ et $R2$,

$$(208) \quad I_r(T) = \frac{V_{be}(T)}{R2(T)} + \frac{V_{ptat}(T)}{R1(T)} = \frac{V_{orig} - QT \cdot \frac{T - T_0}{T_0}}{R2(T)}$$

$$(209) \quad V_r(T) = \frac{-V_{be}(T)}{R2} + \frac{V_{ptat}(T)}{R1} \cdot R30 = \frac{V_{orig} - QT \cdot \frac{T - T_0}{T_0}}{R2} \cdot R30$$

Le comportement en température du signal en mode courant $I_r(T)$ est modifié par la dépendance de la résistance R_2 par rapport à température (voir la section 3.1.2). Ce phénomène peut être exploité pour réduire la valeur du terme quadratique et obtenir ainsi une meilleure linéarité (voir la section 5.2.1.2). En revanche, le signal en mode tension n'est pas affecté par la dépendance de la température des résistances, car il est une fonction d'un rapport de résistances de même nature.

Puisque la valeur de V_{orig} est pratiquement donnée par la largeur de la bande interdite du silicium, ces types de sources à coefficient thermique nul sont normalement appelés *bandgap references*. Ces dernières sont souvent utilisées pour la polarisation des circuits de manière invariante par rapport à la température (par exemple dans le cas des transistors MOS polarisés autour de leur V_{gsZTC}) et comme références de tension (ou de courant) des systèmes asservis, tels que les régulateurs de tension, etc.

Une référence de courant similaire à celle précédemment décrite, mais au coefficient thermique réglable, a été implémentée dans la section 5.2.1.4.d.

4.1.3.2. Compensations par boucles de contrôle

L'implémentation des boucles de contrôle pour la compensation en température s'applique plutôt aux paramètres globaux des circuits qu'aux paramètres des dispositifs. D'une part, parce que l'effort de conception est considérable (Analyse statique : erreur du système de contrôle et stabilité de la référence. Analyse dynamique : stabilité en fréquence et comportement au démarrage (*start-up*) de la boucle de contrôle) et d'autre part, car l'implémentation est normalement très coûteuse en terme de consommation et en terme de surface occupée sur le substrat. De plus, la stabilité en température de la boucle de contrôle doit être soigneusement étudiée.

Il existe deux types de boucles de contrôle : les boucles de contrôle linéaires, où la réponse du système (i.e. la position des pôles et des zéros) ne dépend pas de l'état de la boucle, et les boucles de contrôle non-linéaires, où la réponse varie selon les états du système (i.e. les pôles et les zéros bougent). Dans le deuxième cas, les méthodes d'analyse sont beaucoup plus complexes et les solutions obtenues parfois limitées (i.e. elles ne prennent en compte qu'un nombre, plus ou moins, limité d'états possibles du système).

Quelques circuits bouclés linéaires, fonctionnant à base d'amplificateurs opérationnels résistants aux hautes températures (voir la section 5.1.1.1.e), ont été utilisés pour,

- ∅ L'amélioration de l'impédance de sortie du miroir de courant bipolaire T_6-T_5 , de la Figure 71.
- ∅ La réalisation des filtres actifs et des étapes d'amplification, dans la section 5.1.1.1.d.

Leur fonctionnalité a été vérifiée jusqu'à environ 220°C (voir la section 5.1.1.1.f). Les solutions obtenues (i.e. les circuits implémentés), étant des solutions classiques avec des marges de sécurité élargies pour prendre en compte les variations de la température, ne seront pas abordées en détail.

En revanche, dans la section 5.1.1.1, le problème du contrôle automatique de l'amplitude des oscillations des oscillateurs à cristal, est abordé. La réponse de l'amplitude des oscillations (i.e. le pôle $-1/T_E$ associée à l'enveloppe) dépend fortement des conditions de polarisation de l'étape de gain. La position du pôle de l'enveloppe peut varier idéalement entre ∞ et $-\infty$, 0 étant la valeur du pôle nécessaire pour obtenir des amplitudes stables. Par conséquent le contrôle de l'amplitude de ce système, à travers la polarisation de l'étape de gain, est forcément non-linéaire. Une analyse très détaillée du comportement, statique et dynamique, du circuit de contrôle non-linéaire implémenté, ainsi que les résultats des tests en température, sont présentés dans la section 5.1.1.1.b.

4.1.3.3. Sous-circuits analogiques robustes

Les gradients des paramètres de fabrication mentionnés au début de la section 4.1.1, ne peuvent pas être contrôlés par le concepteur des circuits analogiques. En revanche, les gradients thermiques sont

bien sous le contrôle du concepteur. La bonne conductivité thermique du silicium permet une bonne uniformisation automatique de la température à la surface. Par conséquent, une disposition intelligente des éléments générateurs de chaleur et des éléments sensibles aux gradients thermiques, permet d'obtenir, en haute température, des degrés de similitude ("*matching*") entre composants, équivalents à ceux obtenus à température ambiante. Ainsi, le *matching* reste la propriété la plus intéressante pour l'obtention de degrés de précision élevés, dans les sous-circuits analogiques (i.e. les miroirs de courant, les étapes de gain différentielles, etc.).

Les signaux peuvent être représentés de différentes manières (*modes*) dans un circuit analogique, chacune étant affectée de façon différente par la température, selon chaque application particulière. En guise d'exemple, on peut citer quelques caractéristiques des *modes* les plus courants suivants:

- ∅ Mode tension: Les nœuds de haute impédance, dans les circuits en mode tension en boucle ouverte, doivent être évités pour que les forts courants de fuite n'affectent pas les signaux en tension. L'utilisation des étapes de gain différentielles MOS est un des points forts. Des circuits bouclés en mode tension, avec une entrée différentielle et une sortie en basse impédance, permettent de bien absorber les variations de charge à la sortie (courants de fuite en haute température, etc.) sans dégrader les signaux en tension (voir les filtres actifs en mode tension dans la section 5.1.1.1.d). La consommation et la surface occupée sur le substrat de silicium, augmenteront.
- ∅ Mode courant: L'utilisation extensive de miroirs de courant permet d'implémenter de manière très compacte un nombre élevé de fonctions, et d'ajouter ou de soustraire facilement les signaux (voir la section 5.2.1.2). Les circuits doivent être optimisés pour réduire au minimum les courants de fuite. Ainsi, les signaux en mode courant doivent être surdimensionnés, dans la plus part des cas, afin de s'imposer aux courants de fuite, ce qui entraîne une augmentation de la consommation.
- ∅ Mode stockage de charge: Il s'agit généralement des circuits en mode tension à capacités commutées. Il faut donc prendre en compte les mêmes considérations que pour les circuits en mode tension, afin d'éviter la dégradation de signaux en mode tension, pendant la charge et la décharge des capacités. Ensuite, les valeurs de tension stockées, étant proportionnelles à la charge dans une capacité, se dégradent à cause des courants de fuite dans les interrupteurs (*switches*) en état ouvert. La taille des interrupteurs doit être minimisée (voir la Figure 114), au détriment de la vitesse du circuit.
- ∅ Mode fréquence: Les signaux en mode fréquence sont généralement très insensibles à la température, si la variation en température des capacités parasites dans le circuit intégré n'est pas dominante. L'implémentation des fonctions en mode fréquence (i.e. génération des signaux, mélange des fréquences, filtrage, modulation, etc.) est généralement complexe, et devient plus complexe encore en haute température (voir la section 5.1.1.1), ce qui entraîne une augmentation de la consommation et de la surface occupée sur le substrat de silicium.

4.1.4. Modèles de simulation

Les modèles de simulation des dispositifs passifs et actifs, évoluent en précision et complexité suivant l'évolution des technologies des circuits intégrés et les exigences des applications. Notamment, la conception de circuits analogiques et mixtes a motivé le perfectionnement des modèles des transistors MOS, lesquels continuent à évoluer du fait de la réduction constante des dimensions qui entraîne l'incorporation des nouveaux phénomènes de conduction (i.e. les effets de canal court, étroit).

4.1.4.1. Transistors MOS

Le premier modèle de transistors MOS développé sous SPICE est le modèle *Level 1*, décrivant d'une manière très élémentaire le comportement physique des transistors MOS. Même s'il est maintenant obsolète, il est le seul modèle répandu assez simple pour permettre de faire des calculs à la main des composants et des circuits. Le modèle *Level 2* a été développé quand les effets des petites dimensions ont commencé à prendre de l'importance. Il a été rapidement remplacé par le *Level 3*, à cause des problèmes numériques et de précision.

Le modèle *Level 3* est encore largement utilisée pour la conception de circuits digitaux, du fait de sa simplicité. Cependant, *Level 3* n'est pas bien adapté pour la conception des circuits analogiques, à cause des discontinuités dans la dérivée du courant drain-source, et parce qu'il lui manque une modélisation exploitable de la conduction sous-seuil.

Le modèle *BSIM* représente un changement radical dans la manière de modéliser les transistors MOS. Les équations du modèle sont systématiquement conditionnées pour améliorer la performance des calculs numériques. Par conséquent, la complexité et le caractère empirique du modèle augmentent considérablement dans ce modèle. Ainsi, le nombre de paramètres devient beaucoup plus large, la plupart d'entre eux étant des paramètres de nature empirique. Il en résulte que la qualité d'un jeu de paramètres en particulier, devient fortement dépendante de la qualité des procédés d'extraction des paramètres. Même si la modélisation de la conduction sous-seuil a été améliorée, le modèle *BSIM* a encore quelques discontinuités dans la dérivée du courant drain-source ce qui le rend peu adapté pour le design des circuits analogiques.

Un des plus intéressants modèles pour l'analogique est le *Level 28* de HSPICE. Il contient les équations de *BSIM* améliorées afin d'éliminer les discontinuités dans la dérivée du courant drain-source, et pour améliorer la transition entre les différents régimes d'opération des transistors MOS, générant des simulations très rapides. Mais l'utilisation du *Level 28*, étant un modèle de la propriété de HSPICE, est limitée à l'utilisation du simulateur HSPICE.

Le modèle *BSIM3* prétendait au début simplifier et rendre plus physique les équations et paramètres du modèle. Plusieurs versions et sous-versions ont été livrées, jusqu'à l'actuelle version *BSIM3v3*. Ainsi, le modèle *BSIM3* a évolué pour devenir le plus complexe et empirique des modèles MOS. Il génère des courants dont la dérivée est continue, et permet de bien modéliser toutes les tailles des transistors MOS possibles, avec un même jeu de paramètres. Mais sa complexité mathématique le rend très lent en simulation [66]. Le modèle *MOS Model 9 (MM9)* développé par Philips, et adopté par plusieurs fabricants européens de circuits intégrés, se présente comme le modèle européen concurrent du *BSIM3v3*. Il a été conçu pour la simulation analogique et digitale, particulièrement dans le cas des transistors sous-microniques.

Un autre modèle prometteur, en ce qui concerne la conception des circuits analogiques, est le modèle *EKV*. Il a été spécialement développé pour les applications de faible puissance. Le modèle *EKV* modélise les conductances sans discontinuités, il utilise un jeu réduit de paramètres (la plupart de nature physique), et il est très rapide en simulation. De plus, le modèle est construit de manière hiérarchique, permettant un premier calcul à la main, à travers un jeu réduit de paramètres et d'équations relativement simples, suivi d'un raffinement du design en utilisant le modèle complet.

4.1.4.1.a. Modèles fournis par le fabricant

Les technologies CMOS et BiCMOS choisies, pour les applications développées dans le chapitre 5, sont caractérisées et modélisées pour des applications analogiques et digitales. Deux types de modèles de transistors MOS sont fournis par le fabricant,

- ø Modèle standard: BSIM3v2 adapté aux applications analogiques, à travers une stratégie spéciale pour l'extraction de paramètres, fondée sur un bon ajustement de g_m , g_{ds} , g_{mb} et des points d'opération dans le régime de faible inversion.

- 0 Modèle MOS amélioré: C'est le *MOS15 Model* appartenant au fabricant. Il s'agit d'un modèle fondé sur le *Level 2* standard, avec des équations améliorées pour la mobilité et la tension de saturation, et des dérivées continues dans tous les régimes d'opération. En particulier, la région d'inversion modérée et la transition entre le régime linéaire et le régime saturé, ont été plus soigneusement modélisées. Il en résulte que la convergence est ainsi améliorée par rapport aux modèles standards. Le *MOS15* est particulièrement intéressant pour la conception de circuits intégrés analogiques.

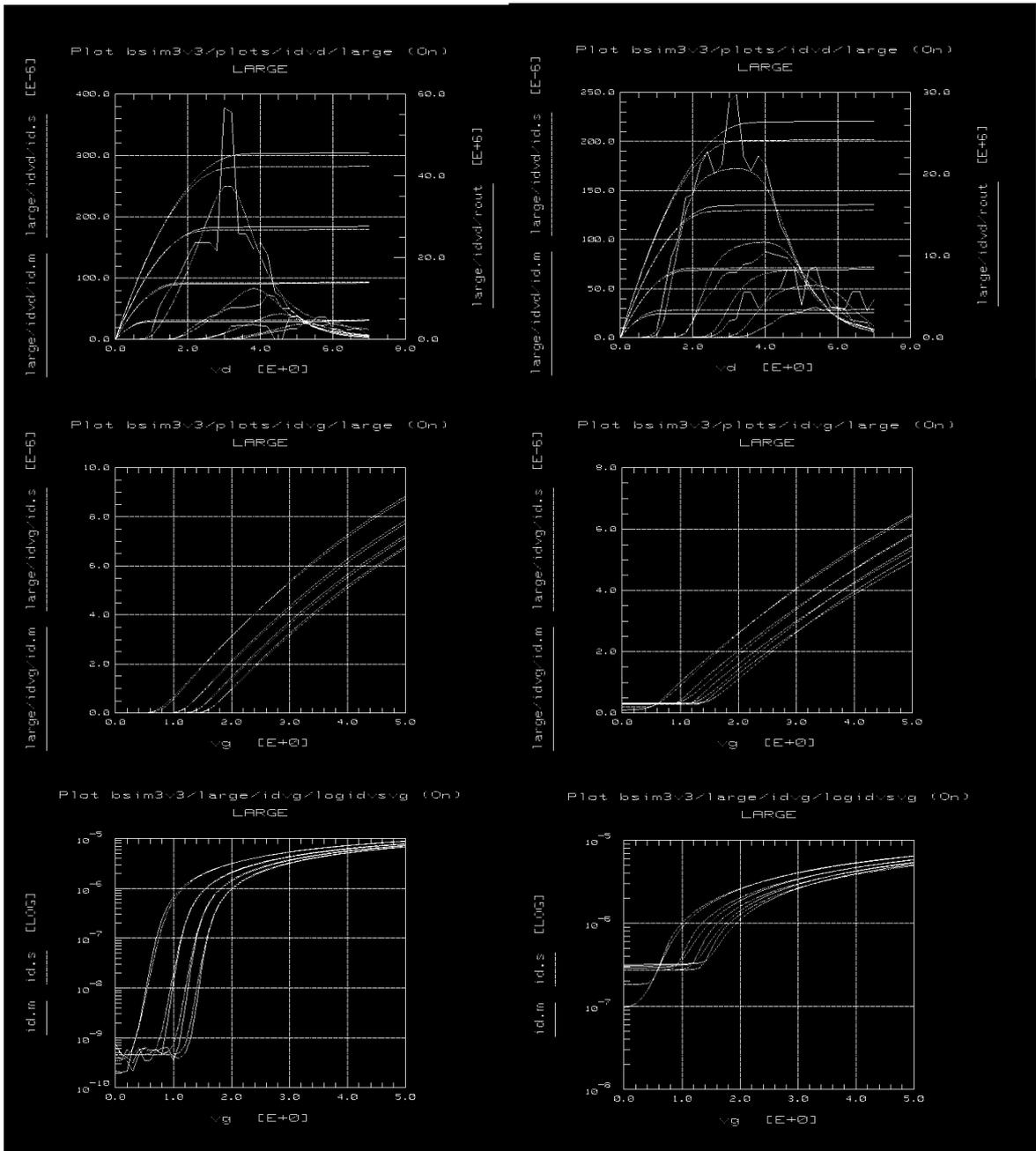


Figure 73. Extraction des paramètres du modèle BSIM3v3 pour les transistors NMOS, à 150°C (à gauche) et à 250° (à droite): courbes mesurées(en trait continu) et simulées (en pointillé).

Les jeux de paramètres de ces modèles fournis par le fabricant, étant caractérisés entre 0 et 100°C , ont été utilisés pour les simulations à températures normales, et à haute température comme première approximation. Particulièrement le *MOS15*, de nature plus physique que le *BSIM3v2*, a donné de meilleurs résultats en haute température.

4.1.4.1.b. Caractérisation et extraction des paramètres

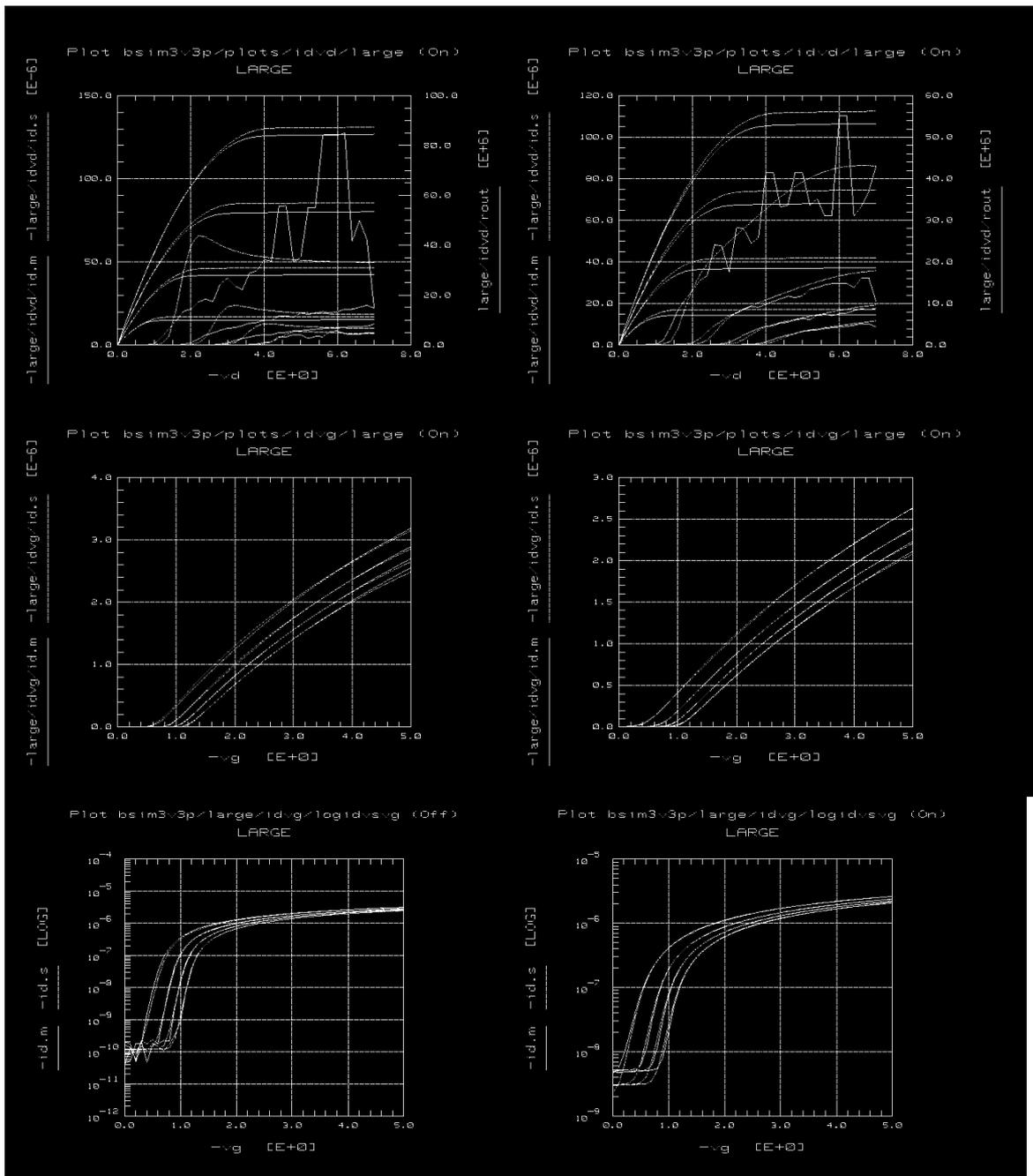


Figure 74. Extraction des paramètres du modèle BSIM3v3 pour les transistors PMOS, à 150°C (à gauche) et à 250°C (à droite): courbes mesurées (en trait continu) et simulées (en pointillé).

Une caractérisation et extraction, en DC, des paramètres des transistors MOS a été faite pour les modèles *BSIM3v3*, avec le logiciel extracteur des paramètres IC-CAP5.0, l'analyseur de paramètres des composants semiconducteurs HP4155A, et les transistors NMOS et PMOS des tailles suivantes: 20/20, 20/3, 20/1.2, 20/0.9, 20/0.8, 3/20, 3/0.8/ 2/20, 2/0.8 μm . Les routines d'extraction, ne permettant pas de faire des optimisations sur une large plage de température, un jeu de paramètres a été extrait pour chacune des températures suivantes: 150°C, 175°C, 200°C, 225°C et 250°C. Dans la Figure 73 et la Figure 74, on peut observer le degré d'ajustement obtenu entre les courbes mesurées et les simulations (en pointillé) correspondant à un transistor long et large. Dans le cas des transistors courts et étroits, l'ajustement est moins serré, mais ces différences ne varient pas fortement avec la température.

4.1.4.2. Transistors bipolaires

Les transistors bipolaires ont été les premiers dispositifs actifs développés et modélisés, mais l'évolution des ces modèles a été moins importante que celle des transistors MOS. Le premier modèle développé est le modèle *Ebers-Moll*, dans lequel le transistor bipolaire est considéré comme une paire interactive de diodes. Il décrit le comportement idéal des transistors, sans prendre en compte la modulation de la largeur de la base, les résistances parasites, les effets de la forte injection des porteurs dans la base, etc. Ces effets, ainsi que d'autres, comme les effets en température, la variation du gain pour des faibles courants, etc., ont été pris en compte dans le modèle de *Gummel-Poon*, plutôt fondé sur l'intégration de la charge dans la région neutre de la base, que dans le comportement des deux diodes.

4.1.4.2.a. Modèles fournis par le fabricant

Les technologies CMOS et BiCMOS choisies, pour les applications développées dans le chapitre 5, sont caractérisées et modélisées pour des applications analogiques et digitales. Le modèle de transistor bipolaire fournis par le fabricant, concernant les transistors NPN verticaux et les PNP latéraux, est un modèle *Gummel-Poon* compatible avec SPICE. Le transistor vertical parasite présent dans la structure des transistors PNP latéraux est incluse uniquement dans le modèle de la technologie CMOS, car dans la technologie BiCMOS le gain de ce transistor parasite est négligeable.

Les jeux de paramètres des modèles fournis par le fabricant, étant caractérisés entre 0 et 100°C, ont été utilisés pour les simulations à températures normales, et à haute température comme première approximation.

4.1.4.2.b. Caractérisation et extraction des paramètres

Une caractérisation et extraction, en DC, des paramètres des transistors bipolaires a été faite pour le modèle *UCB (University of California at Berkeley)*, compatible avec SPICE, des transistors bipolaires. L'équipement de modélisation était constitué par le logiciel extracteur des paramètres IC-CAP5.0, l'analyseur de paramètres des composants semiconducteurs HP4155A. Les transistors bipolaires latéraux et verticaux testés ont une surface d'émetteur respectivement de $4 \times 3.6 \times 0.4 \mu\text{m}^2$ et $0.8 \times 3 \mu\text{m}^2$ (correspondant aux dimensions de transistors utilisés dans les applications du chapitre 5). Les routines d'extraction ne permettant pas de faire des optimisations sur une large plage de température, un jeu de paramètres a été extrait pour chacune des températures suivantes: 150°C, 175°C, 200°C, 225°C et 250°C.

4.2. Techniques d'implémentation

Les techniques d'implémentation concernent les différentes manières de traduire le circuit conçu en masques, pour la fabrication du circuit intégré. C'est une activité fortement liée aux caractéristiques de la technologie. Elle permet d'améliorer substantiellement quelques propriétés des composants intégrés, comme le *matching*, la robustesse aux latchups, etc., parfois au détriment de la taille de la surface occupée sur le substrat de silicium.

4.2.1.1. Maximisation de la précision relative

La performance ou la précision de la plupart des circuits analogiques, est normalement fondée sur le haut degré de similitude ("*matching*") obtenu entre composants de caractéristiques identiques. Pour une technologie donnée, le *matching* entre les composants critiques peut être amélioré, même à de hautes températures, en appliquant les règles d'implémentation suivantes,

- ⊖ *Les dispositifs doivent avoir la même structure.* Par exemple, un diviseur résistif $R1/(R1+R2)$ doit être implémenté avec deux résistances de même nature, d'une part pour assurer que les erreurs de déposition, dopage et gravure soient les mêmes pour $R1$ que pour $R2$, et d'autre part pour que l'évolution en température soit la même pour les deux résistances. De même, dans un inverseur CMOS, le courant de fuite dans le drain du PMOS ne peut pas être compensé par le courant de fuite du drain du NMOS, car leurs variations selon la température sont très différentes.
- ⊖ *Les dispositifs doivent avoir la même température.* Minimiser la distance entre les dispositifs suffit dans les cas où il n'y a pas de forte dissipation thermique, du fait de la bonne conductivité thermique du silicium. Sinon, les dispositifs doivent être placés sur le même isotherme, c'est à dire, de façon symétrique par rapport au point de dissipation de chaleur.
- ⊖ *Les dispositifs doivent avoir la même forme et la même taille.* Par exemple, dans le cas des transistors MOS, les valeurs de W et L doivent être les mêmes (pas uniquement le rapport W/L), et la surface et le périmètre des diffusions (de drain et de la source) doivent aussi être identiques afin d'obtenir des courants de fuite similaires, correspondants aux côtés et aux fonds des jonctions.
- ⊖ *La distance entre les dispositifs doit être minimisée.* Localement les variations des paramètres et de la température sont moins importantes qu'entre des points très distants.
- ⊖ *Utilisation des structures avec le même centre géométrique.* Les gradients spatiaux, des paramètres et de la température, mentionnés peuvent être localement considérés comme constants. Par conséquent, l'influence de ces variations linéaires est négligeable dans le cas des structures identiques qui ont le centre géométrique en commun. Les paires de transistors ou de capacités, formées par quatre éléments disposés en diagonale, deux à deux, autour du centre géométrique, constituent un bon exemple.
- ⊖ *Les dispositifs doivent être placés selon la même orientation.* De cette façon, l'anisotropie de certains phénomènes, comme la mobilité des porteurs, est évitée. Ainsi, les courants drain-source des transistors MOS doivent être strictement parallèles.
- ⊖ *Les dispositifs doivent être entourés par les mêmes structures.* Cela améliore l'uniformité des procédés de déposition et de gravure.
- ⊖ *Ne pas utiliser des tailles minimales.* C'est la meilleure manière de réduire l'influence des erreurs de gravure et d'améliorer le moyennement spatial des paramètres non-uniformes.

Les règles d'implémentation précédentes s'appliquent tout particulièrement aux paramètres contrôlés des dispositifs, du fait d'une certaine homogénéité garantie par le fabricant. Mais, l'efficacité de ces règles peut être réduite, dans le cas des paramètres non-contrôlés variant fortement avec la température (par exemple, les courants de fuites dans les jonctions inversement polarisées).

4.2.1.2. Minimisation des courants de fuite

Les courants de fuite dans les jonctions *pn* affectent le comportement des composants actifs d'une manière incontournable, du fait du rôle actif de ces jonctions dans le fonctionnement du dispositif ou comme isolant. Le concepteur de circuits intégrés, ne pouvant pas modifier les caractéristiques de ces jonctions, a seulement la possibilité d'identifier les jonctions critiques et de réduire au minimum la surface de ces jonctions.

- ◊ *Drain des transistors NMOS et PMOS*: Comme il a déjà été mentionné dans la section 3.4.6, les courants de fuite dans la jonction du drain augmentent exponentiellement avec la température, et entrent en concurrence avec le courant drain-source. La taille de cette jonction peut être pratiquement divisée par deux en modifiant la manière dont le transistor est dessiné (voir Figure 75). De même, la connexion en série ou en parallèle de deux transistors MOS doit être implémentée en partageant les diffusions, du drain ou de la source, à connecter (voir Figure 76).
- ◊ *Caisson correspondant au substrat des transistors PMOS et au collecteur des transistors bipolaires verticaux* : La taille des caissons, ainsi que l'espacement entre caissons, répond à des règles de conception très strictes, car elles ont un lien avec la robustesse aux latchups de la technologie. Il en résulte que les dimensions du caisson peuvent uniquement être réduites dans le cas de l'implémentation de plusieurs transistors avec la même polarisation du caisson (voir Figure 77).

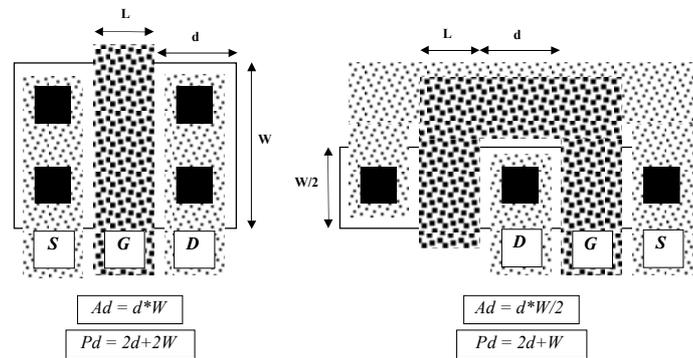


Figure 75. Implémentation normale du transistor MOS (à gauche) et implémentation à jonction du drain réduite (à droite).

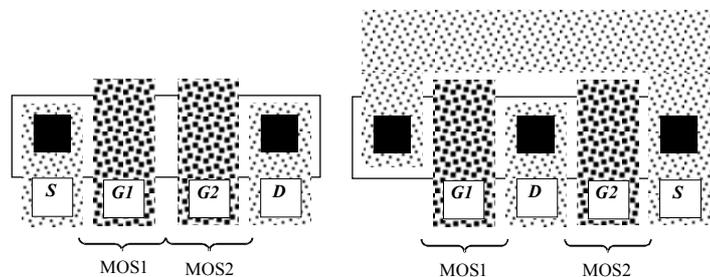


Figure 76. Implémentation d'une connexion de transistors MOS en série (à gauche) et en parallèle (à droite) à jonctions de diffusion réduites.

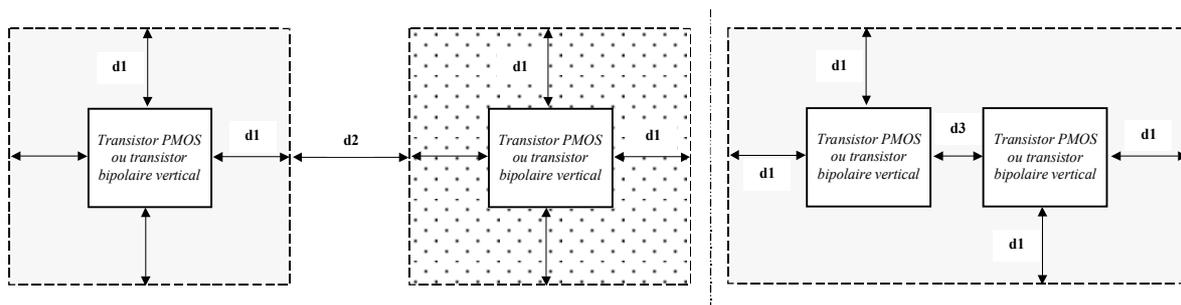


Figure 77. Implémentation de deux transistors avec la même polarisation de caisson, séparés (à gauche) et dans un même caisson de taille réduite (à droite).

4.2.1.3. Robustesse aux latchups

Comme il a déjà été mentionné dans la section 2.3.2.8, les risques de *latchup* augmentent avec la température. Cependant, quelques techniques d'implémentation permettent d'augmenter sensiblement la robustesse aux *latchups*.

- \emptyset *Dégradation du gain en courant βI du transistor bipolaire latéral*: Cela consiste simplement à éloigner le caisson n^- de l'émetteur n^+ , dans la Figure 78, en éloignant le plus possible les transistors PMOS des NMOS. Ainsi, le gain de courant βI du transistor latéral sera réduit par l'augmentation de la largeur effective de sa base.

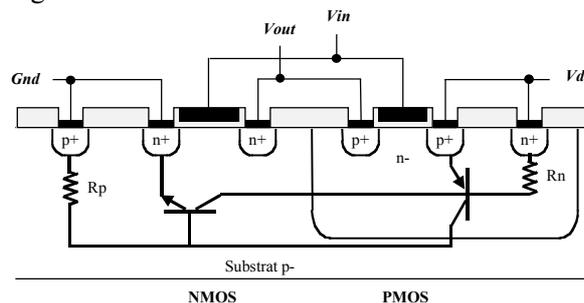


Figure 78. Transistors bipolaires parasites responsables des latchups dans un inverseur CMOS.

- \emptyset *Découplage des transistors par l'insertion de barrières*: Il y a deux types de barrières (voir la Figure 79); des barrières pour les porteurs minoritaires (1) et d'autres pour les porteurs majoritaires (2 et 3):

Les barrières des porteurs minoritaires servent à collecter les porteurs minoritaires injectés dans le substrat, avant qu'ils diffusent dans la base du transistor latéral, en direction du caisson. Si on interpose, entre l'émetteur et le collecteur, une jonction inversement polarisée (voir (1) dans la Figure 79) du type *diffusion n^+* ou *caisson n^-* une partie considérable de ces porteurs minoritaires peut être collectée. Cette barrière est plus effective encore, dans le cas des substrats epitaxiaux. Les barrières sont utilisées pour réduire la différence de potentiel, créée entre la base et l'émetteur du transistor vertical (voir (2) dans la Figure 79) ou latéral (voir (3) dans la Figure 79) par les courants de porteurs majoritaires et les résistances R_n et R_p . L'interposition, entre l'émetteur et le collecteur, d'une barrière de diffusion dopée du même type que la base (et connectée à V_{dd} ou à Gnd dans le cas respectivement d'une prise de caisson ou de substrat) permet de réduire considérablement les résistances effectives R_n et R_p . Dans le cas des substrats epitaxiaux, l'efficacité de ces barrières augmente sensiblement, du fait de la présence d'une région sous-jacente fortement dopée (voir (c) dans la Figure 79).

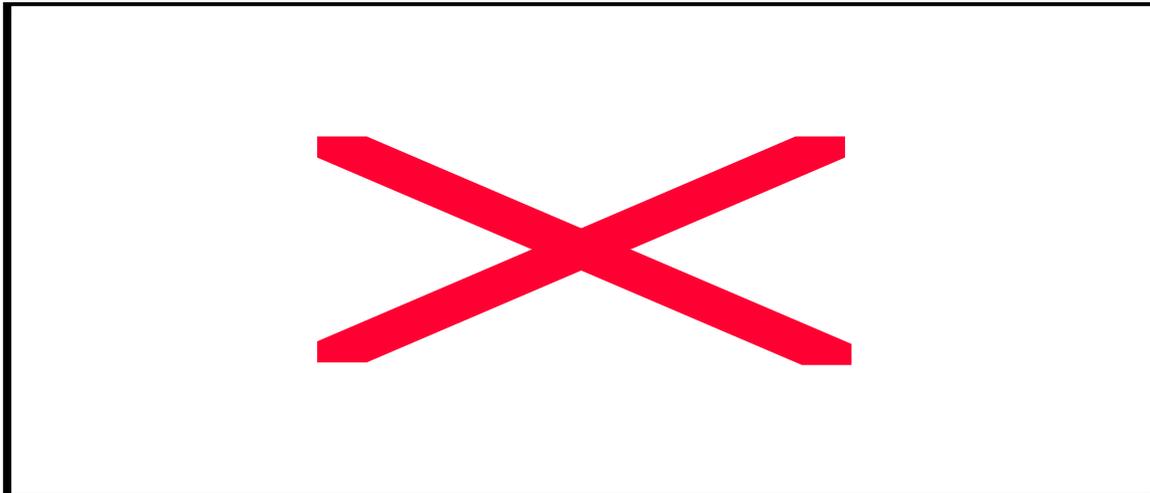


Figure 79. Implémentation de différentes barrières de découplage sur substrat normal (a et b) et sur substrat epitaxiel (c).

L'implémentation des barrières de porteurs majoritaires permet aussi de réduire les variations de tension dans le caisson ou le substrat, générées par les courants de fuite dans les jonctions et par les courants du substrat dus à la génération thermique des porteurs, en haute température. Autrement, ces variations de tension pourraient modifier, par exemple, la valeur de la tension de seuil des transistors MOS et déséquilibrer ainsi l'étape d'entrée différentielle d'un amplificateur opérationnel, etc.

4.3. Fiabilité

En terme général, la fiabilité concerne tous les phénomènes qui réduisent ou imposent une limite à la durée de vie d'un dispositif, d'un circuit ou de tout un système électronique. En ce qui concerne les technologies des circuits intégrés étudiées, la fiabilité est affectée par le phénomène de vieillissement des conducteurs métalliques (dû à l'électromigration) et celui des diélectriques minces (dû aux phénomènes de conduction dans les oxydes minces) et par la sensibilité au latchup. Tous ces phénomènes se manifestent plus fortement au fur et à mesure que la température augmente.

Dans les applications à températures normales ($<70^{\circ}\text{C}$), la prédiction de la "durée de vie" ou "évaluation quantitative de la fiabilité" d'un circuit intégré se fait normalement à travers des tests, sous tension, de la dégradation des circuits intégrés accélérée en température ("burning tests"). Ces tests se font généralement à 125°C , pour des applications à température ambiante. Dans le cas des circuits fonctionnant en haute température (entre 125°C et 250°C), les tests de la dégradation accélérée en température pose un bon nombre de problèmes d'ordre pratique. A savoir, d'une part, la différence entre la température maximale de fonctionnement du circuit et celle du *burning test* (différence nécessaire pour l'obtention d'une dégradation accélérée acceptable) peut compromettre le fonctionnement du circuit lors des *burning tests* et, par conséquent, la validité de la prédiction de la durée de vie. D'autre part, la constitution d'un *back-end* (ensemble d'équipements utilisés pour le test), pour le *burning test*, résistant à des températures bien supérieures à 250°C , est possible mais pas évidente. Dans le cas des applications abordées dans cette thèse, les températures maximales d'opération, imposées par le cahier des charges de chaque application, restent inférieures à 175°C , ce qui nous a permis de programmer des tests de fiabilité à des températures inférieures à 250°C .

5. Applications

Deux applications industrielles, représentatives du marché potentiel des applications des circuits intégrés en haute température, ont permis de vérifier dans la pratique les résultats théoriques obtenus.

Dans les prochaines sections, les applications mentionnées sont décrites en langue anglaise, puisqu'elles ont été rédigées à partir d'articles, publiés dans différents journaux internationaux, et à partir des différents rapports techniques rédigés en anglais, pendant cette thèse doctorale. Un résumé en français des résultats les plus importants obtenus est présenté ensuite.

Application dans le secteur pétrolier

Conception d'un Système de Mesure de Pression et Température, intégré en technologie BiCMOS pour des applications pétrolières. Le système de mesure est un circuit mixte, composé de trois oscillateurs analogiques à cristal (avec un contrôle automatique du gain), de mélangeurs, filtres actifs, buffers digitaux et de plusieurs circuits pour la gestion des alimentations. Le système mesure, en permanence et pendant toute la durée de vie des puits pétroliers, la pression et la température au fond des puits, et la transmet à la surface à travers un câble unique. Par conséquent, la stabilité du système est fondamentale. La résolution obtenue est de $\sim 0.0003 \text{ ppm}$ (0.0007 psi , en termes de pression) entre T_{amb} et 220°C dans le système complet. La stabilité à long terme est de 2.5 ppm/an (0.5 psi/mois). Les oscillateurs analogiques ont été testés et caractérisés séparément du reste du circuit mixte, donnant une meilleure résolution ($\sim 0.00008 \text{ ppm}$) entre T_{amb} et 225°C , et un fonctionnement correct jusqu'à 250°C . L'implémentation de ce circuit mixte, intégré dans un MCM-L (avec d'autres IC numériques, et de puissance), a rendu possible une réduction très importante de la taille et du coût de fabrication de l'instrument complet de mesure, avec une amélioration considérable de la performance et de la fiabilité. Le projet a été réalisé en collaboration avec la société SCHLUMBERGER, importante industrie internationale du secteur pétrolier.

Publications:

θ R. A. Bianchi, J.M Karam and B. Courtois, *ALC Crystal Oscillators based Pressure and Temperature Integrated Measurement System for High Temperature Oil Well Applications*, Joint Meeting of the 13th European Frequency and Time Forum and 1999 IEEE International Frequency Control Symposium, Besançon, France, April 1999.

θ R. A. Bianchi, J.M Karam and B. Courtois, *Analog ALC Crystal Oscillators for High Temperature Applications*, to be published in the IEEE Journal of Solid-State Circuits, January 2000.

Application dans le secteur automobile

Conception d'un capteur de température intégré en technologie CMOS pour des applications automobiles. Deux versions ont été fabriquées et testées, l'une avec une sortie analogique et l'autre avec une sortie digitale. La version analogique, a une compensation in-situ de la linéarité et une précision d'environ 0.34°C rms entre -50 et 150°C . La version digitale (avec une double compensation de la linéarité) a une sortie compatible CMOS, du type modulation du rapport cyclique, qui donne une précision supérieure à 0.25°C rms , entre -60 et 160°C . Les dispositifs conçus ont été transférés à l'entreprise MEMSCAP pour leur commercialisation.

Publications:

- θ R. A. Bianchi et al., *CMOS-Compatible Smart Temperature Sensors*, *Microelectronics Journal* 29 (1998), P 627-636.
- θ R. A. Bianchi et al., *CMOS Compatible Temperature Sensor Based on the Lateral Bipolar Transistor for Very Wide Temperature Range Applications*, *Sensors & Actuators: A. Physical*, Vol. 71/1-2, pp. 3-9, Nov. 1998.
- θ R. A. Bianchi, J. M. Karam, B. Courtois, R. Nadal, F. Pressecq, S. Sifflet, *CMOS Compatible Temperature Sensor with Digital Output for Wide Temperature Range Applications*, *THERMINIC 99' Workshop*, Rome, Oct 1999.

5.1. High Temperature Oil Well Application : Miniature Permanent Gauge (MIPEG) project

The objective, submitted by *EPS-CLAMART (Etudes et Production SCHLUMBERGER)*, was to develop for oilfield applications a new generation of down-hole gauge, which achieves a drastic size reduction through electronics integration. The gauge, called a Permanent Quartz Gauge (PQG), is used for oil-wells permanent monitoring. It provides, through a unique cable connection to the surface, a permanent measurement of the static oil-well bottom pressure (~15000 psi) and temperature (-25 to 150°C steady state, 175°C peak) at producing zones, allowing reservoir management. Recent market constraints, related to the gauge installation in the well, require more and more minimising the gauges diameter and length. In the past, the electronic part of the gauge, based on conventional plated through-hole assemblies and expensive military grade components, occupied a large volume inside an expensive protective mechanic housing (see Figure 80).

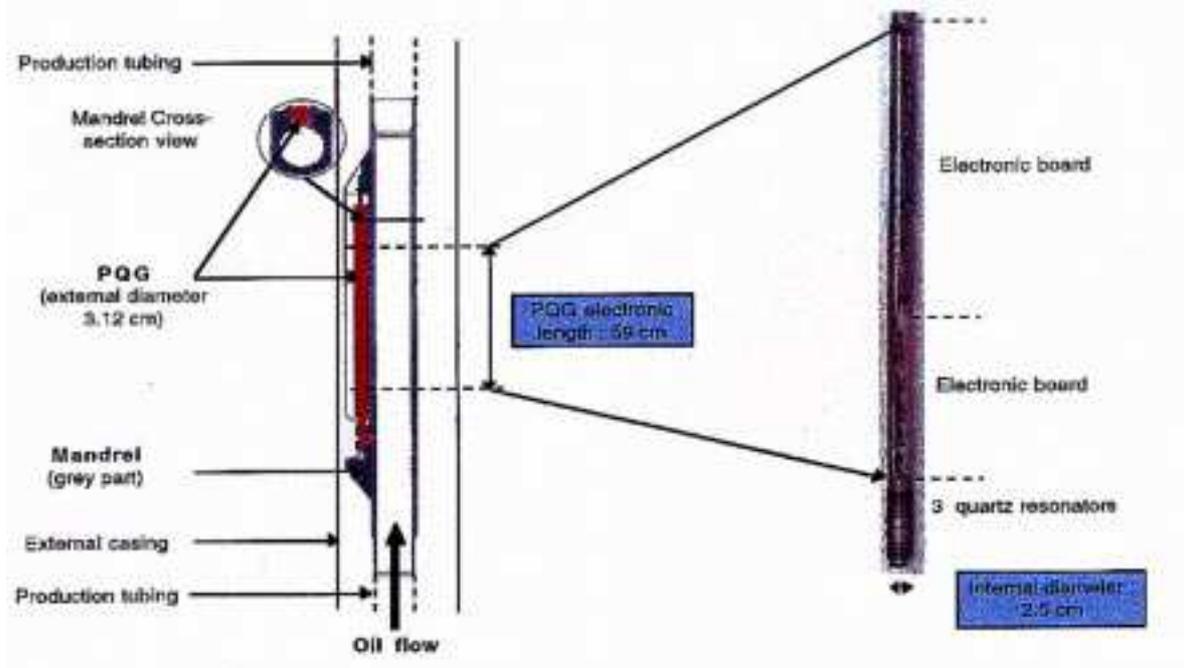


Figure 80. Permanent Quartz Gauge (PQG) measurement tool.

To reach the miniaturization objective an ASIC, named ASIC_MIPEG, has been specially developed, which integrates all the analog functions. Also a direct silicon chips interconnection technique was required for the same purposes (see Figure 81). Thus, the chosen solution was to develop an MCM-L (Multi-Chip-Module) including the already existing dies and the specially developed ASIC_MIPEG.

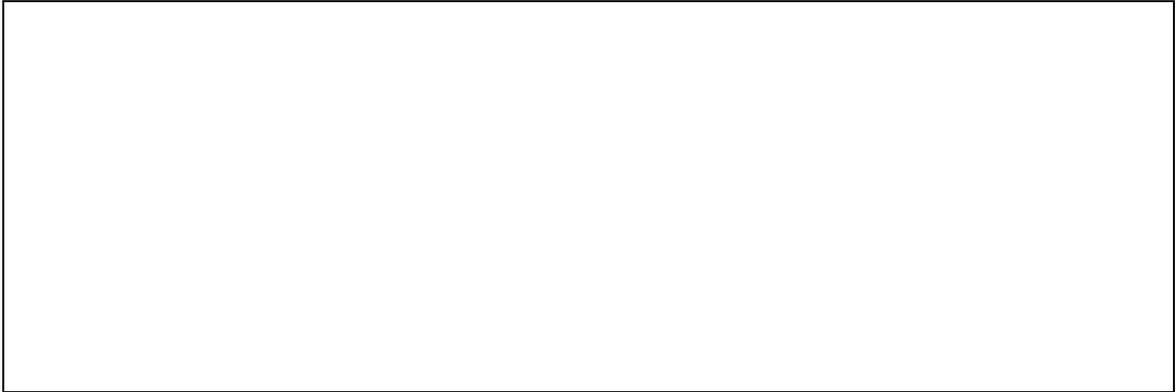


Figure 81. Miniature PERmanent Gauge (MIPEG) size reduction.

The system, integrated in a MCM-L, was partitioned as follows,

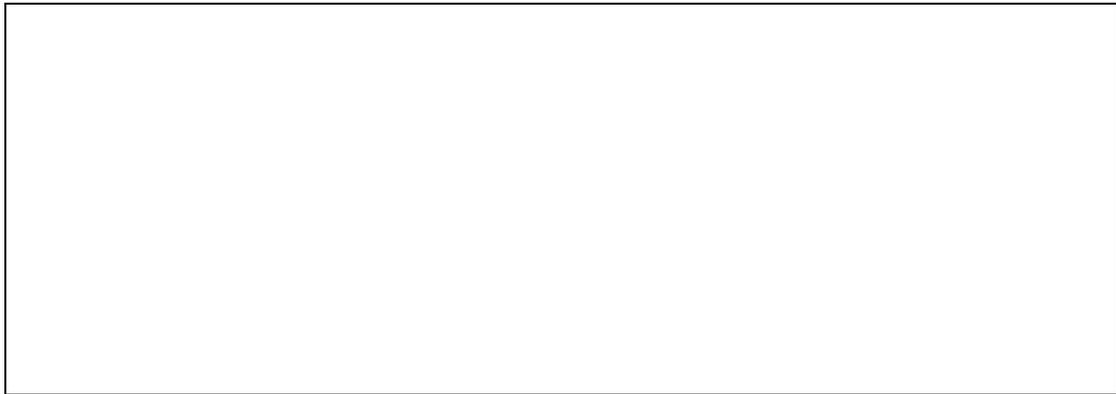


Figure 82. MCM-L layout description

The chips were intentionally placed on the MCM-L in order to divide high frequencies from left to right, and to reduce power dissipation from the right side to the left one.

5.1.1. The ASIC_MIPEG

An application specific integrated circuit for the pressure and temperature measurement at the high temperature oil well bottom conditions, is presented in this section. The system is mainly composed of three high performance Automatic Level Controlled (ALC) oscillator circuits which drive three external crystals (the sensitive elements), and of some mixing and filtering stages. The integrated system was successfully tested up to 220°C, showing a frequency resolution of 0.0003 ppm (0.0007 psi, for the pressure measurement) and a drift of 1.5 Hz/month (0.5 psi/month), concerning respectively the short and long term measurement stability. A drastic reduction of the whole measurement tool size and cost will be allowed by means of this application specific BiCMOS integrated circuit, with an additional increase of the system performance.

As we can observe in the block diagram of Figure 83, the ASIC_MIPEG is mainly composed of :

- ø The sensor interface, that consist of three Automatic Level Controlled (ALC) Crystal Oscillators (the Reference oscillator, Temperature sensitive oscillator and the Pressure sensitive oscillator); two Mixers and two Band Pass Active Filters (which are used to obtain the difference frequency between the reference frequency and the others) and three Digital Buffers to drive the digital outputs.
- ø The communication system, which consists of a Transmitter and a Receiver circuit.
- ø The power supply management circuit, which consists of a Reset detector, a Power Status generator, and a Protection circuit against over current in the cable.
- ø Other elements such as, some $60 \mu A$ pull-up and pull-down transistors, and some 1 Kohms and 5 Kohms integrated resistors, which are not represented in the mentioned diagram.

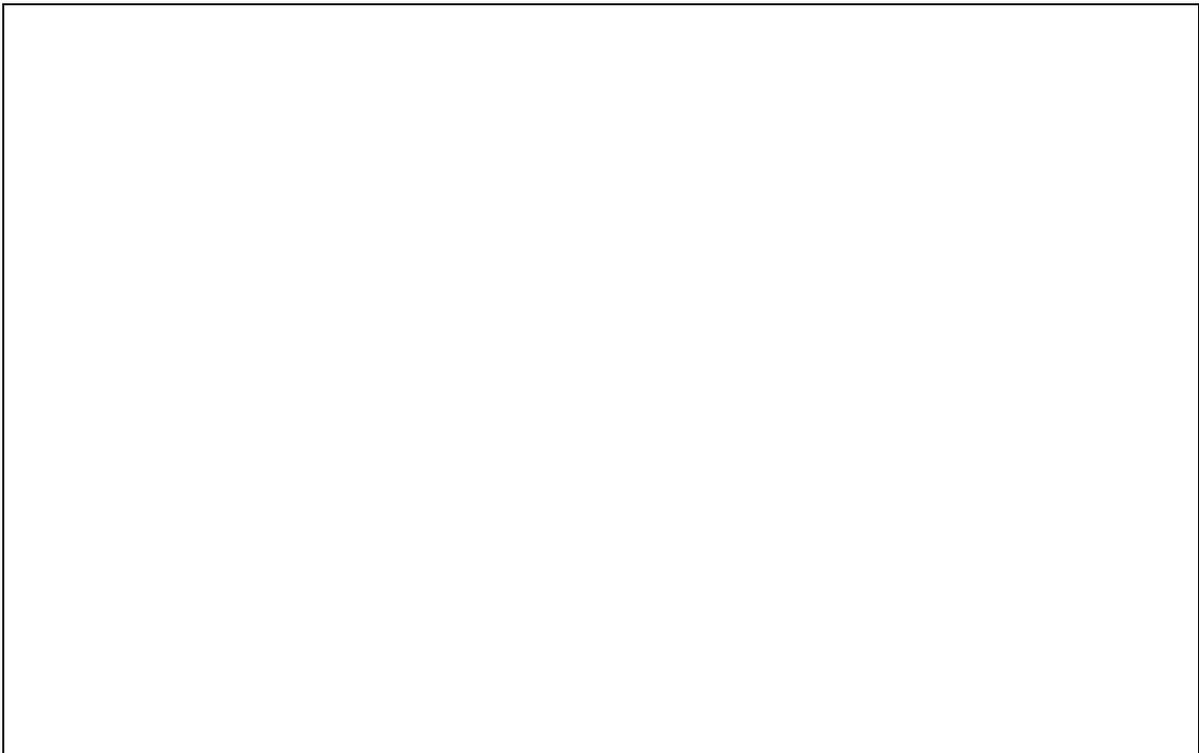


Figure 83. Block diagram of the ASIC_MIPEG

Power is mostly supplied to the ASIC_MIPEG by the 5v voltage regulated source VDD. VDD_prot_ligne provides power supply only to the power control and reset systems which generate RESET_IN, POWER_OK and PROT_LIGNE signals.

The power supplied to the ASIC_MIPEG (only concerning oscillators, mixers, filters and low-frequency digital buffers) can be controlled by the signal SWITCH_N, allowing power consumption savings (the power supply is interrupted in those circuits when SWITCH_N > 4.5v). In order to avoid signal intermodulation through the power supply line due to the on resistance of a single switch, each block was provided of a local switch, all of them controlled by SWITCH_N as already mentioned. Also to avoid the injection of switching noise through the power supply lines the F_REF digital buffer power supply VDD_comp was separated from the oscillator's VDD.

In addition, the Receiver circuit power supply input VDD_receiver was placed separately allowing power consumption savings, when connected to GND, if the Receiver function is not necessary. A circuit, controlled by the signal BOOSTER_N (active when low (<0.5v)), allows the excitation of the pressure quartz oscillator with the Reference frequency, if it is necessary to help it starting up.

The Receiver circuit, integrated in the ASIC_MIPEG, is controlled by the digital signal RST_REC (active when high ($>4.5\text{v}$)) which forces it to the reset condition. REC_DIR is a CMOS output signal at the same frequency than the input signal on the cable ($1.2\text{KHz} \Leftrightarrow 5\text{KHz}$). REC_INT (active when low) detects the presence of a signal from another gauge on the cable.

XMIT is the CMOS digital input of the integrated transmitter circuit. It's frequency ($1.2\text{KHz} \Leftrightarrow 2.4\text{KHz}$) will be forced in the cable by a power transistor. TX_OUT is the CMOS digital output of the integrated transmitter circuit which drives the power transistor.

RESET_IN is a CMOS digital output which is active (5 volts) when VDD goes higher than $4.7 \pm 1\%$ volts, and that falls down (< 0.5 volts) when VDD goes lower than $4.4 \pm 1\%$ volts.

POWER_OK is a CMOS digital output which is active (5volts) when HV (cable signal) goes higher than $9.6 \pm 1\%$ volts, and it falls down (< 0.5 volts) when HV goes lower than $9.2 \pm 1\%$ volts.

PROT_LIGNE is an analog output signal which drives the pass transistor of the power supply pre-regulator, in order to guarantee in HV a permanent level higher than 8.75 volts, for any load or during a short circuit in the system.

F_REF is a CMOS digital output signal that contains the frequency of the Reference Quartz.

The Pressure and Temperature oscillator frequencies are mixed with the Reference oscillator frequency to produce two low-frequency CMOS output signals F_PRES and F_TEMP.

Only the most relevant and important parts of the ASIC_MIPEG will be addressed in detail in the following sections.

5.1.1.1. The sensors interface

This integrated measurement system will provide a permanent measurement of the well pressure and temperature at the well bottom conditions during the whole life of the well. Consequently the stability of the system (i.e. the measurement resolution and particularly the long-term stability) is specially concerned.

The objectives of a drastic size reduction, higher temperature robustness, and better interconnections reliability will be achieved by means of the design of the mentioned BiCMOS Application Specific Integrated Circuit, with an additional increase of the system stability.

The measurement system structure in Figure 83 consists of three integrated oscillator's circuits which drive three different crystals. The Reference one has a reduced temperature sensitivity and is used to generate a stable reference frequency. The Temperature crystal has a quite larger temperature sensitivity which is used for temperature measurement. The external pressure is mechanically transmitted to the Pressure crystal, which is sensitive to pressure and it also has a small temperature sensitivity that must be compensated.

The Reference and Temperature crystals are fundamental mode quartz, but the Pressure crystal is a third harmonic mode one. The generated pressure and temperature signals are mixed with the reference one, filtered and amplified to obtain lower frequency digital signals (F_PRES and F_TEMP) for further digital signal processing.

5.1.1.1.a. Oscillators structure

The general structure of an oscillator is shown in *Figure 84*. It consists of a *Gain Stage* $G(V_p)$ which gain value depends somehow on the resulting oscillation level V_p (passively, because of some amount of non-linearity in the gain stage circuit or in an active way, by means of a control loop); and a *Phase Shift Generator* $\beta(\omega)$, which transfer function depends on the oscillation frequency ω .



Figure 84. General oscillator structure.

The Barkhausen criterion for stable oscillations [10] is given by,

$$(210) \quad \boxed{}$$

Then, the oscillation amplitude dependency of the *Gain Stage* is used to set the desired oscillation level V_p , and the dependence on frequency of the *Phase Shift Generator* is used to select the desired oscillation frequency, setting the required phase shift only at this frequency.

a. A

LC oscillators structure

Both type of oscillators, the fundamental and third harmonic ones, are sinusoidal *Automatic Level Controlled (ALC)* crystal oscillators and have the same general structure. It consists of a *Current Controlled Inverter Gain stage*, a *Phase Shift Generator Circuit* and a *Level Detector & Bias Controller*, as shown in Figure 85. This structure makes the generation of purely sinusoidal signals with a controlled level and a very good frequency stability possible.



Figure 85. ALC oscillator structure.

The need for an ALC structure was determined by several reasons:

- ∅ Oscillations that are limited only by the non-linearity of the gain stage are quite unstable, in the short and in the long term, because of the dependence of this non-linearity on temperature, voltage supply variations, time drift of some transistor characteristics and other not-controlled parameters [53]. Using an ALC structure the influence of most of these parameters on the frequency stability is eliminated.
- ∅ With an ALC structure, the power dissipated in the crystal can be more easily controlled, thus the effects of the power-drive-level sensitivity of crystals and the accelerated ageing due to excessive power dissipation, can be significantly reduced.
- ∅ The level limitation by distortion has also a poor power efficiency, since a significant part of the supplied power is wasted producing harmonics. On the other hand, using an ALC structure pure sinusoidal signals can be generated.

θ In the concerned pressure and temperature measurement system, the generated frequency signals are mixed to obtain lower frequencies which are easier to process, by further *DSP*, etc. Since, to obtain the best performance, most mixers need at their inputs harmonic-free signals with a controlled level, the ALC structure fits very well those measurement systems requirements.

Therefore, in high-performance oscillators, the amount of distortion must be limited fixing the bias point just above the critical condition for oscillation, and the oscillation level must be controlled to help short-term and long-term frequency stability. In the case of low-cost and high-volume integrated circuits design, the high cost and the poor long-term reliability of post-process trimming methods (in order to accurately control the oscillation level) make the implementation of an automatic and integrated level control loop necessary.

b. P

Pierce oscillators

The implemented oscillators are three-point type oscillators, and particularly Pierce type (see Figure 87). The three-point oscillator structure was chosen because, since the crystal acts as a coil, the oscillation frequency will be placed between the shunt and the series resonance frequencies of the crystal (i.e. where the phase shift has the higher frequency sensitivity). As a consequence of this, the pulling capability of the oscillators will be reduced, and their frequency stability will be increased.



Figure 86. Electrical model of crystals

The various possible implementations of the mentioned structure differ on the choice of the AC grounded node. The grounded drain configuration has the advantage of requiring only one additional pin for the crystal connection because the other one is connected to ground. But its major drawback is an increase of $C3$, due to the addition of the pad-ground capacitance and one of the capacitance ($C10$ or $C20$) of the external crystal (see Figure 86), which would degrade the frequency stability (see equations (218) and (219)).

The AC grounded source configuration, also called the Pierce configuration, is therefore preferred for high precision oscillators [53], because pad-ground, $C10$ and $C20$ capacitors add to $C1$ and $C2$ respectively, and not to $C3$. The gain stage was implemented using a single transistor instead of a CMOS inverter biased in its active region. Because of the inherent class AB operation of the inverter, the current increases with the amplitude of oscillation, which creates very strong and undesired non-linear effects, resulting in a very poor frequency stability and a huge waste of power.

MOS transistors were preferred to bipolar ones due to their higher input impedance. Since the used IC technology is N-well type, PMOS transistors placed in a separated N-well were chosen to act as the

gain stage, as their isolated N-well eliminates most of the bulk inter-modulation. In addition, as mentioned in section 3.4.6.2 and in section 2.3.1.1, PMOS transistors has lower leakage currents in the drain and lower hot-carriers injection in the gate, than NMOS ones.

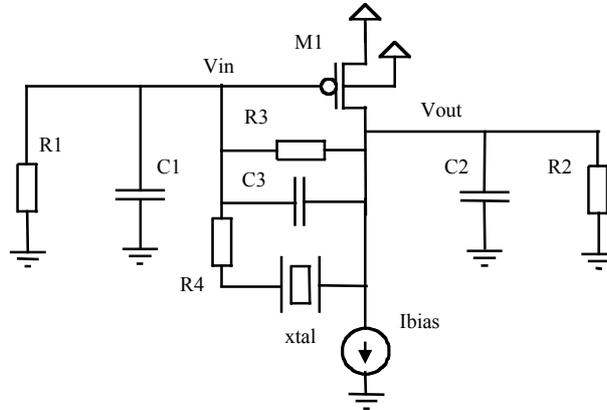


Figure 87. Oscillator equivalent circuit.

The equivalent circuit of the fundamental and the third harmonic oscillators, concerning the desired oscillation frequency ω , is showed in Figure 87. The crystal will be represented by its equivalent electric circuit, shown in Figure 86, where $Ls(i)$, $Cs(i)$, $Rs(i)$ model the resonance mode 'i' of the crystal [10].

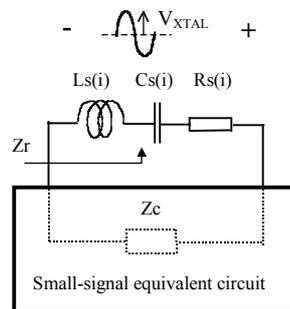


Figure 88. Small-signal linear model

In Figure 87, the total impedance at the input and at the output nodes of the PMOS transistor were represented by $R1$ and $R2$ resistors. $R4$ models the resistance in the wires of the crystal and $R3$ was added for bias purposes. Since all these resistors represent losses in the reactive circuit, which generate the phase shift of the oscillator, their influence must be reduced in order to increase the quality of the reactive circuit and consequently the frequency stability. $C1$ and $C2$ are composed of two integrated Poly-Si capacitors, plus the pad-ground capacitance, plus the $C10$ and $C20$ capacitances of the crystal, respectively. $C3$ is composed of the gate to drain capacitance Cgd capacitance of the PMOS transistor and the $C12$ capacitance of the crystal. The contribution of the mentioned parasitic capacitance to $C1$, $C2$ and $C3$ must be minimised, to ameliorate the long-term frequency stability (see equations (218) and (219)), because of their unknown long-term behaviour.

Considering a linear behaviour of the oscillator circuit for the small-signal case, the whole oscillator circuit can be represented as shown in Figure 88, where Zc represents the impedance (of the small-signal equivalent linear circuit) seen by the selected resonance mode of the crystal [53].

Frequency variations will be expressed in terms of the pulling parameter 'p' as follows,

$$(211) \quad p = \frac{\omega - \omega_s}{\omega_s} \ll 1$$

where ω_s is the series resonance frequency of the selected resonance mode of the crystal, and ω is the resulting oscillation frequency, in rad/sec. Then, the motional impedance Z_r in Figure 88 is given by,

$$(212) \quad Z_r = R_s + j \frac{2p}{\omega C_s}$$

The already mentioned critical condition for oscillations (210) can be simply expressed as [10],

$$(213) \quad Z_c + Z_r = 0$$

Splitting this condition into its real and imaginary components,

$$(214) \quad \text{Re}(Z_c) + R_s = 0$$

$$(215) \quad \text{Im}(Z_c) + \frac{2p}{\omega C_s} = 0$$

These conditions were applied to the circuit defined in Figure 88 to determine the design criteria (concerning $C1$, $C2$, $M1$, I_{bias} , etc.), in order to increase the frequency stability, reduce the start-up time and the power dissipated in the crystals. Even if $R1$, $R2$, $R3$ and $R4$ were carefully taken into account for the design of the oscillators, in the following theoretical analysis they are not going to be considered in order to simplify the analytical expressions.

Power dissipation on crystals

The power dissipated in each crystal is given by the following equation,

$$(216) \quad P_{xtal} = \frac{V_{XTAL}^2}{2|Z_r|} = \frac{V_{XTAL}^2}{2R_s \sqrt{1 + \frac{2p}{\omega C_s R_s}}}$$

V_{XTAL}
'

$$(217) \quad p = \frac{C_s}{2C1C2} \frac{C1C2(C1+C2)}{C1C2+C3(C1+C2)} + gm R_s C3 - \frac{C3(C1+C2)}{C1C2+C3(C1+C2)}$$

where R_s is the motional resistance of the crystal and gm is the transconductance of the PMOS transistor. To reduce the power dissipated in the crystal, V_{XTAL} must be reduced as much as possible, and 'p' must be increased reducing $C1$ and $C2$.

Frequency stability

Since R_s is not precisely known (because of some dispersion in its nominal value, its temperature sensitivity, and the power drive level sensitivity, time drift and ageing of the crystal, etc.) and gm changes following the bias conditions and temperature, the frequency stability of the oscillations will be analysed in terms of the frequency pulling parameter 'p' changes, caused by variations of R_s and gm . From (217) we obtain,

$$(218) \frac{fp}{fRs} = \frac{Cs}{2\omega C1\omega C2} \omega gm\omega C3\omega 1 - \frac{C3\omega(C1+C2)}{C1\omega C2 + C3\omega(C1+C2)}$$

$$(219) \frac{fp}{fgm} = \frac{Cs}{2\omega C1\omega C2} \omega Rs\omega C3\omega 1 - \frac{C3\omega(C1+C2)}{C1\omega C2 + C3\omega(C1+C2)}$$

In order to increase the frequency stability, Cs and $C3$ must be reduced, and $C1$ and $C2$ must be increased.

Minimal start-up time constant

Solving the circuit in Figure 88, the following behaviour of the oscillations envelope is obtained,

$$(220) \quad Vp(t) \propto e^{\frac{t}{T_E}} = e^{-t \frac{(Re(Zc)+Rs)}{2\omega Ls}}$$

Concerning the time necessary to build up oscillations, the minimal possible start-up time-constant can be obtained from (214) using the most negative $Re(Zc)$ value, as follows,

$$(221) \quad T_{UP_MIN} = \frac{-2\omega Ls}{Re(Zc)_{MIN} + Rs} = \frac{2\omega Ls}{\frac{1}{2\omega\omega C3\omega 1 + C3\omega \frac{C1+C2}{C1\omega C2}} - Rs}$$

To reduce oscillators start-up time, $C3$ must be reduced, and $C1=C2$ must be increased.

General design equations

The following equations, used for the design of the oscillators, are applicable to both type of oscillators (see section 5.1.1.1.b.c). Solving the critical condition for oscillation (214), we obtain the following expression for gm_{CR} , the *critical transconductance* of the PMOS transistor,

$$(222) \quad gm_{CR} = \frac{C1\omega C2}{2\omega Rs\omega C3^2} \omega 1 - \sqrt{1 - 2\omega Rs\omega C3\omega \omega 1 + \frac{C3}{C2} + \frac{C3}{C1} \omega 1}$$

The actual and more accurate value of gm_{CR} was obtained with a similar but much larger equation, where $R1$, $R2$, $R3$ and $R4$ were taken into account. However, in the last equation the sensitivity of gm_{CR} , concerning the tolerances and the drift in the values of $C1$, $C2$, $C3$ and Rs , can be more easily determined. By substituting gm in equation (217) with the more accurate expression for gm_{CR} , the critical pulling parameter value can be obtained. Then, its actual sensitivity (i.e. the frequency stability), concerning variations in the values of Cs , $C1$, $C2$, $C3$, $R1$, $R2$, $R3$, $R4$ and Rs , can be also obtained.

The gm_{CR} value and the bias point will define the size of the oscillator transistor, which is always set to its saturated mode of operation,

$$(223) \quad \frac{W}{L} = \frac{gm_{CR}}{KPP\omega(Vdd - Vbias_{SCR} - Vthp)}$$

Where KPP , $Vthp$ and W/L are the PMOS transistor gain factor, threshold voltage and width-length ratio, respectively. $Vdd - Vbias_{SCR} = Vgs(ZTC)$ is the desired steady state operating point for Vgs and Vds

at T_0 (see the Current bias generation section, in section 5.1.1.1.b). Then the current bias for each transistor at the critical condition for oscillation can be calculated as follows,

$$(224) \quad I_{d_{CR}} = \frac{g_{m_{CR}} (V_{dd} - V_{bias_{CR}} - V_{thp})}{2}$$

Similarly, a DC operating point $V_{dd} - V_{bias_{UP}} = 2.5V$ is defined for the beginning of the start-up process. Then the drain to source current $I_{d_{UP}}$ and the transconductance $g_{m_{UP}}$ for the start-up, are calculated as follows,

$$(225) \quad I_{d_{UP}} = \frac{KPP}{2} \cdot \frac{W}{L} (V_{dd} - V_{bias_{UP}} - V_{thp})^2$$

$$(226) \quad g_{m_{UP}} = \sqrt{2 \cdot KPP \cdot \frac{W}{L} \cdot I_{d_{UP}}}$$

Giving a $g_{m_{UP}}$ value at least two times $g_{m_{CR}}$, at T_0 , for the *worst-speed case* of the IC process parameters and for the highest R_s admitted value. In a more typical case the $g_{m_{UP}}$ value is more than four times $g_{m_{CR}}$. The value of $I_{d_{UP}}$ will define the maximal current that the current source I_{bias} will be able to supply, and the value of $I_{d_{CR}}$ will determine the current consumption of the PMOS transistor when the steady state is reached.

Once we know the conditions for start-up, we can calculate its characteristic time-constant, which will determine the order of magnitude of the *AC start-up* process duration. Equations (220), (214) and (226) lead to,

$$(227) \quad T_{UP} = \frac{-2 \cdot L_s}{Re(Z_c)_{UP} + R_s} = \frac{2 \cdot L_s}{\frac{g_{m_{UP}} \cdot C1 \cdot C2}{(g_{m_{UP}} \cdot C3)^2 + \omega^2 \cdot (C1 \cdot C2 + C3 \cdot (C1 + C2))} - R_s}$$

It should be noted that this equation does not take into account the *DC operation points start-up* which is due to other *R-C* time constants of the circuit.

5.1.1.1.b. Implemented circuits

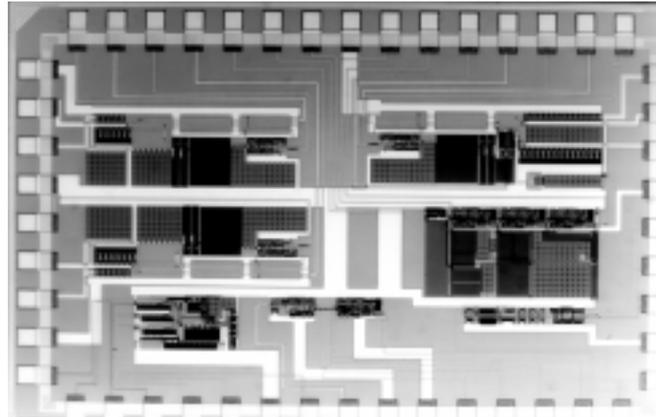


Figure 89. Photomicrograph of the implemented test circuits.

The integrated oscillators were fabricated in a standard $0.8\mu m$ BiCMOS technology. Figure 89 shows a Photomicrograph of the integrated oscillators, the bias generator [50] and other test structures,

- o Upper-left block : 1st harm. Reference oscillator. (1080 x 450 μm)
- o Lower-left block : 1st harm. Temperature oscillator. (1120 x 450 μm)

ø Upper-right block : 3rd harm. Pressure oscillator. (1200 x 470 μm)

The integrated circuit in Figure 89 was fabricated to separately test the performance of each analog circuit.

a. T
emperature robustness

Since temperature acts everywhere in the chip and affects almost every parameter of semiconductor devices, temperature robustness must be studied everywhere, from the system level to the layout level. It concerns the choice of the IC technology, the architecture of the circuit, the implementation of each individual block and the simulation tools (see it in chapter 4).

For the particular case of the ALC oscillators it must be underlined that the performance of the circuit at high temperature mostly relies on the efficient stock of charge in some capacitors, such as $C1$, $C2$, Cg , Cf , Ct , etc. (see next sections) and in the crystal. Thus, their nodes must be connected either to elements with negligible leakage currents or to minimised reversed biased p-n junctions.

b. F
undamental oscillator

In Figure 90 the structure of the fundamental mode oscillator is described. The not-integrated devices (placed outside of the dotted line rectangle) Rg , Cg and $R3$, were used for the test of the dynamic behaviour of the ALC oscillator in different conditions.

Oscillations will not occur at the third, fifth and other oscillation modes of the crystal, which are weaker than the fundamental one, because the gain of the circuit was calculated and limited to avoid exciting these resonance modes. The oscillator output signal should be taken from the Vin node because it has much less amount of distortion than $Vout$ since most of the spurious harmonics are filtered by the crystal.

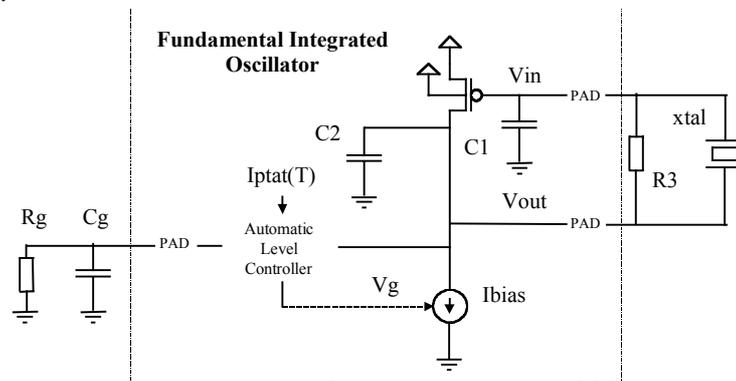
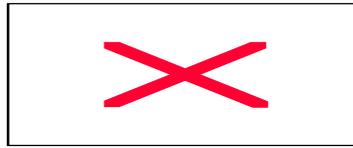


Figure 90. Fundamental oscillator circuit.

c. T
hird harmonic oscillator

Concerning third harmonic oscillator, the above mentioned structure was modified in order to select the desired resonance mode. Another resonant circuit was added (the Lt - Ct circuit in Figure 91) which acts as the appropriated capacitor ($C2$) only at frequencies higher than the mid-value between the first and the third harmonic frequencies. The value of $C2$ corresponding to the oscillation frequency ω can be easily calculated as follows,

(228)



R_t being the resistance in the coil. Thus, replacing the $Lt-Ct$ circuit by $C2$, all the considerations and equations above mentioned are also valid for the third harmonic ALC oscillator case.

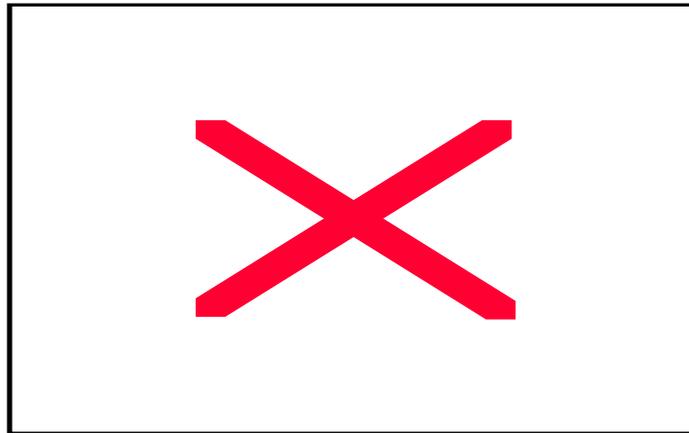


Figure 91. Third harmonic oscillator circuit.

Since at the first harmonic frequency the $Lt-Ct$ circuit does not allow the generation of the necessary phase shift to produce oscillations, the first oscillation mode is then eliminated. In addition, oscillations will not occur at the fifth and other oscillation modes, which are weaker than the third one, because the accurately calculated and limited gain of the circuit is not large enough for those resonance modes. In Figure 91 the structure of the third harmonic oscillator is described. Again, the not-integrated devices R_g , C_g and R_3 , were used to test the response of the ALC oscillator in different conditions.

d. A

LC circuit

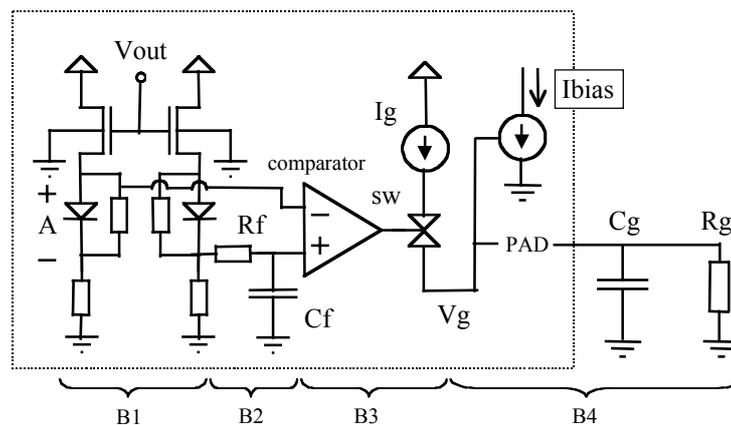


Figure 92. Automatic Level Controller circuit.

The ALC circuit, shown in Figure 92, consists of a *High Input-Impedance Level Shifter Stage (B1)*, a *Reference Level Generator (B2)*, an *Error Comparator (B3)* and the *Control Signal Generator and Voltage Controlled Current Source (B4)*.

Error signal sw generation

In Figure 93 the generation of the *error signal sw* is described. When the oscillation level V_p exceeds the value defined by 'A', an error signal is generated at the output of the comparator, which alternatively switches on and off the current I_g to the external R_g - C_g circuit. As a result of the integration of this current in the R_g - C_g circuit, a very smooth *control signal V_g* is obtained.

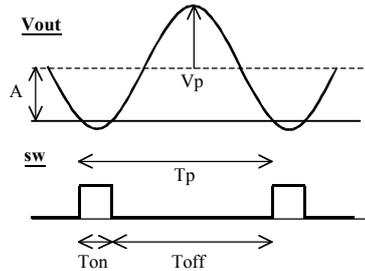


Figure 93. Error Signal generation.

Control signal V_g generation

The transient response of the ALC circuit is shown in Figure 94 for extremely long transients. If $Tp \approx 0.14 \mu sec$ is very small compared to the time constant $R_g \cdot C_g$, the responses are linear and can be described by,

$$(229) \text{ For } sw=0 : \left. \frac{fV_g}{ft} \right|_{t=0} = \frac{V_{g0}}{R_g \cdot C_g}$$

$$(230) \text{ For } sw=1 : \left. \frac{fV_g}{ft} \right|_{t=0} = \frac{I_g}{C_g} \cdot \left(1 - \frac{V_{g0}}{R_g \cdot I_g} \right) \approx \frac{I_g}{C_g}$$

if $R_g \cdot I_g$ is set to be always much larger than the maximal swing of V_g .

Once the steady state value V_{gSS} is reached, the variations of V_g become identical for both values of the error signal sw during each Tp period, as illustrated in Figure 94. Therefore the error of the control system can be expressed in terms of,

(231) X

Note that the error can be reduced setting a $R_g \cdot I_g$ value much larger than the maximal swing of V_g .

The smoothness of the control can be expressed by means of the steady state variations of V_g , as follows,

(232) X

This is the reason why the constant time $R_g \cdot C_g \approx 0.1 sec$ must be much larger than the frequency period Tp . It is important to notice that the smoothness of the control system can be independently trimmed with the external capacitance C_g , and the error of the system can be controlled with the external resistance R_g . Only one external connection, per oscillator, is needed for this issue.

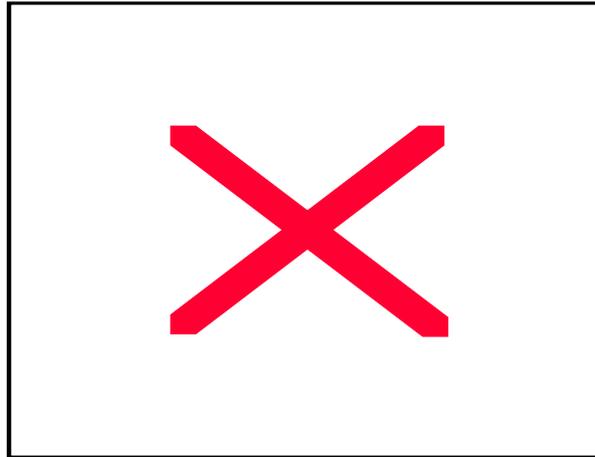


Figure 94. ALC circuit response.

Current bias generation

The gain factor KPP and the threshold voltage V_{thP} , in MOS transistors, decrease when temperature increases. As a result of this, the V_{ds} and V_{gs} bias points and the transconductance g_m of the PMOS transistor in Figure 87 will change following temperature (see equations (234) and (240)). This could have undesired consequences on the ALC loop behaviour due to the limited swing of some signals, such as V_g and particularly the equilibrium point V_{dc} of the oscillations (see equation (240)), which is also affected by the tolerances of the technological parameters KPP and V_{thP} .

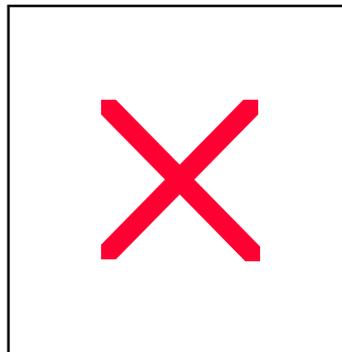


Figure 95. Voltage Controlled Current Source.

More temperature stable bias points and g_m would give to the ALC loop a wider temperature range of operation and a more temperature independent dynamic and steady-state responses. Therefore, the PMOS transistor was biased close to its Zero Temperature Coefficient point $V_{gs}(ZTC) \approx 1.65 \text{volts}$ at the nominal temperature T_0 (150°C), with the following Proportional To the Absolute Temperature (PTAT) current source (see section 4.1.3.1.a),

$$(233) \quad I_{ptat} = I_{ref} \cdot \frac{T}{T_0}$$

The temperature dependency of g_m (see equations (127) and (133)) is given by,

$$(234) \quad \boxed{\text{[Red X]}}$$

Thus, the temperature dependency of gm was significantly reduced to,

$$(235) \quad gm = \sqrt{2 \mu K_{PP}(T_0) \left(\frac{T_0}{T}\right)^{0.42} \frac{W}{L} I_{ref}} \propto T^{-0.21}$$

The *Voltage Controlled Current Source (VCCS)*, in Figure 92, was implemented with the especially designed “*High Input Impedance BiCMOS Translinear Current Mirror*”, shown in Figure 95. The response of the *VCCS* is given by,

$$(236) \quad \boxed{\text{[Red X]}}$$

where ‘ m ’ is a multiplicative factor and $n \cong 0.35$ is a corrective factor added by the PMOS transistors at the inputs, which do not alter the translinear behaviour of the cell. By substituting I_{bias} in equation (234) with equation (236) we obtain the following expression for the transconductance,

$$(237) \quad \boxed{\text{[Red X]}}$$

where gm_{UP} is obtained from equation (226). At the nominal temperature T_0 (150°C) the exponent constant value of gm is about 400mV . Since it is still larger than the expected maximal swing of V_g , the gm equation can be approximated by a linear relation, as follows,

$$(238) \quad gm(T) \cong gm_{UP}(T) \left[1 - \frac{n q \mu g}{4 k T}\right]$$

e. D

ynamic behaviour of the ALC Loop

The dynamic behaviour of the ALC loop must be carefully studied, particularly during start-up, to determine the stability of the oscillator envelope. The ALC loop is a non-linear feedback system, with a moving pole $-1/T_E$ (see equations (220) and (241)) which characterise the behaviour of the oscillator envelope. A circuit-level simulation of the start-up transients (i.e. more than 10^6 oscillation cycles) is not practical, so a numerical solution for the envelope $V_p(t)$ problem is proposed (see Figure 98), taking into account the following considerations.

Figure 96 schematically shows the ALC loop together with its most important time constants. The $R_g.C_g$ error integrator (B4 in Figure 96) responds to the following differential equation,

$$(239) \quad I(t) = \frac{V_g(t)}{R_g} + C_g \frac{dV_g(t)}{dt}$$

Even if the solution for $V_g(t)$ depends on the form of the time-switched current signal $I(t)$, the largest and fastest variations of $V_g(t)$ (see Figure 94) are characterised by the time constant $T_g = R_g.C_g \approx 0.1\text{sec}$. The very large value of T_g allows us to neglect most of the time constants in B5, with the exception of the envelope time constant T_E . Therefore, $V_{dc}(t)$ in Figure 96 can be simply defined by,

(240) 

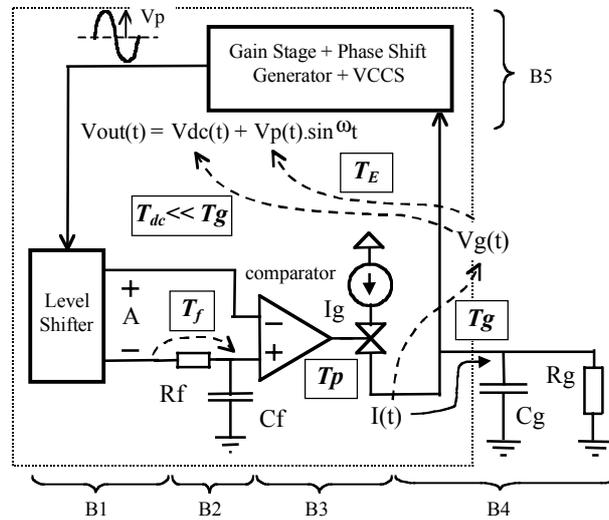


Figure 96. The ALC loop time constants.

The reference level A is generated by the level shifter B1 and the $R_f.C_f$ circuit (B2 in Figure 96) which has a time constant $T_f = R_f.C_f \approx 25\mu\text{sec}$. The value of T_f must be large enough to appropriately filter the oscillation frequency ω , without adding a significant delay at the input of the comparator. If T_f were comparable to the response time of $V_{dc}(t)$ to variations of the error current signal $I(t)$, the equilibrium level $V_{dc}(t)$ could become unstable. Thus, T_f must be much smaller than T_g .

Finally, from equations (220) and (214), the following expression for the envelope time-constant T_e is obtained,

(241) 

Thus, the exponent coefficient of the oscillations envelope $1/T_e$ can be positive, zero or negative, according to the value of $g_m(t)$, leading to growing, stable or damped oscillation, respectively. The minimal positive value of T_e corresponds to the maximal transconductance value g_{mUP} . Since $T_e(g_{mUP}) \approx 1\text{msec}$, and $T_p \approx 0.14\mu\text{sec}$,

(242) 

The last inequality allows us to neglect the switched nature of the error current $I(t)$, because its transitions, having a T_p duration, will not generate transients in $V_{out}(t)$. Thus, the error current $I(t)$ will be described, from Figure 93, as follows,

(243)  if $V_p(t) > A$

(244)  if $V_p(t) \leq A$

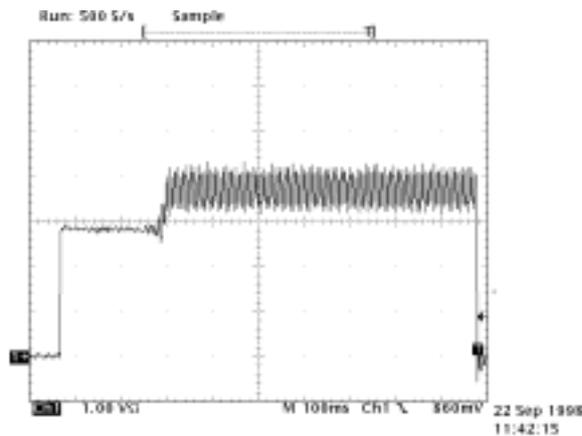


Figure 97. Measured start-up response of the ALC oscillator. (The seeming instabilities of the envelope after start-up are just due to the sampling frequency of the used digital oscilloscope)

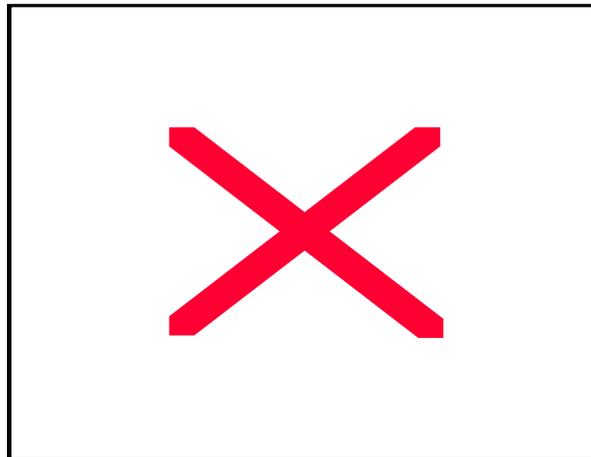


Figure 98. Computed start-up response of the envelope.

No other assumption can be made based on the values of T_e and T_g , because the latter is placed in the trajectory of the former during the start-up of the oscillator. Hence, the dynamic behaviour of the ALC loop was solved computing the equations (220), (237), (239), (240), (241), (243) and (244) with the following initial conditions: $V_g(0)=0$ and $V_p(0)=0$. Using these data, values for R_g and particularly for C_g were chosen to obtain insignificant envelope overshoot (during start-up) and level error ϵ . The final measured and computed start-up responses, of the oscillator signal $V_{out}(t)$, are shown in Figure 97 and in Figure 98 respectively. In both cases, after about $100msec$ AC start-up duration, the oscillation level stabilises with approximately $1.1V_{pp}$ and no amplitude modulation is observed at any temperature.

f. S

teady State analysis of the ALC Loop

We can assume that the steady state values for g_m and V_p are very close to $g_{m_{CR}}$ and A respectively, and we can also consider that the system is reacting linearly against very small perturbations. So, the equations of the different blocks of the ALC loop, can be approximated by linear relations around their steady state values. Some of these linear expressions have been already

described in the last section, so to complete this analysis we need to obtain an expression for the *Error Signal Generator* and the relation between the oscillation level V_p and the transconductance gm .

The behaviour of the *Error Signal Generator* can be approximated considering a triangular input signal, with a peak level V_p , instead of a sinusoidal one. Taking into account the mentioned conditions for the steady state, we obtain,

$$(245) \quad \boxed{\text{X}}$$

The relation between the oscillation level V_p and the transconductance gm depends on how linear the gain stage is. If it is perfectly linear (i.e. the gain is strictly the same for any oscillation level), once the applied gm exceeds the gm_{CR} value the oscillations will grow up to an infinite level. The oscillation level V_p can have finite values only if the gain decreases when the level increases, or if the gain start decreasing beyond a certain value of V_p . Anyway, some amount of non-linearity exists in any gain stage. But the smaller the better for the particular case of the ALC oscillator, because the goal of the ALC circuit is to artificially generate the necessary non-linearity, acting in order to reduce the gain of the oscillator beyond the specified level value 'A'.

As a result of the non-linearity present in any gain stage, the small-signal gain (or the small-signal trans-conductance) must be set to a value slightly larger than the critical gain for oscillations (or the critical trans-conductance gm_{CR}) to obtain a particular oscillation level. This extra small-signal gain value (or small-signal transconductance value) will depend on the desired level value and also on the amount of non-linearity of the gain stage.

To describe the non-linearity of the gain stage we will use the following empirical linear relation between V_p and gm , only valid for small-signal transconductance values slightly larger than gm_{CR} , which is compatible with the already mentioned conditions for the steady state,

$$(246) \quad \boxed{\text{X}}$$

where Qg represents the quality of the gain stage in terms of its linearity (i.e. how independent the gain is, concerning the oscillation level).

Finally, the ALC loop diagram for the steady state is obtained, as shown in Figure 99.

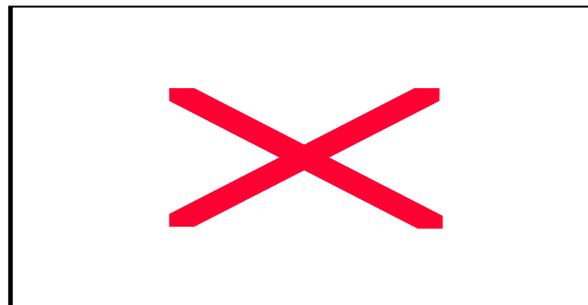


Figure 99. Steady state ALC loop.

The goal of this analysis is to obtain an approximated but useful expression for the steady state error ϵ of the ALC loop, as follows,

$$(247) \quad \boxed{\text{X}}$$

The term inside the square brackets is bigger than unity (see section d). Considering a highly linear gain stage and assuming for the steady state,

$$\begin{array}{l} (248) \quad \boxed{\text{---} \times \text{---}} \\ (249) \quad \boxed{\text{---} \times \text{---}} \end{array}$$

then the steady state error equation can be reduced to,

$$(250) \quad \boxed{\text{---} \times \text{---}}$$

The start-up transconductance gm_{UP} must be larger than the critical one, to guarantee a good start-up for any conditions, but it cannot be too large because of its influence in the steady state error. The temperature dependency of ϵ can be easily reduced replacing I_g by an I_{ptat} current source. Finally, it is important to remark that the steady state error can be controlled setting an appropriated value for $R_g \cdot I_g$, and it can be trimmed by means of the external resistor R_g .

g. M

Measurement results

For the Fundamental Reference Oscillator an SC-cut crystal was used, with its nominal resonance frequency at 7.20 MHz and a small temperature coefficient of about 1.5 Hz/°C. For the Third Harmonic Pressure Sensitive Oscillator an AT-cut crystal was used, which has a nominal third overtone resonance frequency at 7.22 MHz, a temperature coefficient smaller than 55Hz/°C and a pressure sensitivity of about 2.5Hz/psi. The obtained measurement results show good functionality up to the maximal tested temperature, 250°C.

The level constant 'A' was set around 0.5 volts at T_0 (150°C) and it has a small temperature coefficient of about -1.4mV/°C at the mentioned temperature. The generated sinusoidal signals have a very stable level (i.e. no low frequency amplitude modulation was observed) and a quite stable level value up to 200°C, which follows very closely the expected values, as shown in Figure 100. Beyond 200°C and 225°C respectively, the ALC circuit starts losing the control of the oscillation level, as shown in Figure 101.

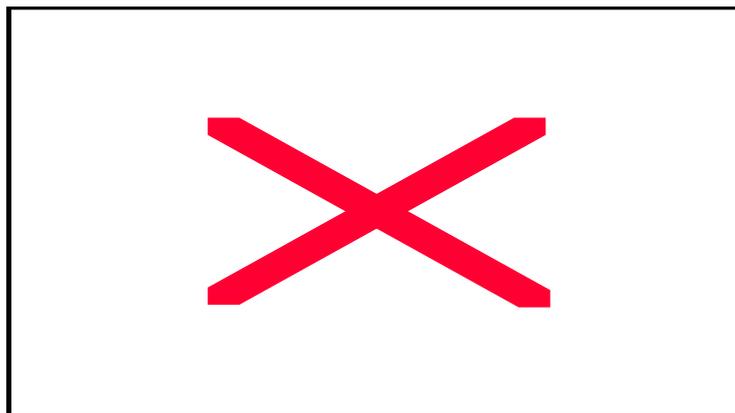


Figure 100. Peak to Peak Oscillation level vs. temperature.

Concerning frequency stability, the short-term (1sec) frequency stability was measured by means of the frequency variations observed with a 12-digit Counter at 1 count/sec rate (see Figure 102). Thus, a short-term frequency stability better than 0.0001ppm/sec was observed between 30°C and 200°C, for both oscillators. At higher temperatures the larger oscillation levels degrade the short-term frequency stability.

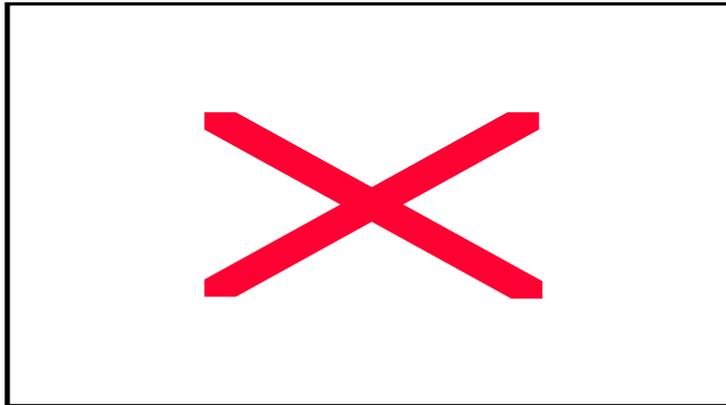


Figure 101. Control Signal V_g vs. temperature.

	Fund. Reference Osc.	3rd harm. Pressure Osc.
Maximal Oscillation Level	1.1 Vpp	1.2 Vpp
Max. Operating Temperature	225°C	200°C
Minimal Voltage Supply Level	4 volts	3.5 volts
Short-term (1sec) Freq. Stability	0.00008 ppm	0.00007 ppm
Maximal 2 nd Harm. Distortion	-60 dB	-53 dB
Phase Noise at 5KHz shift	- 65 dB	- 65 dB
Phase Noise at 50KHz shift	-95 dB	-94 dB
Maximal DC Start-Up time	800 mSec	300 mSec
Maximal AC Start-Up time	300 mSec	150 mSec

Tableau 15. Oscillator's performance

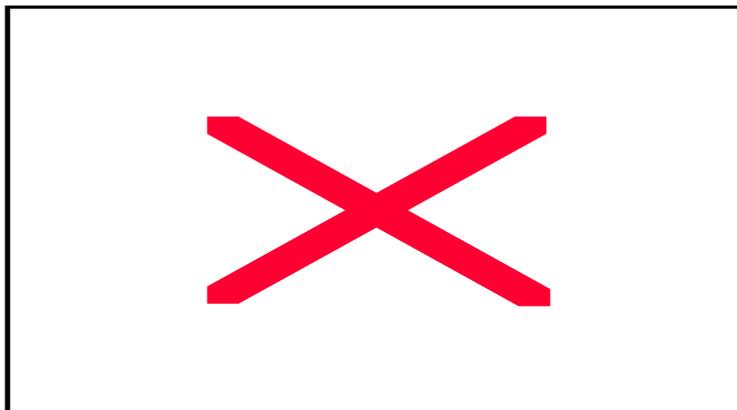


Figure 102. Short-term frequency stability vs. temperature.

The generated sinusoidal signals are almost free of harmonic distortion, as shown in Figure 103 where the ratio between the 2nd and the 1st harmonic is expressed in dB. Note that the 2nd harmonic distortion follows the changes on the oscillation level of Figure 100. Phase noise was measured (see Tableau 15) in terms of the ratio (in dB) between the 1st harmonic and the background noise at 5KHz and 50KHz frequency shifts. The characteristics and measured performance of both types of ALC crystal oscillators are summarised in Tableau 15.

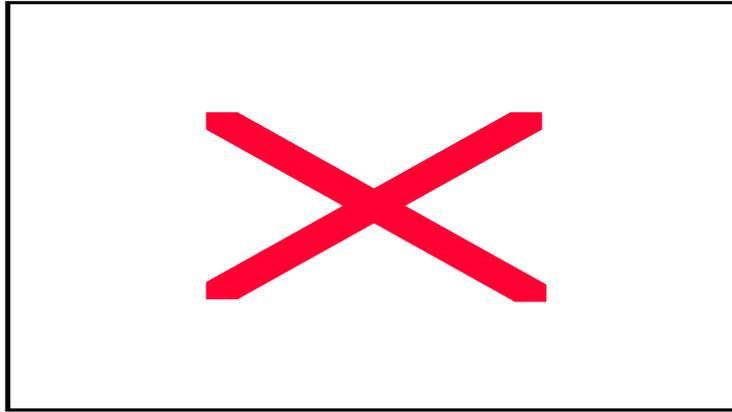


Figure 103. 2nd harmonic distortion vs. temperature.

5.1.1.1.c. Mixer structure

To meet the low-noise and high-linearity requirements for the mixer stage, a *Fully-Balanced Current Mode Switching* mixer structure (see Figure 104) was chosen. It acts as a polarity-switching function of the Temperature and Pressure Oscillator signals in response to the Reference Oscillator signal polarity.

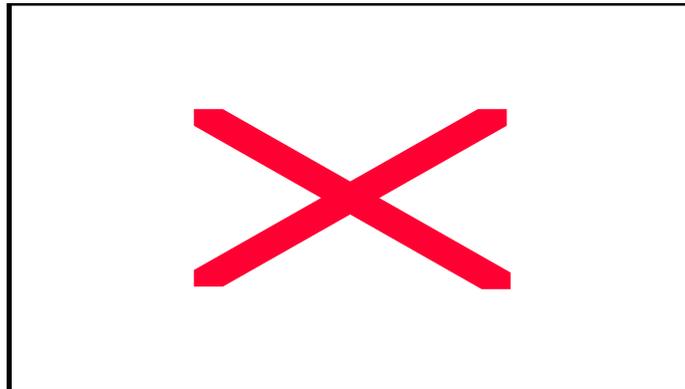


Figure 104. Mixer structure.

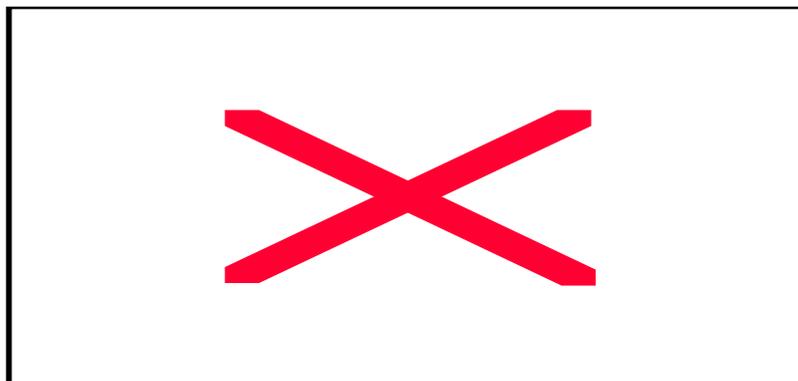


Figure 105. Mixer output frequency spectrum.

Because of the used asymmetric power supply, the polarity switching of signals can be better performed in current mode. So, a high input impedance *Voltage to Current Converter (V/I)*, some *Switches*, and a *Current to Voltage Converter* were implemented for this mixer. To drive those switches a square wave version of the Reference Oscillator signal is then needed. The spectrum of the output signal of this kind of mixers is quite complex, because the product between a sinusoidal signal (TP=Temperature or Pressure oscillator output.) and a unity (+1 ↔ -1) square wave signal (REF), is obtained at the output. The later can be expressed by the following Fourier series,

$$(251) \quad \boxed{\text{Fourier series expansion of a square wave signal}}$$

Thus, the output of the switching mixer will be the product between its S_{TP} input signal,

$$(252) \quad \boxed{\text{Input signal } S_{TP}}$$

and the above series expansion S_{REF} , giving at the output of the mixer,

$$(253) \quad \boxed{\text{Mixer output signal (product of } S_{TP} \text{ and } S_{REF})}$$

If both mixed frequencies are relatively close, the mixer output signal can be described with the spectrum diagram of *Figure 105*, where the desired output signal is,

$$(254) \quad \boxed{\text{Spectrum diagram of the mixer output signal}}$$

So, the higher frequency components must be filtered. If the Temperature or Pressure Oscillator signals have some amount of distortion, specially because of some amount of undesired cross-talk (ω_{CT}) with other signals, low inter-modulation frequencies can be produced in the mixer through the following terms,

$$(255) \quad \boxed{\text{Inter-modulation frequencies produced in the mixer}}$$

which could generate slow variations in the DC level of the mixer output signal. If this DC level modulation is not eliminated, it will generate jitter noise in the comparator output. Special filtering is thus needed also for low frequencies.

5.1.1.1.d. Filters and Buffers

In order to solve the above mentioned problems produced by the high and low intermodulation frequencies, and by the inaccurate DC level present at the mixer output signal, the three stages Band Pass filter (10 to 70KHz) with DC level correction of *Figure 106*, was designed to obtain,

- θ Gain at 10KHz ≅ 20 dB
- θ Gain at 70KHz ≅ 20 dB
- θ Attenuation at ω_{REF} ≅ 100 dB
- θ Attenuation at 10Hz ≅ 40 dB

The first stage is a 2^{nd} Order Active Low-Pass Active Filter specially designed to eliminate the strong ω_{REF} component in the output signal of the mixer. The transfer function of this stage is given by,

(256) 

Which has a $6dB$ Bandwidth of $110KHz$ and an Attenuation slope of $40 dB/dec$. The output DC level is the same of the mixer output signal.

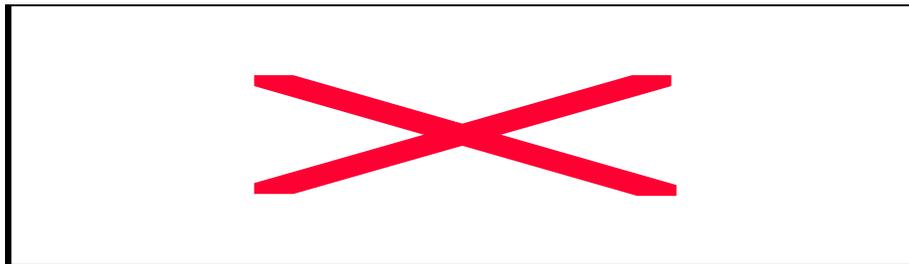


Figure 106. Filter structure

The second stage is a 1^{st} Order Active High-Pass Active Filter and DC level generator specially designed to eliminate the low frequency inter-modulation products and to force the desired DC level $Vdc_reference$. This DC level will be used by the following gain stage and also by the comparator. The transfer function of this stage is given by,

(257) 

Which has an initial zero, a single pole at $10KHz$ that defines its $3dB$ Bandwidth and a slope of $20 dB/dec$. The higher pole ($1/R2.C2 \sim 300KHz$) was added to increase the high frequency rejection and the stability of this stage. The last stage is a Gain stage which pre-amplifies the filtered signal before comparing it to the generated DC level $Vdc_reference$. The gain of this stage is about 20. An additional pole ($\sim 300KHz$) was also added to increase the high frequency rejection and the stability of this stage. Finally, a Comparator provided with about $100mV$ of hysteresis generates the output digital signal.

5.1.1.1.e. Operational Amplifier

The BiCMOS OPAMP, developed for the implementation of the mentioned active filters, is composed of a differential CMOS input stage, a push-pull CMOS gain stage with a Miller frequency compensation capacitor, and a push-pull BiCMOS output stage for current supplying. This structure was chosen due to its very small systematic offset, and the stability of the offset and phase margin over a wide temperature range provided by the PTAT bias current.

<i>Simulated Characteristics</i>	<i>25°C</i>	<i>150°C</i>
Input signal swing	~0.25 to 5 volts	~0.23 to 5 volts
Output signal swing	0 to ~4.35 volts	0 to ~4.3 volts
DC Gain	103 dB	99 dB
Gain at 100KHz	39 dB	38 dB.

Gain-Bandwidth	8 MHz	8 MHz
Phase Margin	60.5°	61°
Gain Margin	25.4 dB	25 dB
Max. Current consumption	100μA	140μA

Tableau 16. Simulated characteristics of the Operational Amplifier.

5.1.1.1.f. Fabricated circuit and measured results

Many layout techniques were considered in order to minimise the high temperature leakage currents [62] [63], to prevent from latch-up occurrences at high temperatures [13] [61], and to reduce the cross-talk between the interconnection layers and through the substrate. For instance, the oscillator's transistors were placed the closest possible to their I/O pads; the power dissipating devices and the digital circuits were placed as far away as possible from the oscillators; and the PTAT bias generator was located in the middle of the circuit.

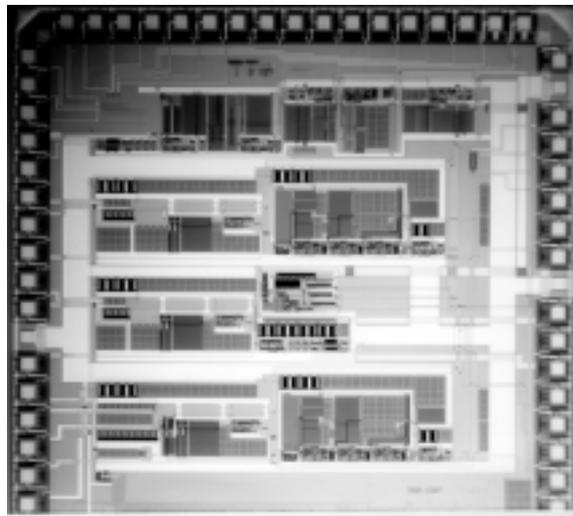


Figure 107. Photomicrograph of the ASIC_MPEG.

The whole integrated system, was fabricated in a standard 0.8μm BiCMOS technology. The Figure 107 shows the Photomicrograph of the fabricated and tested final version (3815 x 3815 μm).

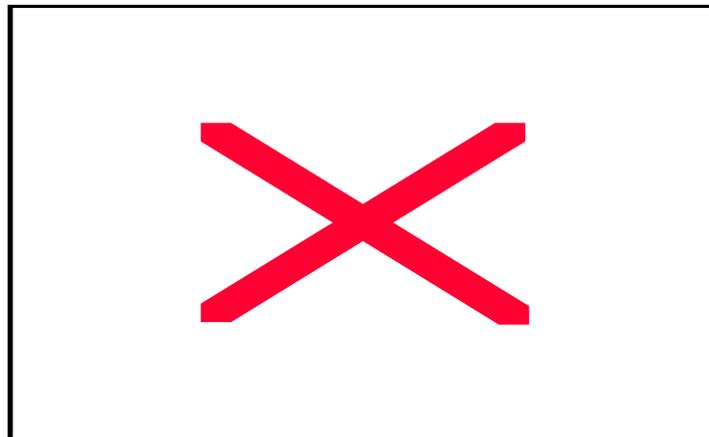


Figure 108. Long term frequency stability

Concerning frequency stability, the short-term (about one second) frequency stability was measured by means of the frequency resolution observed with a 12-digit Counter at 1 count/sec rate. Due to crosstalk and bulk inter-modulation the whole system (mixed-signal integrated system) has a slightly poorer resolution, the analog signals are noisier and contains a larger 2nd harmonic distortion. However, the tested system performance (see Tableau 17) fits largely in the required specifications, showing a very good behaviour, in terms of frequency stability, from ambient temperature to 220°C.

Concerning the long-term frequency stability, we can see in Figure 108 the output signal F_PRES frequency, measured every 10 minutes during almost two months at 185°C and ambient pressure conditions. The temperature variations in the oven were measured in order to correct the small variations of the resulting frequencies, due to the temperature sensitivity of the used crystals. This procedure did not intend to eliminate the influence of the fast (some minutes) temperature variations in the oven but only eliminate the effect of slow temperature drifts (some hours) in order to exclusively evaluate the long-term (several weeks) stability. This is the reason why the corrected plot in Figure 108 looks noisy, if compared to the mentioned short-term frequency stability. The atmospheric pressure variations were also measured and used to correct the frequency variations due to the pressure sensitivity of the crystals.

	<u>Pressure Measurement System</u>
Power Supply Voltage	5 Volts
Maximal Temperature	220°C
Resolution (1/Sec)	0.0003 ppm, 0.0007 psi
Long Term Drift	1.5 Hz/month, 0.5 psi/month

Tableau 17. Pressure measurement system performance

Thus, with stabilised thermal and pressure conditions, the observed drift (see Figure 108) is only due to the long-term drift of the passive devices (i.e. the external coil, the external resistance, the integrated capacitors, etc. Note that the contribution of the parasitic capacitance, mentioned in section 5.1.1.1.a.b, must also be considered), and the characteristics of the crystals.

The reliability tests, showed in Figure 108, were continued during more than one year and at even higher temperatures, showing in general terms an acceptable lifetime.

5.2. Automotive application

Today's customers demand automobiles offering a wide variety of new safety, comfort and performance features. The integration of this features into cars of all categories requires high performance, under harsh environments (i.e. high temperature, strong vibrations, etc.), at low cost. Innovative integrated circuits and microsystems play a key role in satisfying this requirement. It is estimated that by the year 2000, approximately 25% of a car's value will consist of microelectronics or complex microsystems. The automotive industry presents by far the largest potential market for high temperature electronics. Components for the automotive industry are often specified to operate up to 125°C, with the exception of engine management, anti-lock brakes, and power steering applications which have an operational temperature range up to 180°C.

The automotive industry seems to be the main potential market for integrated temperature sensors in the -50 to 150° temperature range, because of the normally big volume of production of this industry. In addition, temperature sensors are needed for a quite large number of applications, such as, tire pressure control, fuel injection, navigation systems, air and fuel admission, heating and air conditioning control, pollutant emissions control, etc. Even if their characteristics and requirements

varies widely following each particular application, important requirements such as low-cost, high-volume, long term reliability and stability, EM noise immunity, and a robust digital interface or communication system, are common to these temperature sensor applications.

The following temperature sensors were designed to fit with most of the mentioned automotive applications requirements.

5.2.1. CMOS Compatible Temperature Sensors

Temperature is an important parameter for many systems and processes, which usually needs to be sensed, limited and/or controlled. Many physical, mechanical and chemical phenomena show an undesired thermal cross-sensitivity. Temperature influences cannot be shielded and consequently they must be measured and compensated; hence, accurate temperature information must be provided by a temperature sensor, having a low economical impact on the cost of the whole system. This thermal sensitivity is also often used to measure other physical quantities (such as flow, radiation, gas pressure, gas type, true RMS values, humidity, heat of chemical reactions, etc.) by means of a thermal translation. Thus, temperature sensors play an important role in many sensing and other microsystems applications. The main characterising properties of a temperature sensor are:

- ∅ Accuracy versus temperature.
- ∅ Sensitivity.
- ∅ Linearity.
- ∅ Temperature range.
- ∅ Long term stability.
- ∅ Low cost.
- ∅ Signal level.
- ∅ Noise immunity.
- ∅ Low power consumption.
- ∅ Absolute calibration.
- ∅ Independence to process tolerances.

Accuracy (within its associated temperature range) is the most important characteristic of temperature sensors in most applications. It will be assured by the good long term stability of the most important technological parameters and by an accurate calibration or trimming procedures, if the used sensing principle is not independent enough concerning process tolerances.

Good sensitivity, linearity, signal level and noise immunity will help and simplify the design of the sensor interface and the digital interfaces. Low power consumption is very important mainly in portable, low voltage and DfTT (Design for Thermal Testability) applications.

The low-cost high-volume requirements demands not only a reduced sensor interface area and standard technologies but also a reduced and robust digital interface and digital signal processing, integrated together with the sensitive element. Since CMOS is still the most extensively used technology, the integration of temperature sensors in high performance CMOS processes is preferred in order to allow signal conditioning and digital signal processing on the same chip (smart sensors).

Other important properties for the next generations of smart temperature sensors are the *adaptability* (bias of amplifiers adapted to the sensor signal, in order to reduce the input amplifier noise; control of the rate of data acquisition, control of power consumption, etc.), and the use of *dynamic ranges* or *autoranging* in order to adapt the sensor sensitivity to the input signal level [38].

5.2.1.1. The possible sensitive elements

The main idea is to find, in standard integrated technologies, a sensing principle which guarantees good linearity and stability over the targeted temperature range. Although both MOS and bipolar transistors have properties that could be used for temperature sensing, the latter has shown much better performance for this purpose [51]. As already mentioned, since CMOS is still the most extensively used technology, the integration of temperature sensors in high performance CMOS technologies is preferred, in order to allow signal conditioning and digital processing on the same chip. In CMOS

technologies there are also available parasitic bipolar transistors, such as the *CMOS Vertical Bipolar Transistor (CVBT)* or the *CMOS Lateral Bipolar Transistor (CLBT)*, which allow an easier way to achieve much better performance.

5.2.1.1.a. Resistors.

Integrated resistors have all an important thermal sensitivity (from 0.05 to 1%/°K at 27°C), but in general they do not show good linearity for high accurate applications. Some of them, such as well, diffused or implanted resistors, have important and sometimes prohibitive leakage currents at high temperatures due to their reverse biased junction isolation.

Polysilicon resistors show excellent long-term stability, especially when passivated with Si_3Na_4 to inhibit impurity migrations and when implanted with boron to avoid carrier segregation. Lightly doped (high resistive) polysilicon resistors show an exponential temperature dependency with a negative temperature coefficient. But highly doped (low resistive) polysilicon resistors follow Ohm's law closely at low electrical fields and show a positive temperature coefficient. A highly linear temperature sensor, with an accuracy of about 2.5°C between -60 and 180°C, using this kind of resistor has been reported [8].

5.2.1.1.b. Thermal diffusion constant of silicon.

The principle is to measure the internal thermal diffusion constant of a single crystal silicon. This diffusion constant is an intrinsic material parameter whose value is fairly independent of process tolerances and has no time instability. It shows a reasonably large temperature sensitivity (-0.57 %/°C) allowing to obtain a temperature resolution of 2%/FS. A thermal delay line integrated in a thermal feedback oscillator using this principle has been reported [42].

5.2.1.1.c. MOS transistors.

To obtain a linear relation between temperature and electrical MOS parameters, it is necessary to use the MOST in weak inversion [43]. The performance of sensors using this principle is mainly limited by leakage currents occurring at high temperatures and the poor reproducibility and large tolerance on the threshold voltages of the MOS transistors, making necessary extensive trimming and calibration operations when high performance is required. Due to the poor long-term stability of MOS transistor parameters, it will be necessary to repeat these calibration procedures after some time, which is highly undesirable.

5.2.1.1.d. Bipolar transistors and diodes.

Silicon diodes and bipolar transistor base-emitter junction potential V_{be} have a property that can be exploited to produce a voltage that is proportional to temperature. The combination of two such voltages yields a voltage that is Proportional To the Absolute Temperature (PTAT). If the transistor and the sensing interface circuit, both monolithically integrated on the same chip, are designed to minimise other temperature effects (such as leakage currents, self-heating, etc.), these voltages will be a highly accurate representation of the substrate temperature over a wide range.

Bipolar devices are then preferred as thermal sensitive elements, and most of the IC temperature sensor have been traditionally designed for bipolar technologies [44] [39] [37] [36] [47] [48] [49]. Because of the additional above mentioned advantages of the CMOS technology, many of those

design have migrated from bipolar to CMOS technologies, using the CLBT [50] [51] [46], or the CVBT [45]. Only for special applications such as DfTT, MOS based temperature sensors in CMOS technologies have been reported (strong inversion [42], weak inversion [43]).

5.2.1.2. Temperature sensor with an analog output signal

A CMOS process compatible wide range temperature sensor, that takes advantage of the lateral bipolar transistor, will be described. Concerning accuracy, a temperature error of 0.34°C rms in current mode (with an on-chip 2nd order temperature compensation) and of 1.86°C rms in voltage mode are the measured performance, without post-fabrication trimming, of this integrated sensor, over the -50°C to 150°C temperature range.

Other important characteristics are the low cost, the less than 1 mW power consumption, the higher than 40 dB PSRR, the relatively small surface, and the output signal swing which is intrinsically referenced to the temperature range, being specially conditioned for further analog to digital conversion in both, current and voltage modes.

The characteristics of this sensor make it especially suitable for low-cost high-volume integrated micro-systems over a wide range of fields, such as automotive, aerospace, oil prospect, etc.

5.2.1.2.a. The sensing principle

Bipolar transistors and diodes have proven to be very interesting devices for temperature sensors applications, especially when low cost, excellent long-term stability and high sensitivity, over a limited temperature range, are required. Their favourable properties are due to the highly predictable and time-independent way in which the base-emitter potential V_{be} in transistors and the forward voltage in diodes are related to the temperature [40]. In addition, their excellent fabrication compatibility, their small dimensions and their low power consumption make them very attractive to be used as temperature sensors, in many applications. It must be mentioned that they also suffer from some imperfections such as self-heating, fabrication tolerances, small drift after thermal cycling and some non-linearity [40].

For temperature sensing, calibrated PN junctions are normally biased by a controlled current, with the forward voltage serving as a measure of temperature. The total diode current consists of not only a diffusion current but also generation-recombination currents in the depletion region and at the surface regions. But these currents, having a different temperature behaviour than purely diffusion currents, can be extracted via the base in a bipolar transistor, so that the collector current of a forward biased transistor is an almost pure diffusion current [9] that gives a better behaviour in temperature sensing applications.

Bipolar transistor base-emitter junction potential V_{be} , when properly biased, can be exploited to produce a voltage which is proportional to temperature (see section 4.1.3.1.b). The combination of two of such voltages yields a voltage V_{ptat} that is Proportional To the Absolute Temperature (PTAT) (see section 4.1.3.1.a). If the transistors and the sensing interface circuit, both monolithically integrated on the same chip, are designed to minimise other temperature effects (such as the temperature variations of the threshold voltage and mobility in MOS transistors, and the increasing leakage currents (see section 3.4)), these voltages will be a highly accurate representation of the temperature over a wide range. In CMOS processes, some parasitic bipolar transistors are available, such as the *CMOS Vertical Bipolar Transistor (CVBT)* or the *CMOS Lateral Bipolar Transistor (CLBT)* (see section 3.5.2). Particularly, the CLBT is an excellent candidate to be used as the temperature sensitive element.

V

be signal

The base-emitter voltage equation of a PNP CLBT biased with a collector current defined by $I_c = I_c(T_0) \cdot [T/T_0]^\alpha$ is expressed by the following second order approximation [52] (see section 4.1.3.1.b),

(258)


where,

(259)
$$V_{orig} = V_{g0} + \frac{a \cdot b \cdot T_0^2}{(T_0 + b)^2} + (\gamma - \alpha) \cdot VT_0$$

(260)
$$ST = V_{g0} - V_{be0} + a \cdot T_0 \cdot \left[1 - \frac{b^2}{(T_0 + b)^2} \right] + (\gamma - \alpha) \cdot VT_0 \cdot \frac{1}{T_0}$$

(261)
$$QT = \frac{a \cdot b^2 \cdot T_0^2}{(T_0 + b)^3} + \frac{(\gamma - \alpha)}{2} \cdot VT_0$$

Where $VT_0 = k \cdot T_0 / q$, T_0 (Mid-range temperature value), V_{g0} (silicon bandgap voltage at T_0), V_{be0} (the base-emitter voltage at T_0), a , b (from silicon bandgap potential equation (1)), α (bias current temperature exponent) are temperature independent parameters.

The value of V_{orig} , is mainly determined by temperature independent parameters (T_0 , V_{g0} , a , b and α), is independent of transistor geometry, and is almost independent of process parameters (only the mobility temperature exponent in γ slightly varies, due to the base doping concentration tolerances).

The base-emitter voltage V_{be0} at T_0 is affected by process parameter tolerances, and it is not possible to predict accurately [9] [40] [39] [37]. However, this will only slightly affect the slope value ST , altering neither the origin value V_{orig} , nor the quadratic coefficient QT . So linearity will not be affected by spreading in technology process parameters [52].

The slope ST of $V_{be}(T)$ can be adjusted, by means of V_{be0} , modifying the transistor geometry and the collector bias current. The quadratic term coefficient QT is responsible for most of the non-linearity of V_{be} . Theoretical calculations showed that this term is minimal when α is positive and close to 3, for the studied technology [52].

The circuit described in section 4.1.3.1.b generates a current mode version of this signal.

P

TAT signal

The difference V_{ptat} between the base-emitter potential of two matched transistors, with different current densities, is Proportional To Absolute Temperature (PTAT), depends on the ratio of the current densities, and is independent of technology parameters (see section 4.1.3.1.a),

(262)


The circuit described in section 4.1.3.1.a generates a current mode version of this signal.

I

intrinsically referenced output signal

Since a PTAT sensor has a very large initial offset at common temperatures, it would be advantageous to have an output signal with its zero T_z at a temperature close to the beginning of the

range of interest, instead of $0^\circ K$, in order to simplify the on-chip A-D conversion. This can be achieved with a combination of V_{ptat} and V_{be} curves.

For this kind of signal processing the use of currents derived from V_{ptat} and V_{be} is preferred (see Figure 109). Voltage signals are transformed into current signals by means of Poly-Si resistors. From equation (84),

$$(263) \quad \boxed{\text{[Red X]}}$$

where α_{T0} is the temperature coefficient of the Poly-Si resistors at $T0$. To avoid dealing with resistor temperature dependencies, it is also preferred to obtain resistor ratio dependent relations, as follows (for voltage mode),

$$(264) \quad \boxed{\text{[Red X]}}$$

replacing V_{be} and V_{ptat} , from equations (258) and (262), in (263), we obtain,

$$(265) \quad \boxed{\text{[Red X]}}$$

$$(266) \quad \boxed{\text{[Red X]}}$$

$$(267) \quad \boxed{\text{[Red X]}}$$

Where the slope S_{out} of $V_{out}(T)$ is the addition of both slopes, giving a better temperature sensitivity. We can obtain almost every desired value of T_z simply by changing the resistor ratio $R2/R1$. Both, T_z and S_{out} are independent of Poly-Si resistor temperature coefficients because they depend on resistor ratios. The non-linearity of the output signal is almost completely contained in the quadratic term V_{qt} [52].

$I_t(T)$

$$(268) \quad \boxed{\text{[Red X]}}$$

replacing $V_{be}(T)$, $V_{ptat}(T)$, $R1(T)$ and $R2(T)$, distributing terms and renaming the resulting coefficients, we obtain,

$$(269) \quad \boxed{\text{[Red X]}}$$

Here, the temperature dependency of the Poly-Si resistors will slightly affect the output linearity of $I_t(T)$, which is expressed in terms of its quadratic term coefficient,

$$(270) \quad \boxed{\text{[Red X]}}$$

The second term of the above described equation is responsible of the 2nd order temperature compensation used for the generation of this current mode temperature signal.

5.2.1.2.b. Implemented Circuit

Since, concerning the high temperature behaviour, some interesting transistor configurations are in current mode, current mode signals were used for signal processing. In addition current mode circuits have almost no signal swing limitations and they have an intrinsically good PSRR. The circuit implemented (Figure 109) consists of two current source generators, I_{ptat} and I_{be} that are combined in order to obtain a current source that is referenced to the origin of the targeted temperature range.

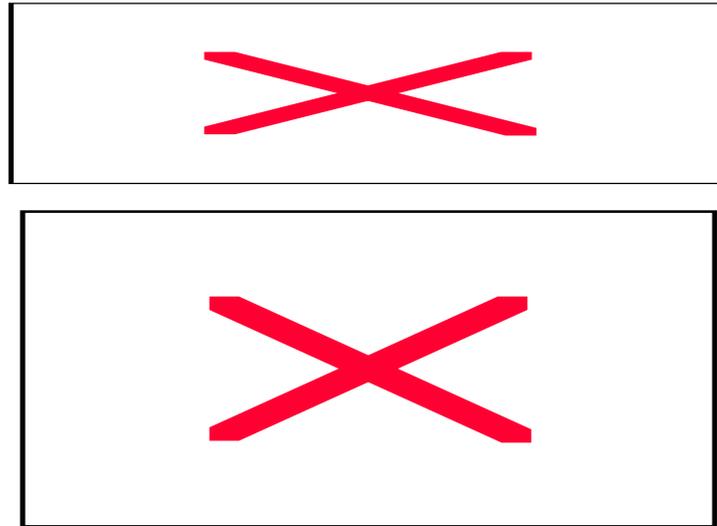


Figure 109. Current sources characteristics

The I_{ptat} and I_{be} current sources were implemented with the circuits described in section 4.1.3.1.a and in section 4.1.3.1.b respectively.

C

urrent adder and output stage

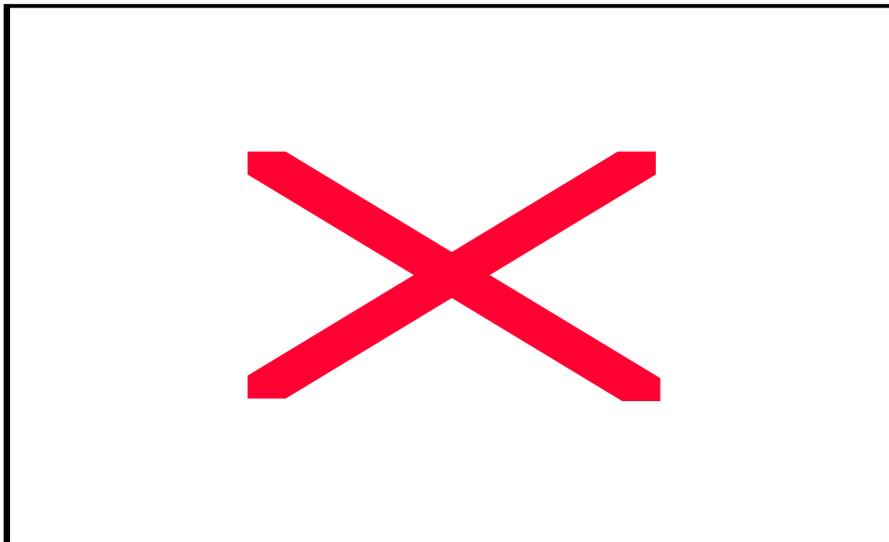


Figure 110. Current adder and output buffer.

The Figure 110 shows the circuit that combines both current sources, to obtain an output signal that is referenced to the origin of the temperature range, in voltage or current mode [52]. The voltage mode output signal, that was performed adding the resistor $R3$ in the output, has an output swing from 0 to 3.5 volts. Here, the influence of the temperature dependency of Poly-Si resistors can be neglected, because the output voltage depends on resistor ratios.

The current mode output signal is obtained by removing the resistor $R3$. It has an output swing from 13 to 90 μA . The influence of the temperature dependency of Poly-Si resistors plays an important role in the temperature behaviour of the output current, because it generates the mentioned 2nd order temperature compensation.

5.2.1.2.c. Measurement results

The Figure 111 shows the Photomicrograph of the integrated temperature sensor ($290 \times 400 \mu\text{m}^2$), and additional test structures, that was fabricated in a standard $0.8 \mu\text{m}$ CMOS technology. Test procedures were performed with a double chamber stable temperature generator using the CNES /TRS31 facilities. The resolution reached at the DUT level is about 0.05°C with a precision of 0.1°C (in the range of -65°C to 200°C), after the correction of the PT100 reference sensor by means of a 4th order polynomial equation.

The measurement results show good functionality from -55°C to 177°C . The most important measured characteristics are shown in Tableau 18 and Tableau 19.

Characteristics	Current mode	Voltage mode
Output swing	13 to 91 μA	0.4 to 3.5 volts
Sensitivity	0.333 - 0.335 $\mu\text{A}/^\circ\text{C}$	12.6 - 12.8 $\text{mV}/^\circ\text{C}$
Accuracy ¹	0.08 - 0.2 %/FS rms	0.57 - 1%/FS rms
Max. Temp. range	-55 to 177°C	
Current consumption	< 180 μA	
Power supply level	4 - 5.5 volts	
PSRR	> 40 dB	
Area	0.112 mm^2	

Tableau 18. Measured performances of the temperature sensor

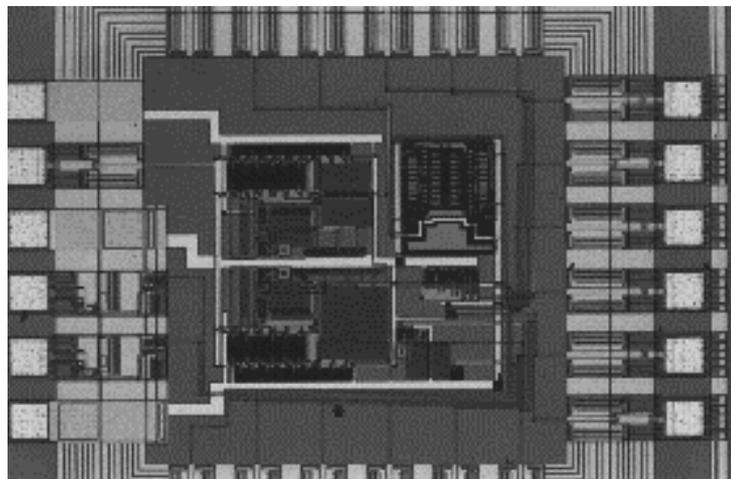


Figure 111. Microphotograph of the Temperature Sensor.

¹ Depends on the considered temperature range (see Tableau 19).

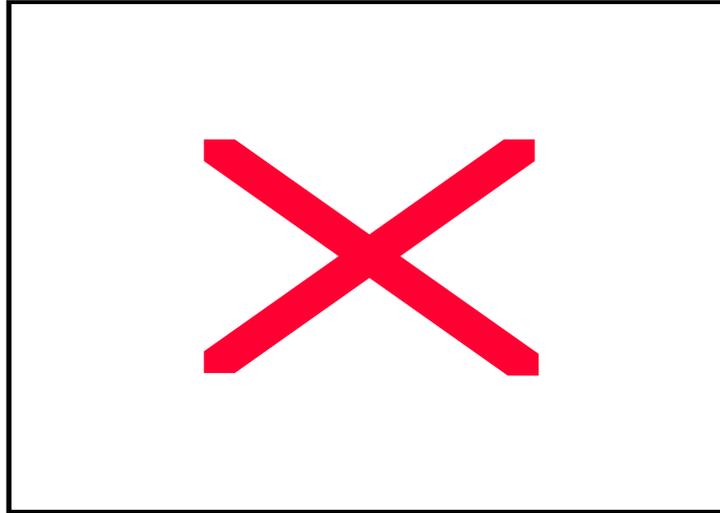


Figure 112. Temperature error characteristics in voltage and current modes

Figure 112 shows that a better performance was obtained in current mode than in voltage mode, according to the already described Poly-Si resistor temperature coefficient influence in the output current. A comparative evaluation between the obtained results and other similar results, corresponding to some integrated temperature sensors and transducers already published, is presented in Tableau 19.

Sensor	Range	Sensitivity	Temp. Error		Area	Power
	[°C]		[1/°C]	[°C]		
I mod	0 to 100	0.333 uA	0.08 rms	0.08 rms	0.116	180
I mod	-55 to 150	0.334 uA	0.34 rms	0.16 rms	0.116	180
I mod	-55 to 175	0.335 uA	0.44 rms	0.2 rms	0.116	180
V mod	0 to 100	12.6 mV	0.57 rms	0.57 rms	0.116	180
V mod	-55 to 150	12.7 mV	1.86 rms	0.9 rms	0.116	180
V mod	-55 to 175	12.8 mV	2.47 rms	1 rms	0.116	180
[45]	-40 to 120	-	± 1	0.63	1.5	28
[51]	180	-	0.5	0.3	4.5	50
[51]	60	-	0.1	0.16	4.5	50
[39]	-20 to 100	-	± 0.5	0.8	5.16	200
[37]	32 to 44	-	± 0.1	0.8	6	200
[48]	-20 to 100	4.6 mV	0.5	0.3	2.1	125
[36]	-50 to 125	170Hz	3	1.7	3	-
[49]	-50 to 125	1 uA	± 0.5	0.3	2.52	190
[46]	0 to 70	0.54 mV	-	-	0.018	0.15

Tableau 19. Performance of some integrated temperature sensors

In Tableau 20 other characteristics of integrated sensors are also considered, such as the type of output signal, the used IC technology and the necessity of calibration procedures to achieve the specified absolute accuracy.

Sensor	Output Signal		Technology	Calibration
	type	characteristics		
TIMA	Analog	Voltage and current	CMOS	yes
[45]	Digital	Σ - Δ converter	CMOS	yes
[51]	Digital	I-F conv. + dig. processing	CMOS	yes
[39]	Digital	Duty-cycle modulated	bipolar	yes
[37]	Digital	I-F converter	bipolar	yes
[48]	Analog	Voltage	bipolar	yes
[36]	Digital	Σ - Δ converter	bipolar	no
[49]	Analog	Current	bipolar	yes
[46]	Analog	Voltage	CMOS	yes

Tableau 20. Performance of some integrated temperature sensors

5.2.1.3. Sensor interfaces and digital interfaces

The chosen bipolar based sensing principle for temperature measurement has been studied and implemented with a standard CMOS technology, resulting in a very accurate, sensitive and low-cost, wide-range temperature sensor. Its characteristics make it especially suitable for low-cost and high-volume integrated microsystems over a wide range of fields, such as automotive, space, oil prospect, biomedical, etc. A digital interface, such as charge balancing AD converters, must be added in order to adapt the analog output of the sensor to Digital Signal Processing requirements, taking advantage of all the potential of the chosen CMOS technology.

5.2.1.3.a. Sensor interfaces

The main responsibility of the sensor interface is to preserve or improve the quality of the signals generated in the sensing element. Two kinds of errors can be encountered : systematic errors, caused by inaccuracy or by the variation of intermediate parameters; and random errors, caused by interference, noise, etc. Systematic errors can be eliminated by calibration or trimming procedures under similar conditions than the actual sensor operation conditions. But these methods are often unsuited solutions from an economical point of view.

A better way to deal with systematic errors is to eliminate the influence of the intermediate parameters, with the exception of those that are stable and accurate enough. Undesired parameters can be eliminated in many ways [38] :

1. Two-signal approach : Both, the reference signal S_{ref} and the measured signal S_m are obtained with the same measurement system. The final result is the ratio $R=S_m/S_{ref}$ or the difference $D=S_m-S_{ref}$ of both signals, eliminating the influence of the multiplicative or additive errors and intermediate parameters, respectively.

2. Three signals approach : Here two reference signals S_{ref1} and S_{ref2} and the unknown signal S_m are measured in an identical way. The final result is the ratio $R=(S_m-S_{ref1})/(S_{ref2}-S_{ref1})$ which is insensitive to both multiplicative and additive errors and parameters.

Smart sensor interfaces can also perform some linearity corrections as proposed in many publications, where the non-linear characteristic of some element [39][51] have been used for linearity compensation.

5.2.1.3.b. Digital interfaces

Smart temperature sensors need to be provided with some kind of output digital signal adapted to microprocessors and digital processing systems. This signal can be of time signal type, where the measurement is represented by the duty cycle or a frequency ratio, or a fully digital code that is sent to the processor in a serial way through a digital bus. Some important restraints, caused by the integration of sensing and digital processing functions on the same chip, are [36],

1. the limited chip area.
2. the tolerances of device parameters.
3. the digital interference.

Small digital interfaces and AD converters are preferable (1). Their most important properties must not rely on the absolute accuracy of devices (2), if trimming is not suited. A serious interference problem in the sensor and the analog circuitry can be expected due to charge injection through the bulk, from the clocked digital blocks. Noisy signals and unwanted synchronisation on switching moments can be produced by this kind of interference.

Since digital interference is highly periodic, it should not be treated as ordinary noise. Periodic sampling of the input signal is dangerous because the interference noise can be correlated with the sample moments. The use of time continuous filters and structures are recommended to avoid this problem.

The synchronisation problem particularly concerns the asynchronous or free-running relaxation oscillators that generates frequency or duty cycle signals. Digital interference can cause the free-running signals to be synchronised with the processor clock [36].

5.2.1.3.c. Suitable AD Converters

The first constraint (1) excludes large and complex converters such as the parallel multithreshold converters (flash ADCs, etc.). The second condition (2) excludes converters that use resistive or capacitive weighting networks, such as successive approximation, flash, charge-redistribution ADCs and any other ADC that uses weighted DACs. One decision level or comparator based AD converters seem to be the better adapted structures concerning the already mentioned constraints. The resolution is given by oversampling and feedback techniques. Charge-balancing converters with an integrator in the forward path for the input and the feedback signal, fit with the above proposed converters. The main design choices are between time-discrete or time-continuous converters, and they will be defined according to the already mentioned periodic noise (i.e. digital interference) constraint.

Time-discrete converters have a sampling character that makes them very sensible to periodic noise at the input. Time-continuous charge-balancing converters deal with this noise feeding it into a time-continuous integrator. Their high normal mode rejection increase their interference robustness. From the synchronisation sensitivity point of view we can identify the best time-continuous converters : free running converters such as frequency ratio and duty cycle, or clock synchronised such as the dual-slope and the sigma-delta converters.

In free-running converters case, important efforts have to be made to avoid the synchronisation of the time signals with the clock. On the other hand they are very interesting because they are small and they provide a single output line to the digital processor.

In clock synchronised converters case, the synchronising problem is completely solved, but the converters and the data communication will be more complex, enlarging the total chip surface. The sigma-delta converted is an exception to this general rule, because its digital output signal is very well adapted to single wire communication without any protocol [36].

5.2.1.4. Temperature sensor with a digital output signal

A CMOS process compatible wide range temperature sensor with a CMOS digital output signal, will be described. The digital interface has been implemented with a free-running converter, which generates a duty-cycle modulated digital output signal. Since no clocked digital blocks were integrated with the temperature sensor, the risks of an undesired clock synchronisation are negligible. Concerning accuracy, a temperature error smaller than 0.25°C rms (with two 2nd order temperature compensations, an on-chip one and an externally trimmable one) are the measured performance of this integrated sensor, over the -50°C to 150°C temperature range.

Other important characteristics are the low cost, the less than 0.55 mW of current consumption, and the digital output signal, which has been specially conditioned for on-chip digital signal processing or for a single-wire connection to an external processor.

The characteristics of this sensor make it especially suitable for low-cost high-volume integrated microsystems over a wide range of fields, such as automotive, aerospace, oil prospect, consumer, etc.

5.2.1.4.a. The sensing principle

To generate the output digital signal a relaxation oscillator [39] is used, where we integrate alternatively a signal proportional to temperature $It(T)$ and a reference Ir one, as shown in Figure 113 and Figure 114, to obtain a duty-cycle modulated digital signal where the influence of multiplicative intermediate parameters and multiplicative errors is eliminated (see section 5.2.1.3).

The duty-cycle and the frequency of the output Digital Signal DS (see Figure 114) are respectively given by,

$$(271) \quad \boxed{\text{X}}$$

$$(272) \quad \boxed{\text{X}}$$

Since DS_{freq} approaches to zero when $It(T) \Rightarrow 0$ or $It(T) \Rightarrow Ir$, the dynamic range of the duty-cycle DS_{d-c} must be limited as follows,

$$(273) \quad \boxed{\text{X}}$$

keeping the output frequency between the following values,

$$(274) \quad \boxed{\text{X}}$$

Where ΔVh and C are respectively the hysteresis level of the Schmitt trigger and the capacitance value, in Figure 114.

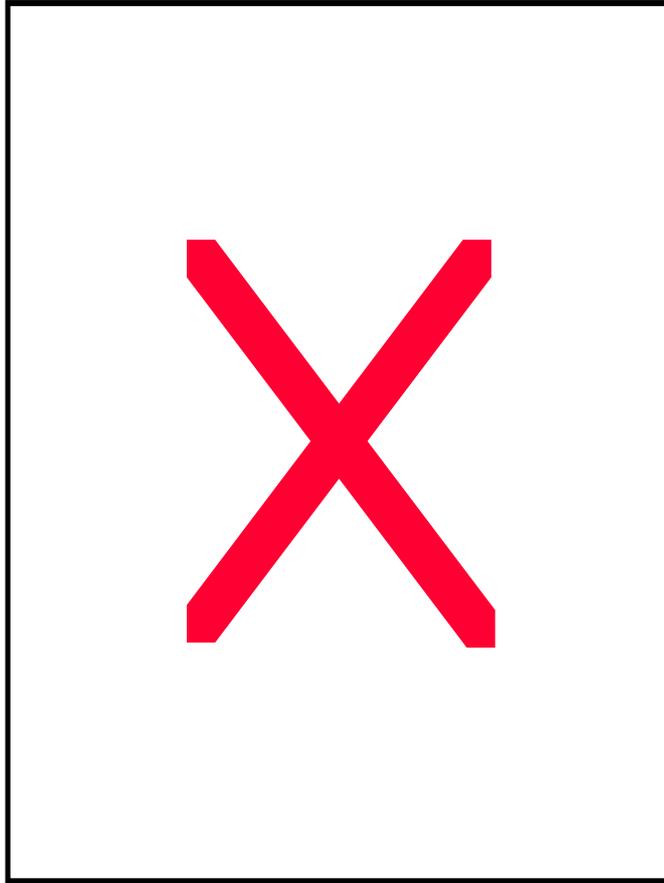


Figure 113. Sensor architecture: Current sources

5.2.1.4.b. The V_{be} signal

The base-emitter voltage equation of a CLBT biased with a collector current defined by $I_c = I_c(T_0) \cdot [T/T_0]^\alpha$ is expressed, as mentioned in section 4.1.3.1.b, by the following second order approximation,

(275)

The slope ST of $V_{be}(T)$ can be easily adjusted, by means of V_{be0} , modifying the transistor geometry (emitter area A_e , base width ωb , or the multiplier factors n and m) and the collector bias current $I_c(T_0)$, as follows,

(276)
$$ST = V_{g0} - V_{be0} + a \cdot T_0 \cdot \left[1 - \frac{b^2}{(T_0 + b)^2} \right] \cdot \sqrt{\gamma - \alpha} \cdot V_{T0} \cdot \frac{1}{T_0}$$

(277)

Replacing the last expression for V_{be0} in (276) and rearranging terms, the slope ST can be expressed by,

(278) 

Where $ST(m)$ represents the temperature slope of the V_{be} of a CLBT with a multiplier factor m constant and equal to the central value of the variable multiplier factor n . The second term shows how the slope can be controlled by means of n . The circuits described in section 4.1.3.1.b and in section 5.2.1.4.f.a, generate current mode versions of this signal.

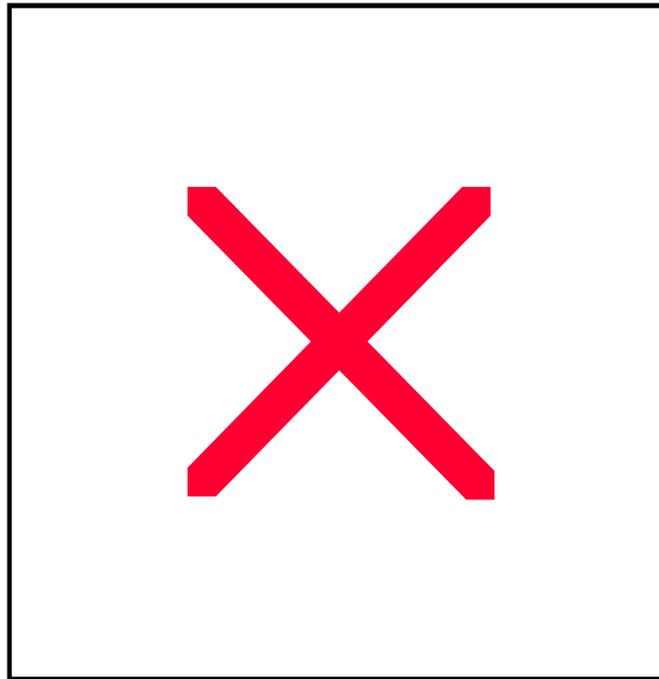
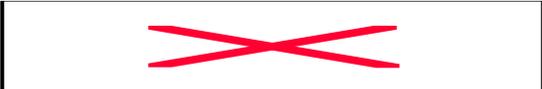


Figure 114. Sensor architecture: digital signal generator

5.2.1.4.c. Intrinsically referenced temperature signal $I_t(T)$

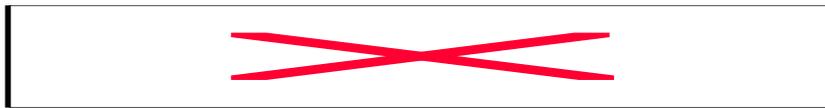
The current mode temperature signal $I_t(T)$, referenced to the origin of the targeted temperature range (-50 to 150°C), has been described in section 5.2.1.2.a.c, and approximated by the following expression,

(279) 

The measurement results of this $I_t(T)$ signal are showed in section 5.2.1.2.c, giving a small and negative $QTIt$ coefficient which is responsible for a small temperature error of about 0.34°C rms (0.16% FS rms) for the -55°C to 150°C temperature range. In the digital output signal temperature sensor, the $I_t(T)$ signal is generated by the circuit in Figure 117.

5.2.1.4.d. The reference signal $I_r(T)$

As shown in Figure 113 the reference signal is generated combining the already described $I_t(T)$ signal and a second V_{be} current mode signal (named I_{be2}) used for slope control, and given by,

(280) 

Thus, from Figure 113,

(281) 

Replacing the above mentioned linear approximations around $T0$ for $I_t(T)$ and $I_{be2}(T)$, in the last equation, we obtain the following expression for the temperature slope of $I_r(T)$,

(282) 

Then we define m in such a way that the slope α_R of $I_r(T)$ becomes zero when $n=m$. We compute consequently the required value for $R3$, according to the mentioned condition,

(283) 

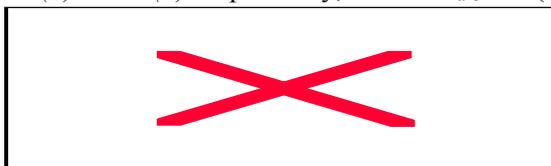
Finally the equation for the control of the slope of the reference current $I_r(T)$, by means of the multiplier factor n , is given by,

(284) 

The slope α_R of $I_r(T)$ becomes zero when $n=m$.

5.2.1.4.e. The output digital signal DS

The output digital signal DS is generated as already shown in *Figure 114*. Replacing the expressions (279) and (281), for $I_t(T)$ and $I_r(T)$ respectively, in the DS_{d-c} one (271), we obtain,

(285) 

Expanding this equation in series to the 2nd order around $T0$, the following duty-cycle temperature slope S_{DSd-c} and quadratic coefficient QT_{DSd-c} , are obtained,

(286) 

(287) 

The temperature slope S_{DSd-c} is slightly affected by the variations of α_R , but the quadratic coefficient QT_{DSd-c} is theoretically eliminated for $\alpha_R \approx -0.0289$, giving to $DS_{d-c}(T)$ a theoretically pure linear behavior. Thus, the final 2nd order compensation is performed by means of an external set-up of

switches which allows to control the multiplier factor n from 1 to 32 (see Figure 116), enabling α_R to be set between -0.071 and 0.023 during calibration procedures.

5.2.1.4.f. *The implemented circuit*

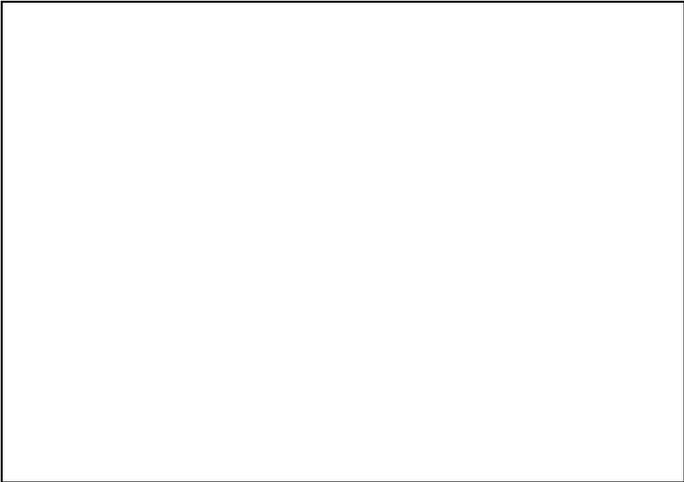


Figure 115. *Photomicrograph of the Temperature Sensor.*

The circuit implemented consists of three current source generators and a relaxation oscillator, combined as shown in Figure 113 and Figure 114. The Figure 115 shows the Photomicrograph of the integrated temperature sensor ($784 \times 776 \mu\text{m}^2$) that was fabricated in a standard $0.8 \mu\text{m}$ CMOS technology.

The $I_{be}(T)$ generators

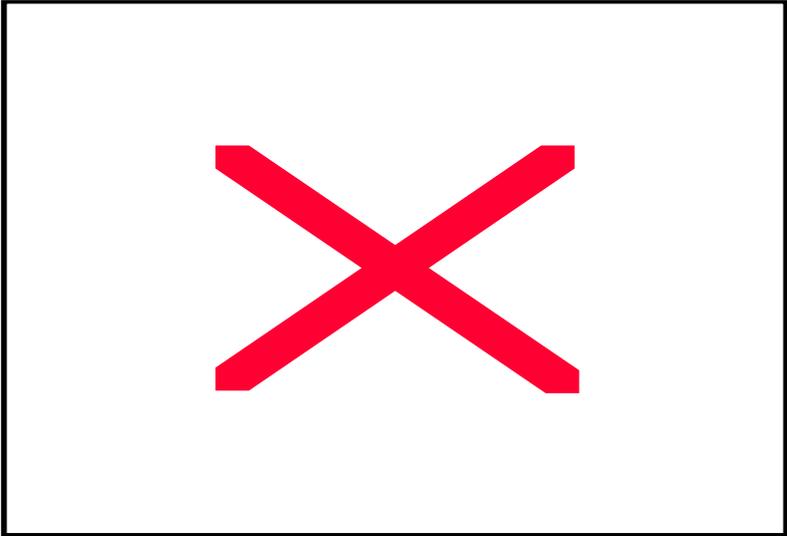


Figure 116. *$I_{be2}(T)$ current source circuit*

The $I_{be1}(T)$ and $I_{be2}(T)$ current sources were implemented as described in section 4.1.3.1.b and in Figure 116 respectively. The I_{be} currents are generated with the Poly-Si resistors $R2$ and $R3$, and the base-emitter voltage of $T1$ and $T2$, respectively. The $R2$ value was calculated for a Tz value slightly smaller than the lower temperature range limit (-50°C), significantly reducing the offset of the initial PTAT signal. The collector currents of $T1$ and $T2$ are PTAT, so they have a temperature coefficient " $\alpha=1$ ". The transistors $M4$ and $M14$ were added in order to increase the output impedance of both generators. The final 2nd order compensation is performed by means of the external set-up of switches shown in Figure 116, which allows to control the multiplier factor n from 1 to 32 during calibration procedures.

I

he $I_t(T)$ generator

The Figure 117 shows the circuit which combines the $I_{ptat}(T)$ and the $I_{be1}(T)$ current sources (from sections 4.1.3.1.a and 4.1.3.1.b respectively), to obtain a temperature signal that is referenced to the origin of the targeted temperature range.

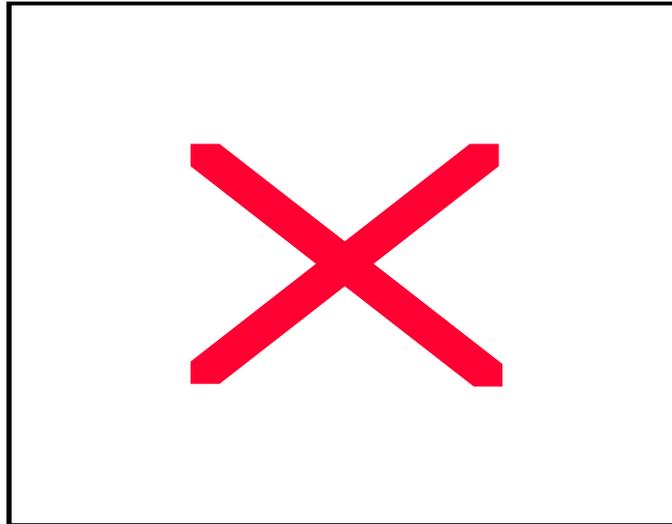


Figure 117. $I_t(T)$ current source circuit.

I

he relaxation oscillator

The free-running oscillator in Figure 114 was simply implemented with one NMOS and one PMOS transistors, acting as the " I_r-I_t " and " I_t " switches, a large PMOS capacitance (see section 3.2.2), and a CMOS comparator with ~ 1 volt of hysteresis (which is quite temperature dependent), acting as the Schmitt trigger.

The Nwell node of the PMOS capacitor was connected to ground (so that, the high temperature leakage currents in the substrate-Nwell junction do not affect the behavior of the circuit) and the Poly-Si plate was connected to one of the comparators inputs. The other input of the comparator was connected to $V_{dd}/2$, in order to bias the PMOS capacitance in the majority carriers accumulation condition. Thus, very large values of capacitance can be obtained occupying a relatively small area. On the other hand, the PMOS capacitor $C(V_c)$ will have some amount of non-linearity (see section 3.2.2.1) that must be taken into account. For any value of V_c , the PMOS capacitance is defined by,

$$(288) \quad \boxed{\text{X}}$$

Therefore, if we consider $I_c(t)$ constant (as described in Figure 114) $V_c(t)$ will not vary in a strictly linear way. We obtain for each value of the output signal DS ,

$$(289) \quad \boxed{\text{X}}$$

$$(290) \quad \boxed{\text{X}}$$

However the charge and discharge processes will be identical, for the same initial and final values of V_c . Rewriting (290), we obtain,

$$(291) \quad \boxed{\text{X}}$$

and,

$$(292) \quad \boxed{\text{X}}$$

as mentioned in section 5.2.1.4.a. The non-linearity of the PMOS capacitance, and the temperature variations of the hysteresis value ΔV_h , do not affect at all the described behavior of the relaxation oscillator.

5.2.1.4.g. Measurement results

Test procedures has been performed with a double chamber stable temperature generator using the CNES /TRS31 facilities. The resolution reached at the DUT level is about 0.05°C with a precision of 0.1°C (in the range of -65°C to 200°C), after the correction of the PT100 reference sensor by means of a 4th order polynomial equation. The measurement results, obtained from two tested devices, show good functionality from -60°C to 160°C . The most important measured characteristics are shown in Tableau 21 and in Tableau 22.

<i>Characteristics</i>	<i>Measurements</i>
Duty Cycle swing	10 to 80 %
Frequency swing	0.6 to 1.8 KHz
Sensitivity	~ 0.25 %/°C
Temperature error	0.24 - 0.37 °C rms 0.11 - 1.7 %/FS rms
Temperature range	-60 to 160 °C
Current consumption	< 110 μA
Area	0.61 mm ²

Tableau 21. Measured performance of the temperature sensor

The *device #2* showed similar characteristics, in terms of accuracy, than the analog version of the temperature sensor. But in the case of the *device #1*, accuracy has been increased as shown in Tableau

22. In Figure 119 and in Figure 120 we can observe how the quadratic error is reduced, due to the linearity compensation. In both cases, the remaining temperature error has a third order characteristic.

<i>n</i>	<i>sel1</i>	<i>sel2</i>	<i>sel4</i>	<i>sel8</i>	<i>sel16</i>	<i>Sel_number</i>	<i>Temperature Error [°C rms]</i>	
							<i>Device #1</i>	<i>Device #2</i>
1	open	open	open	open	open	0	0.48	0.37
2	Vdd	open	open	open	open	1	0.25	0.42
3	open	Vdd	open	open	open	2	0.24	0.54
4	open	open	Vdd	open	open	4	0.35	0.66
9	open	open	open	Vdd	open	8	0.45	0.85
17	open	open	open	open	Vdd	16	0.66	1.05
32	Vdd	Vdd	Vdd	Vdd	Vdd	31	-	1.28

Tableau 22. Accuracy over the -60 to 160°C temperature range.

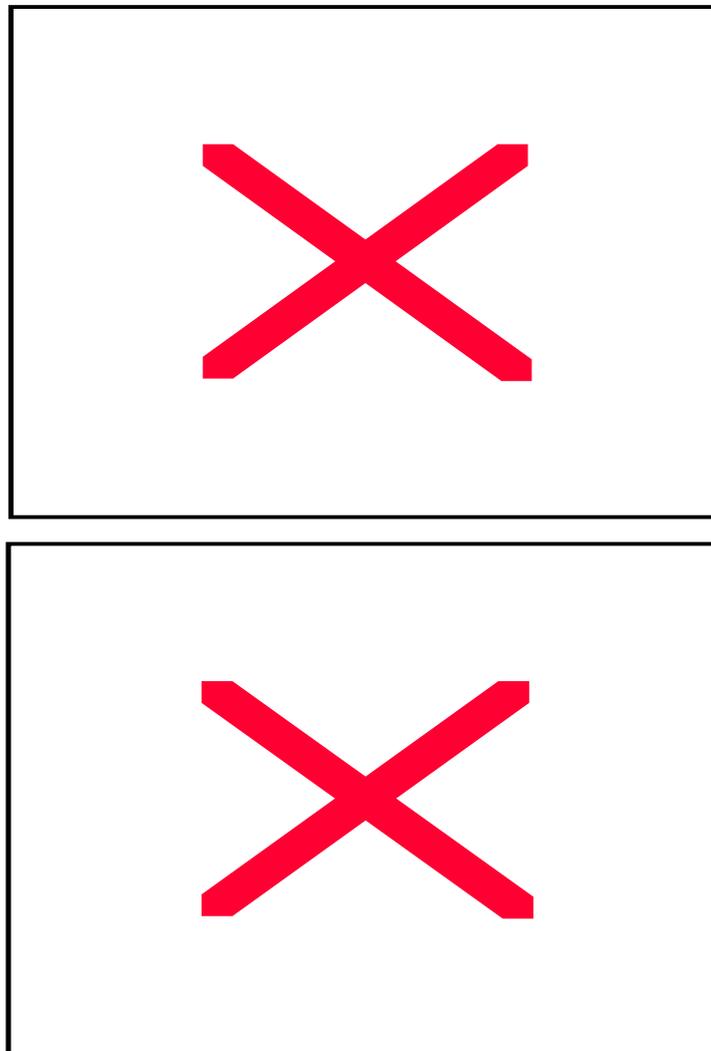


Figure 118. Measured temperature variations of the duty cycle $DSdc(T)$ of the output digital signal (up) and of the reference current signal $I_r(T)$ (bottom)

In Figure 118 we can see the linear response of the duty-cycle modulated digital output signal $DSdc(T)$ of the sensor and the typical slope variations of the reference current signal, measured on an amplified copy of $I_r(T)$ (The non-linearity, beyond 125°C, is only due to the protecting diodes in the output pad of the mentioned amplified copy of $I_r(T)$).

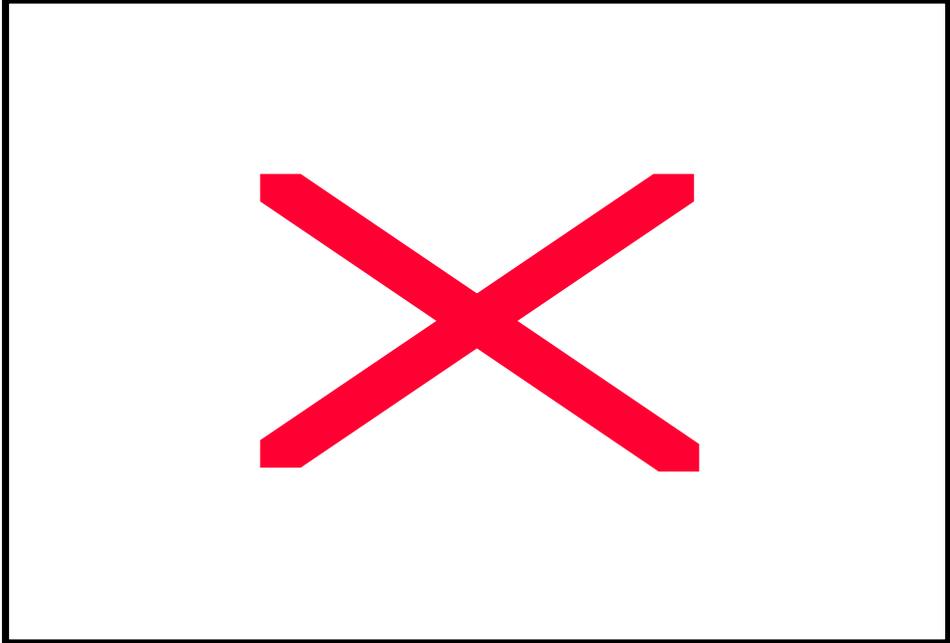


Figure 119. Temperature error vs. temperature for the device #1. The thickest line (sel2) corresponds to the highest obtained accuracy.

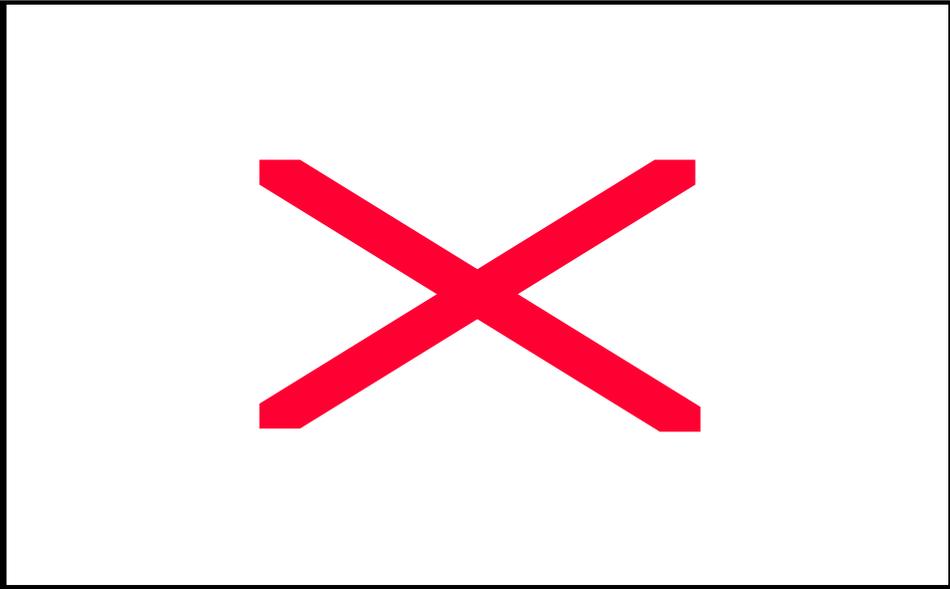


Figure 120. Temperature error vs. temperature for the device #2. The thickest line (sel0) corresponds to the highest obtained accuracy.

6. Conclusions

La problématique et le besoin d'une microélectronique à haute température sont clairement définis par les besoins actuels et futurs de ses différents domaines d'application, notamment l'industrie automobile, l'industrie pétrolière et l'industrie aérospatiale. La majorité des applications actuelles et des dix prochaines années (plus du 70% du total), correspondant à des températures d'opération inférieures à 200°C, ne justifient pas encore l'utilisation des nouvelles technologies de semiconducteurs émergentes, à large bande interdite ou du type SOI couche mince, du fait de la plus haute maturité et compétitivité des technologies standards CMOS et BiCMOS sur substrat de silicium, en termes de coût, de densité d'intégration, et de fiabilité.

En conséquence, le comportement en haute température de ces technologies a été étudié et des techniques de conception et d'implémentation des circuits analogiques (aussi applicables aux circuits mixtes), développées dans cette thèse, sont proposées pour la conception des circuits intégrés fonctionnels jusqu'à 250°C environ, sans introduire de modifications dans les procédés de fabrication. Les études et les tests ont été expérimentés sur une technologie CMOS et une technologie BiCMOS commerciales, les conclusions étant transférables à toute technologie semblable actuelle. De plus, les performances en température tendent à s'améliorer dans le futur, en terme général, car l'augmentation de la densité d'intégration entraîne une augmentation des concentrations des dopants et une réduction de la surface des jonctions d'isolation, autrement dit, une réduction des courants de fuite en haute température.

Deux applications industrielles, représentatives du marché potentiel des applications des circuits intégrés à haute température (d'une part, à cause de leur complexité et d'autre part, du fait qu'elles correspondent aux deux domaines d'applications actuellement les plus importants), ont permis de vérifier dans la pratique les résultats théoriques obtenus. De plus, la conception de ces circuits intégrés étant faite de façon modulaire, une bibliothèque des cellules analogiques paramétrables (i.e. oscillateurs à cristal, amplificateurs opérationnels, mélangeurs, comparateurs, comparateurs avec hysteresis, filtres actifs, buffers digitaux, cellules de polarisation en température) peut être définie à partir des circuits analogiques développés et testés pour les applications précédemment mentionnées.

Les résultats obtenus pendant cette thèse, ont également rendu possible la publication de plusieurs articles dans différents journaux, conférences, workshops et joint meetings internationaux.

7. Références

- [1] S. M. SZE, Physics of Semiconductor Devices, John Wiley & Sons, 1981.
- [2] H.Mathieu, Physique des semiconducteurs et des composants électroniques, Masson, Paris, 1996. ISBN: 2-225-85124-7.
- [3] M.Willander, H.L.Hartnagel, High Temperature Electronics, Chapman & Hall, London, 1997. ISBN: 0-412-62510-5.
- [4] Emiliano Pollino, Microelectronic Reliability, Vol II, Artech House, Norwood, 1989. ISBN: 0-89006-350-8.
- [5] F.P. McCluskey et al, High Temperature Electronics, CRC Press, New York, 1997. ISBN: 0-8493-9623-9.
- [6] J.Y.W. Seto, The Electrical Properties of Polycrystalline Silicon Films, J. Appl. Physics, Vol 46, 1975, pp. 5247-5254.
- [7] G. C. M. MEIJER, Thermal Sensors, Series Editor : B E Jones, 1994.
- [8] Wolfgang RASMUSSEN, Sensor Technology and Devices, Chapter 8, Artech House, Inc, 1994.
- [9] Ljubisa Ristic, Sensor Technology and Devices, Chapter 8, Artech House, Inc, 1994. ISBN: 0-89006-532-2
- [10] B. Parzen and A. Ballato, Design of Crystal and Other Harmonic Oscillators. New York: Wiley-Interscience, 1983. ISBN: 0-471-088196
- [11] R. J. Matthys, Crystal Oscillators Circuits. New York: Wiley-Interscience, 1983. ISBN: 0-471-87401-9
- [12] M. E. Frerking, Crystal Oscillator Design and Temperature Compensation. New York: Van Nostrand Reinhold, 1978. ISBN: 0-442-22459-1
- [13] R. R. Troutman, Latchup in CMOS Technology. Boston: Kluwer Academic Publishers, 1986. ISBN: 0-89838-215-7
- [14] Julian W. GARDNER, Microsensors principles and applications, John Wiley & Sons, 1994.
- [15] Xavier ARREGUIT, Compatible Lateral Bipolar Transistor in CMOS Technology, These N° 817, Ecole Polytechnique de Lausanne, 1989.
- [16] Ravender Goyal, High-Frequency Analog ICs Design, John Wiley & Sons, New York, 1995. ISBN: 0-471-53043-3
- [17] S. Wolf, R. N. Tauber, Silicon Processing for the VLSI Era, Lattice Press, Sunset Beach, 1986. ISBN: 0-961672-3-7
- [18] F. Eynde, Analog Interfaces for Digital Signal Processing Systems, Kluwer Academic Publishers, Boston 1993. ISBN: 0-7923-9348-1
- [19] Donald Pederson, Analog ICs for Communication, Kluwer Academic Publishers, Boston 1991. ISBN: 0-7923-9089-X
- [20] C. Toumazou, Analog IC Design: the current mode approach, Peter Peregrinus Ltd, London, 1990. ISBN: 0-86341-297-1
- [21] M. Sriram, S.M. Kang, Physical Design for Multichip Modules, Kluwer Academic Publishers, Boston 1994.
- [22] Keneth Kundert, The Designer's Guide to Spice and Spectre, Kluwer Academic Publishers, Boston 1995
- [23] S. M. Sze, Semiconductors Sensors, John Wiley & Sons, New York, 1994. ISBN: 0-471-54609-7
- [24] Paul R. Gray and R.G. Meyer, Microsensors principles and applications, John Wiley & Sons, New York, 1993. ISBN 0-471-81454-7
- [25] G. Massobrio, P. Antognetti, Semiconductor Device Modelling with SPICE, McGraw-Hill, New York, 1993.

- [26] Zhenhua Wang, Current-Mode Analog IC and Linearization Techniques in CMOS technologies, Hartung-Gore Verlag Konstanz.
- [27] Robert A. Pease, Troubleshooting Analog Circuits, Butterworth-Heinemann, Boston, 1991. ISBN: 0-7506-9184-0
- [28] N.C. Lu et al., Modeling and Optimization of Monolithic Polycrystalline Silicon Resistors, IEEE Transactions on Electron. Dev., Vol.ED-28, 1981, pp. 818-830.
- [29] E. Obermeir et al., Characteristics of Polysilicon Layers and their Application to Sensors, IEEE Solid-State Sensors Workshop, Hilton Head Is., SC, 1986.
- [30] E. Takeda, H.Kume, T. Toyabe and S. Asai, IEEE Transactions Electron Devices, vol. ED-29, 1982, p. 611.
- [31] F. S. Shoucair, "Scaling, Subthreshold, and Leakage Current Matching Characteristics in High Temperature (25°C-250°C) VLSI CMOS Devices", IEEE Transactions on Components, Hybrids, and Manufacturing Technology, Vol.12, No 2, Dec 1989.
- [32] P.D. Maycock, Thermal conductivity of silicon, germanium, III-V compounds and III-V alloys, Solid-State electronics, V.10, pp 161-168, 1967.
- [33] Properties of silicon, INSPEC, p. 35. The Institute Elec. Engrs., London, 1988.
- [34] F. S. Shoucair and J. M. Early, "High Temperature Diffusion leakage-current-dependent MOSFET small-signal conductance", IEEE Transactions on Electron Devices, Vol.ED-31, No 12, Dec 1984.
- [35] S.E. Nordquist, "High Temperature Leakage Current Suppression in CMOS Integrated Circuits", Electronic Letters, Vol. 25, No 17, 17th August 1989.
- [36] F.R. Riedijk and J. H. Huijsing, An integrated absolute temperature sensor with sigma-delta A/D conversion, Sensor and Actuators, A24 (1992), 249-256.
- [37] G. C. M. MEIJER, An Accurate Biomedical Temperature Transducer with On-Chip Microcomputer Interfacing, IEEE Journal of Solid State Circuits, Vol 23, No 6, December 1988.
- [38] Gerard C. M. MEIJER. Concepts and focus points for intelligent sensor systems. Sensor and Actuators, A41-42 (1994), 183-191.
- [39] G. C. M. MEIJER, A Three-Terminal Integrated Temperature Transducer with Microcomputer Interfacing, Sensors and Actuators, 18 (1989) 195-206.
- [40] G. C. M. MEIJER, Thermal Sensors based on Transistors, Sensors and Actuators, 10 (1986) 103-125.
- [41] J.M. Karam, M. Lubaszewski, B. Courtois, V. Szekely, M. Rencz. Thermal monitoring of self-checking systems, Journal of Electronics Testing, Special issue on On-Line Testing.
- [42] V. Szekely, Cs. Marta, Zs. Kohari, M. Rencz. New temperature sensors for DfTT applications, 2nd THERMINIC Workshop, Sept 1996, Budapest.
- [43] A. Kölling, F. Bak, P. Bergveld, E. Seevinck. Design of a CMOS Temperature Sensor with Current Output, Sensors and Actuators, A21-23 (1990), 645-649.
- [44] Gert van de Horn and Johan H. Huijsing, Smart Temperature Sensor with Integrated Bitstream Calibration, P 234-237, ESSCIRC'95, 1995.
- [45] A. Bakker and J. H. Huijsing, Micropower CMOS Smart Temperature Sensor, P 238-241, ESSCIRC'95, 1995.
- [46] E. Montané, S. A. Bota, J. Samitier, A Compact Temperature Sensor for a 1.0 μm CMOS technology Using Lateral PNP Transistors, 2nd THERMINIC Workshop, Sept 1996, Budapest.
- [47] M. Gunawan, G. C. M. Meijer, J. Fonderi, J. H. Huijsing, A Curvature Corrected Low Voltage Bandgap Reference, Journal of Solid State Circuits, VOL. 28, NO. 6 (1993), 667-670.
- [48] Gerard C. M. MEIJER. A Low Power Easy to Calibrate Temperature Transducer, IEEE Journal of Solid State Devices, VOL. SC-17, NO 3, June 1982.
- [49] Gerard C. M. MEIJER. An IC Temperature Transducer with Intrinsic Reference, IEEE Journal of Solid State Devices, VOL. SC-15, NO 3, June 1980.
- [50] R. A. Bianchi et al., CMOS Compatible Temperature Sensor Based on the Lateral Bipolar Transistor for Very Wide Temperature Range Applications, Sensors & Actuators: A. Physical, Vol. 71/1-2, pp. 3-9, Nov. 1998.

- [51] P. KRUMMENACHER, Smart Temperature Sensor in CMOS Technology, Sensors and Actuators, A21-A23 (1990) 636-638.
- [52] R. A. Bianchi et al., CMOS Compatible Smart Temperature Sensor, Microelectronics Journal 29, pp. 627-636, 1998.
- [53] E. A. Vittoz et al., "High-Performance Crystal Oscillator Circuits: Theory and Application", IEEE Journal of Solid-State Circuits, Vol 23, No 3, June 1988.
- [54] R. G. Meyer and D. Soo, "MOS Crystal Oscillator Design", IEEE Journal of Solid-State Circuits, Vol SC-15, No 2, pp 222-228, April 1980.
- [55] M. Soyuer, "Design Considerations for High-Frequency Crystal Oscillators", IEEE Journal of Solid-State Circuits, Vol 26, No 6, pp 889-893, June 1991.
- [56] M. A. Unkrich and R. G. Meyer, "Conditions for Start-Up Oscillations", IEEE Journal of Solid-State Circuits, Vol SC-17, No 1, pp 87-90, February 1982.
- [57] A. Rusznyak, "Start-Up Time of CMOS Oscillators", IEEE Transactions on Circuits and Systems, Vol CAS-34, No 3, pp 259-268, March 1987.
- [58] J. T. Santos and R. G. Meyer, "A One-Pin Crystal Oscillator for VLSI Circuits", IEEE Journal of Solid-State Circuits, Vol SC-19, No 2, pp 228-235, April 1984.
- [59] Y. Tsuzuki et al., "Fast Start-Up Crystal Oscillator Circuits", 1995 IEEE International Frequency Control Symposium, pp 565-568.
- [60] R. L. Filler and J. R. Vig, "Long Term Aging of Oscillators", 1992 IEEE International Frequency Control Symposium, pp 470-484.
- [61] F. S. Shoucair, "High Temperature Latchup Characteristics in VLSI CMOS Circuits", IEEE Transactions on Electron Devices, Vol 35, No 12, pp 2424-2426, December 1988.
- [62] F. S. Shoucair et al., "Electrical Characteristics of LSI Silicon MOSFET's at Very High Temperatures, Part III: Modeling and Circuit Behavior", IEEE Transactions on Components, Hybrids, and Manufacturing Technology, Vol CHMT-7, No 1, pp 146-153, March 1984.
- [63] F. S. Shoucair, "Design Considerations in High Temperature Analog CMOS Integrated Circuits", IEEE Transactions on Components, Hybrids, and Manufacturing Technology, Vol CHMT-9, No 3, pp 242-251, September 1986.
- [64] Eric A. VITTOZ, MOS transistor Operated in the Lateral Bipolar Mode and their application in CMOS Technology, IEEE Journal of Solid-State Circuits, 1983.
- [65] Marc G. R. DEGRAUWE, CMOS Voltage References Using Lateral Bipolar Transistors, IEEE Journal of Solid-State Circuits, 1985.
- [66] D. Foty, Aspects of MOS Modeling for Analog/Mixed-Signal Design, Analog & Mixed-Signal Applications Conference, July, 1997.
- [67] G. Shorthouse and S. Lande, World Markets for High Temperature Electronics, MST News, February, 1998.