UNIVERSITE MONTPELLIER II
SCIENCES ET TECHNIQUES DU LANGUEDOC

T H E S E

pour obtenir le grade de

DOCTEUR DE L' UNIVERSITE MONTPELLIER II

Discipline: « Électronique et Micro-optoélectronique »

Formation Doctorale: « Électronique: composants et systèmes »

préparée au sein du groupe de Microélectronique du
Laboratoire Européen pour la Recherche Nucléaire (CERN)
dans le cadre de l'École Doctorale:
«Information, Structures, Systèmes»

présentée et soutenue publiquement par

Roberto DINAPOLI

le 23 Mars 2004

Titre:

UN SYSTÈME DE DÉTECTION À PIXELS TOLÉRANT AUX RAYONNEMENTS POUR LES EXPÉRIENCES ALICE ET LHCB AU CERN

A RADIATION TOLERANT PIXEL DETECTOR SYSTEM FOR THE ALICE AND LHCB EXPERIMENTS AT CERN

Directeur de thèse:
M. Jean GASIOT

JURY

M. André TOUBOUL	Professeur, IXL	Rapporteur, Président
M. Raoul VELAZCO	Directeur de recherche, CNRS	Rapporteur
M. Jean GASIOT	Professeur, UMII	Directeur de thèse
M. Walter SNOEYS	Ingénieur de recherche, CERN	Resp. de thèse CERN

UNIVERSITE MONTPELLIER II
SCIENCES ET TECHNIQUES DU LANGUEDOC

N° attribué par la bibliothèque

T H E S E

pour obtenir le grade de

DOCTEUR DE L' UNIVERSITE MONTPELLIER II

Discipline: « Électronique et Micro-optoélectronique »

Formation Doctorale: « Électronique: composants et systèmes »

préparée au sein du groupe de Microélectronique du
Laboratoire Européen pour la Recherche Nucléaire (CERN)
dans le cadre de l'Ecole Doctorale:
«Information, Structures, Systèmes»

présentée et soutenue publiquement par

Roberto DINAPOLI

le 23 Mars 2004

Titre:

UN SYSTÈME DE DÉTECTION À PIXELS TOLÉRANT AUX RAYONNEMENTS POUR LES EXPÉRIENCES ALICE ET LHCB AU CERN

A RADIATION TOLERANT PIXEL DETECTOR SYSTEM FOR THE ALICE AND LHCB EXPERIMENTS AT CERN

Directeur de thèse:
M. Jean GASIOT

JURY

M. André TOUBOUL	Professeur, IXL	Rapporteur, Président
M. Raoul VELAZCO	Directeur de recherche, CNRS	Rapporteur
M. Jean GASIOT	Professeur, UMII	Directeur de thèse
M. Walter SNOEYS	Ingénieur de recherche, CERN	Resp. de thèse CERN

After the torchlight red on sweaty faces
After the frosty silence in the gardens
After the agony in stony places
The shouting and the crying
Prison and palace and reverberation
Of thunder of spring over distant mountains...

T.S. Eliot, *The Waste land*

Table of Contents

Résumé

Le travail présenté dans cette thèse a été effectué au sein du groupe Microélectronique du CERN, le laboratoire européen pour la physique des particules. Il s'agit d'un laboratoire situé près de Genève en Suisse, il a été créé dans les années 50 pour donner aux scientifiques européens les moyens d'étudier la physique des hautes énergies (HEP, High Energy Physics). Grâce aux accélérateurs de particules conçus et réalisés au CERN (en particulier le LEP, Large Electron Positron) il a été possible de développer le « Modèle Standard », une théorie qui essaye d'expliquer la matière en termes de forces et de particules. Ce modèle a été testé avec succès par les expériences de physique des particules, cependant il est incomplet, car il ne prend pas en compte la masse des particules fondamentales. L'idée la plus simple pour inclure cette dernière s'appelle le mécanisme de Higgs. Ce mécanisme implique l'existence de une particule additionnelle, appelée le boson de Higgs, et un type additionnel de force, se manifestant par des échanges de ce boson.

Pour évaluer cette hypothèse, ainsi que plusieurs autres phénomènes et théories, un nouvel accélérateur de particules est actuellement en construction au CERN, il s'agit du Large Hadron Collider (LHC). Le LHC sera l'accélérateur le plus puissant jamais construit.

Le détecteur à pixels décrit dans cette thèse a été conçu pour l'expérience de physique ALICE du futur collisionneur LHC. Les scientifiques pensent qu'il y a eu un «Big Bang» initial duquel tout l'Univers connu a émergé. Quinze milliards d'années après, l'Univers est si grand que la lumière prendrait des milliards d'années à le traverser. Pourtant, au début, tout était contenu dans un volume comparable à celui d'une mouche. Toutes les particules qui forment la matière que nous connaissons aujourd'hui se sont alors formées. Les quarks et les gluons, qui sont les constituants des protons et des neutrons dans notre Univers refroidi, étaient alors trop chauds pour s'associer. Cet état de la matière initiale s'appelle un plasma quark-gluon (QGP). Découvrir et analyser le QGP est l'objectif principal d'ALICE. En effet, ALICE est un détecteur de collisions d'ions lourds conçu pour étudier la physique de la matière en interaction forte et le plasma quark-gluon dans les collisions de noyaux produits par le LHC.

Dans les expériences autour du LHC (ALICE, LHCb, ATLAS, CMS, TOTEM) les particules seront accélérées pour atteindre des énergies de l'ordre du Tera Electron Volt (TeV) et des luminosités très élevées (10^{34} $cm^{-2}s^{-1}$ pour les protons et 1.95 10^{27} $cm^{-2}s^{-1}$ pour des ions de plomb). Cela implique des niveaux de rayonnement qui peuvent être très élevés,

particulièrement pour les détecteurs situés très près du point d'interaction. Pour l'expérience ALICE, en dix ans de fonctionnement du LHC, la dose ionisante totale peut atteindre 2.5 10^3 Gy et la fluence équivalente neutrons 1 MeV peut atteindre 2.95 10^{12} MeV n_{eq}/cm^2. Des niveaux de rayonnement beaucoup plus élevés peuvent être atteints dans les autres expériences.

Ceci pose un problème majeur pour la réalisation de l'électronique située près du point d'interaction, qui est habituellement celle des détecteurs de trajectoires. Cet environnement extrême, et la spécificité de l'électronique des détecteurs de trajectoires, font qu'aucun composant commercial n'est disponible. Le choix du développement d'ASICs dédiés s'est donc imposé. Une possibilité aurait consisté à résoudre le problème de la tolérance au rayonnement par durcissement du procédé technologique. En particulier quelques fondeurs spécialisés fournissent un procédé durci qualifié sous rayonnement. Ces technologies dont la pérennité ne peut être assurée sont toujours très coûteuses et souffrent de plus de divers handicaps liés à la difficulté de production des circuit : performances réduites, stabilité du procédé, l'obtention de rendements acceptables.

Le CERN a donc choisi de soutenir un projet de recherche (RD49) pour évaluer l'intérêt d'utiliser une technologie CMOS standard durcie aux effets des rayonnements par design (Hardening By Design, HBD). L'avantage de cette approche réside, en plus de son coût réduit, dans sa facilité d'adaptation aux nouvelles technologies submicroniques à venir. Le projet de construction du LHC s'échelonne sur une dizaine d'années, dans le même temps les technologies MOS évoluent très rapidement, les premiers essais de durcissement ont été effectués sur des technologies 0.5 μm. Celles-ci seront complètement obsolètes, tout comme les technologies durcies encore disponibles dans le commerce lorsque les approvisionnements de l'électronique pour les expériences du LHC seront réalisés. Jusqu'à ce jour, l'intégration qui accompagne l'évolution des composants s'accompagne d'une amélioration des caractéristiques des composants. D'ailleurs, plusieurs des circuits présentés dans cette thèse, ont été conçus en technologie CMOS standard 0.25 μm, durcis avec des techniques de HBD, ils répondent au cahier des charges pour l'électronique du détecteur à pixel en silicium de l'expérience ALICE (Silicon Pixel Detector, SPD), qui est le plus proche de l'aire de collision des particules.

En particulier, le circuit ALICE1LHCb (ou Circuit Pixel) contient une matrice de 32 par 256 cellules de lecture (pour un total de 13 millions de transistors), mesurant 13.5 par 15.8 mm^2. Des groupes de cinq circuits sont reliés électriquement par une technique de contact entre circuits sur des puces différentes, réalisés au moyen de rangées de billes métalliques microscopiques («bump-bonding»,). On obtient ainsi un grand senseur (160 colonnes par 256 lignes) qui forme le bloc de base qui constitue le SPD de l'expérience ALICE, dénommé «ladder». Un circuit est connecté à un senseur (de même dimensions, 32 colonnes par 256 lignes) pour former un «single», l'élément de base de détection pour le détecteur hybride à

photons (Hybrid Photon Detectror, HPD) de l'expérience LHCb. Le circuit est également employé pour le détecteur de trajectoires de l'expérience NA60.

Il utilise un schéma d'entrée différent des schèmes classiques, qui utilisent l'intégration de charges, la compensation pole-zéro et la mise en forme semi-gaussienne. Le circuit d'entrée réalise une configuration avec trois pôles (deux pôles complexes et un pôle réel, tous avec la même composante réelle), qui a été conçue pour supporter un fort taux d'occupation. Des tests minutieux de la puce, au laboratoire et sous irradiation dans un faisceau de particules, ont montré que ce circuit est entièrement fonctionnel et ce pour des doses allant jusqu'à 300 kGy.

Le **premier chapitre** présente au lecteur les objectifs du CERN, et en particulier l'expérience ALICE (les autres expériences sont décrites dans l'Annexe I). Nous proposons d'abord une courte description du Modèle Standard. Cette théorie inclut l'interaction forte associée à la charge de couleur des quarks et des gluons et à une théorie combinée de l'interaction faible et électromagnétique, plus connue sous le nom de théorie électrofaible, qui introduit les bosons W et Z comme particules porteuses de l'interaction faible, et les photons comme médiateurs des interactions électromagnétiques. Les valeurs très élevées, à des niveaux jamais atteints, de l'énergie et de la luminosité du faisceau du LHC permettent de mieux appréhender les défis technologiques principaux. Certains sont récapitulés dans ce chapitre.

Le **deuxième chapitre** commence par une vue d'ensemble des détecteurs de particules à semi-conducteur les plus importants utilisés pour la physique des hautes énergies avec une description de leurs propriétés, en particulier pour ce qui concerne les détecteurs hybrides à pixels. Ceci permet de mieux comprendre les multiples utilisations dans les expériences au CERN des divers types de détecteurs, et en particulier dans le détecteur de trajectoires ITS (Inner Tracking System) de l'expérience ALICE. Une section de ce chapitre est consacrée aux dommages induits par le rayonnement dans les détecteurs de particules à semi-conducteur. Les dommages les plus communs sont discutés, et les définitions des quantités physiques les plus importantes liées à ces phénomènes sont données (fluence de particules Φ, taux d'augmentation du courant de fuite α, facteur de durcissement K). Ces quantités seront employées pour la description des résultats de l'irradiation des détecteurs de particules d'ALICE.

Le **troisième chapitre** commence par une vue d'ensemble du détecteur de trajectoires ITS d'ALICE, et puis se concentre sur ses deux couches les plus internes, à proximité du faisceau, qui forment le détecteur SPD (Silicon Pixel Detector) d'ALICE. La recherche des trajectoires dans les collisions d'ions lourds au LHC constitue un grand défi, en raison de leur densité extrêmement élevée. L'ITS d'ALICE se compose de six couches de détecteurs cylindriques coaxiaux optimisés pour l'efficacité de détection de trajectoires et la

caractérisation des paramètres d'impact. Les deux couches externes sont équipées des détecteurs SDD (Silicon Drift Detectors), les deux couches intermédiaires avec des détecteurs SSD (Silicon Strip Detectors), et les deux couches les plus internes avec des détecteurs SPD. Ces SPD [ALI99] sont répartis sur deux couches cylindriques coaxiales, de rayons respectifs 3.9 et 7.6 centimètres. Le bloc fonctionnel de base du détecteur SPD d'ALICE est le ladder. Deux ladders son connectées ensemble, puis sont reliées à un Multi Chip Module (MCM) à l'aide d'un bus spécial Kapton-Aluminium pour former une demi-barrette («half stave»). Le bus qui relie le module MCM à l'électronique des capteurs se compose de lignes de données, de contrôle et de puissance.

Les 10 Circuits Pixel d'une demi-barrette sont commandés et lus par un MCM pilote (PILOT MCM). Le module MCM est composé de trois ASICs différents: le circuit APC (Analogue Pilot Chip) pour la polarisation du Circuit Pixel et pour le contrôle de la température et des signaux continus ou faiblement variables sur le MCM; le circuit DPC (Digital Pilot Chip) pour la lecture numérique, et le GOL (Gigabit Optical Serializer) pour la transmission optique des données. Tous les ASICs présents dans le module MCM sont décrits, tout comme la carte consacrée aux tests qui peut également émuler le comportement du logiciel et des composantes électroniques de l'ensemble du SPD.

Deux demi-barrettes forment une barrette, et six barrettes montées sur un support de fibre de carbone constituent un secteur. Le détecteur SPD est formé de dix secteurs assemblés. Les détecteurs et les circuits électroniques produisent une grande quantité de chaleur qui doit être évacuée tout en gardant une bonne stabilité en température. Un soin particulier a dû être apporté à la conception du système de refroidissement et de contrôle de température. La minimisation de la consommation d'énergie par l'électronique a été l'un des éléments les plus importants du cahier des charges. Un système de test spécifique (DAQ), très flexible a été développé, il a pu être employé dans un grand nombre de scénarios de test différents.

Le **quatrième chapitre** traite de la conception et des tests des détecteurs conçus pour les expériences ALICE et LHCb. Les détecteurs à pixels (senseurs) pour les expériences ALICE et LHCb sont constitués par une rangée de diodes p-in-n, réalisées par dopage au phosphore d'un substrat de type n. La valeur finale de l'épaisseur de plaquette sera de 300 μm pour l'expérience LHCb, «détecteurs épais», et de 200 μm pour ALICE, «détecteurs minces». Le détecteur est fabriqué par Canberra Électronique; une description des phases principales du procédé, correspondant à la conception des masques est présentée, ainsi que certaines caractéristiques particulières à la conception de détecteur (anneau de garde, ligne de découpe, structures de test «accordéon», dites aussi « snake test strutures »). La deuxième partie du chapitre est consacrée aux résultats des tests du détecteur. Plusieurs tests ont été réalisés (en plus de ceux effectués par le fabricant) sur un ensemble de plaquettes de pré-série, épaisses et minces, fournies par Canberra. Des tests électriques ont été réalisés sur les diodes de 68 mm^2 présentes sur les plaquettes; ils portent sur la mesure de la tension correspondant à la désertion

maximum (V_{fd}) et sur la mesure du courant de fuite total. A part quelques problèmes sur les premiers lots, qui ont été résolus, toutes les caractéristiques électriques des senseurs répondent à celles données par le fabricant.

Le courant de fuite, mesuré sur une diode de test, est tout à fait uniforme et son maximum est la moitié de la valeur maximale admise. En outre la tension de désertion maximum (8 V), mesurée sur la même diode de test s'est avérée très uniforme et bien au-dessous de la valeur nominale maximum qui est de 20-30 V. Des tests ont été effectués sur plusieurs assemblages pour s'assurer de la qualité du bump-bonding. Sur une partie des derniers véhicules de test examinés avec une source radioactive de strontium 90, la quantité maximum de pixel manquants sur un circuit est inférieur à 0.3% en dessous de la valeur de 1% visé par l'expérience ALICE. Des essais d'irradiation des diodes de 68 mm^2 avec un faisceau de protons de 27 MeV ont été réalisés à Legnaro, Italie, avec des fluences de 0.2 à 25 10^{12} protons/cm^2. La constante α de dégradation du courant de fuite (pour une fluence équivalente neutrons 1 MeV) a été mesurée, $\alpha = 0.62\ 10^{-16}$ A/cm pour les détecteurs minces et $\alpha = 0.83\ 10^{-16}$ A/cm pour les détecteurs épais. Les mesures de recuit montrent que le courant de fuite diminue de 21% après 3 jours à la température ambiante, puis reste constant.

La tension correspondant à la désertion maximum (V_{fd}) a été mesurée après irradiation (à deux fréquences de test différentes, à 1 kHz et à 10 kHertz) pour les senseurs minces et épais. V_{fd} reste au dessous de 100 V pour une fluence équivalente neutrons 1 MeV qui est six fois la fluence attendue pour la couche la plus exposée d'ALICE en 10 ans d'opération. Pour étayer ces mesures, qui n'étaient pas faites dans des conditions de test standard, d'autres irradiations de capteurs ont été faites avec des protons de 24 GeV, et des fluences de 3.36 à 16 10^{12} protons/cm^2. Dans ces conditions, nous avons mesuré la constante $\alpha = 4.52\ 10^{-17}$ A/cm (pour une fluence équivalente neutrons 1 MeV), ce résultat est en bon accord avec ceux présentés dans la littérature. Pour vérifier que l'épaisseur et la courbure de la plaquette sont selon les spécifications pour ALICE, quelques mesures ont été faites dans un laboratoire spécialisé du CERN. Les résultats de ces mesures sont entièrement conformes aux spécifications. En conclusion, nous pouvons assurer que les détecteurs fournis par Canberra sont conformes au cahier des charges, ils font même souvent mieux que ce qui est attendu. Ces détecteurs peuvent être employés sans risque avec les niveaux de rayonnement rencontrés dans l'expérience ALICE.

L'idée centrale qui a motivé ce travail est la possibilité de réaliser un détecteur hybride pour la physique des hautes énergies (HEP) avec une technologie MOS standard en réalisant un durcissement par design (HBD). Dans cette optique, nous exposons au **chapitre 5** les problèmes posés par les environnements fortement radiatifs sur des dispositifs électroniques. Nous décrivons les diverses solutions de durcissement, dont certaines ont été particulièrement étudiées au CERN. Pour ce qui est des effets cumulatifs induits par le rayonnement sur des dispositifs MOS, ils sont regroupés en deux classes : les effets d'ionisation et le déplacement

nucléaire (les composants MOS sont pratiquement insensibles aux effets de déplacement, mais à l'inverse, ils sont vulnérables à l'ionisation). L'effet d'ionisation ou de dose (ou encore TID, Total Ionizing Dose) est un effet cumulatif résultant de l'accumulation de charges induites par le rayonnement qui se piègent dans les oxydes, et se traduit par l'apparition de tensions parasites. Sur les MOS peuvent apparaître, en plus des variations dans la tension de seuil, des modifications dans la pente en inversion faible, de la transconductance, du comportement en bruit, entre autres caractéristiques. L'irradiation ionisante induit aussi différents types de courants parasites, résultant du piégeage de charge dans l'oxyde épais qui sépare des transistors (oxyde de champ). Une dose relativement faible dans un oxyde de champ peut induire suffisamment de piégeage de charges pour causer la défaillance d'un circuit intégré, due à une apparition des courants parasites drain-source et inter-transistor.

Une autre catégorie d'effets indésirables qui peuvent être induits par des rayonnements, est celle des phénomènes non récurrents (Single Event Effects, SEE). Ils sont produits par des particules fortement énergiques, lors de leur passage dans une partie sensible d'un transistor ou d'un circuit intégré. L'erreur la moins pénalisante et la plus courante est l'alea logique (Single Event Upset, SEU). Un SEU est un changement instantané et réversible de l'état logique d'une cellule. Il est produit par une particule ionisante lors de la traversée d'un point mémoire, qui induit (directement ou par des interactions secondaires) une charge suffisante pour modifier l'état logique de la cellule. Un autre effet non récurrent est le verrouillage maintenu (Single Event Latchup, SEL) qui peut (comme des autres phénomènes non récurrents) être à l'origine de la destruction des composants. Le SEL résulte du verrouillage d'un thyristor parasite qui court-circuite l'alimentation d'énergie, ce qui peut détruire le dispositif. Ce phénomène se rencontre dans quelques technologies CMOS; il peut aussi être provoqué lors de mise sous tension, faire suite à une élévation de température trop importante, mais également apparaître lorsqu'une particule ionisante traverse le composant.

Les trois options permettant de disposer de circuits tolérants au rayonnement sont discutées plus en détail: nous abordons le durcissement par le procédé, l'utilisation de composants qui existent dans le commerce, et l'adaptation de technologies standard durcies par HBD. Le principe de base du durcissement HBD consiste, à partir d'une technologie commerciale CMOS disponible et de coût réduit, à appliquer des techniques de conception de circuits propres à améliorer la tolérance au rayonnement. L'avantage principal de cette approche réside dans son prix réduit, et dans la possibilité de l'adapter simplement en fonction de l'évolution des technologies commerciales, dont l'amélioration des performances conduit à leur renouvellement très rapide.

Comme nous l'avons précisé, l'approche de durcissement HBD a été employée pour la conception du circuit ALICE1LHCb. Pour cette raison elle est présentée en détail, et nous évaluons l'impact de la réduction de la taille des composants sur l'effet du rayonnement et sur les performances du circuit.

Lors du passage d'une génération de composants MOS à la suivante, l'épaisseur d'oxyde de grille, t_{ox}, est réduit. Comme l'a montré Saks, le décalage de tension de seuil induit par le rayonnement ionisante diminue avec $(1/t_{ox})^2$ pour des oxydes de grille s'amincissant jusqu'à environ 20 nm puis la dépendance devient beaucoup plus rapide. La technologie CMOS 0.25 μm retenue pour la conception du circuit ALICE1LHCb a une épaisseur d'oxyde environ de 5.5 nm, et le décalage de la tension de seuil induit par le rayonnement est négligeable jusqu'à des doses cumulées de plusieurs dizaines de kGy. Les courants parasites sont diminués lors de la réduction des motifs de la technologie, mais pas éliminés, alors qu'aucun changement notable de la pente en inversion faible n'est induit par rayonnement dans les technologies submicroniques. Tous les phénomènes non récurrents sont atténués (ou éliminés) en technologie submicronique, excepté les SEU, parce que la sensibilité aux SEU est fortement dépendante de la technologie et de la conception.

Le problème principal qui doit être résolu par les techniques de HBD est l'augmentation des courants parasites dans l'état bloqué. La solution au problème des fuites inter-transistor consiste à ajouter un anneau de garde de type p^+ (également appelé « channel stop ») autour de chaque région de type n portée à un potentiel différent de l'alimentation plus baisse, et autour de tous les caissons-n portés à un potentiel différent de l'alimentation plus élevé (ceci se produit par exemple lorsque la source des transistors canal-p est reliée au caisson).

Quelques topologies de transistor qui réduisent ou éliminent les fuites drain-source sont présentées; parmi elles il y a le transistor ELT, «Enclosed Layout Transistor», qui utilise une grille annulaire dans la région active de sorte que le drain (ou la source) est a l'intérieur et la source (ou le drain) est à l'extérieur. Ces transistors occupent une surface supérieure à des transistors normaux, ils sont de plus asymétriques, leur capacité électrique est plus importante, mais ils constituent une solution très efficace pour le durcissement à la dose. D'autres techniques sont décrites pour augmenter la tolérance au rayonnement en agissant par exemple au niveau du système, et en particulier pour améliorer l'immunité aux phénomènes non récurrents.

Les transistors ELT entrent dans la conception du circuit ALICE1LHCb. De ce fait leurs caractéristiques, qui ont été étudiées au CERN dans le cadre du projet RD49, sont présentées en détail. Un modèle précis de l'une des formes possibles des transistors ELT est présenté, qui correspond bien aux résultats des mesures.

L'expression de l'appariement de transistors ELT identiques a une équation légèrement différente de celle des transistors standard (Standard Layout Transistors, SLT). Pour les transistors SLT, l'appariement s'améliore avec la taille du transistor, alors que pour les ELT il sature à une certaine valeur. A part cela, les paramètres sont semblables à ceux des SLT, comme montré dans le tableau 1.

	n SLT	n ELT (drain inside)	n ELT (drain outside)	p SLT
V_{th} Matching equation	$\sigma^2_{\Delta V_{th}} = \frac{A^2_{V_{th}}}{W\,L}$	$\sigma_{\Delta V_{th}} = \sqrt{\frac{A^2_{V_{th}}}{W\,L} + \sigma^2_0}$	$\sigma_{\Delta V_{th}} = \sqrt{\frac{A^2_{V_{th}}}{W\,L} + \sigma^2_0}$	$\sigma^2_{\Delta V_{th}} = \frac{A^2_{V_{th}}}{W\,L}$
$\Delta\beta/\beta$ Matching equation	$\sigma^2_{\Delta\beta/\beta} = \frac{A^2_{\beta}}{W\,L}$	$\sigma_{\Delta V_{th}} = \sqrt{\frac{A^2_{\beta}}{W\,L} + \sigma^2_{0\beta}}$	$\sigma_{\Delta V_{th}} = \sqrt{\frac{A^2_{\beta}}{W\,L} + \sigma^2_{0\beta}}$	$\sigma^2_{\Delta\beta/\beta} = \frac{A^2_{\beta}}{W\,L}$
A_{Vth}	4 mV·µm	4.3 mV	4 mV·µm	3.7 mV·µm
σ_0	-	1 mV	0.5 mV	-
A_β	1 %·µm	1.3 %·µm	1.3 %	1 %·µm
$\sigma_{0\beta}$	-	0.3 %	0.3 %	-

Tableau 1: Performances d'appariement des transistors canal-n et canal-p pour les «Standard Layout Transistors» (SLT) et les «Enclosed Layout Transistors» (ELT), avec le drain à l'intérieur ou a l'extérieur de la grille annulaire.

Une autre caractéristique très importante pour la conception de circuits intégrés est la performance en bruit de la technologie choisie. Des mesures ont été effectuées sur différents types de transistors en technologie 0.25 µm CMOS pour extraire les paramètres de bruit; les résultats sont récapitulés dans le tableau 2. Les valeurs des paramètres sont une fonction de la longueur L des transistors et sont plus élevées lorsque L diminue.

	n ELT s.i.	n ELT m.i.	n ELT w.i.	p SLT s.i.	p SLT m.i.	p SLT w.i.
α	.9 to .98	.9 to .98	.9 to .98	.8 to .9	.8 to .9	.8 to .9
K_a $[10^{-27}C^2/m^2]$	3.8 to 2	3.4 to 1	3.4 to 1	0.8 to 0.6	0.6 to 0.4	0.6 to 0.4
Γ	3.7 to 2.2	1.3 to 1.2	1.2 to 1.1	2 to 1.6	1.2 to 1.1	1

Tableau 2: Performances en bruit de transistors canal-n et canal-p pour les «Standard Layout Transistors» (SLTs) et les «Enclosed Layout Transistors» (ELTs), pour une inversion forte (s.i.), faible (w.i.) ou modérée (m.i.).

Les transistors (ELT et transistors standard) ont été irradiés en utilisant des rayons X de 10 keV, et une source de rayonnement γ, ^{60}Co. Toutes les irradiations ont été exécutées avec une polarisation «pire cas», et les mesures ont été effectuées après recuit. Les ELTs ne montrent aucun courant de fuite même après une dose de 10 kGy (SiO_2). Sur tous les échantillons mesurés la diminution de la transconductance et de la mobilité est inférieure à 6%, et les variations de la pente en faible inversion et de la conductance de sortie sont de l'ordre de quelques pour cent. Le coefficient A_β d'appariement du gain de courant («current gain matching») ne se dégrade pas après une dose de 15 kGy (SiO_2). Le coefficient

d'appariement de tension de seuil A_{vth} pour des transistors canal-p change de quelque pour cents, alors que pour des transistors canal-n A_{vth}, il augmente de 45% (mais la précision de la mesure est très pauvre). L'augmentation du bruit blanc est limitée dans toutes les régions d'inversion : maximum 15% pour les canal-n et 7% pour des dispositifs canal-p. La constante de bruit 1/f (K_a) augmente avec l'irradiation, d'un facteur 2 pour les transistors canal-n et d'un facteur 8 pour les transistors canal-p, pour une dose de 1 MGy (SiO_2).

L'utilisation des techniques de durcissement HBD réduit la densité de composants que l'on peut intégrer, cela se traduit par une augmentation de la consommation. Cette surcharge a été évaluée pour des cellules et pour des systèmes complets. Bien que les résultats soient dépendants de la conception et de la technologie choisie, on peut estimer une pénalité pour la puissance d'environ un facteur 2 et pour la surface d'un facteur 2.5 à 3. Il est important de préciser que dans tous les cas étudiés de durcissement à la conception par HBD, les performances sont toujours supérieures au produit correspondant en technologies durcies disponibles dans le commerce.

L'efficacité de l'approche durcissement HBD par rapport à la tolérance aux SEEs a été examinée au CERN en irradiant plusieurs types de registres à décalage avec des particules de haute énergie. La sensibilité des registres a été mesurée en fonction du transfert linéaire d'énergie (Linear Energy Transfer, LET) jusqu'à 89 $MeVcm^2mg^{-1}$: plus le LET est élevé, plus on a de risque de déclencher des phenomenes non récurrents. On n'a observé aucun effet SEL pendant toute la campagne d'irradiation. Les résultats d'irradiation montrent une meilleure tolérance aux SEU des transistors durcis par HBD que ceux de conception standard. Mais aussi une meilleure tolérance des structures statiques par rapport aux dynamiques. Pour se prémunir des SEU, le fait de charger une cellule avec une capacité de sortie plus élevée donne de meilleurs résultats que l'augmentation de la taille des transistors.

Le **chapitre 6** présente les circuits prototypes dont les développements mènent au circuit ALICE1LHCb. Ce chapitre donne une brève vue d'ensemble de deux des premiers circuits de lecture à pixels destinés à la physique des hautes énergies conçus au CERN. Il s'agit des circuits Omega2 et Omega3, réalisés à partir de technologies commerciales CMOS, respectivement de 3 μm et de 1 μm. Ils ont prouvé la possibilité d'employer des détecteurs à pixel pour la physique des hautes énergies, même si leur tolérance au rayonnement était insuffisante et nécessitait une amélioration sensible.

Le circuit de test ALICE1 était un premier prototype de circuit de lecture à pixels, conçu en technologie commerciale CMOS de 0.5 μm. Il utilisait des techniques de dessin pour améliorer sa tolérance au rayonnement. Il a prouvé le bien fondé de l'approche de durcissement par HBD appliqué a un circuit fonctionnant en mode-mixte (fonctions analogiques et digitales sur le même silicium). Parmi les résultats de test présentés, nous pouvons mentionner le fait qu'il a résisté (selon le type de rayonnement) à une dose totale comprise entre 6 kGy et 17 kGy. Aucun courant de fuite induit par le rayonnement n'a été

observé, et le circuit a finalement été défaillant en raison des effets cumulatifs des décalages de seuil induits par le rayonnement.

La disponibilité pour le CERN d'une technologie plus avancée (avec une lithographie de 0.25 μm) a conduit à la conception du circuit de test ALICE2, basé sur le prototype précédent mais avec quelques changements. Ce circuit, qui accepte à l'entrée des signaux positifs ou négatifs, comporte un préamplificateur, un filtre de mise en forme, un comparateur, une ligne à retard et la logique de lecture. Comme ce prototype n'a pas été prévu pour être connecté à un détecteur, une structure d'entrée a été ajoutée à chaque cellule pour simuler la capacité du détecteur, le couplage entre pixel, et le courant de fuite du détecteur. Quelques détails sur la structure du circuit sont présentés, ainsi que les résultats expérimentaux. Le circuit a été testé avant et après irradiation. Avant irradiation il est entièrement fonctionnel, et présente une tension minimum de seuil de détection du signal d'environ 1500 e^-, avec une dispersion de seuil d'environ 160 e^- rms (sans ajustement de seuil) et 25 e^- rms(après ajustement). Le bruit de pixel est environ 220 e^- rms.

Les résultats les plus importants qui ont été dérivés des tests de ce prototype ont trait à sa tolérance aux rayonnements. En fait ce circuit a été irradié aux rayons X, avec des protons d'énergie élevée et avec un rayonnement γ. L'absence d'une quelconque augmentation de la puissance absorbée avec la dose totale confirmée sur un circuit complet, montre que l'utilisation des transistors ELT canal-n avec anneaux de garde évite les fuites induites par rayonnement. D'ailleurs, le circuit reste entièrement fonctionnel jusqu'à une dose ionisante totale de 300 kGy (SiO_2) et une fluence de 9 10^{14} protons cm^{-2}. L'irradiation a eu lieu au SPS du CERN, avec des protons de 450 GeV/c. Le circuit a été également irradié avec un accélérateur Van de Graaf délivrant des protons de 6.5 MeV/c. Un circuit a été irradié par paliers aux doses de 90, 190 et 480 kGy. Il a cessé de fonctionner à 480 kGy. Les sorties analogiques et la consommation de puissance ont montré un comportement semblable à celui observé dans les autres mesures. Un circuit a été également irradié avec des rayons gamma par palier aux doses de 30, 190, 230 et 260 kGy. Dans ce cas, une légère augmentation de consommation d'énergie dans la partie analogique a été enregistrée. Pour les autres paramètres, les résultats de cette irradiation ont étroitement reflété ceux de l'irradiation aux rayons X discutée ci-dessus.

Ce prototype a prouvé définitivement qu'il était possible de concevoir (en technologie commerciale) un circuit mixte entièrement fonctionnel pour les détecteurs à pixels hybrides. L'utilisation des techniques de durcissement par HBD, permet à ces circuits de résister à des doses de rayonnement et à fluences de particules beaucoup plus élevées que celles prévues pour le LHC.

Le **chapitre 7** présente la conception et les simulations de la puce ALICE1LHCb, et en particulier de son circuit d'entrée. Il utilise un schéma d'entrée différent des classiques, qui utilisent l'intégration de charges, la compensation pole-zéro et la mise en forme semi-

gaussienne. Ce circuit a été conçu pour supporter un fort taux d'occupation du canal, en permettent un retour à l'origine rapide de tous les blocs qui composent la chaîne de lecture, tout en gardant un bruit bas (< 200 e^-) et une faible consommation (de l'ordre de 50-100 μW/canal).

Dans la première section nous récapitulons toutes les exigences de ALICE et LHCb, pour motiver nos choix de conception. La deuxième section est consacrée à une analyse détaillée du choix du design de l'étage d'entrée. L'intégration de charge, plus la compensation pole-zéro et la mise en forme semi-gaussienne sont analysés, et s'avèrent extrêmement difficile à utiliser dans notre cas. En effet, le fort taux d'occupation du canal peut amener à la saturation des étages amplificateurs (pile-up); il serait également nécessaire de concevoir un arrangement différent de l'étage de contre-réaction, pour améliorer la précision de la compensation pole-zéro, et la diminution du temps de retour à l'origine (qui peut être amélioré seulement en augmentant l'ordre du filtre de mise en forme semi-gaussienne). En plus, l'approche analytique proposée par Chang et Sansen [Cha91] est seulement approximative pour un arrangement avec un amplificateur de charge (CSA, Charge Sensitive Amplifier) de basse puissance et un temps de mise en forme rapide.

Ceci nous a conduit à développer un schéma d'entrée différent, ayant des performances en bruit identiques (ou meilleures) à celles obtenues avec un schéma standard, mais avec des temps de retour à l'origine améliorés à la fois pour les sorties du circuit de mise en forme et pour le préamplificateur, ce qui permet d'utiliser ce circuit même dans le cas d'un fort taux d'occupation [Din03, Din01, Din00, Sno01, Sno01a, Wyl99].

L'idée de base est qu'un système à pôles complexes peut avoir un retour à l'origine plus rapide qu'un système ayant des pôles réels. La proposition est que l'amplificateur de charge CSA soit suivi d'un étage simple (avec un gain en continu, A) ayant un pôle avec une constante de temps τ_{p2}, et qu'un courant de contreréaction soit appliqué à l'amplificateur par un étage de contreréaction de transconductance g_{mf}. En choisissant correctement τ_{p2}, A et g_{mf} il est possible de faire en sorte que les pôles soient complexes conjugués et de choisir la valeur de leurs composantes réelles et imaginaires. Des calculs doivent être effectués pour étudier la possibilité de réaliser les performances du système en termes de bruit, de retour à l'origine, de consommation et de surface occupée pour réaliser le circuit. Comme dans notre architecture aucune résistance n'est présente en parallèle avec la capacité de contreréaction, un élément de circuit a été conçu pour absorber le courant de fuite du détecteur, sans interférer avec le comportement à haute fréquence de l'étage d'entrée. Le circuit compense également le décalage (offset) présent à la sortie de l'étage de mise en forme.

Plusieurs éléments indiquent la possibilité d'amélioration du système à deux pôles. D'abord la forme de l'impulsion de sortie présente un rebond qui n'est pas souhaitable. Le retour à l'origine peut être amélioré, le gain doit aussi être augmenté d'un facteur environ 10; pour cela il sera plus adéquat de le réaliser dans un étage suivant pour éviter d'avoir un gain

trop élevé sur l'étage d'entrée. En outre pour ce qui a trait au bruit, les performances sont comparables à celles obtenues dans la mise en forme semi gaussienne, avec une contribution dominante du de bruit de grenaille (shot noise) qui pourrait être réduite. Pour cette raison nous avons essayé d'améliorer l'étage d'entrée en nous orientant vers à un circuit amplificateur et de mise en forme à trois pôles.

Intuitivement, la partie réelle des pôles (dans le plan s) représente la décroissance exponentielle de l'impulsion de sortie, et est dominée par le pôle le plus lent. Si tous les pôles ont la même composante réelle, aucun ne domine le temps de réponse et ceci devrait raccourcir le retour à l'origine du système. Ohkawa et ses collègues proposent un rapport optimum de 1.1 entre le pole réel et la partie imaginaire des deux pôles complexes. En outre, si l'on prend trois pôles ayant la même composante réelle, cela simplifie considérablement les calculs, de sorte qu'ils peuvent être effectués analytiquement. Pour toutes ces raisons nous avons décidé d'analyser la configuration avec trois pôles alignés sur une droite. Toute l'analyse exécutée pour le système à deux pôles conjugués a été refaite pour un système à trois pôles, cette dernière configuration donnant de meilleures performances en ce qui concerne tous les paramètres du circuit.

Une analyse plus réaliste a été conduite en tenant compte du temps de montée du préamplificateur. Le calcul du signal de sortie du système à quatre pôles pour une charge d'entrée Q_{in} = 5000 e^- montre un maximum de 192 mV à 26 ns, alors que le retour à l'origine à 1% intervient après 114 ns. Le préamplificateur a lui même un retour très rapide à l'origine, environ 65 ns (à 1%). Le bruit électronique total est de ENC_T = 117 électrons rms.

L'analyse détaillée du comportement de tous les étages qui composent le circuit d'entrée est présentée, avec les résultats de simulation, qui suivent de près les résultats donnés par l'analyse théorique. Un diagramme simplifié du schéma retenu pour mettre en application le système à trois pôles est montré sur la figure 1. Un étage qui réalise le pôle réel τ_{p3} et un gain A_3 est ajouté au système à deux pôles.

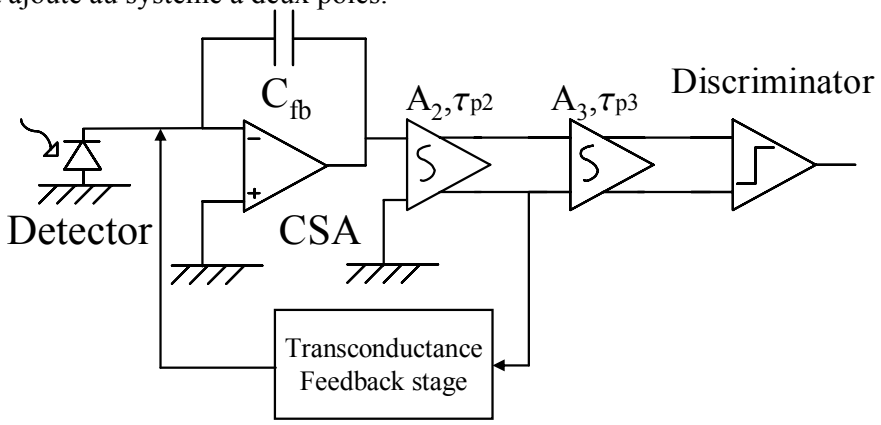

Figure 1: Schéma simplifié du système à trois pôles; un étage qui réalise un le pôle réel τ_{p3} et un gain A_3 est ajouté au système à deux pôles.

Les mesures faites sur les circuits de test ont montré clairement que le minimum du seuil de détection du signal et la performance en bruit sont dominés principalement par l'interférence numérique-analogique. Dans la nouvelle conception cette question a été abordée en priorité: par exemple, l'amplificateur de charge a été conçu avec une entrée différentielle qui rejette le bruit de mode commun (comme le bruit de substrat). D'ailleurs il est réalisé avec des transistors d'entrée canal-p, pour pouvoir être en mesure de court-circuiter le caisson du transistor avec la source et ainsi éviter l'injection de bruit à travers la capacité caisson-grille. Pour la même raison, tous les autres étages dans la chaîne d'entrée sont différentiels. Le système a été également optimisé pour réduire au minimum le couplage capacitif de pixel à pixel (aucune impulsion parasite due à l'interférence n'est observée dans les simulations jusqu'à une charge d'entrée de 55000 e^- pour un seuil à 2300 e^-), et pour récupérer rapidement après un signal d'entrée très grand. En effet, le système est prêt en environ 200 ns à traiter une autre impulsion, après avoir reçu une charge d'entrée de 50000 e^-, pour et un seuil de détection d'environ 2300 e^-. Dans ce cas la composante lente de signal excentre légèrement le seuil; 50000 e^- est la charge la plus élevée prévue dans l'expérience LHCb). L'étage d'entrée a été conçu pour que l'impact d'une variation de la température soit négligeable; une attention particulière a été aussi apportée au contrôle de la phase sur les boucles de contre réaction (72° pour la boucle à haute fréquence, et 82° pour la boucle de basse fréquence) et tout a été fait pour réduire l'injection de bruit par les lignes d'alimentation.

Le discriminateur, qui compare le signal analogique produit par l'étage d'entrée à un seuil global et produit une impulsion numérique si le signal est plus haut que le seuil, a été conçu pour être basse puissance. En outre, il doit réduire au minimum la différence de temps de réponse dû à des impulsions d'entrée d'amplitudes très différentes («time walk»). Le «time walk» à 20 ns a été simulé, et il est meilleur que 140 électrons.

La dispersion calculée du seuil (par rapport au préamplificateur d'entrée) est de 70 e^- rms, mais un circuit de correction a été ajouté dans le discriminateur pour réduire cette dispersion au cas elle serait plus forte que prévu.

Une description de la partie numérique du circuit et des circuits périphériques est donnée dans les dernières sections, de même que quelques considérations sur l'architecture physique, expliquant quelques stratégies importantes de placement des éléments pour améliorer les performances du circuit. Le composant a été envoyé à la fabrication, puis testé intensivement, d'abord non relié à un détecteur, puis assemblé pour former des «singles» et des «ladders». Les résultats expérimentaux essentiels sont présentés au **chapitre 8**.

La première version du circuit (ALICE1LHCb) a montré une limitation de la fréquence d'horloge maximum utilisable (15-18 Mégahertz); cela a conduit à la mise au point d'une deuxième version du circuit (LHCbpix1) pour l'expérience LHCb seulement, qui fonctionne avec une horloge de 40 Mégahertz. Un problème avec le Convertisseur Numerique Analogique (CNA), a été résolu par l'augmentation de la tension d'alimentation de 1.6V à

1.8V et par une séquence appropriée des opérations de test pour la puce ALICE1LHCb. Il a été élimine dans la version LHCbpix1.

Des structures de test sont présentes sur le circuit, pour simuler l'injection de charge lorsque la puce n'est pas encore connectée avec un senseur. Une différence de tension connue V_{in} et appliquée sur une capacité de test C_{in}, ce qui se traduit par une charge d'entrée $Q_{in} = C_{in} V_{in}$.

La calibration du circuit mène à la conclusion que V_{in} = 1 mV correspond à Q_{in} = 100 e^- pour ce qui concerne les tests faits sur LHCbpix1. En raison d'un problème dans l'uniformité du circuit d'injection de charge de test sur le circuit ALICE1LHCb, il n'est pas aussi facile d'estimer le facteur de conversion tension/charge pour le circuit ALICE1LHCb; on a estimé que la valeur moyenne sur le circuit est d'environ 0.5-0.65 mV/100 e^-.

Des résultats sur des circuits sans le senseur sont présentés. Pour un signal d'entrée de Q_{in} = 5000 e^- la sortie différentielle du deuxième étage de mise en forme atteint son maximum après environ 31 ns avec une tension crête d'environ 125 mV. L'étage d'entrée retourne à l'origine 300 ns après une impulsion d'entrée voisine de 100000 e^-, et il faut remarquer que même avec un signal d'entrée si élevé, aucune interférence électrique n'est induite dans les pixels voisins. Le bruit du circuit est d'environ 130 à 140 e^- rms avec une dispersion de 13 e^- rms. Le seuil minimum est à environ1 100 à 1200 e^-, et la dispersion de seuil, moins de 100 e^- rms. Un premier algorithme a permis de ramener cet étalement à environ à 43 e^- rms utilisant le circuit d'ajustement du seuil de pixel à pixel.

Nous avons aussi mis au point une procédure de test des circuits sur un testeur de plaquette. Ce système permet de choisir les bons circuits qui vont être assemblés sur un détecteur.

Le circuit a été irradié avec des rayons X de 10 keV, jusqu'à une dose ionisante de 100 kGy (SiO_2). On n'observe aucun changement significatif des courants jusqu'à la dose d'irradiation maximum. Ceci confirme l'efficacité des techniques de durcissement HBD pour éviter les courants parasites induits par rayonnement. Un léger changement de la valeur du seuil de pixel est induit (environ 100 e^-), il peut être pratiquement ramené à sa valeur de pré-irradiation en agissant sur le CNA qui détermine le seuil global du circuit.

Un circuit LHCbpix1 a été porté à différentes températures entre 30°C à 90°C; l'augmentation du seuil est proche de 600 e^-, mais ce décalage peut être récupéré en jouant sur la valeur du CNA qui fixe le seuil globale du circuit. Le circuit a été irradié à Louvain-La-Neuve avec des ions lourds produisant un LET compris entre 6 et 120 $MeVmg^{-1}cm^2$. Ceci a permis de tracer la courbe de section efficace de SEU en fonction du LET. Le composant a été irradié également avec des protons de 60 MeV, ce qui nous permet de calculer la section efficace pour les SEU à cette énergie, soit 3 10^{-16} cm^2. Pour l'expérience ALICE cela se traduit par moins de un SEU toutes les 10 heures. On n'a observé ni ruptures de grille (SEGR), ni SEL.

Une section de ce chapitre montre les résultats de test de circuits connectés aux détecteurs. Pour un signal d'entrée de Q_{in} = 5000 e^- la sortie différentielle du second étage de mise en forme (circuits LHCbpix1), atteint son maximum après environ 33 ns et la tension maximum est d'environ 131 mV. Le seuil minimum d'un circuit assemblé est légèrement plus élevé que pour un circuit sans détecteurs (~1300 e^-), mais la dispersion de seuil reste en dessous de 100 e^- rms. Le bruit augmente jusqu'à environ 170 e^- rms (dû à la capacité de la cellule pixel du senseur ; la dispersion est de 17 e^- rms). Le «time walk» à 20 ns a été mesuré, mais en raison de la difficulté de la mesure, nous avons seulement estimé une limite supérieure de ~ 250 e^-. En outre le couplage capacitif de charge a été mesuré, aucun signal n'est induit dans un pixel voisin jusqu'à une charge d'entrée d'environ 55000 e^-.

Quelques circuits ont été irradiés avec une source de ^{55}Fe, montrant la sensibilité du circuit aux signaux d'entrée très bas, et avec des rayons X de 10 keV, jusqu'à une dose de 100 kGy (SiO_2). A cause de l'assemblage mécanique, c'est équivalent à 178 kGy (Si) dans le détecteur et à 10 kGy (SiO_2) dans le circuit. Le bruit moyen et sa dispersion ne changent pas à ce niveau d'irradiation. La valeur moyenne du seuil et sa dispersion varient d'environ 20% à la dose maximum d'irradiation.

Toutes les mesures effectuées donnent des résultats conformes au cahier des charges établi par les deux expériences.

Plusieurs assemblages ont été testés pendant quatre campagnes de test au faisceau d'essai dans la ligne de faisceau H4 au CERN, en utilisant des pions de 150 GeV/c, des protons de 350 GeV/c, de l'indium de 158 GeV/c et enfin des protons de 120 GeV/c. Les mesures d'efficacité de détection, de résolution temporelle, balayages de seuil, balayages de polarisation et des mesures à différents angles du détecteur par rapport au faisceau ont été effectuées. Les résultats montrent une fois de plus la pleine fonctionnalité du système.

L'expérience LHCb a testé le circuit LHCBpix1 dans un tube HPD avec un faisceau de particules (pions et électrons), et a détecté les premiers anneaux Cherenkov.

Les circuits ALICE1LHCb et LHCbpix1 exigent un certain nombre de dispositifs auxiliaires externes: le contrôle digital est fait par le circuit pilote digital (Digital Pilot Chip, DPC); les circuits pixel nécessitent également quelques polarisations analogiques externes. De plus il est nécessaire de mesurer quelques signaux analogiques sur le modules électroniques de contrôle MCM des circuits pixel (par exemple les sorties des convertisseurs CNA des circuits pixel, les tensions d'alimentation et la température). Ces fonctions analogues auxiliaires sont exécutées par le circuit pilote analogue (Analogue Pilot Chip, APC). La conception du circuit APC et quelques résultats de tests sont présentés dans le **chapitre neuf**.

L'APC est un circuit intégré fonctionnant en mode mixte avec une puissance totale d'environ de 50 mW, conçu avec les techniques de durcissement HBD pour améliorer la tolérance au rayonnement. Il contient plusieurs blocs qui fournissent les fonctions analogiques demandées. Un prototype du circuit a été conçu et testé, puis une seconde version légèrement

modifiée a été développée et envoyée à la fabrication. Les tests effectués sur les divers blocs montrent la pleine fonctionnalité de la version finale de ce circuit.

Un des éléments du APC est la cellule «bandgap» qui doit fournir une tension de référence précise, avec une sensibilité réduite aux fluctuations de température et des paramètres du procédé. Plusieurs circuits de référence pour le convertisseur analogique numérique (CAN) et les six CNA du circuit APC ont été développés. Les mesures des sorties du «bandgap» et des références montrent des erreurs relatives inférieures à 1%, et une stabilité de température de moins de 10 à 15 mV de décalage pour une variation de température ΔT de 50 °C).

Les six convertisseurs CNA de 8 bits fournissent les tensions de référence du circuit ALICE1LHCB ou du LHCBpix1. La conception du CNA est une version modifiée de celle utilisée pour le Circuit Pixel lui-même; cette version révisée doit supporter une tension d'alimentation plus élevée (2.5V), et les tensions de référence pour ces CNA sont fournies sur le circuit. Un buffer de tension est relié à la sortie de chaque convertisseur CNA. Les mesures montrent que les circuits sont entièrement fonctionnels; une dispersion maximum entre puces (pour un même code) de quelques dizaines de millivolts est observé. Si une précision de l'ordre du millivolt est nécessaire, cela demandera une base de données de calibration, dans laquelle la valeur de tension de sortie de chaque CNA pour chaque code d'entrée est stockée. La sensibilité à la température du CNA est moins de 15 mV de décalage pour une variation de température de 50 °C.

L'APC contient un convertisseur analogique numérique (CAN) de 10 bits, travaillant par approximations successives, dont la consommation est très basse (1 mW) et la modularité élevée. La fréquence d'échantillonnage maximum est d'environ ~ 2.5 Msample/sec (pour une horloge à 30 mégahertz). Le CAN est employé pour la conversion de 16 niveaux continus ou faiblement variables et il utilise une horloge de 10 MHz. La gamme d'entrée du CAN est placée entre 0.513 V et 1.925 V. Un multiplexeur 16 à 1 et un buffer analogique ont été placés devant le convertisseur CAN, pour pouvoir lire les 16 signaux différents. Les mesures faites sur le CAN prouvent qu'il est entièrement fonctionnel, et convertit avec précision les valeurs d'entrée, a part d'un très faible décalage (de l'ordre de quelques millivolts) dû au buffer d'entrée.

L'étage de détection de courant a pour objectif la lecture des sorties des circuits CNA de courant des circuits ALICE1LHCb et LHCbpix1, en respectant les conditions de polarisation des étages de sortie des CNA. Les mesures prouvent que cet étage de détection est capable de lire les courants entrants et sortants sans erreur appréciable. Le circuit contient quatre sources de courant pour la mesure de la température, conçues pour être indépendantes des variations de la température et de l'alimentation du circuit. Leur fonctionnalité a été examinée simulant le comportement d'une résistance thermosensible avec des résistances de valeur connue.

L'étage de détection de courant est précis à quelques pour cent et est très peu sensibles à la résistance de charge, comme prévu.

Le circuit de contrôle numérique JTAG qui fournit tous les signaux numériques nécessaires aux autres blocs est entièrement fonctionnel. Tous les blocs dans le circuit fonctionnent selon les caractéristiques, et le circuit peut être employé dans cette version pour équiper le module MCM de contrôle.

En conclusion, nous pouvons dire qu'un circuit complet a été conçu pour la lecture des détecteurs à pixels en silicium, utilisés pour les expériences ALICE et LHCb. Le circuit a été conçu en technologie CMOS standard 0.25 μm, en employant des techniques de durcissement HBD pour améliorer sa tolérance au rayonnement. Un étage d'entrée a été conçu pour garantir la fonctionnalité du circuit dans le cas d'un fort taux d'occupation du canal.

Deux versions du circuit ont été réalisées: ALICE1LHCb, pour l'expérience ALICE, et LHCbpix1 pour l'expérience LHCb. Les tests montrent le succès de la conception, car le circuit est entièrement fonctionnel et ses caractéristiques répondent au cahier des charges. Son bruit est suffisamment bas (autour 170 e^- rms pour une puce assemblée sur son détecteur), le seuil de détection minimum convient (autour 1300 e^- pour une puce assemblée sur son détecteur) et la dispersion de seuil est faible (autour 100 e^- rms pour une puce assemblée sur son détecteur). Ces caractéristiques permettent d'envisager l'utilisation de ce circuit sur les deux expériences visées. En outre la tolérance aux rayonnements du circuit a été mesurée dans plusieurs environnements. Il peut tenir jusqu'à 100 kGy en dose ionisante totale sans dégradation sérieuse des paramètres, et un ajustage des polarisations des circuits ramène ses caractéristiques pratiquement aux valeurs de pré-irradiation. Ceci prouve l'efficacité de l'approche de durcissement HBD. Le circuit a été exposé à un faisceau de particules produisant un taux d'évènements élevé et le circuit a gardé sa pleine efficacité, montrant la fonctionnalité du nouveau schéma du circuit d'entrée. Une demi-barrette équipée du module MCM contenant les circuits de commande (les circuits pilotes analogiques et numériques, APC et DPC, et la connexion optique, GOL) a été testée avec succès dans un faisceau de particules en montrant une entière fonctionnalité. Enfin, l'expérience LHCb a testé le circuit LHCBpix1 dans un tube HPD avec un faisceau de particules (pions et électrons), et a détecté les premiers anneaux Cherenkov.

Le succès du circuit a donné l'impulsion à plusieurs autres expériences au CERN, qui ont remplacé les dessins en technologie durcie par procédé avec des dessins qui utilisent les techniques HBD.

Introduction

The work presented in this thesis was carried out in the Microelectronic Group at CERN, the European Laboratory for Particle Physics, a laboratory located near Geneva (Switzerland) born in the fifties to give the European scientists the means to study High Energy Physics (HEP). In the past years, by means of particle accelerators conceived and realised at CERN (in particular with the Large Electron Positron) it was possible to develop the Standard Model, a theory which tries to explain matter in terms of forces and particles. This model was tested only to a certain extent, and in particular it does not have an explanation of what causes the fundamental particles to have masses. The simplest idea is called the Higgs mechanism. This mechanism involves one additional particle, called the Higgs boson, and one additional force type, mediated by exchanges of this boson. To test this hypothesis, as well as several other phenomena and theories, a new particle accelerator is at present under construction at CERN, the Large Hadron Collider (LHC). The LHC will be the most powerful accelerator ever built. The ALICE experiment, in particular, will focus on finding and studying the Quark Gluon Plasma (QGP), a state of the matter in which the energy density is so high that gluons and quarks do not stick together to form protons and neutrons. To create QGP in the laboratory, scientists collide heavy-ions into each other at very high energy, squeezing the protons and neutrons together to try and make them melt.

In the LHC experiments (ALICE, LHCb, ATLAS, CMS, TOTEM) particles will be accelerated to reach energies in the TeV range and collided head on at very high luminosities (10^{34} $cm^{-2}s^{-1}$ for protons and $1.95 \cdot 10^{27}$ $cm^{-2}s^{-1}$ for lead ions). For this reason, the radiation levels that can be attained are very high, especially for the inner detectors, very close to the interaction point ($2.5\ 10^3$ Gy of total ionising dose and $2.95\ 10^{12}$ MeV n_{eq} / cm^2 of equivalent neutron fluence for the ALICE experiment in ten years of running of the LHC, but much higher radiation levels can be attained by other experiments).

This poses major problems for the electronics which has to be used close to the interaction point, which is usually the electronics for the particle tracking detectors. For this reason, and for the extreme peculiarity of the tracking electronics, no commercial components are available for this task, and dedicated electronics has to be designed on purpose. A possibility consists in addressing the problem of radiation tolerance at the level of the technological process; in particular some specialised vendors provide a qualified radiation hard process. These technologies suffer from problems related with cost, circuit performance,

process stability, yield and future availability. This lead CERN to set up a research project to investigate the possibility of using a standard commercial CMOS technology, hardened with special design techniques (Hardening By Design, HBD), for HEP integrated circuits. The main advantage of this approach is the low cost, and the possibility to follow the rapid scaling down of commercial technologies. This approach was studied in detail in the CERN RD49 project, and the most interesting results are reported in [Ane00].

Several chips, which will be presented in this thesis, have been designed in a standard CMOS technology employing HBD techniques, to be used in the inner silicon pixel detector of the ALICE experiment. In effect, the two first pixel readout chips designed at CERN, the Omega2 and Omega3 chips (implemented in a commercial 3 and 1 μm CMOS technology respectively), proved the feasibility of using pixel detectors for HEP, even if their low radiation tolerance showed that an improvement was needed in that respect. The ALICE1test chip was a first pixel readout prototype chip designed in a standard 0.5 μm CMOS technology but with layout techniques for radiation tolerance. It withstood a Total Ionising Dose (TID) between 6 kGy and 17 kGy(SiO_2) depending on the type of radiation. No radiation-induced leakage current was observed.

The encouraging results of the RD49 project and of the ALICE1test test results were the starting point of this thesis work. In effect, the availability for CERN of the more advanced 0.25 μm technology lead to the design of the ALICE2test chip, based on the previous prototype but with some changes. This prototype withstood much higher radiation doses, up to an X-ray TID of 300 kGy(SiO_2), again without radiation-induced leakage currents. All the electrical parameters (peaking time, noise, minimum threshold, threshold dispersion) were matching the requirements of two CERN experiments, ALICE and LHCb.

The ALICE2test chip was the starting point for the conception and the development of the ALICE1LHCb chip. The ALICE1LHCb chip is a matrix of 32 columns each containing 256 readout cells (for a total of 13 million transistors), measuring 13.5×15.8 mm^2. Five chips are connected by means of an array of microscopic metallic balls (bump-bonding) to a big detector (160 columns $\times$ 256 rows) to form the ALICE Silicon Pixel Detector (SPD) basic detecting block, *the ladder*. One chip is bump bonded to a detector (of the same dimensions, 32 columns $\times$ 256 rows) to form a *single*, the basic detection element for the LHCb Hybrid Photon Detector, which is also used for the tracker of the experiment NA60.

This thesis reports the work done in the development of the hybrid pixel particle detector for the ALICE SPD and for the LHCb RICH detector. The HBD design approach is described in detail, to understand the challenges and the beneficial effects that it implies. The design and testing of the silicon particle sensors is presented. Most of all, the concept, design and testing of the ALICE1LHCb pixel readout chip are presented.

The chip, designed with HBD techniques, makes use of a different front end scheme with respect to the more classical approach of charge integration plus pole zero cancellation and

semigaussian shaping. This approach has been conceived to be used in high multiplicity environments, and to have a fast return to zero of all the blocks which compose the readout chain, whilst keeping low noise (< 200 e^-) and low power consumption (of the order of 50-100 μW/channel). The chip was tested extensively, both in the laboratory and in a beam of particles.

The ALICE SPD and the latest developments of its components are also presented. This helps also the reader to understand the high level of complexity of the system, which will have in total about 10 million active readout channels (1200 ALICE1LHCb chips).

In particular, the design and testing of an auxiliary chip for the ALICE1LHCb pixel chip, the Analogue Pilot Chip, are discussed.

A description of the contents of the thesis, chapter by chapter, follows.

Chapter 1: This chapter gives an introduction to high energy physics and particle physics at CERN. The basic concepts of the Standard Model, of particle accelerators and of the challenges posed by high energy physics are discussed. The ALICE experiment is presented.

Chapter 2: The core of this thesis work is the design and development of a hybrid pixel detector to be used in the inner layer of the ALICE Inner Tracking System (ITS). For this reason this chapter gives an overview of the most important semiconductor particle detectors for HEP, and in particular of pixel detectors. A section of this chapter is dedicated to radiation damage in semiconductor particle sensors, to define the most important physical quantities related to these phenomena.

Chapter 3: After an introduction about the full ALICE ITS, the chapter focuses on its two innermost layers, which form the ALICE SPD. All the components of the ALICE SPD are presented, as well as some test results.

Chapter 4: A hybrid detector is formed by a silicon pixel sensor bump-bonded to a pixel readout chip. This chapter deals with the design and testing of the sensor of the ALICE SPD and LHCb RICH hybrid detectors.

Chapter 5: The idea which is behind the ALICE1LHCb chip is to realise a full pixel readout chip in a standard CMOS technology, applying special layout techniques (HBD) to obtain radiation tolerance. For this reason the chapter begins with an introduction about the effects of radiation on Integrated Circuits, and then concentrates on the HBD design approach. Its basic building block, the Enclosed Layout Transistor (ELT) is discussed in detail (modelling, noise, matching, radiation tolerance), and results are presented which show the effectiveness of the HBD approach both at transistor level and at system level.

Chapter 6: The first working hybrid pixel detectors for high energy physics were designed in the 1990's with. For example, the Omega2 and Omega3 chips were fully functional but prone to radiation damage. Two prototype chips, built on the experience of the

Omega2 and Omega3 chips were designed and tested, and are presented in this chapter: the ALICE1test and ALICE2test chips.

Chapter 7: This chapter presents the design and simulations of the front-end of the ALICE1LHCb pixel chip. This chip makes use of a different front end scheme with respect to the more classical approach of charge integration plus pole zero cancellation and semigaussian shaping, to be used in high multiplicity environments. In the first section all the requirements resulting in our design choices are summarised. The second section is devoted to an in depth analysis of the choice of the front-end scheme. Different versions of the front-end chosen are discussed (two, three and four poles). The detailed analysis of the implementation of all the stages which compose the front-end chip is presented, with simulation results, which follow closely the theoretical analysis. A description of the digital back-end of the chip and of the chip periphery is given in the last sections, as well as some layout considerations explaining some important layout strategies employed to improve chip performance.

Chapter 8: The chip was sent to fabrication and was then extensively tested, both bare (i.e. not connected to a sensor) and bump-bonded to form "singles" and "ladders". A second version of the chip (LHCbpix1) was designed to meet the LHCb requirement about the clock frequency. The main results obtained testing the two versions of the chip in the lab and during test beams are presented in this chapter.

Chapter 9: The ALICE1LHCb and LHCbpix11 Pixel Chips require a number of external auxiliary features: the digital control is demanded to the DPC, but the Pixel Chip needs also some external analogue biases. Moreover there is the need to measure some analogue signals (i.e. Pixel Chip DACs outputs, voltage supplies and temperature). These auxiliary analogue functions are performed by the Analogue Pilot Chip (APC). The design of this chip and some test results are presented in this chapter.

Conclusions: This final chapter presents the conclusions which can be drawn from this work. The main results are presented and discussed.

Contribution of this work: The pixel detector project has been a team effort, in which I participated. Its success and many of the results published here have been obtained through the efforts of many people. This section helps to clarify the contribution of this thesis work in this large project.

The full thesis in electronic format, as well as some add-ons which are not included in this thesis work for a matter of space can be downloaded at the following web address:

http://dinapoli.home.cern.ch/dinapoli/thesis.htm

Chapter 1
An introduction to particle physics at CERN

At the beginning of the past century achievements in Europe dominated progress in physics, but the conflicts of the 1930s and 40s interrupted this trend, because many scientists were forced to leave their countries. With the return of peace, in the 50s, the Americans started to invest in basic science, building the first powerful particle accelerators. For European research, co-operation was the only possibility. The creation of a European Laboratory was recommended at a UNESCO meeting in Florence in 1950, and less than three years later a Convention was signed by 12 countries of the Conseil Européen pour la Recherche Nucléaire. CERN was born, the prototype of a chain of European institutions in space, astronomy and molecular biology. The laboratory, which is located near Geneva (Switzerland), is today also called European Laboratory for Particle Physics, as the work at the laboratory went beyond the study of atomic nucleus [Web00].

The work presented in this thesis was carried out in the Microelectronic Group at CERN.

1.1 The Standard Model

The Standard Model is the name given to the current theory of fundamental particles and how they interact, which tries to explain matter in terms of forces and particles [Web05, Web06]. It can explain, in a simple and comprehensive theory, all the hundreds of particles and complex interactions with only twelve fundamental particles, six types of quarks and six types of leptons[1]. These particles interact via some basic forces, which are "transported" by four force carrier particles (Table 1.1). All known components of the matter are composites of quarks and leptons, which interact by exchanging force carrier particles; however not all the combinations of quarks and leptons are stable. Some of them can be created for a very short time in particle accelerators, before they decay in some more stable particles [Din04].

All forces between objects are due to interactions. All particle decays are due to interactions. The four fundamental interactions responsible for all observed processes are:

[1] Plus their antimatter counterpart: antiparticles look and behave just like their corresponding matter particles, except they have opposite charges.

1. **Strong interactions**, responsible for forces between quarks and gluons and, as a consequence, nuclear binding;
2. **Electromagnetic interactions**, responsible for electric and magnetic forces and, as a consequence, atom binding;
3. **Weak interactions**, responsible for the instability of all but the least massive fundamental particles in any class;

Gravitational interactions, responsible for forces between any two objects due to their energy (which, of course, includes their mass).

	First generation	*Second generation*	*Third generation*	**Force carriers**
Quarks	**u** : up quark	**c** : charm quark	**t** : top quark	γ : photon
	d : down quark	**s** : strange quark	**b** : bottom quark	**g** : gluon
Leptons	ν_e : electron neutrino	ν_μ : muon neutrino	ν_τ : tau neutrino	**Z** : Z boson
	e : electron	μ : muon	τ : tau	**W** : W boson

Table 1.1: Fundamental particles (columns 1-3) and force carriers (column 4).

Each type of interaction has a characteristic set of force carrier particles associated, with quantum excitation of the force field related to that interaction. The carrier particles either are produced during all processes involving that type of interaction, or appear in intermediate stages of them. Forces between particles can be described in terms of static (unchanging) force fields and exchanges of force carrier particles between the affected particles. One important thing to remember about force carriers is that a particular force carrier particle can only be absorbed or produced by a matter particle which is affected by that particular force.

1. **Gluons** are the carrier particles of strong interactions.
2. **Photons** are the carrier particles of electromagnetic interactions.
3. **W and Z bosons** are the carrier particles of weak interactions.

The Standard Model was the triumph of particle physics of the 1970's, and CERN gave a major contribution to its validation. It incorporated all that was known at that time and has since then successfully predicted the outcome of a large variety of experiments. This model, which is today a well-established theory applicable over a wide range of conditions, explains what the world is and what holds it together.

However, although **gravity** is clearly one of the fundamental interactions, the Standard Model cannot satisfactorily explain it. The name for the carrier particle of gravitational interactions is the **graviton**. The status of this particle is still tentative, because the theory is incomplete and there has been no good experimental evidence that it exists. Moreover, gravity

has a different mathematical structure and no fully satisfactory quantum theory of gravitational interactions via graviton exchange has been identified.

In addition, we do not know what causes the fundamental particles to have masses. The simplest idea is called the Higgs mechanism. This mechanism involves one additional particle, called the **Higgs boson**, and one additional force type, mediated by exchanges of this boson. The Higgs particle has not yet been observed. Today we can only say that if it exists, it must have a mass greater than about 80 GeV/c^2. Searches for a more massive Higgs boson are beyond the scope of the present particle accelerators. Future facilities, such as the Large Hadron Collider at CERN, or upgrades of present facilities to higher energies are intended to search for the Higgs particle and distinguish between competing concepts.

1.2 Particle accelerators

A particle accelerator is a device used to produce high-energy high-speed beams of charged particles, such as electrons, protons, or heavy ions, for research in high energy and nuclear physics, synchrotron radiation research, medical therapies, and some industrial applications [Din04, Web00].

At CERN physicists investigate the constituents of matter at the subatomic level, where the typical distances are of the order of the 10^{-15} m or smaller, so one of the main problems is to find an appropriate way to "probe" such small objects. The simplest way an object can be probed is to irradiate it with waves, and then detect how waves are reflected by the object itself. The problem with using waves to detect the physical world is that the resolution of the image is limited by the wavelength used. As a rough rule of thumb, a wave can only probe down to distances equal to its wavelength. All moving particles of matter have wave properties, so it is possible to use a particle beam as a probe, provided that the wavelength of the particles is short enough. However, most of the particles around us in the natural world have fairly long wavelengths, so physicists apply the principle that a particle's momentum and its wavelength are inversely related. In particle accelerators the momentum of a probing particle is increased, thus decreasing its wavelength.

The second reason for particle accelerators is that quite often physicists want to study massive, unstable particles that have only a fleeting existence (such as the very massive top quark.) However, in the every day world there are only very low-mass particles. In this case the Einstein energy-mass equivalence principle is exploited. To use particles with low mass to produce particles with greater mass, the low mass particles are accelerated, giving them a very high kinetic energy, and then collided. During this collision, the particle's kinetic energy is converted into the formation of new massive (and unstable) particles.

An accelerator usually consists of a vacuum chamber surrounded by a long sequence of vacuum pumps, magnets, radio-frequency cavities, high voltage instruments and electronic circuits. Each of these pieces has its specific function.

The LHC will be a circular accelerator, with two counter-rotating beams of particles that will collide in specific locations of the beam pipe. For each collision, which is called an *event*, the physicist's goal is to count, trace and characterize all the different particles that were produced and fully reconstruct the process.

Each event is very complicated since lots of particles are produced. Most of these particles have lifetimes so short that they go an extremely short distance before decaying into other particles, and therefore leave no detectable tracks, so physicists have to look at particles' decay products, and from these deduce the particles' existence. To look for these various particles and decay products, physicists have designed multi-component detectors that test different aspects of an event. At each collision point along the accelerator beam pipe there is a different multi-component detector, called also *experiment*.

Each component of a modern detector is able to recognize and measure a special set of particle proprieties (e.g. energy, momentum, charge). When all these components work together to detect an event, individual particles can be singled out from the multitudes for analysis. Following each event, computers are used to collect and interpret the vast quantity of data from the detectors and then extrapolate results.

1.3 From the Proton Synchrotron to the Large Hadron Collider

Thanks to the knowledge and expertise gained in the sixties and seventies with the construction of the **Proton Synchrotron,** of the **Intersecting Storage Rings**, and of the **SPS proton synchrotron,** in 1989 CERN's **LEP electron-positron collider** went into operation, and made possible the precise testing of the Standard Model. By 1996, the LEP energy was doubled to 90 GeV per beam in LEPII, opening up an important new discovery domain.

During 2000-2006 the LEP is being removed from its tunnel and replaced by the **LHC (Large Hadron Collider)**, which is scheduled to come into operation in 2006. Figure 1.1 shows an aerial view of the CERN accelerators.

The LHC will consist of two *"colliding" synchrotrons* installed in the 27 km LEP tunnel. They will be filled with protons delivered from the SPS and its pre-accelerators at 0.45 TeV. Two superconducting magnetic channels will accelerate protons to 7-on-7 TeV (about 10 times that of LEP and the Tevatron[2]), providing the experiments with high brightness and high interaction rate collisions, after which the beams will counter-rotate for several hours, colliding at the experiments, until they become so degraded that the machine will have to be emptied and refilled. LHC can also collide beams of *heavy ions* such as lead with a total collision energy in excess of 1,250 TeV, about thirty times higher than at the RHIC[3]. Joint

[2] The Tevatron is the proton-antiproton TeV particle accelerator located at the Fermi National Accelerator Laboratory, in Illinois (USA) [Web01].

[3] The Relativistic Heavy Ion Collider (RHIC) is the heavy ions collider at the Brookhaven National Laboratory (BNL), in New York (USA) [Web02].

LHC/LEP operation can supply proton-electron collisions with 1.5 TeV energy, some five times higher than presently available at HERA[4].

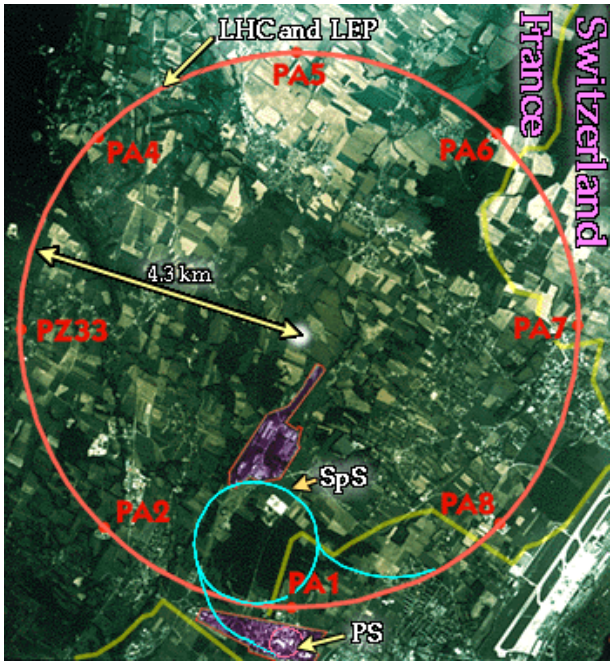

Figure 1.1: Aerial view of the CERN accelerators.

1.4 A Large Ion Collider Experiment (ALICE)

Five experiments are foreseen to take place at the LHC: ALICE, CMS, ATLAS, LHCb and TOTEM [Web04]. In this section the ALICE experiment will be presented in more detail.

1.4.1 The Quark Gluon Plasma

Scientists believe there was a Big Bang from which everything in the Universe emerged. Fifteen billion years later, the Universe is so huge that it would take light billions of years to cross. Yet in the beginning everything was squeezed into a tiny volume no bigger then a flea. All the particles which make up everyday matter, had yet to form. The quarks and gluons, which in today's cold Universe are locked up inside protons and neutrons, would have been too hot to stick together. Matter in this state is called **Quark Gluon Plasma**, QGP. Finding and studying QGP is ALICE's goal [ALI99, Din04, Web11].

To understand the first moments of the Universe's life, scientists must create QGP in the laboratory. To do so, they collide ions, atoms stripped of electrons, into each other at very high energy, squeezing the protons and neutrons together to try and make them melt.

[4] The Hadron-Electron Ring Accelerator HERA is the largest particle accelerator (electron-proton) at the German Electron Synchrotron in Hamburg (DESY, Deutsches Elektronen-Synchrotron) [Web03].

Physicists believe that the high energies at the LHC will be ideal for making QGP. A little bit of energy can knock atoms out of molecules or electrons out of atoms. With a little more energy, it is possible to knock protons and neutrons out of atomic nuclei. But no matter how much energy they have, it appears to be impossible to knock an individual quark or gluon out of its proton or neutron cage. This confinement poses a problem for studying quarks and gluons. One approach is to increase the volume in which quarks and gluons are confined, so they behave as if they were free, or deconfined. Deconfinement is a step on the way to making QGP, a mixture of quarks and gluons which has existed long enough for all the quarks and gluons to reach the same temperature, as it takes time for deconfined matter to thermalize.

Theorists predict different effects as matter heats up from its normal state into a deconfined one and cools down again. Some “signatures” of the QGP formation should be the reduced production of J/Ψ particles (composed of charm quarks and antiquarks), increased production of electron pairs and particles which contain strange quarks. Several of these signatures have been individually detected in previous CERN experiments, but the temperatures currently achieved appear to be only just enough to reach deconfinement. At the LHC lead ion collisions should heat matter up to temperatures at which QGP production becomes routine.

1.4.2 Recreating and detecting QGP

In fixed target experiments (like NA49 at CERN) heavy nuclei that are accelerated to near light velocity are bombarded onto nuclei in a thin metal foil (target). In effect, colliding heavier particles at higher energies raises the energy density and temperature of the collision, increasing the chances of deconfinement.

A head-on collision of a Lead (^{208}Pb) projectile with a Lead target nucleus, at the SPS beam energy of 160 GeV per nucleon in the Pb projectile, may compress and heat the nuclear matter contained in the two nuclei. It may thus reach the required energy density (20-fold higher than that of the initial nuclei) in a short-lived "fireball" volume. After about $8 \cdot 10^{-23}$ s this state expands, cools down and emits hadrons into the detector system.

There are about 1500-2000 charged particles created in each of theses collision events (and at the LHC, this will go up to 50,000). The detector system thus has to have an extreme spatial resolution to separate the particle tracks and record electronically the track of ionisation (and the ionisation strength) of each traversing particle. In addition, the flight time of the particles has to be measured. In ALICE this will allow to identify and determine momentum of all charged particles produced in Pb+Pb head-on collision. It is also possible to identify neutral "strange particles" by their secondary decay into charged particles.

1.4.3 The ALICE detector

ALICE is a general-purpose heavy-ion detector designed to study the physics of strongly interacting matter and the quark-gluon plasma in nucleus-nucleus collisions at the LHC. The experiment currently includes more than 750 physicists and ~70 institutions in 27 countries. Figure 1.2 shows a pictorial representation of the ALICE experiment [Din04, Web04].

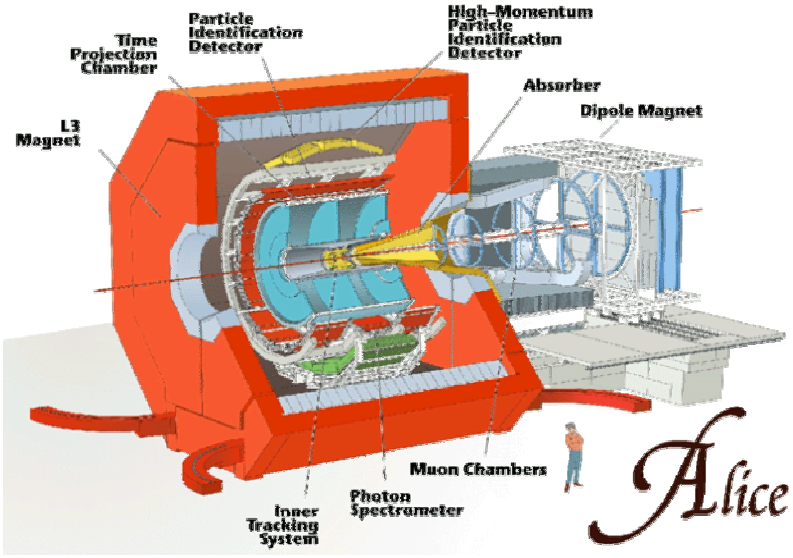

Figure 1.2: Pictorial representation of the ALICE experiment [Web04].

The detector is designed to cope with the highest particle multiplicities anticipated for Pb-Pb reactions (dN/dy ~ 8000)[5] and it will be operational at the start-up of the LHC. In addition to heavy systems, the ALICE Collaboration will study collisions of lower-mass ions, (this is a way to vary the energy density), and protons (both pp and p-nucleus), which provide reference data for the nucleus-nucleus collisions.

ALICE consists of a central part, which measures event-by-event hadrons, electrons and photons, and a forward spectrometer to measure muons. The central part, which covers polar angles from 45° to 135° ($|\eta| < 0.9$) over the full azimuth, is embedded in the large L3 solenoidal magnet. It consists of an inner tracking system (ITS) of high-resolution silicon tracking detectors, a cylindrical Time Projection Chamber, two particle identification arrays of time-of-flight and Cerenkov counters and a single-arm electromagnetic calorimeter. The forward muon arm (2-9° , η = 2.5-4) consists of a complex arrangement of absorbers, a large

[5] dN/dy is the particle density per unit rapidity. The rapidity y is defined as $y = 0.5 \ln [(E+p_l)/(E-p_l)]$, where E and p_l are the energy and longitudinal momentum of the given particle, respectively. When calculated in the center of mass system, values of rapidity around zero tag particles coming from the central fireball, which can be distinguished from the particles coming from the target and projectile nuclei. In the limit E>>m the rapidity can be approximated with the pseudo-rapidity η. η is an angular variable defined by $\eta = 0.5 \ln [(p+p_l)/(p-p_l)]$ which is related to the angle between the particle being considered and the undeflected beam. η can be used (instead of y) when the particle is not identified.

dipole magnet, and 14 planes of tracking and triggering chambers. The set-up is completed by a set of zero-degree calorimeters located far downstream in the machine tunnel, and a forward multiplicity detector which covers a large fraction of the phase space ($|\eta| < 4$).

1.5 Summary

The work presented in this thesis was carried out in the Microelectronic Group at CERN the European Laboratory for Particle Physics, a laboratory located near Geneva (Switzerland) born in the fifties to give the European scientists the means to study High Energy Physics.

The Standard Model is the name given to the current theory of fundamental particles and how they interact, which tries to explain matter in terms of forces and particle. It can explain, in a simple and comprehensive theory, all the hundreds of particles and complex interactions with only twelve fundamental particles, three forces and four force carriers. Gravity is clearly one of the fundamental interactions, but the Standard Model cannot satisfactorily explain it. Moreover, we do not know what causes the fundamental particles to have masses. The simplest idea is called the Higgs mechanism. This mechanism involves one additional particle, called the **Higgs boson**, and one additional force type, mediated by exchanges of this boson. The main task of the particle accelerator at present under construction at CERN, the Large Hadron Collider (LHC) is to search for the Higgs particle and distinguish between competing concepts.

A particle accelerator is a device used to produce high-energy high-speed beams of charged particles, such as electrons, protons, or heavy ions, for research in high energy and nuclear physics, synchrotron radiation research, medical therapies, and some industrial applications. An accelerator usually consists of a vacuum chamber surrounded by a long sequence of vacuum pumps, magnets, radio-frequency cavities, high voltage instruments and electronic circuits. The LHC will consist of two *"colliding" synchrotrons* installed in the 27 km LEP tunnel. They will be filled with protons delivered from the SPS and its pre-accelerators at 0.45 TeV. Two superconducting magnetic channels will accelerate protons to the extremely high energies (7-on-7 TeV) needed to study Higgs physics. LHC can also collide beams of *heavy ions* such as lead with a total collision energy in excess of 1,250 TeV.

Five experiments are foreseen to take place at the LHC: ALICE, CMS, ATLAS, LHCb and TOTEM. A brief description the ALICE experiment was presented, starting from its research goal. At very high energy densities, such as those present in the first moments after the Big Bang, the quarks and gluons, which in today's cold Universe are locked up inside protons and neutrons, would have been too hot to stick together. Matter in this state is called Quark Gluon Plasma, QGP. Finding and studying QGP is ALICE's goal. ALICE is a general-purpose heavy-ion detector designed to study the physics of strongly interacting matter and the quark-gluon plasma in nucleus-nucleus collisions at the LHC. The detector is designed to cope with the highest particle multiplicities anticipated for Pb-Pb reactions.

Chapter 2
Semiconductor particle detectors for High Energy Physics

In the 1980's it became clear that there was a need for compact, high spatial resolution, high speed radiation detectors, with direct[1] and efficient conversion of the radiation energy. This was especially true for high energy physics trackers. To replace photographic films, gas filled wire chambers and various other types of arrays, several solutions based on semiconductor detectors were conceived in these years.

This chapter will give the reader an idea about the most important semiconductor particle detectors for High Energy Physics. This will help to better understand the different use done in the CERN experiments of the various types of detectors, and in particular of the hybrid pixel detectors, used in the inner layers of the ALICE Silicon Pixel Detector.

Although silicon is the most widely used semiconductor material for particle detection, due to its room temperature operation and wide availability, some other semiconductors have been employed and tested such as germanium, gallium arsenide, diamond, cadmium telluride or mercuric iodide. An overview of the semiconductor detectors basic properties can be found in [Mil71, Sze81, Din04].

2.1 Semiconductor detectors for trackers in high energy physics

Several different types of silicon detectors are used in high energy physics, to match the different needs in terms of physics requirements, technical specifications, cost and time to develop the solution chosen. The following sections give an overview of the main types of semiconductor detectors used as trackers in high energy physics.

[1] In semiconductor detectors the particle energy is deposited in the detector, and directly sensed. In several other types of detectors the particle energy is first transformed (for example in light) and then sensed. In wire chambers for example the particle ionises the chamber gas, producing visible light that was "detected" by eye.

2.1.1 Microstrip detectors

Silicon microstrip detectors were proposed in the 80's as an alternative to Multiwire Proportional Chambers (MPC), drift chambers or scintillation counters. Although it was not possible to cover large areas, microstrip detectors presented much better performance in terms of precision, rate and multiplicity capability [Hei80, Hei88, Wyl97, Rie98].

A microstrip detector, schematically represented in Figure 2.1, is a thin (usually about 300 μm), high resistivity silicon substrate (n-type in the picture) where thin strips are implanted (p^+ in the picture) to form the detecting diode structures. The strips can have a pitch from 10 μm to some hundreds of microns, a length up to several cm, and each can be connected to a dedicated channel on a readout chip. The precision that can be achieved depends on the strip pitch and on the readout technique, but can be as small as a few microns.

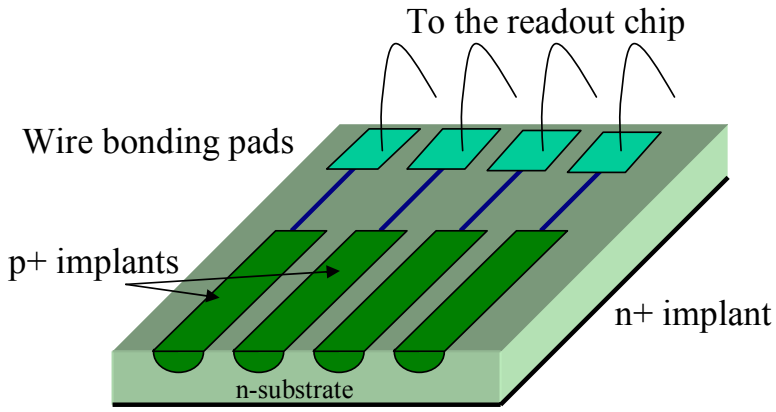

Figure 2.1: Schematic view of a single sided strip detector. The sensing diodes are realized implanting a set of parallel p^+ strips on high resistivity n-substrate.

This kind of layout can provide only one dimension tracking information; to overcome this problem several layers of angled microstrip detectors can be used, or a single substrate can be segmented on both sides with tilted strips (as shown in Figure 2.2).

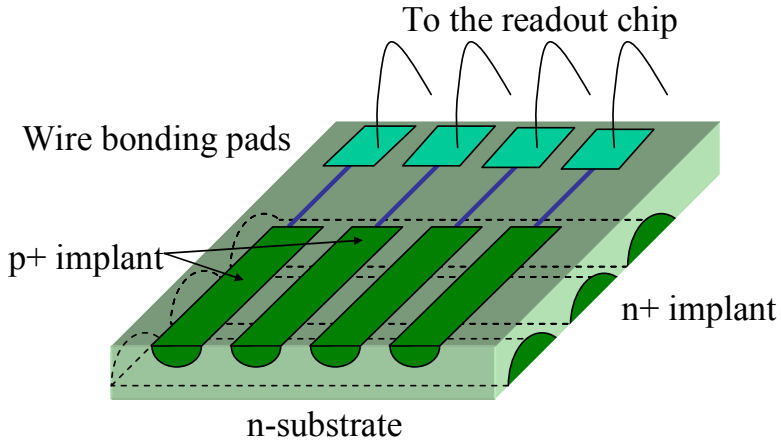

Figure 2.2: Schematic representation of a double sided microstrip detector.

2.1.2 Pad Detectors

One of the first solutions proposed for tracking with semiconductor elements was to use arrays of discrete silicon diodes connected to individual readout channels. Pad detectors are a development of this idea [Wei96], where the detection diodes are implemented on the same segmented thin substrate. The sensor is segmented in a two-dimensional array of pads (thus providing two-dimensional tracking information). Each single pad is individually routed via a metal connection on one surface of the silicon substrate to an array of wire bonding pads, as illustrated in Figure 2.3.

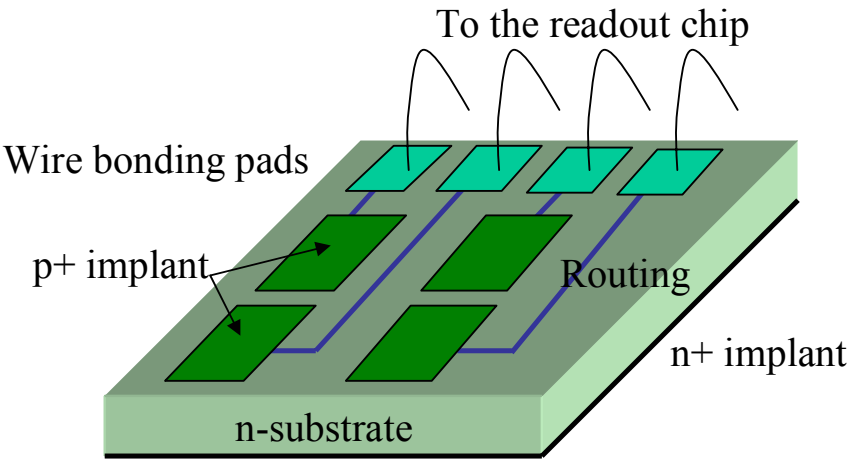

Figure 2.3: Schematic view of a pad detector, which is segmented in a two-dimensional array of pads.

This allows wire bonding of each diode with a dedicated front-end readout channel, hosted in a separate chip, outside the detecting area. In order to fully exploit the low capacitance of the sensor cells (< 1pF/pad) a readout chip with low input capacitance has to be used. Each detection element has a size of the order of 1 mm^2, so although this is an improvement compared with discrete diode arrays, its resolution may be too coarse for many tracking applications. The main advantages are the reduced constraints on layout and implementation of the front-end electronics.

2.1.3 Semiconductor drift detectors (SDD)

Silicon drift detectors were proposed by Gatti and Rehak [Gat84, Reh86], and are realized on a high resistivity and high purity silicon substrate. Figure 2.4 shows a schematic view of a silicon drift detector: on the two sides of a weakly n-doped wafer two sets of p^+ doped strips are implanted, with a pitch of about 150-200 μm. At the edge of the detector an n^+ strip is implanted, which has to act as a low capacitance collecting anode.

If a high enough potential is applied on both sides of the detector, the substrate can be fully depleted. In this case a potential minimum is created in the semiconductor. If an ionising particle traverses the biased device, the generated holes drift toward the p^+ electrodes, while

the electrons migrate towards the minimum of the potential well. By means of a proper biasing of the drift chamber, a transverse electric field (up to 1000 V/cm) is created, which pushes electrons towards the collecting anode, where they are read out. This type of detector can give a one-dimensional information about the hit position by measuring the drift time of the electrons to the anode, with an accuracy down to 5 μm. True two-dimensional information (with an accuracy of the order of the 10 μm) can be obtained segmenting the collecting anode.

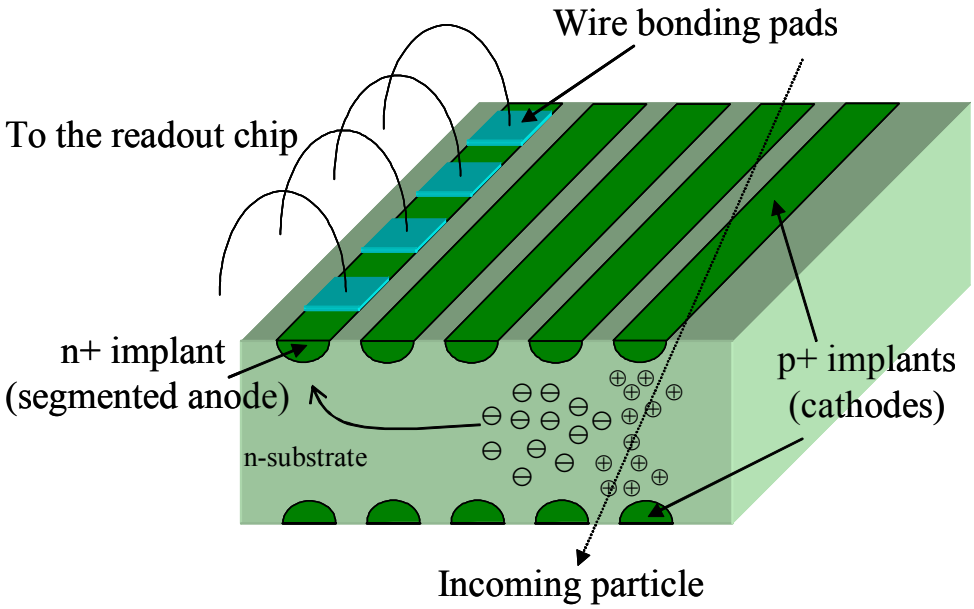

Figure 2.4: Schematic view of a linear semiconductor drift detector. Electrons (⊖) flow towards the anode, holes (⊕) are collected by the cathodes.

The very low anode capacitance results in very low electronic noise, and it is independent from active area. Moreover, if using a SDD for position measurements[2], although spatial resolutions obtained are similar to microstrip detectors, the number of readout channels is much lower. If used for energy measurements, SDDs can have wide active area[3] and high energy resolution.

The main limitations of this kind of detector are the long drift time of electrons (of the order of several microseconds), their diffusion while drifting[4] and a very high sensitivity to temperature and parameter variations (to keep a high resolution a temperature control within 0.1 °C can be required), so they cannot be used in high multiplicity or high rate environments.

[2] SDDs can be used both for position measurements (tracking) and for energy measurements. In the first case we are interested in the impact position of the impinging particle, and a segmented anode has to be used. In the second case we are interested in the total energy lost in the detector by the impinging particle, so all the charge has to be collected (and integrated) on a single anode.

[3] A cylindrical SDD of 55 cm^2 was designed to be used for forward tracking in the heavy ion experiments NA45 and WA98 at CERN SPS [Hol96].

[4] The charge cloud generated by the impinging particle widens while drifting towards the anode, worsening the spatial resolution.

SDD detectors are used in the two central layers of the ALICE ITS.

2.1.4 Charge Coupled Device (CCD) Detectors

A CCD is a silicon substrate on which a rectangular array of MOS capacitors is created (as schematically represented in Figure 2.5). To make the detector suitable for particle detection, a high resistivity substrate is used.

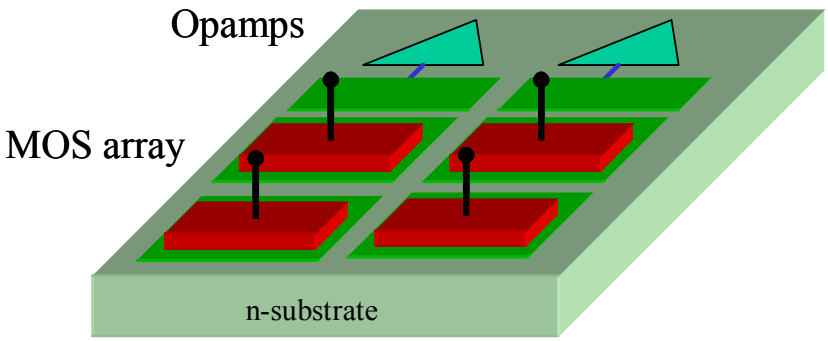

Figure 2.5: Schematic representation of a CCD detector. A rectangular matrix of MOS capacitors is created on a high resistivity silicon substrate.

A tiny potential well is created underneath each MOS capacitor, so that when an electron/hole pair is generated in the semiconductor due to an ionising particle crossing the detector, charge is stored in this potential well (which then acts as a charge integrator). At the end of the sensing phase, a readout phase is started, during which the charge underneath each capacitor is shifted from one MOS to the neighbouring one, and the last MOS in the row (or column) is read out by a charge amplifier. This serial readout scheme results in readout times of the order of the millisecond, even for readout frequencies of tens of MHz, thus restricting the possibility to use this kind of detectors to low rate experiments [Dam83]. Moreover, CCDs are very sensitive to radiation damage.

The size of the pixel, which can go down to 5 μm × 5 μm, determines the spatial resolution of the detector, which of course is able to provide true two-dimensional information on the hit position.

2.1.5 Pixel detectors

The term “pixel” has been introduced in image processing for the smallest discernable element in a given process or device. In CCDs, for example, “pixel” indicates the basic cell. The pixel (also called micropattern) particle detector can be defined then as the semiconductor device that consists of a matrix of contiguous particle sensing elements with dimensions between ~10 μm and ~500 μm (so much smaller than pad detector sensing elements) each of which are connected to an individual processing chain (which for example is composed, for

hybrid pixel detectors, by: amplifier, signal shaper, comparator, memory and logic) and a hierarchical readout architecture such that the user is provided with a pattern (true bi-dimensional information) of the particle hit position at each chosen timeframe, even at an interaction frequency in the MHz range [Hei89, Del92, Hei94].

The incorporated active signal processing at microscopic level and the digital information storage/treatment are peculiar of this kind of device (and is not present in pad detectors). The high granularity of the device results in a very high two-dimensional spatial and two-track resolution, and in low pixel capacitance (down to about 20 fF for hybrid detectors with small pixel size), which allows a very good signal to noise ratio[5].

As all the electronics has to sit in an area not bigger than the area of the sensor pixel, and has to be replicated thousands of times, several major constraints have to be fulfilled by the designer in terms of complexity of the design and power consumption. All these constraints are partially relaxed as technology improves, providing designers with smaller and faster transistors. This is why working pixel detectors, meeting the needs of high energy physics (especially for experiments with high multiplicities and high track densities) have become available in recent years.

The crucial point in realising pixel detectors is the integration of the pixel sensor with the corresponding electronic cell. Depending on the way this integration is done, four main types of pixel detectors can be described.

2.1.5.1 Monolithic pixel detectors

In monolithic pixel detectors the sensor and the corresponding electronics are integrated onto the same substrate. This results in a thinner (thus reducing the particle energy loss and scattering in a complex high energy particle detector) but more robust device, if compared with a hybrid approach where a silicon sensor and an electronics chip have to be connected with microscopic metallic balls (see section 2.1.5.4). Moreover, lower stray capacitance at the preamplifier input reduces electronic noise.

Different architectures for monolithic pixel detectors have been proposed [Sno92]. One of the solutions proposed is shown in Figure 2.6. In this approach the readout electronics is implanted in a doped well next to the signal collection electrodes in the sensor substrate. The doped wells shield the electronics and the sensor bulk against mutual influences. The main problem of this monolithic approach is that double-sided fabrication processes on special wafers are needed, which are not standard in microelectronics industry. Moreover, the performance of the sensor and of the electronics cannot be optimised separately.

[5] The ALICE1LHCb hybrid chip, which will be discussed in detail in Chapters 7 and 8, for example, has a noise of ~ 100 e^- for a typical signal of 5000 e^- (in the LHCb RICH, 17000 e^- in the ALICE SPD), which gives a S/N of about 50.

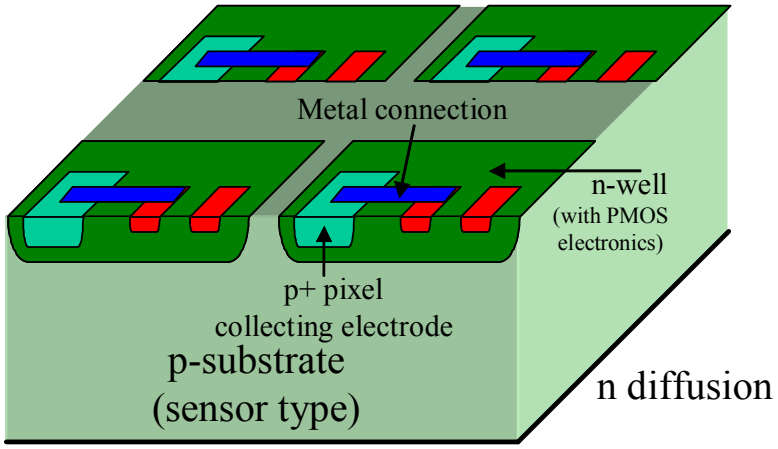

Figure 2.6: Schematic representation of a monolithic detector (bulk technology): the sensor and the corresponding electronics are integrated onto the same substrate.

Another, more recent, approach in monolithic detectors is to use a standard CMOS process both for the sensor and for the readout electronics [Tur01, Deb02, Gor02]. In the early 90's monolithic pixel sensors have been proposed as a viable alternative to CCDs in visible imaging. These sensors are made in a standard VLSI technology, usually CMOS, which is the reason why they are often called CMOS imagers. Two main types of sensors exist: the Passive Pixel Sensor (PPS) or the Active Pixel Sensor (APS). In the former, a photodiode is integrated in a pixel together with selection switches, which connect the photodiode directly to the output line for readout. In the latter, an amplifier integrated in each pixel directly buffers the charge signal. Today most CMOS imagers have an APS structure because of its better performance. The straightforward application of a CMOS sensor in a tracking detector is nevertheless not possible because of its poor fill factor[6]. Because of the transistors and metal interconnections, it is usually in the range of 20-30%, an unacceptable low value for particle tracking applications.

In order to overcome this limitation, it was proposed to integrate a sensor in a twin-tub process with an n-well/p-substrate diode [Tur01]. This technique has already proved its effectiveness in visible light applications reducing the blind area only to the routing metal lines, which absorb the visible light but have no effect on the detection of minimum ionising particles. It takes full advantage of the substrate structure of most modern VLSI technologies, which feature twin (p and n) tubs, implanted in lightly doped, p-epitaxial silicon, grown on a highly doped p^{++} substrate. The photodiode is made of the junction existing between the n-well and the p-type epitaxial layer. Because the epi-layer doping is a few orders of magnitude smaller than that of the p-well or the p^{++} substrate, potential barriers exist at its boundaries, which act like mirrors for the excess electrons (minority charge carriers).

[6] For visible light applications the fill factor is defined as the fraction of the pixel area that is sensitive to the radiation.

This architecture allows a 100% fill factor and small pixels of 20 μm × 20 μm, thus leading to a very low noise (down to ~12 e^- rms) due to the reduced input capacitance (estimated to about 10 fF) and a good signal to noise ratio (up to ~40), with practically 100% efficiency.

One of the main problems of this approach is its strong dependence on the technology chosen, and in particular on the quality and thickness of the epitaxial layer. Moreover, as charges created by the impinging particles rely not only on the drift mechanism but also on the diffusion mechanism to move towards the electrodes, the charge collection time can be much longer than in a fully depleted junction (in the order of 100-150 ns, depending on the thickness of the epitaxial layer).

2.1.5.2 Silicon on insulator pixel detectors

An alternative process to planar technology to fabricate integrated circuits is the Silicon On Insulator (SOI) process. In this technology transistors are grown on an insulating substrate, thus reducing the stray capacitances to the substrate and increasing circuitry speed if compared to planar technology (for a constant power budget). Another advantage of SOI is that it can also be made intrinsically radiation tolerant.

As a high resistivity layer is available underneath the SOI oxide, it was proposed [Pen96] to use this layer as the detecting volume (Figure 2.7).

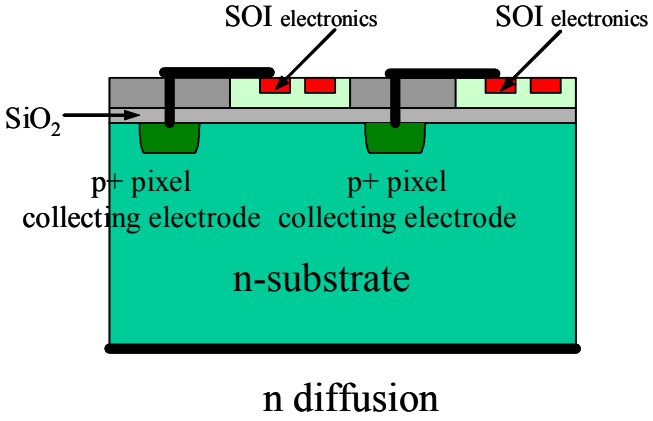

Figure 2.7: Schematic representation of a SOI monolithic pixel. In this technology transistors are grown on an insulating substrate.

The pixel implants on the sensor should then be connected with the electronics using vias through the silicon dioxide, which helps to reduce crosstalk between the detector and the electronics. This approach requires non-standard processing steps, which can result in a much higher price per chip, if compared with standard planar technology.

2.1.5.3 Hydrogenated Amorphous Silicon pixel detectors

A different approach, but that results in a scheme quite similar to a SOI detector, is the Hydrogenated Amorphous Silicon (a-Si:H) Pixel Detector. In this case the electronics is a standard CMOS process, where on top of each pixel a large metal pad (with an opening in the polyimide[7]) is placed. A layer of amorphous silicon, which acts as the detecting element, is then deposited on the chip.

In the very beginning this kind of detectors needed to be coupled with a converting scintillating material[8] (as CsI). An effort is now being made to use a-Si:H as a direct detector itself, trying to develop a low noise high gain amplifier, able to detect the low signal produced by the impinging particles in the detecting layer. More information can be found in [Jar03, Des03].

2.1.5.4 Hybrid pixel detectors

In a hybrid pixel device the readout chip and the sensor are processed separately, and then each readout channel is connected to its corresponding detecting element through a microscopic metallic ball, usually with a diameter of ~20 μm, as shown in Figure 2.8.

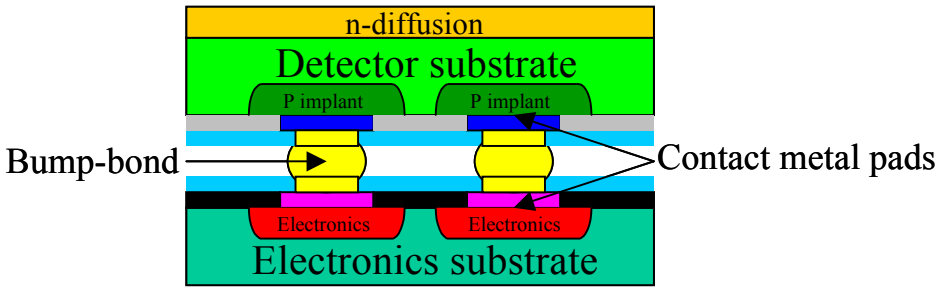

Figure 2.8: Schematic cross section of a hybrid pixel detector: the readout chip and the sensor are processed separately, and then each readout channel is connected to its corresponding detecting element through a microscopic metallic ball.

This process is called *flip-chip bump-bonding*, and a three-dimensional schematic view of the process is shown in Figure 2.9. Depending on the vendor, the bump-bonding can be realised in different ways, for example with indium or with lead-tin solder bumps.

The great advantage of this approach is that standard technologies can be used for both detector and electronics. Both of them can be optimised separately, as the best silicon substrates for particle detection and for high speed/high quality electronics are very different. Moreover, the detector material can be chosen for a specific application. This makes hybrid detectors very interesting for medical imaging, where different types of detectors (for example

[7] Polyimide is a protecting layer which can be deposited on top of the wafer in planar technologies.
[8] A scintillating material is a material which emits light in the visible spectrum when it is hit by an energetic particle.

GaAs, CdTe, which have higher absorption efficiency than silicon for X-rays with energy >20KeV) can be bump-bonded to the same front-end chip. An example of hybrid detectors used in the medical field is the MEDIPIX chip, a single photon counting system for X-ray radiography, realised by the Medipix collaboration[9] [Llo02, Mik00].

This type of sensor is used in the two inner layers of the ALICE ITS.

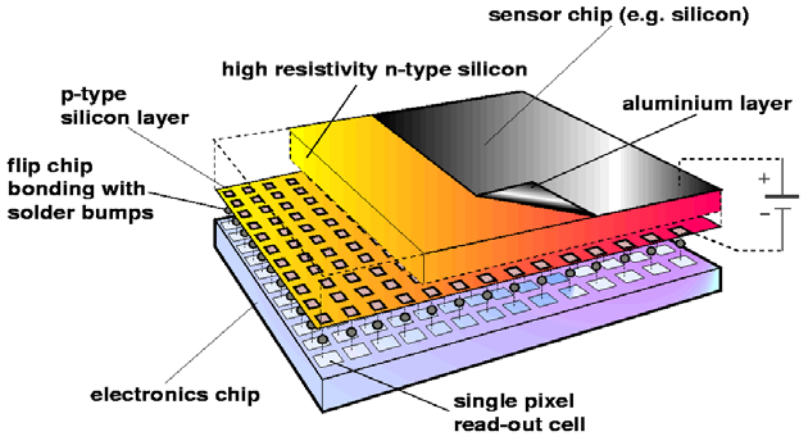

Figure 2.9: Schematic three-dimensional representation of the bump-bonding process.

2.2 Radiation damage in semiconductor particle sensors

Sensors can be damaged by radiation. There are three main consequences of the radiation damage created in the sensor substrate (bulk damage): increase of the leakage current, change of the effective doping concentration and decrease of the charge collection efficiency. Less important effects are related to the reduction of the inter-segment resistance and the increase of the inter-segment capacitance.

Due to the creation of energy levels within the sensor band-gap, which act as charge carrier generation centres, the bulk *leakage current per unit volume* (diode current density) increases. It is usually measured at full depletion voltage, and is defined as:

$$I_{Vol} = \frac{I}{A\ W} \qquad (2.1)$$

where I is the diode reverse leakage current, A is the diode area and W is the diode thickness. I_{Vol} is proportional to the particle fluence Φ[10], the proportionality factor being the current-related damage rate $\alpha(T,t)$, which depends on the temperature T and on the annealing time t:

[9] The Medipix collaboration was created by CERN, University of Friburg (Germany), University of Glasgow (Scotland), Universities and INFN sections of Pisa and Napoli (both Italy) to spin off ideas and technologies conceived at CERN for high energy physics into the medical imaging domain.

[10] The particle fluence Φ is the number of particles per unit area which cross the given area. It is often expressed as the 1 MeV equivalent neutron fluence Φ_{eq}.

$$I_{Vol} = I_{Vol0} + \alpha(t,T)\cdot\Phi \qquad (2.2)$$

where I_{Vol0} is the diode current density before irradiation (this value, of the order of some tens of nA/cm^3, can usually be neglected).

The diode reverse leakage current is strongly temperature dependent, so it is common practice to scale it to a reference temperature T_R (usually 20 °C) with the following formula [Bar93]:

$$I_{Vol}\ (T_R) = I_{Vol}(T)\left(\frac{T_R}{T}\right)^2 \exp\left(\frac{-E}{2\,k_b}\left(\frac{1}{T_R}-\frac{1}{T}\right)\right) \qquad (2.3)$$

where T is the absolute temperature during the measurement, T_R is 293.16 K, k_B is the Boltzmann constant, and E is related with the silicon bandgap, and can be assumed E = 1.23 eV (as suggested in [Bar93]).

According to the non-ionising energy loss (NIEL) mechanism, the dependence of the volumetric leakage current increase rate $\alpha = \alpha(X, E_X)$ on the radiation type X and energy E_X scales with the ratio:

$$\frac{\alpha(X,E_X)}{\alpha(Y,E_Y)} = \frac{K(X,E_X)}{K(Y,E_Y)} \qquad (2.4)$$

where $K(X,E_X)$ ($K(Y,E_Y)$) is the hardness factor of radiation type X (Y) at energy E_X (E_Y).

Irradiation can induce vacancies in the sensor material, which in turn may react with doping atoms and decrease the actual donor density N_D in an n-type semiconductor sensor substrate. At the same time irradiation increases the acceptors density N_A. The combination of the two effects ends up in a *reduction of the actual n-doping* of the n-type substrate, which eventually becomes p-type (*sensor type inversion*) at a certain fluence (*inversion fluence*). Even if a sensor reached the inversion fluence, it can still work because the junction moves from the p^+ segment implant to the n^+ back-plane connection. This has a direct effect on the sensor depletion voltage.

Damage induced defects in the sensor bulk material can act as trapping centres causing a *reduction of the charge collection efficiency*, i.e. a smaller signal is generated for the same impinging radiation. This effect is rather small if compared to the previous one, and can be compensated by increasing the detector bias voltage.

The type inversion of an n-type substrate detector generates a *reduction of the inter-segment resistance*[11], and this effect degrades the sensor resolution. Moreover, the trapping of charges in the passivation layer at the sensor surface can deform the electric field between the

[11] This is the resistance of the sensor in between the sensitive p^+ implantations.

active sensor segments, leading to an *increase of the inter-segment capacitance* (and thus of the capacitance of the overall detector).

Surface effects are also present, but can be neglected for p-implant on n-substrate particle sensors.

2.3 Summary

This chapter gave an overview of the most important particle detectors used for High Energy Physics. This is to better understand the different use done in the CERN experiments of the various types of detectors, and in particular in the ALICE ITS presented in Chapter 3.

Particularly interesting for our purposes are pixel detectors, because hybrid pixel detectors are used for tracking in the inner layer of the ALICE ITS. In a hybrid pixel device the readout chip and the sensor are processed separately, and then each readout channel is connected to its corresponding detecting element through a microscopic metallic ball (flip-chip bump-bonding). The design and testing of the detector and of the front-end chip of the ALICE hybrid sensor are presented in Chapters 4, 7 and 8.

A section is dedicated to bulk radiation damage in semiconductor particle sensors. The more common radiation induced damages in particle sensors are discussed, and the definitions of the most important physical quantities related to these phenomena are given (particle fluence Φ, leakage current increase rate α, hardness factor K). These quantities will be used while describing the results of the irradiation of the ALICE particle sensors, presented in Chapter 4.

Chapter 3
The ALICE Silicon Pixel Detector

Track finding in heavy-ion collisions at the LHC presents a big challenge, because of the extremely high track density [ALI99]. In order to achieve a high granularity and a good two-track separation ALICE uses three-dimensional hit information, wherever feasible, with many points on each track and a weak magnetic field. The ionization density of each track is measured for particle identification. The need for a large number of points on each track has led to the choice of a Time Projection Chamber (TPC) as the main tracking system. In spite of its drawbacks, concerning speed and data volume, only this device can provide reliable performance for a large volume at up to 8000 charged particles per unit of rapidity. The minimum possible inner radius of the TPC (of about 90 cm) is given by the maximum acceptable hit density. The outer radius (of about 250 cm) is determined by the minimum length required for a dE/dx resolution better than 10%. At smaller radii, and hence larger track densities, tracking is taken over by the Inner Tracking System (ITS).

In this chapter we will give an overview of the ALICE ITS, and then focus on its two innermost layers, which form the ALICE Silicon Pixel Detector (SPD). All the basic components of the SPD are presented, as well as the dedicated test system.

3.1 The ALICE Inner Tracking System

The system consists of six cylindrical layers of coordinate-sensitive detectors, covering the central rapidity region ($|\eta| \leq 0.9$) for vertices located within the length of the interaction diamond ($\pm 1\sigma$), i.e. 10.6 cm along the beam direction (z). The detectors and front-end electronics are held by lightweight carbon-fibre structures. A general view of the ITS is shown in Figure 3.1 (left).

The number and position of the layers are optimized for efficient track finding and impact parameter resolution. In particular, the outer radius is determined by the track matching with the TPC, and the inner one is the minimum compatible with the radius of the beam pipe (3 cm).

The silicon detectors feature the high granularity and excellent spatial precision required. Because of the high particle density, up to 90 per cm^2, the four innermost layers ($r < 24$ cm) must be truly two-dimensional devices. For this task Silicon Pixel Detectors (SPD) and Silicon Drift Detectors (SDD) were chosen.

The two innermost layers of the ITS are fundamental in determining the quality of the vertexing capability of ALICE (determination of the position of the primary vertex, measurement of the impact parameter of secondary tracks from the weak decays of strange, charm and beauty particles).

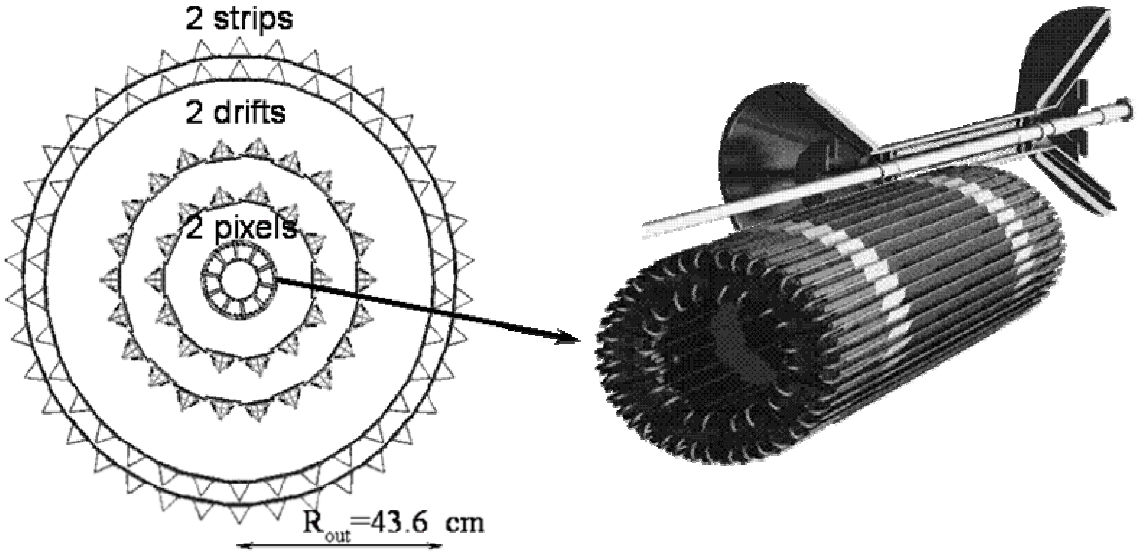

Figure 3.1: General view of the Alice Inner Tracking System (cross section, left). It consists of six cylindrical layers of silicon detectors. Three dimensional pictorial view of two inner layers of the ITS, the ALICE SPD (right) (INFN Padova).

Several motivations led to the choice of equipping ALICE with a barrel of two layers of **Silicon Pixel Detectors** (Figure 3.1, right). A silicon detector with a two-dimensional segmentation combines the advantages of unambiguous two-dimensional readout with the geometrical precision, double-hit resolution, speed, simplicity of calibration and ease of alignment characteristics of silicon microstrip detectors. In addition, a high segmentation leads naturally to a low individual diode capacitance, resulting in an excellent signal-to-noise ratio at high speed. The SPD will be described in more detail in the next section.

Silicon Drift Detectors have been selected to equip the two intermediate layers of the ITS, since they couple a very good multi-track capability with dE/dx information.

The outer two layers at a radius of about 45 cm, where the track densities are below 1 per cm^2, will be equipped with double-sided **Silicon Strip Detectors** (SSD) with a small stereo angle. Double-sided microstrips have been selected rather than single-sided ones because they introduce less material in the active volume. In addition they offer the possibility to correlate the pulse height read out from the two sides, thus helping to resolve ambiguities inherent in the use of detectors with projective readout.

With the exception of the two innermost pixel planes, all the ITS layers will have analogue readout for particle identification via a dE/dx measurement in the non-relativistic

region. This will give the inner tracking system a stand-alone capability as a low-p_t particle spectrometer.

The large number of channels in the layers of the ITS requires a large number of connections from the front-end electronics to the detector and to the readout. The requirement for a minimum of material within the acceptance does not allow the use of conventional copper cables near the active surfaces of the detection system. Therefore Tape Automated Bonding (TAB) aluminium multilayer microcables are used.

The outer four layers of the ITS detectors are assembled onto a mechanical structure made of two end-cap cones connected by a cylinder placed between the SSD and the SDD layers. Both the cones and the cylinder are made of lightweight sandwiches of carbon-fibre plies and Rohacell[1]. The carbon-fibre structure includes also the appropriate mechanical links to the TPC and to the SPD layers.

The latter are assembled in two half-cylinder structures, specifically designed for safe installation around the beam pipe. The end-cap cones provide the cabling and cooling connection of the six ITS layers with the outside services.

3.2 The Silicon Pixel Detector (SPD)

The ALICE SPD [ALI99] is composed of two concentric barrel layers of pixel detectors, with radii of 3.9 and 7.6 cm. Figure 3.1 (right) shows a three dimensional pictorial view of the ALICE SPD, while Figure 3.2 shows a cross section of it [Din02, Cho03].

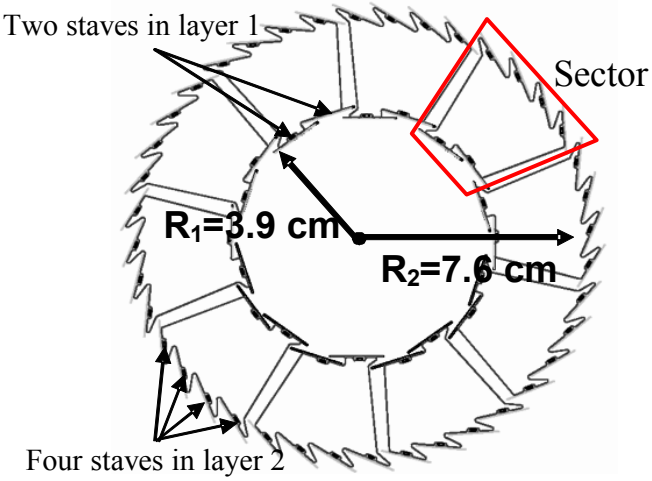

Figure 3.2: Cross section of the SPD detector, showing the ten sectors assembled together to form the two pixel detector layers with radii 3.9 cm and 7.6 cm.

[1] ROHACELL® is a polymethacrylimide- (PMI-) hard foam, that is used as a core material for sandwich constructions. It shows outstanding mechanical and thermal properties. In comparison to all other foams it offers the best ratio of weight and mechanical properties as well as highest heat resistance. It is produced by Röhm GmbH & Co. KG [Web09].

The basic building block of the ALICE SPD system is the ladder, a high resistivity p^+ on n silicon detector of around 13 × 71 mm^2 that is bump-bonded to five ALICE1LHCb pixel readout chips. Two ladders are connected together and then to a Multi Chip Module (MCM, described in section 3.2.3.1) by means of a special Kapton-Aluminium Pixel Bus (described in section 3.2.4) to form a half stave, shown in Figure 3.3.

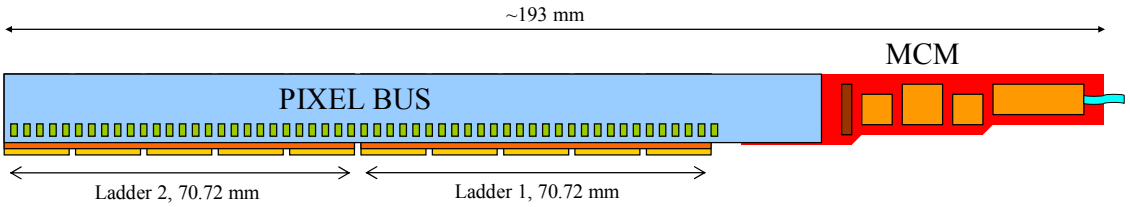

Figure 3.3: Two ladders connected to an MCM by means of a Pixel Bus forming a half stave [Mor03].

The MCM hosts three different ASICs: one for the biasing of the ALICE1LHCb chip and for temperature and DC bias monitoring (the Analog Pilot Chip, APC); one for digital read-out (the Digital Pilot Chip, DPC [Klu01]) and one for data transmission (The Gigabit Optical Serializer, GOL [Mor00]). All the ASICs are designed in a commercial CMOS process but employing special layout techniques (HBD) which improve their radiation tolerance, as will be explained in Chapter 5.

Two half staves form a stave, and six staves are mounted on a carbon fibre support to make a sector, with two inner staves and four outer staves. Ten sectors assembled together, as depicted in Figure 3.2, form the SPD detector. Figure 3.4 (left) shows a cross section of a sector, and in Figure 3.4 (right) shows a picture of the carbon fibre support.

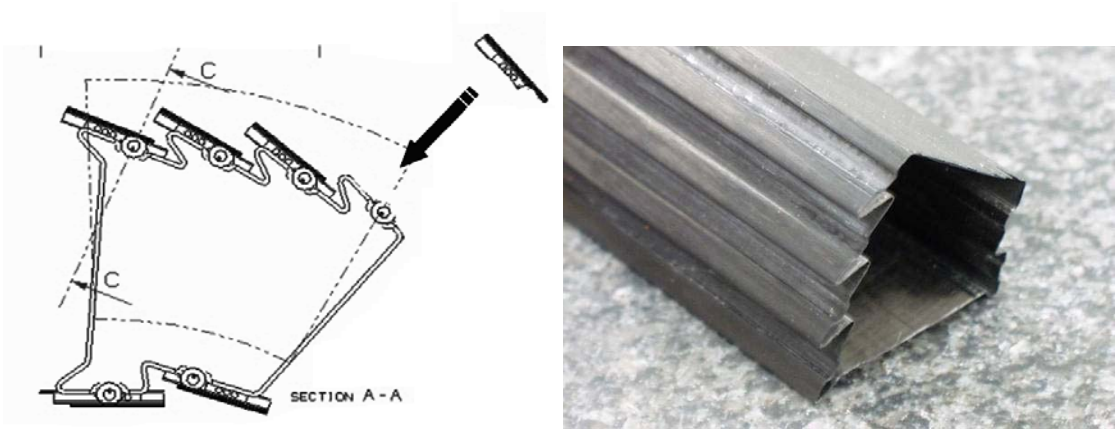

Figure 3.4: Cross section of a sector of the ALICE SPD (left). Four staves are mounted on the outer carbon fibre support (right) and two on the inner side (INFN Padova).

Once all the components are mounted on the carbon fibre support, the sector will look like in Figure 3.5. The SPD will be equipped with 120 half-staves (1200 Pixel Chips), which means a total number of about 10 million active readout channels.

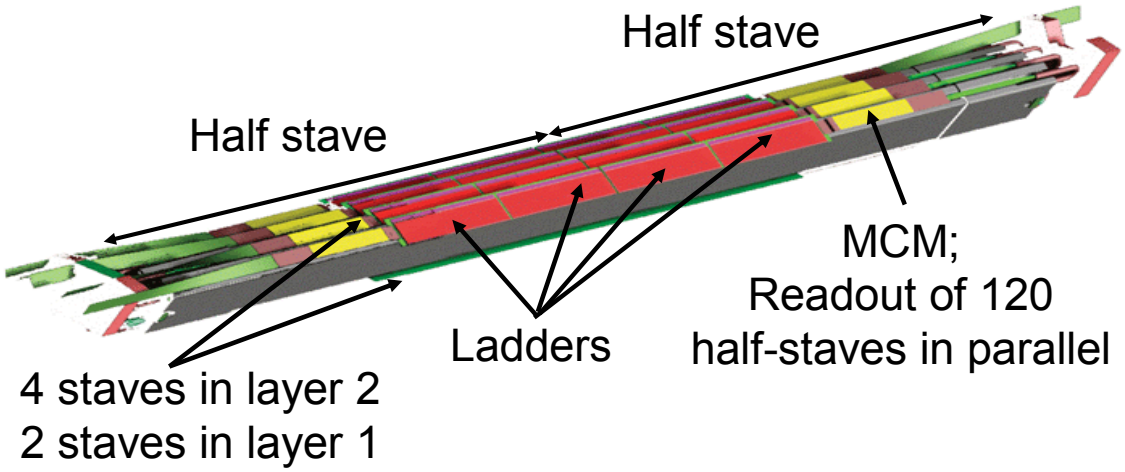

Figure 3.5: Drawing of a complete sector of the ALICE SPD (INFN Padova).

A problem which impacted directly on the ALICE1LHCb chip design was the material budget. The momentum and impact parameter resolution for particles with small transverse momenta are dominated by multiple scattering effects in any existing tracking detector. Therefore the amount of material in the active volume has to be reduced as much as possible. The thickness of the silicon detectors used to measure ionization densities was foreseen to be approximately 300 μm in the first version of the ITS TDR [ALI99] but is fixed at present at 200 μm to decrease the contribution due to silicon to the total amount of material.

The total material budget for the 2 layers of the SPD is ~2% X_o (radiation length) , where 0.36% is due to silicon, 0.3% to the cooling system, 0.1% to the support and 0.17% to the interconnection aluminium-kapton bus (the bus is described in more detail in section 3.2.4).

The resulting relative momentum resolution is better than 2% for pion momenta between 100 MeV/c and 3 GeV/c.

3.2.1 The front-end chip

The design and testing of the ALICE1LHCb chip, briefly summarised here, are presented in detail in Chapters 7 and 8.

The pixel detector readout chip (the ALICE1LHCb Pixel Chip) was developed to serve two experiments, ALICE and LHCb [Din03, Din02, Din01, Din00, Sno01, Sno01a, Wyl99]. The chip contains a matrix of 32 columns × 256 rows (8192 readout cells), measuring 13.5 × 15.8 mm^2. The analogue front-end of the pixel cell is composed of four main blocks: a charge preamplifier, a first shaping stage, a current feedback stage and a second shaping stage, which feeds a low time-walk discriminator. The first three stages constitute a system with two complex conjugate poles and a real pole. This solution was chosen instead of a standard charge integration and pole-zero cancellation scheme to achieve low noise and fast

return to zero of the amplifier, avoiding pile-up effects in high rate environments. As a consequence the next hit on the same pixel can be processed after less than 200 ns. A low-frequency feedback is also present, to compensate for detector leakage current and correct for offset at the second shaping stage output.

The discriminator output is connected to logic blocks which store a time stamp for the duration of a trigger latency. When a stored event coincides with a trigger the event, it is stored in a FIFO. Up to 4 triggered events can be stored before readout is initiated. The cell also contains 3 bits to finely adjust the threshold on a pixel-by-pixel basis. An on-chip pulser allows electrical tests to be performed on the chip without a detector being attached. Measurements done with this pulser indicate a minimum operating threshold of ~1000 e^-, with a dispersion of less than 200 e^- rms (unadjusted), and a noise below 120 e^- rms.

The radiation tolerance of the pixel chip has been extensively evaluated. The effects of Total Ionising Dose (TID) were assessed by exposing the chip to 10 keV X-rays at a rate of 6 kGy/hour. As the beam spot was smaller than the chip two different positions of the chip were used, for a TID of 120 kGy in some zones of the chip, and 240 kGy in some others. After the irradiation the minimum threshold at which the chip can operate is unchanged, as well as the pixel noise. The power consumption of the chip is unaffected by the irradiation.

The cross section for a Single Event Upsets (SEU) in the memory cells of the chip was measured in two different ways, with heavy ions with a Linear Energy Transfer between 6 and 120 $MeVmg^{-1}cm^2$ and with a 60 MeV Proton beam and a total fluence of 6.4×10^{12} cm^{-2}. In the first case we extrapolated a SEU cross-section of 9×10^{-16} cm^2 for protons with energy of 60 MeV, in the second case we measured a cross section of 3×10^{-16} cm^2. No Single Event Gate Ruptures or Single Event Latch-ups were observed.

3.2.2 The assemblies and the ladders

In a hybrid pixel device the readout chip and the sensor are processed separately, and then each readout channel is connected to its corresponding detecting element through a microscopic metallic ball. This process, which is a critical step for hybrid detectors, is called *bump-bonding* (see section 2.1.5.4). The bumps have to provide electrical contact and mechanical stiffness, and have to withstand a certain amount of thermal cycling. Two different vendors, each with full 8" wafer processing capability are being assessed. One of the vendors provides led-tin solder bump-bonding and the other uses indium. The quality of the bump-bonding process is in continuous improvement; ladders with chips with more than 99.7% working pixels on each chip are available, but more effort has to be made to meet the ALICE yield requirements of having > 99% working bump-bonded pixels.

Details on the SPD particle sensors, assemblies and ladders, are presented in Chapter 4 and Chapter 8.

3.2.3 The On-detector read-out pixel Pilot System (OPS)

The 10 pixel chips of one half stave are controlled and read out by one PILOT multi chip module (MCM) [Klu01]. Figure 3.6 shows a block diagram of the entire system. The PILOT MCM, described in the next section, transfers the data to the control room. In the control room 20 9U-VME-based router cards, two for each detector sector, receive the data. One router card contains three 2-channel link receiver daughter boards, a channel for each half stave in the sector. The link receivers process the data and store the information in an event memory. The router merges the hit data from 6 half staves into one data block, processes the data and stores them into a memory where the data wait to be transferred to the ALICE data acquisition (DAQ) over the detector data link DDL [Rub98].

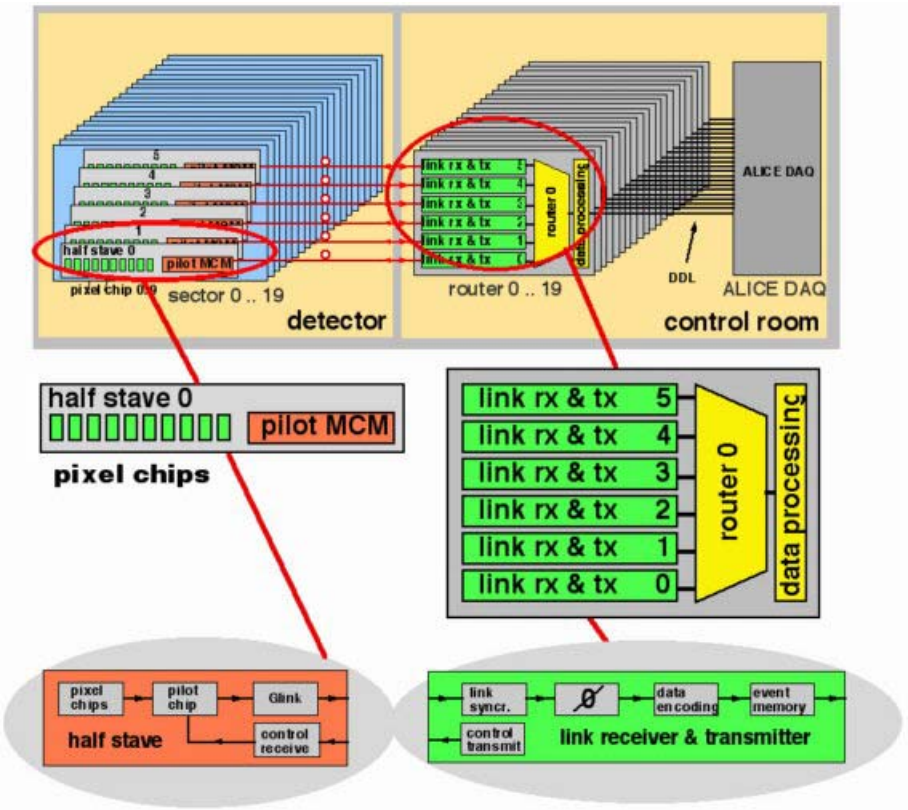

Figure 3.6: System block diagram of the ALICE OPS [Klu01].

The link from the control room to the MCM is done via two fibres, one for the data and the other for the clock. The use of this second fibre avoids the need for clock reconstruction on the MCM. On the data fibre the control room sends both configuration and test signals for all of the chips (mainly JTAG [IEEJt]) and trigger information for the Pixel Chips.

3.2.3.1 The PILOT Multi Chip Module (MCM) and the Digital Pilot Chip (DPC)

A block diagram of the complete SPD readout scheme is presented in Figure 3.7. The block on the left is the aluminium-kapton interconnection Pixel Bus with the two half staves

(10 pixel chips) and the temperature sensors added to monitor online the temperature variations. The centre block is the PILOT MCM. The third block shows a schematic view of the Control Room section.

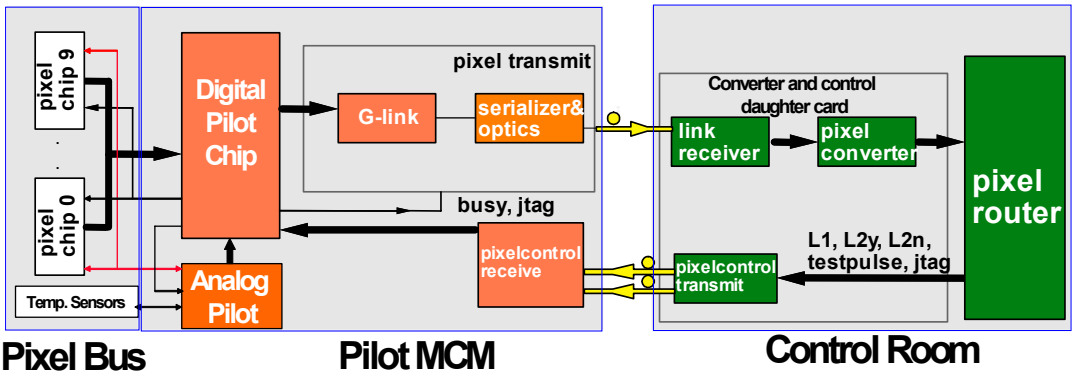

Figure 3.7: Block diagram of the complete ALICE SPD readout scheme.

The core of the PILOT MCM (and so of the OPS) is the Digital Pilot Chip (DPC, Figure 3.8, left), as it controls via a JTAG interface all of the chips on the half stave, and the data readout from the Pixel Chips [Klu01].

When the ALICE data acquisition issues a Level 1 trigger signal, the pixel router forwards the signal to the DPC, which asserts a strobe signal to all Pixel Chips. They store the delayed hit information into a multievent buffer, waiting for Level 2 (L2) trigger. If the pixel router sends an L2n, meaning that the events stored in the buffer are not interesting, hit data are discarded from the Pixel Chips. If an L2y is asserted, meaning that the events stored in the buffer are interesting and have to be read out, the DPC starts the readout procedure of the 10 Pixel Chips one after the other. The 256 rows of 32 pixels of a pixel chip are presented sequentially on a 32-bit bus. The read-out clock frequency is 10 MHz. As a result, the read-out of 10 chips takes about 256 μs.

The Pixel Chip data are then sent to the Gigabit Optical Link (GOL, see section 3.2.3.2; Figure 3.8, right) for encoding, serialization and transmission to the control room.

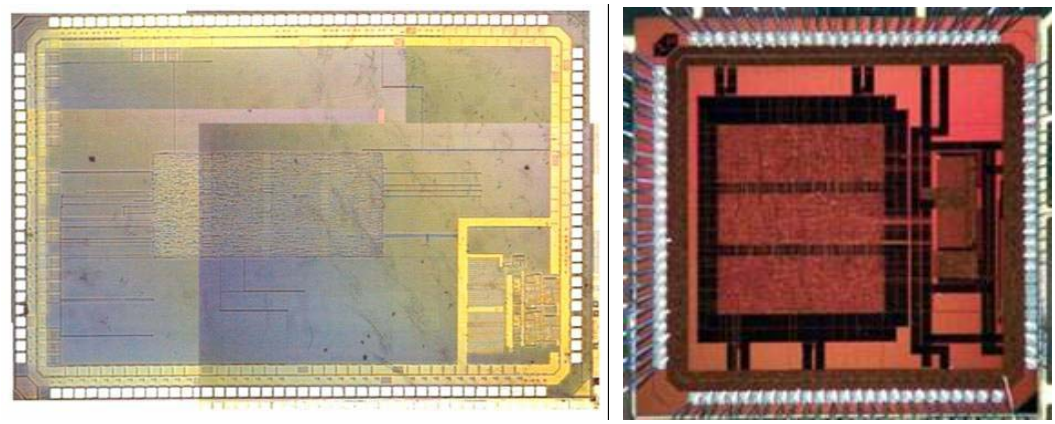

Figure 3.8: Photograph of the Digital Pilot Chip (left) and of the Gigabit Optical Link (right) chips.

The PILOT logic performs no zero-suppression and no data processing but directly transmits data to the control room. This approach has several advantages. The first is that the on detector PILOT-ASIC architecture is simple. Secondly, the system becomes more reliable as the complex data processing units are accessible during operation in the control room. Finally, if the detector hit occupancy increases in the future, data compression schemes can be adapted in the FPGA based control room located electronics.

The pixel data stream arrives from the pixel chips at the PILOT chip on a 32-bit bus in 100 ns cycles. That means that the transfer bandwidth of the GOL is twice as high as required. The 100 ns pixel data cycle is split up into four 25 ns GOL transmission cycles. In two consecutive GOL cycles, 16 bits of pixel data are transmitted. The remaining two transmission cycles are used to transmit data control (for example error codes or event numbers) and signal feedback (all trigger and configuration data sent from the control room to the detector are sent back to the router for error detection).

A dedicated board was designed and produced to test the DPC (Figure 3.9). The board was tested and is fully functional.

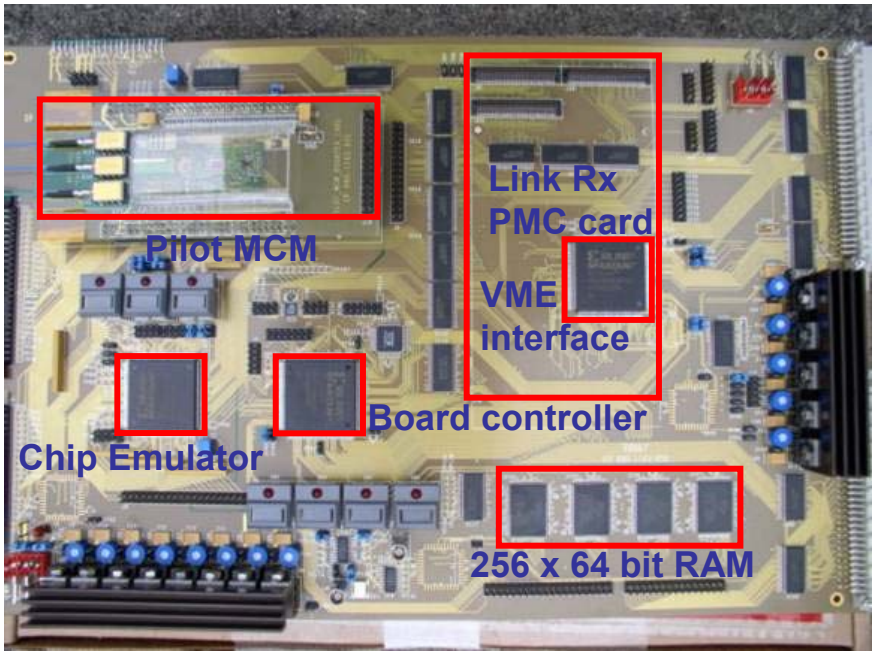

Figure 3.9: Picture of the Readout Test System Board (RUDOLF board).

The core of the test system is an FPGA that emulates the control room, all the other components of the half stave (except the APC) are hosted on the test board. The FPGA can send/read control/hit data, and check data integrity at the end of the full test chain. Tests performed on the complete chain show that the system is fully functional.

3.2.3.2 The Gigabit Optical Link (GOL)

The connection between the MCM and the control room is done via a fast optical link. Hit and control data are generated by the DPC chip, which sends them to the GOL chip (Figure 3.8, right). The GOL allows the transmission of 16 bit data words every 25 ns into a G-link compatible [Wal92] 800 Mbit/s stream of data on an optical link to the control room. This chip was already developed and tested at CERN, and is fully functional [Mor00].

3.2.3.3 Link receiver daughter card

The serial-parallel converter receives the G-link data stream and recovers the 40 MHz transmission clock using a commercial component, the AGILENT HDMP-1034 [Web08]. The implementation of the link receiver is based on a commercial FPGA and storage devices. Figure 3.10 shows a block diagram of the link receiver data converter [Klu01].

Figure 3.10: Block diagram representation of the Link Receiver Data Converter [Klu01].

The received data are checked for format errors and zero suppression is performed before the data are loaded into a FIFO. The expected occupancy of the detector will not exceed 2% [ALI99]. As a result, it is economic to encode the raw data format.

In the raw data format the position of a hit within a pixel row is given by the position of a logic '1' within a 32-bit word. The encoder transforms the hit position into a 5-bit word giving the position as a binary number for each single hit and attaches chip and row number to the data entry [Gra99]. The output data from the FIFO are encoded and stored in an event memory in a data format complying with the ALICE DAQ format [Klu00]. There it waits until merged with the data from the remaining five staves by the router electronics.

3.2.3.4 Pixel control transmitter and receiver

The pixel control transmitter and receiver are responsible for the transmission of the trigger and configuration signals from the control room to the detector [Klu01]. This includes the following signals: L1, L2y, L2n trigger signals, reset signals, a test pulse signal and JTAG signals. The data have to arrive at the detector in a 10 MHz binning, since the on detector PILOT system clock frequency is 10 MHz. The link is unidirectional since the return path for the JTAG system (TDO) uses the G-link data link. The data protocol has to be simple in order to avoid complex recovery circuitry on the detector in the PILOT chip. The data transmission is performed using two optical fibres, one carrying the 40 MHz clock and the other the actual data. The pixel control transmitter translates the commands into a serial bit stream. A priority encoder selects the transmitted signal in case two commands are active at the same time.

3.2.3.5 The Analogue Pilot Chip (APC)

The Pixel Chip requires six external bias voltages; two are for reference for the GTL input/output logic, two provide reference voltages to the DACs and two are used for electrical testing of the chip. These biases are provided, for all the ten chips mounted on the same half stave, by the Analogue Pilot Chip (APC). The APC is a mixed-mode IC containing the following blocks:

- Six 8-bit DACs, providing reference voltages to the ALIC1LHCb Chip. The design is a modified version of the DACs used for the ALIC1LHCb Chip itself. However, the reference voltages for these DACs are derived from an on-chip band-gap reference circuit [Web07, Kui73].
- Four current sources designed to be independent of temperature and power supply variations. These provide currents which can be used by PT1000 devices[2] for monitoring of the temperature of the stave.
- A 16-input analogue multiplexer followed by a 10-bit ADC [Riv01]. These inputs are used to enable in-situ scanning and monitoring of the Pixel Chip and APC DACs. They can also be connected to the PT1000 outputs for temperature monitoring.
- A JTAG-controlled digital block, providing all the necessary control signals to the other blocks.

More details on the APC will be presented in Chapter 9.

3.2.3.6 The Multi Chip Module implementation

The PILOT MCM is implemented on a special printed circuit board (PCB). The MCM PCB has five metal layers, two dedicated to power supplies and three for signal routing. It hosts three different chips (the APC, DPC and GOL), various passive Surface Mount Devices (SMDs) for biasing and decoupling, and one optical package containing two pin diodes and a laser diode. The chips can be directly glued and bonded onto the MCM without a package.

Due to mechanical constraints, the final version of the MCM must not exceed 100 mm in length and 14 mm in width[3]. Moreover, it has to be less than 2 mm thick. This is why a special optical package has been designed, which is less $1.2 \times 17 \times 5.5$ mm^3 and houses two pin diodes and a laser diode.

The complete MCM module (shown in Figure 3.11) was tested in the lab and during a test beam and is fully functional. In particular, the jitter after the GOL (at 800 MHz) was measured to be 10 ps, and the jitter after the DPC to be 23 ps.

[2] The PT1000 are special platinum resistors, whose value changes (in first approximation) linearly with temperature.

[3] The final MCM will have a "stair" shape with different widths; the largest part of the MCM will be in the order of 14 mm, the smallest around 8 mm.

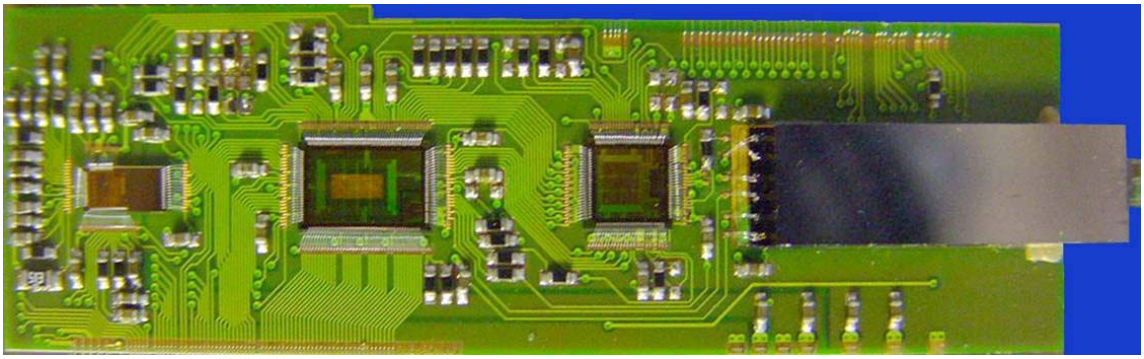

Figure 3.11: Photograph of a populated prototype of the PILOT MCM [Rie03].

3.2.4 The Pixel Bus

The two ladders and the MCM of a half stave are connected together by means of the pixel bus, which provides data, control and power lines between the MCM and the chips. A special aluminium-kapton bus has been designed, which is a complex technical development in the EST-DEM workshop at CERN. It uses aluminium metal layers to reduce the material budget. Figure 3.12 shows a schematic cross section representation of the Pixel Bus.

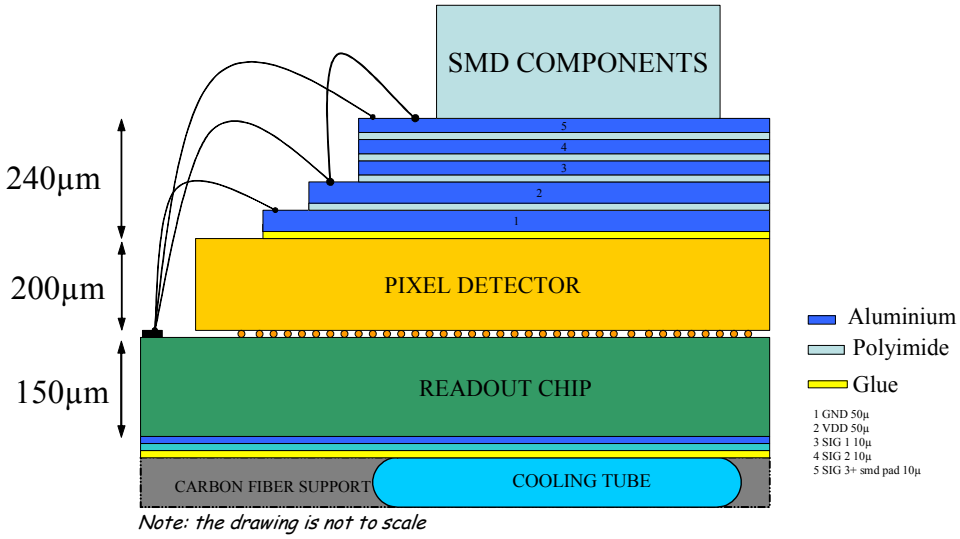

Figure 3.12: Schematic cross section representation of the Pixel Bus [Mor03].

The flexible bus has a total of five layers of aluminium[4] (two dedicated to power supplies and three for signal routing) and a total thickness of about 240 µm. Figure 3.13 (left) shows a photograph (top view) of the bus prototype already populated with SMD

[4] The first prototype of the pixel bus had 7 layers of metal.

components. The bus is glued on top of the detector, and the electrical connection with the Pixel Chips is done with wire bonds. There will be ~1100 wire bonds per half-stave, done with a 25 μm diameter wire. The bonding pads on the bus have an area of 80 × 300 μm^2. A picture of the wire bonding from the ALICE pixel chips to a prototype bus is shown in Figure 3.13 (right). SMDs will be placed on the bus for decoupling of power supplies. The "staircase structure" of the five layers of metal at the border of the bus (visible in Figure 3.13) has been foreseen for ease of wire bonding.

A prototype version of the pixel bus with 10 chips mounted on it has been successfully tested in the laboratory. Moreover, a ladder was mounted on a prototype bus and this was used in a test beam in 2002, and a fully equipped prototype half stave was tested in a test beam in 2003 (with both it was possible to track particles, see Chapter 8).

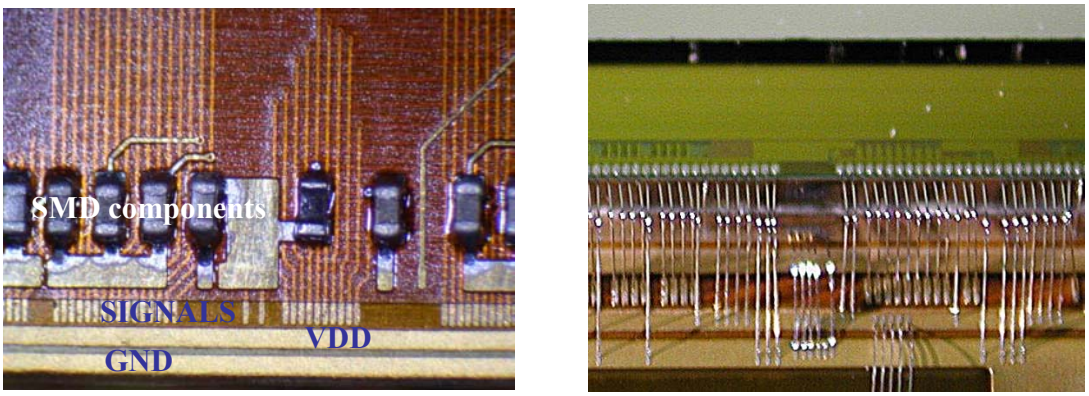

Figure 3.13: Photograph of a top view of the bus prototype (left); wire bonding from the ALICE pixel chips to a prototype bus (right) [Mor03].

3.2.5 The cooling system

The detectors and their front-end electronics produce a large amount of heat, in the order of 1.5 kW, which has to be removed while keeping a very high degree of temperature stability. In particular, the SDDs are sensitive to temperature variations in the 0.1°C range. For these reasons, particular care was taken in the design of the cooling system and of the temperature monitoring. A C_4F_{10} (evaporative) cooling system at room temperature (25 °C) is the solution chosen for all ITS layers. For temperature monitoring dedicated integrated circuits are mounted on the readout boards and specific calibration devices are integrated in the SDDs. A photograph of a cooling test measurement done using a resistive network and infrared camera is shown in Figure 3.14.

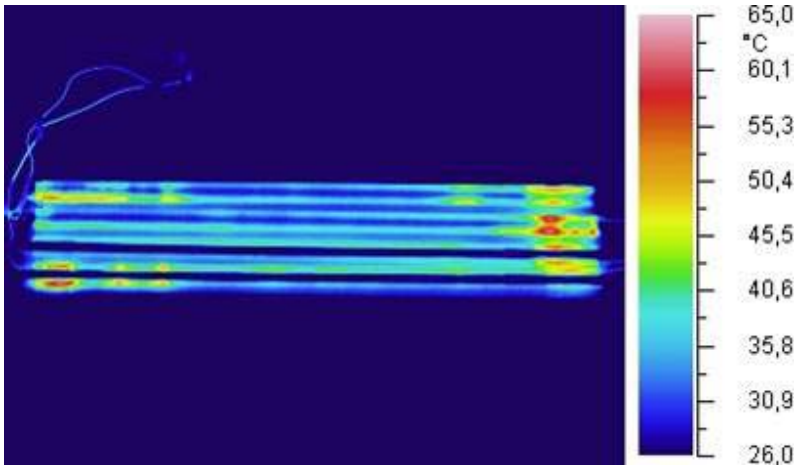

Figure 3.14: Photograph of a cooling test measurement done with an infrared camera, using a resistive network to simulate the chip thermal dissipation.

3.2.6 The Chip Test System

The test system [Bur01] has been designed around a PC, hosting the testing and control software, connected with a MXI connection to a VME crate. A Readout Controller (PILOT VME Module) has been developed in the VME standard to control the readout of the Pixel Chip. Figure 3.15 shows the layout of the basic test system. The main characteristic of the Pilot Test System (PTS) is its flexibility, as it has been conceived and realized to be used in several different testing scenarios (chip tests, bus and ladder tests, wafer probing and test beam). Both the hardware and the software share this characteristic.

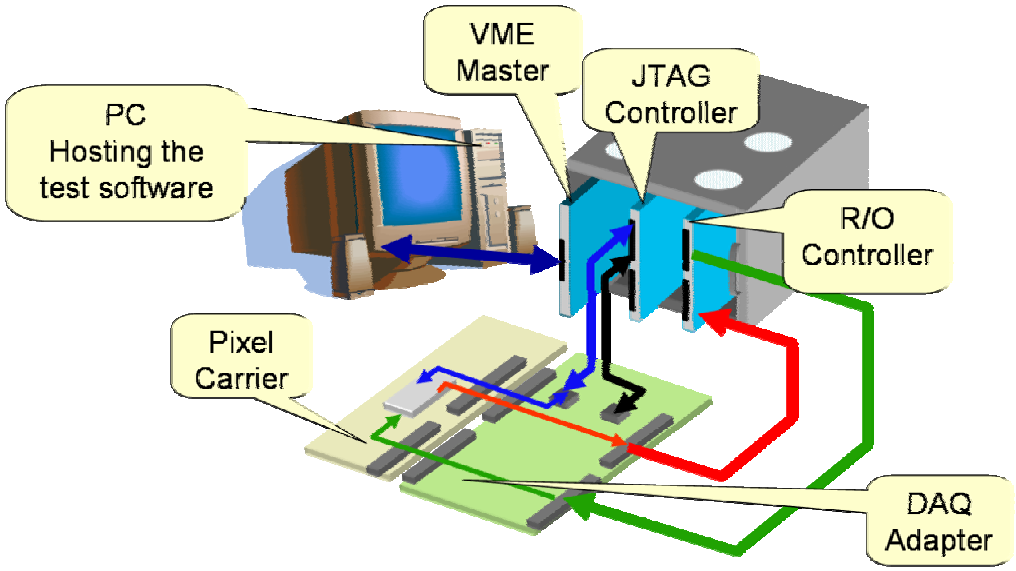

Figure 3.15: Layout of the basic test system. The system is modular, and can be adapted to different testing scenarios (chip tests, bus and ladder tests, wafer probing and test beam) [Bur01].

As the Pixel Chip is configured and controlled by JTAG, JTAG is also used to control a DAQ Adapter board that is situated close to the Pixel Chip under test. Both of them are addressed via a two-channels JTAG controller.

The DAQ Adapter Board (Figure 3.16) is situated between the Pilot VME Module and Pixel Chip under test and serves as an interface between the two environments. It houses the line drivers and receivers necessary for the DAQ connection, and the Gunning Transceiver Logic (GTL) drivers necessary for the Pixel Chip bus connection. It also houses the circuitry to derive the necessary power and bias supplies. Monitoring of both applied voltage and consumed current are possible.

The individual Pixel Chips are wire bonded to the carrier board (Figure 3.16), which may be connected to either the IC tester, or the DAQ Adapter board. Test points have been included on all bus signals. Facilities have been provided to allow observation of various internal nodes of the device. Different versions of the carrier board were designed for both wafer probing tests and evaluation of the prototype bus structure currently under design for use in the Pixel Detector.

Differential connections between the modules installed in the VME crate and the DAQ Adapter board allow the use of long interconnecting cables making the system suitable for use where the readout/test system has to be sited far from the Pixel Chip or chips.

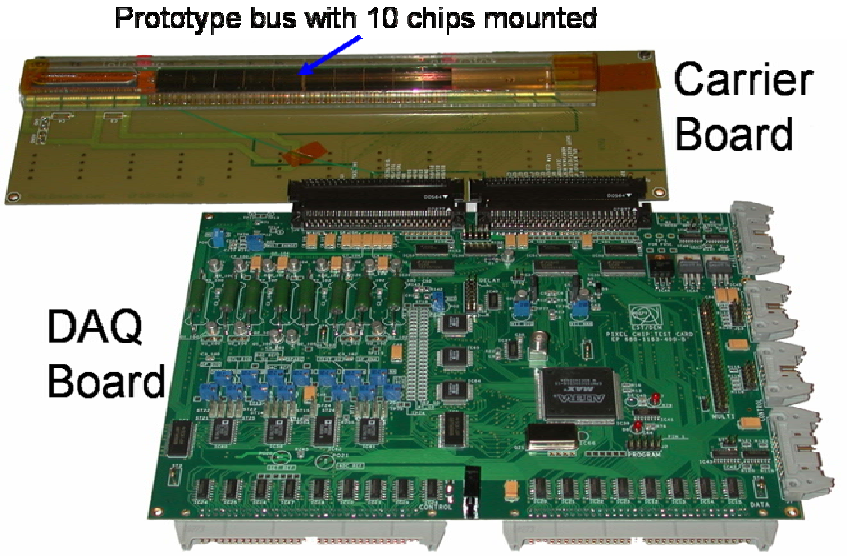

Figure 3.16: Photograph of the DAQ adapter board connected with the Carrier Board (the carrier board shown here is the version used to test the prototype bus).

The test software architecture (based mainly on Windows and LabView) reflects the flexibility of the hardware. Its modularity and architectural structure guaranties that the system can be used with different hardware configurations without the need of rewriting the software core. The test beam DAQ and monitoring program is the most powerful part of the

PTS. A single application enables the acquisition of data in a variety of conditions, and some plug-ins are loaded on demand to perform specialised tasks (e.g. checking of trigger efficiencies). The test system proved to be extremely robust and flexible, and was used with great advantage in a number of tests: single chip electrical tests, single chip assembly tests with radioactive source, ladder testing, wafer probing, pixel bus studies and test beams.

3.3 Summary

The chapter begins with an overview of the ALICE ITS, and then concentrates on its two innermost layers, which form the ALICE Silicon Pixel Detector (SPD). The ALICE ITS consists of six cylindrical layers of coordinate-sensitive detectors optimized for efficient track finding and impact parameter resolution. The two outer layers are equipped with Silicon Drift Detectors, the two intermediate layers with Silicon Strip Detectors and the two innermost layers with SPDs.

The ALICE SPD [ALI99] is composed of two concentric barrel layers of pixel detectors, with radii of 3.9 and 7.6 cm. The basic building block of the ALICE SPD system is the ladder, a high resistivity p^+ on n silicon detector of around 13×71 mm^2 that is bump-bonded to five ALICE1LHCb pixel readout chips. A section is dedicated to a summary of the characteristics of these readout chips , which will be described in much more detail in Chapters 7 and 8.

Two ladders are connected together and then to a Multi Chip Module by means of a special Kapton-Aluminium Pixel Bus to form a half stave. The bus provides data, control and power lines between the MCM and the chips.

Two half staves form a stave, and six staves are mounted on a carbon fibre support to make a sector. Ten sectors assembled together form the SPD detector.

The 10 pixel chips of one half stave are controlled and read out by one PILOT multi chip module (MCM). The MCM hosts three different ASICs: one for the biasing of the ALICE1LHCb chip and for temperature and DC bias monitoring (the Analogue Pilot Chip, APC); one for digital read-out (the Digital Pilot Chip) and one for data transmission (the Gigabit Optical Serializer). All the ASICs hosted by the MCM were described, as well as a dedicated test card which is also able to emulate the behaviour of the control room-based software and hardware.

The detectors and their front-end electronics produce a large amount of heat which has to be removed while keeping a very high degree of temperature stability. For these reasons, particular care was taken in the design of the cooling system and of the temperature monitoring.

A very flexible test system was specially developed to be used in several different testing scenarios, and was used with great advantage in a number of tests.

Chapter 4
The ALICE SPD and LHCb RICH hybrid assemblies

We have seen in Chapter 3 that the basic building block of the ALICE SPD system is the "ladder", a high resistivity p^+ on n silicon detector of around 13×71 mm^2 that is bump-bonded to five ALICE1LHCb pixel readout chips. For LHCb [LHC98] a single chip is bump-bonded to a single detector ("single", see also section 7.1.1 for more details). The front-end chip (which will be discussed in more detail in Chapters 7 and 8) is produced in a standard 0.25 μm CMOS process (a discussion on the beneficial effects of using a sub-micron technology for designing ICs for High Energy Physics is presented in Chapter 5). The chip is bump-bonded to a p-in-n sensor produced by Canberra Électronique [Add01]. Two different vendors have been selected for bump-bonding, VTT [Add02] and AMS [Add03].

This chapter gives an overview of the design of the ALICE and LHCb hybrid assemblies[1] and of the main issues related to bump-bonding. Electrical and irradiation tests of the sensors and of the bump-bonding process are presented, showing the behaviour within the specs of the sensor, and the big improvement achieved on the bump-bonding process.

4.1 The ALICE SPD sensors

The ALICE SPD sensor is an array of p-in-n diodes (the basic pixel is shown in Figure 4.1 and Figure 4.2), realised with doping of an n-type substrate, with a resistance of about 15 KΩ·cm (this can change in the future, but all the test results presented in this thesis refer to this value). The default value of the wafer thickness is 300 μm, but (apart from some wafers in the earlier test phases) they will be thinned (for ALICE only) down to 200 μm for material budget reasons.

Although sensor structure sizes are still fairly coarse in comparison with microelectronics, the requirements on technology in some aspects are more severe. This concerns in particular the processing of ultrapure silicon in a way that does not cause its properties to deteriorate by, for example, the introduction of impurities or the creation of

[1] If referring to hybrid pixel systems, we will define from now "sensor" the (passive) detecting part of the hybrid, while with "assembly" we will define a structure composed by a sensor and one or more front-end chips bump-bonded to it.

defects, the need of processing both wafer sides without damaging the opposite surface, and the production of wafer-sized defect-free detectors. These requirements demand that microelectronics technology and processing cannot be applied in a straightforward manner, and that special equipment and procedures have to be applied.

4.1.1 The basic pixel

For the design of the sensor five layers are used. They are shown in Figure 4.1, where a layout view of two basic pixels of the matrix is depicted (Figure 4.1, top) together with a cross section of a pixel (Figure 4.1, bottom).

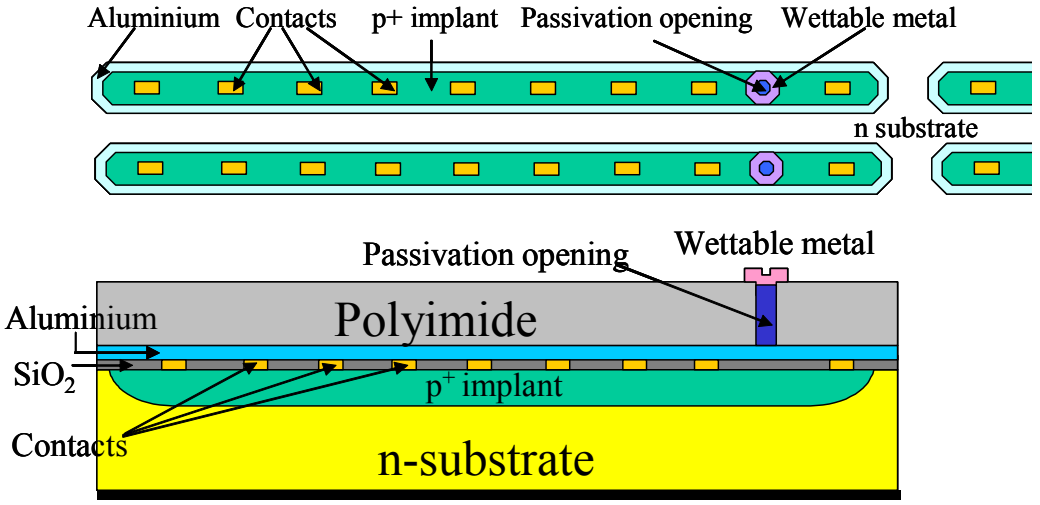

Figure 4.1: The basic pixel of the sensor: layout representation, with an indication of all the design layers used (top); schematic cross section representation (bottom).

The first step for detector fabrication consists in oxidizing wafers (a typical temperature is ~1000 °C) to have the whole surface passivated. Using photolithographic and etching techniques, windows are then opened in the oxide to enable ion implantation in the desired areas using appropriate masks. Doping of silicon can be performed by either ion implantation or diffusion. The p^+ pixels are then implanted by means of the IPD1 mask (p^+ implant in Figure 4.1). An annealing (a typical annealing is at ~600 °C for about 30 min.) is carried out after ion implantation to reduce the damage caused by the heavy ion irradiation.

For the backside contact of the LHCb wafers, ions of arsenic or phosphor are used. To minimize the energy lost by the 20 kV photoelectrons striking the backside, this ohmic surface is formed by a thin 150 nm n^+ implant, a standard fabrication option offered by the manufacturer.

The next step consists in depositing a layer of insulator (silicon dioxide) on the surface of the sensor, which will separate the silicon surface from the interconnecting metal. In some cases this step is not performed, either because the metallization mask is the same as the

implant mask, or because the annealing is enough to generate a layer of native silicon oxide. Photolithographic and etching techniques are used (if needed) to open windows in the oxide, in the position where the interconnection metal has to contact the silicon sensor. This corresponds to the CS1 mask ("Contact" in Figure 4.1).

A layer of aluminium is deposited on top of the insulator, and using the appropriate masks the desired interconnection pattern is realised, that will carry the signal. This corresponds to the ALD1 mask ("Aluminium" in Figure 4.1). The standard process covers also the backside of the wafer with aluminium. This is done for the ALICE1LHCb sensor which has to sense high energy particles, but not for the LHCbPIX1 one because this layer of aluminium could decrease the energy of the electrons impinging on the detector. In this case only a ring of aluminium is left, surrounding the chip backside.

A layer of polyimide is then deposited on the whole surface of the sensor, in order to protect it from mechanical or chemical damage. To allow the contact to the aluminium metallization, some openings are made in the polyimide layer (PAP1 mask, "Passivation opening" in Figure 4.1). In the ALICE and LHCb sensors the openings correspond to the contacts for the bump-bonding ball. For this reason an additional metal layer is deposited (WTME mask, "Wettable metal" in Figure 4.1), which has to acts as an interface metal between the pixel cell and the bump-bonding ball.

A more detailed description of the dimensions of the basic pixel is presented in Figure 4.2 (top).

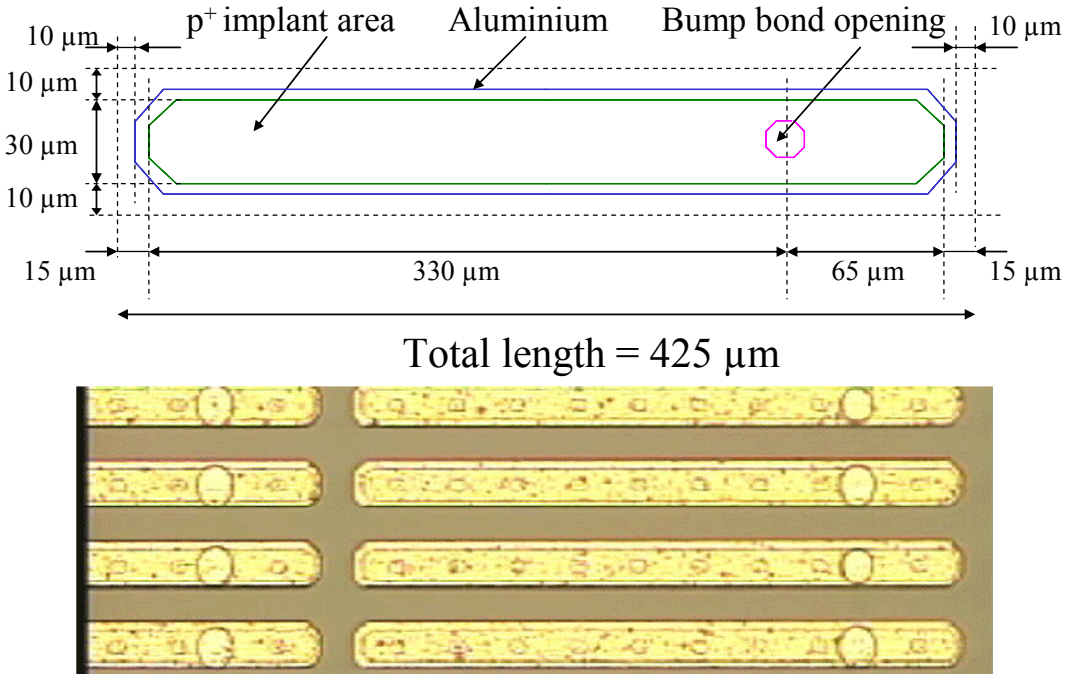

Figure 4.2: Basic pixel of the sensor. Detailed description of the basic pixel dimensions (top); microscope photograph (bottom).

The total pixel length is 425 μm, because it has to match the dimensions of the pixel on the electronics chip.

4.1.2 The single and the ladder

Two different sensors were designed, using as a basic element the pixel shown in Figure 4.1 and Figure 4.2.

One is an array of 32 columns of 256 basic pixels (which matches exactly a readout chip of the same size) for a total active area of 12.8 × 13.6 mm^2. This assembly is called "single", and is not an element of the final ALICE SPD system. Nevertheless, it has been designed to be used in the LHCb RICH detector, in the NA60 experiment pixel detector and for testing purposes. The second type of assembly (ladder) is an array of 160 columns of 256 pixels. The ALICE sensor is much bigger than a single, as it has an active area of 12.8 × 69.6 mm^2, and a total number of 40960 pixels. It is bump-bonded to five read-out chips.

At the periphery of each readout chip some space has to be provided for routing and dicing, so it would not be possible to bump-bond five readout chips on a sensor which is just laid out as a matrix of 160 columns of 256 basic pixels. This is why two columns of "junction" pixels are placed at the boundary of two chips, as depicted in Figure 4.3 (top). A junction pixel is similar to a basic pixel, but is 625 μm long, as shown in Figure 4.3 (bottom). The extra length provides the margin for the dead area present at the readout chip edges.

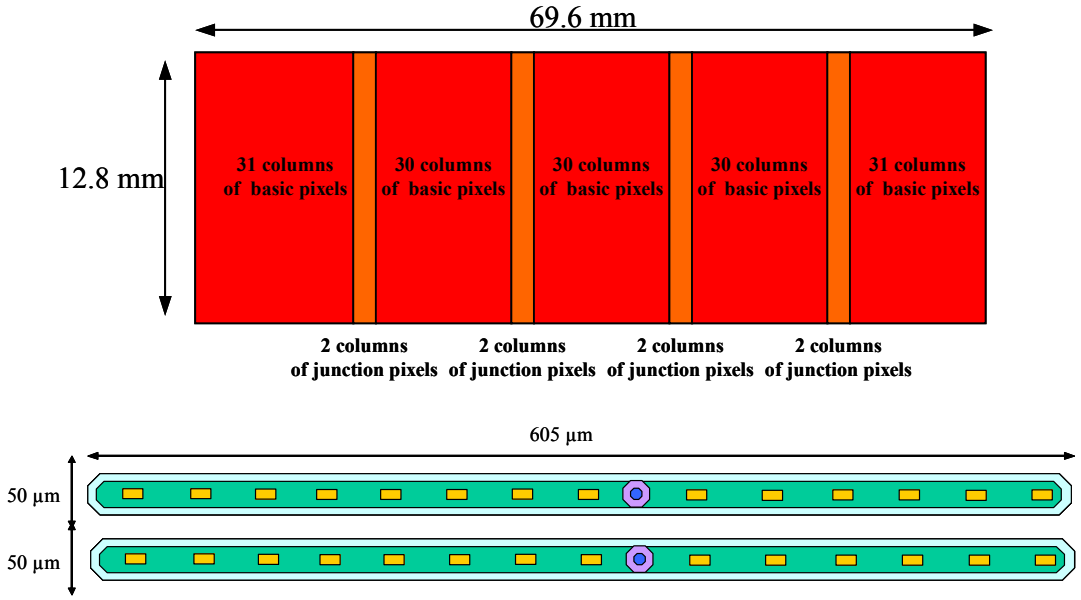

Figure 4.3: Schematic representation of a ladder sensor (top) and layout representation of two "junction" pixels (bottom). A junction pixel is similar to a basic pixel, but is 605 μm (+20 μm for pixel to pixel spacing) long. Two columns of junction pixels are placed at each boundary between two chips.

4.1.3 The sensor periphery

As shown in Figure 4.4 and Figure 4.5 (left), the pixel matrix is surrounded by a guard ring (150 µm wide).

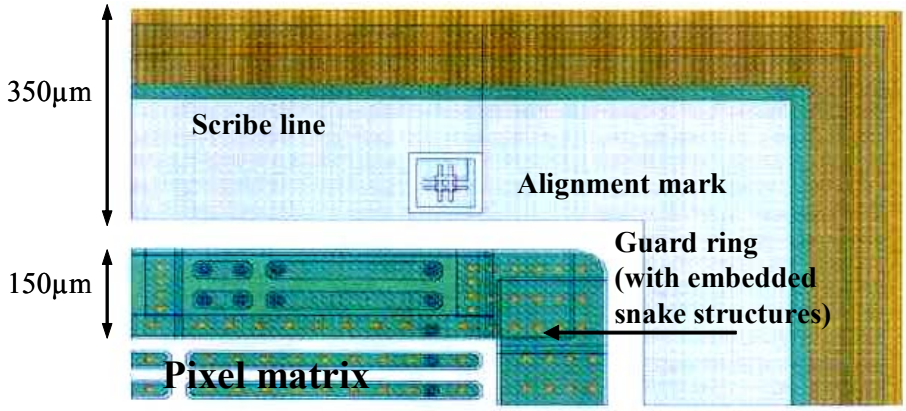

Figure 4.4: Layout of a corner of the pixel sensor. All around the pixel matrix there is a guard ring (with embedded "snake" structures). The outer ring is the scribe line, which is used to define the dicing path, and which embeds an alignment mark for bump-bonding.

The guard ring is a large diode which has the aim of collecting all the charge which is not collected by the pixel matrix and reduce the active sensor area leakage current. In effect it protects the active diode matrix from additional leakage, coming from the back side of the detector or generated by the influence of defects (cracks) introduced when cutting the wafer into chips (dicing). Moreover it reduces the electric field distortions at the edges of the pixel matrix.

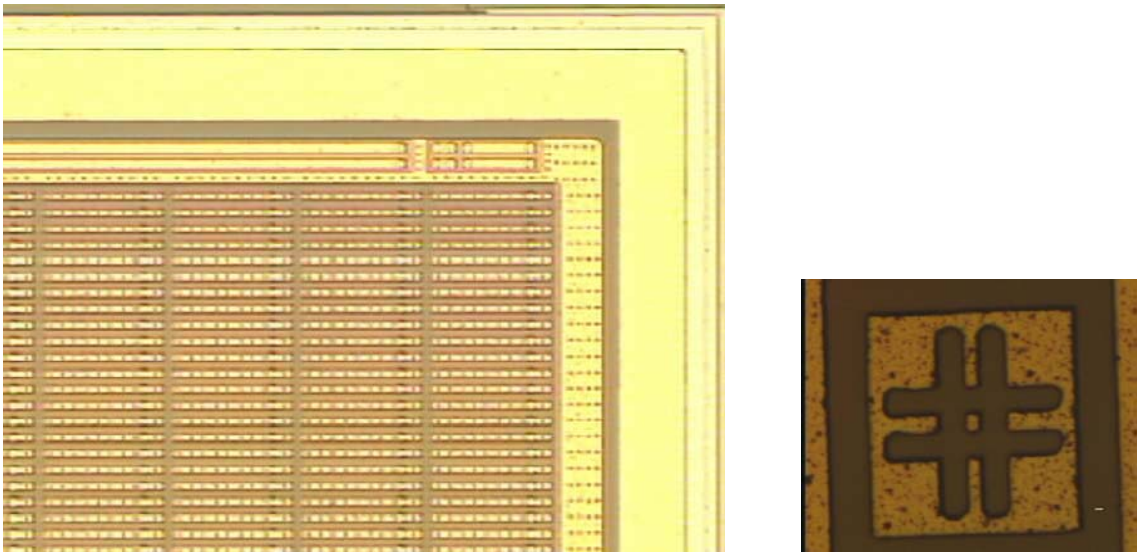

Figure 4.5: Microscope photograph of a corner of the pixel sensor (left). The pixel matrix and the guard ring with the embedded snake structure are evident. Microscope photograph of one of the alignment marks embedded in the scribe line that can be used during bump-bonding to correctly align the sensor chip with the readout chip (right).

The outer ring of the sensor (350 μm wide) is the scribe line, which defines the position where the chip will be diced. Some alignment marks are embedded for use during bump-bonding to correctly align the sensor chip with the read-out chip. The mask alignment guaranteed by the producer is better than 3 μm. A microphotograph of the alignment mark is shown in Figure 4.5 (right).

A special structure, the "snake" (Figure 4.4, Figure 4.5 (left) and Figure 4.6), has been embedded in the sensor guard ring, both for the single and for the ladder. This structure is composed of two halves: a series of non-contiguous metal strips on the detector (Figure 4.6, top), and a complementary series of metal strips on the readout chip. During bump-bonding these two sets of metal strips are connected together to form a long metal path which runs half on the sensor and half on the electronics chip (Figure 4.6, bottom).

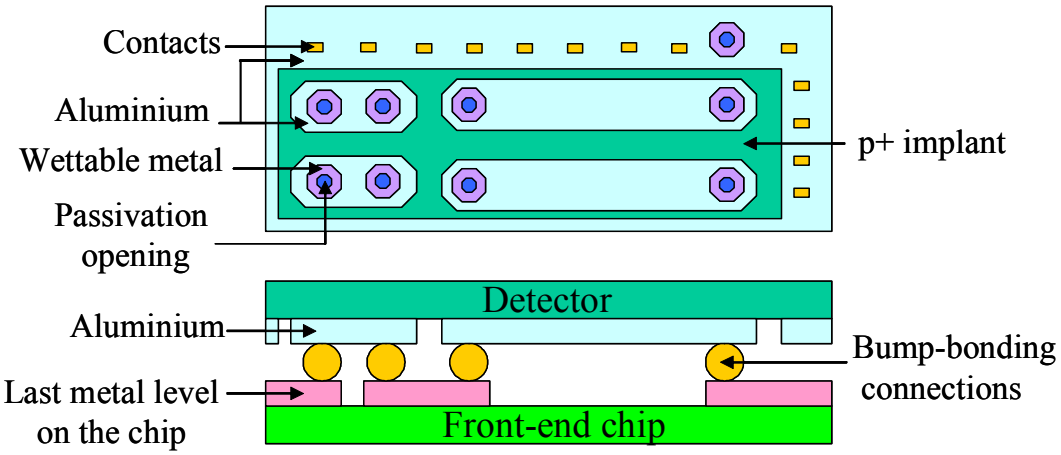

Figure 4.6: Layout of a section of the two parallel "snake" structures embedded in the guard ring. Top view, detector only (top) and cross section, detector and front-end chip bump-bonded (bottom). Only the relevant metal layers are shown in the cross section view. With this structure it is possible, using a probe-station, to verify the electrical continuity of the two snakes and the absence of snake-to-snake shorts.

At both ends of the snake there are metal pads, and two parallel snakes are embedded in each guard ring. In this way it is possible, using a probe-station, to verify the electrical continuity of the two snakes and the absence of snake-to-snake shorts.

This kind of test is a very demanding one for the bump-bonding process, as it requires 100% of good bump-bonds in the snake (and in a region of the assembly which is particularly prone to bump-bonding defects, i.e. the sides and especially the corners) and 0% of bump-to-bump shorts. This would allow the on-line monitoring of the bump-bonding quality. In effect, up to now the assemblies are shipped without any testing. The functionality test can only be performed with a full test system, irradiating the sensor with a radioactive source (as explained in section 4.2.2), so it is carried out here at CERN. This imposes a delay of at least one or two weeks between the production of the assembly and the possibility to feedback

information to the vendor about a problem in the bump-bonding process, which would then affect all the assemblies produced in the meanwhile. The possibility to perform the snake test already by the vendor would reduce this delay to only a few hours. Moreover, if the bump-bonding process quality reaches values close to 100% the snake test could be performed as an acceptance test already by the bump-bonding vendor.

Figure 4.7 shows a picture of a wafer. The sensors for the ladders are placed in the centre, and all around the sensors for the singles. Some test structures are placed at the four sides, which can be used to monitor the process quality.

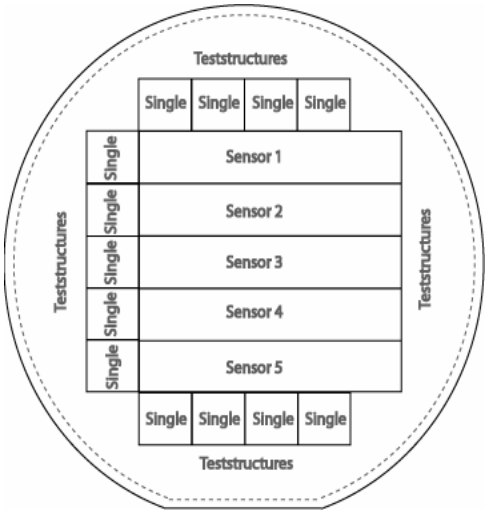

Figure 4.7: Picture of a wafer. The sensors for the ladders are placed in the centre, and all around the sensors for the singles. Some test structures are placed at the four sides, which can be used to monitor the process quality.

4.2 Tests of the ALICE SPD sensor

To validate the sensor for the final experiment, several tests have been performed on a pre-series set of Canberra wafers, both thick (300 μm) and thin (200 μm).

The manufacturer carries out a first set of tests on the wafer or on the on-wafer test structures. First of all it performs a microscope visual inspection, looking for implant or aluminium shorts and breaks. Also the wafer thickness is checked, as a wafer thickness uniformity of ±15 μm and a sensor thickness uniformity of ± 5 μm have to be guaranteed. Also the bow/wrap is guaranteed to be less than 30 μm.

4.2.1 Electrical tests

Electrical tests are performed on the 68 mm^2 test diodes present on the wafer; namely the full depletion voltage (V_{fd}) and the total leakage current are measured. The nominal electrical parameters of the sensor are summarised in Table 4.1.

Maximum full depletion voltage (V_{fd})	20-30 V
Operating voltage	$2 \times V_{fd}$
Minimum breakdown voltage	>100 V
Maximum leakage current @ V_{fdmax} (20 °C)	5 nA/cm^2

Table 4.1: Nominal electrical parameters of the sensor, as specified by Canberra (300 µm thick detectors).

Some additional tests, for example oxide thickness or flat band voltage, are done from time to time on some wafers. The results, reported in Figure 4.8, are from two different deliveries: one of 15 thick wafers (2001) and one of 10 thin wafers (2003).

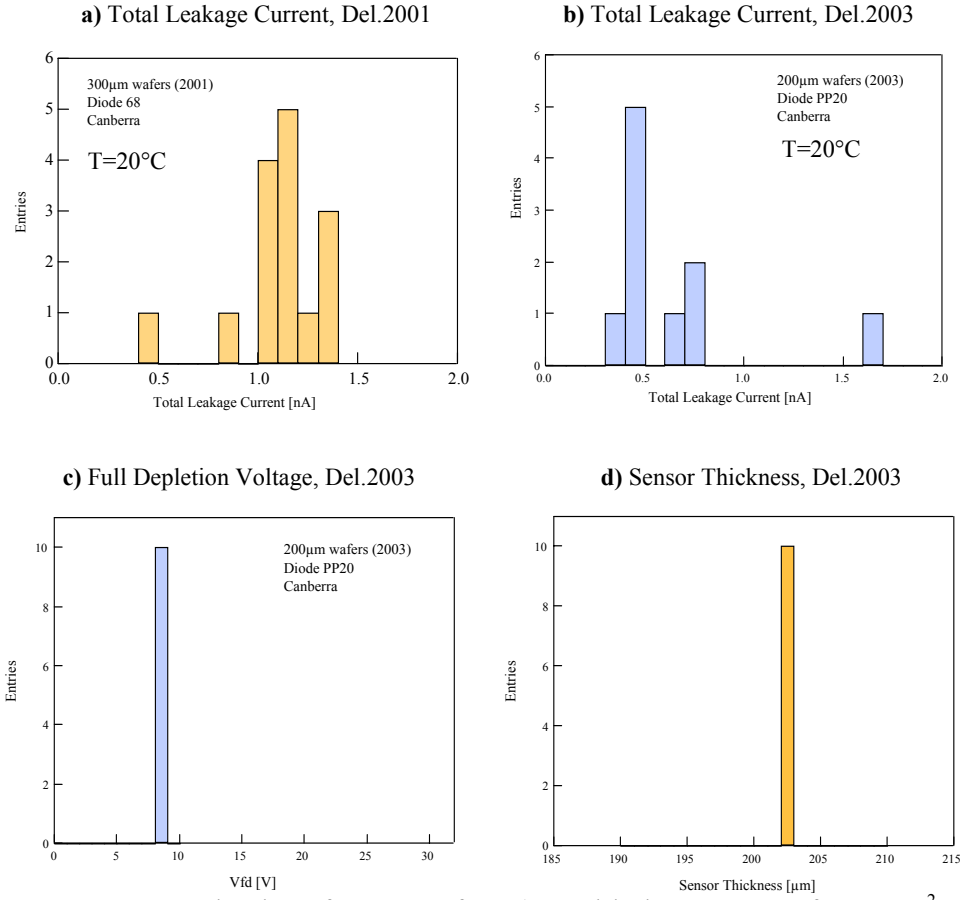

Figure 4.8: Tests on two batches of sensor wafers. a) Total leakage current of a 68 mm^2 test diode, delivery 2001 (15 wafers 300 µm thick); b) Total leakage current of a 68 mm^2 test diode, delivery 2003 (10 wafers 200 µm thick); c) Full depletion voltage of a 68 mm^2 test diode, delivery 2003; d) Sensor thickness, delivery 2003 [Rie03b].

The nominal leakage current of 5 nA/cm^2 gives a total leakage current of 3.4 nA for the 68 mm^2 of the test diode. Both on thick wafers (Figure 4.8 (a)) and on thin wafers (Figure 4.8 (b)) the leakage current (< 1.7 nA) is quite uniform and well below the maximum nominal value. Also the full depletion voltage (8 V), measured on the same test diode (Figure 4.8 (c), delivery of 2003 only), is very uniform and well below the maximum nominal value of 32 V.

Figure 4.8 (d) shows the sensor thickness (delivery of 2003 only), which is within 1.5% of the nominal value of 200 µm. Most importantly of all it is extremely uniform, as all the measured sensors have exactly the same thickness.

Some tests on these pre-series wafers have been carried out also at CERN and at Legnaro[2]. Current-Voltage (I-V) curves measured on some test diodes are reported in Figure 4.9. They are measured on 8 thin wafers, 1 delivered in 2000 and 7 delivered in 2003. Though the wafer from 2000 has a higher I-V curve, and one of the wafers from 2003 has an I-V characteristic which has a shape quite different from the others, the leakage current measured at full depletion voltage is within the specifications for all the wafers. Similar tests were performed after dicing, and are fully compatible with the results presented in Figure 4.9.

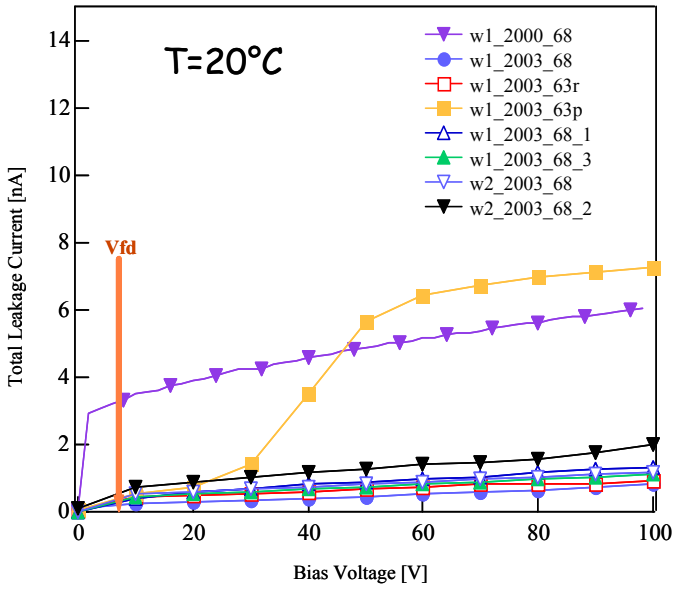

Figure 4.9: Total leakage current measured on the test diodes of 8 thin wafers, 1 delivered in 2000 (thick) and 7 delivered in 2003 (thin). The leakage current measured at the full depletion voltage is within the specifications for all the wafers [Rie03b].

4.2.2 Bump-bonding tests

The two bump-bonding vendors use different processes. VTT uses a lead-tin bump at eutectic concentration for ALICE and with high lead concentration for LHCb. The reason for this is that LHCb assemblies have to stand much higher temperatures due to the subsequent

[2] INFN Laboratori Nazionali di Legnaro, Padova, Italy.

process steps required for the fabrication of the hybrid phototube [LHC98]. AMS uses a low temperature indium deposition process. A SEM photograph of a Pb-Sn bump and of an array of indium bumps is shown in Figure 4.10.

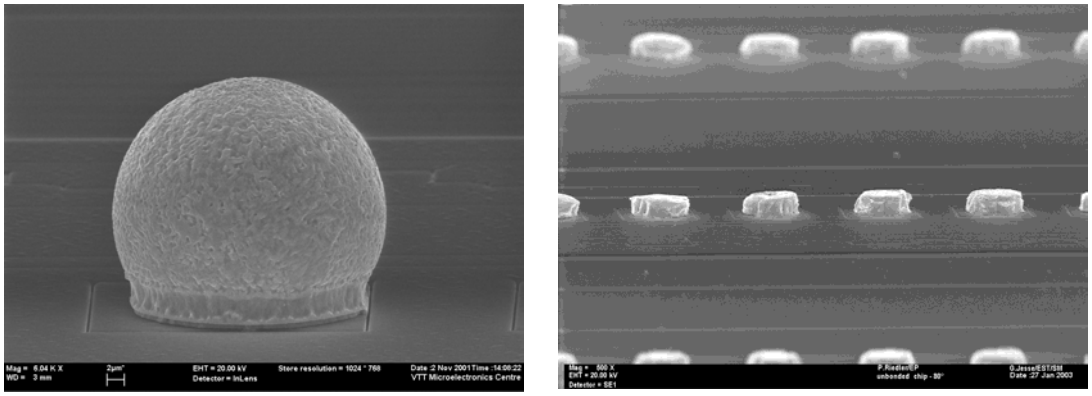

Figure 4.10: SEM photograph of a Pb-Sn bump (VTT process, left; courtesy of VTT) and of an array of indium bumps (AMS process).

During bump-bonding, each single pixel on the sensor is connected with each single readout channel on the electronics chip. So, the most efficient way to test the quality of the process is to expose a bump-bonded assembly to a highly energetic radioactive source, leaving the assembly exposed for a time long enough that each pixel in the matrix is hit several times. In this way pixels that have a missing or inefficient bump-bonding connection show up as never being hit.

Figure 4.11 and Table 4.2 show the results of exposing a VTT thin ladder to a strontium 90 radioactive source. As can be seen, on this ladder results are very good, since the maximum amount of missing pixels on a chip is <0.3%. The target fixed by the ALICE experiment is 1%.

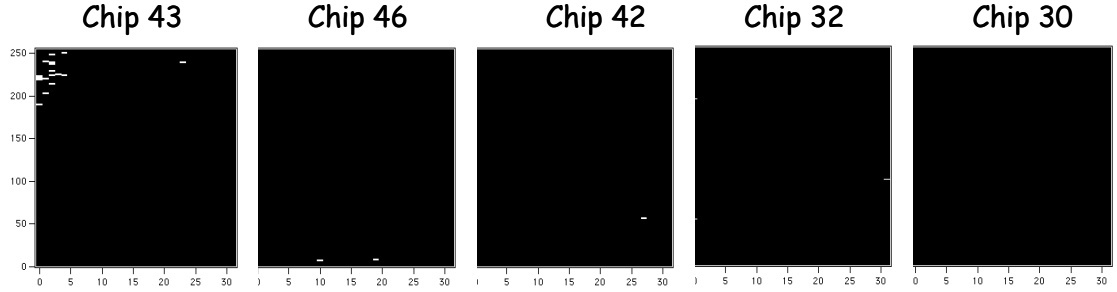

Figure 4.11: Results obtained exposing a thin VTT ladder to a Sr-90 radioactive source. The non-responding pixels are marked white. Numerical details are given in Table 4.2 [Rie03b].

	Chip43	Chip46	Chip42	Chip32	Chip30
Working pixels	99.7%	99.95%	99.98%	99.98%	100%
Missing pixels	28	4	2	2	0

Table 4.2: Numerical details of working bump-bonding connections relative to the thin VTT ladder shown in Figure 4.11 [Rie03b].

The quality of the bump-bonding process has not been always as good, and has improved with time from assemblies with corners missing to the present high-quality process. Reducing the bump bonding pitch to 50 μm decreased the volume of the bump bonding metallic ball by a large amount, and this created more problems than expected. Some other problems had to be addressed by the vendor: an interesting example is shown in Figure 4.12, which plots the leakage current versus bias voltage curves of some of the ladders received throughout 2001-2003. The plot shows a leakage current which is much higher than the nominal one. After many trials to understand the source of this leakage current (which for example could be decreased by an order of magnitude baking the wafers for 40 hours at 120 °C and 136 hours at 140 °C), it was traced back to a contamination of the sensor surface during deposition, due to an improper etching of the field metal (interface metal between the aluminium of the sensor and the Sn-Pb bump). The addition of a final plasma etching step to the process cured the problem, taking back the leakage current to the nominal value.

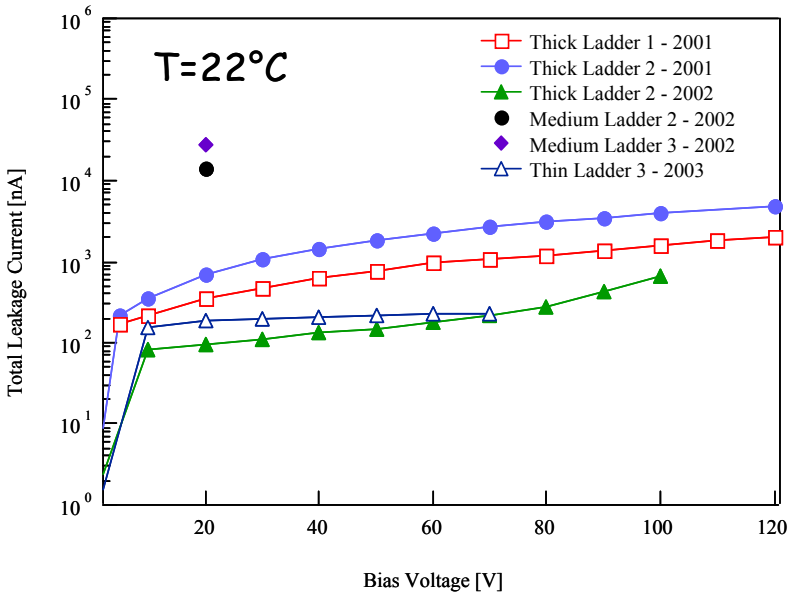

Figure 4.12: Leakage current of some of the ladders received throughout 2001-2003, as a function of the applied bias voltage. Its value is much higher than the nominal one. The problem was traced back to sensor surface contamination during bump-bonding [Rie03b].

From 9 VTT ladders of different thickness delivered since December 2001 only four out of the seven tested met the ALICE requirements of having >99% working bump-bonded pixels (which results in a overall yield of 42%). Though this is quite far away for the target value of >80% after bump-bonding, more tests have to be carried out to increase statistics. Moreover, a new jig has been purchased by VTT which is better optimised for ALICE ladders. With the old bonding machine, the sensor was first placed on the jig, and then three chips were bump-bonded to it. The sensor had then to be displaced to bump-bond the last two chips. With the upgraded jig all the chips can be bump-bonded without displacement of the sensor, and this should result in an improvement of the process yield.

4.2.3 Irradiation tests

Several irradiation tests have been performed on sensors alone and on bump-bonded assemblies, here we will report tests on sensors. Tests on the bump-bonded assemblies are presented in Chapter 8, together with the results of the front-end chip.

4.2.3.1 Tests with the 27 MeV Legnaro proton beam

Irradiation tests of the test 68 mm^2 diodes with a 27 MeV proton beam were performed at Legnaro, with fluence from 0.2 to 25×10^{12} protons/cm^2. The beam was focussed in a 5×5 cm^2 area and an array of 3×3 Faraday Cup[3] was used to read the current and calibrate the beam uniformity. The measurements were performed under a vacuum. The precision reached for the measurements here reported is about 7%. The nominal temperature was set at 25 °C, and it was read with a temperature sensitive resistor connected to the chip frame. For this reason the temperature during irradiation was not well controlled. Moreover, the sensor guard ring was connected with the diode, so the diode volumetric leakage current contribution and the surface leakage current contribution could not be separated. Both Capacitance-Voltage (C-V) and Current-Voltage (I-V) measurements were performed.

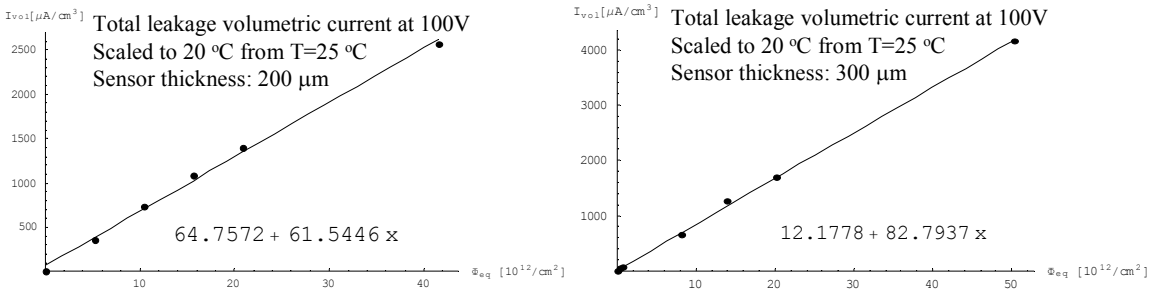

Figure 4.13: Sensor leakage current as a function of the particle equivalent fluence. 200 μm detector, $\alpha = 0.62\ 10^{-16}$ A/cm (left); 300μm detector, $\alpha = 0.83\ 10^{-16}$ A/cm (right) [Raw data after Rie03b].

According to equation 2.2, if we plot the leakage current as a function of the proton fluence, we can estimate the current damage rate constant α (also called volumetric leakage current increase rate). To do this we fit the points with a straight line, and extract the slope. If we now normalize the proton fluence to 1 MeV Neutrons using equation 2.4, with the hardness factor k = 2.08 given in [Bis01] for 27 MeV protons, we can plot the leakage current as a function of the equivalent fluence end extract the damage rate constant α. Figure 4.13 plots the sensor leakage current both for a thin detector (left) and for a thick detector (right). The α values (calculated scaling the I_{vol} at 20 °C with equation 2.3) are respectively $0.62\ 10^{-16}$ A/cm and $\alpha = 0.83\ 10^{-16}$ A/cm. These values, if taking into account the uncertainty of the temperature and the guard ring additional current are compatible with measurements which can be found in literature (around $0.4\text{-}0.45\ 10^{-16}$ A/cm [Can01, Mol02, Can03]).

[3] A Faraday Cup allows a beam of charged particles (electrons, ions) to be measured accurately. It collects all the particles which enter it, and leads them to an amperometer.

After exposing the sensor to the maximum fluence, it was then annealed at 25 °C to monitor the leakage current behaviour with time (V_{bias} = 100V). Figure 4.14 shows the leakage current of a thin sensor as a function of the annealing time. As can be seen, it starts at 75 µA immediately after irradiation, and then it decreases about 7% in the first two hours and about 21% in the first three days. It stays almost constant from the third day on.

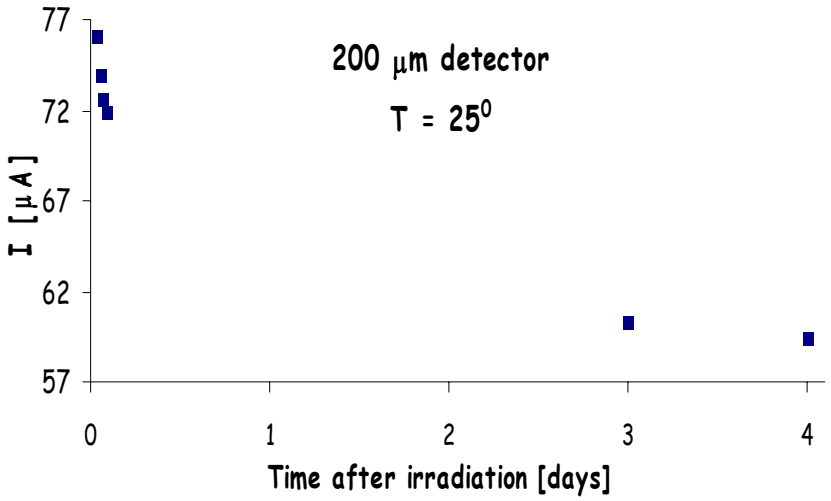

Figure 4.14: Thin sensor annealing at 25 °C. Leakage current as a function of time. It starts at 75 µA immediately after irradiation, and then it decreases about 7% in the first two hours and about 21% in the first three days, than it stays almost constant from the third day on [Rie03b].

Figure 4.15 plots the leakage current as a function of the applied bias voltage; each curve is an I-V curve taken after 50, 80, 105, 125 minutes and after 3 days from the irradiation.

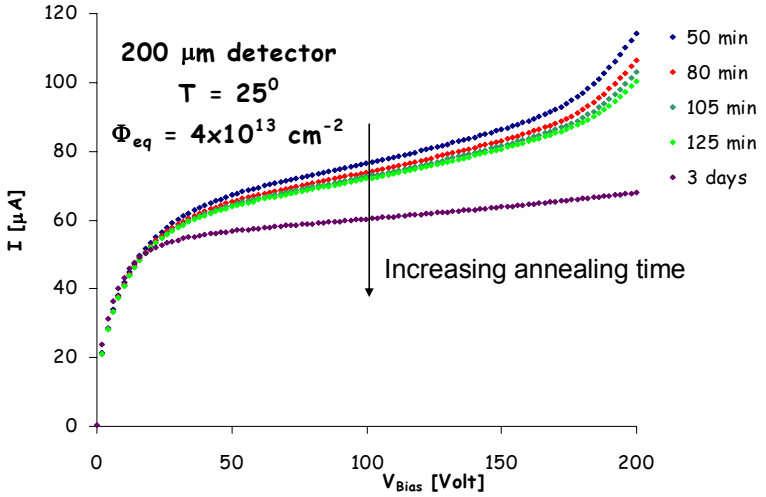

Figure 4.15: Thin sensor annealing at 25 °C. Leakage current as a function of the applied bias voltage, after 50, 80, 105, 125 minutes and after 3 days (bottom). The plot shows an evident reduction of the leakage current with annealing time (for high bias voltages) [Raw data after Rie03b].

The plot shows that there is an evident reduction of the leakage current with annealing time (for high bias voltages). Results for thick ladders show a similar behaviour, but starting at higher leakage current values.

The full depletion voltage (V_{fd}) was measured after irradiation; the results are reported in Figure 4.16. Figure 4.16 (left) shows the depletion voltage of a thin sensor as a function of the equivalent fluence, measured at two different frequencies, 1KHz and 10 KHz. Results are quite similar and indicate that type inversion of the sensor would happen at an equivalent fluence of about $8 \cdot 10^{12}$ cm^{-2}, and that V_{fd} still stays lower than 100 V for an equivalent fluence which is more than twelve times the expected dose for the inner tracking layer of ALICE in 10 years of operation [Pas03]. Very similar results are obtained for thick wafers, as shown in Figure 4.16 (right), where the depletion voltage of a thick sensor as a function of the equivalent fluence is plotted for three different samples.

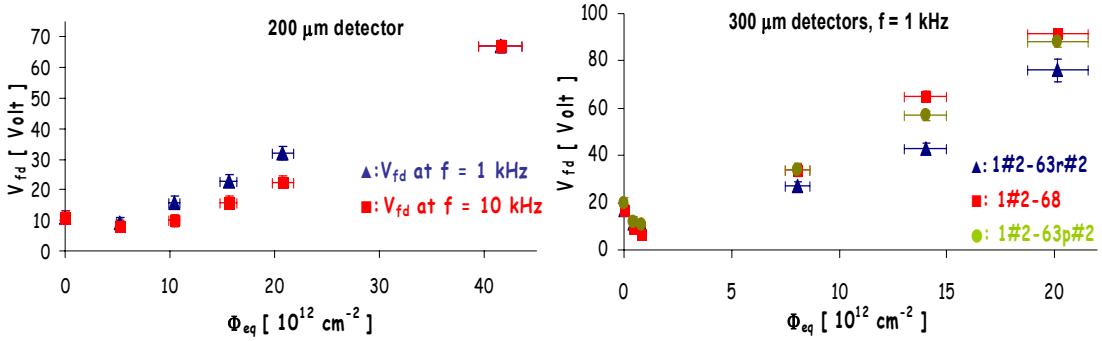

Figure 4.16: Depletion voltage as a function of the equivalent fluence: for a thin sensor with measurements done at 1 and 10 KHz (left) and for 3 different samples of thick sensors (right). Measurements are done at 1KHz. V_{fd} still stays lower than 100 V for en equivalent fluence which is more than twelve times the expected dose for the inner tracking layer of ALICE in 10 years of operation (and is even higher for thin detectors) [Raw data after Rie03b].

4.2.3.2 Tests with the 24 GeV CERN proton beam

More sensor irradiations were performed at the CERN T7 beam line, with 24 GeV protons. The same type of 68 mm^2 diodes (300 µm thickness only) was irradiated, with fluence from 3.36 to 16×10^{12} protons/cm^2. The measurements were not performed under a vacuum; the temperature was monitored during the experiment with a temperature sensitive resistor, and remained constant at 30.4 °C. The sensor guard ring was not shorted with the diode and the measurements were done immediately after each irradiation step. The sensors were biased at 50 V during irradiation (the full depletion voltage before irradiation is ~20V).

Figure 4.17 plots the pad leakage current versus bias voltage for different proton fluences, the behaviour is consistent with previous measurements.

To estimate the current damage rate constant α we fit the points with a straight line with a least mean square algorithm, and extract the slope.

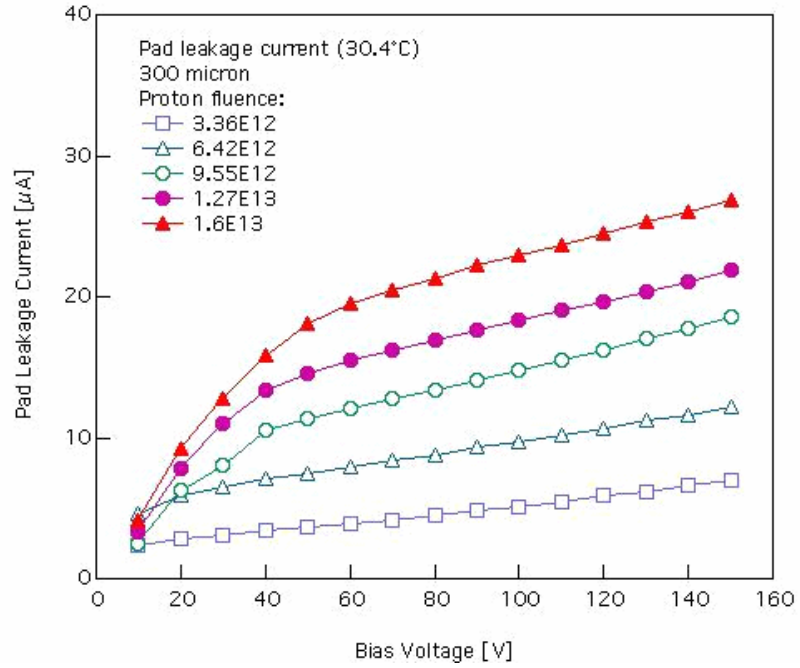

Figure 4.17: Total pad leakage current versus bias voltage for different fluences [Rie03a].

The value of the constant α for 24 GeV protons (fluence not normalised) is $2.53 \cdot 10^{-17}$ A/cm for a 70 V bias and $2.8 \cdot 10^{-17}$ A/cm at 100 V, very close to the value of $2.62 \cdot 10^{-17}$ A/cm cited in [Can01] and $2.67 \cdot 10^{-17}$ A/cm cited in [Ruz99]. If we now normalize the proton fluence to 1 MeV Neutrons using the hardness factor of 0.62 given in [Mol02], we can plot the pad leakage current as a function of the equivalent fluence (as shown in Figure 4.18).

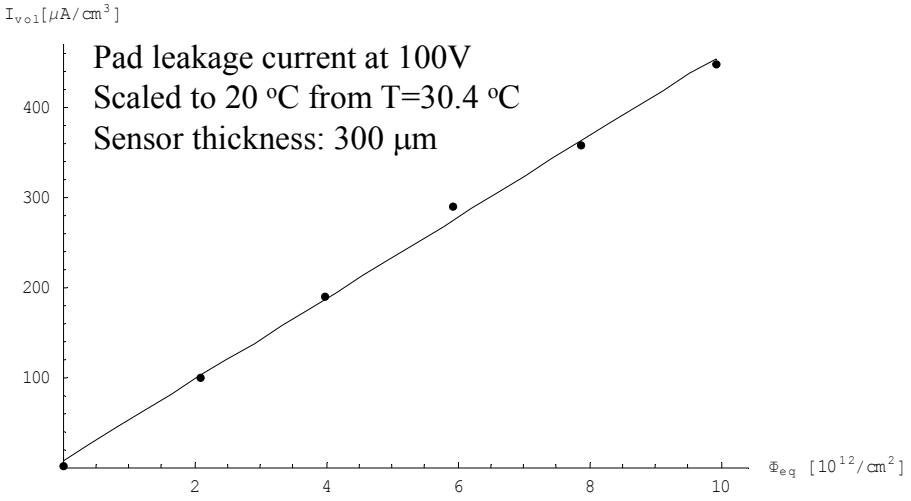

Figure 4.18: Pad leakage current as a function of the equivalent fluence. The hardness factor used is k = 0.62. We can extract the damage rate constant $\alpha = 4.52 \times 10^{-17}$ A/cm.

We can then extract the damage rate constant $\alpha = 4.52 \times 10^{-17}$ A/cm which is again in good agreement with measurements done in the already cited works. The bias voltage of 100 V is a likely value for the ALICE detector bias voltage because, as we have seen, it ensures full depletion of the sensor also after high radiation fluence.

4.2.4 Metrology measurements

To check that also the wafer thickness and bend are within the ALICE specifications, some metrology measurements were done at CERN at the metrology lab. The flatness of the wafer was measured both along the sensors, in three different positions as shown in Figure 4.19 (left) and across the sensors in four different positions, as shown in Figure 4.19 (right).

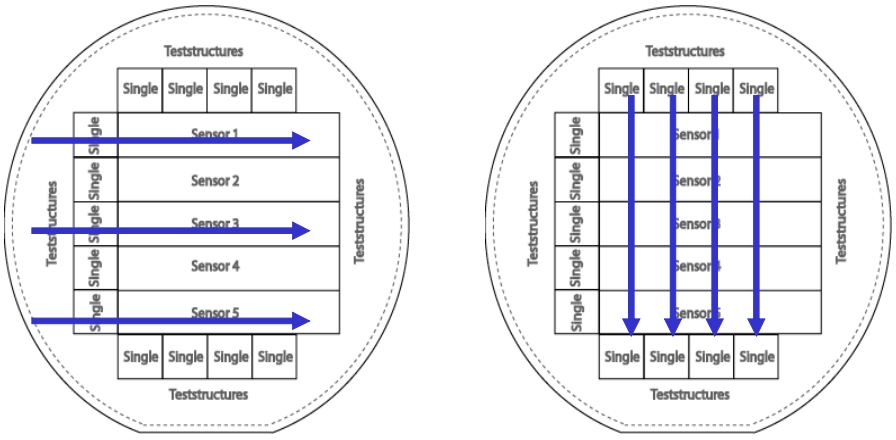

Figure 4.19: Metrology measurements done on the sensor wafers: along the sensors (left) and across the sensors (right).

The wafer flatness measured along the sensors (on a thin wafer delivered in 2001) is better than 4.43 µm, 9.9 µm and 9.11 µm for each of the three directions shown in Figure 4.19 (left). An example of such a measurement is shown in Figure 4.20.

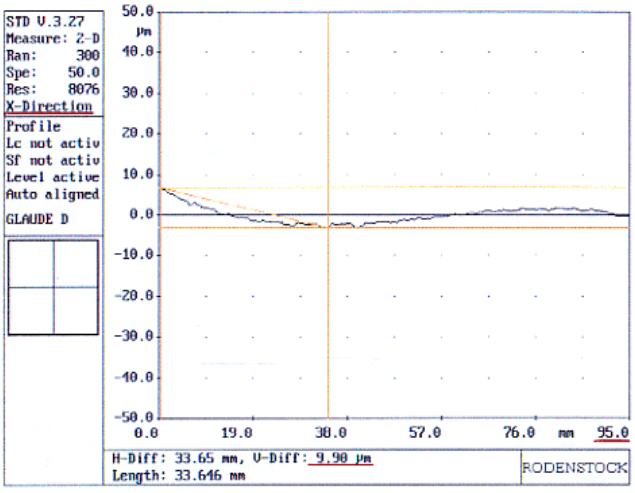

Figure 4.20: Example of a metrology measurement. In this case the flatness is better than 9.9 µm.

The flatness of the same wafer measured across the sensors is also better than 10 μm. The results of these measurements are fully within the ALICE specifications, as for bump-bonding the flatness of the sensor has to be better than 30 μm.

4.3 Summary

The ALICE and LHCb silicon pixel sensor is an array of p-in-n diodes, realised with phosphor doping of an n-type substrate. The final value of the wafer thickness will be 300 μm for LHCb ("thick detectors") and 200 μm for ALICE (for material budget reasons, "thin detectors"). The design of the silicon sensor was presented in this chapter, as well as some test results.

The sensor is manufactured by Canberra Electronique; a description of the main phases of the process, corresponding to the design masks, was given.

The basic pixel is 425 μm long and 50 μm wide; a "single" assembly is an array of 256×32 basic pixels bump-bonded to a readout chip (the singles will be used by LHCb, for testing purposes and by NA60). The ALICE sensor is much bigger than a single, as it has an active area of 12.8×69.6 mm^2, and a total number of 40960 pixels. It is bump-bonded to five read-out chips. At the periphery of each readout chip some space has to be provided for routing and dicing, so it would not be possible to bump-bond five readout chips on a sensor which is just laid out as a matrix of 160 columns of 256 basic pixels. This is why two columns of "junction" pixels are placed at the boundary of two chips. A junction pixel is similar to a basic pixel, but is 625 μm long.

The pixel matrix is surrounded by a guard ring (150 μm wide), a large diode which has to protect the pixel array from additional leakage current. A special structure, the "snake", has been embedded in the sensor guard ring, which could allow testing the bump-bonding process quality already by the bump-bonding vendor.

To validate the sensor for the final experiment, several tests have been performed (in addition to the tests done by the manufacturer) on a pre-series set of Canberra wafers, both thick and thin. Electrical tests are performed on the 68 mm^2 test diodes present on the wafer; namely the full depletion voltage (V_{fd}) and the total leakage current are measured. Apart from some problems (which were understood and solved) on the first lots, all the electrical characteristics of the sensor meet the manufacturer specifications.

The leakage current, measured on a test diode, is quite uniform and its maximum is half the nominal value. Also the full depletion voltage (8 V), measured on the same test diode showed to be very uniform and well below the maximum nominal value of 32 V.

Tests were carried out on assemblies to asses the quality of the bump-bonding process. On some of the latest VTT thin ladders, tested with a strontium 90 radioactive source, the maximum amount of missing pixels on a chip is <0.3%, below he target fixed by the ALICE experiment (1%).

Irradiation tests of the test 68 mm^2 diodes with a 27 MeV proton beam were performed at Legnaro, Italy, with fluence from 0.2 to 25×10^{12} protons/cm^2. The current damage rate constant α (for the equivalent 1 MeV Neutrons fluence, calculated scaling the I_{vol} at 20 °C) was measured to be $\alpha = 0.62\ 10^{-16}$ A/cm for thin detectors and $\alpha = 0.83\ 10^{-16}$ A/cm for thick detectors. Annealing measurements show that the leakage current decreases after 3 days of annealing at room temperature of 21%, and then stays constant.

The full depletion voltage (V_{fd}) was measured after irradiation at two different frequencies, 1KHz and 10 KHz, and for thin and thick sensors. V_{fd} stays lower than 100 V for en equivalent dose which is twelve times the expected dose for the inner tracking layer of ALICE in 10 years of operation (and even higher for thin detectors).

These measurements suffered from some imprecision, so more sensor irradiations were performed with 24 GeV protons, with fluence from 3.36 to 16×10^{12} protons/cm^2. In this case we measured a damage rate constant (for the equivalent 1 MeV Neutrons fluence) $\alpha = 4.52\ 10^{-17}$ A/cm which is in good agreement with measurements found in literature.

To check that also the wafer thickness and bend are within the ALICE specifications, some metrology measurements were done at CERN at the metrology lab. The results of these measurements are fully within the ALICE specifications.

As a conclusion, we can state that the Canberra sensors perform according to (or even better than) the specifications, that they can stand the ALICE radiation levels, so that they can be safely used in the experiment.

Chapter 5
Electronics for High Energy Physics

In the LHC experiments particles will be accelerated to reach energies in the TeV range and collided head on at very high luminosities (10^{34} $cm^{-2}s^{-1}$ for protons and $1.95 \cdot 10^{27}$ $cm^{-2}s^{-1}$ for lead ions). For this reason, the radiation levels that can be attained are very high, especially for the inner detectors, very close to the interaction point. Table 5.1 presents the total ionising dose and the equivalent neutron fluence for the ALICE experiment. They were determined with the FLUKA simulator by B. Pastircak [Pas03], taking into account also beam-gas interactions and radiation originating from particle production at the interaction point, for the foreseen 10 years of running of the LHC: 10^8 seconds of proton-proton collisions, $5 \cdot 10^6$ seconds of lead-lead collisions and 10^6 seconds of Argon-Argon collisions. The statistical error of the simulation is about 30%.

	Dose [Gy]	**Dose [rd]**	**Fluence [1MeV n_{eq}/cm^2]**
SPD Layer 1	2.5×10^3	$2.5\ 10^5$	2.95×10^{12}
SPD Layer 2	6.9×10^2	$6.9\ 10^4$	1.72×10^{12}

Table 5.1: Total ionising dose and neutron fluence for the ALICE experiment, in the two SPD layers. They were determined with FLUKA taking into account also beam-gas interactions and radiation originating from particle production at the interaction point, for the foreseen 10 years of running of the LHC [Pas03].

A safety factor of two is applied to the previous data, so that the working values for the ALICE SPD are 500 krd for TID and $6 \cdot 10^{12}$ for the total fluence.

These values (although lower than the corresponding values for other experiments at LHC, such as CMS or ATLAS, for example) point out that the readout electronics (in particular for the inner trackers) has to be able to stand high radiation doses. This chapter deals with IC irradiation-induced problems and their possible solutions.

The physical effects of radiation on MOS devices are presented. These translate in a radiation-induced change of the electrical parameters of the MOS devices, described in section 5.2. Single Event Upsets are then explained, to complete the discussion about detrimental radiation-induced effects in ICs.

The three possible solutions to IC irradiation-induced problems are presented. One is to select commercial components (components off the shelf, **COTS**) and test them, setting up a database of selected standard electronic components which have been characterised and qualified as robust against radiation effects.

A second possibility is to design and fabricate the electronics in **qualified radiation hard processes** provided by some specialised vendors (e.g. Atmel, with DMILL technology).

These two solutions have several disadvantages; this lead CERN to set up a research project[1] to investigate the possibility of using a **standard commercial CMOS technology**, hardened with special design techniques, for High Energy Physics ICs. For this reason the advantages and drawbacks of scaling down to submicron technologies, both for what concerns radiation induced effects and circuit performance are presented in 5.5. Several advantages of submicron technologies are point out, but some of the detrimental effects related to radiation are not eliminated by technology scaling. The solution to these problems comes from Hardening By Design (HBD) techniques, which are presented in section 5.6. This type of approach, which proved to be very successful, has been used to design the ALICE1LHCb chip. For this reason the characterisation of the basic building block of the HBD, the n-channel edgeless transistor, is presented in detail, as well as an evaluation of the impact of HBD techniques at system level.

5.1 Radiation effects on MOS devices [Ane00, Boe85, Lac03, Mcl89, Win89]

Looking at the effects of radiation on matter (which are summarised in [Din04]), we can group them in two classes: ionisation effects and nuclear displacement. These phenomena can be generated directly by the incident particle or by the secondary particles, and represent the vast majority of the events that happen in irradiated matter.

5.1.1 Atomic displacement

If the particle passing through the medium has enough kinetic energy, it can give origin to an atomic displacement, generating a neighbouring atom and a vacancy (this is called a Frenkel pair). At room temperature 90% of the Frenkel pairs recombine within a minute after the end of the irradiation. Moreover, MOS devices (Figure 5.1) are almost insensitive to atomic displacement damages, since the conduction principle is not based on the properties of the bulk (which can be damaged by atomic displacements, for example reducing the minority carriers lifetime), but on the flow of majority carriers in a very thin region below the silicon dioxide (SiO_2) - silicon (Si) interface.

[1] RD49 – Study for radiation tolerance of ICs for LHC.

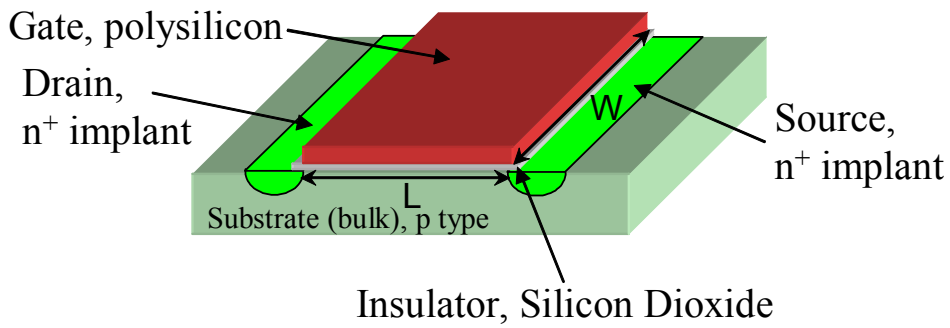

Figure 5.1: Schematic drawing of an n-channel MOS transistor. The current flows from drain to source underneath the SiO_2-Si interface, and is modulated by the gate voltage. The SiO_2 under the gate is called "gate oxide".

5.1.2 Ionisation

As an ionising particle traverses a medium, it produces electron-hole pairs. The amount of pairs generated is proportional to the energy deposited in the material, and the total damage is also roughly proportional to the total dose received by the material. So, even if different particles interact with different media - provided that they interact mainly by ionisation - we can refer to the Total Ionising Dose (TID)[2] as the only important quantity.

The part that is more sensitive to ionising radiation in a MOS device is the SiO_2. A simplified description of the most important phenomena which can take place due to irradiation is shown in Figure 5.2.

When an ionising particle passes through a MOS structure (Figure 5.2 (1)), it produces electron-hole pairs (Figure 5.2 (2)). These pairs quickly disappear in the gate (metal or polysilicon) and in the substrate, due to the little resistance of these materials, while in the oxide, that is an insulator, they behave quite differently. Some of the electron-hole pairs recombine immediately after their generation, the others are split apart by the electric field. The percentage of recombined pairs depends on the LET[3] of the incident radiation and on the electric field applied to the oxide.

Two different recombination models can be applied to the recombination of electron-hole pairs in SiO_2, the columnar model and the geminate model. The electron-hole pairs generated by the impinging particle are characterised by two parameters: the equilibrium distance r_t of the electron-hole separation once they reach thermal equilibrium, and the mean separation λ between electron-hole pairs. In SiO_2 r_t is thought to be ~5 nm. The columnar model is

[2] The Total Ionising Dose, which is a measure of the total amount of energy absorbed by the material, is often measured in rads (Si) or rads (SiO_2), symbol rd (as suggested by the IEEE), a unit equal to 100 ergs absorbed per gram of the material specified. The SI unit is the Gray (Gy), 1Gy=100 rds.

[3] The Linear Energy Transfer expresses the linear transfer of energy to the material by the incident (charged) particles. For photons, which are not charged, the energy transmitted to the target is expressed by the *absorption coefficient*.

applicable if $\lambda << r_t$, that is the case for a particle with high LET. In this case we can imagine to have a dense column of pairs with diameter r_t. If on the contrary $\lambda >> r_t$ the geminate model can be applied. The effect of the electric field on the recombination process is to separate electrons and holes, which will result in less recombination. More details about the two models and the recombination process can be found in [Mcl89].

Due to the huge difference of their mobility in the oxide electrons can drift towards the gate[4] (Figure 5.2 (3)) in a time of the order of some picoseconds or less[5], whereas holes (with much lower mobility in the oxide[6]) drift towards the SiO_2-Si interface in a much longer time (Figure 5.2 (4)).

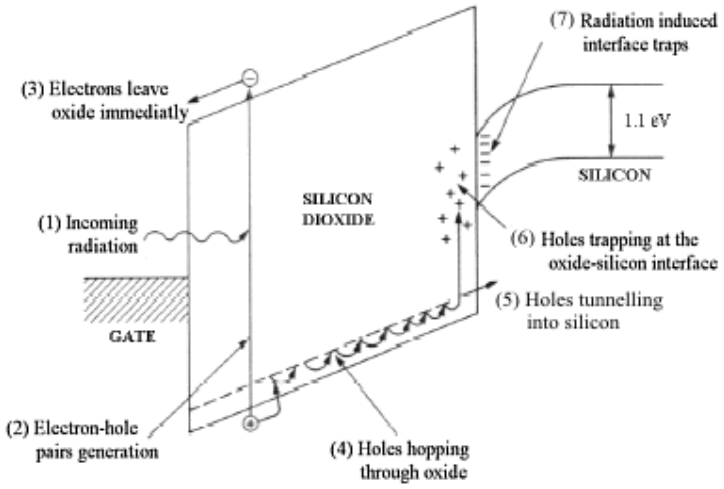

Figure 5.2: Schematic band diagram representation of the effects induced by an ionising radiation in a MOS structure (positive gate bias). The incoming ionising radiation (1) creates electron-hole pairs in the silicon oxide (2) which are separated by the electric field. Electrons leave the oxide in some picoseconds (3), holes hopping transport through localised states in the oxide is much slower (4). Close to the SiO_2-Si interface holes can tunnel into silicon (5) or be trapped (6). At the SiO_2-Si interface radiation can induce deep interface traps (7).

The hole transport phenomenon takes place over many decades of time, and can last up to several seconds at room temperature. The basic concept used to build up a mathematical description of the mechanism is the *small polaron hopping*, and the mathematical model based on it is called Continuous-Time Random Walk (CTRW). This model was developed to describe the dispersive transport phenomena in a disordered solid, and agrees perfectly with the polaron hopping transport mechanism. In the polaron hopping model, as the positively charged hole moves through the oxide, it causes a distortion of the localised potential field of the SiO_2 lattice. The effect of this distortion is to increase the trap depth at the localized site,

[4] This example refers to the case of a positively biased gate.

[5] The electron velocity in SiO_2 saturates at about $v_{sat} = 10^7$ cm/s for electric fields higher than 2 MeV/cm; if we suppose an oxide thickness of 10 nm and a velocity $v_{sat,}$ the transit time is 0.1 ps.

[6] Electrons and holes can have mobilities in the oxide that differ form five to twelve orders of magnitude.

which tends to confine the hole. As the hole moves through the oxide, the polaron distortion travels with the hole, slowing down the transport phenomenon. [Mcl76, Mcl77].

When the holes reach the vicinity of the SiO_2-Si interface, three processes can take place: charge trapping in the oxide, charge trapping in the interface states, direct tunnelling.

5.1.2.1 Oxide trapped charge

Some holes, migrating in the oxide from the generation point to the SiO_2-Si interface, may be trapped close to the interface, giving origin to a fixed positive charge in the oxide (Figure 5.2 (6)). In effect, at the SiO_2-Si interface there are a large number of oxygen vacancies due to the out-diffusion of oxygen from the oxide and the large lattice mismatch between SiO_2 and Si. These vacancies can act as deep-hole traps.

The neutralization of these traps can happen both for tunnelling of electrons from the silicon and for thermal emission of electrons from the oxide valence band. The traps closest to the interface are neutralized first by electron tunnelling, so that the phenomenon can be thought as a "tunnelling front" moving logarithmically with time in the SiO_2. If a hole is not within some nanometers from the interface, it is practically no longer within the reach of the front.

The thermal emission of electrons is thermally activated, and the traps with energy closest to that of valence band are neutralised first, so the process can be thought as a thermal emission front. This mechanism is field dependent (the electric field lowers the energy barrier height between the trap energy and the oxide valence band) and independent of the spatial distribution of the traps.

The oxide trapped charge generates a negative voltage shift of the MOS threshold voltage which is not sensitive to the silicon surface potential and which can stay for a period of time varying from milliseconds to years. As the amount of trapped holes is proportional to the number of defects present in the oxide, one of the fundamental steps for the fabrication of radiation hardened technologies is the control of the gate oxide quality.

5.1.2.2 Interface states

Another effect of radiation on MOS devices is the increase by several orders of magnitude of the trap density at the interface SiO_2-Si (Figure 5.2 (7)). Interface traps are electronic levels located at the SiO_2-Si interface that can capture or emit electrons or holes. These electronic levels arise because of the lattice mismatch at the interface, disconnected chemical bonds or impurities.

The phenomenon [Ma89, Win89] has been studied for many years, and several different models have been developed to explain the dependencies on several parameters, such as electric field, time, temperature and total dose. The radiation induced traps at the SiO_2-Si interface increase the absolute value of the threshold voltage both for n- and for p-channel MOS transistors.

5.1.2.3 Direct tunnelling

When a hole reaches the vicinity of the SiO_2-Si interface without being localised in a deep trap, it can be neutralised due to tunnelling of the electrons from the silicon (Figure 5.2 (5)) and be collected as substrate current at the silicon body contact. The process is weakly temperature dependent and results in a decrease in the net positive charge remaining in the oxide. The distance from the interface over which tunnelling can occur is of some nanometers.

5.2 Radiation effects on the electrical parameters of a MOS device

The most important effects of irradiation on the electrical parameters of a MOS device are the threshold voltage shift, the increase of the subthreshold and parasitic currents, and the decrease of the mobility and transconductance.

5.2.1 Threshold voltage shift

In a MOS, the total radiation induced threshold shift ΔV_T can be split in two components: one is related to holes trapped in the oxide (ΔV_{ox}) and the other is related with the charge state of the traps at the oxide-silicon interface (ΔV_{it}); so that $\Delta V_T = \Delta V_{ox} + \Delta V_{it}$.

5.2.1.1 Oxide trapped charge

The threshold shift in the case of charge trapped in the silicon oxide can be calculated with the following formula [Mul86, p.401]:

$$\Delta V_{ox} = -\frac{1}{C_{ox}} \cdot \int_0^{t_{ox}} \frac{x}{t_{ox}} \rho(x) dx \qquad (5.1)$$

where C_{ox} is the oxide capacitance per unit area, t_{ox} is the oxide thickness, and $\rho(x)$ is the charge distribution in the oxide per unit volume as a function of the distance from the gate oxide/gate polysilicon interface x. As ρ is the charge distribution due to the holes trapped in the oxide, it will always give a positive contribution to the integral in equation (5.1), so that ΔV_{ox} has always a negative value. This means that this contribution will always increase threshold voltage absolute value for a p-channel MOS, and decrease it for an n-channel MOS. From equation (5.1) it is also evident that, due to the fact that $\rho(x)$ is weighted by x, charges trapped close to the SiO_2-Si interface contribute more to the threshold shift. This means that a positive gate bias, which helps in pushing holes towards the SiO_2-Si interface, will give a larger ΔV_{ox} than a negative gate bias.

5.2.1.2 Interface states

The case of a threshold shift due to the charges at the oxide-silicon interface can be treated as before. Since the charge distribution is highly localised at $x = t_{ox}$, it can be expressed as:

$$\Delta V_{it} = -\frac{\Delta Q_{it}}{C_{ox}} \tag{5.2}$$

where ΔQ_{it} is the difference of charge per unit area which fills the interface states before and after irradiation. Unlike ΔV_{ox} , ΔV_{it} can have positive or negative values because the traps can act both as donors or as acceptors. This results in a positive threshold shift for an n-channel transistor and in a negative shift for a p-channel transistor (i.e. always in an increase of the absolute value of the threshold). Also for interface states generated threshold voltage shifts the worst case bias condition for maximising ΔV_{it} is to have the gate positively biased.

5.2.1.3 Overall threshold voltage shift

The overall threshold voltage shift $\Delta V_T = \Delta V_{ox} + \Delta V_{it}$ will then always increase the threshold voltage absolute value for a p-channel transistor, while it can be both positive or negative for n-channel transistors. In old technologies, it is in general negative for n-channel transistors. On the contrary, for newer technologies with thin gate oxide (roughly <10 nm) it is positive for n-channel transistors, because there is a smaller volume where oxide holes can be generated, and because the annealing of the trapped holes becomes more probable[7].

It is very important to point out another major difference between the two phenomena which generate threshold voltage shifts: the interface states increase is slower than the build-up of positive charge, so that ΔV_{it} starts to play a role during irradiation later than ΔV_{ox}. This difference in the timing of the two phenomena can also explain the *rebound* effect. Since in an n-channel transistor the two contributions are opposite, the threshold shift as a function of the annealing time can be negative at the beginning, when the component ΔV_{ox} dominates. As ΔV_{it} starts to play a role the threshold shift can reduce in absolute value, and eventually become positive. This is why annealing has to be carefully carried out in transistor and system radiation hardness testing.

5.2.2 Subthreshold slope

We define *subthreshold slope n* as the slope of the curve log I_d versus V_{GS} in the region where the transistor is in weak inversion (in weak inversion the drain current is an exponential function of V_{GS}). Its inverse is called *subthreshold swing*.

[7] One of the phenomena which reduce the amount of holes trapped in the silicon is the *tunnel annealing*, where electrons can tunnel from the surface to the oxide, neutralizing some of the traps in the oxide. The probability of an electron tunneling is exponentially related to the hole-surface distance, and then to the oxide thickness.

The subthreshold slope[8] is not affected by the presence of oxide-trapped charge in the device, but decreases with the introduction of interface trapped charge. This allows also the separation of the effects of irradiation on MOS devices, as from the change in the subthreshold slope we can find ΔQ_{it} (and ΔV_{it}) and from ΔV_T we can then extract ΔV_{ox} (and ΔQ_{ox}).

5.2.3 Off-state and parasitic currents

The "off-state" current[9] in a MOS transistor can be affected by irradiation in three ways: the decrease of the subthreshold slope, the threshold voltage shift and the generation of parasitic currents (edge leakage and inter-transistor leakage).

5.2.3.1 Effect of the subthreshold slope change and threshold voltage shift

As we have seen, irradiation can induce a threshold voltage shift. This, in turn, can affect the off-state current as shown in Figure 5.3 (a). The figure plots the pre-irradiation curve $\log I_d$ versus V_{gs} for an n-channel MOS in which we suppose there is a decrease of the threshold value due to irradiation.

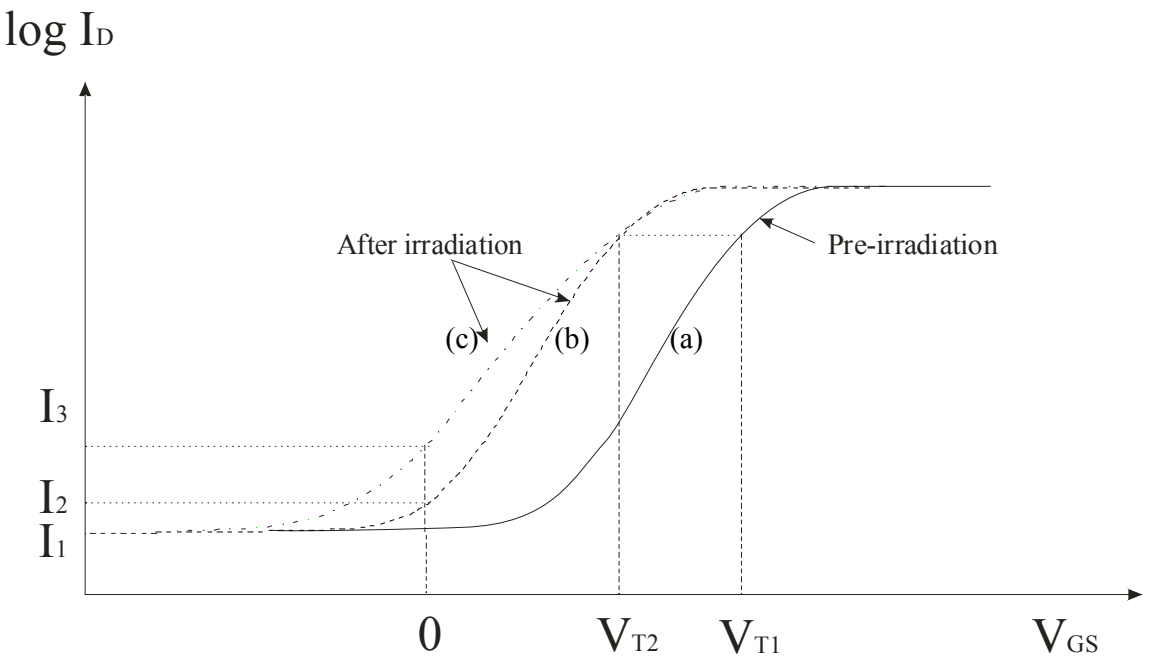

Figure 5.3: Representation of two main effects of radiation on the electrical parameters of a MOS transistor: the pre-irradiation curve (a) shifts in (b) due to the threshold voltage shift; this causes the threshold voltage V_{T1} to decreases to V_{T2} and the off-state current I_1 to increases to I_2. If the subthreshold slope decreases (c), the off-state current I_2 increases to I_3.

[8] The precise equation for n can be found for example in [Enz95]. For small signal, $n = g_{ms} / g_m = 1 + C_{bs} / C_{gs}$, where g_{ms} is the source transconductance, g_m the gate transconductance C_{bs} the bulk to source capacitance and C_{gs} the gate to source capacitance.

[9] The "off-state" current is defined as the current which flows in a MOS device from drain to source (I_{DS}) when the gate to source voltage (V_{GS}) equals zero.

The pre-irradiation curve shifts left (Figure 5.3 (b)), and the subthreshold current increases from I1 to I2. The second effect can be seen in Figure 5.3 (c): due to the irradiation, the subthreshold slope decreases, and the off-state current increases, going from I_2 to I_3. We have supposed that the threshold shift is negative, but we have seen that n-channel transistors can have an overall positive threshold shift after annealing; in this case the decrease of the threshold voltage can help to compensate the decrease in the subthreshold slope.

5.2.3.2 Edge leakage currents

Even though the radiation hardness of commercial gate oxides may improve as the IC industry tends towards ultra-thin oxides, field oxides (i.e. oxides used for device isolation) of advanced commercial technologies will still be relatively thick, in the range of 100 nm to 1000 nm, and may still be very soft to ionizing radiation. A relatively small dose in a field oxide (~10 krd(Si) for many commercial devices) can induce sufficient charge trapping to cause field-oxide induced IC failure. Unlike gate oxides, which are routinely grown by thermal oxidation, field oxides are produced using a wide variety of deposition techniques. Thus, the trapping properties of a field oxide may be poorly controlled and can be considerably different than for a gate oxide. Even for thermally grown thick oxides, the buildup of charge in gate and field oxides can be qualitatively different [Sch02, Boe85].

In technologies down to 0.35 μm, the device isolation technique employed is the local oxidation of silicon (LOCOS), while in newer technologies this has been replaced by shallow-trench isolation (STI) that, with respect to LOCOS, allows higher transistor densities. In technologies employing STI, to isolate transistors a "trench" is etched in the silicon substrate, and then it is filled up with thick isolation oxide, called field oxide (FOX). A schematic representation is shown in Figure 5.4.

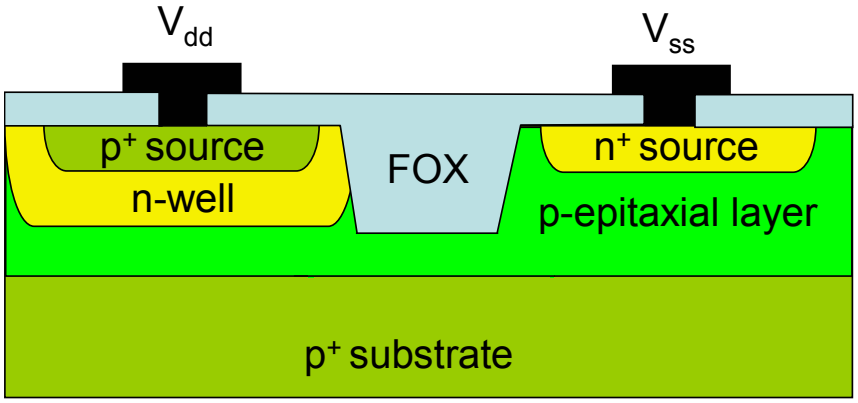

Figure 5.4: Schematic representation of the inter-transistor isolation for technologies employing STI. A "trench" is etched in the silicon substrate, and then it is filled up with isolation field oxide.

In technologies employing LOCOS, a thick field oxide is grown at the end of the gate to precisely define the gate area. The transition region between the gate oxide and the field oxide is called “bird’s beak” (Figure 5.5).

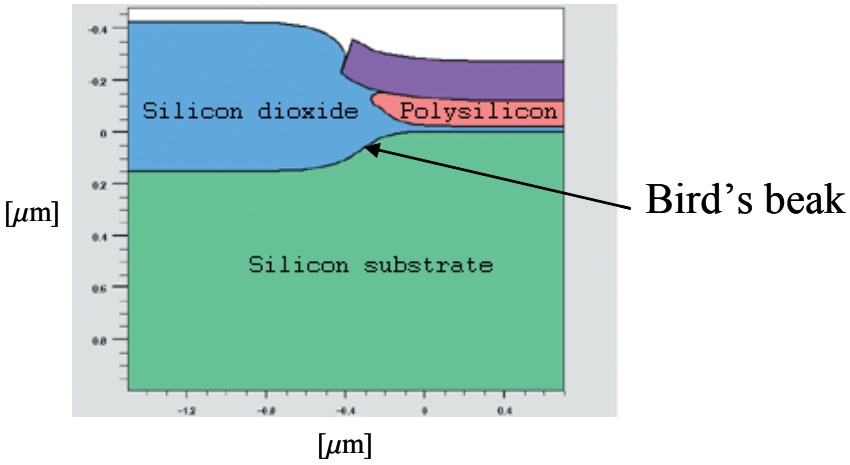

Figure 5.5: Schematic representation of a cross section of a MOS transistor. A parasitic transistor is present in the bird’s beak region, under the thick field oxide.

This region is particularly prone to radiation damage, both for the thicker oxide and because the silicon dioxide in the bird’s beak region is under mechanical stress produced by the dynamics of the oxide growth process and the transition from thin to thick oxide. The transition region oxide is of variable thickness and experiences a relatively high electric field from the combination of poly gate bias and the fringing fields from the source to drain bias.

This phenomenon generates in standard n-channel transistors parasitic paths from drain to source (not under the gate oxide), which can increase the off-state current. Figure 5.6 shows a schematic representation of an n-channel MOS transistor with possible radiation induced leakage paths. A similar phenomenon happens also for technologies with STI.

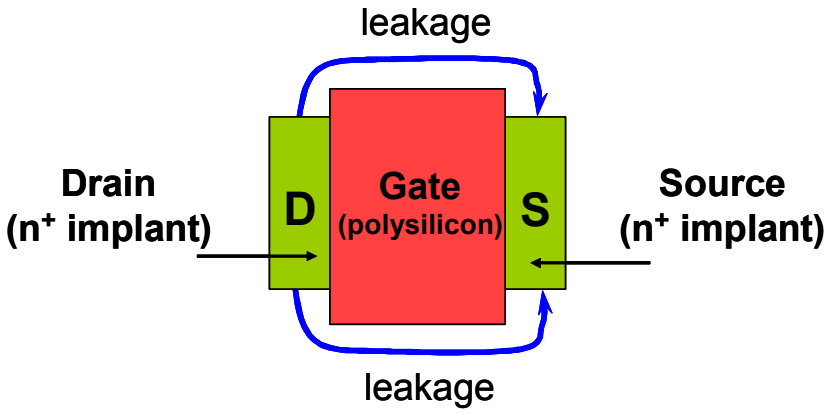

Figure 5.6: Schematic representation of an n-channel MOS transistor with possible radiation induced leakage paths.

The edge leakage can be qualitatively represented as a combination of several "elementary" parasitic transistors in parallel with the main one, each one having a width ΔW and about the same length, but with an increasing gate oxide thickness [Bri96]. The gate capacitance of the transistors with higher oxide thickness is lower, so the threshold voltage is higher[10]. Without irradiation, these transistors can have a threshold voltage of several volts, so they are normally off. During irradiation, their thick oxide can trap a large amount of charge, and their threshold voltage can shift also by several volts, thus switching them on[11]. The final shape after irradiation of the characteristic curve I_d versus V_{gs} of the transistor strongly depends on the total dose absorbed and on the quality and type of the field oxide. Edge leakage can result in signal corruption, reduced noise margins and additional power supply currents (with the associated increase in the voltage drops along the lines).

The contributions to the off-state current coming from the parasitic edge currents and from the subthreshold currents have been measured for several technologies, with oxide thickness ranging from 0.7 μm to 0.25 μm[12] [Ane97]. For all of them the contribution due to the parasitic currents dominated over the contribution due to the subthreshold current.

An example of the effect of irradiation on traditionally laid-out MOS transistors in shown in Figure 5.7.

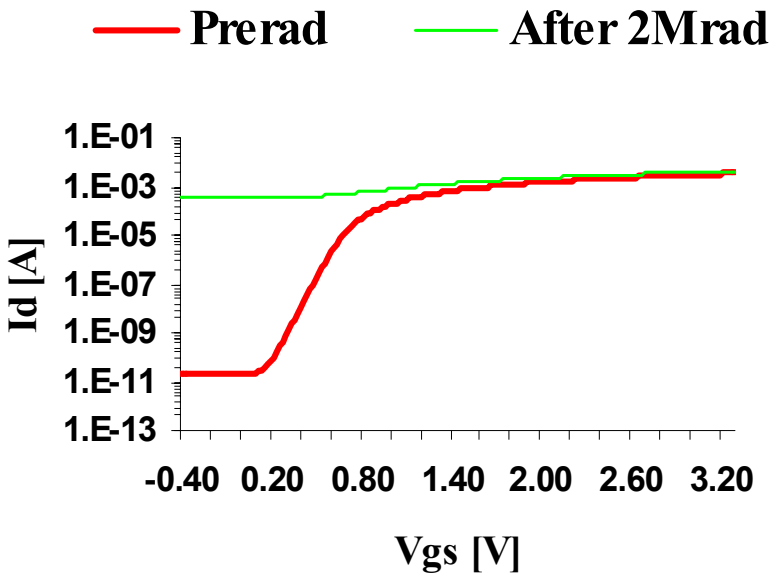

Figure 5.7: Log I_d versus gate voltage before (thick line) and after 2 Mrd irradiation (thin line) for a traditionally laid out n-channel device with a W/L of 10/0.5 μm. The curve after irradiation shows an unacceptable leakage current. The leakage was measured to be unacceptable already at 40 krd [Sno00].

[10] The oxide capacitance per unit area is given by $C_{ox} = \epsilon_{ox} / t_{ox}$ and the threshold voltage increases with $1/C_{ox}$.

[11] Moreover the transistors with a higher threshold voltage (lower C_{ox}) will also have the higher threshold shift (thicker oxide).

[12] The transistors measured had minimum gate length.

An n-channel device with a W/L of 10/0.5 μm is irradiated with X-rays up to a TID of 2 Mrd. The pre-irradiation curve is the thick one, and shows normal behaviour; the curve after 2 Mrd shows on the contrary an unacceptable leakage current. The leakage was measured to be unacceptable already at 40 krd [Sno00].

This kind of leakage is not a problem for p-channel transistors, since the effect of oxide trapped charge and interface trapped charge is to increase the threshold voltage of the parasitic transistor (in absolute value). This has been verified experimentally, and an example is reported by Lacoe and coworkers [Lac99]. A p-channel transistor fabricated at Chartered Semiconductors in a 0.35 μm CMOS technology has been irradiated under worst case bias condition up to 300 krd (Si) without any visible change of the transistor I_D/V_G characteristic.

5.2.3.3 Inter-transistor leakage

The other problem that needs to be solved to use a commercial CMOS technology in a radiative environment is the creation of inter-transistor parasitic currents. The problem is schematically represented for an n-well CMOS technology in Figure 5.8. In this structure we can recognise a parasitic transistor which has its drain on the n-well of a p-channel MOS (connected at the positive power supply), the source at the source of an n-channel transistor (connected at the negative supply), and the gate at the inter-transistor field oxide. Due to the irradiation, the threshold of this parasitic transistor decreases and the device can eventually turn on (generating the parasitic currents). If the radiation dose is high enough to invert the silicon, the parasitic transistor is always on and a permanent low resistance path is created, but if metal or polysilicon lines run on the field oxide their electric field can switch on the parasitic transistor even before its threshold voltage reaches zero. There is also another leakage path between the n^+ source/drain regions of adjacent n-channel transistors, and between n-wells of p-channel transistors, if they are biased at different potentials.

The use of Polysilicon Buffered LOCOS (PBL) or STI instead of LOCOS does not solve the problem [Sha98].

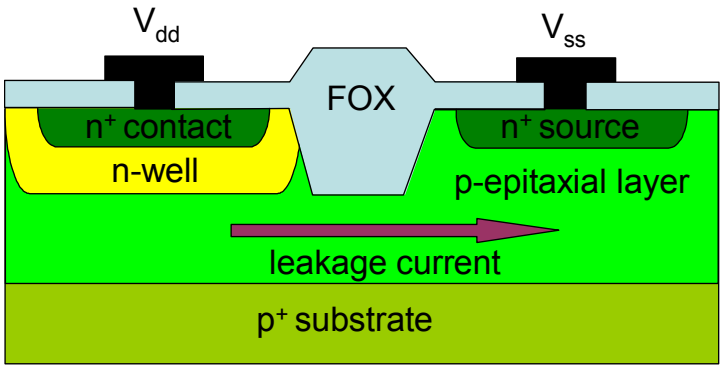

Figure 5.8: Schematic representation of the radiation induced inter-transistor leakage current for an n-well CMOS technology.

Special field-oxide characterization test vehicles have been designed, fabricated and tested as a function of total dose exposure to evaluate the radiation hardness of commercial CMOS isolation oxides. Interesting results derived from the measurements of large area capacitors fabricated using the LOCOS isolation under worst case bias conditions for the Chartered Semiconductor 0.35 μm process are presented in [Lac99]. Before irradiation the threshold voltage is about 30 V, while at 100 krd (Si) it is approaching 0 V. Special field-oxide transistors were fabricated also in the TSMC 0.25 μm process [Lac00] with a polysilicon gate over STI field oxide between adjacent n-wells. The pre-irradiation value of the FOX transistor threshold is ~42 V and the threshold voltage after 100 krd (Si) (in worst-case bias conditions) is reduced down to 0.5 V.

5.2.4 Mobility and transconductance

The conduction in a MOS transistor is due to the carrier motion very close to the SiO_2-Si interface. As a consequence the irradiation induced mobility degradation is not much affected by the bulk damage, but by the increase of the interface traps. An empirical simplified formula has been proposed by Sexton and Schwank [Sex85] to express the mobility as a function of the irradiation induced surface traps:

$$\mu = \frac{\mu_0}{1+\alpha \cdot \Delta N_{it}} \tag{5.3}$$

where μ_0 is the pre-irradiation mobility, ΔN_{it} is the radiation-induced increase of the interface traps, and α is a technology-dependent parameter. This formula neglects the dependence on the oxide-trapped charge, but after some seconds the majority of oxide-trapped charge near the interface has been already neutralised by tunnelling and the remaining charge has a negligible scattering cross-section for carriers in the channel. The transconductance of a MOS transistor decreases with μ both in linear region and in saturation, so that the mobility degradation reduces also the device transconductance.

5.3 Single Event Effects (SEE)

Single Events Effects are generated by highly energetic particles crossing a sensitive part of a transistor or of an integrated circuit (see for example [Mus01]). They are usually divided in *soft errors* (which are reversible and non-destructive, as they do not cause a permanent damage to the device) and *hard errors* (which are non-reversible[13] and can be destructive under certain conditions).

[13] A non-reversible event interrupts the chip functionality; if the event was non-destructive the chip functionality can still be recovered (usually with a power-off power-on cycle).

5.3.1 Soft errors

Soft errors can be divided in: Single Event Upsets, Single Event Transients, Multiple Bit Upsets and Single Event Function Interrupts.

5.3.1.1 Single Event Upset (SEU)

If a highly energetic particle passes through a memory cell, it can induce (directly or by secondary interactions) an instantaneous and reversible change of the logic state of the cell. This is due to the charge generated along the particle track in the circuit, which can upset the memory cell sensitive node. This phenomenon is called Single Event Upset (SEU) [Die82, Roc99]. SEUs are reversible, so if the system is able to detect them, and if the information has been duplicated, they can be recovered. In memories the concerned information is rewritten, in CPUs the algorithm being executed has to be restarted. For each sensitive node there is a minimum charge (the *critical charge*) which can generate an SEU. Since the charge generated in a material is proportional to the LET of the incoming particle (see footnote 3), the critical charge translates into a critical LET: if the incoming particle has a LET which is higher than the critical LET, the node is vulnerable to SEU. Although strictly related, critical charge is usually determined by computer simulations, while critical LET experimentally, bombarding an integrated circuit with highly energetic particles.

5.3.1.2 Multiple Bit Upset (MBU)

In this case more than a single bit can be upset at the same time. If a particle impacts on the integrated circuit in a direction which is almost parallel to the surface, it can cross the sensitive volume of several devices and upset more than one sensitive node. Also if the particle hits the IC perpendicularly to the surface, but is able to deposit a sufficient amount of energy, an MBU can be generated. Another possibility is that two different particles upset two adjacent nodes at the same moment[14]. This kind of error is more difficult to detect and recover than an SEU, so since it is usually also much more unlikely to happen, circuits are often SEU protected but not MBU protected.

5.3.1.3 Single Event Function interruption (SEFI)

A SEFI can be considered as an SEU that is generated in a particular sub-system of a complex circuit. In complex memories, for example, the memory cells and the periphery circuits are controlled by the Error Detection and Correction (EDAC) circuitry. If one of the sensitive nodes of the EDAC is upset, the error can influence the functionality of the whole circuit.

[14] This usually means, for digital circuits, in the same clock cycle.

5.3.1.4 Single Event Transient (SET)

Single Event Transients refer to errors that can results from an energetic particle strike on non-latched elements, such as combinatorial logic, clock lines and global control lines. The incident particle charges or discharges the output load of a combinatorial logic element, inducing a voltage transient. The voltage transient can propagate through the combinatorial logic and reach a latching element. If the signal is in coincidence with the clock edge that latches data into the register, a wrong state can be induced in the register itself.

The SET propagation distance through the circuit depends on the speed of the circuit and on the width of the transient voltage spike. The *critical width* is the minimum width of the transient pulse required for the SET to propagate through an infinitely long chain of inverters. If the transient width is narrower than the critical width, the transistors do not have the speed to respond to the transient, and it will be attenuated and die out after passing through only a few gates.

Since the probability of a SET being captured depends on the number of falling edges arriving to the latch per unit time (i.e. the clock frequency), it is expected that the SET upset rate will depend linearly on the clock frequency.

The total circuit upset rate, the Soft Error Rate (SER), is the combination of SEU and SET rates.

5.3.2 Hard errors

Hard errors can be divided in: Single Event Latchups, Single Event Snapbacks, Single Hard Errors, Single Event Gate Ruptures and Single Event Burn Outs.

5.3.2.1 Single Event Latchup (SEL)

In some planar CMOS technologies there is a parasitic thyristor which can be switched on by electrical transients, high temperature and improper sequencing of power supply biases, but also by a highly energetic ionising particle. This phenomenon is called Single Event Latchup (SEL) [Joh96][15]. A thyristor is a PNPN device [Sze81, p. 190] that exhibits a bistable characteristic and can be switched between a high-impedance, low current OFF state and a low-impedance high-current ON state. In a MOS device it is composed of two parasitic bipolar transistors in positive feedback, as shown in Figure 5.9. If they are forward biased, their gains are high enough and the power supply is able to deliver a current higher than the characteristic hold current of the device I_H at a voltage higher than the characteristic hold voltage V_H, the thyristor can switch on and short the power supply lines.

[15] SOI technologies are latch-up free because the parasitic thyristor is not present.

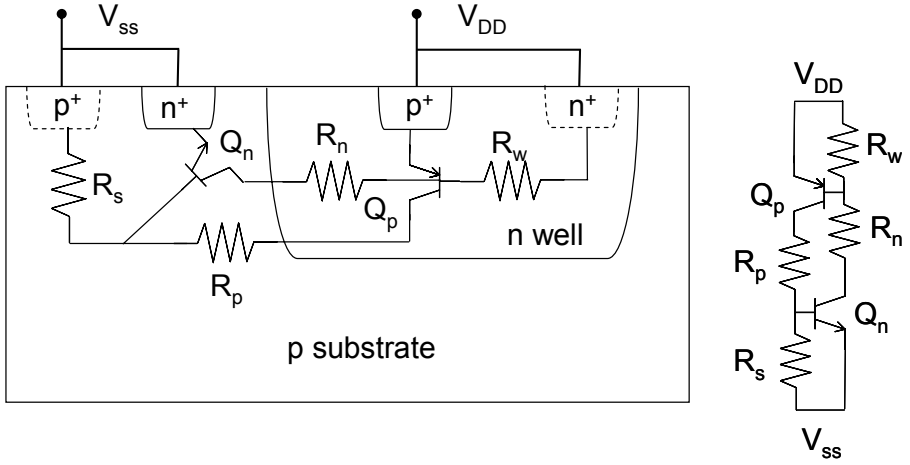

Figure 5.9: Schematic representation of the parasitic thyristor in an n-well CMOS technology. If the thyristor switches on it can short the power supply lines.

A single heavy ion or high energy proton passing through either the base emitter junction of the parasitic n-p-n transistor, or the emitter-base junction of the p-n-p transistor can initiate the regenerative action. If it is not interrupted, the high current flow generated by the SEL can damage the integrated circuit. Also for SEL a critical (or threshold) LET can be defined: if the incoming particle has a LET which is higher than the critical LET, the node is vulnerable to SEL.

This is the most likely hard error which can happen in modern planar submicron technologies.

5.3.2.2 Single Event Snapback (SES)

Single Event Snapback, like SEL, is also a regenerative current mechanism, but does not require an N-P-N-P structure [Bei88]. It can be triggered in n-channel MOS transistors with large currents, such as in IC output driver devices, by a single event hit-induced avalanche multiplication near the drain junction of the device. In this case the hit-induced charge acts as a base current of the NPN bipolar transistor which has as emitter, base and collector respectively the source, substrate and drain of the MOS device. The parasitic bipolar transistor in turn induces an electron flow from source (emitter) to drain (collector), reinforcing the avalanche mechanism and closing the feedback loop. This phenomenon was present in old technologies; in modern sub-micrometric technologies it is not present anymore thanks to the lower power supplies and to the fact that a SEL is easier to induce than a SES.

5.3.2.3 Single Hard Error (SHE)

A highly energetic particle can deposit locally in the gate oxide of a MOS transistor a dose sufficiently high to induce a threshold voltage shift large enough to make the transistor unusable. This phenomenon, called SHE [Duf92], is present especially in high integration submicrometric memories (> 256 kbits), where it can generate “stuck bits”: in an SRAM cell,

the n-channel transistor subthreshold current can become high enough to block the cell in the high or low logic state. Newer technologies are much less sensitive to this problem thanks to the reduced gate oxide thickness.

5.3.2.4 Single Event Gate Rupture (SEGR)

A Single Event Gate Rupture is the destruction of the gate oxide of a MOS device due to a highly energetic particle. When the particle passes through the gate oxide, it can form a highly conductive plasma filament between the bulk and the oxide. If the energy is high enough, this can generate a localised heating of the oxide that can locally melt the oxide, all along the particle track.

This problem is usually strictly bound with high electric fields in the gate oxide, as in power MOSFETs or in non-volatile memories, such as EEPROMs[16] during write or erase operations, when a voltage much higher than the normal operation power supply is applied to the gate [All95]. Relations between submicron technologies and SEGR are discussed in section 5.5.

5.3.2.5 Single Event Burn Out (SEBO)

N-channel power MOSFET devices, which have large applied biases and high internal electric fields, are susceptible to single event burnout (SEB or SEBO) [Hoh87]. The penetration of the source-body-drain region by the charge deposited by a heavy ion can forward bias the thin body region under the source. If the terminal bias applied to the drain exceeds the local breakdown voltage of the parasitic bipolar transistor, the single event induced pulse can initiate avalanching in the drain depletion region. Local power dissipation due to the large drain-source current leads to destructive burnout. A similar effect has been seen in power bipolar devices.

5.4 Design for radiation tolerance

To increase the radiation tolerance of a circuit, it is possible to operate with two different approaches: at foundry (process) level, and at design level. The best results can be achieved facing the problem of radiation at both levels, if possible. A third possibility to obtain radiation tolerance at chip level consists in selecting adequate Components Off The Shelf.

5.4.1 Hardening by process

This consists of addressing the problem of radiation tolerance at the level of the technological process. It is possible to modify some steps of a normal technology to improve radiation tolerance, or to conceive a technology in a way that is radiation hardness oriented.

[16] Electrically Erasable Programmable Read Only Memories

The most radiation sensitive part of a MOS transistor is the oxide, both the gate oxide and the isolation oxide ("field oxide"). This means that in radiation tolerant technologies the control of the oxide quality is of crucial importance. As an example, nitrided oxides or reoxidised nitrided oxides improve gate oxide total dose radiation tolerance.

The replacement of LOCOS in newer technologies by STI significantly shrinks the area needed to isolate transistors while offering superior latch-up immunity, smaller channel width encroachment and better planarity. At 180 nm, STI processes replace LOCOS isolation in virtually all devices. However, this new isolation technique does not eliminate the problem of radiation induced leakage currents [Sha98]. Several solutions have been proposed (examples are the double layer field oxide [Neu94] or growing a thin layer of oxynitride as a membrane for controlled diffusion of O_2 and oxidation of Si at high temperatures with low thermal budget [Man02]), but all of them make the process more complex (and then more expensive) and lower the yield and the component density.

For what concerns SEE tolerance, the main actions that can be taken to improve it are: reduction of the epitaxial layer thickness (SEL and SEU), increase of the substrate doping (SEL and SEU), reduction of the minority carriers life time (SEL), tailoring of the doping density and of the doping profile of the wells and increase of the minimum allowed distance between p^+ and n^+ diffusions.

SOI/SOS technologies are inherently more tolerant to SEL and SEU than CMOS technologies, as the parasitic thyristor is not present (no SEL possible), sensitivity to SEU is reduced (there is no epitaxial layer) and they do not have radiation-induced leakage between devices.

Some specialised vendors (e.g. Atmel, with DMILL technology) provide a qualified radiation hard process. These technologies were developed mainly for space and military applications. This approach is expensive, as the volume required by the market is very small compared to standard commercial processes. Moreover, due to the high costs needed to keep a technology up-to-date, radiation hardened technologies are also less advanced (e.g. with bigger and slower devices) as more development is required. In some cases problems related with low yield and wide parameter variation from wafer to wafer and from lot to lot were found. A major problem related with radiation-hardened processes is their future availability: the small market tends to push semiconductor vendors to more high volume commercial applications. Due to the drop of demand for this kind of electronics, some radiation hardened technologies processes will be discontinued, as is the case for example with DMILL.

5.4.2 Hardening by design (HBD)

The basic principle of hardening by design is to use a readily available, low cost CMOS technology and to apply some special layout and design techniques to increase radiation tolerance. The main advantage of this approach is of course the low cost, and the possibility to

follow the rapid scaling down of commercial technologies driven by a market that is much larger than the HEP market can be of major importance, especially where a high density-high speed circuitry is needed, as for example in pixel detectors.

The definition of HBD proposed by R. Lacoe in [Lac03] is:

"*HBD is an approach to producing radiation-hardened components and systems using innovative design and layout techniques at the transistor level to assure performance and radiation-hardness requirements are met. The fabrication of HBD components is at commercial microelectronic foundries using standard commercial processes and process flow*".

HBD does not allow any modification to the standard commercial process flow, and has to be non-invasive to the foundry, to not depend on the willingness of the foundry to change its process or process flow. This HBD approach has been used for the design of the ALICE1LHCb chip, so it will be presented in more detail in section 5.6; experimental results of systems designed completely with these HBD techniques are presented in section 5.7.

5.4.3 Components Off The Shelf

This approach consists in selecting commercial components (components off the shelf, COTS) and test them, setting up a database of selected standard electronic components which have been characterised and qualified as robust against radiation effects. This approach is applicable especially for components that do not require a high radiation tolerance. The advantages of this approach include a large variety of available components, high performance and low component costs. The disadvantages include short availability lifecycles, the possibility of process changes that can change the radiation tolerance characteristics, lack of traceability, costs associated with qualification and the inability to get reliable data from the manufacturer. Moreover, for some applications an available COTS device with the required radiation tolerance could not be available (this is, for example, the case for the front-end electronics for the inner trackers of the LHC experiments).

The present trend in commercial technologies goes towards improved components resistance to total ionising dose and single event latch-up. COTS also take profit of the technology scaling that makes components intrinsically more tolerant to TID and SEL. For example, Howard [How01] showed that the commercial Pentium III (biased under irradiation) can stand a TID higher than 400 krd(Si).

For what concerns SEUs, technology scaling has lead to an increase of the normal operation SEU rate to a point where for high reliability microelectronic applications (e.g. server systems) SEU mitigation has proven necessary. This has raised the awareness of the microelectronics industry interest to this problem and increased the chances to find, in the near future, more and more COTS devices which can stand much higher irradiation levels.

5.5 Deep submicron CMOS technologies for HEP

We have seen that process hardened technologies advance much slower than commercial technologies, and cannot profit from all the beneficial effects of the technology scaling. In the next sections we will briefly present the most important advantages (and drawbacks) of technology scaling, both for what concerns radiation induced effects and circuit performance.

5.5.1 Total ionising dose

To move from one generation to the next one, with the aim of reducing the device dimensions (thus increasing its performance) by a scaling factor α without introducing major disturbing side effects, several different approaches can be used. In all of them the gate oxide thickness t_{ox} is scaled down with α[17].

The beneficial effects of the oxide thickness reduction with the consequent reduction of the radiation induced charges trapped in the oxide and interface states were studied and described by Saks and coworkers in [Sak84, Sak86], and are reported in the next sections. This means that at each technology step, devices became more intrinsically radiation tolerant.

5.5.1.1 Oxide trapped charge

Saks and coworkers measured the radiation induced flat-band voltage shift at 80 K in MOS capacitors with oxide thickness ranging from 6.0 nm to 50 nm [Sak84]. The interest in studying this effect at low temperature is the following. At low temperature nearly all holes created by the ionising radiation in an oxide are trapped, which leads to a very large (thus easily measurable) flat-band voltage shift. In addition, the formation of radiation induced interface states is completely suppressed and time dependent annealing is much less severe than at room temperature.

Results are reported in Figure 5.10, which plots the variation of the flat-band voltage (normalised to the total dose) as a function of the gate oxide thickness.

The figure shows how, down to oxide thicknesses of about 20 nm, the flat-band voltage shift is proportional to t_{ox}^2, as measured by Boesch and McGarrity [Boe76] (who also explained it mathematically considering the charge build-up and removal process in the oxide [Boe86]). For oxide thicknesses going from about 20 nm to about 10 nm this law starts to lose validity, approaching a much faster dependence for $t_{ox} < 12$nm. This change in the flat-band voltage dependency has been explained by tunnelling of electrons (see footnote 7) [Ben85].

[17] Only in the "Constant voltage scaling" t_{ox} is scaled less drastically than linearly by a scaling factor α' ($1 < \alpha' < \alpha$), to avoid an excessively high field in the gate oxide.

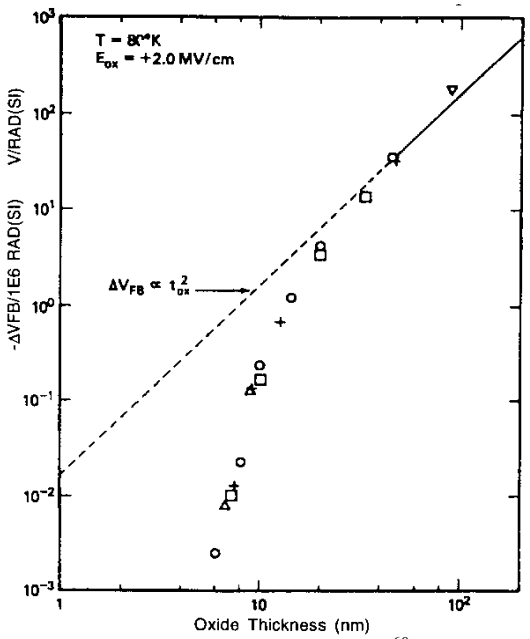

Figure 5.10: Flat-band voltage shift ΔV_{fb} per Mrd dose (Co^{60}) for MOS capacitors at 80 K as a function of the oxide thickness. The dashed line is the Boesch prediction of $\Delta V_{fb} \propto t_{ox}^2$. It can be seen that for thinner oxides ΔV_{fb} decreases even faster [Sak84].

5.5.1.2 Interface states

The other mechanism (shown in Figure 5.11) which can cause a threshold voltage shift (ΔV_{it}) is the creation of traps at the SiO_2-Si interface (ΔD_{it}), where $\Delta V_{it} \propto \Delta D_{it}$.

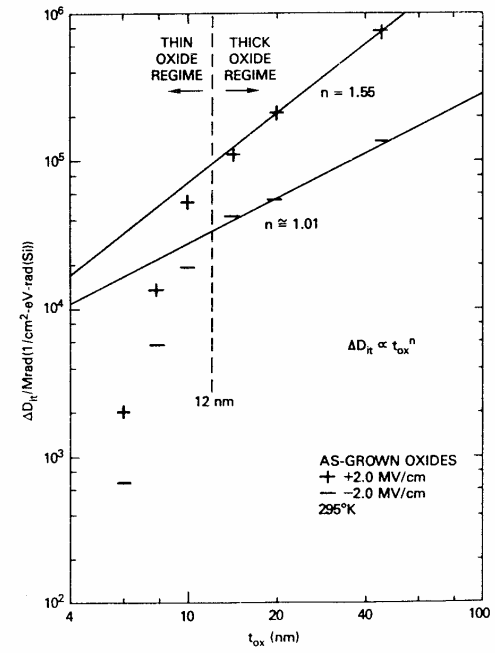

Figure 5.11: Interface trap density D_{it} creation rate (ΔD_{it}) as a function of the gate oxide thickness for as-grown oxides at ±2MV/cm, per Mrd dose [Sak86]. An effect similar to Figure 5.10 is present.

Figure 5.11 plots the interface trap density D_{it} creation rate (ΔD_{it}) as a function of the gate oxide thickness for as-grown oxides at ±2MV/cm, per Mrd dose [Sak86]. It can be seen that there is a drastic reduction of ΔD_{it} if the gate oxide thickness is less than 12 nm. The mathematical explanation of this effect is not straightforward, because it depends on the model chosen to explain the interface state build-up.

5.5.1.3 Total threshold voltage shift

The total irradiation threshold voltage shift is $\Delta V_T = \Delta V_{ox} + \Delta V_{it}$ so it will also be strongly influenced by the gate oxide thickness. Since 1996, ICs in standard 0.7 μm, 0.5 μm, 0.35 μm and 0.25 μm CMOS technologies have been designed in the Microelectronics Group at CERN.

Results of measurements done on these technologies are presented in Figure 5.12, which shows the threshold voltage variation ΔV_T per Mrd dose as a function of the oxide thickness of the mentioned technologies. Four more points were added, taken from [Osb98] (1.6 μm, 1.2 μm, 0.8 μm and 0.5 μm technologies). These measurements confirm the drastic impact of the gate oxide thickness on the radiation induced threshold voltage shift.

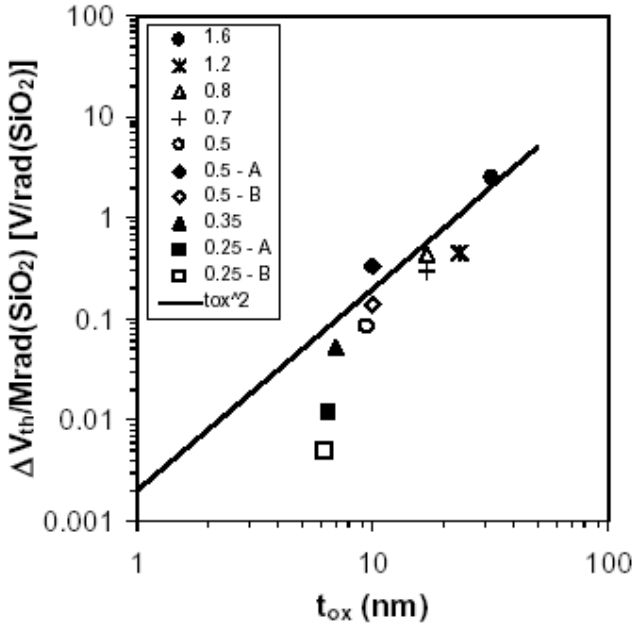

Figure 5.12: Threshold voltage variation ΔV_T per Mrd dose as a function of the oxide thickness [Ane00, p.75]. The points are taken from measurements done in the microelectronic group at CERN [Ane97] (except for four points, whose data are taken from [Osb98]). The legend gives the minimum gate length for the technologies measured (in microns). The solid line shows the $\Delta V_T \propto t_{ox}^2$ trend. The large deviation of ΔV_T from the t_{ox}^2 trend for $t_{ox} < 10$ nm is evident.

5.5.1.4 Off-state currents

As we have seen in section 5.2, radiation can induce edge leakage and inter-transistor leakage. Scaling down to submicron technologies improves radiation tolerance also with respect to these parameters, even if leakage is not eliminated.

Although the use of STI, which replaces LOCOS in deep submicron technologies, does not solve these problem, it is interesting to note that the transition from LOCOS isolation to STI isolation in conjunction with scaling results in processes that show increased hardness to edge effects. Lacoe supposes [Lac03] that the absence of the bird's beak could result in a higher quality oxide. In a typical STI process there is a narrow region of thin isolation oxide before the trench. This makes the continuum of parasitic transistors very thin, reducing the threshold voltage shifts that are responsible for edge leakage. In addition, the STI thickness is likely to be considerably thinner than that of the LOCOS isolation. It should be remarked that enhanced electric fields, with a related degradation of radiation hardness, can be generated at the steep transition gate oxide/trench.

5.5.1.5 Subthreshold slope

The subthreshold swing changes with the formation of interface states (and is insensitive to the oxide trapped charge). For processes with gate oxides thicknesses less than the critical tunnelling length (in the order of some nanometers), the number of interface states created is insignificant, and the change in the subthreshold swing is negligible.

Lacoe presented measurements [Lac01] done on a thin gate process (TSMC 0.18 μm). Edgeless transistors – to avoid edge leakage effects – were irradiated up to a dose of 20 Mrd(Si) without any change of the subthreshold swing. Similar results (presented in section 5.7) were found by Anelli and coworkers for a 0.25 μm technology [Ane99].

5.5.2 Single event effects vulnerability

Scaling down has also an impact on the single event vulnerability, both for soft and for hard errors.

5.5.2.1 Single event upset

SEUs are generated when the charge deposited in the IC by ionising particles exceeds a minimum *critical charge*. Scaling reduces circuits dimensions (and node capacitances) and power supply voltages, so in principle it could worsen the circuit sensitivity to SEU ($Q = C \cdot V$). However, due to the increase of the capacitance per unit area, charges per unit area do not scale, or do not scale much (depending on the scaling method used), so it is possible to keep the same SEU sensitivity (for important nodes, for example) not scaling the transistor dimensions. Moreover, the reduction of the epitaxial layer thickness in advanced technologies

(see section 5.5.2.3) reduces the charge collection efficiency, so that for the same impinging particle the charge generated in the sensitive node is lower.

5.5.2.2 Single event transient

Technology scaling increases the SET vulnerability of a circuit. The critical width is determined by the ability of the circuit to respond to the fast signal induced by the impinging particle. As technology advances and technology feature size shrink, the critical width narrows, so that narrower and narrower transients can propagate through the circuit. In [Mav02] is presented the critical transient width versus feature size for unattenuated propagation. For a 0.25 μm technology the critical width is about 70 ps.

Baze pointed out that while a transient pulse greater or equal to the critical width would propagate infinitely, a pulse of half its width would terminate in the first gate [Baz97].

5.5.2.3 Single event latch-up

An important beneficial effect in submicrometric technologies is the decrease of SEL vulnerability. In effect several modifications have been introduced in the process to help reducing the vulnerability to electrically induced latch-up, and as a consequence to SEL.

The epitaxial substrate is a lightly doped p-layer a few microns thick, grown on a highly doped p^+ substrate[18]. In this way the body of the transistor profits from the low doping and the highly doped substrate reduces the parasitic resistances which contributes to SEL occurrence. Epitaxial substrates limit the charge collection region after a particle strike, therefore limiting the induced current flow, and this results in improved SEL hardness. Thin epitaxial substrates can increase significantly the holding voltage of the parasitic thyristor for equal anode-cathode spacing.

Retrograde wells are wells with a profile specially tailored to have much higher doping near the bottom of the well. This lowers parasitic resistances and, since the depth of the well is smaller than conventionally doped wells, allows thinner epitaxial layers.

STI reduces drastically the gain of the parasitic lateral bipolar transistor structure, thus reducing SEL sensitivity.

Another beneficial effect of scaling comes from the decrease of the power supply voltages that are approaching and will eventually become smaller than the holding voltages.

5.5.2.4 Single event gate rupture

It has been suggested that, given the industry trend toward increasing electric field as oxide thickness and feature size scale down in advanced technologies, SEGR could become more and more important, in particular for special applications (e.g. space applications). The results published in [Sex97] suggest that advanced technologies will be correspondingly more

[18] In some technologies starting from the 0.13 μm technology node the epitaxial substrate has been dismissed. For these technologies STIs and very low voltage supplies should be enough to prevent SELs.

SEGR resistant at a given electric field, and that SEGR may be a significant concern only for devices that operate with gate oxide electric field significantly above 5 MV/cm. If we suppose an oxide thickness of 5.5 nm (which is the oxide thickness in our target technology) this results in critical gate voltage significantly above[19] 2.5V. Results shown by [Joh98] give lower values for the gate rupture critical field but confirm the trend presented by Sexton. More recent results in this direction can also be found in [Mas01], who shows that the critical gate voltage to breakdown should stay well below the maximum allowed power supply voltage for the near future technologies.

5.5.3 Circuit performance

Scaling is also beneficial for circuit performance, since it improves area usage, speed, transconductance, weak inversion characteristics, noise and matching properties of the transistors.

5.5.3.1 Transconductance

The most important parameter of a MOS device is its transconductance, which is related to the device driving power. For a transistor in strong inversion and in saturation [Lak94]:

$$g_m = \sqrt{2 \frac{K}{n} \frac{W}{L} I_{DS}} \tag{5.4}$$

where W and L are the width and the length of the transistor, I_{DS} is the drain current, $K = \mu C_{ox}$ (μ is the mobility of electrons or holes in silicon and C_{ox} is the gate capacitance per unit area) and n is the weak inversion slope factor, which does not change significantly with the scaling. From equation (5.4) we can say that for a transistor of the same size and for a given drain current, g_m increases with the technology scaling (K is inversely proportional to t_{ox}). An example of how K increases with technology scaling is shown in Table 5.2 [Riv99].

L_{min} [μm]	t_{ox} [nm]	K [$\mu A/V^2$]
1.2	24	68
0.8	14	90
0.5	10	134
0.25	5	280

Table 5.2: Oxide thickness and K for an n-channel transistor in different technologies. L_{min} is the length of the minimum transistor that can be drawn in the corresponding technology [Riv99].

5.5.3.2 Velocity saturation

In a deep submicron process, however, it is not always possible to increase the transconductance decreasing L, following equation (5.4). In effect for very high drain to

[19] The power supply of the ALICE1LHCb chip is 1.6-1.8V.

source (longitudinal) electric fields[20] (and so for short channel lengths) the carrier velocity in the channel does not increase anymore linearly with the electric field. This effect is called velocity saturation, and limits the transconductance that can be achieved in a given technology to:

$$g_{m,max} = W\, C_{ox}\, v_{sat} \tag{5.5}$$

which is independent on the transistor length.

5.5.3.3 Weak inversion region operation

In weak inversion the transconductance is given by:

$$g_{m,wi} = \frac{I_{DS}}{n\, U_t} \tag{5.6}$$

where U_t is the thermal voltage, so for a given current the reduction of the transistor length does not help to increase the transistor driving capability.

On the other hand, if we consider the ratio g_m/I_{DS}, which is a measure of the power efficiency of the device, this stays constant in weak inversion and goes with the inverse of the square root of I_{DS} in strong inversion. This means that in weak inversion the power efficiency is maximised. Moreover in weak inversion the V_{DS} needed to make the transistor work in saturation is of the order of 100-200 mV and is independent of the transistor current. This feature allows the stacking of several transistors, even with low power supplies.

In submicron technologies it is easier to operate in the weak inversion region. The boundary of the strong and weak inversion is given by the *specific current* I_s [Enz01, Enz95]:

$$I_S = 2\, n\, K \frac{W}{L} U_T^2 \tag{5.7}$$

If the *inversion factor (or coefficient)* $i_f = I_{DS}/I_S$ is higher than ~5, the transistor can be considered in strong inversion, if i_f is lower than ~1/5 the transistor can be considered in weak inversion. From equation (5.7) we can see that I_S increases with technology scaling (for a constant W/L), making it easier to work in weak inversion. Moreover, advanced technologies allow a shorter L and thus increased I_S.

5.5.3.4 Electronic noise

Another important effect of scaling is its impact on the electronic noise of the devices[21]. In effect the input referred power spectral density of the thermal noise is inversely proportional to the transconductance g_m, and the input referred power spectral density of the

[20] In the order of 1 V/μm for electrons and 3 V/μm for holes.
[21] See Appendix I and Section 5.7.3 for a more detailed discussion about noise in submicron technologies.

flicker noise is inversely proportional to the gate capacitance per unit area C_{ox}^2. This means that in principle technology scaling should improve also noise performance of the device. In practise, the two noise contributions depend also on empirical coefficients which can change severely in different processes, and on other physical effects such as hot carriers or weak avalanche. So, even if in principle a noise reduction for a constant power budget is expected, no conclusion can be drawn until experimental measurements on the target technology are performed.

5.5.3.5 Matching

Matching of the threshold voltage is also affected by technology scaling[22]. It can be shown that for a constant device area, for a mature technology, the threshold voltage matching characteristic parameter improves linearly with the gate oxide thickness [Tui01].

5.5.3.6 Reduced voltage margins

One of the problems related with the scaling in the submicron range, for the analogue designer, is the fact that the market is driven by digital circuits, and as a consequence technologies are optimised for digital applications. The major problem is that since the subthreshold slope is not technology dependent, to avoid high off-state currents the threshold voltage is not scaled in the same way as the power supply voltage. The available margin for analogue signal swing then decreases, making analogue design on the same physical substrate of a digital circuit more and more difficult for advanced technologies. This problem is somehow mitigated by the better matching characteristics of scaled technologies. In effect the analogue designer has to apply "safety" design margins which take into account transistor mismatch, so that a better matching effectively results in increased design margins.

5.6 Hardening By Design

5.6.1 Hardening by layout

This is the part of the HBD approach which consists in using special layout techniques to improve radiation hardness characteristics of a commercial technology.

5.6.1.1 Threshold voltage shift

As we have seen in section 5.5, the problem of radiation induced threshold voltage shift is drastically reduced thanks to the much higher radiation tolerance of thin oxides. The target technology CMOS 0.25 μm chosen for the design of the ALICE1LHCb chip has an oxide thickness of about 5.5 nm, which is much less than the cited limit of ~20 nm, where the

[22] See Section 5.7.2 for a more detailed discussion of matching in submicron technologies.

dependence of the threshold voltage shift with the gate oxide thickness starts to be much faster than $1/t_{ox}^2$. The only design-related way to cope with this problem is at system level, as we will see in section 5.6.2.1.

5.6.1.2 Edge leakage

The main problem that has to be solved by adequate layout techniques is the increase of the parasitic off-state currents. Some solutions to this problem [Hat85, Hat86] (already explained in 5.2.3) are shown in Figure 5.13 and Figure 5.14. The first one, shown in Figure 5.13 (left) consists of increasing the width of the gate at the edges of the transistor. This increases the length (and the threshold) of the parasitic device. This solution is the least intrusive but also the least effective.

Another possibility, shown in Figure 5.13 (right), consists of using an additional p^+ mask in the standard n-channel transistor process. In this way the limits of the thin oxide are defined by the p^+ implant. With this solution the geometry of the transistor is almost identical to that of a standard device, but cannot be classified as HBD as it usually violates the design rules of commercial CMOS processes. Moreover, some parasitic leakage paths can still be created for very high doses.

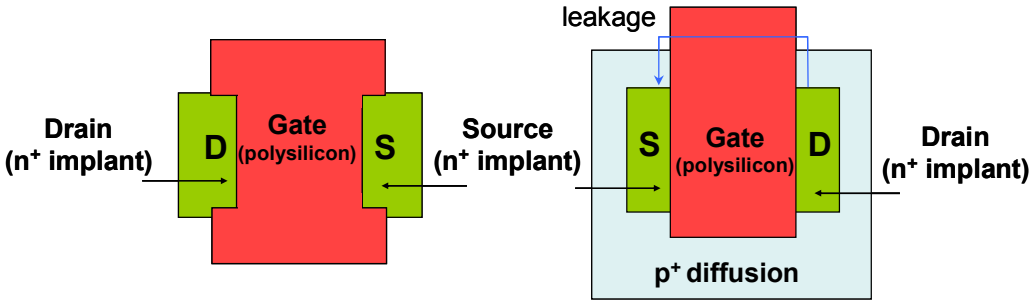

Figure 5.13: Two layout solutions which can mitigate edge leakage effects. The first (left) consists of increasing the width of the gate at the edges of the transistor. The second (right) of using an additional p^+ mask to the standard n-channel transistor process.

Other two types of solutions, much safer from the radiation hardness point of view, are shown in Figure 5.14. The first solution was proposed by Nowlin and co-workers in [Now03] (Figure 5.14, left). It consists of enclosing only the source (or the drain) of the n-channel transistor with an annular gate, so that radiation-induced source-to-drain leakage paths are eliminated. In this case, rather than being radial as in the solution shown in Figure 5.14 (right), the channel is more nearly transverse as in standard devices. With this solution the resulting transistor is very close to a standard device, if we exclude the high gate-to-source/gate-to-drain capacitance. Moreover the excess polysilicon can have a significant resistance associated with it, especially in large devices.

The solution shown in Figure 5.14 (right), consist of drawing an annular gate in the active region so that the drain (or the source) is inside and the source (or the drain) is outside.

Also for this solution the parasitic path under the field oxide is not present, because all the current which flows from drain to source has to pass underneath the gate. This means that no irradiation induced off-state currents can be generated. This type of transistor is called Enclosed Layout Transistor (ELT) or "edgeless transistor", and is the one which has been used for the design of the ALICE1LHCb chip. For this reason in section 5.7 the characteristics of this type of transistor will be described in detail.

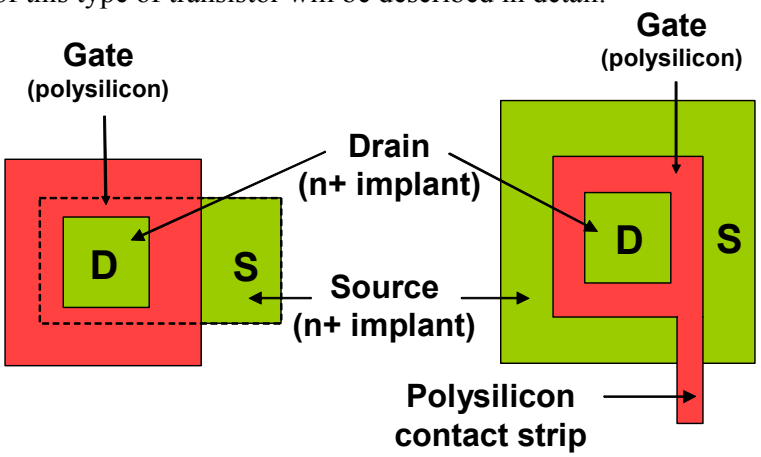

Figure 5.14: Two layout solutions which can eliminate edge leakage effects. The first (left) consists of enclosing only the source (or the drain) of the n-channel transistor with an annular gate. The second (right) consists of drawing an annular gate in the active region.

Although this solution is very efficient for radiation hardness, it presents several disadvantages. First of all, a transistor laid out in this way has a much bigger area than a rectangular transistor, and this decreases the area gain of using a standard CMOS technology instead of a rad-hard process (more details on area and performance loss due to the use of ELTs are given in section 5.8). The increased size generates also additional capacitance, both gate capacitance and source (or drain[23]) capacitance. This worsens the time response characteristics of the device.

The enclosed geometry generates a transistor which is not symmetric, having a different behaviour if the source is connected inside or outside the gate (for example the output conductance is higher if the drain is connected inside). Moreover, it is not evident how to model the device, as the models provided with the circuit simulator by the foundry are designed and extracted for rectangular transistors. In particular, it is difficult to model the most important parameter, which is the W/L ratio. In addition to this, this parameter cannot be chosen freely, because there is a minimum W/L which cannot be reduced. For all the currents that can be used in a low-power design this results in all the n-channel transistors in the weak or moderate inversion region, posing problems especially for the precision of current sources.

All these peculiarities of the ELT require modifications of the CAD tools at several levels to make them able to sensibly process an ELT design. The first problem is at simulator

[23] The terminal which is at the outside of the gate has a much higher capacitance than the other one.

level, because as we have seen the transistor cannot be properly modelled with the normal rectangular transistor model provided by the foundry. The layout is also complicated by the use of ELT transistors. Special scripts are also needed in the extraction phase[24], for example to recognise a round and enclosed shape as a transistor gate.

The use of ELTs as radiation hard devices was already known and proved in the early days of CMOS [Din77a, Din77b, Nap82, Ale96]. To design a full radiation tolerant chip with ELTs, it was fundamental to study its properties in detail, as for example modelling, noise, matching, and design density issues.

Other potential layout solutions to mitigate edge leakage were proposed, but they never gained enough acceptance as an effective approach [Ale96].

5.6.1.3 Inter-transistor leakage

The problem of inter-transistor leakage was investigated, together with ELTs, in the frame of the RD49 CERN research program. Some special transistors were designed, called FOXFETs[25], and are shown in Figure 5.15, compared with a standard transistor. These structures are built to measure the leakage current underneath the LOCOS or the STI as a function of the TID or of the gate material. This structure simulates the worst case condition for the inter-transistor leakage, that is a biased interconnection line running on top of the field oxide.

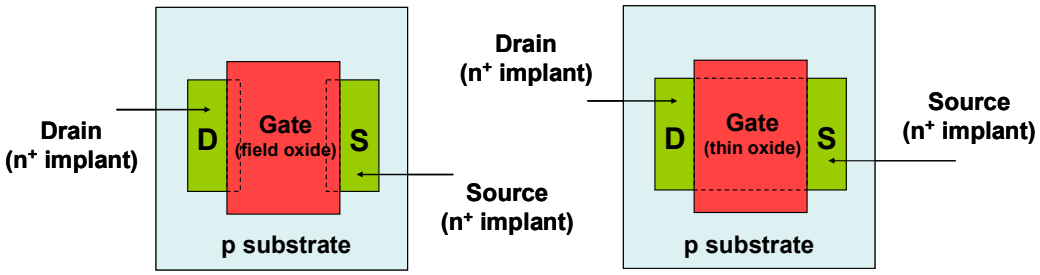

Figure 5.15: Schematic representation of the FOXFET test structure (left) compared with a standard transistor (right).

The solution to the problem is shown in Figure 5.16. It consists of adding a p^+ guard ring (also called channel stop) around each n-type region at different potential than the negative power supply, and around any n-well at a different potential than the positive power supply (this for example happens for p-channel transistors with the source connected to the bulk). The p^+ guard ring stops any possible leakage towards any n-well at a different potential. In

[24] In the extraction phase the layout is processed and an equivalent schematic diagram is produced automatically by the CAD tool. The output of the extraction tool can be fed to another checking program, the Layout Versus Schematic (LVS) which compares the extracted circuit with the original schematic version of the circuit, to check if there is any difference between them.

[25] Field Oxide Field Effect Transistors, because they are FETs realized with the field oxide instead than the gate oxide.

effect it prevents the inversion of the field oxide at that location by adjusting the local threshold voltage to a very high value. This maintains the integrity of the isolation between adjacent n-channel devices and eliminates the n-well to n^+-source leakage paths.

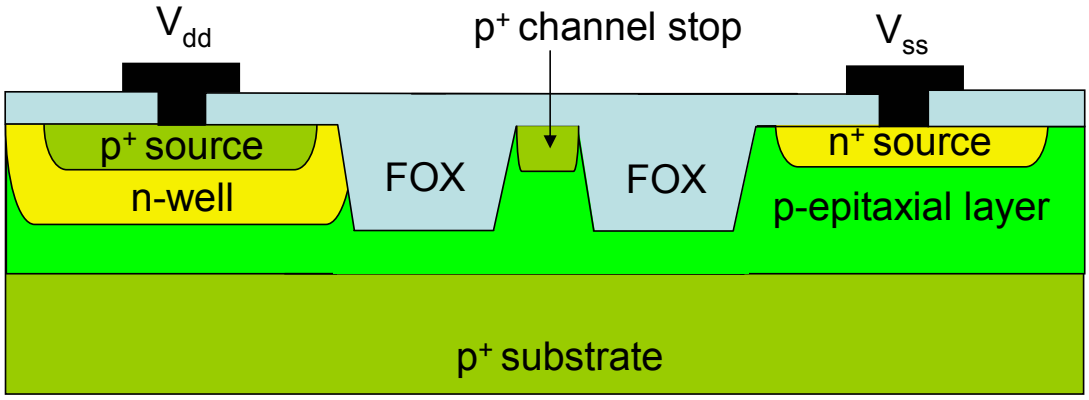

Figure 5.16: Schematic representation of the interposition of a p^+ guard ring to prevent inter-transistor leakage.

This can be performed as part of a standard CMOS process flow and does not require the insertion of additional masks or process steps. The effectiveness of this technique is shown in Figure 5.17.

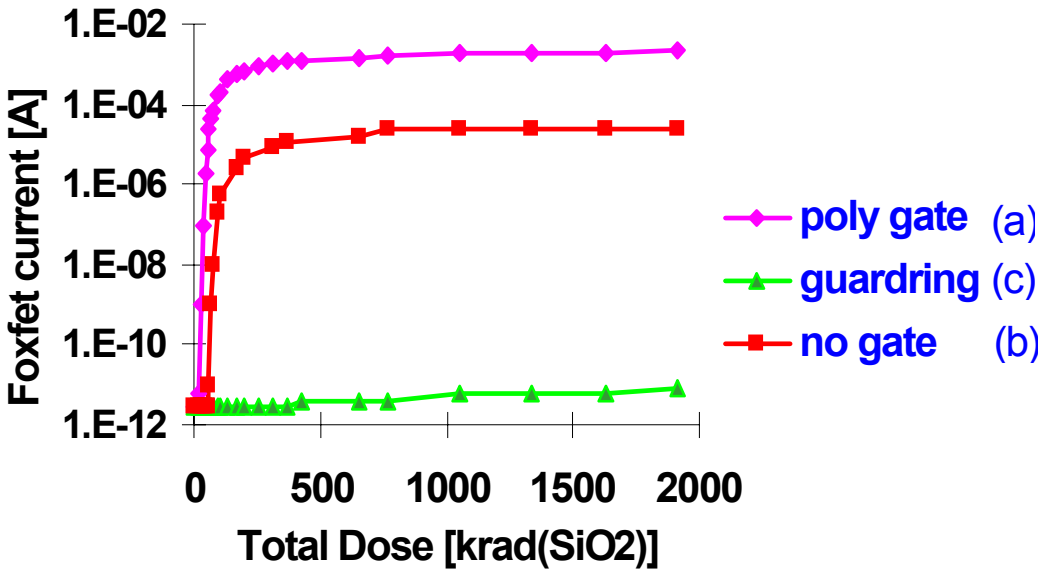

Figure 5.17: Example of the effectiveness of guard rings in preventing inter-transistor leakage. The current for three different FOXFETs (laid out in a 0.5 μm technology) is shown as a function of the total dose. The highest curve is relative to a FOXFET with a polysilicon gate, the central one to a FOXFET without gate, and the lowest one to a FOXFEX with a guard ring isolating the source from the drain [Sno00].

The current for three different FOXFETs (laid out in a 0.5 μm technology) is shown as a function of the total dose. The highest curve is relative to a FOXFET with a polysilicon gate (a), irradiated with a gate-source voltage equal to V_{dd} (the positive power supply voltage; this is the worst case bias condition). The central curve refers to a FOXFET without gate (b), and the lowest one to a FOXFET with a guard ring isolating the source from the drain. In all the cases the separation was the minimum as determined by the 0.5 μm technology. In the case of (a) and (b) the leakage starts at 30 and 70 krd respectively, while in case (c) no leakage is observed up to 2 Mrd (SiO_2) [Sno00].

The use of p^+ guard rings has some drawbacks. The first is that a local polysilicon interconnect cannot be used in signal routing when channel stops are employed. The reason is that if the polysilicon gate of a transistor crosses a guard ring, this blocks the p^+ implant and creates a gap in the channel stop. This can provide a potential leakage path.

The second drawback of guard rings is that they require additional area. The amount of extra area used is strongly dependent on the technology layout rules and on the shape of the transistor which is enclosed by the guard ring.

5.6.1.4 SEL vulnerability

Advanced technologies make already use of low supply voltages, epitaxial layers, retrograde wells and STIs, which reduce SEL sensitivity. However, some other layout precautions can be applied to further reduce SEL vulnerability.

Referring to Figure 5.9, the n^+ and p^+ contacts must be sufficiently far to prevent the β of the lateral parasitic npn transistor being too high (this increases the length of the base of the parasitic transistor). The well (substrate) must be connected via n^+ (p^+) contacts at several points sufficiently close to each other, so that R_W (R_S) can be kept low. The use of p^+ guard rings around n-channel transistors and n^+ guard rings around p^+ transistors also reduces the gain of the parasitic lateral npn and npn transistors modifying the doping profile in the base of the transistors. For example, Osborne [Osb98b] showed that quadrupling the minimal distance (set by the technology design rules) between well edges and active regions results in a latchup threshold increase of a factor ~2 in an HP 0.5 μm process. It should be noted that quadrupling this distance could result in a minimal area waste, because diffusions approach well edges only in a small fraction of the total area of a typical CMOS circuit.

5.6.2 Hardening by circuit and system architecture

5.6.2.1 Total Ionising Dose

The total ionising dose tolerance of a circuit is addressed mainly at layout and process level. However, the designer can take into account the foreseen change in the device

parameters, taking care that the design can tolerate these changes without any major functionality problems.

To help this procedure some dedicated CAD tools can be employed. One approach starts from first principles, so that the device parameter variations with irradiation can be predicted on the basis of the physical structure of the MOS transistor and on the type of radiation interactions that will happen during its lifetime. The usefulness of this approach is strictly connected with the possibility of knowing the technological details of the transistors to be employed, which could not be available from the foundry.

A different approach consists of extracting the models of sample transistors during and after irradiation; this does not require the knowledge of the technological details but the accuracy of this technique depends on the consistency of radiation-induced changes in transistors behaviour across a lot, and has to be re-evaluated from lot to lot to ensure that process variations or changes do not affect radiation behaviour.

Circuit simulations with post-irradiation transistors have to ensure that the design is robust enough for the foreseen TID received during its lifetime.

For digital circuits the synchronous mode of operation is preferred over the asynchronous, since synchronising the logical states with a clock limits the sensitivity to transition time variations and to drifts in the electrical parameters.

5.6.2.2 Single Event Upset

The three main techniques to decrease SEU sensitivity are charge dissipation techniques, spatial redundancy and error correction techniques. Static architectures are more SEU-tolerant than dynamic ones (see section 5.9).

Charge dissipation techniques

These techniques aim to improve the cell ability to dissipate or absorb the charge created by the impinging particle. A first approach consists of **increasing the width of the transistors driving the sensitive node**. This increases their conductance and their driving capability, so that when a particle strikes the sensitive node the pull-up transistor is able to supply the additional current needed to maintain the original latched logic state. This technique also increases the node capacitance, so that the voltage swing caused by the SEU injected charge is reduced. An **additional capacitive load** can be added by design, but while the increase of the transistor width does not slow down the circuit, additional node capacitance reduces the logic gate speed.

The **resistive hardening** consists of adding feedback resistances between the cross-coupled inverters of a SRAM cell or latch, as shown in Figure 5.18.

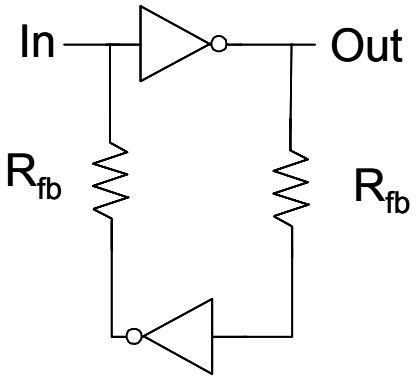

Figure 5.18: Schematic representation of the resistive hardening technique on a D-Latch.

The additional resistors increase the RC delay of the cell, so that the pull-up transistor has enough time to restore the original voltage on the node before it is latched by the cell. This approach has severe drawbacks because it increases the delay of the cell and the cell area.

Spatial redundancy

The techniques which make use of the spatial redundancy reproduce the information stored in the cell in different positions on the chip, so that if the information gets corrupted somewhere, it is restored thanks to the same information stored somewhere else.

An in depth presentation of SEU-hard cells can be found in [Hei99]. An elegant example of an area-efficient storage cell that uses a 4-node redundant structure is the *Dual Interlock storage CEll (DICE)*, proposed by Calin [Cal96]. This cell is suitable for replacing SEU-soft latches and flip-flops and to implement SEU-hardened SRAMs, avoiding the main drawbacks of other solutions proposed earlier. This SEU-hardening technique has been applied in the design of the most important latches (the pixel configuration latches) of the ALICE1LHCb chip.

The most straightforward implementation of a spatial redundancy technique is the *Triple Modular Redundancy (TMR)*, shown in Figure 5.19. It consists of replacing a single unhardened latch with three unhardened latches; each of them stores the same information. Each of these latches is connected to the same clock and to the same data line. The output of the latches goes into a majority voting logic circuit.

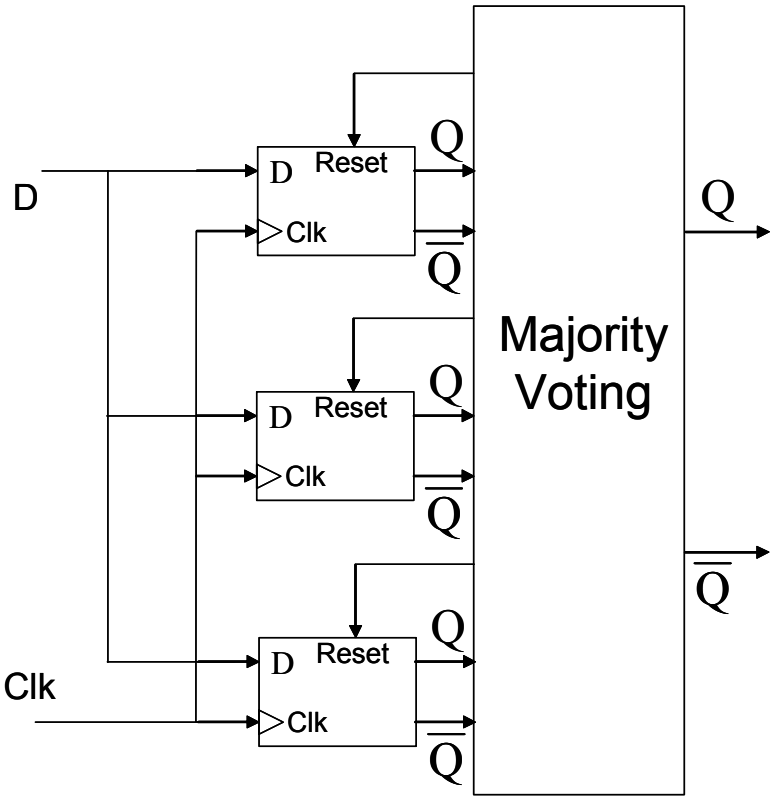

Figure 5.19: Schematic representation of a Triple Modular Redundancy cell.

If a particle generates an upset in one of the latches, the other two still retain the correct information. The majority voting scheme will still read two correct inputs over three, and provide the correct output. The area and power penalties associated with this scheme are approximately a factor 3, which makes this technique not suitable for large SRAMs. A problem associated with this cell is the possibility of a double error in the TMR latches, which results in a wrong output state. If the cell refresh rate is not high in comparison with the expected upset rate, errors can accumulate and give origin to double errors. For this reason the cells in the TMR latch should be refreshed periodically to avoid accumulation of errors. Also a particle which is energetic enough to upset two latches at the same time, or which impacts on the IC almost parallel to the surface, can create a double error.

Another scheme to obtain the SEU-hardness of a complex logic circuit is shown in Figure 5.20. In this case the full state machine is replicated three times, and the outputs are fed to the voting circuitry. The output of the voting circuitry is then fed back to the state machine, to recover the wrong states induced by the SEU. This SEU-hardening technique is applied in the DPC (see section 3.2.3.1).

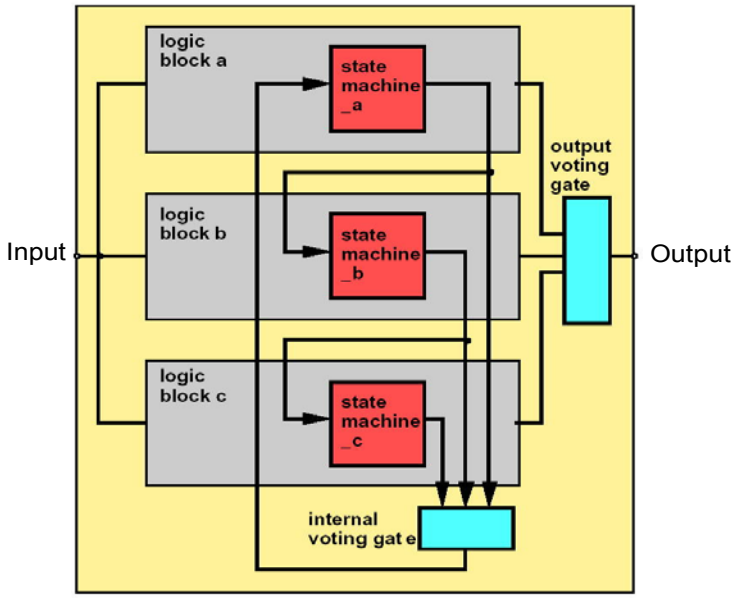

Figure 5.20: Different implementation of a triple redundancy scheme. In this case the full state machine is replicated three times, and the outputs are fed to the voting circuitry.

The voting logic can be designed with only combinatorial logic which is immune to SEU, but which is not immune to SET: special design techniques to mitigate SET should be used to implement the voting logic.

Error correction techniques (ECTs)

Error correction techniques consist in hardening a memory element checking for errors in a set of unhardened memory elements, ignoring or rewriting elements that have been corrupted.

Several types of codes, called Error Detection And Correction codes (EDAC) or Error Control Codes (ECC) have been developed to detect and if possible correct corrupted bits (TMR is a particular EDAC).

The simplest ECC is the **parity checking**, which consists of adding a parity bit to a string of bits, so that the total number of zeroes or ones is always even (or odd, depending on the type of parity chosen). If the parity check is odd, this means that an error has occurred in one of the bits of the string. Parity check cannot reveal multiple errors (which occur in an even number) and cannot correct the wrong string; moreover the possible upset of the parity bit introduces an additional source of error.

A widely used EDAC is the **Hamming technique**. The code uses extra redundant bits to check for errors, and performs the checks with special check equations. Each check bit has a corresponding check equation that covers a portion of all the bits of the string, but always includes the check bit itself. This code allows the correction of single errors, because the parity check equations, in case of an error, will flag the position of the corrupted bit. In case

of a double error the code is still able to detect it, but the information stored in the parity bits is not enough to find the position of the wrong bits. This code has a good balance between safety of the information stored and overhead (number of additional bits needed to implement the code). The (relative) overhead decreases with the increasing size of the word to protect: for a 16 bits word we need 6 additional bits (38% overhead), for a 57 bit word we need 7 additional bits (for a total of 64 bits; 12% overhead).

5.6.2.3 Single Event Transient

Single event transients can be mitigated using HBD techniques which are very similar to those used to mitigate SEUs. One of the possibilities is to slow down the circuit speed, increasing the critical width. This can be obtained by increasing the capacitive load at the logic cell output. Alternatively it is possible to increase the drive of the transistors in the logic cells, decreasing in this way the transient width and its voltage change; this approach also increases the capacitance on the output node.

A redundancy technique can also be used for SET, but in this case a temporal redundancy has to be applied (Figure 5.21). In the triple voting scheme the signal is stored in three parallel D-latches, which are fed with delayed clock input edges. The latch outputs are fed to an asynchronous majority voting scheme. If a SET occurs, it will arrive at the same moment at all the latch inputs, but will be stored at most in one of them (the one which has the clock edge in coincidence with the SET). The asynchronous majority voting circuit will then be able to produce the correct output.

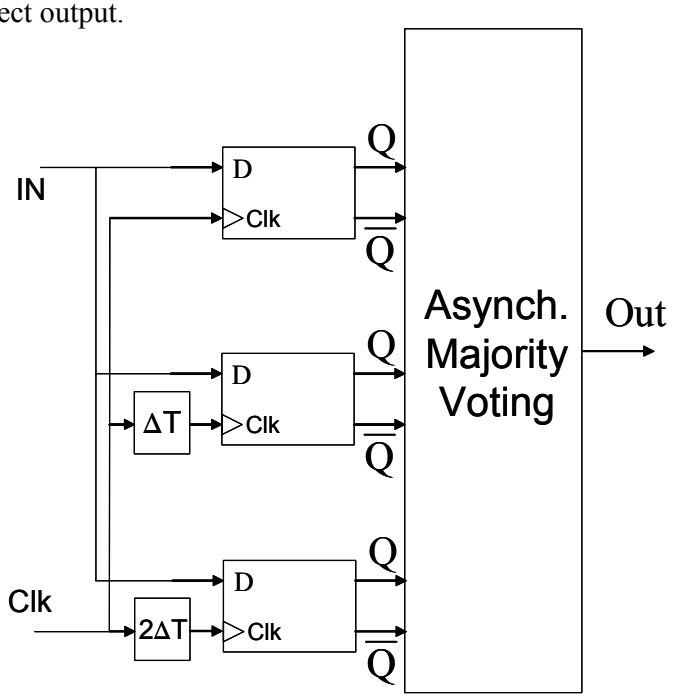

Figure 5.21: Schematic representation of a SET hardened latch design using temporal redundancy.

5.7 The Enclosed Layout Transistor (ELT)

The Enclosed Layout Transistor is the basic building block used to make a standard CMOS technology radiation tolerant. It is fundamental to know in detail its characteristics to be able to design a full system: a reliable model, its matching and noise characteristics and its behaviour under irradiation. All this was studied in the frame of the RD49 program and presented in [Ane00].

5.7.1 Modelling of the enclosed layout transistor

Some of the possible shapes for an ELT are shown in Figure 5.22.

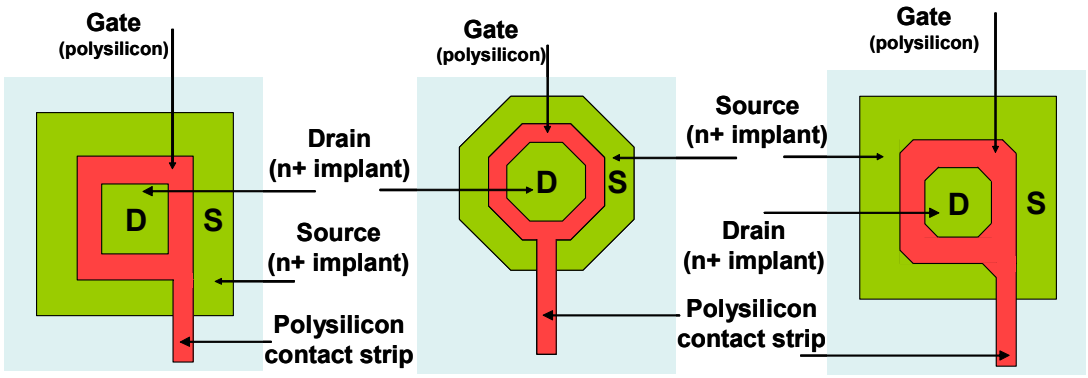

Figure 5.22: Three possible shapes for an ELT. (a) square; (b) octagonal; (c) square with corners cut at 45°. Drain is marked inside the transistor and Source outside, but they can be reversed.

In principle the best thing to do from a design point of view is to choose different transistors in different blocks of the circuit, and this is what has been done (to some extent) in the design of the ALICE1LHCb chip.

Shape (a) violates design rules of the target technology chosen for the design, so it was not used for the chip design.

The transistor shown in Figure 5.22 (b) minimises the gate capacitance for the same W and L if compared to the transistor (c). However, shape (c) allows designing a minimum size transistor smaller than what could be realised with shape (b), due to the shape (square) of the metal contact needed to connect the inner diffusion. Information about modelling of generic ELTs can be found in [Gri82, Gir98, Gir00]. Note that the shape and the position of the polysilicon strip (which is necessary to integrate the gate contact outside the thin gate oxide region) can change the effective parameters of the device, especially for small transistors.

A shape which has been studied in detail is shape (c), so the guideline used for the chip design was to use transistor of the type (c) when a precise modelling of the W/L ratio was required, and to use type (b) if speed of the transistor was an issue.

Studies of the electric field under the gate of the device, supported by simulations and measurements, lead to the following formula, based on the "decomposition" of the annular transistor gate in several transistors in parallel:

$$\left(\frac{W}{L}\right)_{eff} = \overbrace{4 \cdot \frac{2\alpha}{\ln\frac{d'}{d'-2\alpha \cdot L_{eff}}}}^{T1} + \overbrace{2K \cdot \frac{1-\alpha}{\frac{1}{2}\sqrt{\alpha^2+2\alpha+5} \cdot \ln\frac{1}{\alpha}}}^{T2} + \overbrace{3 \cdot \frac{\frac{d-d'}{2}}{L_{eff}}}^{T3} \qquad (5.8)$$

where c,d,d'=d- $c \cdot \sqrt{2}$ and α are shown in Figure 5.23.

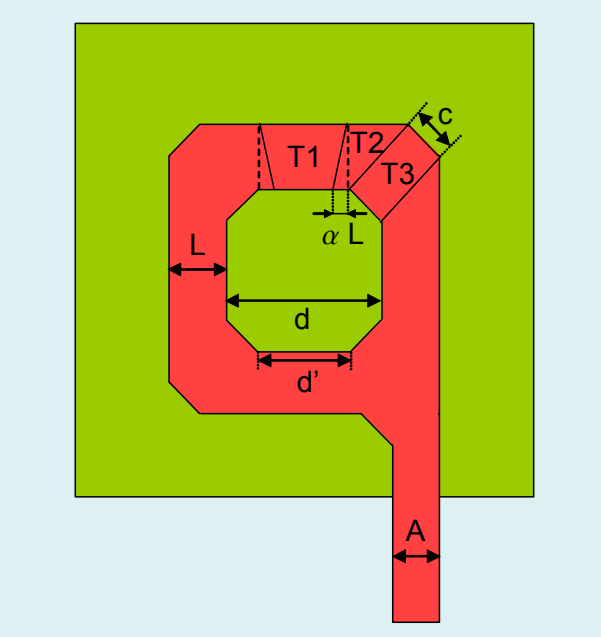

Figure 5.23: Shape of an ELT transistor with all the dimensions used in equation (5.8). The transistor can be thought as composed by the parallel combination of transistors of three different shapes, labelled T1, T2, T3 in the picture an corresponding to the three terms in equation (5.8).

L_{eff} is the so called *effective gate length*: due to the diffusion of the drain and source under the gate (underdiffusion) and photolithography and etching imprecision, the gate length which is physically realised on the device is slightly shorter than the designed one. This correction has little effect on long devices, but can affect calculations if the transistor is nearly minimum size. K is a geometry dependent parameter, which takes into account the number of transistors T2 present in the ELT. α is a fitting parameter, related to the position of the border line between transistors of type T1 and T2. The value of α has to be found experimentally for

each technology measuring several test structures and fitting the results with equation (5.8). A value of $\alpha = 0.05$ is cited in [Gir00] as the best fit for several technologies from 2.5 μm to 0.25μm.

Each term of equation (5.8), representing one of the transistors shown in Figure 5.23, has a multiplying factor which depends on the number of times the transistor is present in the ELT shape. The polysilicon strip "hides" a transistor T3 and part of a transistor T2, this is why the third term is multiplied only by three and the second one by a geometry dependent parameter K[26].

Table 5.3 shows the good agreement between equation (5.8) and the measured values [Ane99]. An ELT with a given L (left column) is biased at a given gate overdrive $V_{gs} - V_{th}$, and the drain current is measured. The ratio of this current to the current of a standard transistor with the same gate bias allows extracting the effective W/L of the ELT under test.

L_{drawn} [μm]	Calculated $(W/L)_{eff}$	Extracted $(W/L)_{eff}$
0.28	14.8	15
0.36	11.3	11.2
0.5	8.3	8.3
1	5.1	5.2
3	3	3.2
5	2.6	2.6

Table 5.3: Calculated and extracted $(W/L)_{eff}$ for ELTs of different drawn lengths. The table shows the good agreement of the experimental values with the calculated ones [Ane99].

One of the main problems related to ELTs is the limitation in the achievable aspect ratio. To obtain large W/L ratios it is enough to stretch the basic ELT shape in one or two dimensions. This increases W and leaves L constant, increasing W/L. Unfortunately it is not possible to obtain low values for W/L in such a simple way. There is a minimum size for the inner terminal (i.e. there is a minimum d in Figure 5.23), so the only way to decrease W/L is to keep the minimum d and increase L. In this way, due to the annular shape of the ELT W increases also. Increasing L decreases the contributions due to T1 and T3, so W/L decreases until T2 (which does not depend on L) becomes dominant. For the shape in Figure 5.23 the minimum achievable W/L is ~2.26, which is already almost attained for L = 7 μm. Moreover, aspect ratios close to these lead to the design of ELTs with a significant area waste in comparison with a standard device, and should be avoided by using different circuit topologies.

The asymmetry of the device has also to be taken into account while designing with ELTs. The output conductance of the device, for example, is significantly different if the drain contact is placed inside the ELT or outside. This effect is shown in Table 5.4: for a given L

[26] K is about 7/2 for short transistors (L ≤ 0.5 μm) and 4 for long channel devices.

the output conductance of an n-channel ELT transistor is measured considering as the drain the inner terminal (G_{di} column 2) or the outer terminal (G_{do} column3).

L_{drawn} [μm]	G_{di} [μs]	G_{do} [μs]	Difference [%]
0.28	11.89	9.62	19
0.36	7.17	5.55	23
0.5	4.10	2.73	33
1	1.68	0.79	53
3	0.57	0.17	70
5	0.41	0.10	75

Table 5.4: Output conductance of an n-channel ELT of different gate lengths designed in our 0.25 μm technology, measured with the inner diffusion as drain (G_{di}) or with the outer diffusion as drain (G_{do}) [Ane99].

The difference between G_{di} and G_{do} (column four, Difference = $(G_{di} - G_{do}) / G_{di}$) increases with the gate length L and can reach values as high as 75%. The value of the output conductance G_{dn} for a standard (rectangular) device is always in between G_{di} and G_{do}: for $L \leq 0.5$ μm $G_{di} \approx G_{dn}$, for longer devices $G_{dn} \approx (G_{di} + G_{do})/2$. This means that it is possible to obtain lower (i.e. better) values for G_d with an ELT, and this feature can be exploited during the design, for example to obtain a higher gain in an amplifier. Due to the different size of the drain and source diffusions, the capacitance associated with the terminals is also different, and is higher for the outer terminal. This has to be taken into account during the design: for example in pass-gates charge injection towards the outer terminal is bigger, or this can add capacitance in a sensitive node of an amplifier or of a logic gate. On the other hand, we have seen that extra capacitance added on a sensitive node of a logic circuit helps reduce SEU sensitivity.

5.7.2 Matching properties of ELTs [Pel89, Ane00, Tui01]

Mismatch of transistors is the process that causes time-independent random variations in physical quantities of identically designed and biased devices, closely spaced and surrounded by an identical environment. Mismatch is a limiting factor in the design of several circuits or blocks, both in analogue and in digital design (for example current sources, bandgap voltage references, D/A and A/D converters, switched capacitor circuits). For this reason matching characteristics of ELT devices in our target 0.25 μm technology were studied [Ane00].

The mismatch of two devices can be generated by two main classes of effects: systematic and stochastic effects.

5.7.2.1 Systematic effects

Systematic effects can be generated by some technological and geometrical causes. The most typical are related to real physical dimensions, device orientation, dopant non uniformities, contact and via resistance non-uniformities, charging damage and mechanical stress[27]. Systematic effects can also be related with electrical mismatch, as for example bias differences, parasitic components or supply line currents. This kind of effects can be addressed (and solved) both by the foundry and with design or layout techniques. Good matching practices to avoid systematic effects are:

- design identical transistors (neither rotated nor mirrored);
- place them at small distance (<100 μm);
- connect them exactly in the same way;
- bias them identically;
- not overlay them with other components (as metal lines or other circuit elements); if this cannot be avoided identical overlaying shapes should be used;
- take care that they have exactly the same neighbouring (for example using "dummy neighbours" up to 20-40 μm around the critical areas);
- place them at a distance higher than 40 μm from the rest of the circuit and 200 μm from the chip edge;
- be careful about voltage drops, delays and temperature gradients on the chip.

5.7.2.2 Stochastic effects

While cancelling systematic effects is a matter of good engineering, stochastic effects cannot be avoided. If we consider a generic transistor parameter P, in general the value of the parameter is composed of a fixed part and a randomly varying part, resulting in different values of P at different coordinates on the wafer. The randomly varying part is related with statistical fluctuations of the physical parameters of the transistor: dopant fluctuations [Sto98], local mobility fluctuations, polysilicon gate granularity [Tui97], oxide charges and interface states fluctuations.

A common mathematical approach to all these mismatch sources can be followed, provided that the phenomena under study have these characteristics:

- the mismatch of the parameter under study is caused by many single events of the mismatch generating process;
- the effect of a single event on a parameter is so small that the contributions of many single events can be summed;

[27] Each of these causes can be related to several other sub-causes, for example a variation of the real physical dimensions can be generated by photomask offsets or raggedness, photoresist crowding or proximity effects, while dopant non-uniformities can be caused by non-uniformities of the furnace temperature, gas depletion, lateral auto-doping or ion implantation beam striping.

- the events have a correlation distance which is much smaller than the active area of the components.

With these assumptions the occurrences of the events are mutually independent, and the Poisson statistic is applicable. This results in a Gaussian distribution of the random mismatch amplitude ΔP of the parameter P that has a variance [Pel89]:

$$\sigma^2_{\Delta P} = \frac{A^2_P}{W\,L} \tag{5.9}$$

where W and L are the effective width and length of the device, and A_p is the area proportionality constant for the parameter P.

The assumption of a short correlation distance implies that no relation exists between matching and spacing between two devices. This is not true in reality: this second class of mismatch is a deterministic process, but as the original placement of dies on a wafer is unknown after packaging, the effect on the parameter under study can be modelled as an additional stochastic process with a long correlation distance. This contribution is negligible at small distances.

This means that we can re-write equation (5.9) for the case of the matching of the two most important parameters for a MOS device, the threshold voltage V_{th} and the current factor β [28]:

$$\sigma^2_{\Delta V_{th}} = \frac{A^2_{V_{th}}}{W\,L} \tag{5.10}$$

$$\sigma^2_{\Delta\beta/\beta} = \frac{A^2_{\beta}}{W\,L} \tag{5.11}$$

The mismatch of the threshold voltage is expressed on the basis of the difference between the threshold voltages $\Delta V_{th}=V_{th1} - V_{th2}$, while the current factor is based on the relative difference between the β of the two transistors of the pair, that is:

$$\frac{\Delta\beta}{\beta} = 2 \cdot 100 \cdot \frac{\beta_1 - \beta_2}{\beta_1 + \beta_2} \tag{5.12}$$

A_{Vth} [V·m] is the threshold voltage matching performance of the given technology, and is generated by the sum of several physical effects. The major contribution is given by the random dopant fluctuation in the depletion region underneath the channel; another important contribution is due to the variation of the charges at the silicon-oxide interface.

---

[28] The current factor β is given by: $\beta = \mu\, C_{ox} \frac{W}{L}$ where μ is the carrier mobility, C_{ox} is the gate capacitance per unit area and W and L are the effective gate width and length of the transistor.

A_β [%·m] is the current factor matching performance of the given technology, which is expressed by equation (5.11) if the main contribution to the current factor mismatch is due to the mobility and oxide capacitance fluctuations (as is the case in the modern technologies).

It can be calculated that [Sto98] the threshold voltage matching factor related with the random dopant fluctuations is proportional to $t_{ox} \cdot \sqrt[4]{N}$ (where N is the channel dopant concentration). With technology scaling t_{ox} decreases and N increases, but due to the fourth root in the formula the overall effect is a decrease of A_{vth}, which means that transistor matching improves with scaling (for a constant area device). The overall matching of minimum size transistors can be degraded with technology scaling, but if the parameters are scaled sensibly it is possible to keep the matching of minimum size pairs almost unchanged.

A parameter that is often used as a benchmark for technologies is A_{vth}/t_{ox} [V]: it was found, measuring several processes with gate oxide thicknesses ranging from 50 nm to 5 nm [Tui01], that a value of 1 mV·μm / nm is a target value for a mature technology. For what concerns the current factor matching, A_β has been measured to be from 1% to 3% for many different processes [Tui01].

A campaign of measurements on a statistically relevant sample of matched pair transistors (7000 transistors) of different dimensions was done at CERN [Ane00], to estimate A_{vth} and A_β for standard transistors and for n-channel ELT transistors in our target 0.25 μm technology. The shape used for n-channel ELTs is the one shown in Figure 5.23, with a single contact inside the annular gate, or with four contacts inside the gate. Some transistors are laid out as a fourth of a normal ELT transistor, to investigate a possible mismatch component related to the corner of the enclosed device.

The measurements were done at wafer level, to avoid dicing, bonding and packaging. The transistors were also irradiated with X-rays to check the post irradiation behaviour of the matching parameters. Results of the threshold voltage matching properties are shown in Figure 5.24 and Figure 5.25.

Figure 5.24 plots the threshold voltage matching performance of ELT pairs, measured with the drain inside the annular gate (ELT_di), compared with the matching performance of the standard n-channel devices. The numbers on the plot close to the points indicate the gate length. The two circles on the plot represent the $\sigma_{\Delta Vth}$ values for the pairs with four contacts in gate, measured with the drain inside the gate. The figure shows that standard n-channel transistors follow the (5.10), and the value that can be extracted for A_{Vth} is 4 mV·μm. From a similar plot for p-channel transistors it was extracted a value for A_{Vth} of 3.7 mV·μm.

For standard n-channel and p-channel transistors the β current factor mismatch was also measured. The measured values of $\sigma_{\Delta\beta/\beta}$ are below 1 % for all the measured pairs, and A_β is around 1 %·μm, though the level of confidence of the fit is rather low.

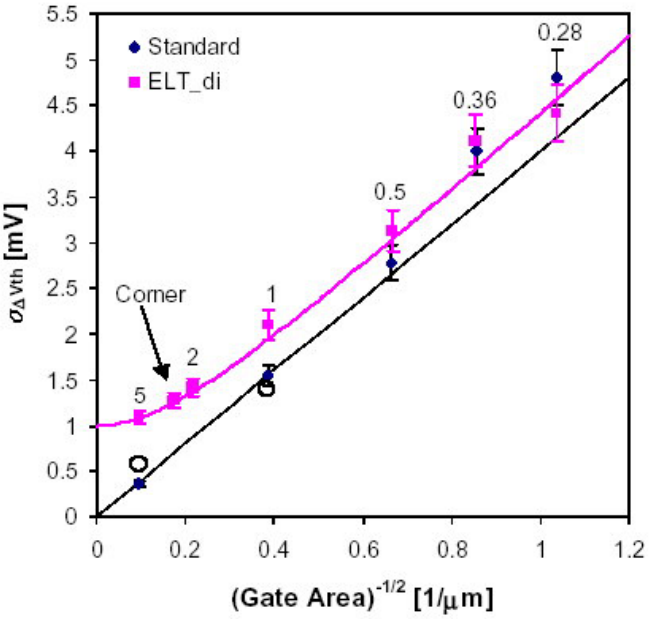

Figure 5.24: Threshold voltage matching performance of ELT pairs, measured with the drain inside the annular gate (ELT_di), compared with the matching performance of the standard n-channel devices. The numbers on the plot close to the points indicate the gate length. The two circles on the plot represent the $\sigma_{\Delta Vth}$ values for the pairs with four contacts in gate, measured with the drain inside the gate. The error bars of these two points are not indicated. "Corner" indicates the corner transistors described in the text [Ane00].

This is not true for ELT transistors that do not follow[29] equation (5.10) but can be fitted with a curve which has the following equation:

$$\sigma_{\Delta V_{th}} = \sqrt{\frac{A_{V_{th}}^2}{W L} + \sigma_0^2} \qquad (5.13)$$

which points to an additional source of mismatch independent from the causes that generate mismatch already considered in equation (5.10), and independent from the gate area. No proven physical explanation has been found for the phenomenon yet, even if the fact that the two transistors with a bigger inner diffusion (circles in Figure 5.24) lay on the same line of standard transistors can indicate a correlation with this area. The experimental fitting values for A_{Vth} and σ_0 are 4.3 mV·μm and 1 mV, respectively.

Another phenomenon which is still unexplained is shown in Figure 5.25. If the measurements are done reversing the drain and source with respect to Figure 5.24 the measurements still can be fitted with the (5.13) but with different values A_{Vth} and σ_0 (A_{Vth} = 4 mV·μm and σ_0 = 0.5 mV).

[29] In effect the points which represent the measurements done on ELTs are not on a straight line.

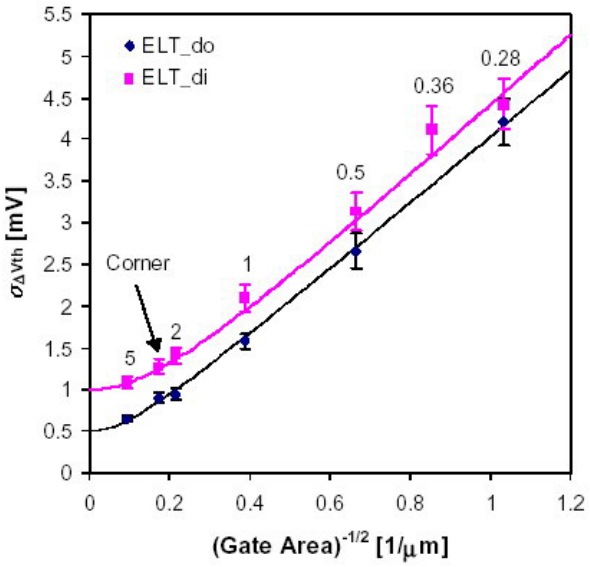

Figure 5.25: Comparison between the threshold voltage matching performance of ELT pairs measured with the drain inside (ELT_di) or outside (ELT_do) the annular gate. The numbers on the plot close to the points indicate the gate length [Ane00].

A similar effect is present also in the current factor matching. It can be seen that it does not follow equation (5.11) but have the same expression of equation (5.13); the parameters of the fit are A_β = 1.3% and $\sigma_{0\beta}$ = 0.3%. The current factor matching is not influenced by the choice of the drain diffusion (i.e. inside or outside the gate).

5.7.3 Noise performance of ELTs

Another very important characteristic to know for IC design is the noise performance of the target technology [Ane00, Ane01]. For this reason, measurements were performed on different types of transistors of our target technology to extract noise-related parameters. In particular, n-channel transistors were laid out as ELTs. The measurements were done on p- and n- channel transistors of different dimensions (W = 2000 μm[30]; L = 0.36, 0.5, 0.64, 078 and 1.2 μm) and in saturation in weak (I_D = 30 μA), moderate (I_D = 500 μA) and strong (I_D = 20 mA) inversion.

From transistors noise spectra measurements, the two 1/f white noise parameters K_a and α were extracted (equation I.1, in Appendix I). For what concerns the parameter α the measured values vary from .9 to .98 for n-channel transistors and from .8 to .9 for p-channel transistors. The 1/f noise constant K_a is plotted in Figure 5.26 as a function of the device gate length for n-channel and p-channel transistors in moderate (m.i.) and strong (s.i.) inversion. The values in weak inversion are not shown since they are very close to those in moderate inversion. It can be seen that K_a is higher for n-channel than for p-channel transistors, that it is

[30] The use of wide transistors helps obtaining more precise measurements.

higher for strong inversion than for weak inversion and that it increases with decreasing gate length.

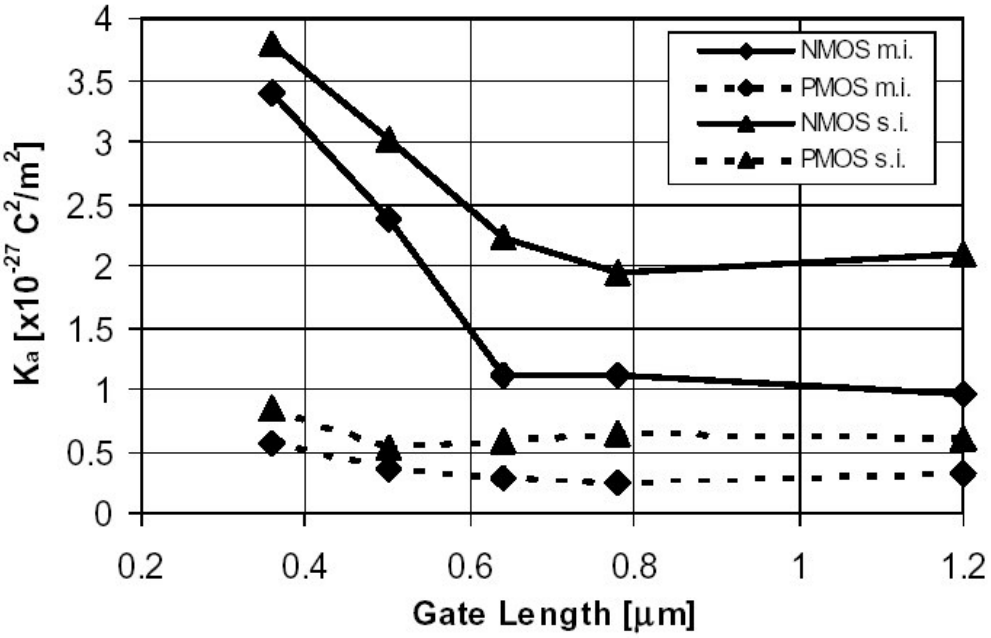

Figure 5.26: 1/f noise constant K_a as a function of the device gate length for n-channel and p-channel transistors in moderate (m.i.) and strong (s.i.) inversion. The values in weak inversion are not shown since they are very close to those in moderate inversion [Ane00].

The white noise excess factor (see section I.1, in Appendix I) was also measured. Figure 5.27 plots the noise excess factor Γ as a function of the device gate length for devices in weak, moderate and strong inversion. The values are calculated also taking into account the substrate resistance noise. The values for the p-channel transistors in moderate inversion are not well visible being almost coincident with those of the n-channel transistors in weak inversion.

The white noise is higher for n-channel transistors than for p-channel transistors and it increases with the current. Moreover, it tends to increase for shorter gate lengths. Both for the white noise and for the 1/f noise this can be attributed to imprecise or incomplete noise modelling.

Some measurements were also performed varying the drain to source voltage[31]. A small increase of the white noise (of the order of a few %) with the V_{DS} has been observed.

[31] For devices with L = 0.36 μm and L = 0.5 μm, in strong inversion and in saturation, with V_{DS} varying from 0.5 to 2.5V.

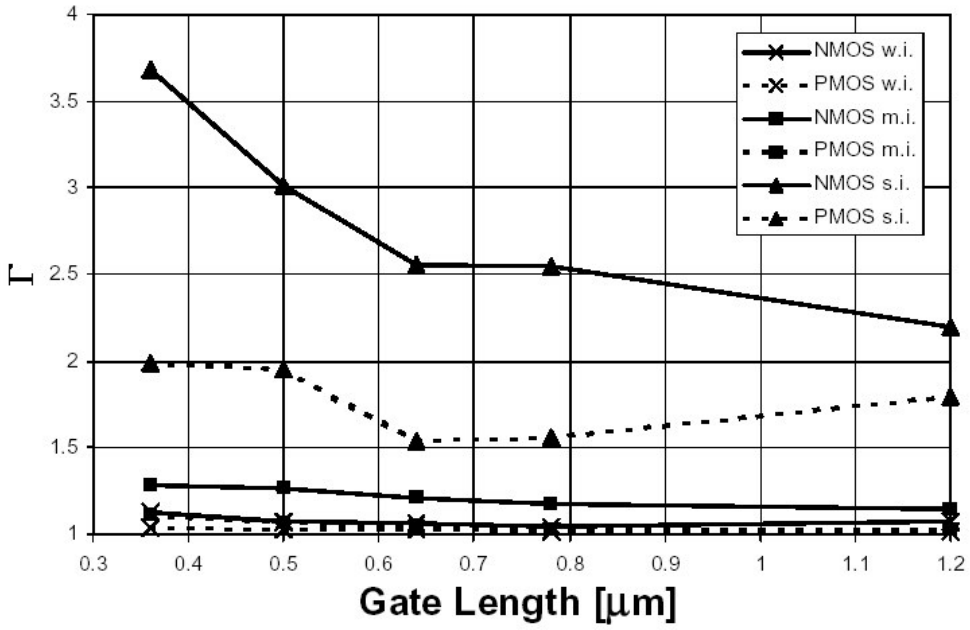

Figure 5.27: White noise excess factor as a function of the device gate length for devices in weak (w.i.), moderate (m.i.) and strong (s.i.) inversion. The values are calculated taking also into account the substrate resistance noise. The values for the p-channel transistors in moderate inversion are not well visible being almost coincident with those of the n-channel transistors in weak inversion [Ane00].

5.7.4 Radiation tolerance of ELTs

All the properties of ELTs have been checked during and after irradiation, to validate the possibility to use them in highly radiative environments.

The transistors (both ELTs and standard devices) were irradiated using 10KeV X-rays, with dose rates from 5 to 25 krd/min and a ^{60}Co γ-ray source, with a dose rate of 3 krd/min. All irradiations were performed under worst case bias[32] [Fac98, Ane99, Ane00], and transistors were measured also after an annealing of 24 h at room temperature and 168 h at 100°C (under worst case bias).

Figure 5.28 and Figure 5.29 summarize some important results. It can be seen that the standard n-channel transistor exhibits a very high leakage current already after 1 Mrd (SiO_2). On the contrary, the edgeless transistor does not show any leakage even after 30 Mrd (SiO_2).

Also threshold shifts are very small: for n-channel transistors the threshold shift is 15 mV after 10 Mrd (SiO_2) and 35 mV after 30 Mrd (SiO_2), measured immediately after irradiation. After 30 Mrd (SiO_2) and annealing, due to the creation of new interface states, ΔV_{th} increases to 45 mV. For for p-channel transistors the threshold shift is of -30 mV after 10 Mrd (SiO_2) and -70 mV after 30 Mrd (SiO_2), measured immediately after irradiation. After 30 Mrd (SiO_2)

[32] With worst case bias we mean the bias conditions in a circuit that maximize the radiation effects.

and annealing, due to the detrapping of charge trapped in the oxide, ΔV_{th} decreases (in absolute value) to -55 mV

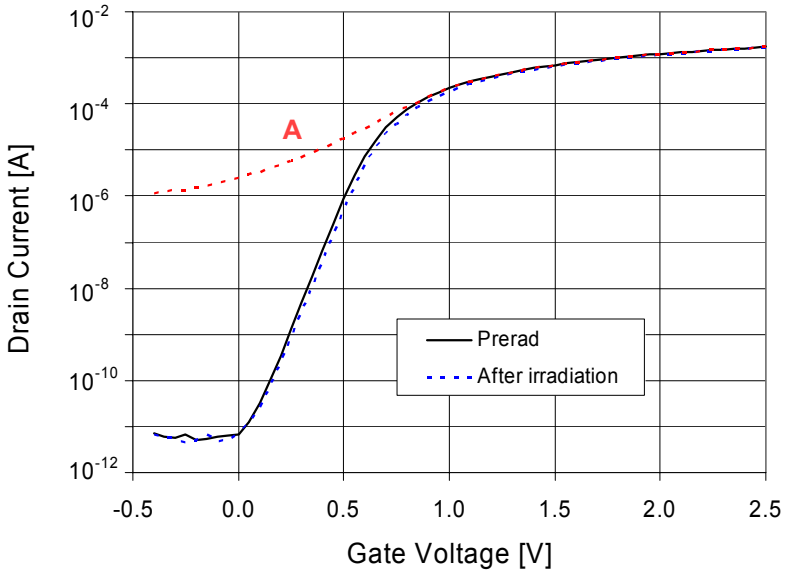

Figure 5.28: Drain current (log scale) before and after 30 Mrd (SiO_2) for an edgeless n-channel device (L = 0.28 µm) and drain current for a standard n-channel device (L = 0.28 µm) after 1Mrd (SiO_2) (curve A), as a function of the applied gate voltage [Ane99].

The curves which refer to the edgeless transistor show a minor degradation of all the parameters. In effect, on all the samples measured the transconductance and the mobility decrease is less than 6% (see Figure 5.29), and also the subthreshold swing and the output conductance change are in the order of a few percent.

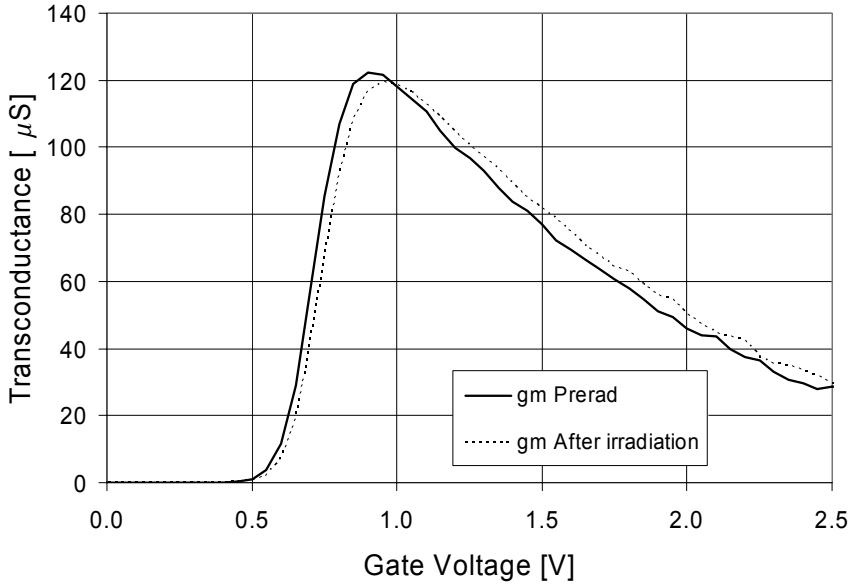

Figure 5.29: Transconductance before and after 30 Mrd (SiO_2) for an edgeless n-channel device (L = 0.28 µm) as a function of the applied gate voltage [Ane99].

For what concerns post-irradiation matching factors, irradiations were performed up to 1.5 Mrd (SiO_2). These measurements show that the current voltage matching coefficient A_β does not degrade after an irradiation of 1.5 Mrd (SiO_2). The threshold voltage matching coefficient A_{Vth} for p-channel transistors changes of a few percent, while for n-channel transistors[33] A_{Vth} increases up to 5.4 mV·μm (+45%). These measurements were affected by some technical problems, and are not fully coherent with the minor radiation-induced shift of all the other transistor parameters, both for p- and for n- channel transistors. However, the value of 5.4 mV·μm after 1.5 Mrd SiO_2 (which is a dose 6 times higher than expected for the inner layer of the ALICE SPD [Pas03]) is still reasonably good.

To better understand the behaviour of the noise performance of ELTs, two pairs of devices (L=0.5 and 0.78 μm) were irradiated up to 100 Mrd (SiO_2). The results on the noise parameter K_a are reported in Figure 5.30.

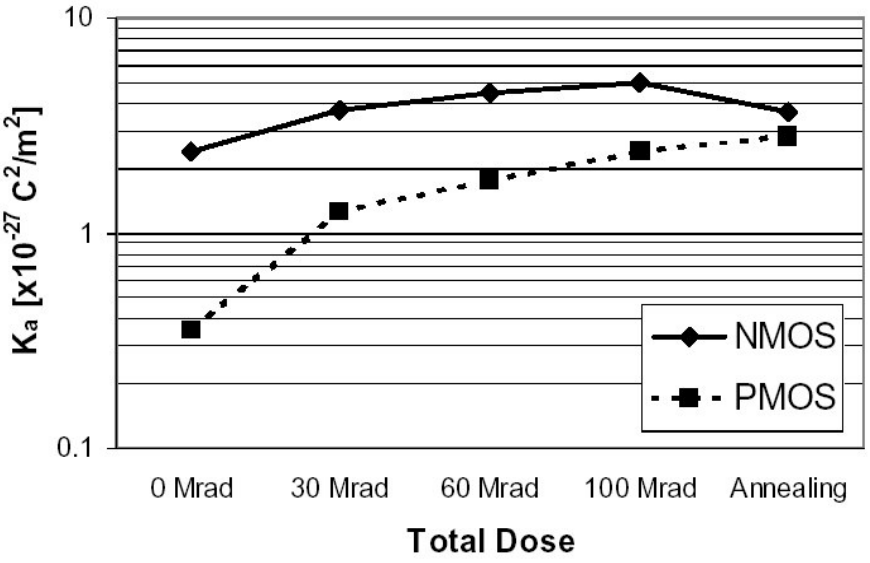

Figure 5.30: 1/f noise constant K_a after irradiation and annealing as a function of the total dose.

The white noise increase is small in all inversion regions: maximum 15% for n-channel and 7% for p-channel devices (the noise being expressed in $V/\sqrt{Hz}$); the noise levels after annealing stay roughly constant for the p-channel and increase of a few percent for the n-channel transistors.

5.8 HBD area and performance penalties

The use of HBD techniques reduces the component density that can be achieved on a given area, which results in some same area waste compared with standard layout.

[33] Measured on the special ELT shape with 4 contacts inside the annular gate.

For what concerns analogue design, the main density loss is related with long edgeless devices, so the area penalty is highly design-dependent.

For digital design, several attempts have been made to estimate the area and performance penalties related with the use of HBD techniques. The area loss can be described in two different ways. We can define the *percentage area penalty* (for a cell or for a circuit) as the ratio of the area needed to design the cell with HBD techniques or with default layout rules. We can also define the *generational area penalty* G_P as follows:

$$G_P = \frac{A_{HBD} - A_{ST}}{A_{PR,ST} - A_{ST}} \tag{5.14}$$

where A_{HBD} is the area needed to design the cell using HBD techniques; A_{ST} is the area needed to design the cell with standard layout and $A_{PR,ST}$ is the area needed to design the cell with standard layout in the previous technology node. The ratio is measured in "generations" (which can be abbreviated in Gen). Nowlin and co-workers evaluated the area and generational penalty associated with the design of a two input NOR cell at the 0.18 μm technology node (the previous technology node is the 0.25 μm node) for edgeless transistors and for enclosed source transistors [Now03]. The n-channel transistors have a W/L = 1.2 / 0.2 except in the design with edgeless devices, where W/L = 3.8 / 0.2. In all cases the p-channel devices are standard-edged with W/L = 1.2 / 0.2. The percentage and generational area penalty are evaluated respectively to 125% and 0.2 Gen for the enclosed source cell and to 175% and 0.7 Gen for the edgeless cell.

A similar analysis done on a two input NAND cell designed in 0.35 μm (the previous technology node is the 0.5 μm node) with edgeless transistors gives a percentage area penalty of 1.7 and a generational area penalty of 0.7 Gen, similar to what was found for the NOR cell [Lac00, Lac01].

For the same NAND cells also the performance penalty was evaluated. The performance comparison is presented in Table 5.5.

	0.35 μm	0.35 μm HBD	0.5 μm
Cell area [μm^2]	51	88	114
W/L (n)	5	12	5
W/L (p)	9	20	9
Supply voltage [V]	3.3	3.3	5
Propagation delay [ns] (fanout = 2)	0.1	0.09	0.14
Power dissipation [μW/MHz]	0.29	0.64	1.1
Power delay product [aJ/MHz]	38	80	168
Gates density [Mgates/cm^2]	1.96	1.13	0.87
Maximum operating frequency [MHz]	376	402	328
Throughpout per Watt [Mgates MHz/cm^2/μW]	6.82	1.75	0.79

Table 5.5: Comparison of the performance for three NAND cells: standard design 0.35 μm, HBD 0.35 μm and standard design 0.5 μm [Lac00].

The performance was evaluated under certain simplifying assumptions (e.g. the output is loaded by a capacitance equivalent to a fanout of two inputs). All the values of the comparison parameters for the 0.35 μm HBD cell are better than the corresponding values for the 0.5 μm standard cell.

Snapp [Sna03] evaluated the area and performance penalties related with the use of HBD techniques at the 0.25 μm technology node, using 77 standard cells available in the foundry commercial library, and full custom cells designed using ELTs and channel stops. The transistors in the custom library were sized to have similar propagation delays to the standard foundry cells. Under these assumptions Snapp found an average area penalty of about 2.6 and a power penalty from 2.5 to 3.5 (the power consumption comparison was limited to the inverter, NAND and NOR cells only).

At CERN a special radiation tolerant digital library was designed in a 0.25 μm technology using HBD techniques [Mar98, Klo98]; the cells from this library were compared with cells designed with standard layout and with cells designed in standard layout in a 0.6 μm technology[34]. The percentage area penalty was found to be between 1.5 and 3.5, but still the cells designed in 0.25 μm with HBD techniques were up to 3.2 times smaller than the cells designed with the 0.6 μm technology. Also performances of the HBD cells are still better than the 0.6 μm cells. For example, the delay of an inverter in 0.25 μm HBD biased at 2V with fan-out of 1 is 2.4 times less than that of the 0.6 μm standard design (biased at 3.3) and consumes 10 times less power.

Although with slightly different results, all these studies confirm that the area and performance penalty due to the use of HBD techniques is small enough (if compared to standard technologies) to encourage their use if compared with the available radiation-hardened technologies, which have always a worse performance (they are usually 2 or 3 generations behind commercial technologies).

5.9 SEE tests on systems employing HBD techniques

The effectiveness of the HBD approach has been tested at CERN [Fac98, Fac99, Fac99a] irradiating several types of shift registers with particles with a LET up to 89 $MeVcm^2mg^{-1}$.

The shift registers represent a convenient test vehicle to understand the SEU mechanism, as the test of the device and the interpretation of the results is quite simple.

Five different types of shift registers, based on different D-Flip Flops (DFFs), were designed in our target 0.25 μm technology and all of them were designed using ELTs and guard rings. One used the standard static architecture for the DFF; another type used a dynamic architecture and a third type was static but implemented using a dedicated SEU-

[34] This technology (2 generation older) was chosen to have comparable performance with the more advanced radiation hardened technologies available.

tolerant design ("hardened" cell). The standard static DFF was then modified, trying to improve its SEU-tolerance implementing two of the charge dissipation techniques discussed in section 5.6.2.2. One approach consists in oversizing some of the transistors to increase their driving capabilities and at the same time the node capacitance ("oversized" cell). The second approach consists in increasing the capacitance of the sensitive nodes adding a metal-to-metal capacitance on top of the cell ("overloaded" cell). In both cases the overall area of the DFF cell is unchanged with respect to the standard DFF.

5.9.1 SEL sensitivity

The SEL sensitivity of the first three types of shift registers[35] was tested irradiating them with heavy ions and monitoring the current consumption during irradiation [Fac98]. No SEL was observed during the whole irradiation campaign, at an applied supply voltage of 2.5 V, up to the maximum LET available of 89 $MeVcm^2mg^{-1}$.

A simulation study has shown that the maximum energy deposition occurring with some probability in the LHC radiation environment will correspond locally to a LET of about 15 $MeVcm^2mg^{-1}$ [Fac98a]. Therefore, the measured threshold for SEL indicates that latch-up will not be a threat in circuits designed using radiation tolerant layout.

5.9.2 SEU sensitivity

All the shift registers were irradiated with protons or heavy ions ad different LET values. For each test the errors were counted until the desired particle fluence was reached. The ratio between the errors and the fluence, normalized to the number of memory points (i.e. the number of DFF cells in the shift register) gives the *cross section* value σ for the given heavy ion LET. The measured σ is then plotted as a function of the particles LET. It is common practice to fit the experimental points with a Weibull curve, which has the following expression:

$$\sigma = \sigma_{sat}\left(1 - e^{-\left(\frac{LET - LET_{th}}{W}\right)^S}\right) \qquad (5.15)$$

where LET_{th} is the threshold under which the circuit is not sensitive to upsets, W and S are two shape factors without physical meaning. σ_{sat} is the saturation cross-section, that is the value at which the cross-section does not change anymore with the LET of the incoming particle, and is related with the total SEU-sensitive area of the memory cell.

The static register was tested in two modes of operation. In the "unclocked mode", a pattern was written in the register at a frequency of 30 MHz, then the clock was stopped

[35] The static, dynamic and hardened.

during a variable irradiation time until another train of clock pulses was applied to read out the pattern for error detection. In the "clocked" mode the clock was constantly applied during the whole irradiation test at a frequency between 460 KHz and 30 MHz[36].

Figure 5.31 shows the measurement results.

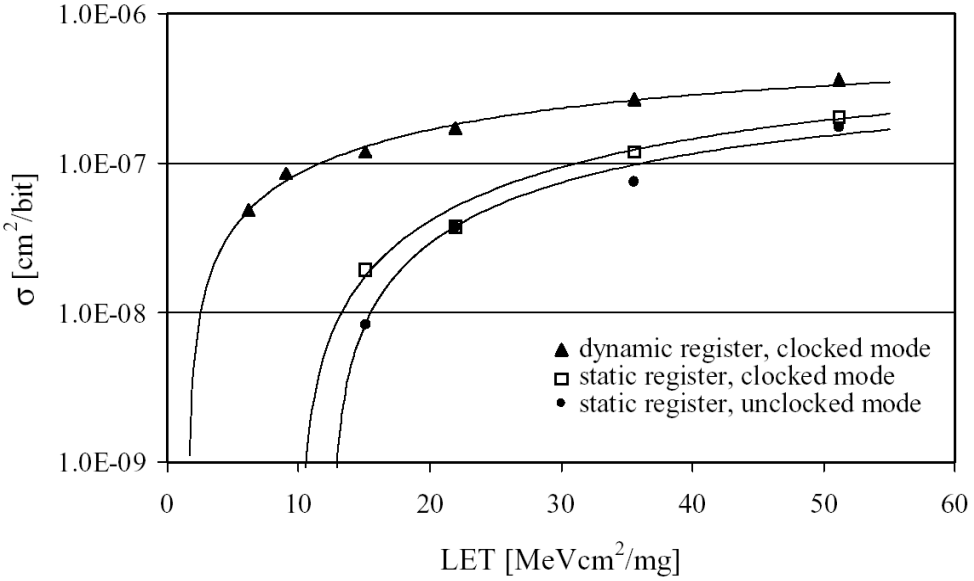

Figure 5.31: Measured cross-section for dynamic and static shift registers. Experimental data are fitted with a Weibull curve in all cases. The frequency of the clocked test was 30 MHz [Fac99].

The static register has a much lower sensitivity (higher threshold) if compared to the dynamic one, and is more SEU-tolerant if operated in the unclocked mode. The LET_{th} for the static register in unclocked mode is 15 $MeVcm^2mg^{-1}$; this value is rather high for a 0.25 μm design. This difference is attributed to the increased transistor size and transistor driving capabilities which are unavoidable in HBD designs using ELTs, which then have an intrinsic higher charge dissipation capability.

Figure 5.32 shows the results obtained irradiating the “overloaded” and the “oversized” cells, compared with the standard static register. Both techniques improve SEU-tolerance, but the “overloading” is the more effective. For the “hardened” memory cell a threshold LET of around 89 $MeVcm^2mg^{-1}$ has been measured at a power supply voltage of 2 V.

[36] For the “dynamic” shift register the clocked mode is the only one applicable.

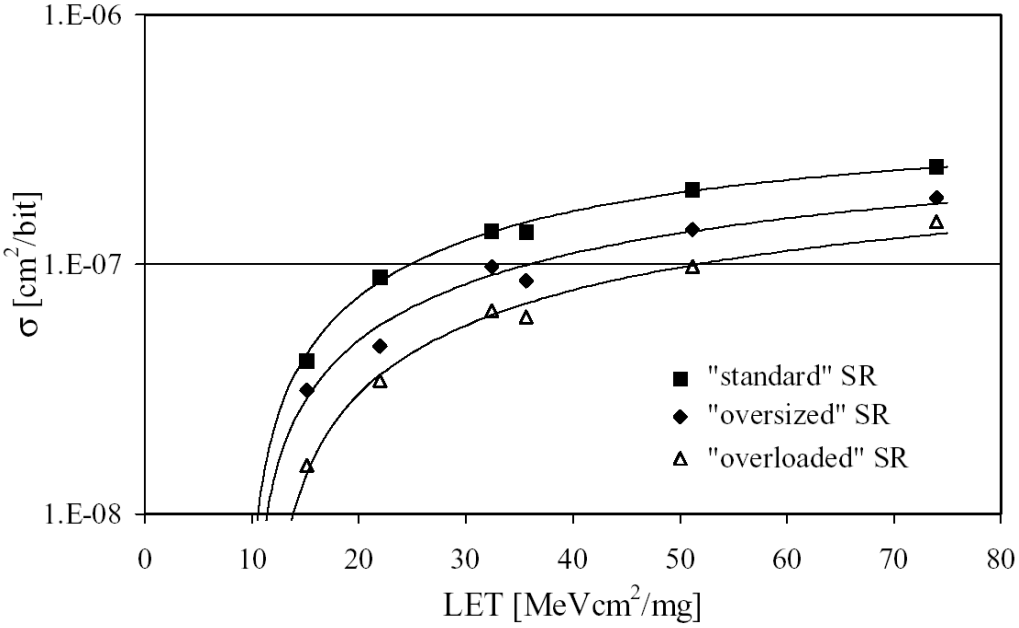

Figure 5.32: Measured cross-sections for different static DFFs irradiated with heavy ions in the clocked mode [Fac99].

5.10 Summary

In the LHC experiments the radiation levels that can be attained are very high, especially for the inner detectors, very close to the interaction point. The working values for the ALICE SPD are 500 krd for TID and $6 \cdot 10^{12}$ MeV n_{eq} / cm^2 for the total fluence. This means that the readout electronics has to be able to stand this radiation doses. This chapter dealt with the effects of radiation on MOS devices and on the solutions available for designing radiation tolerant ICs for High Energy Physics, as well as their effectiveness.

The physical effects of radiation on MOS devices can be grouped in two classes: ionisation effects and nuclear displacement. MOS devices are almost insensitive to atomic displacement damages, while are sensitive to ionising radiation. When an ionising particle passes through a MOS structure it produces electron-hole pairs; holes can damage the silicon oxide. Holes migrate in the oxide from the generation point to the SiO_2-Si interface; if they do not tunnel out of silicon oxide they may be trapped close to the interface, giving origin to a fixed positive charge in the oxide. Another effect of radiation on MOS devices is the increase by several orders of magnitude of the trap density at the interface SiO_2-Si. Interface traps can capture or emit electrons or holes.

These two phenomena generate two components of the total radiation-induced threshold shift ΔV_T. The subthreshold slope is not affected by the presence of oxide-trapped charge in the device, but decreases with the introduction of interface trapped charge. The threshold

voltage shift and the subthreshold slope variation induce a change in current which flows in the transistor when no gate bias is applied (the off-state current). Another detrimental effect of radiation is the damage in the oxide which separates transistors (field oxide). A relatively small dose in a field oxide can induce sufficient charge trapping to cause field-oxide induced IC failure, both for drain-to-source parasitic currents and for inter-transistor parasitic currents. Some other transistor parameters can be degraded by radiation, such as the transconductance, the noise behaviour and the matching characteristics.

Other detrimental effects that can be induced by radiations are the Single Event Effects (SEEs). SEEs are generated by highly energetic particles crossing a sensitive part of a transistor or of an integrated circuit. They are usually divided in *soft errors* (which are reversible and non-destructive, as they do not cause a permanent damage to the device) and *hard errors* (which are non-reversible and can be destructive under certain conditions). The most important soft error is the Single Event Upset (SEU). A SEU is generated by a highly energetic particle which passes through a memory cell, if it induces enough charge (directly or by secondary interactions) to cause an instantaneous and reversible change of the logic state of the cell. Single Event Transients refer to errors that can result from an energetic particle strike on non-latched elements. In some planar CMOS technologies there is a parasitic thyristor which can be switched on by electrical transients, high temperature and improper sequencing of power supply biases, but also by a highly energetic ionising particle. The thyristor shorts the power supply and destruct the device. This phenomenon is called Single Event Latchup (SEL). A Single Event Gate Rupture is the destruction of the gate oxide of a MOS device due to a highly energetic particle.

To increase the radiation tolerance of a circuit, it is possible to operate with two different approaches: at foundry (process) level, and at design level with Hardening By Design (HBD) techniques.

Hardening by process consists of addressing the problem of radiation tolerance at the level of the technological process; in particular some specialised vendors provide a qualified radiation hard process. These technologies suffer from problems related with cost, circuit performance, process stability and yield. A major problem related with radiation-hardened processes is their future availability: due to the drop of demand some radiation hardened technologies processes will be discontinued.

The basic principle HBD is to use a readily available, low cost CMOS technology and to apply some special layout and design techniques to increase radiation tolerance. The main advantage of this approach is the low cost, and the possibility to follow the rapid scaling down of commercial technologies.

A third possibility to obtain radiation tolerance at chip level consists in selecting adequate Components Off The Shelf (COTs) and test them, setting up a database of selected standard electronic components which have been characterised and qualified as robust against

radiation effects. The advantages of this approach include a large variety of available components, high performance and low component costs. The disadvantages include short availability lifecycles, the possibility of process changes that can change the radiation characteristics, lack of traceability, costs associated with qualification and the inability to get reliable data from the manufacturer. Moreover, for some applications an available COTS device with the required radiation tolerance could not be available.

The HBD approach has been used for the design of the ALICE1LHCb chip. For this reason it was presented in detail, starting with the impact of technology scaling both for what concerns radiation induced effects and circuit performance.

Moving from a MOS generation to the next one the gate oxide thickness t_{ox} is reduced. As showed by Saks [Sak84], the total radiation-induced threshold voltage shift decreases with $1/t_{ox}^2$ for oxides down to ~20 μm, then the dependence becomes much faster. The target technology CMOS 0.25μm chosen for the design of the ALICE1LHCb chip has an oxide thickness of about 5.5 nm, so the radiation induced threshold voltage shift is negligible up to several Mrd of radiation dose.

Off state currents are decreased by technology scaling, but not eliminated, while no relevant change in the subthreshold slope is induced by radiation in submicron technologies.

All the SEEs are attenuated (or eliminated) in submicron technologies, except SEUs, because SEU sensitivity is highly technology and design dependent.

The most important effects of technology scaling on circuit performance were also presented.

The main problem that has to be solved by adequate Hardening By Layout techniques is the increase of the parasitic off-state currents. The solution to the problem of inter-transistor leakage consists of adding a p^+ guard ring (also called channel stop) around each n-region at a different potential than the negative power supply, and around any n-well at a different potential than the positive power supply (this for example happens for p-channel transistors with the source connected to the bulk). The p^+ guard ring stops any possible leakage towards any n-well at a different potential.

Some transistor topologies reducing or eliminating edge leakage were presented; among them there is the Enclosed Layout Transistor (ELT), that consists of drawing an annular gate in the active region so that the drain (or the source) is inside and the source (or the drain) is outside. Although ELTs have several drawbacks, such as bigger area, additional capacitance and asymmetry of the device, and most of all limitation in the achievable aspect ratio, this solution is very efficient for radiation hardness.

Several techniques can also be used at system level to increase radiation tolerance, and in particular to improve immunity to SEEs, and were described in section 5.6.2.

ELTs were extensively used for the ALICE1LHCb pixel chip design, so their characteristics, which were studied at CERN in the framework of the RD49 project, were

presented in detail in section 5.7. A precise modelling of one of the possible ELT shapes was explained, which fits measurements quite precisely.

Matching of ELTs has a slightly different equation from standard layout transistors (SLTs). For SLTs matching improves with transistor size, while for ELTs it saturates at a certain value. Apart from that, matching parameters are similar to those of SLTs, as shown in Table 5.6.

	n SLT	n ELT (drain inside)	n ELT (drain outside)	p SLT
V_{th} Matching equation	$\sigma^2_{\Delta V_{th}} = \frac{A^2_{V_{th}}}{W\,L}$	$\sigma_{\Delta V_{th}} = \sqrt{\frac{A^2_{V_{th}}}{W\,L} + \sigma^2_0}$	$\sigma_{\Delta V_{th}} = \sqrt{\frac{A^2_{V_{th}}}{W\,L} + \sigma^2_0}$	$\sigma^2_{\Delta V_{th}} = \frac{A^2_{V_{th}}}{W\,L}$
$\Delta\beta/\beta$ Matching equation	$\sigma^2_{\Delta\beta/\beta} = \frac{A^2_{\beta}}{W\,L}$	$\sigma_{\Delta V_{th}} = \sqrt{\frac{A^2_{\beta}}{W\,L} + \sigma^2_{0\beta}}$	$\sigma_{\Delta V_{th}} = \sqrt{\frac{A^2_{\beta}}{W\,L} + \sigma^2_{0\beta}}$	$\sigma^2_{\Delta\beta/\beta} = \frac{A^2_{\beta}}{W\,L}$
A_{Vth}	4 mV·µm	4.3 mV·µm	4 mV·µm	3.7 mV·µm
σ_0	-	1 mV	0.5 mV	-
A_β	1 %·µm	1.3 %·µm	1.3 %·µm	1 %·µm
$\sigma_{0\beta}$	-	0.3 %	0.3 %	-

Table 5.6: Matching performance of n- and p-channel Standard Layout Transistors (SLTs) and Edgeless transistors (ELTs).

Another very important characteristic to know for IC design is the noise performance of the target technology. Measurements were performed on different types of transistors of our target technology to extract noise-related parameters; results are summarised in Table 5.7. The parameters are a function of the transistor length L and are higher for a shorter L.

	n ELT s.i.	n ELT m.i.	n ELT w.i.	p SLT s.i.	p SLT m.i.	p SLT w.i.
α	.9 to .98	.9 to .98	.9 to .98	.8 to .9	.8 to .9	.8 to .9
K_a $[10^{-27}C^2/m^2]$	3.8 to 2	3.4 to 1	3.4 to 1	0.8 to 0.6	0.6 to 0.4	0.6 to 0.4
Γ	3.7 to 2.2	1.3 to 1.2	1.2 to 1.1	2 to 1.6	1.2 to 1.1	1

Table 5.7: Noise performance of n- and p-channel Standard Layout Transistors (SLTs) and Edgeless transistors (ELTs) in strong (s.i.) moderate (m.i.) and weak (w.i.) inversion.

The transistors (both ELTs and standard devices) were irradiated using 10KeV X-rays, and a ^{60}Co γ-ray source. All irradiations were performed under worst case bias, and transistors were measured also after annealing. Edgeless transistors do not show any leakage even after 30 Mrd (SiO_2). On all the samples measured the transconductance and the mobility decrease

is less than 6%, and also the subthreshold swing and the output conductance change are in the order of a few percent. The current voltage matching coefficient A_β does not degrade after an irradiation of 1.5 Mrd (SiO_2). The threshold voltage matching coefficient A_{Vth} for p-channel transistors changes of a few percent, while for n-channel transistors A_{Vth} increases up to 5.4 mV·μm (+45%). The white noise increase is small in all inversion regions: maximum 15% for n-channel and 7% for p-channel devices.

The 1/f noise constant K_a increases more with irradiation, around a factor 2 for n-channel transistors and a factor 8 for p-channel transistors (after 100 Mrd (SiO_2)).

The use of HBD techniques reduces the component density that can be achieved on a given area, and increases the area consumption. This performance penalty was evaluated both for single cells and for full systems. Although results are rather design dependent, a rule of thumb of a power penalty of about 2 and area penalty of 2.5-3 can be given. It is important to point out that in all the cases under study the performance of the HBD design was always superior to the corresponding design in rad-hard technologies.

The effectiveness of the HBD approach has been tested at CERN irradiating several types of shift registers with particle with a LET up to 89 $MeVcm^2mg^{-1}$. No SEL was observed during the whole irradiation campaign. The irradiation results for SEUs show a better tolerance of HBD than standard layout designs to irradiation, as well as a better tolerance of static structures versus dynamic ones, and that loading a cell with more output capacitance is more effective than increasing the transistor sizes to prevent SEUs.

Chapter 6
The ALICE1test and ALICE2test prototype chips

The journey towards the design and production of working hybrid pixel detectors for high energy physics started in the 1990's with, for example, the use of the Omega2 and Omega3 chips in the CERN WA97 experiment, installed in the Omega spectrometer. Although fully working, their main limitation was related to radiation tolerance, as they started to suffer severe degradation at a total ionising dose of some tens of krd.

The front-end electronics for LHC experiments has to stand much higher radiation doses, and this lead to the RD49 project to study radiation tolerance of submicron CMOS technologies, and to the design of the first chips in a commercial CMOS technology, but with radiation tolerant layout techniques.

6.1 The Omega2 and Omega3 chips

The Omega2 and Omega3 chips were implemented in a commercial 3 and 1 μm CMOS technology, respectively. The **Omega2** [Cam94] chip was the first chip developed by the RD-19 collaboration which was installed in a full pixel system in an experiment [Ang92].

It contains a matrix of 63 rows × 16 columns of 75 × 500 μm^2 active pixels. Each front-end cell contains about 80 transistors with a single metal layer available for interconnect. The charge preamplifier and comparator are followed by a digital delay line, a coincidence unit and readout logic. If the delayed output produces a pulse when the strobe signal is high, the hit is latched for readout. The readout is done by configuring the data registers in one column as a shift register, and by shifting the data down to the bottom of the column. The output of the delay line is used to reset the front-end.

The delay line is implemented using a chain of current deprived invertors, where the limiting current defines the delay of the delay line. Any variation of this current across the chip (due to transistor threshold or power supply variations) will cause corresponding delay variations.

The top cell in every column has an electrical test input to inject a well-defined charge into the front end. The rest of the chip is inaccessible for testing. The single metal layer made it impossible to provide a metal shield on top of the readout electronics to avoid parasitic coupling into the detector. To prevent the front end from having to absorb the full detector

leakage current, a dummy detector cell at the bottom of each column was implemented to measure the leakage current. This current was then reproduced in every pixel with the opposite sign. For uniform detector leakage current, no current has to be absorbed by the front-end itself.

The **Omega3-LHC1** chip [Hei96] was a further improvement after the Omega2: it was implemented in a 1 μm technology with two metal layers, which allowed a reduced pixel size (50×500 μm^2, instead of 75×500 μm^2), and this with more than four times the number of transistors (~380 per pixel). The matrix consists of 16 columns and 128 rows, so twice the number of pixels as Omega2. Every cell is accessible by an electrical test input, and the content of a test flip–flop in the pixel determines whether the analogue test pulse is applied to the pixel or not. A mask flip–flop within each pixel allows switching off the pixel in case it is noisy or defective.

The delay line in the pixel is again based on current deprived invertors, but the number of stages was increased from 3 to 36. The reset for the front-end is the output of the fourth stage in the delay line. In addition, a three bit delay adjust, also based on current deprived invertors, was added to the delay. This allowed to fine-tune the delay of every pixel individually to obtain a more uniform delay across the chip (30 ns instead of 180 ns peak to peak), or across the system (for an array of 24 chips 75 ns peak to peak instead of 220 ns) [Can98].

Due to the availability of three metal layers, the top metal could be used as a shield between the readout electronics and the detector. However, after a first production, the chip showed a large top-down variation of the pixel charge threshold due to the resistive drop in the power supply of one of the bias lines along the column. This was diagnosed by using the possibility to individually address every pixel for electrical test, and corrected by taking some part of the metal shield for the power supply of this bias line to decrease the resistive drop.

Another additional feature in the Omega3 chip is the ‘fast OR’. When the discriminator in a pixel fires, a current pulse is immediately sent to the bottom of the column, where a low input impedance circuitry detects it and converts it to a digital signal. This signal is then logically OR-ed for the full chip and sent to an output pad. This pad hence indicates whether the chip detected a hit immediately after it occurred, and gives the possibility to run this chip in a ‘self-triggering’ mode. This signal is also useful for diagnosis, as any activity in the chip can be detected without a full readout.

6.2 The design of the ALICE1test and ALICE2test chips

The first test chip designed in a CMOS technology with HBD techniques was the **ALICE1test** chip, a prototype chip for the future ALICE silicon pixel detector. It was designed in the 0.5 μm Mietec technology, and proved the feasibility of a full mixed-mode chip with the radiation tolerant layout approach. In effect, as will be explained in section 6.3, it stood a TID between 600 krd and 1.7 Mrd depending on the type of radiation [Sno00]. No

radiation-induced leakage current was observed, and the circuit failed ultimately because of the cumulative effects of radiation-induced threshold shift which, for a technology with an oxide thickness of about 10 nm, is still in the range of about 100 mV/Mrd. Transistor irradiation measurements indicate that without special layout precautions the chip would have died at about 50 krd due to excessive power consumption from radiation induced leakage.

Although these were very encouraging results, the density penalty of these layout techniques proved to be too large for application in the ALICE pixel detector. The possibility for CERN to access a 0.25 μm CMOS technology allowed the design and production of the **ALICE2test** chip [Sno98, Cam99, Sno00a], based on a design almost identical to the previous chip, but taking profit of all the beneficial effects of a deep-submicron technology. The only relevant design differences with ALICE1test are the discriminator, that was redesigned, and the implementation of a digital delay line in between the discriminator and the chip logic. The chip is a matrix of two columns each containing 65 identical cells, it occupies 10 mm^2, and contains about 50000 transistors.

The circuit builds on the experience obtained with previous pixel readout chips [Cam90, Ang92, Cam94, Hei96]. It is a matrix of two columns each containing 65 identical cells. The chip occupies 10 mm^2, and contains about 50 000 transistors. The readout cell is shown in Figure 6.1.

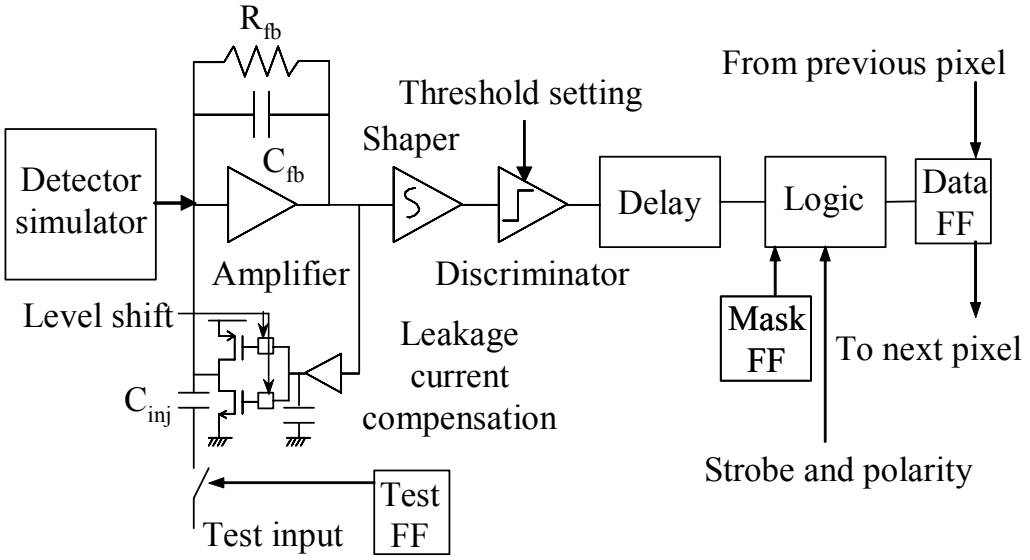

Figure 6.1: Schematic representation of the front-end cell of the ALICE2test chip. It comprises a preamplifier, a shaper filter, a comparator, a delay line and readout logic. The only relevant design differences with ALICE1test are the discriminator (redesigned) and the delay line (added).

The circuit, which can work both with positive and negative input charge, comprises a preamplifier, a shaper filter, a comparator, a delay line and readout logic. As this prototype was not intended for bump-bonding to a detector, an input structure has been added to each cell to simulate detector capacitance, coupling between pixels, and detector leakage current.

The preamplifier is a capacitively fed back operational transconductance amplifier, or integrator. The preamplifier feedback circuit is a modification of a circuit first proposed by Krummenacher [Kru91], shown in Figure 6.2 (a), to allow both polarities of leakage current. It allows setting DC values of preamplifier input and output independently, and it has a low frequency feedback which adjusts itself to absorb the leakage current coming from the detector. The scheme was modified to absorb both signs of leakage current by introducing both a p- and an n-channel MOS instead of only an n-channel MOS. To avoid extra loading of the preamplifier output node, the feedback is linked to the source of the input transistor of the shaper (Figure 6.2 (b)), which follows the preamplifier output at frequencies below $(2\,\pi\,\tau)^{-1}$ where τ is the shaper time constant.

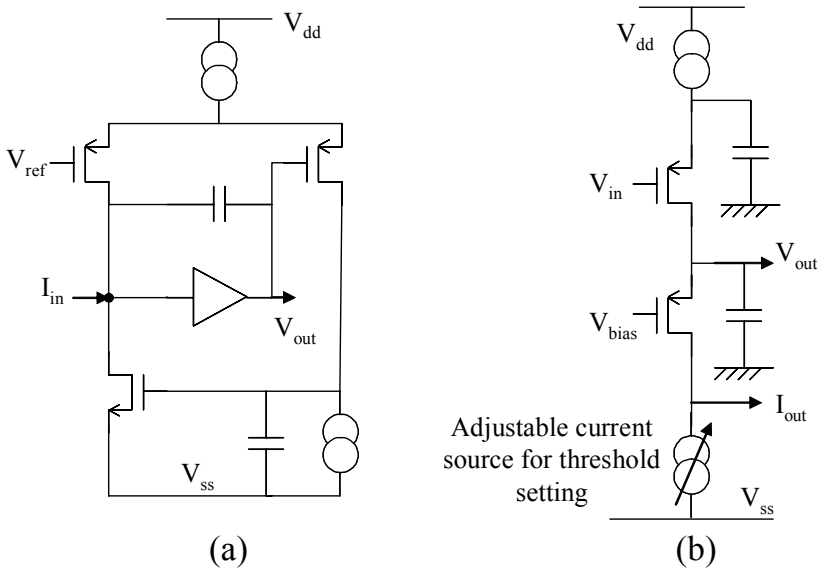

Figure 6.2: (a) schematic representation of the Krummenacher scheme [Kru91] on which are based the ALICE1test and ALICE2test front-ends. The low frequency feedback path provided by the NMOS linked to the input allows potential detector leakage to be absorbed by the NMOS and not by the differential pair. In the ALICE2test chip the scheme has been modified to be able to absorb both leakage current polarities. (b) Shaper configuration for the ALICE1test chip. The n-channel based adjustable current mirror at the bottom, used to set the pixel charge threshold, was a major source of charge threshold spread, and was removed in the ALICE2test version.

The feedback consists of a higher frequency part (time constant around 1 microsecond, still much slower than the 25 ns of the shaper time constant τ), and a very low-frequency part near dc. The higher frequency part is implemented as follows: if a voltage imbalance occurs at the differential pair controlled by a reference voltage Vref and the voltage on V_{out} (or the shaper output for ALICE2test), a corrective current is injected into the input of the preamplifier. To avoid stability problems the differential pair is operated at relatively high current (200 nA), but the noise contribution of the feedback is contained by sending only part of the output current of the differential pair (10 nA) to the preamplifier input. The noise

contribution of the feedback is therefore limited to the noise of the transistor conducting those 10 nA.

For what concerns the low frequency part of the amplifier feedback, the output of the preamplifier is amplified and filtered (by a high capacitance on a high impedance node) to make a low pass filter. This node controls the gate of an n- and p-channel transistor which have their drain connected to the preamplifier input, and therefore controls the current to the preamplifier input. So, if the output of the preamplifier starts to drift due to a change in detector leakage current, the voltage on the high impedance node will change, and the current in the n- and p-channel devices will change to absorb the change in leakage current. This scheme has the advantage that not all detector leakage current has to pass through the feedback resistor. A level shift prevents the detector-leakage-current-absorbing n- and p-channel transistors from turning on together to avoid an unnecessary noise increase. Since this leakage current compensation circuit acts at frequencies much lower than the frequency of the shaper pulse, detector leakage current (although it will add noise) should not alter the shape of this pulse, and hence not the pixel charge threshold.

The shaper used for these two chips is shown in Figure 6.2 (b). It provides both a current output (drain of second transistor) and a voltage output (source of second transistor). In the 0.5 μm test chip it was used with the current output [Sno00] while in the 0.25 μm the voltage output was used with some further modification to increase the shaper gain [Sno00a].

In the ALICE1test chip, the threshold is defined by a current which is subtracted from the shaper output current. The resulting current is sent to a current comparator which has its threshold set at zero. The problem with this is that the pixel charge threshold in this scheme critically depends on the accuracy of a current defined by an NMOS current mirror, which was made very large to obtain good current matching, and which was laid out in edgeless geometry for radiation tolerance. However, as we have seen in Chapter 5, the matching of enclosed geometry devices is worse in some cases than for traditional devices (Figures 5.24 and 5.25; problems related to the design of n-channel transistor current mirrors will be explained in more detail in the next chapter). Because increasing the size of this current mirror (which already occupied ~90×40 μm^2) would not yield an improvement, the voltage output of the shaper was taken for the 0.25 μm design.

The comparator has a 3 bit threshold fine-adjust. In the ALICE2test version of the chip it is controlled by a 3 bit bus directly linked to the outside. In the complete version of the chip, the ALICE1LHCb, the fine adjust is controlled by flip-flops implemented on the pixel.

The delay element consists of an 8 bit counter and some control logic. A flag tells the delay control logic which polarity of the comparator output corresponds to a logic one, depending on whether one collects positive or negative input charge. If the comparator fires the counter in the delay is started. The counter can be preset with an arbitrary value, and the carry of the most significant bit is used to generate the “end of count signal”. If the “end of

count" signal is in coincidence with the externally applied strobe or trigger, a one is written into the data flip-flop. In order to compare different designs of the counter, and in particular its SEE sensitivity, it was implemented in static logic in one column and in dynamic logic in the other. The clock is propagated up the column using full swing differential CMOS logic.

Changing the test flip-flop pattern allows to address one or several pixels simultaneously in an arbitrary way during testing. The content of the test flip-flop determines whether or not an analogue input signal is applied to the preamplifier input across an injection capacitance. This injects a known amount of charge in the preamplifier, emulating the detector, and allowing an electrical test of the chip even if it is not physically connected to a sensor. The capacitance was realised as a parasitic metal-to-metal capacitance. A mask flip-flop allows the disabling of a pixel in case it is noisy or completely non-functional.

Each cell consumes ~50 μW and the chip operates on an analogue supply voltage of 2.2 V and a digital supply voltage of 1.6 V. The pitch in the short dimension is 50 μm as this was the dimension specified for the experiment. The cell layout is based on the ALICE1test chip, implemented in the 0.5 μm technology. The smaller technology feature size and the elimination of the large current mirror used to set the pixel threshold reduced the dimensions of the front end from 265×50 μm^2 to about 125×50 μm^2.

The space made available this way was used to implement a counter-based synchronous delay, consisting of static logic in the left column, and dynamic logic in the right column. The counter in the delay takes 40 by 60 μm^2 for the static case and 40 by 35 μm^2 for the dynamic case. The delay control logic measures about 20 by 25 μm^2. The rest of the digital part was directly taken from the 0.5 μm chip and was not shrunk to 0.25 μm design rules.

6.3 Experimental results form the ALICE1test and ALICE2test chips

As not all the pixels in the chip have an analogue test output, the threshold value and the noise value for each pixel have to be extrapolated by the chip digital output. The method, described in the next section, makes use of the so called *s-curve* and is based on the possibility to inject in the front-end a known, variable input charge. Since the two chips were not bump-bonded to a sensor, the injection capacitance could not be calibrated, but only estimated from data on layer-to-layer capacitances provided by the vendor. All numbers given in absolute electron charge (e^-) are based on the estimate that each millivolt pulsed across the injection capacitance ($C_{inj} = 16$ fF) is equivalent to about 100 electrons injected in the preamplifier.[1]

[1] From $\Delta Q_{Coulombs} = C \cdot \Delta V$ we can derive $1.6 \cdot 10^{-19}\, \Delta Q_{electrons} = C\, \Delta V$. The injection capacitance implemented on the chip is $C_{inj} = 16$ fF, so: $\Delta Q_{electrons} = 10^5\, \Delta V$.

6.3.1 The s-curve method

In binary chips[2] the analogue information about the signal amplitude is lost after discrimination. The consequence is that there is no means to evaluate the pixel noise with analogue techniques at the shaper output. The information has to be deducted by the pixel digital output, for example with the s-curve method.

An s-curve (Figure 6.3) shows, for a particular threshold setting, the number of times that a given pixel, pulsed up to several thousand times with the same input charge Q_{inj}, responds to an increasing value of the input charge. To form an s-curve, a fixed threshold is chosen. The input charge is scanned across a range of values (x axis). For each value, the pixels are pulsed (triggered) many times (~ 100) and the number of digital output hits is recorded for each pixel. This, divided by the number of triggers, gives the efficiency of the pixel for this value of input charge (y axis).

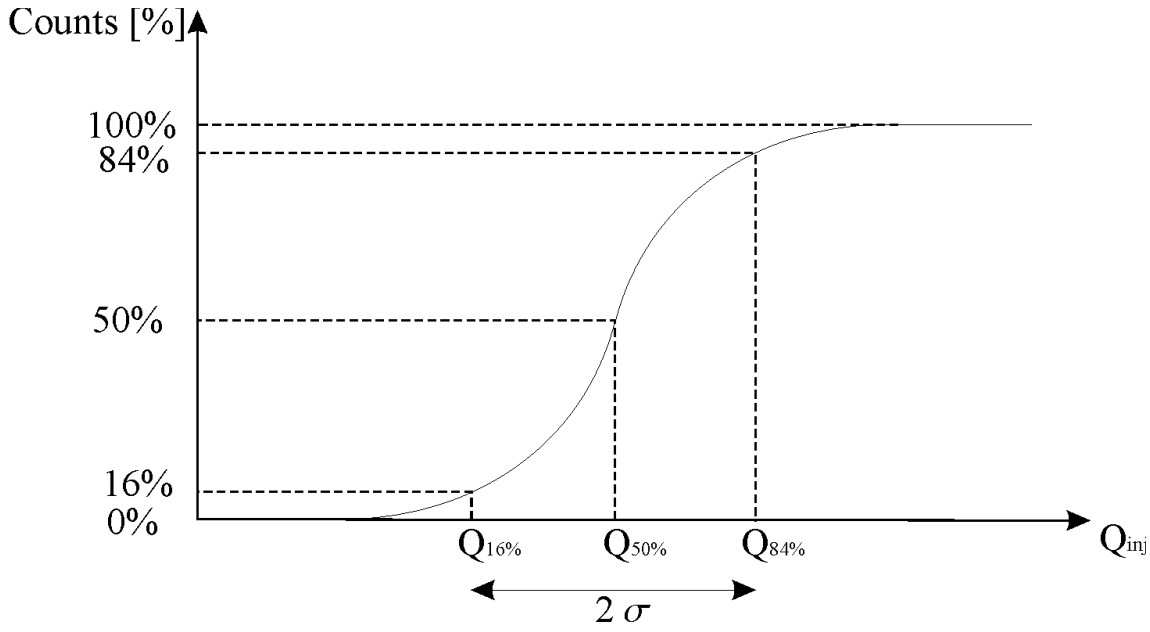

Figure 6.3: Example of an s-curve; the value $Q_{16\%}$-$Q_{84\%}$ is an estimation of 2 σ, while $Q_{50\%}$ is an estimation of Q_{Th}.

For input charges quite below the threshold charge, no hit is recorded. For input charges sufficiently above the threshold charge, each input pulse fires the discriminator, so each hit is recorded. In the region close to the threshold, the electronic noise which is superimposed on the preamplifier output pulse can fire the discriminator even if the input charge is below the threshold or, on the contrary, prevent the discriminator to fire even if the input charge is

[2] With binary chips we mean chips where the analogue front-end is followed by a discriminator, so that the only hit information that is retained is if a hit was present or not in each time slot.

above threshold. This shows that the slope of the s-curve is related with electronic noise, while the point at 50% is an estimate of the pixel threshold, as shown in Figure 6.3.

We can suppose that the charge amplitude distribution $Q_{indisc}(Q_{inj}, \sigma)$ seen at the discriminator input is Gaussian, with a mean value Q_{inj} and variance σ^2. If the number of pulses is high enough, the probability that $Q_{indisc} > Q_{inj}$ is given by the complement to one[3] of the cumulative probability of the Gaussian distribution $Q_{indisc}(Q_{inj},\sigma)$.

This means that if we count the number of events in which $Q_{indisc} > Q_{inj}$, we are estimating the cumulative probability function of $Q_{indisc}(Q_{inj} ,\sigma)$. Since we know the shape of the cumulative probability function of a Gaussian distributed process, we can estimate the variance σ^2 from the s-curve. In effect, $Erf(\sigma) \sim 0.16$ and $1\text{-}Erf(\sigma) \sim 0.84$. As a consequence, if we measure the ΔQ on the s-curve from where the 16% of hits occur and where the 84% of the hits occurs we have an estimation[4] of 2 σ.

Examples of measured s-curves are shown in Figure 6.5 (b) and Figure 6.10.

6.3.2 Timewalk estimation

Another important parameter to evaluate in binary chips is the timewalk of the circuit, which is mainly related to the discriminator performance. The response delay of the discriminator can be rather different for an input signal which is slightly above threshold than for an input signal which is much higher than the threshold. If the time delay between the slowest and the fastest discriminator output pulse is too high, some hits can be wrongly synchronised and be recorded by the on-chip logic as having occurred in the subsequent clock cycle. To measure the timewalk, the input threshold charge has to be measured (Q_{th} in Figure 6.4). For a signal infinitesimally smaller than Q_{th} the response time of the discriminator can be considered as infinite. If we apply increasing input charge signals, the discriminator response time decreases, and then tends to saturate to a minimum value d_{min}. The value at which the discriminator response time is d_{min}+25 ns is indicated as Q_{25} in Figure 6.4. The value $Q_{25} - Q_{th}$ is called *timewalk* (at 25 ns).

The timewalk can be defined at any delay; for what concerns LHCb the bunch crossing and the system clock have a period of 25 ns, so that timewalk at 25 ns is strictly related to the possibility of a small input charge pulse falling in the subsequent clock cycle. In particular, for the design of the ALICE1LHCb chip we decided to use a more stringent definition of timewalk (at 20 ns) to have 5 ns contingency for the propagation delay of the signals across the chip.

[3] The complement to one of x is 1-x.
[4] Slightly different algorithms can be used to extract the noise, but all of them are based on this principle.

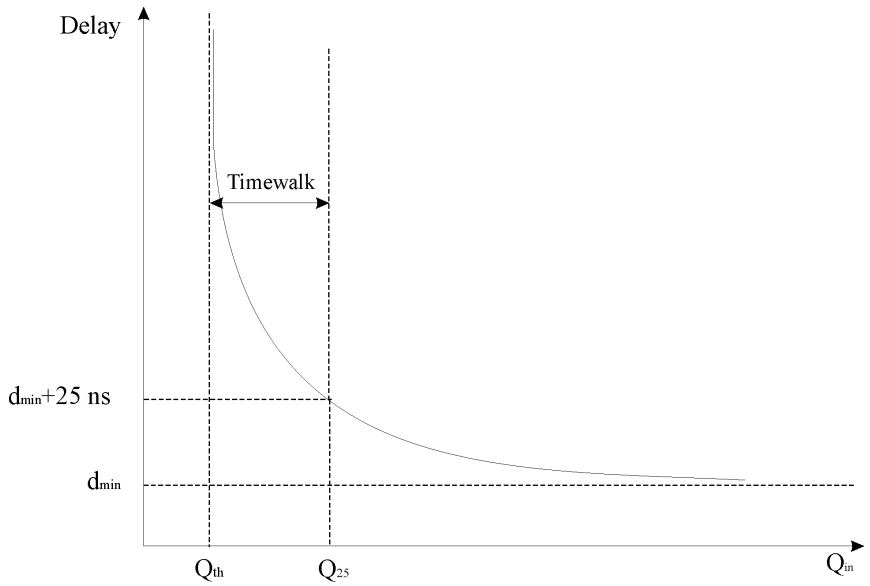

Figure 6.4: Time walk effect. If we apply increasing input charge signals above the threshold charge Q_{th}, the discriminator response time decreases and then tends to saturate to a minimum value d_{min}. The value at which the discriminator response time is d_{min}+25 ns is indicated as Q_{25} in the picture. The value Q_{25} - Q_{th} is called timewalk (at 25 ns).

6.3.3 Experimental results on the ALICE1test prototype chip

Figure 6.5 (a) shows how the average threshold charge (over all 130 cells) can be changed by the threshold setting voltage, for the ALICE1test chip.

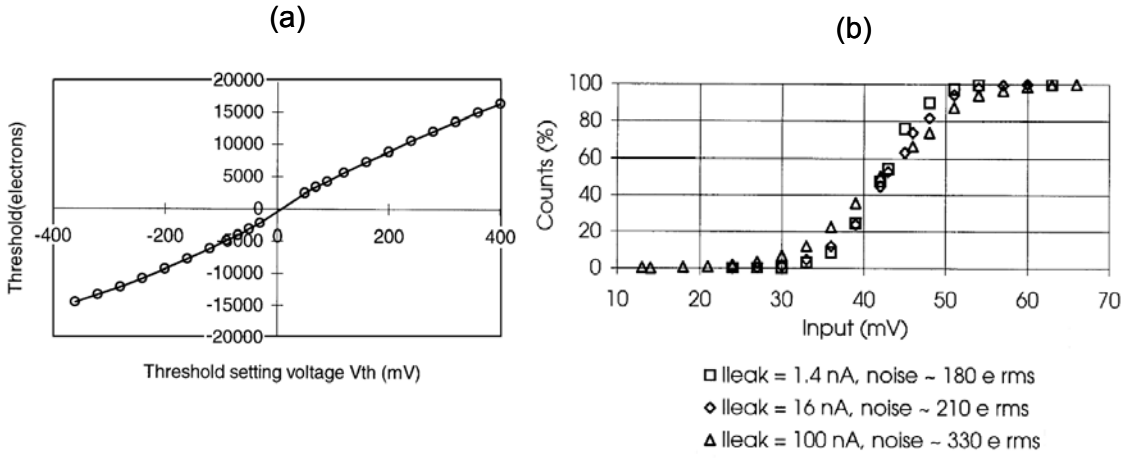

Figure 6.5: ALICE1test chip. (a) Relation between the average threshold charge (over all 130 cells) and the threshold setting voltage. Some non-linearity related to the response of the differential pair controlling the threshold is visible. (b) Example of s-curves for the ALICE1test chip, with different leakage currents injected in the preamplifier.

Some non-linearity related to the response of the differential pair controlling the threshold is visible. The observed threshold spread of 400-500 e^- rms[5] over the full chip is higher than expected. This is probably due to a poor modelling of the mismatch of the large n-channel MOS devices in the comparator current mirror. Solutions which avoid current mirrors based on large n-channel MOS have been implemented in the subsequent chips.

Figure 6.5 (b) shows an example of s-curves for the ALICE1test chip, with different leakage currents injected in the preamplifier. The threshold did not vary by more than 1% for leakage currents varying from -200 to +200 nA. The average noise is about 200 e^- rms at low-detector leakage current. The input structure adds about 100 fF to the preamp input to simulate the detector capacitance. Care was taken to inject the leakage current into the frontend with the proper noise spectral density (2qI) for a detector. A negative 200 nA detector leakage current is absorbed by a long narrow p-channel transistor, and increases the average noise to 350 e^- rms. Because a positive detector leakage current is absorbed by a short wide edgeless n-channel in weak inversion the noise increases to 400 e^- rms at 200 nA.

The timewalk performance of the ALICE1test discriminator is sufficiently good: due to the fast shaping all hits more than a few hundred electrons above threshold fall within a 25 ns time window.

The chip has been irradiated with X-rays, γ-rays, high-energy particles and protons, to measure the radiation tolerance of the chip for different irradiation sources.

As a first measurement of the radiation tolerance of the chip X-ray irradiations were carried out using a dedicated machine (Seifert RP149). The target material used in the tube was Tungsten and the X-ray energy peaked at 10 keV. The dose rate was 4 krd(SiO_2)/min.

The pixel comparator threshold and its spread start to degrade significantly only after a TID of 600 krd. The analogue power consumption remains unchanged during irradiation, and the digital power consumption decreases. The latter can be explained by the radiation induced threshold shifts. This indicates that on a full circuit scale enclosed n-channel transistors and guard rings prevent radiation induced leakage.

The gamma irradiation was carried out at the National Health Institute[6] in Rome, Italy, using a standard source (Gammacell 220) of 1.173 and 1.332 MeV γ-rays from ^{60}Co. In this case the dose rate was 610 rd/min. The evolution of the comparator threshold and its dispersion with accumulated dose is similar to the one observed for X-rays. Severe degradation sets in at about 1 Mrd where the chip is anyhow still fully functional. At a total accumulated dose of 1.5 Mrd only 10% of the pixels respond. Partial recovery happened during the annealing. The supply currents showed a similar behaviour as for the X-ray irradiations.

[5] Electrons root mean square
[6] Istituto Superiore di Sanita', (ISS), Roma, Italia.

For what concerns high energy particle irradiation, the chip was irradiated primarily by electrons with an energy of 1 MeV or above. Also in this case the behaviour is quite similar to that after X-rays irradiation, and a severe degradation of the threshold, threshold spread, electronic noise and non-responding pixels sets in at 1.7 Mrd. After an irradiation of 2.6 Mrd no pixels were responding below 20000 electrons, but again with annealing partial recovery of the chip can be induced.

The tolerance to charged hadronic particles was investigated using 6.5 MeV protons at the Van de Graaf accelerator in the National Laboratory of Legnaro, Italy. The estimated dose rate is about 12 krd/min. Serious degradation again occurred only above an estimated TID of about1 Mrd and partial recovery was seen during the annealing phase.

6.3.4 Experimental results on the ALICE2test prototype chip

The availability for CERN of the more advanced 0.25 μm technology lead to the design of the ALICE2test chip, based on the previous prototype but with a redisigned comparator. A photograph of the chip is shown in Figure 6.6.

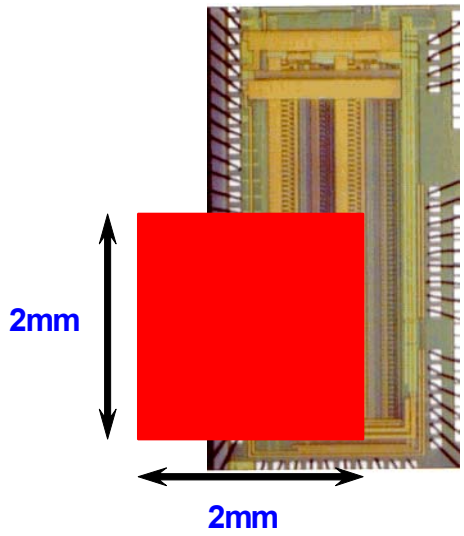

Figure 6.6: Die photograph of the 130-channels chip ALICE2test; The superimposed square shows the approximate shape and position of the proton beam during irradiation tests.

6.3.4.1 Pre-irradiation tests

As explained at the beginning of section 6.3, for this chip the injection capacitance could not be calibrated, but only estimated. The estimate is the same as for the ALICE1test chip (i.e. each millivolt pulsed across the injection capacitance is equivalent to about 100 electrons injected in the preamplifier). The chip was characterised electrically prior to irradiation, and is fully functional. As for ALICE1test, it was possible to set the global threshold of the chip with a dedicated bias (V_{th}).

Figure 6.7 shows how the average pixel threshold can be linearly adjusted using an external bias voltage from about -20000 e^- to +20000 e^-. The nonlinearity for higher threshold values is due to the clipping of the signal in the preamplifier to avoid circuit saturation and excessive recovery times for large input signals. The minimum threshold, ~1500 electrons[7], is determined by crosstalk between the analogue and digital parts of the circuit.

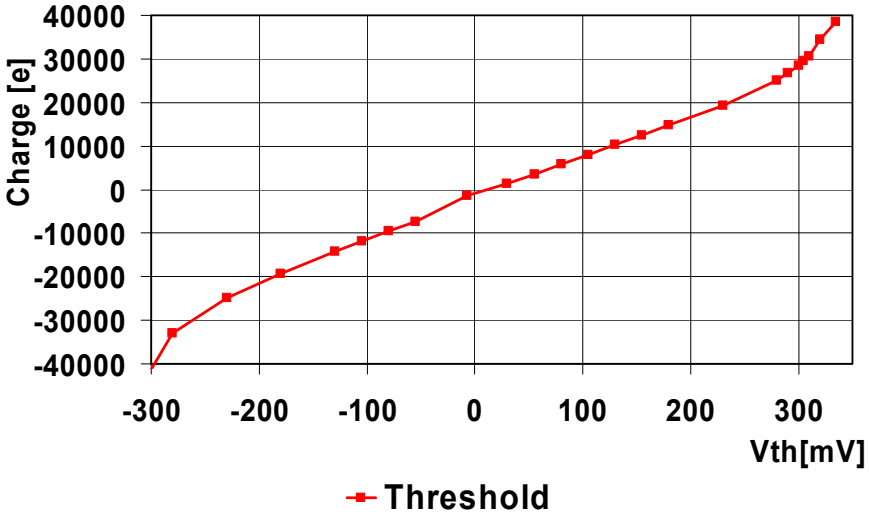

Figure 6.7: Evolution of the average pixel threshold as a function of the bias V_{th} controlling the threshold. The nonlinearity for higher threshold values is due to the clipping of the signal in the preamplifier.

Figure 6.8 shows how the threshold varies across the chip.

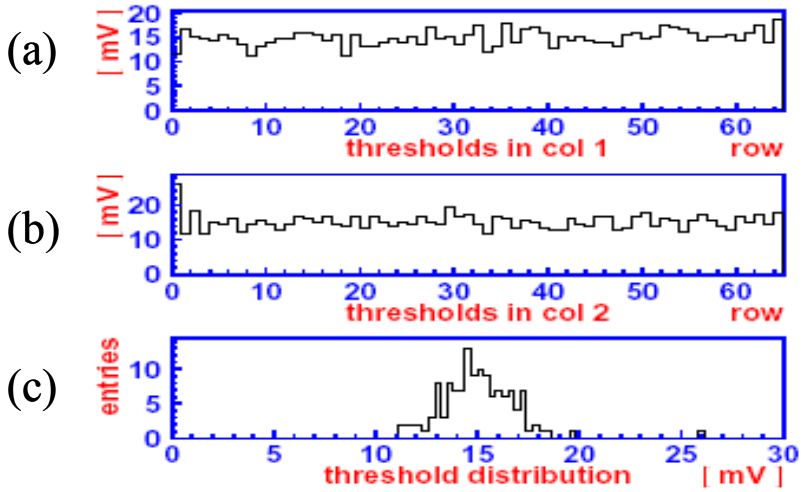

Figure 6.8: Distribution of the pixel threshold for all 130 pixels. (a) and (b) show the threshold of the pixels in column 1 and 2 respectively, as a function of the position (row number). In (c) is presented the threshold distribution. For all the plots 1 mV corresponds to 100 electrons input charge.

[7] As minimum threshold we define the minimum value of the threshold at which no spurious hits due to noise are detected by the full chip matrix.

Figure 6.8 (a) and (b) show the threshold of the all the pixels in column 1 and 2 respectively, as a function of the position (row number).

Figure 6.8 (c) shows the threshold distribution for all the pixels in the chip. There is no systematic dependence of the threshold on the position of the pixel within the chip. At this V_{th} threshold setting the average threshold charge over all 130 cells is about 1500 e^- and the spread is 160 e^- rms. The pixel noise is ~220 e^- rms. The behaviour is almost identical for both polarities of the circuit.

Figure 6.9 shows the evolution of the channel-to-channel rms variation of the average threshold shown in Figure 6.7, i.e. the σ of the Gaussian fitting the threshold distribution shown in Figure 6.8 (c), as a function of the threshold bias voltage V_{th}.

Figure 6.9 shows also the evolution of the average pixel noise as a function of the threshold bias voltage V_{th}. For -200 mV < V_{th} < +200 mV both threshold dispersion and pixel noise show little sensitivity to V_{th}.

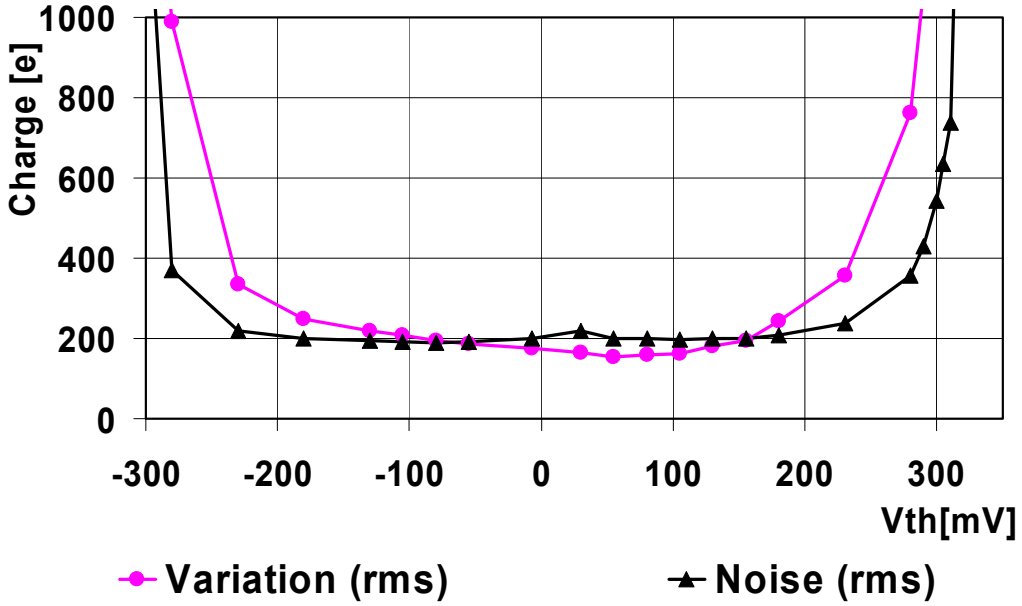

Figure 6.9: Evolution of the channel-to-channel rms variation of the average threshold shown in Figure 6.7, and of the average pixel noise, as a function of the threshold bias voltage V_{th}.

As previously mentioned, the charge threshold of every individual pixel can be fine-tuned using a 3-bit DAC controlled by a 3-bit register in each pixel. The pixel charge threshold can be extracted for every setting of this register by using the electrical input to the pixel to generate an s-curve like the one shown in Figure 6.10. It shows the s-curves for two different values of injected detector leakage current. A curve fit yields 2690 electrons for the threshold (50% point) in both cases and 220 and 250 electrons rms for the noise.

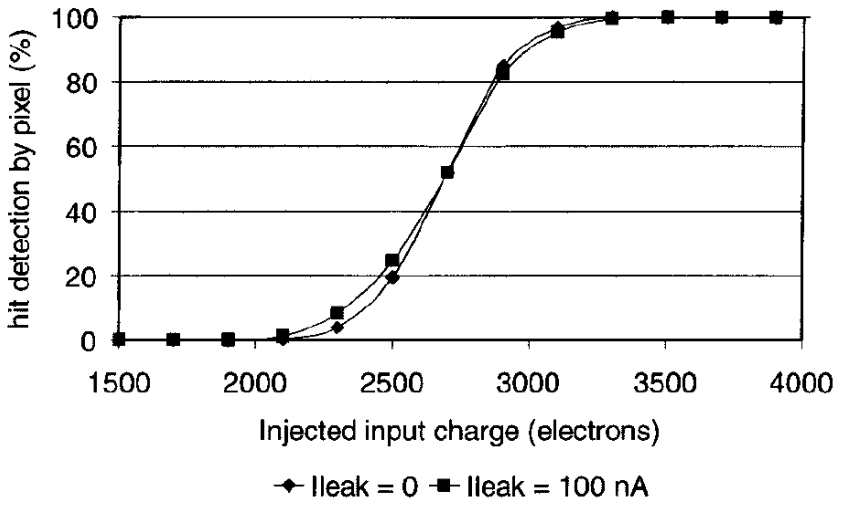

Figure 6.10: Pixel response in percent of input signal counts versus signal amplitude (s-curves) for two different values of injected detector leakage current. A curve fit yields 2690 electrons for the threshold (50% point) in both cases and 220 and 250 electrons rms for the noise.

Figure 6.11 shows three threshold distributions: one with the register controlling the threshold adjust DAC set to zero, one with the register set to 7 (maximum for three bits), and one with the optimal DAC setting to minimizes the pixel charge threshold spread.

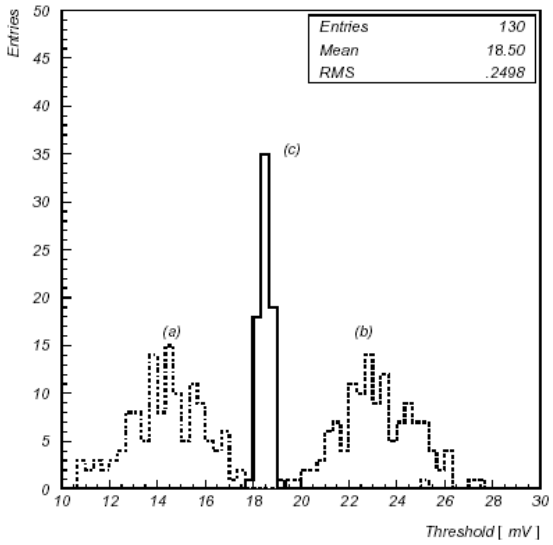

Figure 6.11: Distribution of the threshold for (a) minimum added threshold, (b) maximum added threshold, and (c) tuned threshold. As usual, 1mV corresponds to ~100 e^-.

The shift between the first two shows the total range of the adjustment (which itself can be modified externally).

The plot demonstrates that the threshold adjust can reduce the threshold spread from 160 e^- rms down to 25 e^- rms. Further refinement of the tuning algorithm may lead to even smaller values of threshold variation.

The sensitivity of the circuit behaviour to detector leakage current was measured. As can be seen in Figure 6.12, there was no change in average threshold for leakage currents of +/-100 nA per pixel. Threshold variation was almost unchanged and noise degraded by less than 20%.

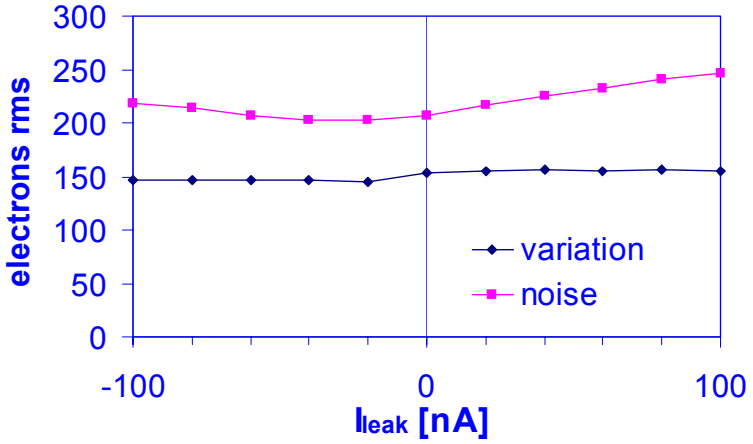

Figure 6.12: Noise (top curve) and threshold variation (bottom curve) for different injected leakage currents.

Pixel detectors normally have a larger capacitance between neighbouring elements than to ground. The input structure has the option of connecting neighbouring pixels together using a 30 fF capacitor or connecting the input node to ground using a 60 fF capacitor. In order to verify the sensitivity of the circuit to capacitive cross-coupling neighbouring pixels were connected together, the average threshold of the array was set to 1500 e^- and a pixel was stimulated whilst its neighbour was observed. On average it was necessary to inject 29000 e^- in one pixel in order to produce a false hit in the neighbour. The measurement was repeated without the coupling capacitors and this time 30500 e^- were necessary for the neighbour to react. This indicates that the front-end is rather insensitive to capacitive cross-coupling and that probably much of what was measured was due to parasitic effects in the electrical injection or coupling through power supplies.

6.3.4.2 Irradiation tests

The chip has been irradiated with X-rays, γ-rays, high-energy particles and protons.

X-rays

The irradiation was done on the dedicated Seifert RP149 X-ray machine, in the same conditions described for the irradiation of the ALICE1test chip. Figure 6.13 shows the evolution of the power supply currents with increasing X-ray dose. The absence of any increase in power consumption with total dose confirms once more on a full circuit scale that the use of enclosed n-channel transistors and guard rings prevents radiation induced leakage.

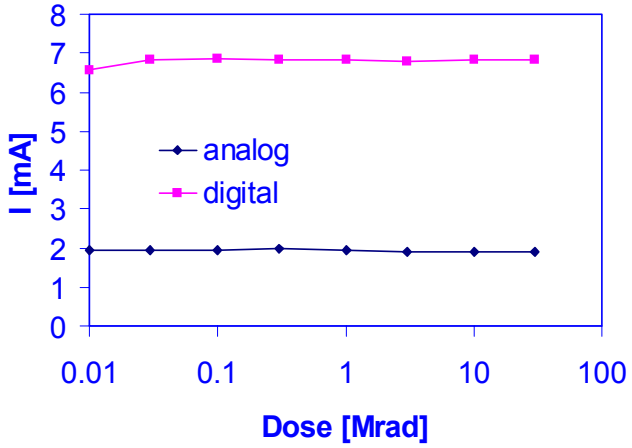

Figure 6.13: The evolution of the analogue and digital power supply currents with X-ray dose. Note the logarithmic scale on the x-axis.

Figure 6.14 shows the evolution of the average pixel threshold, the threshold variation and the pixel noise for the same irradiation. For this particular chip a minor bias adjustment was necessary after 30 Mrd(SiO_2) to prevent premature signal clipping in the preamplifier. Apart from this all other biases were kept constant. These results illustrate that the chip remains fully functional up to 30 Mrd(SiO_2).

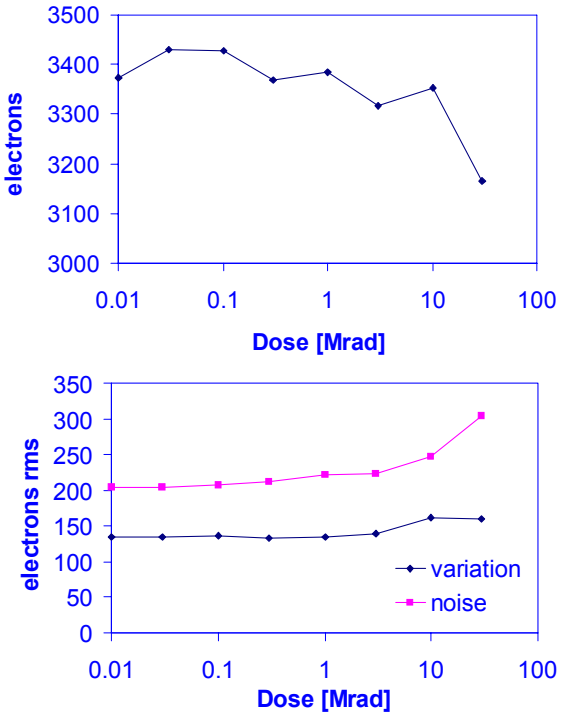

Figure 6.14: The upper plot shows the evolution of the average pixel threshold with total X-Ray dose. The lower plot indicates how threshold variation and noise evolve with total X-Ray dose. Note the logarithmic scale on the x-axis.

After 24 hours under bias at room temperature the parameters were unchanged. Following a subsequent anneal for one week at 100°C the average threshold remained the same, the threshold variation degraded slightly to 190 e^- rms, and the pixel noise returned to 230 e^- rms. Further annealing under bias at room temperature did not affect the circuit parameters.

High energy protons

A further test was made with high energy protons at the NA50 experiment of the CERN SPS machine. Figure 6.6 indicates how the chip was placed in the high intensity beam of 450 GeV/*c* protons, with the beam focussed on a roughly square 2 mm × 2 mm area. In total, the chip received 3.6×10^{13} protons, over a 12 hour period, representing an equivalent of 9×10^{14} protons/cm^2 in the target area.

The chip was kept under bias the whole time and read out between spills of the machine. Figure 6.15 shows the evolution of the thresholds during irradiation and anneal. During irradiation the threshold of the hit pixels was reduced and the noise increased to ~1000 e^- rms by the end of exposure.

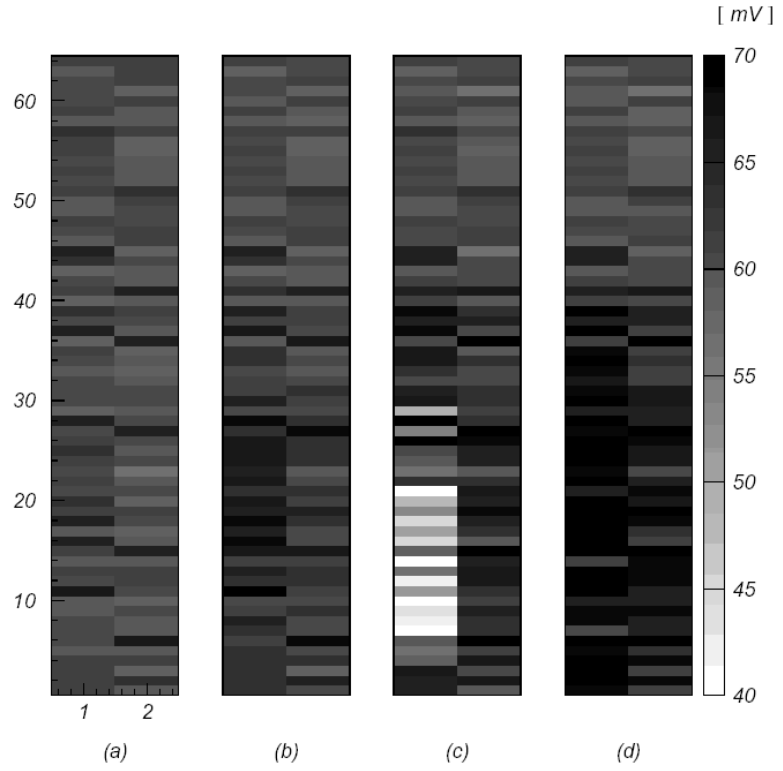

Figure 6.15: Pixel thresholds during and after the proton irradiation. (a) Thresholds before irradiation, (b) thresholds after 8×10^{12} protons, (c) thresholds after 6×10^{13} protons and 4 hour anneal, (d) thresholds after 6×10^{13} protons and 20 hour anneal. 1 mV corresponds to 100 e^-.

During annealing at room temperature, the threshold recovered and even increased slightly, whilst the noise returned to its pre-irradiation value. The pixels outside the target region remained unchanged throughout the test. In addition, there was no increase in power consumption.

Charged hadronic particles

As for the earlier prototype chip, the tolerance to charged hadronic particles was investigated using 6.5 MeV protons at the Van de Graaf accelerator in the National Laboratory of Legnaro, Italy. Chips were irradiated with doses of up to 48 Mrd. One chip was irradiated in steps to 9, 19 and 48 Mrd. It ceased to function at 48 Mrd. A second chip was irradiated to 37 Mrd. The evolution of the power consumption with the dose is shown in Figure 6.16. The analogue outputs and the power consumption showed a behaviour similar to that observed in the other measurements.

Gamma-rays

The chip was also gamma-ray irradiated. The irradiation was carried out at the National Health Institute in Rome, Italy, using the same source used for the irradiation of the Alice1Test chip. The dose rate was 540 rd/min. The chip was irradiated in steps to doses of 3, 19, 23 and 26 Mrd. The results for the power consumption are indicated in Figure 6.16. In this case, a slight increase in the analogue power supply was recorded. For the other parameters, the results of this irradiation closely mirrored those of the X-ray irradiation discussed above.

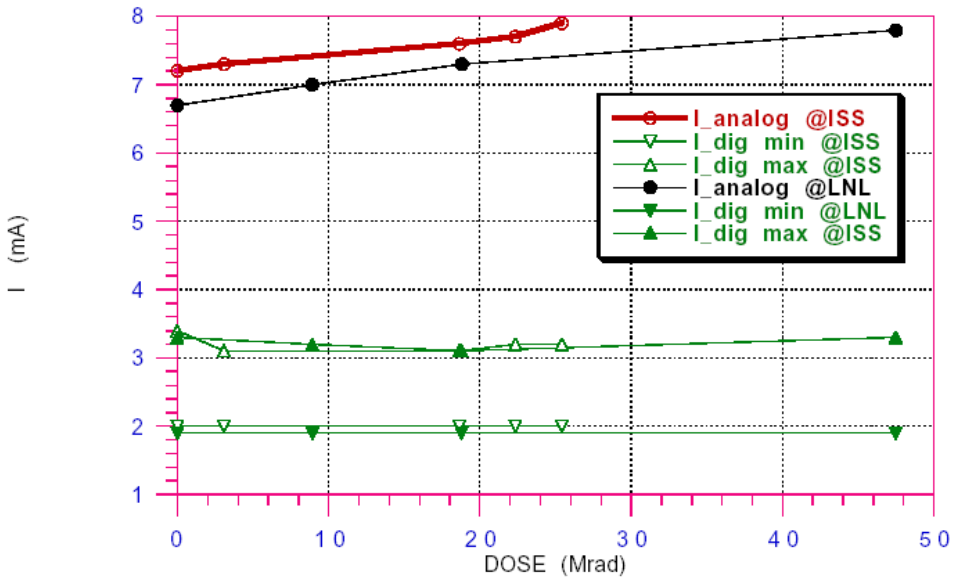

Figure 6.16: Analogue and digital supply currents during and after the proton and ^{60}Co irradiation. The results indicated with ISS are those obtained with the ^{60}Co source. Those indicated with LNL were obtained with the 6.5 MeV proton source. The digital power consumption is unchanged with dose, whilst the analogue consumption increases by around 10%.

6.4 Summary

The chapter begins with a short overview of two of the first pixel readout chips for high energy physics designed at CERN, the Omega2 and Omega3 chips, implemented in a commercial 3 and 1 μm CMOS technology respectively. They proved the feasibility of using pixel detectors for HEP, even if their low radiation tolerance showed that an improvement was needed in that respect.

The RD49 collaboration at CERN started to test radiation tolerance of conventional and edgeless transistors, and the effectiveness of all the layout techniques proposed to improve radiation tolerance of commercial technologies. In that framework the ALICE1test chip was a first pixel readout prototype chip designed in a commercial 0.5 μm CMOS technology but with layout techniques for radiation tolerance. It proved the feasibility of a full mixed-mode chip with the radiation tolerant layout approach. In effect it withstood a TID between 600 krd and 1.7 Mrd depending on the type of radiation. No radiation-induced leakage current was observed, and the circuit failed ultimately because of the cumulative effects of radiation-induced threshold shifts. Some test results were presented in section 6.3.3.

The availability for CERN of the more advanced 0.25 μm technology lead to the design of the ALICE2test chip, based on the previous prototype but with some changes. The circuit, which can work both with positive and negative input charge, comprises a preamplifier, a shaper filter, a comparator, a delay line and readout logic. As this prototype was not intended for bump-bonding to a detector, an input structure has been added to each cell to simulate detector capacitance, coupling between pixels, and detector leakage current. Some details of the circuit scheme were presented in section 6.2, and experimental results in section 6.3.4. The chip was tested before and after irradiation. Before irradiation it was fully functional, and had a minimum threshold voltage of about 1500 e^-, with a threshold spread of 160 e^- rms without threshold fine adjust and 25 e^- after adjustment. The pixel noise is about 220 e^- rms.

The most important results that have been derived from the tests of this prototype concern its radiation tolerance. In effect this chip was tested for radiation tolerance with X-rays, high energy protons and γ-rays. The absence of any increase in power consumption with total dose confirms once more on a full circuit scale that the use of enclosed n-channel transistors and guard rings prevents radiation induced leakage. Moreover, the chip stays fully functional up to a X-ray TID of 30 Mrd(SiO_2) and up to a fluence of 9×10^{14} protons/cm^2, when irradiated at CERN SPS with 450 GeV/c protons. The chip was also irradiated using 6.5 MeV protons at the Van de Graaf accelerator. One chip was irradiated in steps to 9, 19 and 48 Mrd. It ceased to function at 48 Mrd. The analogue outputs and the power consumption showed a behaviour similar to that observed in the other measurements. The chip was also gamma-ray irradiated in steps to doses of 3, 19, 23 and 26 Mrd. In this case, a slight increase in the analogue power supply currents was recorded. For the other parameters, the results of this irradiation closely mirrored those of the X-ray irradiation discussed above.

This prototype proved definitively that it was possible to design a fully working mixed-mode readout chip for hybrid pixel detectors in a commercial technology, which, thanks to the special layout techniques adopted, could withstand radiation doses and particle fluences much higher than the ones expected at LHC.

Table 6.1 shows a comparison of the most important parameters of the four chips.

	Omega2	Omega3	ALICE1test	ALICE2test
CMOS Technol. [feature size]	3 μm	1 μm	0.5 μm	0.25 μm
Chip array [rows ×colums]	16 × 63	16 × 127	65 × 2	65 × 2
Pixel size [μm^2]	75 × 500	50 × 500	50 × 420	50 × 420 (frontend:125 μm)
PowerSupply [V]	3.3	3.3	2.2 (Analogue) 1.6 (Digital)	2.2 (Analogue) 1.6 (Digital)
Analogue power consumption [μW/pixel]	30	34	37	50
Max. Clock [MHz]	20	40	40	40
Peaking time [ns]	50	80	25	25
Noise [e^- rms] no detector	100	100	200	220
Noise [e^- rms] with detec.	170	220		
Minumum Threshold [e^-]	5000	2000	1500	1500
Threshold dispersion [e^- rms]	750	400 (before adjust) 160 (after adjust)	450	~150 (before adjust) ~25 (after adjust)
Minimum radiation tolerance [Mrd]	~0.03	~ 0.05	.6	30

Table 6.1: Comparison of the most important parameters of the four chips.

Chapter 7
The ALICE1LHCb chip: concept and implementation

This chapter presents the design and simulations of the front-end of the ALICE1LHCb chip, a read-out chip for hybrid pixel detectors (described in Chapter 2). It makes use of a different front end scheme with respect to the more classical approach of charge integration plus pole zero cancellation and semigaussian shaping. This approach has been conceived to be used in high multiplicity environments, and to have a fast return to zero of all the blocks which compose the readout chain, whilst keeping low noise (< 200 e^- rms) and low power consumption (of the order of 50-100 μW/channel).

In the first section we summarize all the requirements resulting in our design choices. The second section is devoted to an in depth analysis of the choice of the front-end scheme. The charge integration plus pole zero cancellation and semigaussian shaping approach is analysed, and shown to be extremely impractical for our case. For this reason a different scheme with two complex conjugate poles is proposed, and improved by going to a three poles scheme, adding a real pole with the same real part to the two-pole scheme. A more realistic analysis is carried also out taking into account the rise time of the preamplifier.

The detailed analysis of all the stages which compose the front-end chip is presented, with simulation results. A description of the digital back-end of the chip and of the chip periphery is given in the last sections, as well as some layout considerations explaining some important layout strategies employed to improve chip performance.

7.1 Requirements and specifications

The ALICE1test chip, the first prototype of the ALICE1LHCb chip (section 6.2) was designed to meet the requirements of the ALICE SPD only. Successively it appeared clear that with an advanced CMOS technology as the 0.25 μm it was possible to fulfil the requirements also of the LHCb RICH detector. For this reason the chip was designed using the more stringent requirements from both experiments.

7.1.1 LHCb requirements

The concept of encapsulating a pixel sensor and readout chip within a vacuum tube to form a Hybrid Photon Detector (HPD) has been demonstrated with a number of prototypes

[Ale99, Gys95]. This has led to the development of larger area pixel HPDs in the framework of LHCb [LHC98]. Figure 7.1 shows schematically the concept of the pixel HPD [Wyl99].

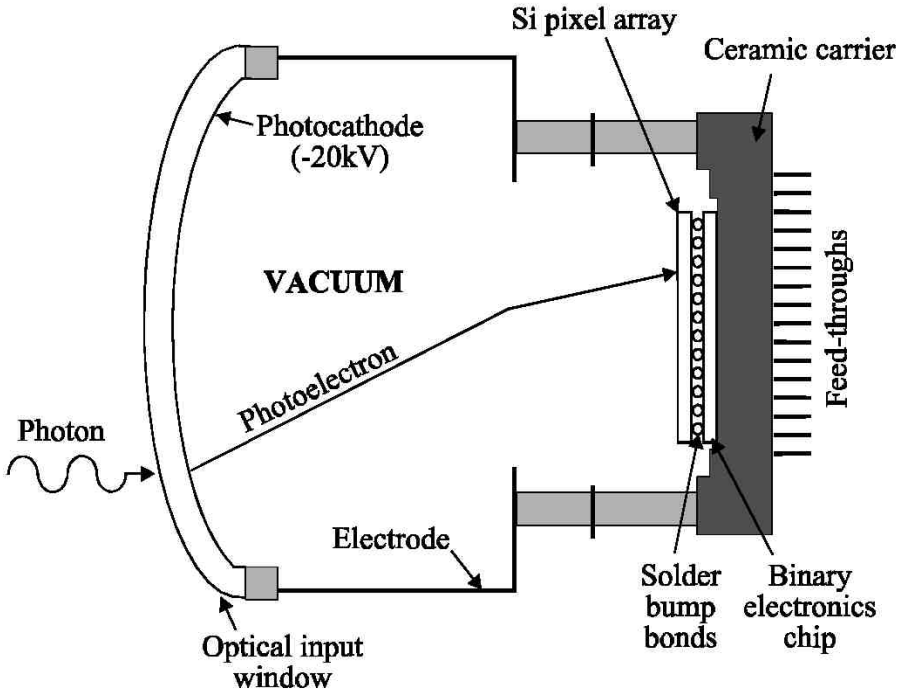

Figure 7.1: Schematic representation of a Pixel Hybrid Photon Detector (HPD) [Wyl99].

Table 7.1 summarizes all the specifications that the chip has to meet, both for what concerns the general LHCb specifications for Level-0 electronics (described in [Chr99]) and the requirements that are more particular to the RICH detector [Gys00].

Photons incident on an optical input window release a photo-electron from a photo-sensitive cathode layer deposited on the inner surface [Wyl00, Bib98]. These photo-electrons are accelerated within the vacuum by a high potential and electrostatically focussed onto an anode which in this case is the pixel sensor and chip. Data from the chip are transmitted out of the device by means of vacuum-tight feed-throughs. The encapsulation of the electronics within the vacuum envelope means that the chip and its packaging must be compatible with all the steps in the manufacturing process of the HPD[1]. Additionally, since the vacuum will present difficulties to the removal of heat, the chip must consume minimal power.

The electrostatic focussing de-magnifies an image on the input window by a factor of 5 in each direction, so the 2.5mm × 2.5mm channel size maps to a 500 μm × 500 μm granularity on the pixel sensor. Triggered events are de-randomised by buffering on the chip,

[1] In particular, the HPD has to undergo some high temperature steps and the tube has vacuum inside. This means that the bump-bonding process has to stand such high temperatures without degradation, and that it has not to outgas once placed in the HPD.

and a 40 MHz clock is then applied for the readout. A complete detector system for the RICH of LHCb would consist of 500 HPDs.

With an accelerating voltage of 20 kV at the HPD photocathode, the most probable signal size created in the sensor is 5000 e^-, but charge-sharing effects may reduce this to 2500 e^- in a single channel. The Level-0 electronics must therefore be sensitive to this level of signal, with an operational threshold of or below 2000 electrons. Large signals can be created in the sensor by minimum ionising particles (MIPs, ~ 24,000 e^-) or by Cherenkov light produced in the input window of an HPD. Thus the front-end electronics has to be able to accept large signals without saturating and recover in a reasonable time after such an event. Noise has to be minimised to enhance the pattern recognition, and a 1% level of noise occupancy is the maximum which can be tolerated [For98]. The chip must be able to detect a hit with a time resolution of 25 ns, dictated by the LHC bunch-crossing rate. This requirement have to include any effects due to the timewalk of the system where the time to reach threshold is a function of the magnitude of the input pulse.

Operational threshold	≤ 2000 e^-
Typical input signal	5000 e^- (2500 e^- with charge sharing)
Dynamic range	Linear between 0 – 5000 e^- with recovery for large signals
Maximum noise occupancy	1%
Time resolution	25 ns
Return to zero time	150-200 ns
Channel size	500 μm × 500 μm
Maximum occupancy	8%
Maximum time-averaged occupancy	4%
Bunch crossing rate	40.08 MHz
Average Level-0 trigger rate	1 MHz
Level-0 latency	4 μs
Level-0 derandomiser depth	16
Maximum readout time, including data headers	900 ns
Maximum radiation total dose for 10 years	30 krd

Table 7.1: Specifications of RICH binary Level-0 electronics.

The effective chip pixel size has to match that of the sensor (500 μm × 500 μm). This size determines directly the channel occupancy. According to simulations, the maximum probability that a pixel is hit during an event with one interaction is 8%. This is for channels in RICH1 which correspond to tracks nearest the beam-pipe. The time-averaged occupancy,

defined as the probability of a pixel being hit in each bunch crossing, is 4%. This is calculated by weighting the 8% occupancy with the probability of having 0, 1, 2 or 3 interactions per crossing according to data in the LHCb Technical Proposal [LHC98, p.99 Figure 12.2. The nominal luminosity for LHCb is 2×10^{32} $cm^{-2}s^{-1}$]. The occupancy is, however, distributed non-uniformly across the RICH detectors, and drops to below 1% across 80% of RICH-1 and all of RICH-2.

The electronics has to be able to accept an average Level-0 trigger rate of 1 MHz and, to maximise the efficiency of the system, all Level-0 electronics in LHCb must be able to accept triggers in two consecutive bunch crossings. Any hits detected by the front-end have to be delayed by 4 μs to match the Level-0 latency. Events accepted by the Level-0 trigger need to be de-randomised to comply with a periodic readout structure. The optimal depth of a buffer for this task is 16 if readout is completed within 900 ns. The chip must also be able to withstand total-dose radiation levels of 30 krd across 10 years of operation and, whilst the levels of heavily-ionising particles in the RICH are still under investigation, immunity to single-event upset (SEU) is desirable.

7.1.2 ALICE requirements

For what concerns ALICE, the original requirements, presented in the ALICE ITS Technical Design Report (TDR) [ALI99], are listed in Table 7.2.

Cell size	50 μm ($r\varphi$) × 300 μm (z)
Number of cells	256 ($r\varphi$) × 32 (z)
Minimum threshold	below 2000 e^-
Typical input signal	25000 e^-
Threshold non uniformity	< 200 e^-
Power consumption	< 1 W/chip
Strobe (LVL1) latency	up to 10 μs
Strobe duration	200 ns
Clock frequency	10 MHz
Derandomiser depth	4
Readout time	400 μs
Hit occupancy	1%
Radiation tolerance	>500 krd
Individual cell threshold adjust	Yes
Individual cell mask	Yes
JTAG controls	Yes

Table 7.2: Specifications of ALICE SPD front-end electronics.

The nature of the heavy ion collisions will generate events of high multiplicity, and the system must be able to cope with a hit occupancy of 1%.

Any hit detected by the readout chip has to be delayed until the arrival of the Level-1 trigger, which has a maximum latency of 10 μs and a rate of a few kHz depending on the types of particles being collided in ALICE. The duration of the trigger signal is one clock period of the 10 MHz system clock, so the delay of hits must be accurate to 100 ns across this 10 μs. The readout of the pixel chips is initiated by a Level-2 trigger, which has a maximum latency of 100 μs and a rate of a few kHz. To keep the deadtime under the specified 10%, the readout of an entire event from the pixel detector system must be completed within 400 μs. The final requirement is that the system has to be tolerant to an ionising radiation dose of 500 krd, integrated across 10 years of LHC operation [Pas03] (see also Chapter 5).

An additional constraint imposed by the ALICE SPD geometry is related to the dead area[2], which has to be kept at a minimum. In particular the maximum space available at the bottom of the chip was 2.5 mm and about 500 μm at the top of the chip. This is why the chip has pads only at one end, which makes an efficient power supply distribution more complex.

7.1.3 Chip requirements

After investigation it turned out that it was not possible to shrink all the electronics in a 50×300 μm^2 pixel, as required by the original ALICE TDR, so it was increased to 50×425 μm^2. Moreover, it was decided to reduce the ALICE sensor thickness from 300 μm to 200 μm for material budget reasons, so the typical ALICE input signal is now 16700 e$^-$.

As the requirements for the on-chip logic are somehow different for the two experiments, the logic can be reconfigured thanks to a special bit which selects the mode of operation, ALICE or LHCb (as explained in section 7.4.4).

The requirements coming from the two experiments are summarized in the Table 7.3.

Cell size	50 μm ($r\varphi$) × 425 μm (z) 400 μm ($r\varphi$) × 425 μm (z) in LHCb mode
Number of cells	256 ($r\varphi$) × 32 (z) 32 ($r\varphi$) × 32 (z) in LHCb mode
Dynamic range	Linear between 0 – 5000e$^-$ with recovery for large signals
Typical input signal	Up to 16700 e$^-$
Power consumption	< 1 W/chip
Maximum noise occupancy	1%
Threshold uniformity	200 e$^-$
Maximum occupancy	8%
Maximum time-averaged occupancy	4%
Peaking time	25 ns

[2] The dead area is the area of the chip which is not covered by the active pixel matrix.

Return to zero time	150-200 ns
Internal programmable delay	Up to 10 μs
Clock frequency	10 MHz and 40 MHz
Radiation tolerance	> 500 krd
Maximum readout time, including data headers	900 ns (with a 40 MHz clock)
On chip FIFO (derandomiser) depth	4 16 in LHCb mode
Individual cell threshold adjust	Yes
Individual cell mask	Yes
JTAG controls	Yes
Reconfigurable logic	Yes

Table 7.3: Specifications of ALICE1LHCb front-end electronics, taking into account the needs of both ALICE and LHCb.

7.2 Choice of the front-end scheme

7.2.1 The Charge Sensitive Amplifier with Semigaussian shaping

The scheme which makes use of a Charge Sensitive Amplifier (CSA) with semigaussian (SG) shaping of order N is studied in detail in [Cha91]. Figure 7.2 shows a schematic representation of this type of front-end.

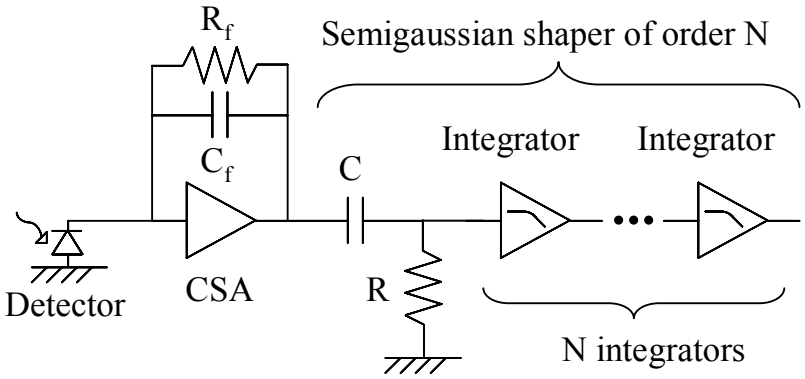

Figure 7.2 : Schematic representation of the Charge Sensitive Amplifier (CSA) with semigaussian (SG) shaper of order N.

The working principle of this circuit is based on a first integration of the input signal coming from the detector (that can be approximated with a δ pulse of charge with area Q_{in}) performed by the CSA. It is assumed that the rise time of the CSA is small, so that its output can be considered to be a voltage step. This step is then differentiated and integrated several times by a semigaussian shaper of order N to decrease the noise content of the signal.

The calculations of all the noise contributions are presented by Chang and Sansen in [Cha90] and [Cha91], and summarised here. Figure 7.3 shows all the noise sources which are taken into account in the calculation.

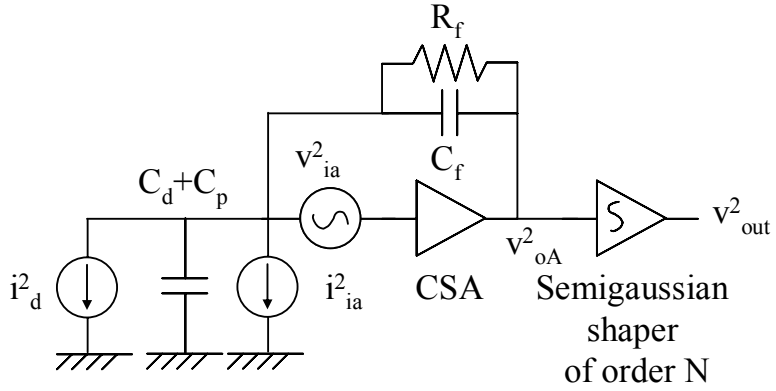

Figure 7.3: Representation of the CSA with SG shaping with all the relevant noise sources. v^2_{ia} and i^2_{ia} are respectively the equivalent input voltage noise and current noise generators of the CSA. i^2_d is the noise current associated with the detector leakage current and its associated bias network. The detector capacitance is C_d. The parasitic capacitance of the interconnections between the detector and the amplifier input is represented by C_p.

In Figure 7.3 i^2_d is the noise current associated with the detector leakage current and its associated bias network, while v^2_{ia} and i^2_{ia} are respectively the equivalent input voltage noise and current noise generators of the CSA. The detector capacitance is C_d. The parasitic capacitance of the interconnection between the detector and the amplifier input is represented by C_p.

Neglecting the bias network associated with the detector, the noise current associated with the detector can be expressed as:

$$i^2_d = 2\,q\,I_o \tag{7.1}$$

where I_o is the detector leakage current and q the electron elementary charge $1.62 \cdot 10^{-19}$ C. In this component the noise coming from the feedback network can also be added.

It can be shown that the noise power spectrum at the CSA output node is given by:

$$v^2_{oA} = \left| \frac{C_d + C_p + C_f + C_G}{C_f} \right|^2 v^2_{ia} + \left| \frac{1}{s\,C_f} \right|^2 i^2_d \tag{7.2}$$

where C_G is the gate capacitance[3] of the input MOS transistor.

To obtain the noise power spectrum at the output of the shaper we have to weight v^2_{oA} by the shaper transfer function:

[3] Note that in [Cha91] $C_G = C_{GS} + C_{GD}$ and C_{GB} is neglected. This is a reasonable assumption in strong inversion, but in weak inversion C_{GB} should not be neglected. In our calculations C_{GB} was taken into accout.

$$\int_0^{\infty} \left| v_{oA}(2j\pi f) \right|^2 \left| H(2j\pi f) \right|^2 df \tag{7.3}$$

where the shaper transfer function is, in case of a semigaussian shaper of order N and with a time constant τ_o both for the differentiator and integrators, and a DC gain A:

$$H(s) = \left[\frac{s\tau_o}{1+s\tau_o} \right] \left[\frac{A}{1+s\tau_o} \right]^n \tag{7.4}$$

In order to calculate the various noise contributions at the output of the shaper in equivalent noise electrons, we have to calculate the maximum of the shaper signal amplitude generated by a charge of one electron at the CSA input. This value is given by:

$$V_{outp} = \frac{q\, A^N\, N^N}{C_f\, N!\, e^N} \tag{7.5}$$

If we suppose that the main noise contribution comes from the input transistor of the CSA, as is usually the case, we can evaluate the different contributions to the ENC of the circuit (v^2_{ia} and i^2_{ia} can be found in Appendix I):

$$ENC_d^2 = \frac{8}{3} k\, T \frac{1}{g_m} C_t^2 \frac{Beta[1.5\,,\, N-0.5]\, N}{q^2\, 4\pi\, \tau_s} \frac{(N!)^2\, e^{2N}}{N^{2N}} \tag{7.6}$$

for the thermal noise component;

$$ENC_f^2 = \frac{Kf}{C_{ox}^2\, W\, L} C_t^2 \frac{1}{q^2\, 2N} \frac{(N!)^2\, e^{2N}}{N^{2N}} \tag{7.7}$$

for the flicker noise component and

$$ENC_o^2 = 2\, q\, I_o\, \tau_s \frac{Beta[0.5,\, N+0.5]}{q^2\, 4\pi\, N} \frac{(N!)^2\, e^{2N}}{N^{2N}} \tag{7.8}$$

for the shot noise component, where k is the Boltzmann constant, T the absolute temperature, g_m the device (gate) transconductance, C_t is the total input capacitance defined as: $C_t = C_D + C_P + C_f + C_G$, τ_s is the shaper peaking time defined as: $\tau_s = N\,\tau_o$, C_{ox} is the gate oxide capacitance per unit area, W and L are the width and the length of the input transistor,

Kf is the flicker noise constant (also called Ka) as defined in Appendix I and Beta[x,y] is the Beta function of arguments x and y.

This noise analysis was done targeting mainly front-end chips for silicon strip detectors in MOS technologies with feature size above 1 μm. Though this analysis is rather general, some changes need to be applied to match the design of front-end circuits for pixel detectors in submicron technologies, especially if applying Hardening By Design (HBD) techniques.

Equation (7.8) does not depend at all on the CSA parameters, but only on the shaper parameters. It is still valid for our design, provided that in I_o all the contributions of the parallel noise are accounted for.

Some imprecision in equation (7.6) arises from the fact that the gate-source g_m is used, neglecting the bulk-source transconductance g_{mb}. In our calculations we took into account also g_{mb}, as explained in Appendix I.

The other difference with respect to our design is related to the assumption that the input MOS device of the CSA is in strong inversion and in saturation. Although it is reasonable to assume it is in saturation, it may no longer be in the strong inversion region. The trend of the power supply voltage with scaling is to decrease faster than the threshold voltage, thus reducing the margin for transistors operating in strong inversion. Moreover, the input transistor of the CSA is always sized with minimal (or almost minimal) length for noise and speed performance, and the power budget for pixel front-ends is strictly limited (usually to some tens of μA). If we add that n-channel transistors have to be designed edgeless for radiation tolerance, this means that n-channel transistors are never in strong inversion. Most likely an optimised input p-channel transistor will also end up in weak or moderate inversion. However, the best region to operate the input transistor to maximise the ratio performance/power consumption is the brink between strong and moderate inversion [Oco02].

The other factor which is not accounted for in Chang's formulae is the fact that for a real device measured noise is often higher than what can be predicted using the standard equation for noise (equation I.7, Appendix I). This is usually taken into account adding the *noise excess factor* Γ in equation I.8 (Appendix I). In our formulae we took into account the Γ factor, using as an estimation of its value the results shown in Figure 5.27.

An estimation of the various noise contributions, even with the correction factors already explained, can give us an insight of the noise level that it is possible to reach with a standard configuration employing a CSA with semigaussian shaping.

Thermal noise

If we use the equations given in Appendix I and in [Enz95, Enz01] for the interpolation of g_m and for C_G from weak to strong inversion, using equation (7.9) we can calculate the ENC due to the thermal noise (see Appendix I; formulae for ENC_f and ENC_0 stay the same).

$$ENC_d^2 = 4\,k\,T\,\frac{n\gamma}{g_m}\,C_t^2\,\frac{Beta[1.5\,,\,N-0.5]\,N}{q^2\,4\,\pi\,\tau_s}\,\frac{(N!)^2\,e^{2N}}{N^{2N}} \qquad (7.9)$$

Neglecting the dependence of the slope factor n with the gate voltage (i.e. assuming n constant in equation (7.9)), γ (equation I.8, Appendix I) is a function of the inversion factor (i.e. of I_D, W and L), g_m is a function of I_D, W and L, and C_t is again a function of I_D, W and L.

To obtain a first estimate of the thermal noise contribution, we can make some assumptions. First of all, we can estimate the detector capacitance to be about 100 fF, which should be approximately the capacitance of a bump-bonded pixel cell, and the feedback capacitance about 20 fF. This value is based on the fact that a value which is too high degrades the noise performance and reduces the CSA gain and speed, and can be difficult to integrate in the pixel cell, while a value which is too low could result in an excessively high CSA gain (which could lead to instabilities and oscillations) and is also more difficult to implement in the layout with reasonable precision and reproducibility.

For power budget reasons a maximum drain current I_D (per input transistor) of about 10 μA can be used. Due to the limited area and power available for the design, it would be rather difficult to implement a semigaussian shaper with a value of N > 1. The shaping time τ_s is fixed by the experiment to be 25 ns. For what concerns the transistor length L, it is usually chosen to maximize g_m and maximise the device speed, so the best choice is the minimum L allowed by the technology. In our technology it is 0.24 μm, but to avoid excessive short channel effects a design value of 0.28 μm is safer[4]. This means that in equation (7.9) all the parameters are fixed, except W. Figure 7.4 shows the ENC_d as a function of W for a p- and an n-channel transistor. For the p-channel transistor the value for n was chosen to be 1.35 and, as most probably the transistor would end up in moderate inversion, Γ was chosen from figure 5.27 to be 1.15. For the n-channel transistor the value for n was chosen to be 1.4 and Γ was chosen from figure 5.27 to be 1.35. To compute Figure 7.4 we took into account the gate to drain, gate to source and gate to bulk overlap capacitances ($C_{gso} = C_{gdo} = 3 \cdot 10^{-10}$ F/m; $C_{gbo} = 3.2 \cdot 10^{-11}$ F/m).

The minimum for the p-channel MOS is 51 e^- at W = 23 μm, and the ENC_d is still only ~51 e^- at W = 30 μm, while for the n-channel MOS the minimum is 45 e^- at W = 15 μm and the ENC_d is still 47 e^- at W = 30 μm. It can be seen that there is not a big difference so, as far as thermal noise is concerned, both a p and an n channel transistor would be almost equivalent. In addition, if 5 μm < W < 80 μm is chosen, the variation of ENC_d with W is negligible for our purpose. The small difference is due mainly to the fact that for transistors

[4] Note that the design value for L and W is different from the effective one. For L, the formula is: $L_{eff} = L_{drawn} - \Delta L$, where $\Delta L = 0.06$ μm in our case. A similar formula applies to W, with $\Delta W = 0.02$ μm. For short (or narrow) transistors, it is very important to use the effective length (or width) instead of the drawn one.

working in weak/moderate inversion, the g_m of an n-channel transistor is not much higher than the g_m of a p-channel transistor (as it would be the case in strong inversion).

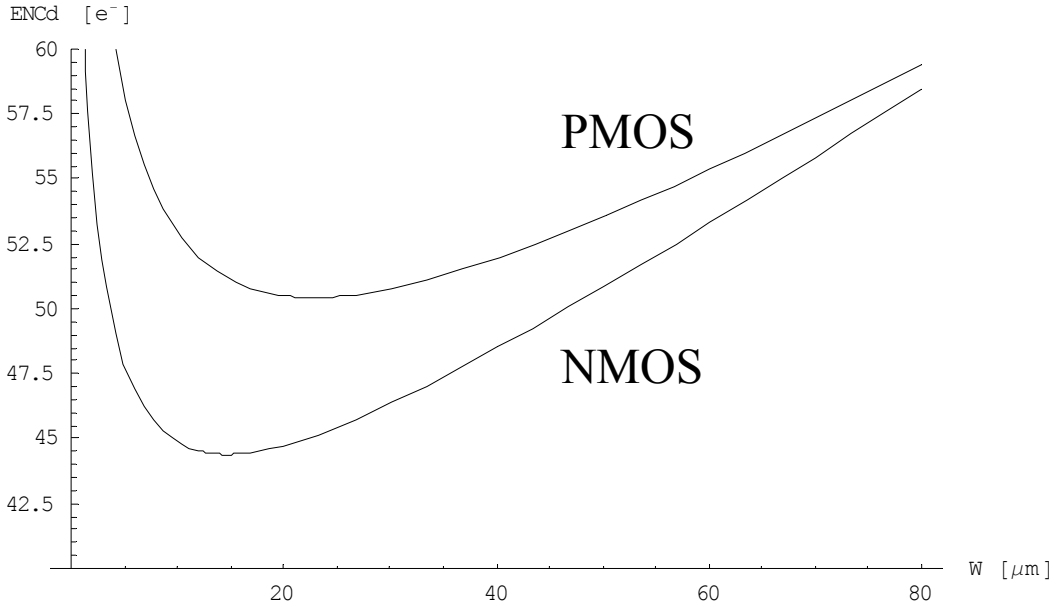

Figure 7.4: ENC_d as a function of W for a p- and an n-channel transistor. For the p-channel transistor n = 1.35 and Γ = 1.15. For the n-channel transistor the value for n is 1.4 and Γ = 1.35.

Flicker noise

A similar analysis can be done for the flicker noise contribution ENC_f. The values for the flicker noise constant K_f can be found in Figure 5.26 (in the figure it is named K_a). Even if they are not constant from weak to strong inversion, they do not vary by a large amount, so we can assume a constant value of $0.6 \cdot 10^{-27}$ C^2/m^2 for p-channel transistors and $3.5 \cdot 10^{-27}$ C^2/m^2 for n-channel transistors. The result is shown in Figure 7.5.

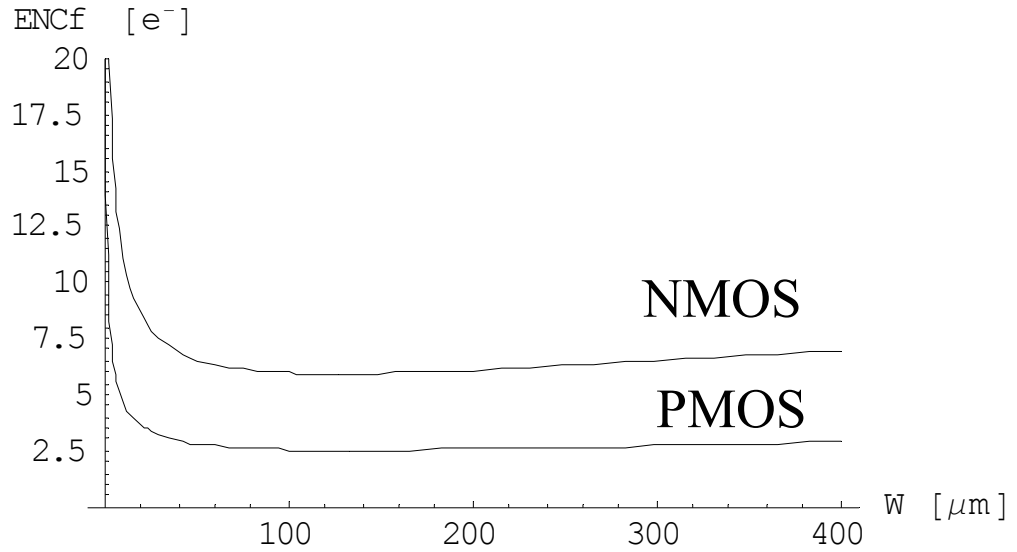

Figure 7.5: ENC_f as a function of W for a p- and an n-channel transistor. For the p-channel transistor $K_f = 0.6 \cdot 10^{-27}$ C^2/m^2 and for the n-channel transistor $K_f = 3.5 \cdot 10^{-27}$ C^2/m^2.

The minimum for the p-channel transistor is 2.5 e^- at W=133 μm, and the ENC_f is 3.2 e^- at W = 30 μm; the minimum for the n-channel transistor is 6 e^- at W=125 μm, and the ENC_d is 6.1 e^- at W = 30μm. This contribution is then negligible, and the optimisation of the front-end can be done without taking it into account.

Shot and total noise

For what concerns the shot noise, the leakage current coming from the detector can be evaluated from the data presented in table 4.1. The maximum leakage current guaranteed from the vendor is 5 nA/cm^2; as the pixel area is 50×425 μm^2 this results in a maximum leakage current $I_o \sim 1$ pA.

Irradiation should generate a per-pixel leakage current of about 2 nA (cfr. Figure 4.18) at twice the maximum simulated ALICE fluence, so the feedback current of the front-end (which will be in the several tens of nA range) should always be the major contribution to shot noise. To have an idea of the shot noise we can set $I_0 = 50$ nA, which results in an ENC_o of 119 e^-. This does not depend on the CSA but only on the shaper, whose characteristics are fixed.

The total noise will be then (choosing W = 30 μm):
$ENC_t = \sqrt{ENC_d^{\,2} + ENC_f^{\,2} + ENC_o^{\,2}} = 130\ e^-$ for a p-channel transistor and 128 e^- for an n-channel transistor.

It is very important to remark that, though the analysis both for a p- and for an n-channel transistor is presented here, a p-channel input transistor is preferred for our purposes. The g_m of an n-channel transistor is not much higher than the g_m of a p-channel transistor, if working in weak/moderate inversion so that the penalty in power consumption is reduced (if we assume the same g_m in both cases).

The main reason to choose a p-type input is related to substrate noise. In an n-well technology all the n-type transistors share the same p-type substrate. This means that if digital transistors generate a disturbance in the substrate, this can be fed back to the analogue section of the chip. The preamplifier, in particular, has a very high gain, so that a small substrate signal could induce a high signal at the preamplifier output. This is especially true in a scheme without a differential input, where (through the transistor g_{mb}) the substrate noise generates directly a drain current, while in a differential scheme the substrate noise is common mode signal and is rejected (according to the CMRR of the preamplifier). Moreover, also in a differential amplifier the substrate noise can influence the input signal, as it can be coupled to the input through the gate to bulk transistor capacitance C_{gb}. C_{gb} acts as an injection capacitance, so that a disturbance ΔV in the substrate injects a charge $\Delta Q = C_{gb}\,\Delta V_{sub}$ in the input node.

These kinds of couplings degrade the noise performance of the circuit, and could even start an oscillating behaviour of the full chain. Using a p-input transistor reduces drastically

these problems. In effect, to implement a p-channel transistor an n-well has to be implanted first, so that the p-type MOS can be created in the proper n-type substrate. This allows the designer to separate the different n-wells. The first beneficial effect is the possibility to have different digital and analogue wells, with a reduction of the digital-to-analogue crosstalk. Having the possibility to use separate wells for separate transistors allows also to connect the well of the transistor to its source[5]. In a scheme with a differential input with one of the gates of the input pair tied to a reference (as in our case) any disturbance in the well is coupled through the gate to bulk transistor capacitance only to the input which is connected to the signal (because the other is kept constant by the voltage reference). This signal is not common mode, and thus is not rejected. Shorting the well and the source of the input transistors "moves" the substrate disturbance to the common source, so that the noise generates only common mode signal. C_{gb} is now connected to a low impedence node, and in practise it is shorted[6].

Using a p-channel input transistor has some other beneficial effects. Since the value of the flicker noise constant K_a is never known with precision, and can change by a large amount even in the same lot due to process variations, the possibility to neglect the flicker noise contribution is an important point in designs where an excessive noise could kill the performance of the full detector system. The lower flicker noise constant helps to keep the flicker noise level well below the thermal and shot noise, so that it can always be neglected. Another reason for a p-channel input transistor is layout related. As an n-channel transistor would have to be layouted edgeless, this would result in some area waste (the transistors are relatively wide) but most of all in imprecision of the transistor modelling (especially input capacitance and output conductance).

Let now consider the CSA transfer function. We can suppose that in Figure 7.2 the stage has an equivalent transconductance g_m and an input capacitance C_{in}, and that the output node can be modelled as a load resistance R_L in parallel with a load capacitance C_L. The current to voltage transfer function, calculated supposing $g_m R_L >> 1$ and $g_m R_f >> 1$ is:

$$H(s) = \frac{g_m}{g_m G_f + s C_f gm + s^2 C_t (C_l + C_f)} \qquad (7.10)$$

where $G_f = 1/R_f$.

If we also suppose that the two poles are widely spread, the circuit can be modelled as a two poles system, where the time constants of the two poles are:

[5] This is not possible for n-channel transistors, since all of them share the same substrate.
[6] Furthermore, C_{gb} in this case is in parallel with C_{gs} which is dominant.

$$\tau_f = R_f\, C_f$$
$$\tau_r = \frac{C_t\,(C_l + C_f)}{C_f\, g_m} \tag{7.11}$$

τ_r is responsible for the fast rise time of the CSA, and is usually of the order of some nanoseconds or less, while τ_f is related with the fall time of the signal, and can be of the order of several microseconds. In effect, as R_f acts as a source of parallel noise at the CSA input, from the noise point of view a high value (several megaohms or more) is recommended. The CSA output in the time domain is given by equation (7.12):

$$V_{out}(t) = \frac{Q_{in}\,\tau_f \left(e^{-\frac{t}{\tau_r}} - e^{-\frac{t}{\tau_f}} \right)}{C_f\,(\tau_r - \tau_f)} \tag{7.12}$$

Figure 7.6 plots the output voltage of the CSA (normalised to Q_{in} / C_f) for a rise time constant of 10 ns, and a fall time constant of 10 μs and of 0.5 μs.

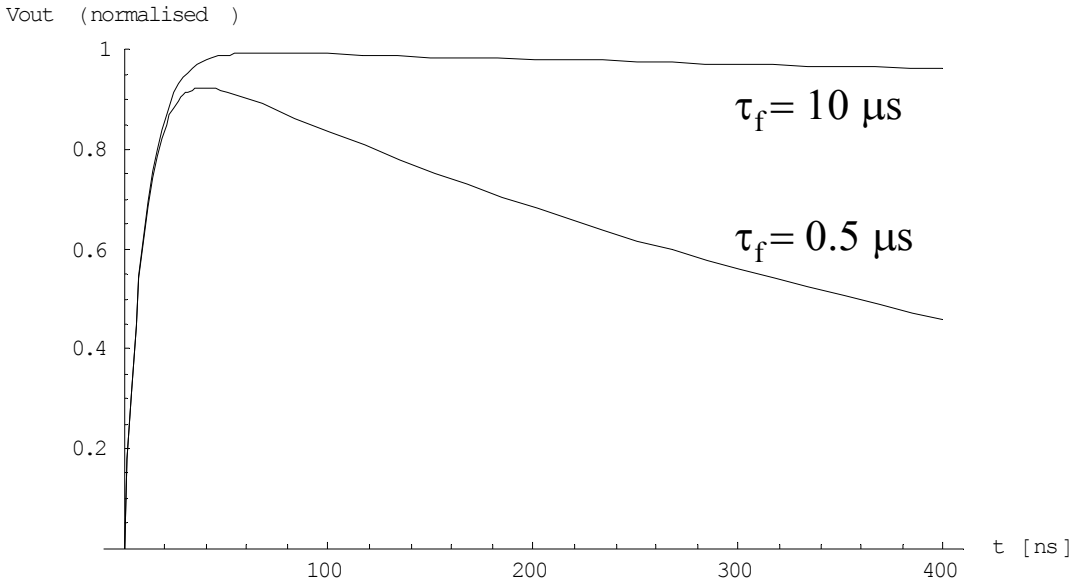

Figure 7.6: Plot of the output voltage of the CSA (normalised to Q_{in} / C_f) for a rise time constant of 10 ns and a fall time constant of 10 μs and for a fall time constant of 0.5 μs.

Figure 7.7 shows how the maximum of equation (7.12) (for a constant input signal and a fixed rise time) varies with τ_f. If τ_f decreases, approaching τ_r, the value of the maximum decreases too, and the circuit loses efficiency. The same effect is also evident in Figure 7.6.

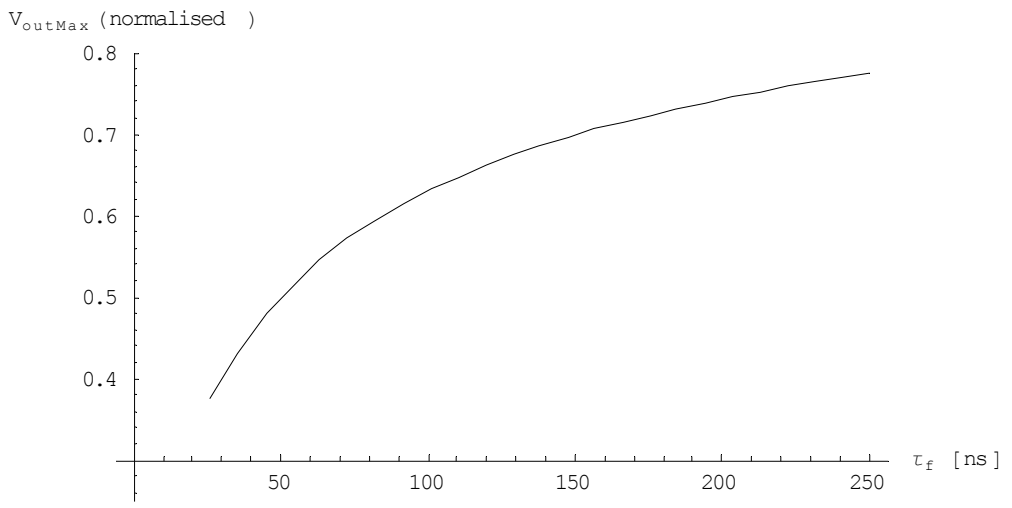

Figure 7.7: Maximum value of the output (normalised to Q_{in}/C_f) of the CSA as a function of the fall time of the CSA. τ_r is fixed at 25 ns. If τ_f decreases, approaching τ_r, the value of the maximum decreases too, and the circuit loses efficiency.

Although it is possible to recover the signal to the original height by increasing the gain of the subsequent stages, this is an undesirable side effect. In effect, it can require additional power and it could lead to a degradation of the noise performance, if the gain of the second stage increases so much that its noise contribution becomes comparable to that of the front-end.

The slow decaying tail of the preamplifier will result in a negative undershoot of the CSA output, and the duration of the undershoot is determined by the decay time of the step signal at the CSA output.

If the differentiator in between the CSA and the shaper in Figure 7.2 is modified adding a resistance in parallel with the capacitance C, the RC network adds a zero in the transfer function. If the RC network time constant is chosen to be equal to τ_f, the zero introduced cancels the slow pole of the CSA, feeding the shaper input with a signal which is much closer to an ideal step. The effect is that the slow return to zero at the output of the CSA is cancelled at the output of the shaper. This technique is called *pole-zero (PZ) cancellation*, and leads to a shaper output that is given by equation (7.13), if the rise time of the step signal at the CSA output is much shorter than the peaking time of the shaper:

$$V_{ShaperOut} = \frac{Q_{in} A^N}{C_f N!} \left(\frac{t}{\tau_s} \right)^N e^{-\frac{N t}{\tau_s}} \qquad (7.13)$$

where A is the DC gain of one shaping stage and $\tau_s = N\,\tau_o$ is the peaking time of the cascade of the N shaping stages, each with a time constant τ_o. The plot of the shaper output for $Q_{in}/C_f = 1$, A = 1, N = 1 and $\tau_s = 25$ ns is shown in Figure 7.8.

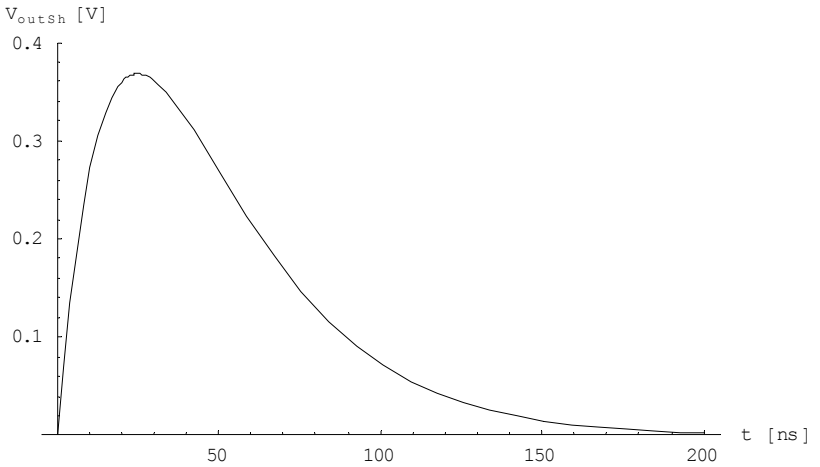

Figure 7.8: Plot of the shaper output of a CSA followed by a semigaussian shaper of order 1, for $Q_{in}/C_f = 1$, A = 1 and $\tau_s = 25$. The peaking time is at 25 ns, the return to zero time 191 ns.

One of the possible ways to define the return to zero time is the time at which the signal has reached 1% of its peak value. In this case[7] it depends only on parameters that are already set (N = 1, $\tau_s = 25$ ns), its value is 191 ns. This value is very close to the upper limit imposed by LHCb, and cannot be changed if N and τ_s are fixed.

All the calculations done up to now suppose that the rise time of the step signal at the CSA output is much shorter than the peaking time of the shaper. If we assume a factor ten, this would lead to a rise time of 2.5 ns and a time constant[8] of 1.1 ns. The power required to realise such a low time constant is too high, and a preamplifier with a higher time constant will have to be used. This means that in any case all the formulae presented up to now are only approximate, and that more calculations should be carried out for precise results.

The main problem related with the CSA scheme with semigaussian shaping is present at the CSA output. The cancellation has no effect on the CSA and its output, therefore if another particle hits the pixel before the signal caused by the first one is vanished, at the CSA output we have a superposition of the two (or even more, if τ_f is very high) signals. This effect is called *pile-up.* An example of pile-up of five subsequent input signals is shown in Figure 7.9, where the rise time is 25 ns, the fall time 2 μs and the period of the input signal is 1 μs.

---

[7] In this case the peak is at $t = \tau_s$, and is equal to: $V_{outPeak} = \dfrac{Q_{in} A^N N^N}{C_f N! e^N}$.

[8] It can be calculated from the (7.13) that the rise time t_r of the preamplifier, defined as the time period between which the output signal rises from 10% to 90% of the maximum amplitude, is $t_r = 2.2\,\tau_r$.

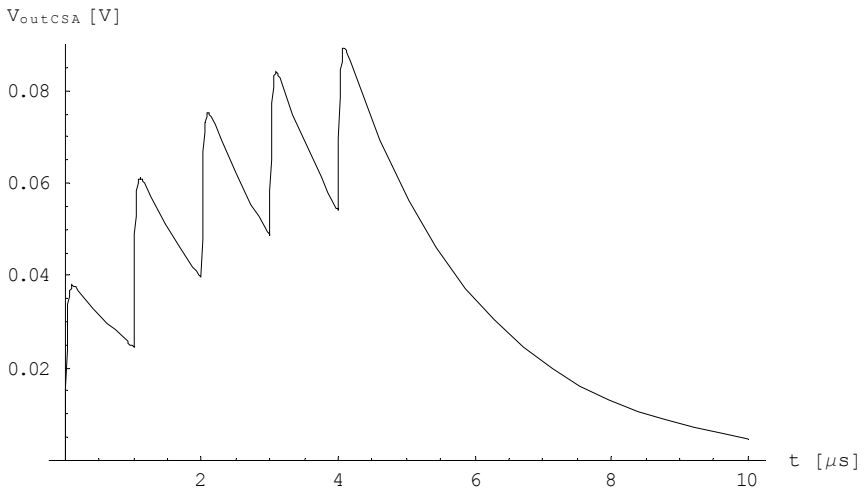

Figure 7.9: CSA output for five subsequent input pulses. $\tau_r = 25$ ns, $\tau_f = 2$ μs, the input pulse period is 1 μs.

Although in principle[9] no pile-up will be present at the shaper output, the plot shows that the CSA output increases with time, and this can saturate its output stage. In case of saturation of the CSA the system is not linear anymore, and also the shaper output will be affected. This problem is of course of paramount importance in experiments with high multiplicity, as for example LHCb where the maximum multiplicity is of the order of 8%.

Another problem related with the pole-zero cancellation scheme is the precision of the cancellation. To realize the feedback resistor R_f of high value a MOS transistor biased at very low current (tens of nA) has to be used, thus in weak or moderate inversion region. The resistance generated in this way has to be matched with a corresponding transistor in the pole-zero cancellation network. In the classic implementation of the PZ scheme the matching precision is rather poor, so this means that the slow pole of the CSA is not perfectly cancelled by the zero of the pole-zero cancellation. This generates a reduction of the gain and an oscillation of the signal, with a remaining "tail" which goes back to zero much slower than the signal with a proper pole cancellation. This tail is equivalent to a time-variable shift of the pixel threshold, which can become quite important if several hits arrive in a relative short time, so that more than one tail piles up at the same time.

7.2.2 The ALICE1LHCb front-end concept: two poles system

Several elements pointed out in the previous section indicate that a standard CSA with pole-zero cancellation would be quite impractical to realize in our case and would lead to unnecessary system degradation. In effect pile-up can be a problem in high multiplicity environments; it would also be necessary to conceive a different scheme for the feedback

[9] i.e. if the pole-zero cancellation is perfect.

stage to improve the precision of the pole-zero cancellation, and the return to zero time can be improved only by increasing the order of the shaping. Moreover, the analytical approach proposed in [Cha91] is only approximate for a scheme with a low-power CSA and a fast shaping time.

This lead us to try to develop a different front-end scheme, with noise performance comparable with the standard scheme, but with improved return to zero time both at the shaper output and at the preamplifier output, to be able to safely operate even in high multiplicity environments [Din03, Din01, Din00, Sno01, Sno01a, Wyl99].

The basic idea is that a system with complex poles could have a faster return to zero than a system with only real poles. Supposing that the CSA is followed by a simple stage realising a pole with time constant τ_{p2} (and with a DC gain A), and that a feedback current is fed back to the amplifier input by means of a transconductance feedback stage g_{mf}, choosing in the proper way τ_{p2}, A and g_{mf} it is possible to realize complex conjugate poles and to chose their real and complex part. Calculations have to be carried out to investigate the feasibility of the system in terms of noise, return to zero, power consumption and area required to implement the layout. The block diagram representation of the system is shown in Figure 7.10. Note that in this case no feedback resistance is present in parallel with the feedback capacitor, so a different scheme has to be used to absorb the detector leakage current. A scheme for leakage current and offset compensation will be presented in section 7.3.6, but its contribution to the signal formation can be neglected as it acts only at very low frequency.

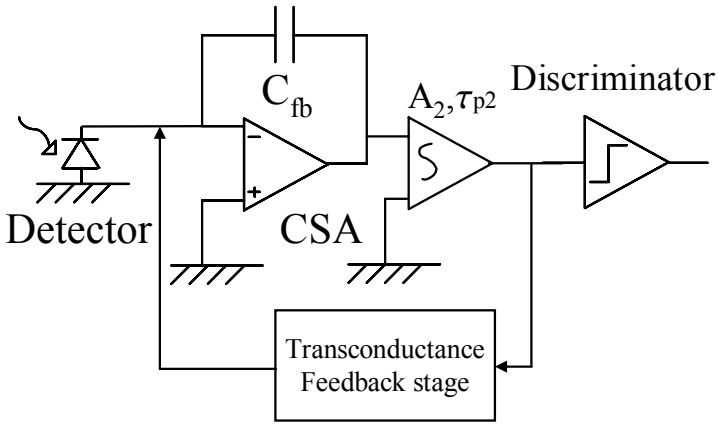

Figure 7.10: Block diagram representation of the system with two complex poles, with differential input CSA and shaper.

7.2.2.1 The CSA

Regardless of the internal structure of the CSA, it is reasonable to approximate it with an input capacitance C_{in}, a transconductance g_m and an output impedance composed of the

parallel combination of an output capacitance C_{out} and an output resistance r_{out}. Solving the following system we can find the CSA transresistance transfer function V_{out}/I_{in}:

$$\begin{cases} I_{in} = s\, C_{in}\, V_{in} + s\, C_{fb}\,(\, V_{in} - V_{out}\,); \\ V_{out}(1/r_{out} + s\, C_{out}) + g_m\, V_{in} - s\, C_{fb}\,(\, V_{in} - V_{out}\,) = 0; \end{cases} \tag{7.14}$$

After a straightforward calculation and supposing[10] $g_m\, R_{out}\, C_f >> C_{fb} + C_{in}$, the CSA transresistance can be written as:

$$H_{CSA}(s) = \frac{-1}{s\, C_{fb}} \frac{1 - \frac{C_{fb}}{g_m} s}{1 + s \frac{C_{fb}(C_i + C_{out}) + C_i\, C_{out}}{g_m\, C_f}} \tag{7.15}$$

This means that the CSA can be modelled as an ideal integrator plus a zero τ_z (positive real) and a pole τ_r (negative real) placed at:

$$\begin{aligned} \tau_z &= \frac{C_{fb}}{g_m} \\ \tau_r &= \frac{C_{fb}(C_i + C_{out}) + C_i\, C_{out}}{g_m\, C_f} \end{aligned} \tag{7.16}$$

For our purposes, the zero is usually of the order of 100 ps, while the pole in the order of some nanosecond. This means that usually the zero can be neglected, and that if the rise time of the subsequent stage is high enough, the pole can also be neglected, so that the CSA can be considered as an ideal integrator with a transfer function $-1/(s\, C_{fb})$.

7.2.2.2 The first shaping stage

The second stage in the Figure 7.10 can be represented by a one pole (τ_{p2}) function with a DC gain A_2, so that its transfer function V_{out}/V_{in} can be written as:

$$A_{Sh1}(s) = \frac{A_2}{1 + s\, \tau_{p2}} \tag{7.17}$$

[10] This hypothesis is usually verified, as $g_m\, R_{out}$ is in general >>10.

This is a one pole shaping function, so we will call the block "shaping stage" even if in this case it has a slightly different meaning than that for the semigaussian shaper discussed earlier. The overall direct transconductance transfer function is then $H_{Direct} = H_{CSA}(s) \cdot A_{Sh1}(s)$.

7.2.2.3 The system in closed loop

For the feedback stage we can assume an ideal transconductance behaviour, $I_{out}/V_{in} = -g_{mf}$ (where V_{in} is the shaper output voltage and I_{out} is the current fed back to the input considered as positive if leaving the input node), and we can employ the well known formula for systems with feedback [Gra84, p.467]:

$$\frac{S_o}{S_i} = \frac{a}{1 + a\, f} \tag{7.18}$$

where a is the direct transfer function of the system, and f is the feedback transfer function (the feedback signal is subtracted from the input signal). If we define an "effective feedback pole" $\tau_{fb} = C_{fb}/ (A_2\, g_{mf})$, the closed loop transfer function is:

$$\tau_{fb} = \frac{C_{fb}}{A_2\, g_{mf}} \tag{7.19}$$

$$H_{cl} = \frac{-1}{g_{mf}\,(1 + \tau_{fb}\, s + \tau_{fb}\, \tau_{p2}\, s^2)}$$

This is the V_{out}/I_{in} transfer function, so its inverse Laplace transform is the time response of the system to an input pulse of current (of unitary area), and also the response to a step of charge[11] of unity height. The roots of $1 + \tau_{fb}\, s + \tau_{fb}\, \tau_{p2}\, s^2 = 0$ are the poles of the closed loop system. If they are real we do not have any beneficial effect with respect to the PZ cancellation scheme, so we have to impose the Δ of the previous equation to be negative, i.e.: $\tau_{p2} > \tau_{fb}/4$.

The inverse Laplace transform of equation (7.19) is:

$$Vout(t) = \frac{2\, e^{-\frac{t}{2\tau_{p2}}}\, Sin\left[\frac{t\sqrt{-\tau_{fb}+4\,\tau_{p2}}}{2\sqrt{\tau_{fb}\,\tau_{p2}}}\right]}{g_{mf}\sqrt{\tau_{fb}}\,\sqrt{-\tau_{fb}+4\,\tau_{p2}}} \tag{7.20}$$

[11] The Laplace transform of the step function is 1/s and $I = s\, Q_{in}$, so $V_o = s\, Q_{in}\, H_{cl}\, /s = Q_{in}\, H_{cl}$.

which can also be written in terms of two parameters $\tau_s = 1/y$ and $r = x/y$, where x and y are respectively the real and the imaginary part of the two complex conjugate poles:

$$V_{out}(t) = \frac{(1+r^2)\, e^{-\frac{r\,t}{\tau_s}}\ \mathrm{Sin}\left[\frac{t}{\tau_s}\right]}{g_{mf}\,\tau_s} \tag{7.21}$$

The relationship between the four parameters is given in equation (7.22).

$$\tau_{fb} = \frac{2\,\tau_s\,r}{1+r^2} \qquad \tau_{p2} = \frac{\tau_s}{2\,r} \tag{7.22}$$

An example of the output for an input of 5000 e^-, assuming $g_{mf} = 1.4$ μS, $\tau_{fb} = 15.5$ ns and $\tau_{p2} = 25$ ns is shown in Figure 7.11.

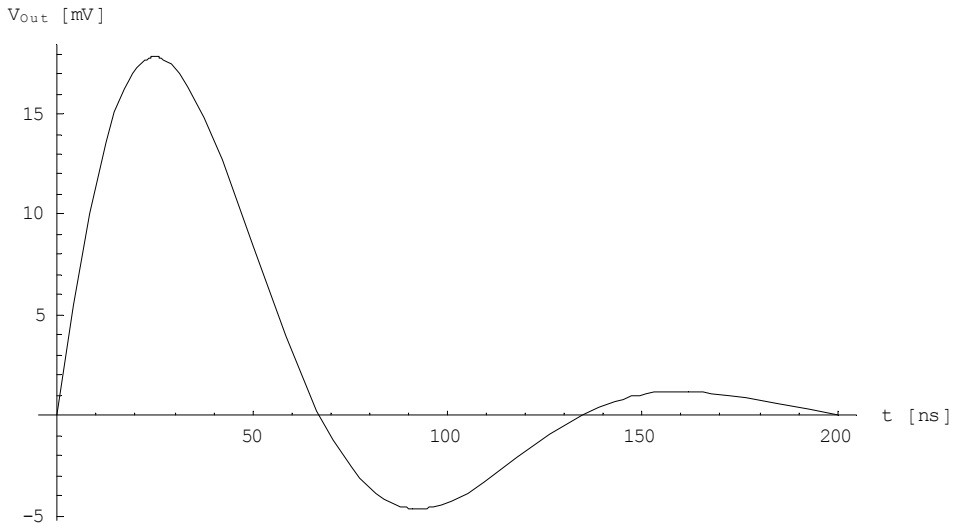

Figure 7.11: Example of the system output for 5000 e^-, assuming $g_{mf} = 1.4$ μS, $\tau_{fb} = 15.5$ ns and $\tau_{p2} = 25$ ns.

Both couples of parameters r and τ_s, τ_{fb} and τ_{p2} are not independent, because if one parameter is given the other has to be found imposing a peaking time at 25 ns. The use of r and τ_s is very helpful when studying the peaking time and return to zero, as the relation which links them to the given peaking time t_o is very simple: all the extremes of V_{out} are given by equation (7.23) and the value of the first maximum (the peak, for K = 0) by equation (7.24):

$$t_{maxk} = \left(\text{ArcTan}\left[\frac{1}{r}\right] + K\pi\right)\tau_s \qquad (7.23)$$
$$\text{for}\{K = 0 .. \infty\}$$

$$V_{omax}(t) = \frac{e^{-r\,\text{ArcTan}\left[\frac{1}{r}\right]}\sqrt{\frac{1+r^2}{\tau_s^2}}}{g_{mf}} \qquad (7.24)$$

From equations (7.20) and (7.21) it can be deduced that, as can also be seen in Figure 7.11, an undershoot is always present. Moreover, although the first return to zero of the function is very fast, the function oscillates. In this case we should define the return to zero time (at 1%) as the time t_{rz} which is needed for the absolute value of the pulse to be less than 1% for each $t > t_{rz}$.

As an approximation (overestimated) of the return to zero time, we can assume the moment in which the envelope of the maxima reaches 1% of the peak value. Figure 7.12 shows the return to zero at 1% as a function of the parameter r; a minimum of about 170 ns exists at $r \sim 2$, which is already better than the system with PZ cancellation. Moreover, the preamplifier output goes back to zero faster than the shaper output, so it would not suffer from pile-up in a high multiplicity environment.

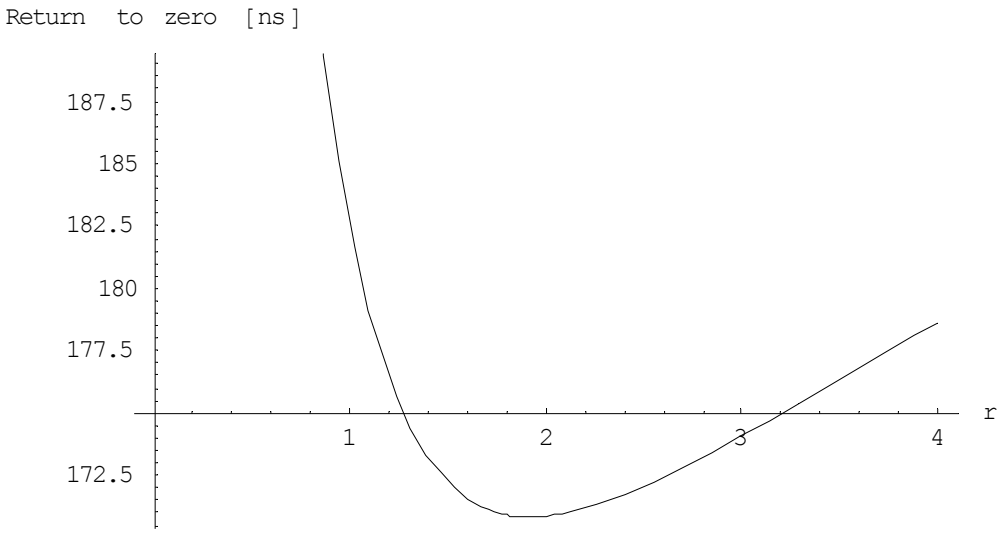

Figure 7.12: return to zero at 1% as a function of the parameter r; a minimum of about 170 ns exists at $r \sim 2$.

The same kind of analysis can be done for the maximum; the maximum of the response to an input charge step of 5000 e^- is reported in Figure 7.13 as a function of r.

If $0 < r < 2$ the value of the peak decreases very rapidly, then tends to saturate to an almost constant value.

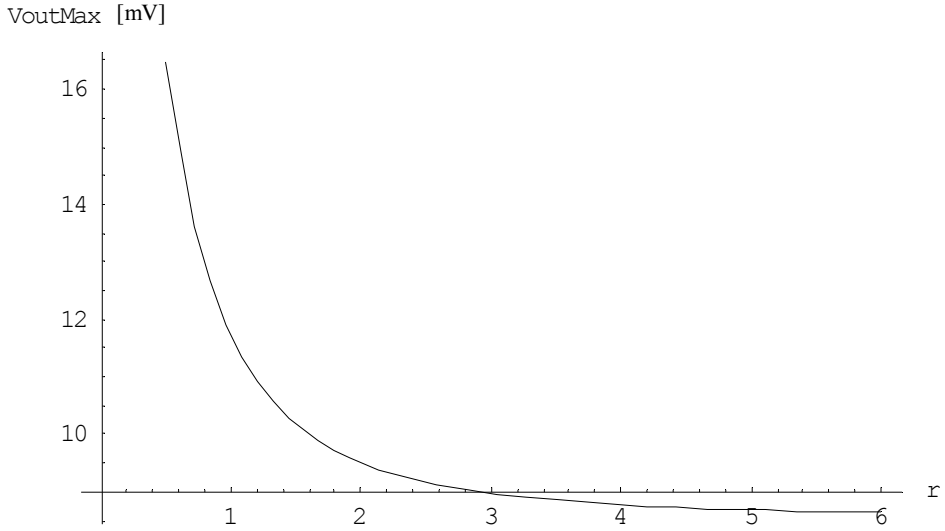

Figure 7.13: Maximum of the response function to an input charge step of 5000 e^- as function of r. If $0 < r < 2$ the value of the peak decreases very rapidly, then tends to saturate to an almost constant value.

7.2.2.4 Noise analysis

For our system, which is not composed by a CSA followed by a semigaussian shaper in open loop, equation (7.2) is not valid anymore. The correct approach in our case is to calculate the input equivalent current noise $|i_{in}|^2$ (because equation (7.19) is a transimpedance transfer function) and use equation (7.25) to calculate the total voltage noise at the output of the system (where the transfer function H is given by equation (7.19)).

$$v_{oA}^2 = \int_0^\infty \left|i_{in}(2j\pi f)\right|^2 \left|H(2j\pi f)\right|^2 df \tag{7.25}$$

It can be shown that the input equivalent current noise $|i_{in}|^2$, calculated for a CSA with a differential input stage (but with one input to a reference potential), is given by the following:

$$\left|i_{in}\right|^2 = 2\, v_{ia}^2 \left|s\right|^2 \left(C_d + C_p + C_f + \frac{C_{IN}}{\sqrt{2}}\right)^2 + i_d^2 \tag{7.26}$$

where i_d^2 is defined in the (7.1) and v_{ia}^2 is defined in Appendix I, and C_{IN} is the CSA input capacitance.

Substituting in equation (7.25) equations (7.19) and (7.26), we can calculate the three noise contributions. The results for the shot noise are given by equation (7.27).

$$ENCo^2 = \frac{e^{\frac{2\,\mathrm{ArcTan}\left[\frac{\sqrt{-\tau_{fb}+4\tau_{p2}}}{\sqrt{\tau_{fb}}}\right]\sqrt{\tau_{fb}}}{\sqrt{-\tau_{fb}+4\tau_{p2}}}}\, Io\,\tau_{p2}}{2\,q} \qquad (7.27)$$

If we consider Io as constant, ENCo is a function of one parameter only; Figure 7.14 shows the value of ENCo as a function of τ_{p2}, for Io = 50 nA. A minimum of about 102 e^- is present at τ_{p2} = 25 ns. This value is less than the 119 e^- found for the PZ cancellation scheme.

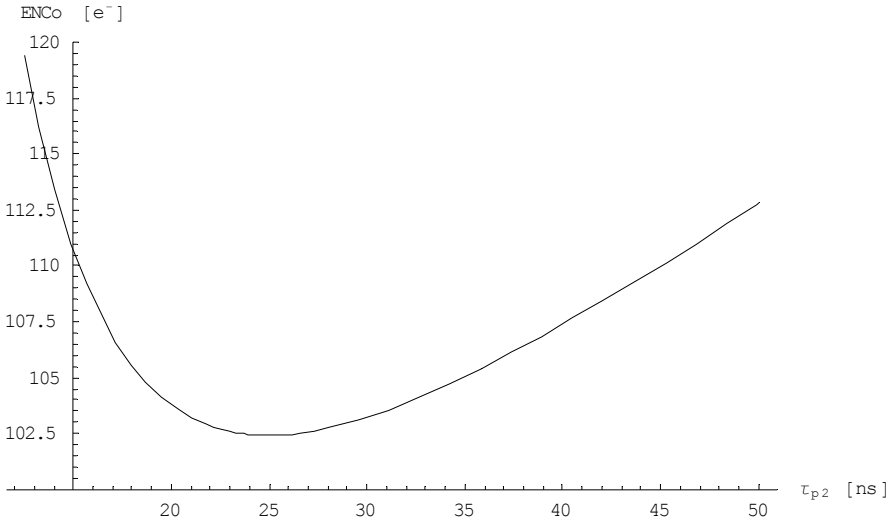

Figure 7.14: ENCo as a function of τ_{p2}, for Io = 50 nA. A minimum of about 102 e^- is present at τ_{p2} = 25 ns.

The results for the thermal noise are given by equation (7.28).

$$ENCd^2 = \frac{8\,k\,T\,n}{q^2}\,\frac{e^{\frac{2\,\mathrm{ArcTan}\left[\frac{\sqrt{-\tau_{fb}+4\tau_{p2}}}{\sqrt{\tau_{fb}}}\right]\sqrt{\tau_{fb}}}{\sqrt{-\tau_{fb}+4\tau_{p2}}}}\,\gamma\,C_{tdiff}^2}{4\,g_m\,\tau_{fb}} \qquad (7.28)$$

where: $C_{tdiff} = C_d + C_p + C_f + \frac{C_{IN}}{\sqrt{2}}$

where the differential front-end has been taken into account with C_{tdiff}. A plot of ENC_d as a function of the two free parameters W and τ_{p2} is shown in Figure 7.15. The lowest noise (of about 65 e^-) is achieved for the minimum[12] τ_{p2} (12.5 ns) and for a W again in a wide range at about 20 μm < W < 40 μm.

[12] The minimum $\tau_{p2} = t_o / 2$ can be found imposing $\tau_{fb} = 4\,\tau_{p2}$ in the (7.23) where r and τ_s are expressed as functions of τ_{fb} and τ_{p2} with the (7.22), using the Taylor expansion around 0 of Arctan [x] = x.

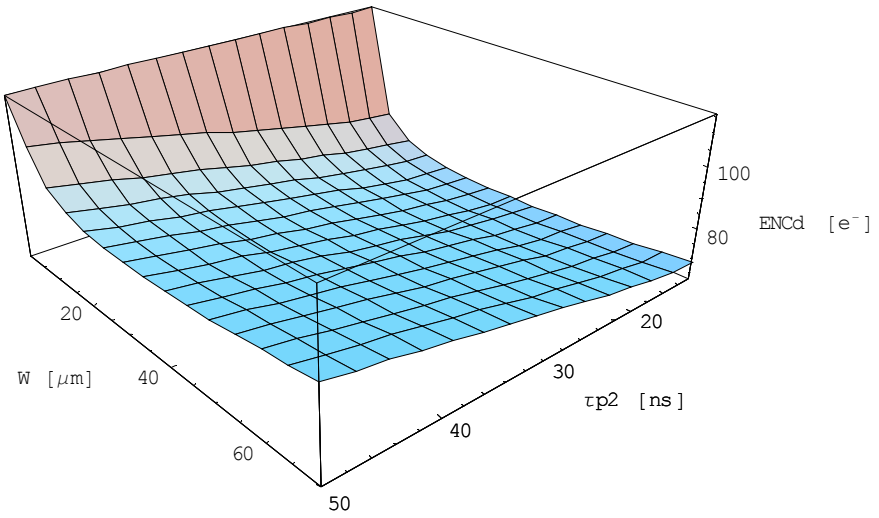

Figure 7.15: ENCd as a function of the two free parameters in equation (7.28) W and τ_{p2}.The lowest noise (about 65 e^-) is achieved for the minimum τ_{p2} (12.5 ns) and again for a W in a wide range at about 20 μm < W < 40 μm.

This value has to be compared with the ENC_d for the PZ cancellation scheme with a differential front-end, which is 69 e^-, and again is not the dominant contribution.

The results for the flicker noise are given by equation (7.29), and shown in Figure 7.16.

$$ENCf^2 = \frac{4\,Ct^2\, e^{\frac{2\,\mathrm{ArcTan}\left[\frac{\sqrt{-\tau_{fb}+4\tau_{p2}}}{\sqrt{\tau_{fb}}}\right]\sqrt{\tau_{fb}}}{\sqrt{-\tau_{fb}+4\tau_{p2}}}}\, Kf\, \mathrm{ArcTan}\left[\frac{\sqrt{-\tau_{fb}+4\tau_{p2}}}{\sqrt{\tau_{fb}}}\right]\tau_{p2}}{L\,q^2\,W\,C_{ox}^2\,\sqrt{\tau_{fb}}\,\sqrt{-\tau_{fb}+4\tau_{p2}}} \qquad (7.29)$$

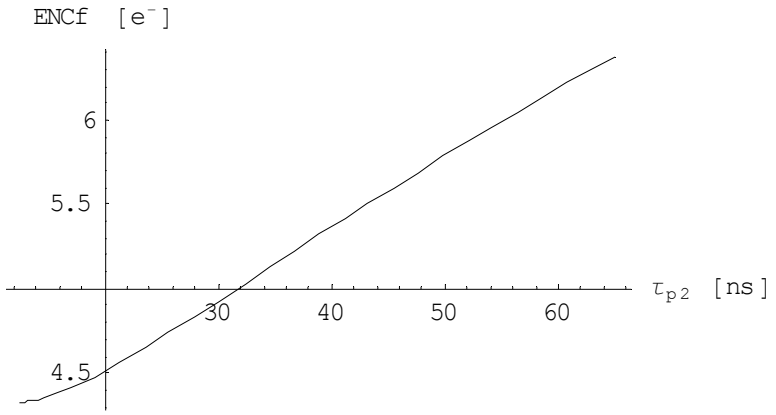

Figure 7.16: Plot of ENC_f as a function of τ_{p2} for W = 30 μm, L = 0.22 μm, I_D = 10 μA; the minimum is about 4.5 e^- for τ_{p2} = 12.5 ns.

Also in this case the flicker noise contribution is negligible; if we plot ENC_f for $W = 30\ \mu m$, $L = 0.22\ \mu m$, $I_D = 10\ \mu A$ the minimum is about $4.5\ e^-$ for the minimum $\tau_{p2} = 12.5$ ns.

If we choose $\tau_{p2} = 25$ ns to minimise the shot noise contribution, which is the dominant one for $I_o = 50$ nA, the total output noise is $124\ e^-$. This has to be compared with the corresponding value for the semigaussian shaping stage, but again taking into account a differential front-end, which gives $138\ e^-$.

Several elements point towards a possible improvement of the two poles system. First of all the shape of the output pulse presents an undershoot, which is usually not wanted. The return to zero time can be improved and the gain has to be increased of about a factor 10, which would be better to realize in a subsequent stage to avoid having a too high gain in the front-end stage. Also for what concerns noise, the performance is comparable to the PZ cancellation scheme, with a dominant shot noise contribution that could be reduced. We will see in the next section that the circuit can be improved significantly by simply adding a third stage with a real pole.

7.2.3 The ALICE1LHCb front-end concept: three poles system

The noise performance of several types of networks has been studied in detail [Tsu61, Bla66, Rad68, Kon68, Hat68], comparing them with the "optimum" filter, which is found to be the cusp function [Now70, Rad88]:

$$H_{opt}(t) = e^{-|t|/\tau} \tag{7.30}$$

In most of the cases the optimisation is relying also on the possibility to choose the peaking/shaping time, and on the possibility to cascade several stages. In our case the (fast) peaking time is fixed by the experiment, and power consumption and area limitations constrain the possibility to use multiple cascaded stages. An interesting analysis is presented in [Ohk76], where a different way to approximate a semigaussian filter is considered. Instead of using a differentiator and a cascade of integrators with the same time constant, he proposes a different Hurwitz factorization, which results in a constellation of complex poles (if they are even) or complex plus a real pole (if they are odd). For a generic order n of the system he calculates the optimum real and imaginary part of each pole. In the case $n = 3$ there are 2 complex conjugate poles and a real pole, and the ratio of the respective real parts is 1.1.

Intuitively, the real part of the poles (in the s plane) represents the exponential decrease of the output pulse, and is dominated by the slowest pole. If all the poles have the same real part, none is dominating the time response and this should improve the return to zero capabilities of the system. The assumption of having three poles with the same real part simplifies greatly the calculations, so that they can be carried out analytically and a closed form for all the formulae can be found. Moreover, the optimum ratio suggested by Ohkawa and coworkers is 1.1, so that the real pole is close to the real part of the complex conjugate

poles. For these reasons we decided to analyse the configuration with three poles on a straight line. This means that we have to place the real pole τ_{p3} at: $\tau_{p3} = 2\,\tau_{p2}$. The (closed loop) poles constellation is shown in Figure 7.17.

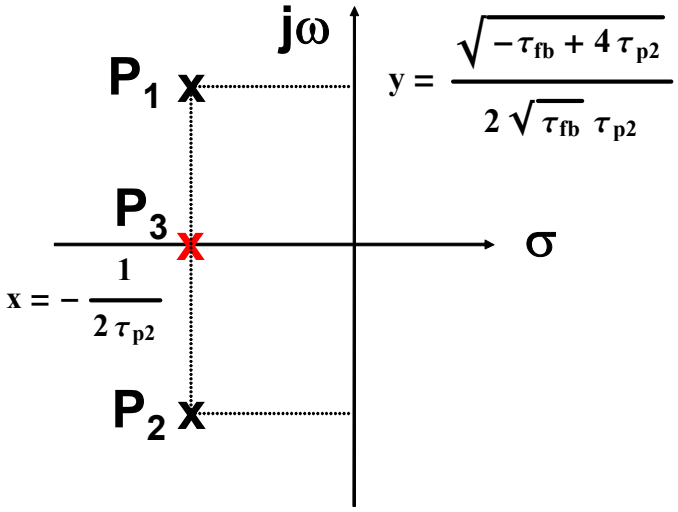

Figure 7.17: Position of the three (closed loop) poles of the system in the s-plane. $P_3 = x$; $P_1 = x + j\,y$; $P_2 = x - j\,y$. x and y are calculated from equation (7.19).

7.2.3.1 The three poles system in closed loop

A schematic diagram of the scheme used to implement the three poles system is shown in Figure 7.18: a stage which realises a real pole τ_{p3} and a gain A_3 is added to the two poles system. The figure is based on the final circuit implementation, which uses full differential shapers (even if one of the inputs of the first shaping stage is tied to a reference voltage).

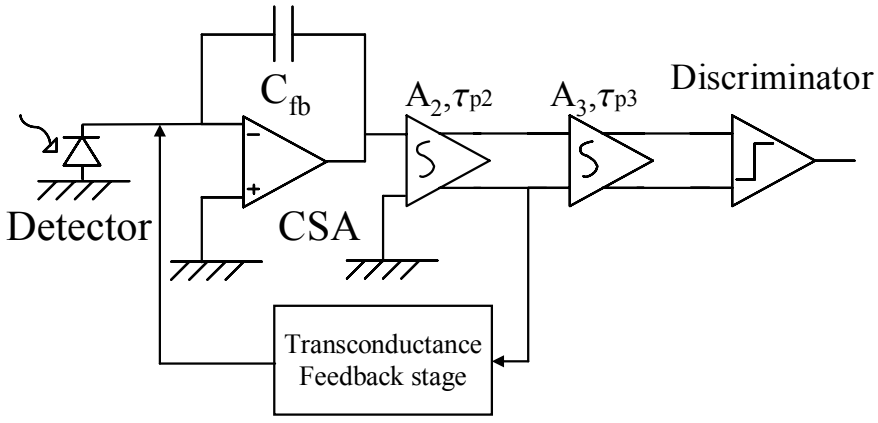

Figure 7.18: Schematic diagram of the three poles system; a stage which realises a real pole τ_{p3} and a gain A_3 is added to the two poles system.

The overall closed loop transimpedance transfer function (defined using the same conventions used for the two poles system) of the system is:

$$H_{cl}(s) = \frac{A_3}{g_{mf}(1 + 2\,s\,\tau_{p2})(1 + s\,\tau_{fb} + s^2\,\tau_{fb}\,\tau_{p2})} \qquad (7.31)$$

Note that in the DC gain A_3 the fact that only one branch of the first shaping stage is fed back to the input has also to be taken into account, as well as the fact that the second shaping stage has a differential output (as shown in Figure 7.18).

Its inverse Laplace transform is:

$$V_{out}(t) = -\frac{2\,A_3\,e^{-\frac{t}{2\tau_{p2}}}\left(1 - \mathrm{Cos}\left[\frac{t\sqrt{-\tau_{fb}+4\tau_{p2}}}{2\sqrt{\tau_{fb}\;\tau_{p2}}}\right]\right)}{g_{mf}\,\tau_{fb} - 4\,g_{mf}\,\tau_{p2}} \qquad (7.32)$$

The position of all the maxima and the minima of V_{out} are given by equations (7.33) and the value of the first maximum (the peak) by equation (7.34):

$$t_{maxk} = \frac{4\,K\,\pi\,\sqrt{\tau_{fb}\;\tau_{p2}}}{\sqrt{-\tau_{fb}+4\,\tau_{p2}}} + \frac{4\,\mathrm{ArcTan}\left[\frac{\sqrt{-\tau_{fb}+4\,\tau_{p2}}}{\sqrt{\tau_{fb}}}\right]\sqrt{\tau_{fb}\;\tau_{p2}}}{\sqrt{-\tau_{fb}+4\,\tau_{p2}}}$$

$$t_{mink} = 2\,K\,\pi\,\frac{\sqrt{\tau fb}\;\tau p2}{\sqrt{-\tau fb+4\,\tau p2}} \qquad (7.33)$$

$$\text{for}\,\{K = 0\,..\,\infty\}$$

$$V_{outmax} = \frac{A_3\,e^{-\frac{2\,\mathrm{ArcTan}\left[\frac{\sqrt{-\tau fb+4\,\tau p2}}{\sqrt{\tau fb}}\right]\sqrt{\tau fb}}{\sqrt{-\tau fb+4\,\tau p2}}}}{g_{mf}\,\tau_{p2}} \qquad (7.34)$$

An example of the output for 5000 e^-, assuming $A_3 = 14.4$, $g_{mf} = 1.4\ \mu S$, $\tau_{fb} = 21.9$ ns and $\tau_{p2} = 6.7$ ns is shown in Figure 7.19. It can be remarked that the undershoot that was present in the two poles system is no longer present and, from equation (7.32), that no undershoot can be present for any value of τ_{fb} and τ_{p2}.

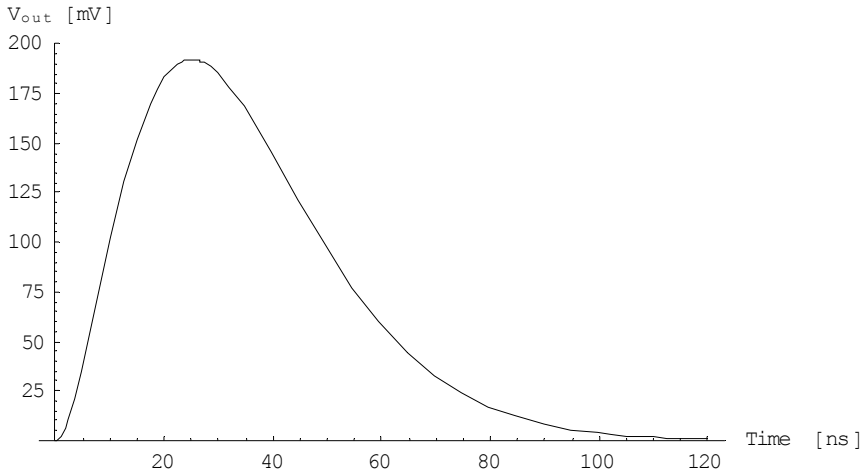

Figure 7.19: Example of the output for 5000 e^-, assuming g_{mf} = 1.4 μS, τ_{fb} = 21.9 ns and τ_{p2} = 6.7 ns, τ_{p3} = 13.4 ns.

Also for the three poles system we can define the two parameters $\tau_s = 1/y$ and $r = x/y$, where x and y are respectively the real and the imaginary part of the two complex conjugate poles, so that we can obtain a simple relation which links the two parameters for a fixed peaking time t_o:

$$t_o = 2\,\mathrm{ArcTan}\left[\frac{1}{r}\right]\tau_s \qquad (7.35)$$

If we define the return to zero time as in the case with two poles, using again the envelope of the maxima, we can plot t_{rz} as a function of r, as shown in Figure 7.20. A minimum of about 107 ns exists for r ~ 1.5.

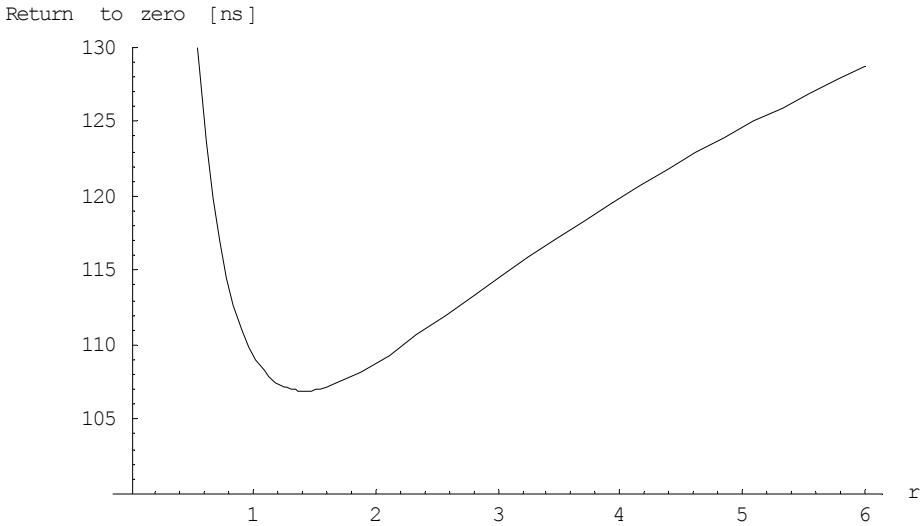

Figure 7.20: Return to zero time for the three poles system as a function of the parameter r.

Figure 7.21 plots the peak value of the output as a function of the parameter r, for an input charge of 5000 e^-, with g_{mf} = 1.4 μS. It stabilises at ~ 200 mV for r higher than about two.

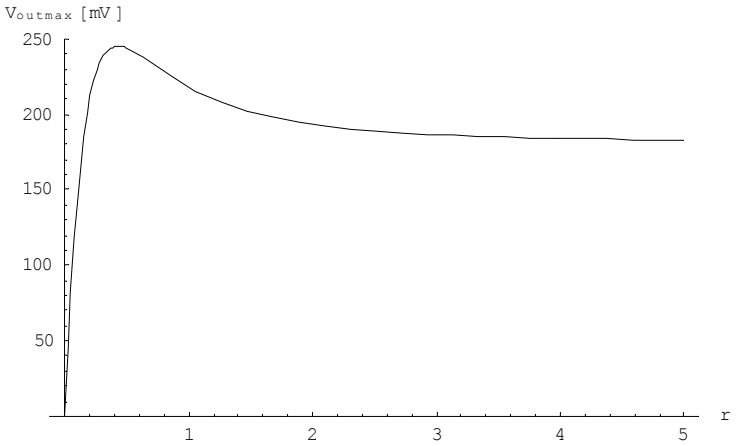

Figure 7.21: Peak value of the output as a function of the parameter r, for an input charge of 5000 e^-, with g_{mf} = 1.4 μS and A_3 = 14.4.

7.2.3.2 Noise analysis

The results of the calculations for the shot noise component are given by equation (7.36).

$$ENCo^2 = \frac{3\,e^{\frac{4\,\mathrm{ArcTan}\left[\frac{\sqrt{-\tau_{fb}+4\,\tau_{p2}}}{\sqrt{\tau_{fb}}}\right]\sqrt{\tau_{fb}}}{\sqrt{-\tau_{fb}+4\,\tau_{p2}}}}\,Io\,\tau_{p2}^2}{6\,q\,\tau_{fb}+8\,q\,\tau_{p2}} \qquad (7.36)$$

Figure 7.22 plots the shot noise component as a function of τ_{p2}, supposing I_o = 50 nA. A minimum of about 83 e^- is present at τ_{p2} ~ 18 ns.

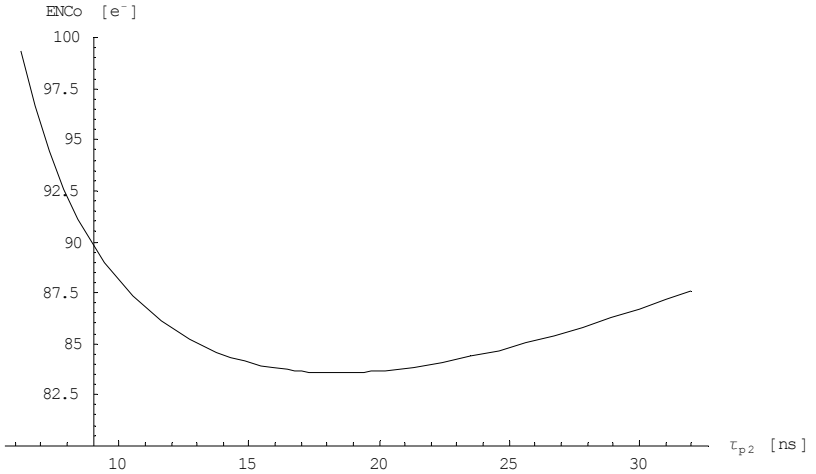

Figure 7.22: Plot of the shot noise component as a function of τ_{p2}, supposing I_o = 50 nA. A minimum of about 83 e^- is present at τ_{p2} ~ 18 ns.

The results of the calculations for the thermal noise component are given by equation (7.37).

$$ENC_d^2 = \frac{2\,e^{\frac{4\,\mathrm{ArcTan}\left[\frac{\sqrt{-\tau_{fb}+4\,\tau_{p2}}}{\sqrt{\tau_{fb}}}\right]\sqrt{\tau_{fb}}}{\sqrt{-\tau_{fb}+4\,\tau_{p2}}}}\; k\,n\,T\,\gamma\,C_{tdiff}^2\,\tau_{p2}}{3\,gm\,q^2\,\tau_{fb}^2 + 4\,gm\,q^2\,\tau_{fb}\,\tau_{p2}} \qquad (7.37)$$

Figure 7.23 plots the thermal noise component as a function of τ_{p2} and W. A wide minimum of about 67 e^- is present for the minimum[13] τ_{p2} = 6.25 ns and 20 μm < W < 40 μm.

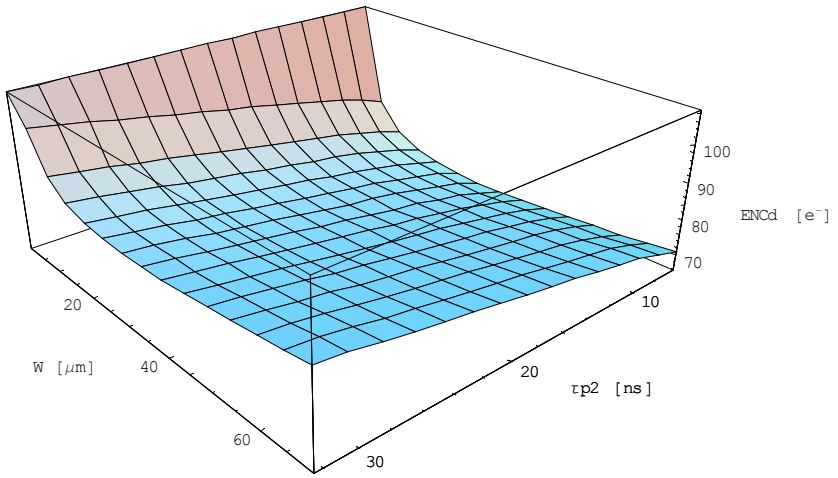

Figure 7.23: ENC_d as a function of the two free parameters in (7.37) W and τ_{p2}.The lowest noise (of about 67 e^-) is achieved for the minimum τ_{p2} (6.25 ns) and for a W in a wide minimum at about 20 μm < W < 40 μm.

The flicker noise contribution can be evaluated with equation (7.38).

$$ENC_f^2 = \frac{2\,e^{\frac{4\,\mathrm{ArcTan}\left[\frac{\sqrt{-\tau_{fb}+4\,\tau_{p2}}}{\sqrt{\tau_{fb}}}\right]\sqrt{\tau_{fb}}}{\sqrt{-\tau_{fb}+4\,\tau_{p2}}}}\; Kf\,C_{tdiff}^2\,\tau_{p2}^2\left(\mathrm{Log}\left[\frac{4\,\tau_{p2}}{\tau_{fb}}\right]\sqrt{\tau_{fb}} + \mathrm{ArcTan}\left[\frac{\sqrt{-\tau_{fb}+4\,\tau_{p2}}}{\sqrt{\tau_{fb}}}\right]\sqrt{-\tau_{fb}+4\,\tau_{p2}}\right)}{L\,q^2\,W\,C_{ox}^2\,\sqrt{\tau_{fb}}\,(4\,\tau_{p2}-\tau_{fb})\,(3\,\tau_{fb}+4\,\tau_{p2})} \qquad (7.38)$$

The contribution of the flicker noise is negligible; if we plot ENC_f as a function of W for L = 0.22 μm, I_D = 10 μA for the minimum τ_{p2} = 6.25 ns, the noise is already less than 4 e^- for W > 20μm.

[13] The minimum $\tau_{p2} = t_o/4$ and can be found as explained in footnote 12.

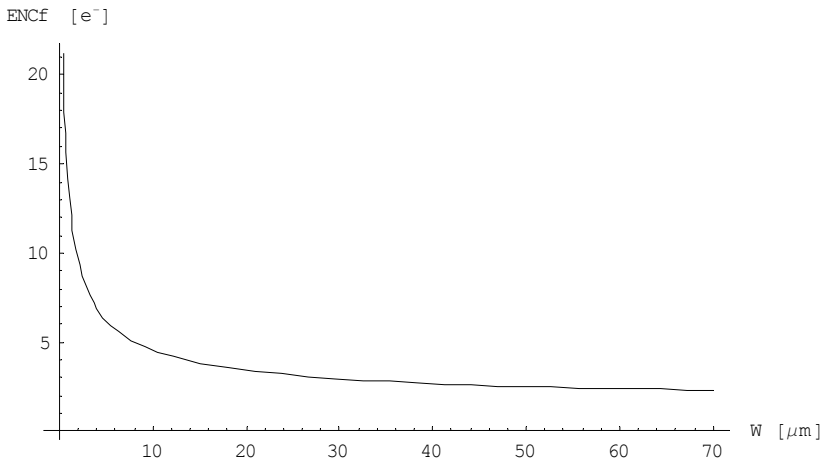

Figure 7.24: Plot of ENC_f as a function of W for $L = 0.22$ μm, $I_D = 10$ μA for the minimum $\tau_{p2} = 6.25$ ns. The noise contribution is already less than 4 e^- for W > 20μm.

7.2.3.3 Sensitivity analysis

A very important aspect in a design is the sensitivity of the circuit (e.g. gain, peaking time, noise) to the variations of the physical parameters (e.g. threshold voltage and β of the transistors) and their mismatch.

The relative sensitivity of a quantity F to a design parameter x is defined as:

$$S_x^F = \frac{x\,\partial F}{F\,\partial x} \tag{7.39}$$

and is the relative change of F per unitary relative change of x. It can be a function of x and of the other variables on which F depends.

In our case we want to be sure that a little mismatch in the position of the open loop poles (which are the poles that we physically place in the circuit) does not result in a major shift of the closed loop poles. In particular, we want to be sure that the imaginary complex poles do not become real, as this would degrade significantly the performance (especially in terms of return to zero time) of the circuit. The imaginary part of the two complex poles is:

$$y = \frac{\sqrt{-\tau_{fb} + 4\,\tau_{p2}}}{2\sqrt{\tau_{fb}}\,\tau_{p2}} \tag{7.40}$$

and the relative sensitivities of y to τ_{fb} and τ_{p2}:

$$S_y^{\tau_{fb}} = \frac{\tau_{fb}}{y}\,\frac{\partial y}{\partial \tau_{fb}} = \frac{2\,\tau_{p2}}{\tau_{fb}^2 - 4\,\tau_{fb}\,\tau_{p2}} \tag{7.41}$$

$$S_y^{\tau_{p2}} = \frac{\tau_{p2}}{y}\,\frac{\partial y}{\partial \tau_{p2}} = \frac{\tau_{fb} - 2\,\tau_{p2}}{-\tau_{fb} + 4\,\tau_{p2}}$$

Solving $S_y^{\tau_{fb}} < 1$ gives as constraints on the open loop poles: $\tau_{fb} < 3\,\tau_{p2}$; also imposing $S_y^{\tau p2} < 1$ results in the constraint $\tau_{fb} < 3\,\tau_{p2}$.

Equation (7.41) shows the importance of this kind of analysis: several elements in the previous calculations would suggest the choice $\tau_{fb} = 4\,\tau_{p2}$ (three real coincident poles), but we can see that $S_y^{\tau_{fb}}$ and $S_y^{\tau p2}$ are infinite for $\tau_{fb} = 4\,\tau_{p2}$.

If we set $\tau_{fb} = 3\,\tau_{p2}$ and the peaking time at 25 ns, from equation (7.35) (where τ_s and r are expressed as functions of τ_{fb} and τ_{p2}, see equation (7.22)) we obtain the position of the poles, that is: $r = \sqrt{3}$, $\tau_{fb} = 20.7$ ns and $\tau_{p2} = 6.9$ ns. Once we have found the poles, we can calculate again all the noise contributions $ENC_o = 96\ e^-$, $ENC_d = 67\ e^-$, $ENC_f = 3.4\ e^-$, $ENC_{tot} = 117\ e^-$. The return to zero time is $t_{rz} = 107$ ns and the maximum is 197 mV (all the parameters used for the calculations are set as previously stated).

7.2.4 The ALICE1LHCb front end concept: the four poles system

The system with three poles described in the previous section meets all the requirements for the ALICE1LHCb design, and in particular has a fast return to zero both of the shaper output and of the preamplifier output, so that it can be used in high multiplicity environments. However, since we have fixed $\tau_{p2} = 6.9$ ns, the hypothesis that the CSA rise time constant $>> \tau_{p2}$ is no longer valid. This is why an analysis of a four poles system is needed.

The four poles system transfer function is given by equation (7.42):

$$H_{cl}(s) = \frac{A_3}{g_{mf}\,(1 + s\,\tau_{p3})\,(1 + s\,\tau_{fb} + s^2\,\tau_{fb}\,(\tau_{p2} + \tau_r) + s^3\,\tau_{fb}\,\tau_{p2}\,\tau_r)} \qquad (7.42)$$

where τ_{p3} is the pole of the second shaping stage, and τ_r is the time constant of the parasitic pole of the CSA. The difference between equation (7.31) and equation (7.42) is that a third order polynomial is present at the denominator, so that to find the closed loop poles we have to solve a third order equation which involves also τ_r. Equation (7.42) will be used for a precise numerical evaluations of all the parameters of interest, but it is interesting to point out that, as a second order approximation, the term in s^3 can be neglected in the band of our system, so that equation (7.42) becomes identical to equation (7.31) making the substitution $\tau_{p2}' = \tau_{p2} + \tau_r$. For our circuit the rise time pole can be estimated to be about 2 - 2.5 ns. So all the analysis done in the previous section is still approximately valid as long as we set in the circuit $\tau_{p2} = 6.9$ ns - 2 ns = 4.9 ns.

After the circuit implementation and simulations described in the next sections, the final values for the four poles of the circuit are: $\tau_r = 2$ ns, $\tau_{fb} = 22$ ns, $\tau_{p2} = 5.1$ ns, $\tau_{p3} = 15.1$ ns, $A_3 = 14.4$. Figure 7.25 plots the output of the four poles system for $Q_{in} = 5000\ e^-$, assuming $g_{mf} = 1.4\ \mu S$. The peak is 192 mV at 26 ns, while the return to zero time at 1% is 114 ns.

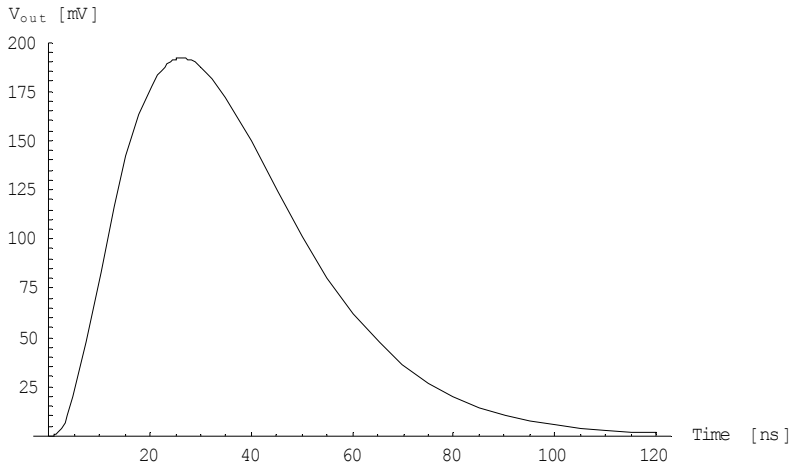

Figure 7.25: Example of the output of the four poles system for Q_{in} = 5000 e^-, assuming g_{mf} = 1.4 μS, τ_r = 2 ns, τ_{fb} = 22 ns, τ_{p2} = 5.1 ns, τ_{p3} = 15.1 ns, A_3 = 14.4.

Figure 7.26 shows the preamplifier output for the same parameters used to plot Figure 7.25. Features to note are the fast peaking time (~7 ns) and return to zero time at 1% (~65 ns) of the preamplifier, that were the main reasons to discard the choice of a standard CSA with semigaussian shaping scheme.

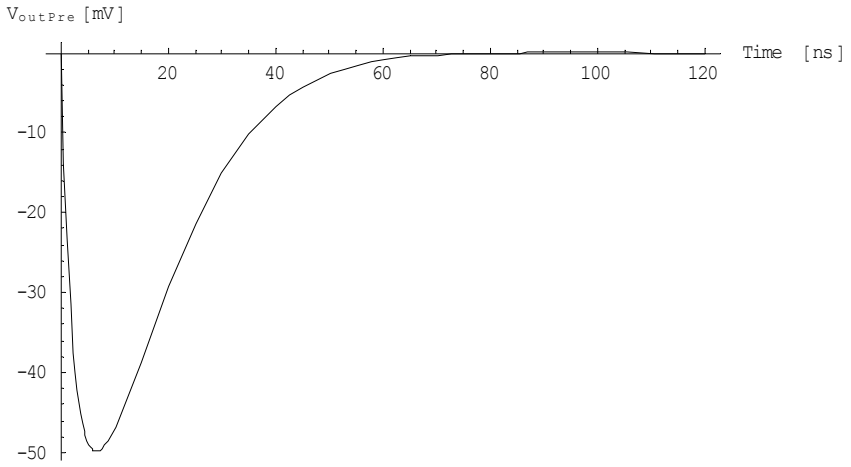

Figure 7.26: Preamplifier output (for the same parameters used to plot Figure 7.25). The peaking time is ~7 ns and return to zero time at 1% is ~65 ns.

7.2.4.1 Sensitivity analysis

The sensitivity of the peaking time and of the peak value with the positions of the four poles of the system is presented in the following figures.

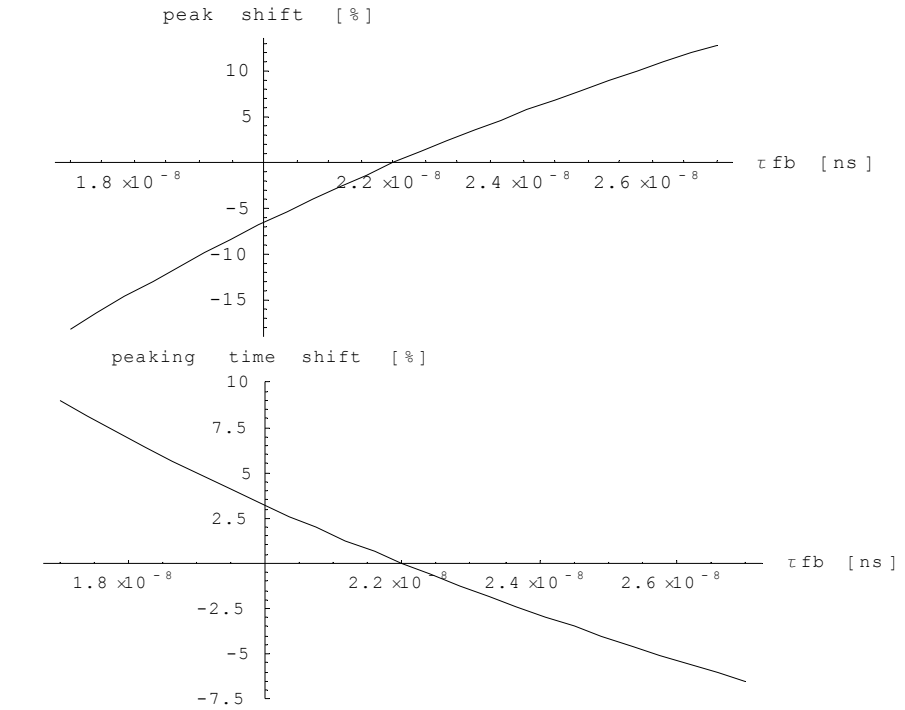

Figure 7.27: Sensitivity analysis of the peak value (top) and of the peaking time (bottom) to the position of the pole τ_{fb}.

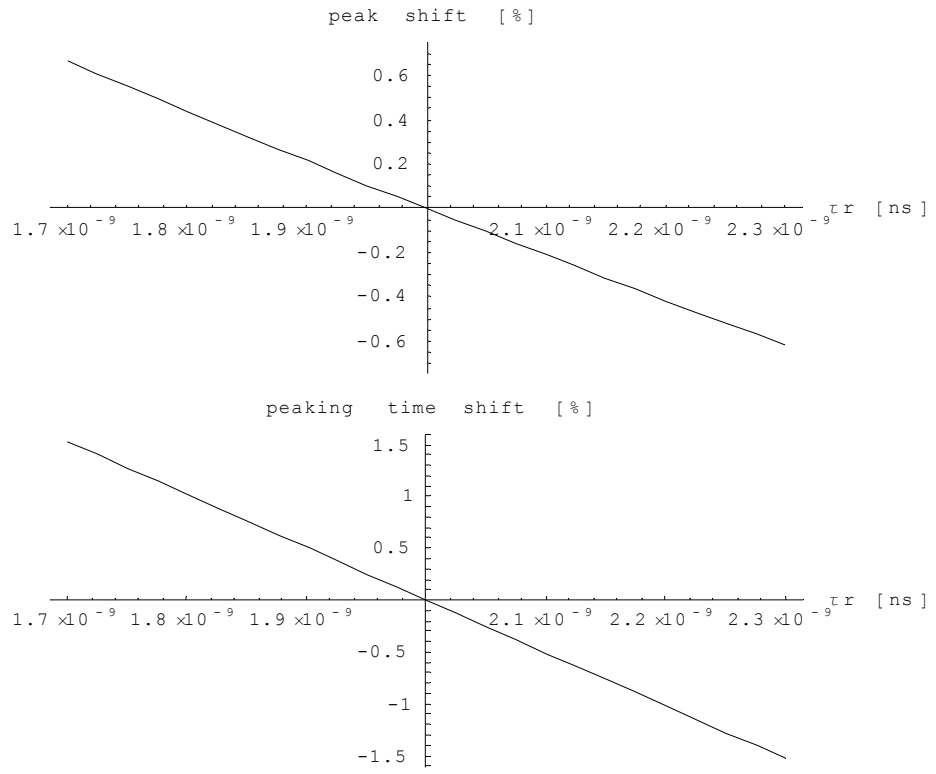

Figure 7.28: Sensitivity analysis of the peak value (top) and of the peaking time (bottom) to the position of the pole τ_{r}.

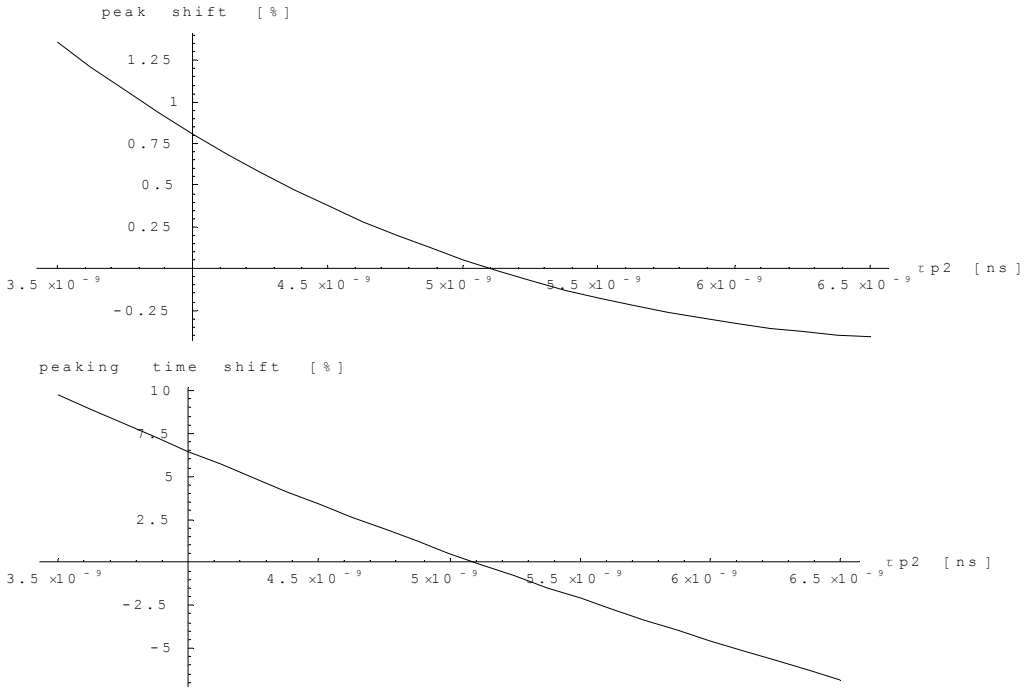

Figure 7.29: Sensitivity analysis of the peak value (top) and of the peaking time (bottom) to the position of the pole τ_{p2}.

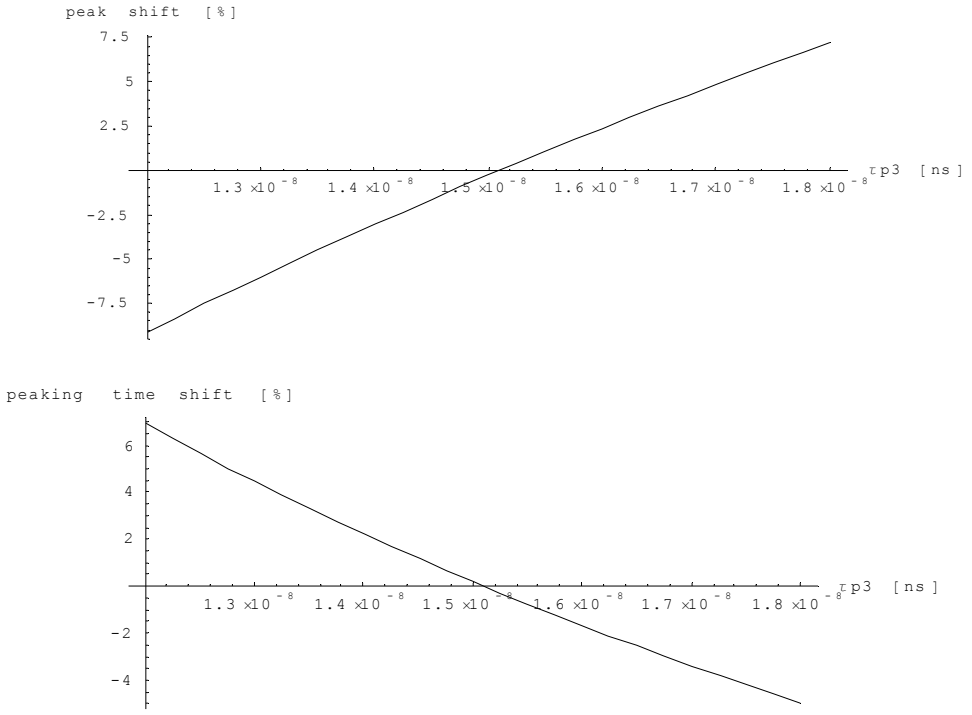

Figure 7.30: Sensitivity analysis of the peak value (top) and of the peaking time (bottom) to the position of the pole τ_{p3}.

The relative sensitivity (as defined in equation (7.39)) of the peaking time and of the peak value with the positions of the four open loop poles (calculated at their nominal values) is shown in Table 7.4. The higher sensitivity is of course the sensitivity to the position of the feedback pole.

	τ_{fb}	τ_r	τ_{p2}	τ_{p3}
S_R peak value	66 %	-4 %	-3 %	41 %
S_R peaking time	-33 %	-10 %	-28 %	-29 %

Table 7.4: Relative sensitivity (as defined in equation Figure 7.39) of the peaking time and of the peak value with the positions of the four open loop poles (calculated at their nominal values).

7.2.4.2 Noise analysis

The noise analysis performed numerically on the 4-poles system has to take into account an additional noise contribution present in the final layout. The transistor which realises the leakage current compensation is connected to the preamplifier input, so its noise contribution should not be neglected. The complete scheme can be found in Figure 7.33, Figure 7.38 and Figure 7.43, but the working principle for the complete noise calculation is shown in Figure 7.31.

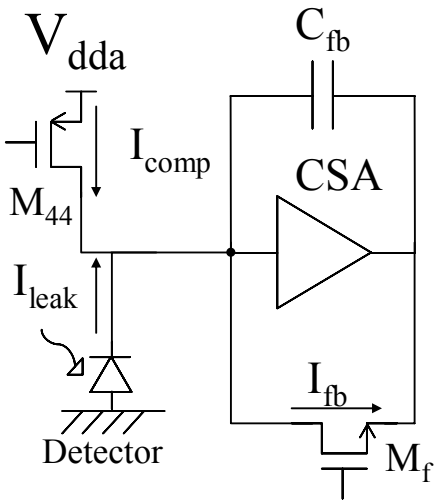

Figure 7.31: Schematic representation of the currents present at input node.

The DC current flowing in M_f is constant I_{fb} = 40 nA. As will be explained in 7.3.6, M_{44} provides a current which compensates for the unknown leakage current coming from the detector, so that: $I_{comp} = I_{fb} - I_{leak}$. M_f is in weak inversion, so its noise contribution is 2 q I_{fb}, while the contribution of the detector is 2 q I_{leak}. The contribution of M_{44} comes from its drain current noise. Regardless of the inversion region of M_{44} we can write:

$$i^2_n = 2\ q\ (2U_t\ n\ g_m\ \gamma) \qquad (7.43)$$

where Ut is the thermal noise and γ is a factor that takes into account the excess noise factor Γ and the region of inversion of the transistor (as explained in Appendix I).

This means that we can group all the current noise contributions in one term:

$$i^2_{ntot} = 2\,q\,(I_{fb} + I_{leak} + 2U_t\,n\,g_m\,\gamma) \qquad (7.44)$$

The shot noise contribution to the output noise for $I_{leak} = 0$ (i.e. $I_{fb} = 40$ nA) is 95 e^-; the contribution of the thermal noise is 68 e^- and the flicker noise is 3 e^-. The total noise is then $ENC_T = 117\ e^-$. The total noise at the output of the four poles system as a function of the leakage current is plotted in Figure 7.32.

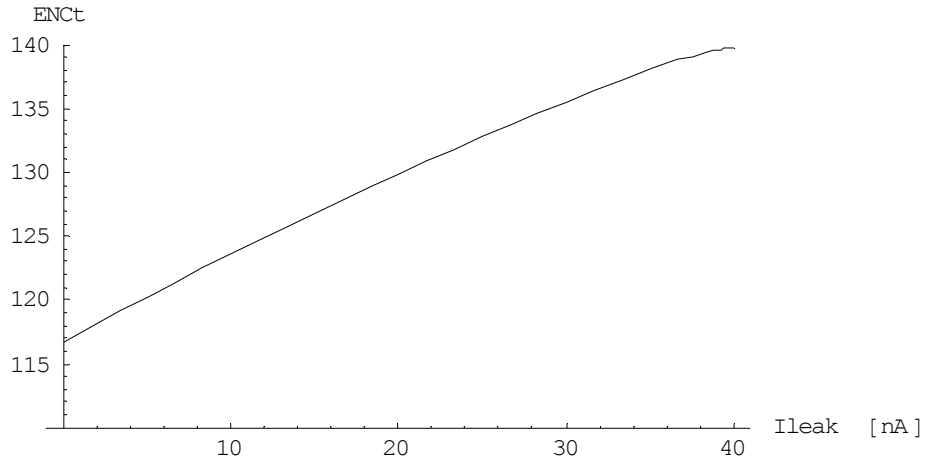

Figure 7.32: Total output noise of the four poles system as a function of the detector leakage current. The total noise for $I_{leak} = 0$ is $ENC_T = 117\ e^-$.

The figure shows that the ENC_T does not change much for increasing leakage currents. This is due to the fact that the contribution due to the leakage current itself increases, but the current in M_{44} decreases and its noise contribution decreases.

7.3 The ALICE1LHCb front-end implementation

Once the best front-end scheme for our needs has been chosen and the most important design parameters calculated, it is possible to start the design phase, which consists of designing and optimising each single circuital block separately and in the system as a whole. This phase of the IC design is greatly simplified if it is supported by an analytical analysis which can give the designer an idea of the behaviour of the system for the different parameters.

7.3.1 The Charge Sensitive Preamplifier

A classical implementation for a CSA is a cascoded or folded cascoded scheme. This scheme has the advantage of having a high gain while still keeping a relatively wide band and

allows to precisely define the feedback capacitance C_{fb}. In effect this was the configuration used for the two ALICEtest chips. From the measurements done on these test chips it became clear that the minimum threshold and the noise performance of the chip were dominated mainly by the digital-to-analogue crosstalk, so in the new design this issue was addressed with high priority. The calculations done in the previous sections give a noise performance of the chip which is better than what is needed for the experiments, so that the electronic noise can be sacrificed to some extent to improve digital-to-analogue crosstalk sensitivity. Figure 7.33 shows the detailed circuit diagram of the CSA.

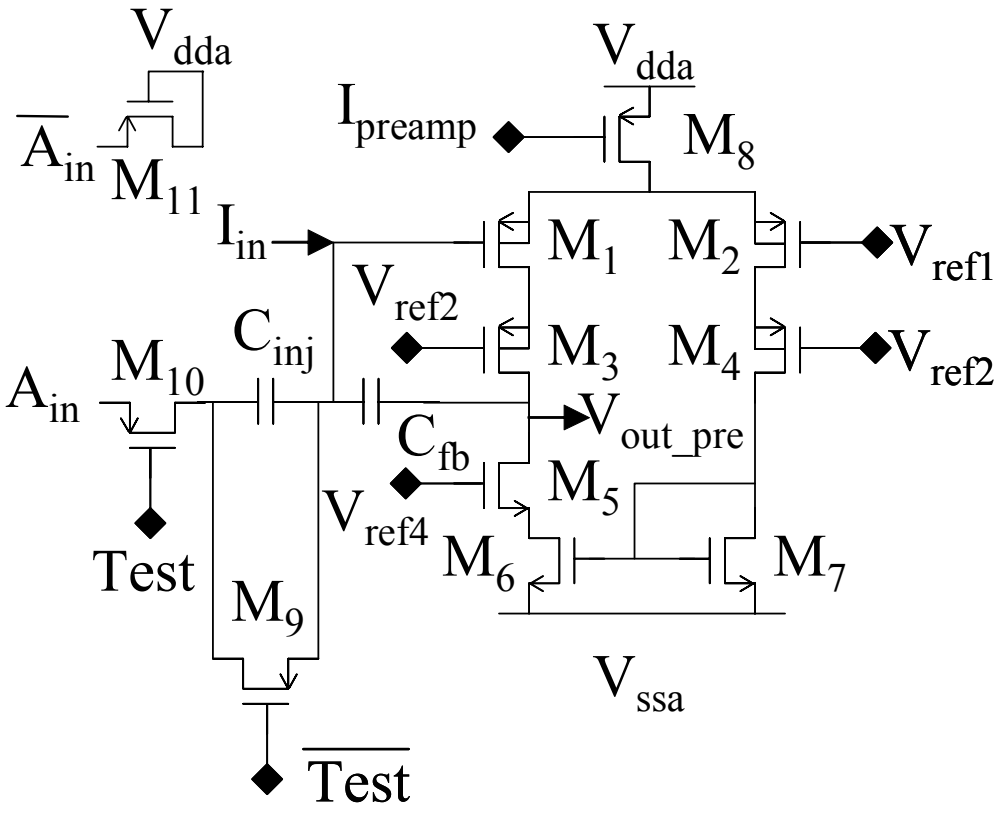

Figure 7.33: Circuit diagram of the ALICE1LHCb CSA. The substrate of all the n-channel transistors is connected with a dedicated power supply called "sub" (not shown in the picture). The well of all the p-channel transistors is connected with the dedicated supply voltage "wella" (not shown in the pictures); in the few cases where this is not the case the transistor well terminal is drawn and connected to the corresponding potential (e.g. transistors M_1 and M_2).

In all the circuit diagrams the substrate of all the n-channel transistors is connected with a dedicated power supply voltage called "sub" (not shown in the pictures). The well of all the p-channel transistors is connected with the dedicated supply voltage "wella" (not shown in the pictures). In the few cases where this is not the case the transistor well terminal is drawn and

connected to the corresponding potential. The power distribution strategy will be explained in more detail in section 7.6.3

The type of the input transistor is then chosen to be p-channel, as explained in section 7.2.1. The size of the input transistor can be chosen on the basis of the analysis done in the previous sections. An almost minimum $L = 0.28\ \mu m$ was chosen. The thermal noise contribution has a minimum at $20\ \mu < W < 40\ \mu m$, while the flicker noise (whose contribution is negligible) does not change much for $W > 30\ \mu m$. For these reasons and to keep area to a minimum we chose $W = 30\ \mu m$.

To try to reduce the sensitivity to analogue-to-digital crosstalk, all the front-end chain was designed to be differential. Unfortunately, it was not possible to design a fully differential[14] CSA because the sensor cannot feed the preamplifier input with a differential signal. Anyhow the choice of a differential scheme is effective in reducing the substrate coupling, because any signal in the substrate (which would couple into a single-ended cascoded stage as input signal) is a common mode signal for a differential pair, and is then greatly attenuated at the output (the degree of attenuation depends on the Common Mode Rejection Ratio (CMRR) of the preamplifier). The use of a differential input results in an increased electronic noise (for a constant power budget) which is however still below the design specifications. The power budget for the CSA cannot exceed 20 μA, so we fixed the preamp current at this value.

The input transistor has then a DC current of about 10 μA, which results in an inversion factor (defined in section 5.5.3.3) $i_f = 0.75$, so it is in moderate inversion. The CSA, as the rest of the chip, is powered[15] at $V_{dd} = 1.6$ V. This fixes the preamplifier power consumption to 32 μW.

The value of the feedback capacitance C_{fb} is, as we have seen, a trade-off between several factors, namely speed of the circuit, stability, gain, and pixel-to-pixel gain stability. For these reasons in our case it has to be of the order of tens of femto Farads. Moreover, together with the gain of the first shaping stage (A_2) and the transconductance of the feedback transistors (g_{mf}) it contributes to set the feedback pole $\tau_{fb} = C_{fb} / (A_2\ g_{mf})$. The final value was the result of the design optimisation procedure, and ended up to be $C_{fb} = 15$ fF.

The transistors M_3 and M_4 are cascode transistors, with the aim of decreasing the impedance seen by the drain of the input transistors. The transistor M_5 increases the output resistance of the stage, thus increasing its gain. Moreover, increasing the output resistance of the n-channel load avoids having it as the dominant contribution to the preamp gain. In effect it is not good design practise to rely on the precise modelling of the output conductance of an enclosed transistor, because it is not well modelled by the simulator.

[14] A stage which has a differential input and a differential output is called fully differential.
[15] We will see in Chapter 8 that during chip tests it was decided to increase the power supply to 1.8 V.

The transistors M_9 and M_{10} are switches, and are driven by an on-pixel flip flop (TEST). If the logic signal TEST is high, the injection capacitance C_{inj} is connected at the preamplifier input and at the Analogue Test Input signal A_{in}. If TEST is low, the capacitance is shorted and disconnected from $A_{in.}$ The transistor M_{11} is a dummy transistor used to balance the capacitive load seen by the test pulse generator, and is connected to the inverted analogue test pulse $\overline{A_{in}}$. The transistor M_8 is the CSA current source. More details about the on-chip test structures will be discussed in section 7.6.6. Figure 7.34 shows the CSA output[16] for a 5000 e^- input pulse.

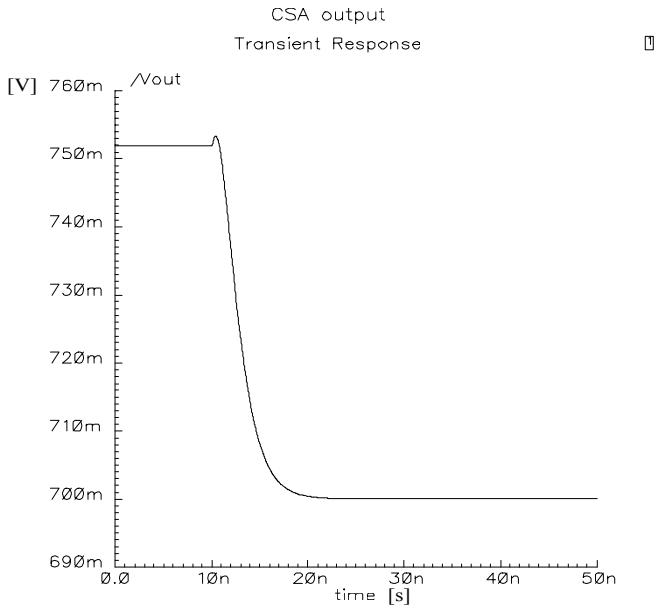

Figure 7.34: Plot of CSA output as a function of time for a 5000 e^- input (C_{fb} = 15 fF). In all the figures the input is applied at t = 10 ns.

The rise time is 4.5 ns, so we can estimate the preamplifier time constant to be $\tau_r \sim 2$ ns. In all the figures the input signal is applied at t = 10 ns.

7.3.2 The first shaping stage

The first shaping stage, whose circuit diagram is depicted in Figure 7.35, has the aim of adding the open loop pole τ_{p2} in the system, so that it is possible to realize two complex poles in closed loop. Moreover its DC gain contributes to the position of the feedback pole τ_{fb}.

For these reasons it was extremely important to have a stage with a well controlled gain and output impedance. Several circuit solutions were tested (e.g. unity gain buffer); the best (especially for what concerns circuit stability) is the solution shown in Figure 7.35, which does not have local feedback. In this configuration the gain of the stage is known with much more precision, as it is given by the ratio of the transconductance of the input transistors (M_{12}

[16] The CSA is simulated alone, without the rest of the circuit.

and M_{13}) to the transconductance of the output transistors (M_{14} and M_{15})[17]. Both the input and the output transistors are p-channel, so we can rely on the precise model coming from the foundry. Moreover, a change of the DC current in the circuit will affect the two transconductances in the same way, decreasing the circuit sensitivity (both of the gain A_2 and of the pole τ_{fb}) to the bias current.

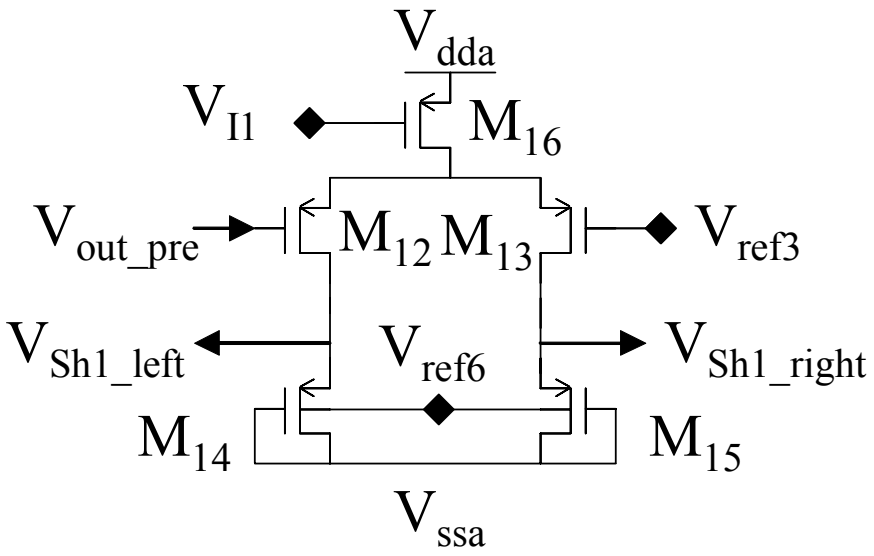

Figure 7.35: Circuit diagram of the ALICE1LHCb first shaping stage.

Again, the current budget for this stage was rather limited, as the major current use had to be done at preamplifier level. The bias current (generated by the current source M_{16}) was set at $I_1 = 3$ μA; so the power consumption of the first shaping stage is 4.8 μW. The circuit is differential for noise rejection reasons.

Since the input transistor is not critical for the circuit performance, its length was chosen to be L = 0.42 μA, larger than the minimum one as suggested by the technology design manual. The pole τ_{p2} is given by the resistance seen at the shaper output multiplied by the capacitance seen at the same node. The main contribution to this capacitance is due to the input capacitance of the second shaping stage, so the design of this stage had to be tailored at the end of the full front-end design, once the input capacitance of the second shaping stage and of the feedback transistors was known. The total capacitance is $C_{p2} = 155$ fF. The conductance seen at the output node is the source transconductance g_{ms_out} of the load transistors (M_{14} and M_{15}) plus the source transconductance g_{mf} of the feedback transistors, $g_{p2} = g_{ms\_out} + g_{mf} = 30.4$ μS. The pole is then $\tau_{p2} = C_{p2} / g_{p2} = 5.1$ ns.

Figure 7.36 plots the first shaping stage output as a function of time for a 10 mV input step[18].

[17] The simulator calculates the g_m of a transistor much more precisely than its g_{ds}.
[18] Simulated alone in open loop.

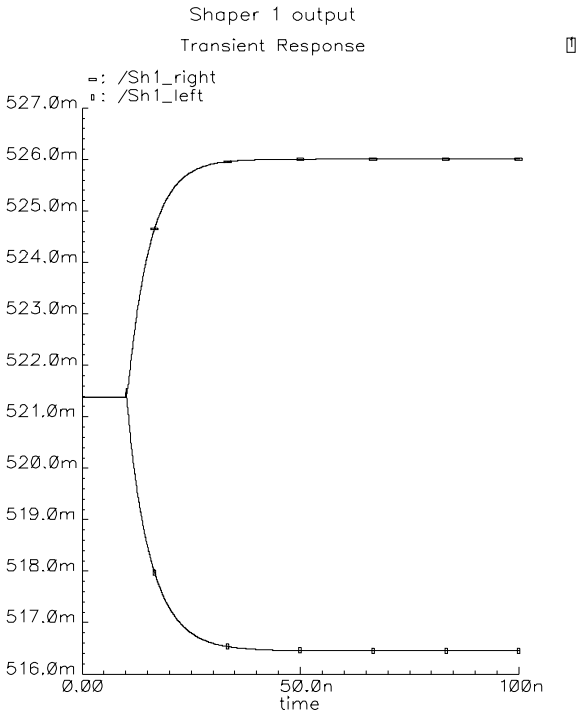

Figure 7.36: Plot of the first shaping stage transient outputs (left and right branches) as a function of time for a 10 mV input step.

Figure 7.37 plots the first shaping stage AC outputs (left and right branches) as a function of the input frequency. From the -3 dB point we can find the dominant pole to be at τ_{p2} = 5.3 ns, which is in good agreement with the analytical calculations.

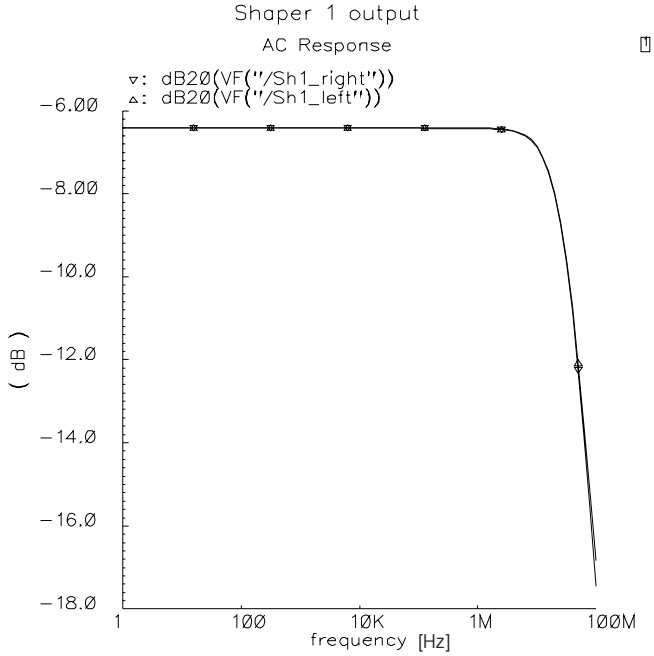

Figure 7.37: Plot of the first shaping stage AC outputs (left and right branches) as a function of the input frequency. We can extract the dominant pole to be at τ_{p2} = 5.3 ns and the differential gain of the stage A_2 = 0.957.

The differential gain of the stage is $A_2 = g_{mM12}/g_{msM14} = 0.964$. From Figure 7.37 we can extract $A_2 = 0.957$ which again is in good agreement with the analytical calculations.

The well of the load transistors is connected to a reference voltage $V_{ref6} = 400$ mV instead of the V_{ddWell}. The reason for this is that we needed enough voltage margin to keep the feedback transistor in saturation, even in case of a threshold voltage shift of M_{15}. Reducing the well potential (but still in a regime where no current flows through the well-to-substrate junction) decreases the DC voltage at the output of the first shaping stage to about 500 mV, so that the drain to source DC voltage across the feedback transistor is ~ 250 mV.

7.3.3 The feedback stage

The feedback stage is realised with a pair of n-channel MOS transistors in parallel, with the source connected to the output of the first shaping stage and the drain to the input of the CSA. The gate is connected to a dedicated bias circuitry which will be discussed in 7.3.5. Figure 7.38 shows the circuit diagram of the ALICE1LHCb feedback stage. The stage on the right is a dummy load which is used to balance the capacitive and resistive load seen by the two outputs of the shaper. M_{d1} and M_{d2} are identical to M_{f1} and M_{f2}, and M_{d3} is used to recreate a bias condition as similar as possible between the two transistor pairs.

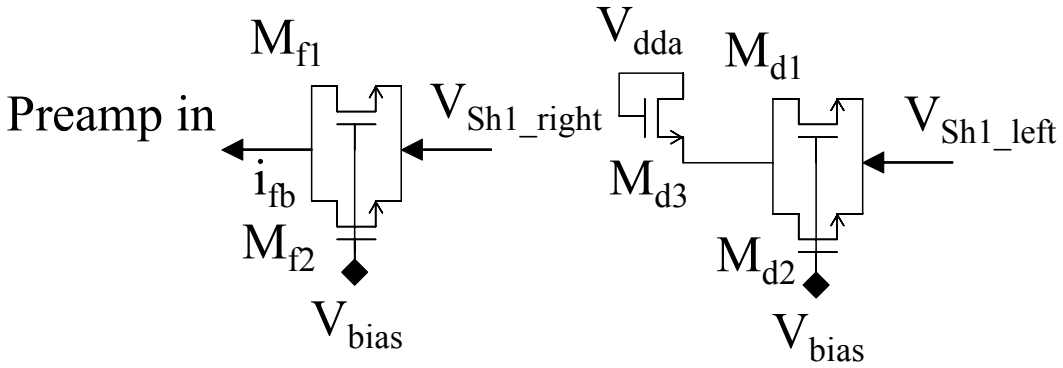

Figure 7.38: Circuit diagram of the ALICE1LHCb feedback stage. The left stage is the real feedback stage; the right stage is a dummy load to balance the shaper output load.

The transfer function of the feedback stage is:

$$i_{fb} = g_{mf}\, v_{sh1\_right} \tag{7.45}$$

where g_{mf} is the sum of the source transconductances $g_{mf} = g_{mf1} + g_{mf2}$ of the (identical) feedback transistors M_{f1} and M_{f2}, v_{sh1_right} is the small signal voltage at the right output of the first shaping stage, and i_{fb} is the feedback current injected into the preamplifier input node.

Together with the feedback capacitance, g_{mf} is the main parameter for positioning the feedback pole τ_{fb}. The feedback DC current is 18.2 nA in each feedback transistor, which results in a feedback total source transconductance $g_{mf} = 1.4$ μS. From chosen values $C_{fb} = 15$ fF and $A'_2 = A_2/2 = 0.482$ we can calculate $\tau_{fb} = Cfb/(g_{mf} A_2') = 22$ ns. A_2' is the

direct gain of the first shaping stage towards the feedback. In effect the first shaping stage has a differential output towards the second shaping stage, but a single-ended output towards the feedback.

The very low feedback current, which is needed to keep the shot noise contribution as small as possible, causes the feedback transistors to work in the weak inversion region. This can be a problem from a transistor matching point of view, as a change in the transistor parameters could cause a shift of the feedback pole and a modification of the output signal characteristics. To be sure that this would not have been a problem for the design, a sensitivity analysis of the circuit to the position of all the open loop poles was carried out on the final system (already presented in section 7.2.4.1).

7.3.4 The second shaping stage

The second shaping stage has the aim of realizing the third (open loop) pole τ_{p3}. Moreover, we can add more gain to the system and also set the global chip threshold in this stage. The basic structure is rather simple, as it resembles the first shaping stage and consists of a differential input pair with a p-channel low resistance load. Figure 7.39 shows the circuit diagram of this shaping stage.

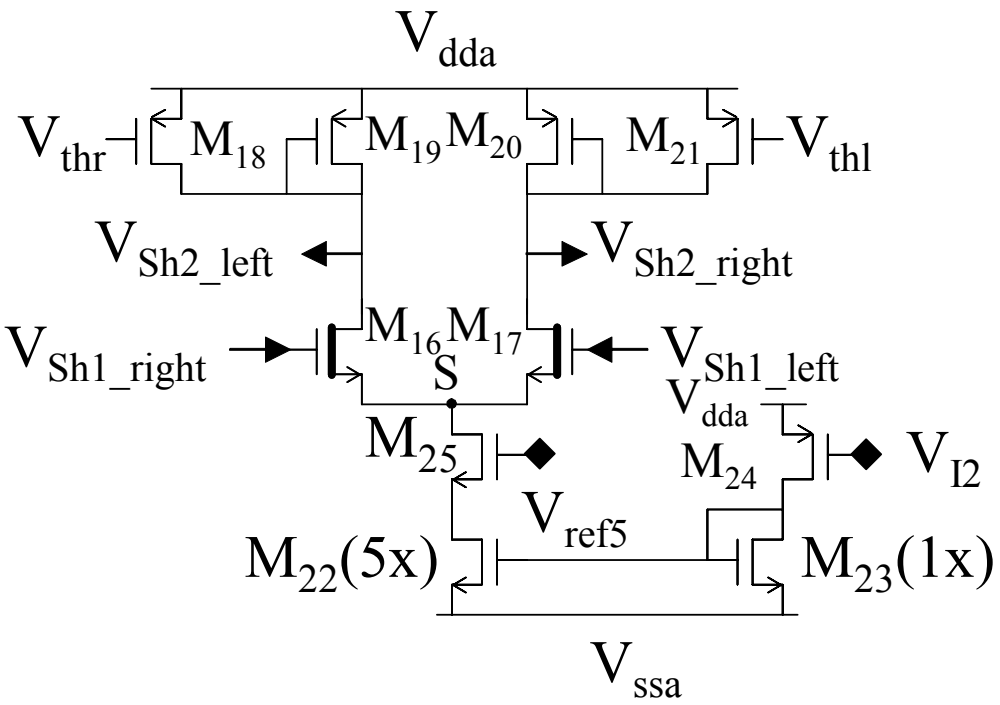

Figure 7.39: Circuit diagram of the ALICE1LHCb second shaping stage. M_{16} and M_{17} are Zero-Vt transistors.

The scheme is fully differential and profits from a special device provided by our technology, which is the Zero-Vt device. This is an n-channel transistor where no ion

implantation has been performed with the result that its threshold value is close to zero. The advantage is that the DC V_{gs} voltage drop of this type of transistor is reduced, if compared with an identical standard transistor identically biased, by the value of the n-channel threshold voltage V_{th} (~ 600 mV). This allows us to use an n-type input differential pair, which would not have been possible with a standard device at such a low input DC voltage (~ 500 mV).

The circuit bias current $I_2 = 1.5$ μA is generated by M_{24}, and is then multiplied by five in M_{22}, due to the ratio of the transistor sizes 5:1. The total power consumption of the second shaping stage is then 14.4 μm.

The need for local current mirroring is an example of the design penalties deriving from the use of enclosed devices. In effect, the constraints imposed by the ELT design result in a serious limitation for what concerns the precision of n-type current mirrors. In a current mirror it is common practice to try to keep transistors in strong inversion and with a high overdrive[19]. The reason is that the relative dispersion of the mirrored current is inversely proportional to the overdrive in strong inversion, while it is independent from the bias in weak inversion. It can be shown that a mirror in strong inversion is more precise than a mirror in weak inversion (for the same transistor areas). Edgeless transistors tend to be always in weak inversion when biased at our low current levels, thus providing a worse current matching. For this reason, if an n-current mirror cannot be avoided, the current is mirrored locally to the pixel (and this generates also an additional power consumption) to avoid big current mismatch due to long distance transistor mismatch.

Only 4/5 of the current is injected into the differential pair, because 1/5 is sunk by the feedback bias circuit from the node S, as will be explained in the next section.

The differential gain A_3 is the ratio of the input transistors transconductance $g_{m16\text{-}17} = 75$ μS to the transconductance of the output transistors $g_{m19\text{-}20} = 10.1$ μS, and is $A_3 = 7.43$. The input "effective" capacitance seen at the first shaping stage output can be calculated taking also into account the Miller effect which acts on the gate-to-drain capacitance of the input transistors, and is $C_{inSh1} = 79$ fF.

Similarly to the first shaping stage, the pole τ_{p2} is given by the resistance seen at the shaper output multiplied by the capacitance seen at the same node. The total capacitance (which comprises the capacitance of the input transistors of the leakage current compensation stage, described in 7.3.6, and the capacitance of the input transistors of the discriminator) seen at this node is $C_{p3} = 174$ fF, and the transconductance is $g_{p3} = 11.4$, so $\tau_{p3} = 15.2$ ns. Figure 7.40 shows the AC response of the second shaping stage (differential output, simulated alone). From it we can extract the DC gain $A_3 = 7.05$ and the pole $\tau_{p2} = 14.9$ ns, which are consistent with the analytically calculated values.

[19] The transistor overdrive is defined as $V_{gs} - V_{th}$.

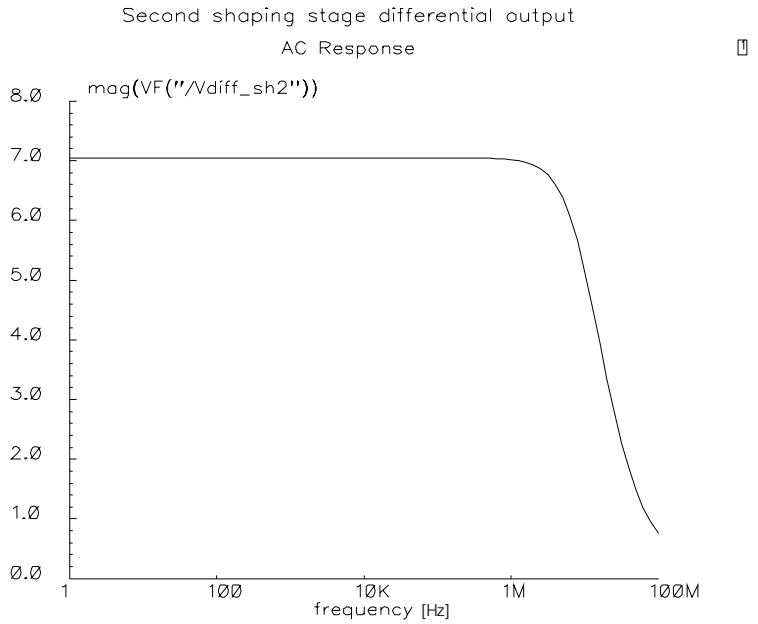

Figure 7.40: AC response of the second shaping stage (differential output); we can extract the DC gain $A_3 = 7.05$ and the pole $\tau_{p2} = 14.9$ ns.

The second shaping stage differential voltage output response to a differential input step of 10 mV is shown in Figure 7.41.

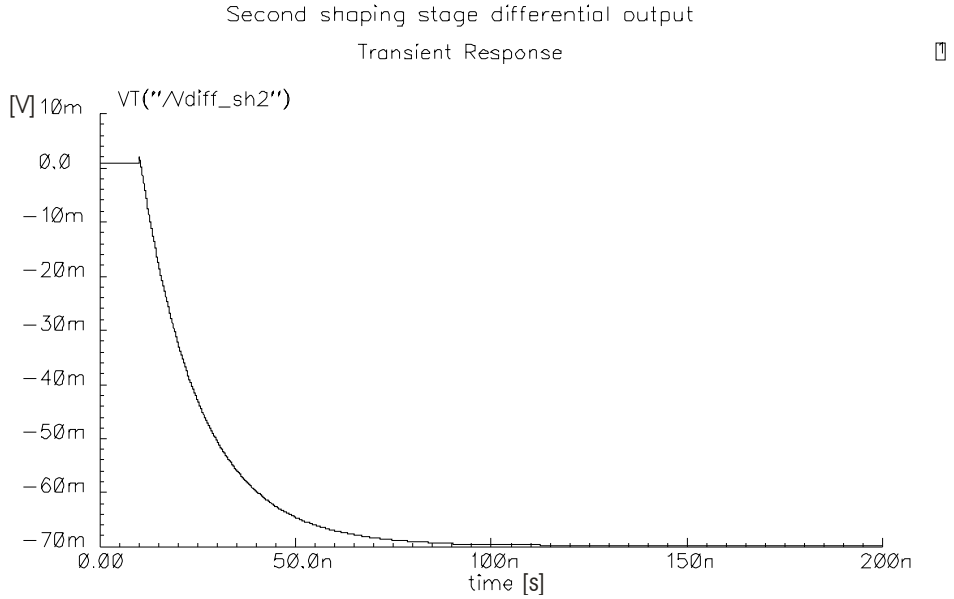

Figure 7.41: Transient response of the second shaping stage (differential output) to a differential input voltage step of 10 mV.

The transistor M_{22} is a large area transistor for matching reasons (five copies of the same transistor with L = 1.5 μm and W = 51.3 μm), which results in a high parasitic capacitance at the node S. The problem is reduced with the insertion of the cascode transistor M_{25}.

Transistor M_{18} and M_{21} have a two-fold aim, namely to increase the stage gain and to set the global threshold[20]. Two reference voltages, V_{thl} and V_{thr} have to be applied to the gate of the two transistors. The two biases are generated with a dedicated DAC which takes as a reference the voltage V_{I2}, and produces a differential imbalance on the two gates with respect to V_{I2}. The five transistors M_{18} to M_{21} and M_{24} are identical, so that when $V_{thl} = V_{thr} = V_{I2}$ the gate voltage of the four transistor M_{18} to M_{21} is the same and $V_{Sh2\_left} = V_{Sh2\_right} = V_{I2}$. In this condition each of them absorbs an equal current, i.e. ¼ of the bias current, so that the load transistors M_{19} and M_{20} are biased with half of the current which flows in the input transistors M_{16} and M_{17}, decreasing their transconductance and increasing the stage gain.

By means of the dedicated DAC its possible to generate a voltage imbalance on V_{thl} and V_{thr} so that $V_{thr} + V_{thl} = V_{I2}$ and $V_{thr} - V_{thl} = V_{th}$ where V_{th} is the global threshold voltage that we want to set (i.e. on each gate is applied a voltage $V_{thl\text{-}r} = V_{I2} \pm V_{th} / 2$). Supposing a linear behaviour of the four transistors[21], the increase of the M_{18} gate voltage to $V_{thr} = V_{I2} + V_{th} / 2$ will generate a current $I_2 + \Delta I_{Vth}$ in M_{18} and a current $I_2 - \Delta I_{Vth}$ in M_{19} (their sum has to be constant and equal to 2 I_2). Due to the fact that the two transistors are identical, this will result in the gate voltage of M_{19} being $V_{Sh2\_left} = V_{I2} - V_{th} / 2$. In the same way decreasing V_{thl} to $V_{I2} + V_{th} / 2$ results in the gate voltage of M_{20} being $V_{Sh2\_right} = V_{I2} - V_{th} / 2$. This means that a differential voltage imbalance V_{th} has been generated at the shaper output.

7.3.5 The feedback bias stage

Figure 7.42 shows the circuit diagram of the feedback bias stage.

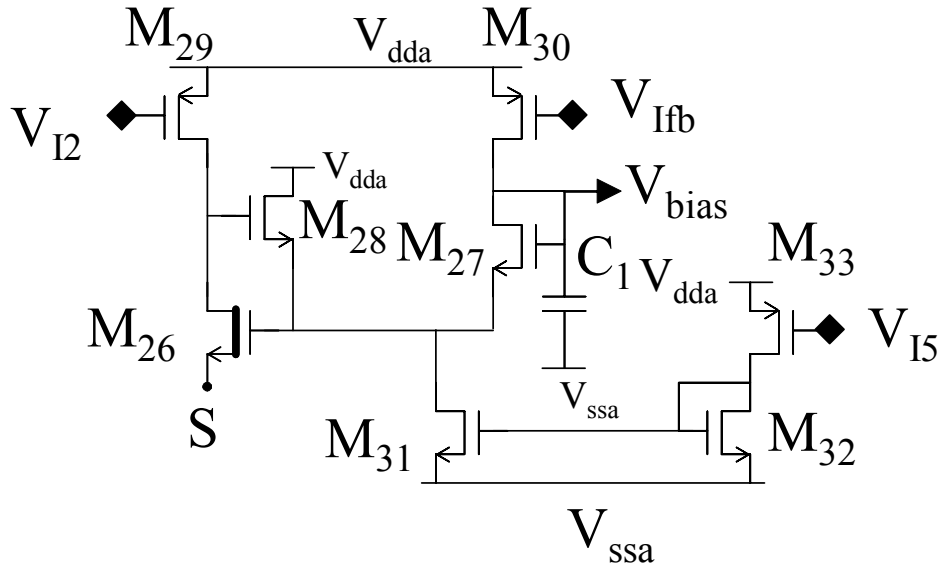

Figure 7.42: Circuit diagram of the feedback bias stage. The output of the circuit is the voltage V_{bias}. V_{I2} and S are in common with the second shaping stage (see also Figure 7.39).

[20] In a binary chip it is required the possibility to vary the global threshold of the chip, i.e. the minimum input charge that has to be injected into the preamp to have the discriminator firing.

[21] We are supposing also, for simplicity, that the shaper input voltages are equal: $V_{Sh1\_left} = V_{Sh1\_right}$.

The aim of the circuit[22] is to generate the bias voltage V_{bias} to correctly bias the feedback transistors. The Zero-Vt transistor M_{26}, which has the same length of M_{16} and M_{17} but half of their W, shares with M_{16} and M_{17} the same source (node S). It is biased by the current I_2, generated by the transistor M_{29} which is identical to M_{24}. M_{16} and M_{17}, which have a double W, are biased by $2 \cdot I_2$, so they have the same gate voltage of M_{26}, which has the source connected with the gate of M_{16}. This means that the transistor M_{27}, which is a copy of the feedback transistors, has the same DC source voltage of the feedback transistors. By means of the transistor M_{30} and its reference voltage V_{Ifb}, which can be set by a dedicated on-chip DAC, we can force a current I_{D27} (the design value is about 20 nA) through M_{27}, which generates the bias voltage V_{bias}. M_{27} and the feedback transistors (also the dummy copies of the feedback transistors) have the gates shorted together and the same V_S, this means that they have the same V_{GS} and the current I_{D27} is copied by M_{f1} and M_{f2} (and also by M_{d1} and M_{d2}).

The advantage of this scheme is that if there is a parameter variation of the Zero-Vt transistor or of the standard n-channel transistor, this does not have an impact on the feedback current, provided that all the transistors experience the same change. Layout consideration on how to minimise transistor mismatch will be presented in section 7.6.

The transistor M_{28} provides local negative feedback to M_{26} to hold its gate voltage V_{G26}. If, for example, the current in M_{26} increases (and M_{26} gate voltage follows), the inverter-like structure M_{26}-M_{29} decreases M_{28} gate voltage, which acts as a follower, trying to reduce V_{G26}. Transistors M_{31}, M_{32} and M_{33} provide the bias current of 500 nA (with local n-mirroring) for M_{28}. The power consumption of this stage is 1.6 μW. The capacitance C_1 (of about 320 fF) acts as a low pass filter, filtering the disturbances that could perturb V_{bias}.

7.3.6 The leakage current and offset compensation scheme

No DC path is foreseen in our CSA to allow the flow of the detector leakage current. Without a scheme able to absorb this leakage current, it would be absorbed by the feedback transistor and would modify the feedback current and thus the position of the CSA poles.

The circuit diagram of the leakage current compensation scheme is presented in Figure 7.43. The core of the scheme is a very low bandwidth differential amplifier, with a double differential input stage. The bandwidth is kept low using relatively big input and load transistors (M_{34} to M_{37} are 20 μm^2, M_{38} and M_{39} are 40 μm^2) with very little bias current (5 nA for each transistor), and an additional capacitance C_2.

The reason for the double differential stage input is that we have to set an imbalance at the second shaping stage differential output, which is needed for the chip global threshold setting, but we do not want the leakage current compensation scheme to correct for this

[22] Refer also to Figure 7.38 and Figure 7.39.

imbalance. The outputs of the second shaping stage V_{Sh2_right} and V_{Sh2_left} are presented at one of the inputs of the two differential stages (i.e. the gates of M_{37} and M_{34}), and are compared with V_{thl} and V_{thr}, which are fed to the other two inputs of the differential stage (i.e. the gates of M_{35} and M_{36}). This means that a difference is generated at the differential stage output only if V_{Sh2_right} or V_{Sh2_left} are different, respectively, from V_{thr} and V_{thl}. In this case the differential stage output is sensed by the transistor M_{44} which generates a feedback current at the preamplifier input.

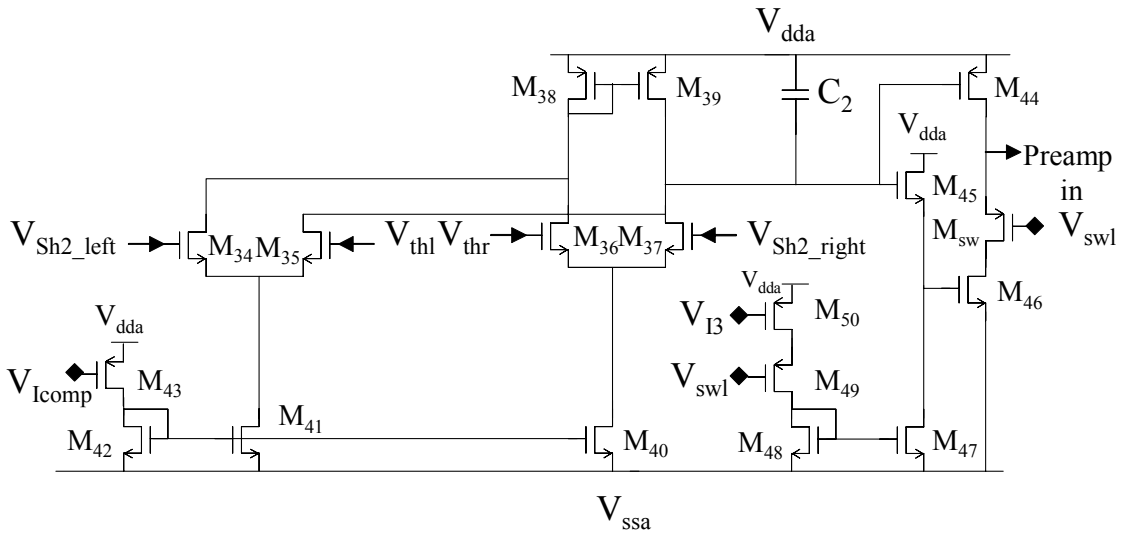

Figure 7.43: Circuit diagram of the leakage current compensation scheme.

A leakage current entering the amplifier input node is equivalent to a very low frequency signal, so it generates an imbalance at second shaping stage output. This difference is sensed by the leakage current compensation stage, which generates a feedback current into the amplifier input which cancels the leakage current.

The system is such that if $V_{Sh2\_right} = V_{thr}$ and $V_{Sh2\_left} = V_{thl}$, and a leakage current I_{leak} is injected into the CSA input node, M_{44} delivers a current equal to $I_{feed} - I_{leak}$ (see also Figure 7.31).

Figure 7.44 shows the second shaping stage differential output (simulated in closed loop) for a leakage current of 0 and 30 nA ($V_{th} = 50$mV). The high frequency behaviour of the circuit is practically unchanged. The total power consumption of this stage is negligible.

In normal operation mode the voltage V_{swl} is set at 1.6 V, so M_{49} and M_{sw} switch off all the transistors from M_{45} to M_{50} (and M_{sw}). This reduces the current consumption but most of all eliminates an additional possible noise path towards the preamplifier input while running in normal operation mode. If I_{feed} exceeds I_{leak}, M_{44} switches off and the system cannot absorb the excess leakage current. In this case M_{45} and M_{46} (and their corresponding biasing transistors M_{47} to M_{50}) can be switched on to absorb $I_{leak} - I_{feed}$.

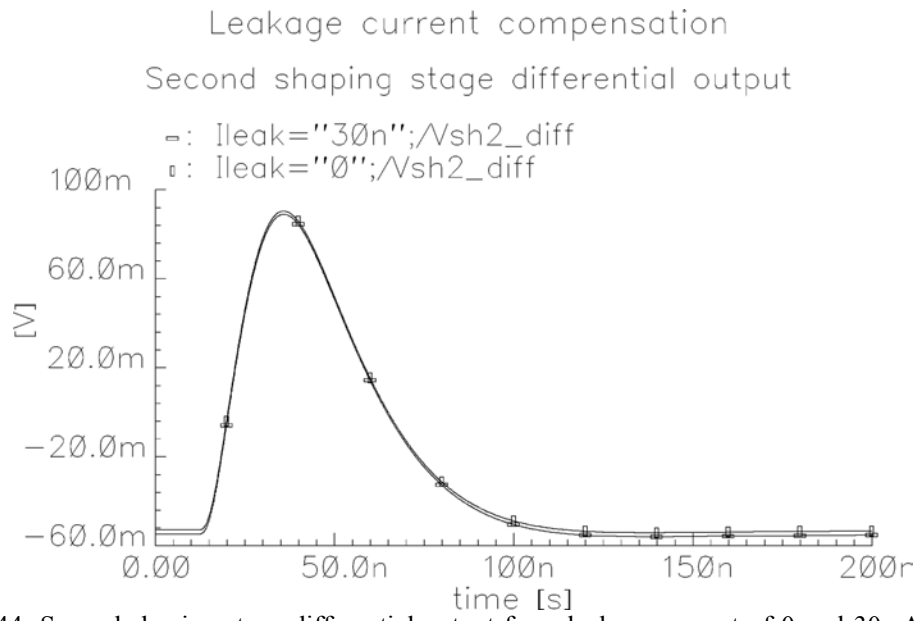

Figure 7.44: Second shaping stage differential output for a leakage current of 0 and 30 nA: the high frequency behaviour of the circuit is practically unchanged.

Any mismatch in the parameters of the transistors of the front-end can generate an offset (imbalance) at the front-end output. However, the leakage current compensation scheme corrects for any imbalance present at the shaper output, regardless of its origin. This means that with this scheme we can also correct for the offset present on V_{Sh2_right} and V_{Sh2_left}. For this reason, the offset presented at the discriminator input will be related only to the mismatch of the transistors of the leakage current compensation scheme.

7.3.7 Full front-end simulations and analysis

After the analysis of each single block considered in open loop, we will now go through the simulation analysis results of the full front-end chain.

7.3.7.1 Transient response

Figure 7.45 plots simulation results showing the preamplifier output voltage (left) and the first shaping stage differential output voltage (rigt) for an input charge of $Q_1 = 5000\ e^-$ and $Q_2 = 16000\ e^-$ (the global threshold bias voltage is set at $V_{th} = 50$ mV, which is equivalent to about 2300 e^- at the preamplifier input).

The CSA peaking time is 7.7 ns for Q_1 and 8.6 ns for Q_2. The return to zero time is 106 ns for Q_1 and 103 ns for Q_2. This means that the amplifier is ready to process another input signal after less than 110 ns. The first shaping stage peaking time is 13.3 ns for Q_1 and 14.6 for Q_2; the return to zero time is 113 ns for Q_1 and 120 ns for Q_2. The CSA gain is about 10 μV / e^- for Q_1 and 9.6 μV / e^- for Q_2, while the first shaping stage gain (differential output) is 7 μV / e^- for Q_1 and 5 μV / e^- for Q_2.

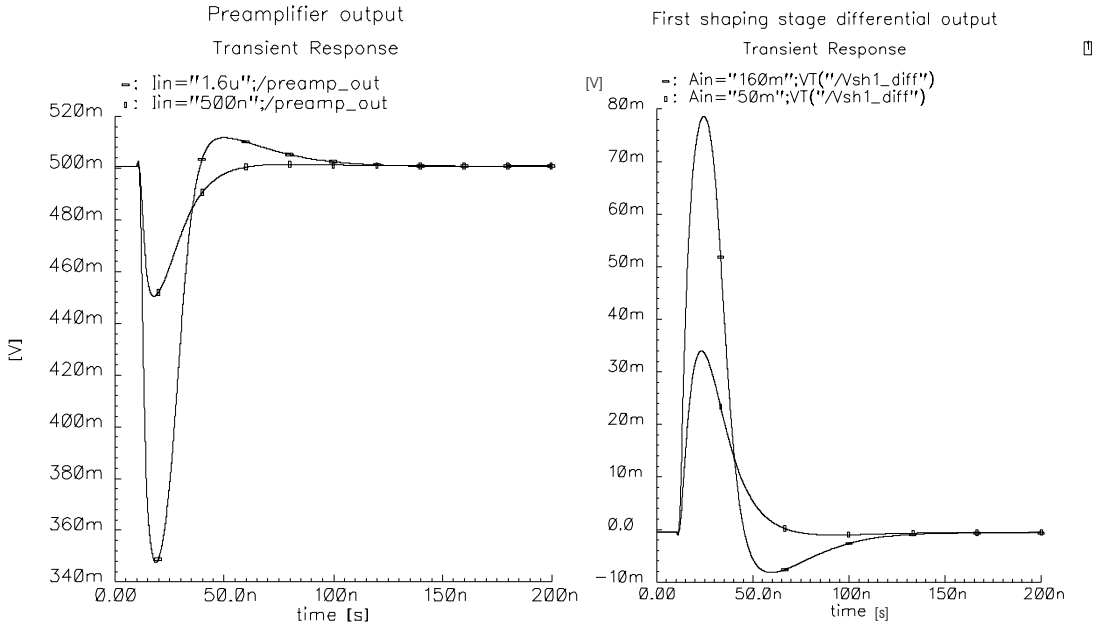

Figure 7.45: Simulation results showing the preamplifier output voltage (left) and the first shaping stage differential output voltage (right) for an input charge of 5000 e^- and 16000 e^- (V_{th} = 50 mV, ~2300 e^-).

Figure 7.46 plots simulation results showing the second shaping stage differential output voltage for an input charge of Q_1 = 5000 e^- and Q_2 = 16000 e^- (V_{th} = 50 mV, ~2300 e^-).

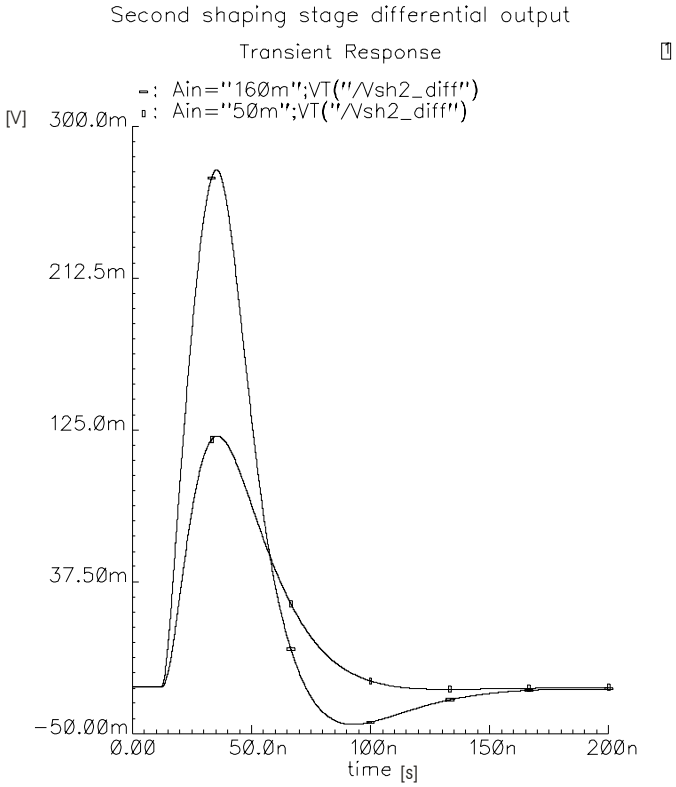

Figure 7.46: Simulation results showing the second shaping stage differential output voltage for an input charge of 5000 e^- and 16000 e^- (V_{th} = 50 mV, ~2300 e^-).

The second shaping stage peaking time is 25.6 ns for Q_1 and 25.2 ns for Q_2. The return to zero time is 106 ns for Q_1 and 158 ns for Q_2. The big increase in the return to zero time in this case is due to the undershoot, whose value is higher than 1% (the value at which we have defined the return to zero). This means that for the typical LHCb signal the return to zero time of the full front-end is less than 110 ns. The gains are reported in the next section. Some of the simulated values do not match exactly what was found with the analytical calculations. This is due to the system nonlinearity; simulations done at very low input signals confirm this hypothesis. In particular, the preamplifier return to zero time is increased with respect to calculations due to the undershoot, whose value is again higher than 1%.

7.3.7.2 Peaking time and gain variations

The second shaping stage differential output voltage as a function of the input charge is shown in Figure 7.47 (left, 1 mV = 100 e^-). The variation of the peaking time at the differential output of the second shaping stage with the input charge is shown in Figure 7.47 (right, 1 mV = 100 e^-).

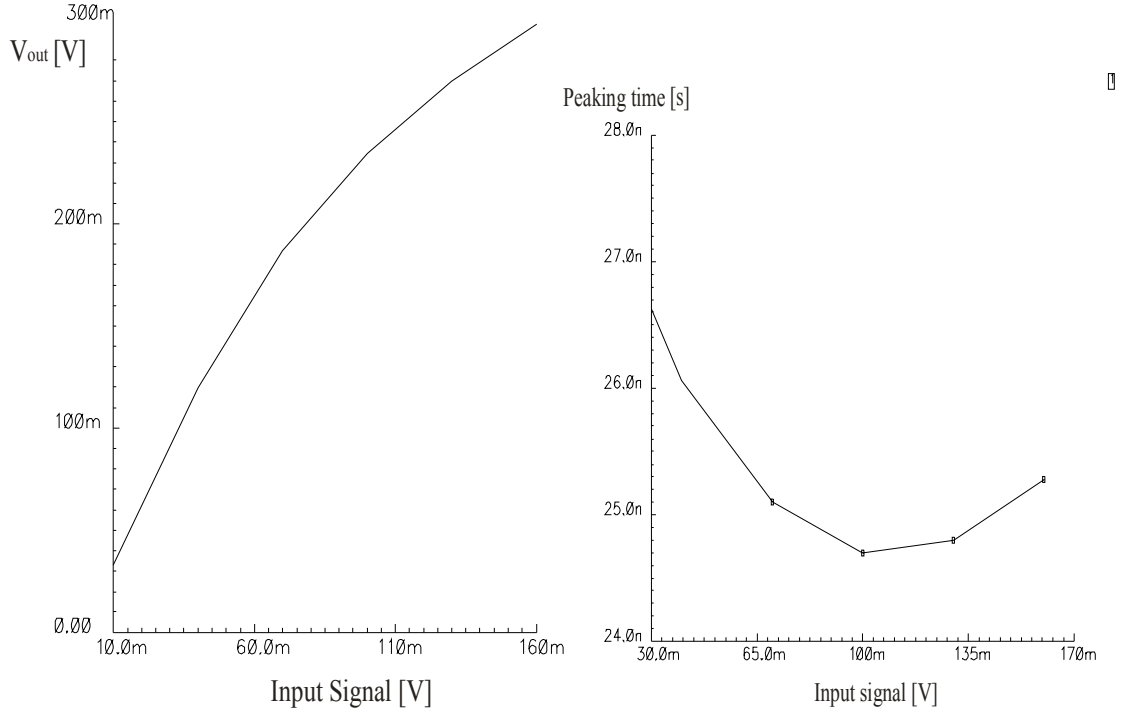

Figure 7.47: Second shaping stage differential output voltage as a function of the input charge (left, 1 mV = 100 e^-). Variation of the second shaping stage peaking time (differential output) with the input charge (right, 1 mV = 100 e^-). The gain at Q_{in} = 5000 e^- is 28.7 μV/e^- while at Q_{in} = 17000 e^- it is 17.7 μV/e^-.

7.3.7.3 AC response

Figure 7.48 (left) plots the current to voltage AC response of the second shaping stage differential output. The band of the system goes from about 10 KHz to about 6 MHz. Figure 7.48 (right) plots the Discrete Fourier Transform of the signal at the same output for an input charge of Qin = 5000 e^-. The major signal content is up to about 20 MHz.

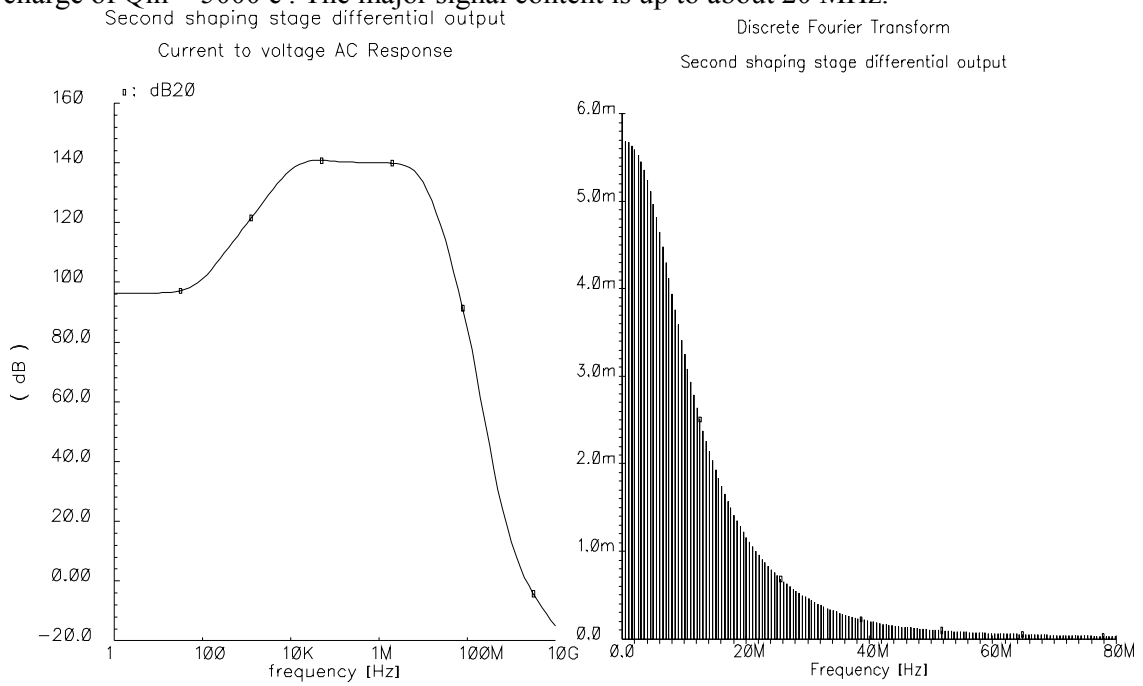

Figure 7.48: Current to voltage AC response of the second shaping stage differential output (left) and Discrete Fourier Transform of the signal at the same output for an input charge of Qin = 5000 e^- (right).

7.3.7.4 Crosstalk

Due to the large pixel-to-pixel capacitance a very important issue to address is *crosstalk*. If a sensor cell receives a big input charge the neighbouring cells can see an input signal due to the capacitive coupling between pixels which could in principle be enough to be detected as a hit. We performed some simulations where three cells were capacitively coupled via a 50 fF capacitor, and had a capacitance to ground of 25 fF. Only the central cell was pulsed with increasing input charge, until a hit was detected on the neighbouring pixels.

Figure 7.49 (left) shows the second shaping stage differential output of a neighbouring cell for an input charge of the pulsed cell of 55000 e^- (V_{th} = 50 mV, ~2300 e^-), which is the minimum charge which triggers the discriminator. The discriminator output of the same pixel is shown in Figure 7.49 (right). The front-end design makes the system rather insensitive to crosstalk. Crosstalk depends on the ratio of the preamplifier rise time versus shaping time. In effect, the faster the preamplifier, the sooner it is able to re-establish virtual ground at its input after the hit. The signal input seen at the neighbouring channels is a fast bipolar voltage

transient, which is reduced in amplitude by the shaping action of the front-end. This means that the slower the shaper, the more the fast signal produced by crosstalk is shaped (reduced). Moreover, it produces at the discriminator input a signal which has a peaking time much faster than the peaking time produced by a real input signal (15 ns instead of 25 ns), and this reduces the chance that the discriminator can fire on such a fast signal. The peaking time was fixed by the circuit specifications, and the preamplifier was designed as fast as possible with the power budget given.

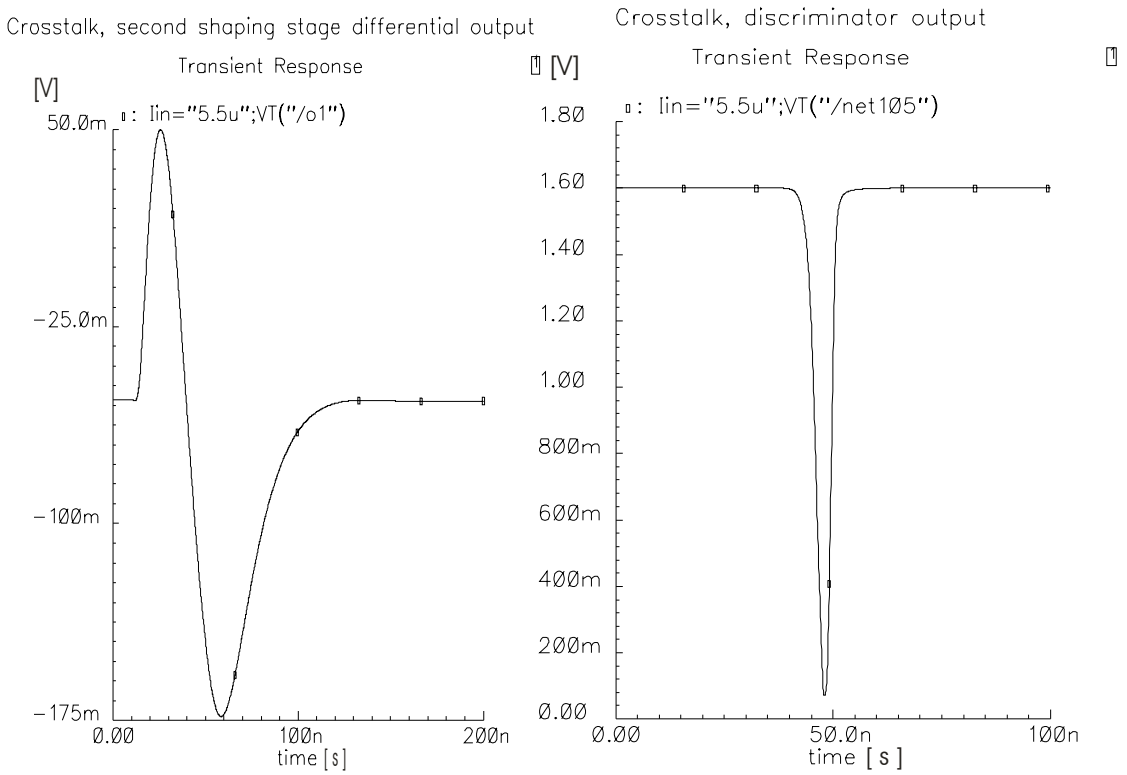

Figure 7.49: Crosstalk simulations, second shaping stage differential output of a neighbouring cell for an input charge of the pulsed cell of 55000 e^- (left), which is the minimum charge which triggers the discriminator. The discriminator output of the same pixel is shown on the right (V_{th}=50 mV, ~2300 e^-).

7.3.7.5 Response to a huge input signal

Although the system does not need to be linear up to high values of the input charge, an important feature that it needs to have is to be able to recover in a reasonable time after that a big charge signal has hit the sensor. Figure 7.50 (left) shows the second shaping stage differential output for an input charge of 50000 e^- which is the highest charge foreseen in LHCb (V_{th} = 50 mV, ~2300 e^-). After 150 ns the discriminator is back to 1.6 V (Figure 7.50 right) and after ~200 ns the system is ready to process another pulse, even if the slow signal tail offsets the pixel threshold. At 200 ns the residual signal is less than 3% of the signal

generated by a 5000 e^- input charge. The behaviour for 1000000 e^- is very similar[23]; after 600 ns the discriminator is back to 1.6 V and the system is ready to process another pulse; after 1.5 μs the remaining signal is about 10 % of the signal generated by a 5000 e^- input charge.

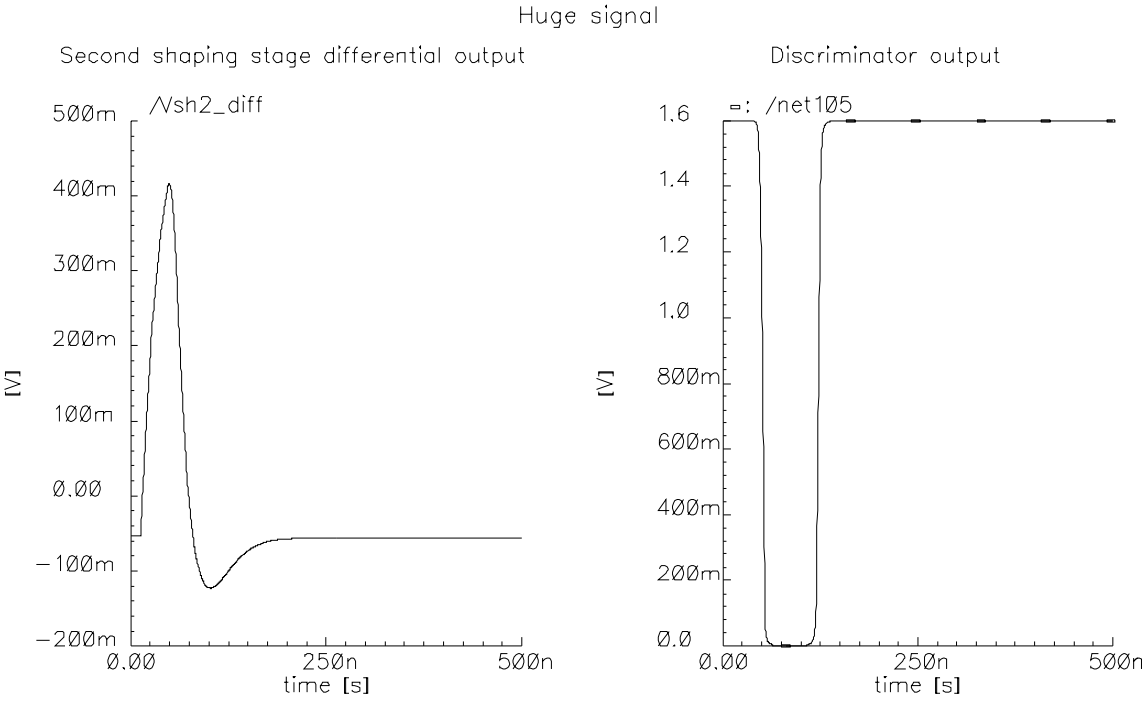

Figure 7.50: Second shaping stage differential output for an input charge of 50000 e^- (left) and corresponding discriminator output (right).

7.3.7.6 Phase margin

To check for possible stability problems of the front-end, we simulated the phase margin of the circuit both for the high frequency loop (preamplifier, first shaping stage and high frequency feedback) and for the low frequency loop (the full system with the low frequency leakage current compensation block).

This was possible thanks to the design of a special circuit block, a "virtual loop cutter", which allows to cut the loop in a certain position of the circuit chain only while performing the AC analysis.

In this way the simulator can properly calculate the closed-loop DC operating point, and use this operating point to evaluate the open-loop response of the system. Figure 7.51 shows the results; the phase margin for the high frequency loop is 72° and for the low frequency loop it is 82°.

[23] Such a huge signal can be deposited only by a very high energy particle crossing the sensor; this event is extremely unlikely to happen at pixel level.

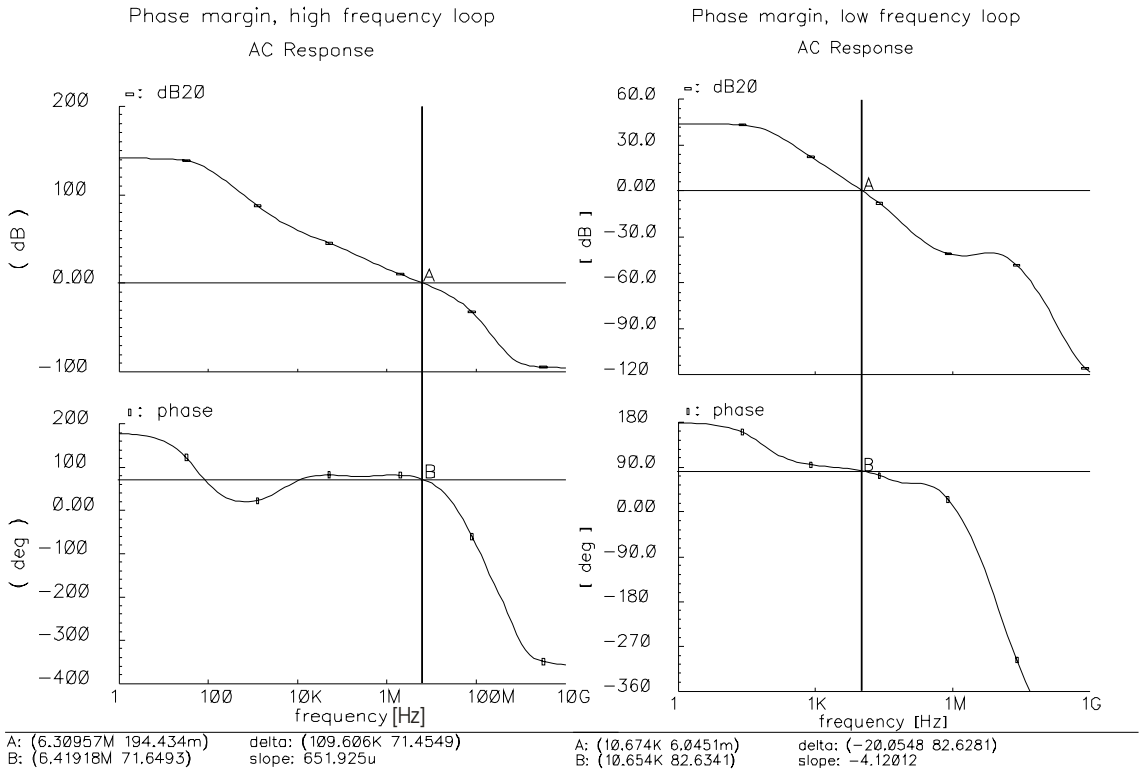

Figure 7.51: Phase margin analysis; the phase margin for the high frequency loop is 72° (left), the phase margin for the low frequency loop is 82° (right).

7.3.7.7 Power supply noise coupling

Noise on the power supply lines can couple into the circuit and give disturbance at the output. Given the extremely high gain of the preamplifier, it is good practice to check that noise injected on the power supply lines is not amplified by a very large factor.

Figure 7.52 and Figure 7.53 show simulations checking the AC second shaping differential output with an AC voltage source of 1V on V_{dda}, V_{wella}, V_{ssa} and V_{sub}. It can be seen that for V_{dda} and V_{wella} the gain factor is less than three (i.e. one millivolt noise on these power supplies generates three millivolts at the output), while for V_{ssa} is less than 1.5 in the entire frequency domain. For what concerns V_{sub} the gain factor is about eight. This is due to the noise coupling through the feedback transistors, and cannot be further reduced given the circuit topology. For this reason extreme care has been put in trying to keep the substrate noise as low as possible. If using the amplifier gain at $Q_{in} = 5000\ e^-$, 3 mV at the second shaping stage differential output correspond to about 104 e^-, while 8 mV correspond to 279 e^-.

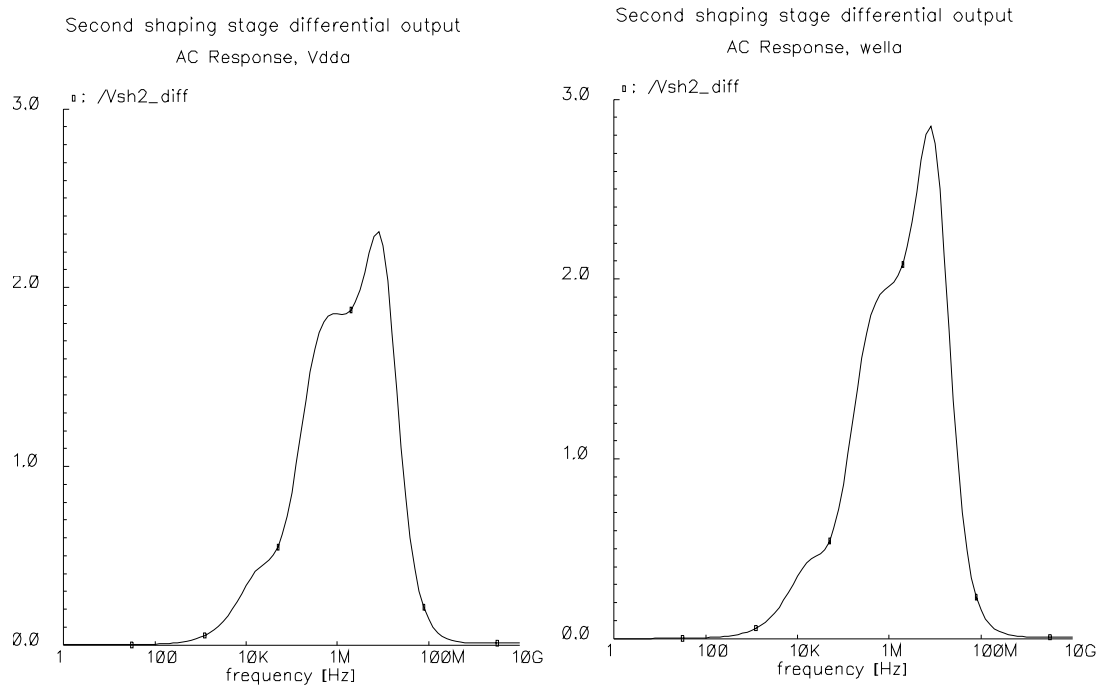

Figure 7.52: Power supply noise coupling: simulations showing the AC second shaping differential output with an AC voltage source of 1V on V_{dda} (left) and V_{wella} (right).

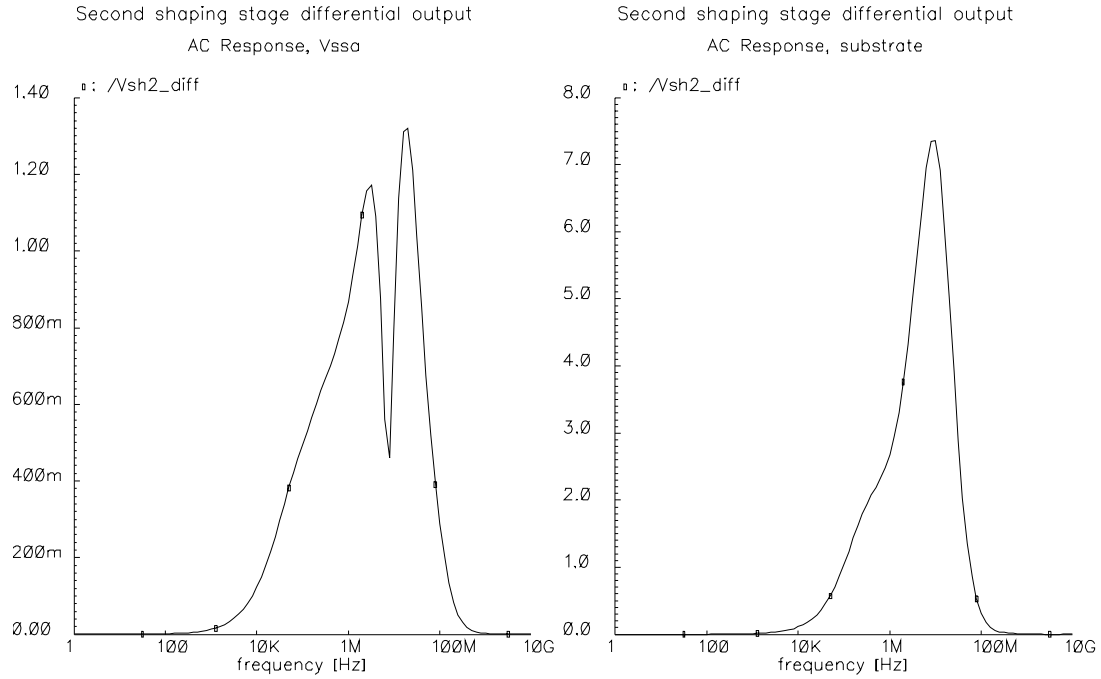

Figure 7.53: Power supply noise coupling: simulations showing the AC second shaping differential output with an AC voltage source of 1V on V_{ssa} (left) and V_{sub} (right).

7.3.7.8 Temperature sensitivity

A temperature sensitivity analysis is shown in Figure 7.54. The front-end is simulated with temperatures ranging from 0 to 100 °C (the input charge is 5000 e^- and V_{th} = 50 mV). No

major variations of the peaking time and of the gain are present (Figure 7.54, left). The discriminator output falling edge (the following on chip logic is sensitive to this edge only) is almost insensitive to temperature variations (Figure 7.54, right). For ALICE the temperature will be controlled to be constant within one °C, For LHCb the temperature will be about 30 °C at start up, then increase up to about 50 °C and then stay constant. For both the experiments, the degree of temperature sensitivity shown in Figure 7.54 will not be a problem.

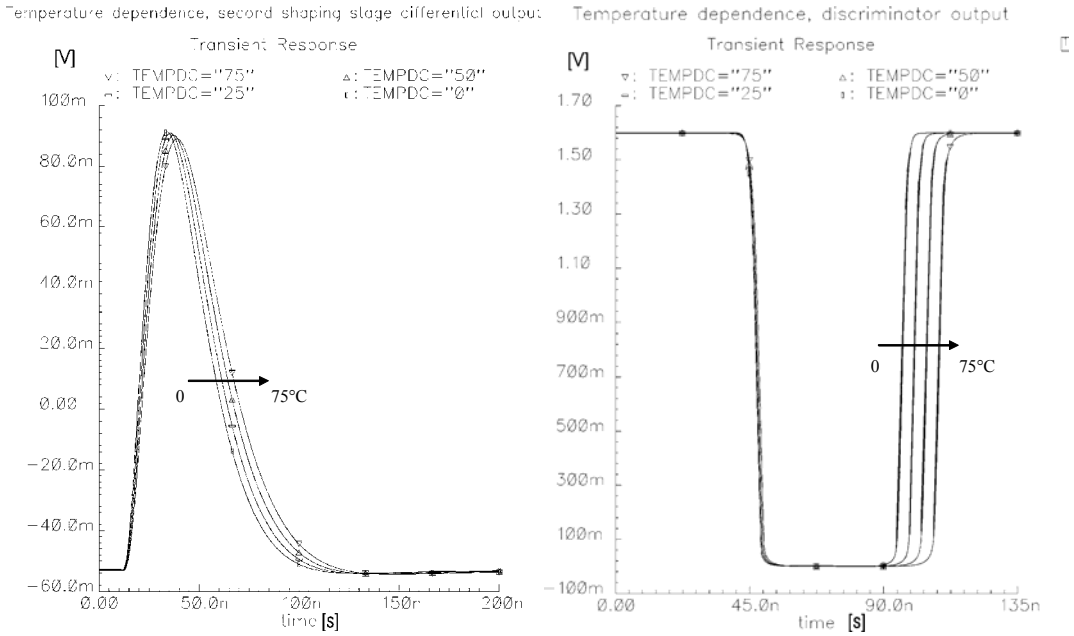

Figure 7.54: Sensitivity of the second stage differential output voltage to temperatures variations ranging from 0 to 100 °C (left) and its corresponding discriminator output (the input charge is 5000 e^- and V_{th} = 50 mV).

7.3.8 The discriminator

The shaper differential output feeds into a discriminator, which transforms the analogue pulse into a binary signal. A block diagram of the structure of the discriminator is shown in Figure 7.55. The discriminator input stage is an Operational Transconductance Amplifier (OTA), which makes a voltage-to-current conversion. The discrimination is then performed in current mode by the subsequent non-latching current discriminator. The outgoing voltage pulse is then squared and adapted to the correct digital voltage levels. A NAND gate is used to mask the pixel in case of malfunctioning or excessive noise.

To increase pixel-to-pixel uniformity and decrease threshold dispersion, a 3-bit register and a corresponding digital-to-analogue converter (DAC) are used in every pixel to finely adjust the pixel threshold. The total power consumption of the discriminator is ~7 mW.

The discriminators in the pixel matrix provide also a fast-OR signal, which is foreseen for diagnostic purposes during testing or for self-triggering[24].

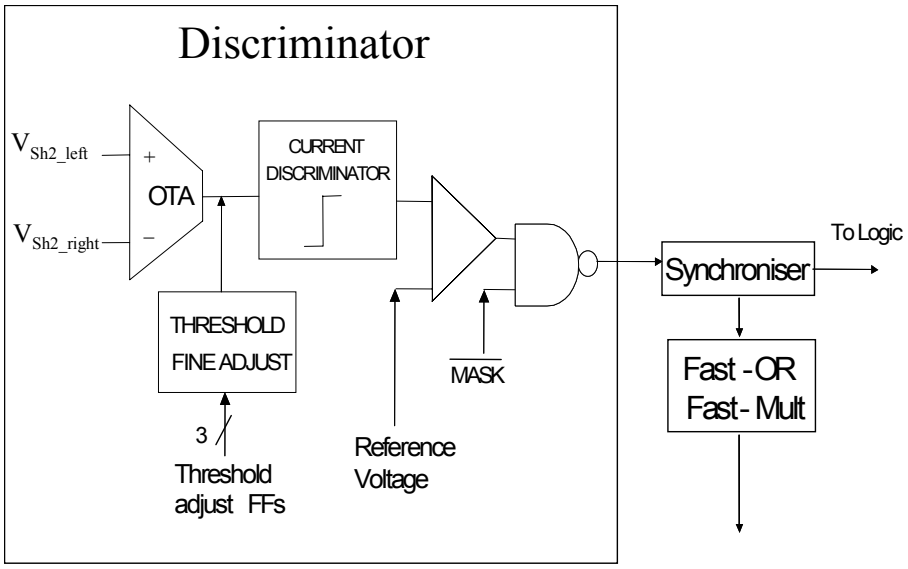

Figure 7.55: Schematic block diagram of the discriminator.

7.3.8.1 The Operational Transconductance Amplifier (OTA)

The circuit diagram of the OTA is shown in Figure 7.56.

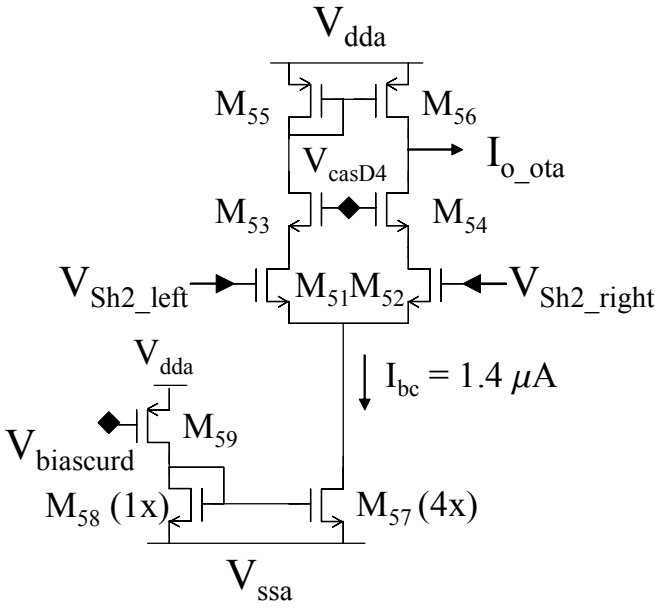

Figure 7.56: Circuit diagram of the discriminator input block, the OTA.

[24] During the experiment only the "interesting" events have to be read out, and they are selected with the external trigger. While testing the chip all the events are interesting and have to be read out. The fast-OR can generate a trigger signal in coincidence with any hit on the chip.

Transistors from M_{57} to M_{59} provide the bias current (of 1.4 μA) with local mirroring (with a multiplication factor of four for power consumption reasons). This stage converts the differential input signal V_{Sh2_left} - V_{Sh2_right} into a current, which is then fed to the following stage, the current discriminator.

The choice of the OTA input transistors is critical for the full front-end performance. The minimum chip threshold is influenced by several contributions, namely electronic noise, digital-to-analogue crosstalk and threshold dispersion. The electronic noise is minimised in the analogue section of the front-end, and the digital-to-analogue crosstalk with special layout techniques and power distribution schemes. For the threshold dispersion the problem can be targeted mainly at the discriminator level. The mismatch of the OTA input transistors has to be added in squares to the mismatch which is generated by the analogue front-end section, and it should not be the dominant component to the total threshold dispersion. This means that the input transistors should be large enough to reduce their mismatch to an acceptable level (i.e. comparable to the threshold dispersion due to the front-end). On the other hand an excessive size could degrade the OTA performance due to the additional capacitance. For this reason, the input transistors M_{51} and M_{52} are rather large[25] (W/L = 25 μm / 1 μm), but to increase the OTA bandwidth we inserted the two cascode transistors M_{53} and M_{54}.

The p-load transistors M_{55} and M_{56} can also contribute to the threshold mismatch. The total mismatch can be calculated with the following formula:

$$\sigma_{Tot}^2 = \sigma_N^2 + \sigma_P^2 \frac{g_{m,p}}{g_{m,n}} \tag{7.46}$$

where σ_N^2 is the mismatch of the input differential pair, σ_P^2 is the mismatch of the p-load transistors (both of them can be calculated with equations 5.10 or 5.13), $g_{m,p}$ is the transconductance of the p-channel transistor pair M_{55}-M_{56} and $g_{m,n}$ is the transconductance of the p-channel transistor pair M_{51}-M_{52}. For the discriminator OTA σ_{Tot}^2 = 1.96 mV^2. We can calculate the total threshold dispersion adding (in squares) this contribution to the contribution coming from the front-end.

We have seen that the leakage current compensation scheme corrects for the offset present on V_{Sh2_right} and V_{Sh2_left}. For this reason, the offset presented at the discriminator input will be related only to the mismatch of the transistors of the low frequency feedback. Repeating the calculations already done for the discriminator, we can calculate σ_{Tot}^2 = 2.2 mV^2. This results in a σ_{chip} = 2 mV which corresponds to an equivalent charge at the input of the preamplifier of about 70 e$^-$ (using the gain at Q_{in} = 5000e$^-$). A higher value should be expected in reality, so an additional block for the reduction of this threshold spread has been added, and will be described in 7.3.8.3.

[25] The load they represent was taken into account in the calculations of the front-end poles.

7.3.8.2 The current discriminator

Once the second shaping stage differential voltage output has been converted in a current by the OTA, the current is sensed by the current discriminator depicted in Figure 7.57. Due to the global threshold setting which creates an imbalance at the second shaping stage differential voltage output, the OTA in "idle" mode (i.e. waiting to process an event) is imbalanced, so that a DC current has to be absorbed by the OTA. This current is provided by the transistor M_{60}, while M_{61} is off.

Transistors M_{62} and M_{64} form the inverting stage depicted Figure 7.57 (right), and are biased with 1.2 μA (provided by the local mirror M_{66}-M_{65}). The overall negative feedback regulates the gate voltage of M_{60} such that it provides exactly the current absorbed by the OTA and $v_o = - i_{in} / g_{mN}$.

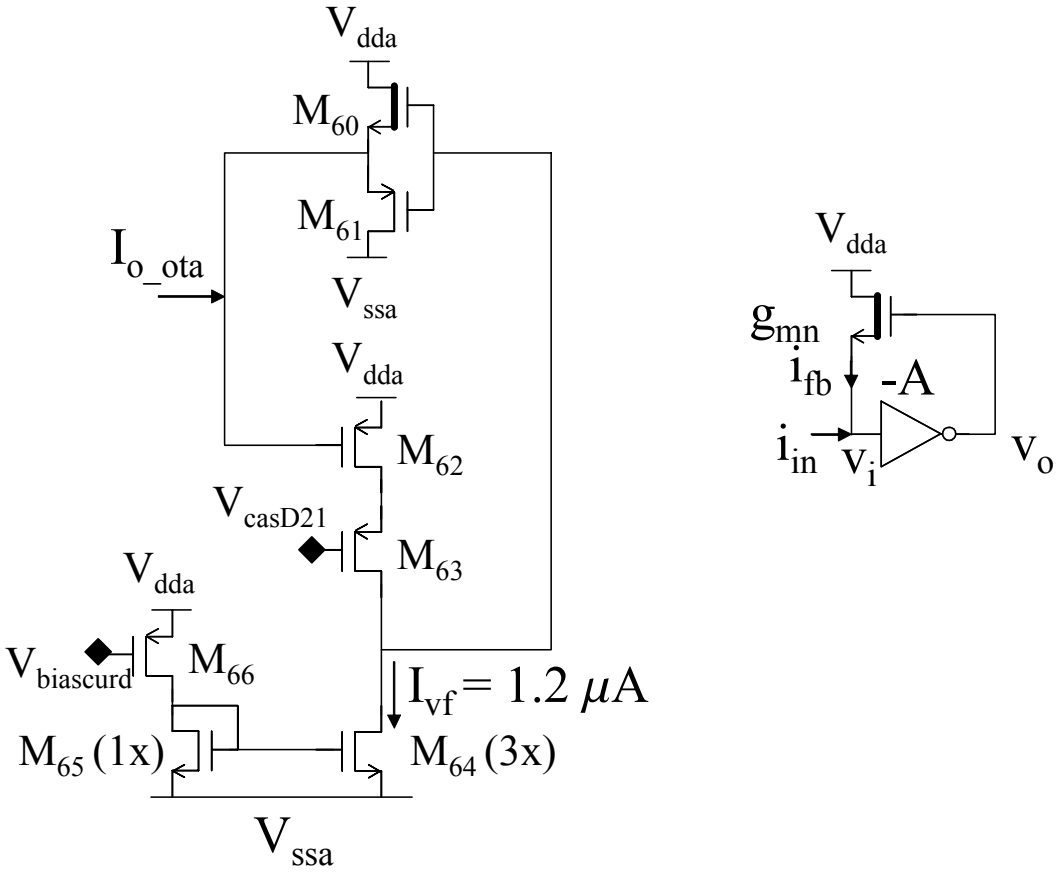

Figure 7.57: Circuit diagram of the current discriminator (left) and simplified block diagram of its working principle (right) when M_{60} is ON.

A low input charge creates a signal at the discriminator input which reduces the current absorbed by the OTA. This means that the g_{mN} of M_{60} decreases and the stage gain increases. When the input current is almost zero, the system is in open loop and has a voltage gain -A, so that the current discriminator output decreases very rapidly until it reaches its "low" state.

If the signal at the preamplifier input is above threshold, the OTA starts to source current toward the current discriminator, the transistor M_{61} switches on and a feedback path is restored, and the system works in the same way as before (but with M_{61} acting as feedback transistor). When the signal produced by the shaper goes back to zero the input current reduces and eventually changes sign, switching again M_{61} off and M_{60} on. During the current zero crossing the current discriminator output switches back to its "high" state. Figure 7.58 shows the current discriminator output for an input signal of 5000 e^- and a threshold of about 2300 e^-.

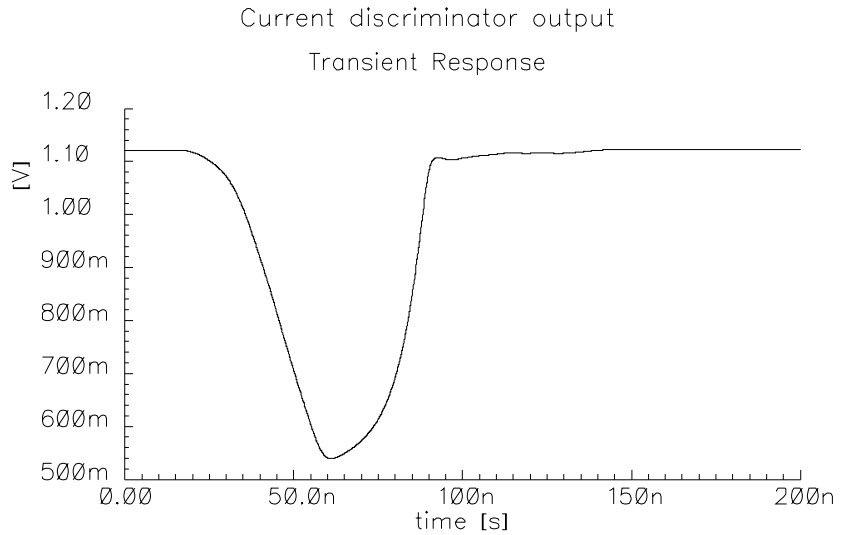

Figure 7.58: Current discriminator output for an input signal of 5000 e^- and a threshold of about 2300 e^- (V_{th} = 50mV).

7.3.8.3 The threshold fine adjust block

A block (shown in Figure 7.59) was added at the OTA output to further reduce the threshold dispersion. A mismatch of the voltage threshold at the OTA input stage is translated in a current mismatch at its output.

A base current I_b, whose value can be set by means of a dedicated DAC, is generated by M_{78}; M_{79} generates a current which is twice as much and M_{80} a current that is four times the base current. These currents are injected into the OTA output node if, respectively, b_0, b_1 or b_2 are set to a logic 1 (i.e. 1.6V; note the inverted bits in Figure 7.59, left). Three on-pixel registers store the values of these bits.

It is possible to measure the threshold of each pixel and the chip average threshold with the s-curve method described in section 6.3.1. With appropriate algorithms the fine threshold adjust current can be set for each pixel to make the pixel threshold approach the average threshold, decreasing the effective threshold spread. The maximum threshold spread which can be corrected for is about 900 e^- in our case.

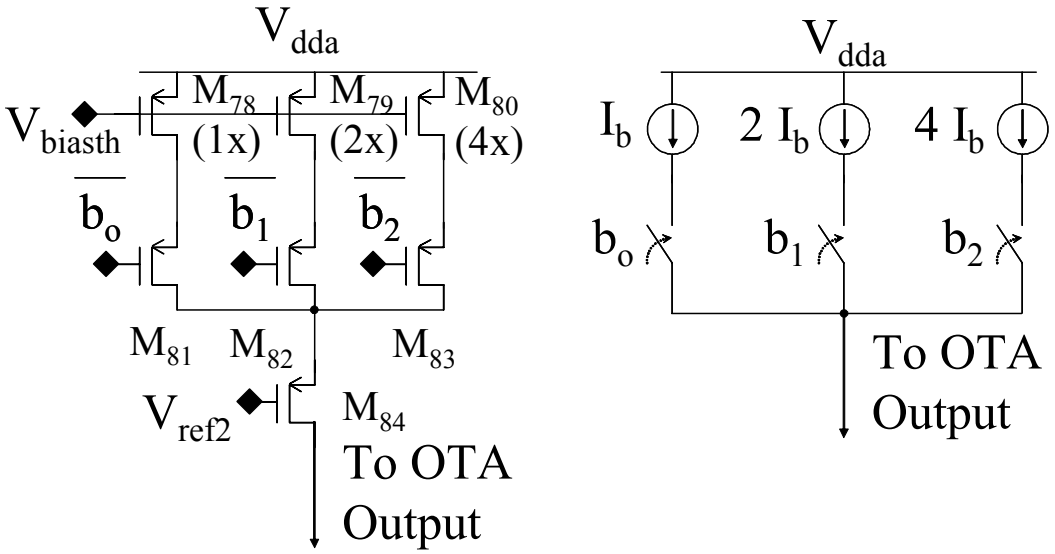

Figure 7.59: Circuit diagram of the threshold fine adjust stage (left, note the inverted values of the bits b_0, b_1 and b_2) and schematic representation of its working principle (right).

7.3.8.4 The digital conversion and the pixel masking

The signal produced by the current discriminator is not able to drive properly a logic gate. For this reason (as shown in Figure 7.60) a simple voltage comparator (transistors M_{67} to M_{74}) has been added in between the current discriminator output and the first logic gate. A signal with faster edges is generated by the inverter M_{76}-M_{77}.

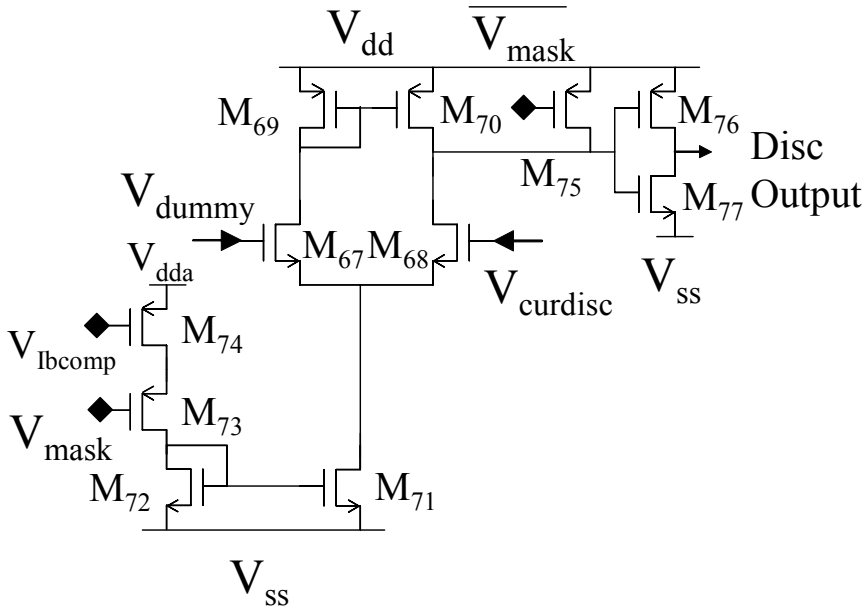

Figure 7.60: Circuit diagram of the voltage comparator with inverting stage.

The pixel can be masked if it is malfunctioning or excessively noisy setting V_{mask} to a high logic state. This shorts, thanks to the transistor M_{75} which acts as a switch, the gates of M_{76} and M_{77} to the positive power supply voltage. For more safety the transistor M_{73} switches off the voltage comparator bias. Three examples of discriminator outputs for a small, medium and huge input signal are shown in Figure 7.61.

This is the first stage that is powered with the digital power supplies V_{dd} and V_{ss} (the power supply strategy is discussed in more detail in section 7.6).

7.3.8.5 Time walk estimation

The time walk at 20 ns, defined as explained in section 6.3.2, has been simulated. Figure 7.61 shows the simulation results.

The threshold is set at V_{th} = 30 mV. The minimum signal over threshold is 1390 e^-. The pulse generated by a huge signal (50000 e^-) crosses 0.8 V (= V_{dd} / 2) at 39 ns (note that the input signal is applied at t_0 = 10 ns). The signal which crosses 0.8 V at 59 ns is generated by an input pulse of 1530 e^-. This means that the time walk at 20 ns is 140 e^-.

The choice to define the time walk at 20 ns comes from the fact that the clock period for LHCb is 25 ns, so we decided to apply a more stringent requirement than 25 ns to the time walk to leave 5 ns of contingency for the signal delay propagation.

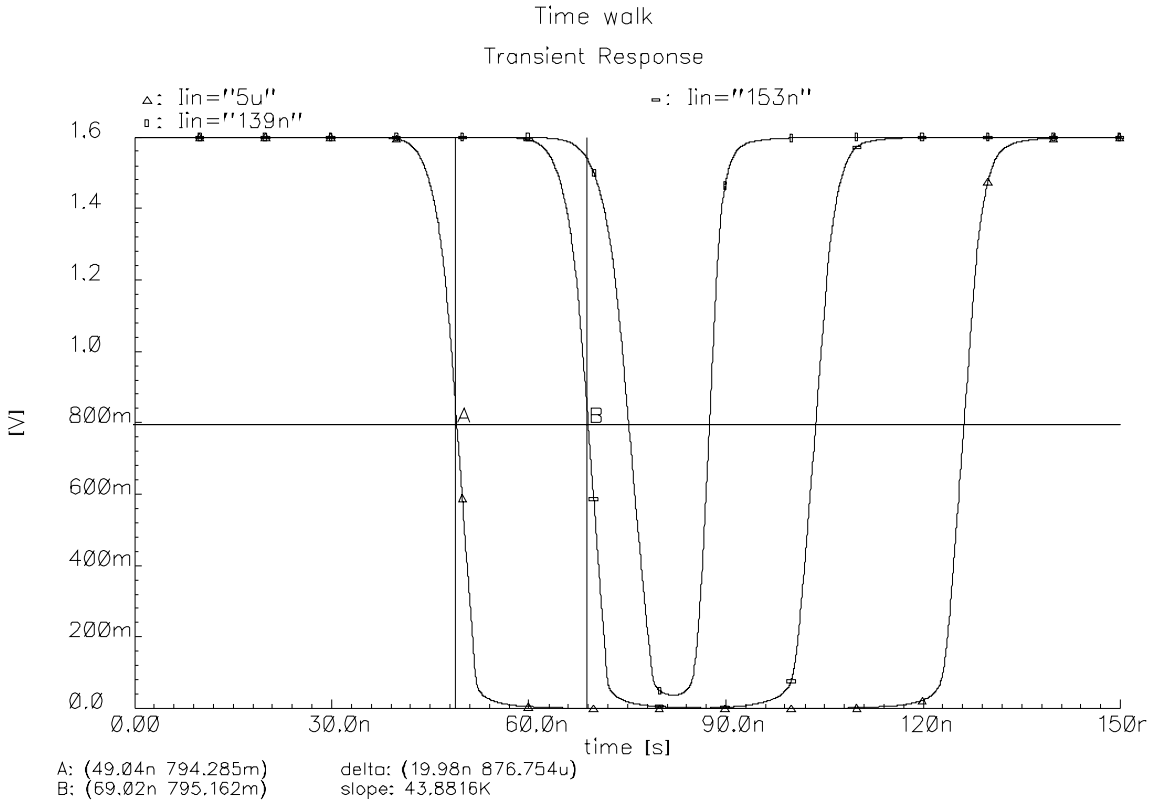

Figure 7.61: Time walk simulations. The time walk at 20 ns is 140 e^- (1 nA input current = 10 e^-, V_{th} = 30 mV; the input signal is applied at t = 10 ns).

7.3.8.6 The fast-OR and fast-multiplicity circuitry

The discriminators in the pixel matrix provide also a *fast-OR* signal. This means that if a discriminator fires, a current pulse synchronised with the clock is sent on a bus which is common to all the pixels in a column. At the bottom of the column the signal is discriminated and OR-ed with the signals coming from the other columns, so that the information that at least one pixel in the chip has been hit is provided on the fast-OR pad after 1 clock cycle. The fast-OR signal can be useful for diagnostic purposes during testing or for self-triggering.

A synchronised current pulse is also sent on a *fast-multiplicity* bus. The fast multiplicity circuit, at the periphery of the chip, generates a current pulse whose magnitude tracks the number of hits in each clock period. The output current is proportional to the number of hit pixels. Nominally the current/hit pixel is set to 5 μA. Adjusting the corresponding DAC, this value can be varied between close to 0 μA and 10 μA. For a large number of hit pixels some saturation will be present (there is some dependence on whether all hit pixels were located in one column or spread more evenly across the chip).

7.4 The digital back end

The full digital back end is composed by the synchroniser and the delay unit selection, the delay units, the FIFO buffer and the readout logic [Wyl01].

The design of the digital back-end makes use of Grey encoding to minimise the logic state transitions, and thus the digital noise.

7.4.1 The synchroniser and the delay unit selection

The synchroniser circuit accepts the signal directly from the discriminator, and synchronises this to the chip clock. The circuit is shown in Figure 7.62 (left). The leading edge of the discriminator output sets a flip flop. The output of this is then latched by the clock into a second flip-flop, which resets the first.

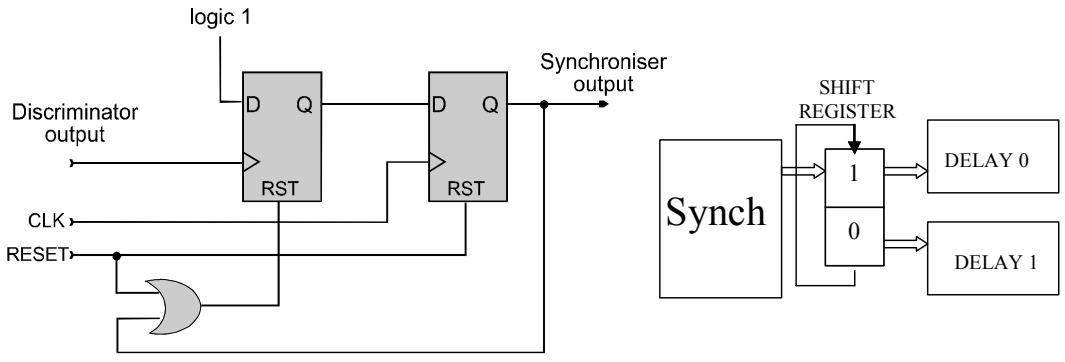

Figure 7.62: Schematic block diagram of the synchroniser (left) and of the delay line selection (right).

Each pixel cell (in the basic mode of operation; the two modes of operation are described in section 7.4.4) has 2 delay lines (Figure 7.62, right), so that the possibility that a hit needing to be delayed finds a delay line which is still busy processing the previous event is divided by two. A 2-level roll-over shift register contains the address of the delay line that was used the previous hit. The output of the synchroniser clocks the shift register, so that the other delay line is chosen, and the delay count is started.

7.4.2 The delay units

The next stage consists of two digital delay units, whose purpose is to store a hit for the duration of the trigger latency. Figure 7.63 shows a schematic block diagram of the delay units.

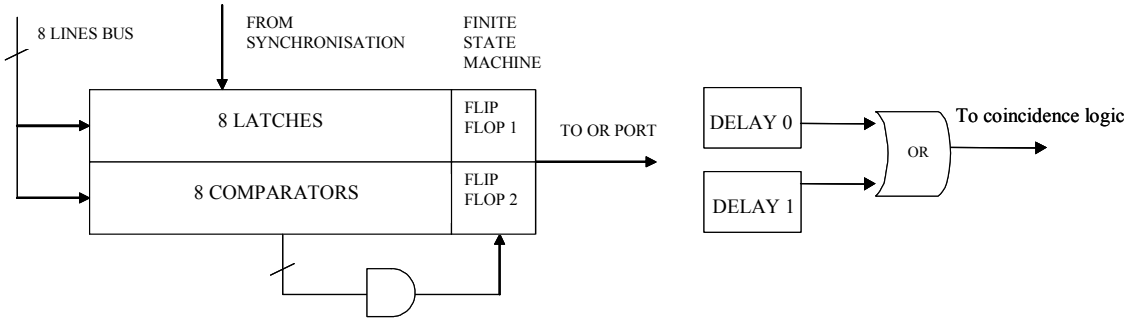

Figure 7.63: Schematic block diagram of the delay units (left). The output of the two on-pixel delay units is OR-ed before being sent to the coincidence logic (right).

Each delay unit consists of an 8-bit latch that, on receipt of a hit from the discriminator, latches the bit-pattern present on an 8-bit bus. This pattern is the Gray-encoded content of an up-down counter whose state changes synchronously with the clock and has an adjustable modulo-n. The time structure of the pattern is shown in Figure 7.64.

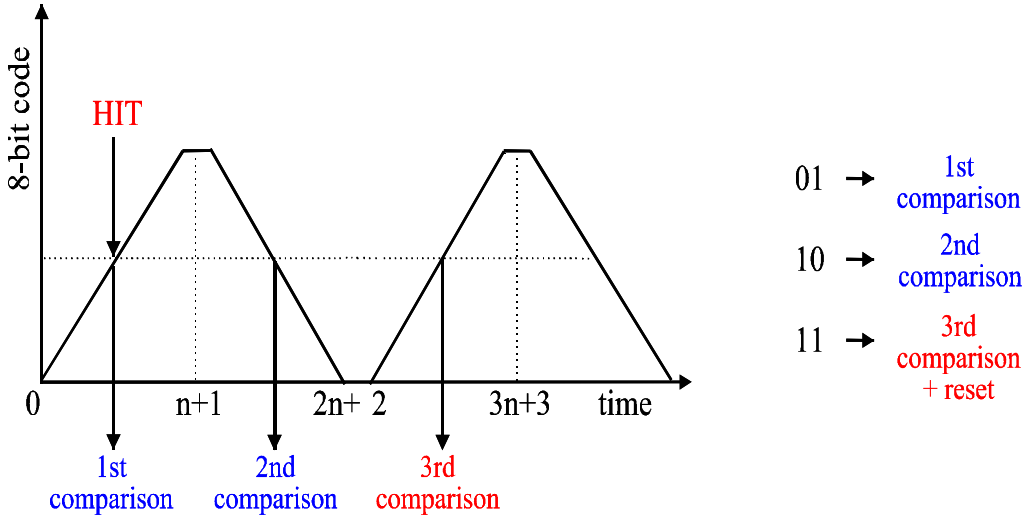

Figure 7.64: The time structure of the pattern used to timestamp hits in the delay units.

The contents of each latch are compared with the pattern on the bus and, on each positive comparison, a 2-bit counter is incremented. The third positive comparison occurs 2n+2 clock ticks after the hit, and a logic 1 is then presented to the coincidence logic. This state also resets the delay unit. In this case, 2n+2 clock ticks represent the trigger latency and this can be adjusted to meet the requirements of the experiment.

The outputs of the two on-pixel delay units are OR-ed before being sent to the coincidence logic (Figure 7.63, right).

All the digital logic in the pixel cell is under current control to limit the switching speed and minimise noise injected into the analogue circuitry. Biases are used to control the switching speed for each digital block.

7.4.3 FIFO buffer and Readout logic

The result of the trigger coincidence is loaded into the next-available cell of a 4-event FIFO, which acts as a multi-event buffer and de-randomiser. This FIFO is read/write addressable by means of two 4-bit busses that again carry Gray-encoded patterns.

Each channel contains four static latches, which are the memory of the buffer, and two decoders. These decode the 2-bit words which address the latch to be written to or read from. The 2-bit addresses are generated in a single block in the chip periphery and bussed up the pixel columns.

The contents of the FIFO cells waiting to be read out are loaded into a static flip-flop by a NEXT-EVENT-READ signal (external to the chip). The flip-flops of each column form a shift register, and data are shifted out using the system clock. During readout, this flip-flop is clocked by a signal derived from the system clock supplied from outside. A schematic representation of the FIFO and the readout logic is shown in Figure 7.65.

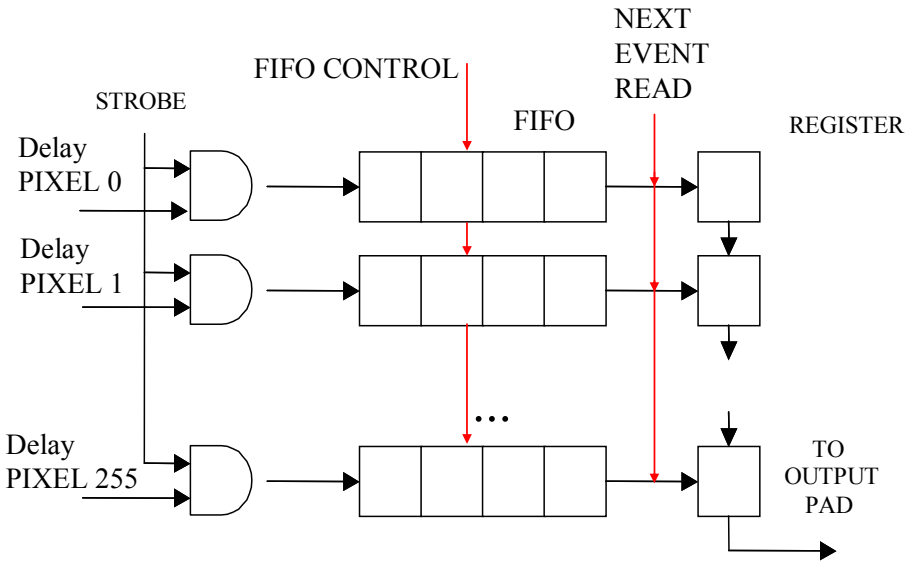

Figure 7.65: Schematic representation of the FIFO and of the readout logic.

7.4.4 The two operational modes

The chip can be operated in two modes, selected by an external control signal: the ALICE mode and the LHCb mode.

7.4.4.1 The ALICE mode

In ALICE mode, each pixel cell acts as an individual channel and the full matrix of 256×32 cells is read out. Using the two delay units, each cell has the capability of simultaneously storing two hits for the trigger latency. Figure 7.66 shows the configuration of the pixel cells in ALICE mode. In this operational mode the chip is clocked at 10 MHz.

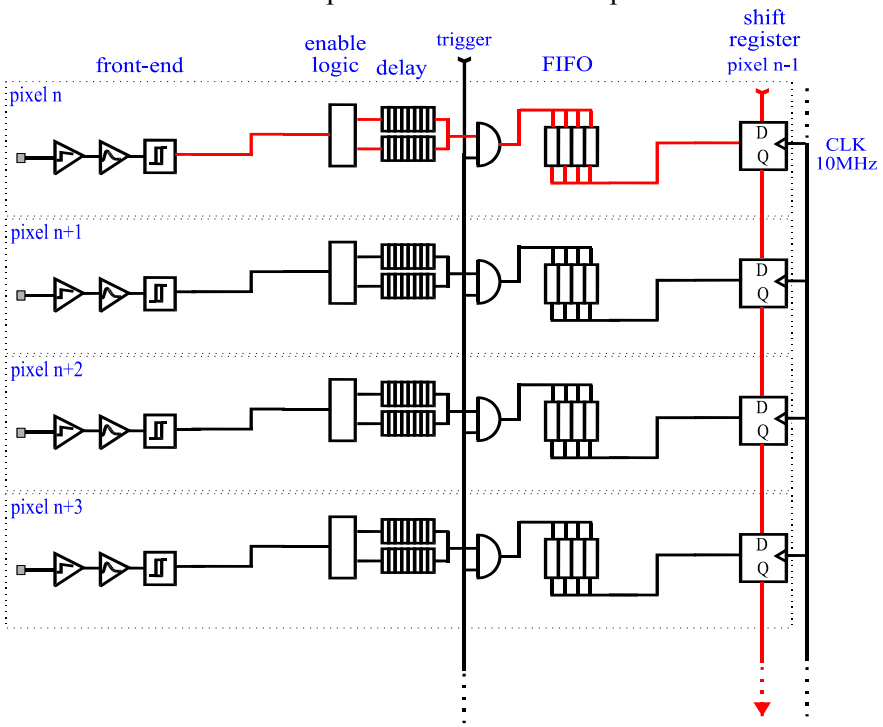

Figure 7.66: The configuration in ALICE mode.

7.4.4.2 The LHCb mode

In LHCb mode, eight pixels in the vertical direction are configured as a 'super-pixel' of 400 μm × 425 μm. The discriminator outputs within a super-pixel are OR-ed together and the sixteen delay units of these eight cells are configured as an array. Four of the 4-event FIFOs are connected together to form a 16-event FIFO, which can be written to by any of the sixteen delay units. The FIFO output is loaded into the flip-flop of the top pixel in the group, which bypasses the other seven during readout. This scheme reduces the matrix to 32×32. Figure 7.67 shows the configuration of the pixel cells in LHCb mode. Note that the glue logic which performs the OR-ing operations is again current-controlled by means of specific biases. In this operational mode the chip is clocked at 40 MHz.

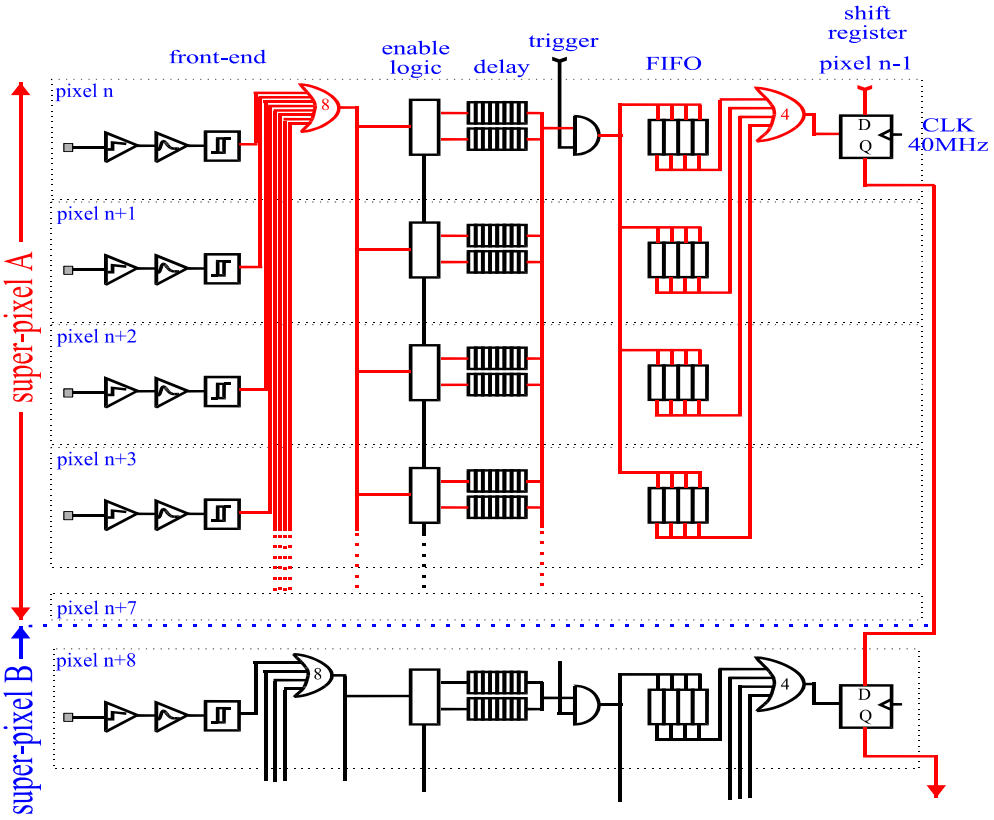

Figure 7.67: The configuration in LHCb mode.

7.4.4.3 Asynchronous operation

In addition to the two experiments ALICE and LHCb some other experiments have expressed interest in using this chip. In the case of NA60 [Na6I] the events will not be synchronous to a bunch-crossing clock. The chip will synchronize the trigger pulse to an externally applied clock. For these applications one has to make the chip sensitive for a sufficiently long time after the event to ensure full efficiency by having the trigger pulse last over a few clock cycles if necessary.

The important parameter here is the timewalk of the front end. The requirement in LHCb to associate the pixel hits to the correct bunchcrossing or 25 ns interval also puts limitations on this timewalk. The simulated time walk of 140 e^- at 20 ns ensures the required full efficiency with conservation of position information (i.e. charge sharing information) for particle hits where charge is shared between several pixels in the chip.

7.5 The chip peripheral circuitry

7.5.1 Peripheral Control Logic

The periphery of the chip contains a number of circuit blocks to control the operation of the chip during data acquisition. These are shown in the block diagram of Figure 7.68.

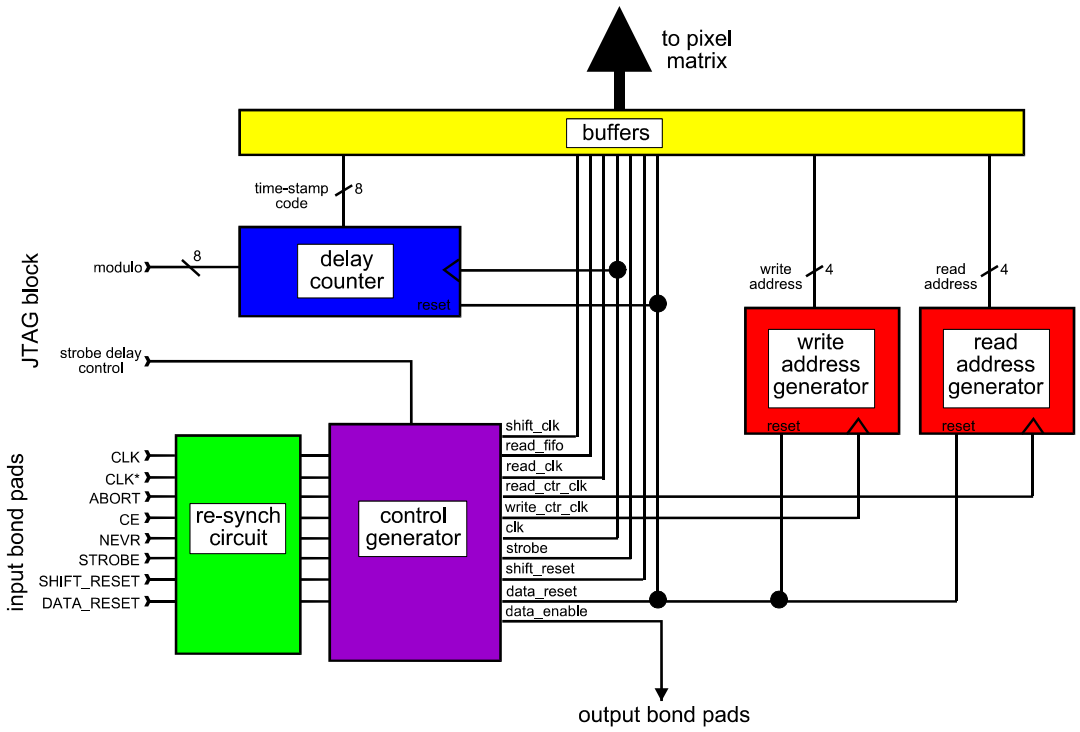

Figure 7.68: Block diagram of the peripheral control logic.

Re-synchronisation circuit: all the synchronous digital control signals applied to the chip are re-synchronised to the clock.

Control generator circuit: this accepts the external control signals and converts them to the format required inside the chip.

FIFO address generators: two blocks generate the address pointers for writing and reading the FIFOs. Note that the outputs of these blocks are 4-bit addresses. In ALICE mode, only the two least-significant bits are used for addressing the FIFO elements. In LHCb, all four bits are used. These counters are configured appropriately with the choice of operating mode (ALICE or LHCb).

Delay counter: the pattern which is used to timestamp hits in the pixels, as described in 7.4.2, is generated by a counter in a peripheral block. The 8-bit modulo (n) of this counter can be programmed via the JTAG interface. The value of n after chip reset is 10000000.

Trigger-strobe delay bit: a pixel hit is delayed by $2n + 2$ clock periods, i.e. an even number of clock periods. If the optimal delay is in fact an odd number of clock ticks, then the

trigger (strobe) received from outside can be delayed by one period on the chip. This delay can be switched on or off via the JTAG interface. It is switched off by a reset of the chip.

Buffers: the control signals are sent to the pixel matrix in two stages. One set of buffers distributes the signals from the central control block to the bottom of the pixel columns. Sets of buffers at the bottom of each column drive the signals up to the pixels. The switching speed of these buffers is controlled by a dedicated DAC.

The first three blocks here described are transparent to the user. The last three can be configured to meet the requirements of the system.

7.5.2 The Digital-to-Analogue Converters (DACs)

There are 42 8-bit DACs included in the chip. Their role is to provide bias (voltages and currents) to the analogue and digital circuitry within the pixel cells. All are configurable via the JTAG interface.

Figure 7.69 shows the principle of the current DAC [Cor00]. An operational amplifier circuit is used to guarantee the correct ratio between the currents in two branches [Sno01].

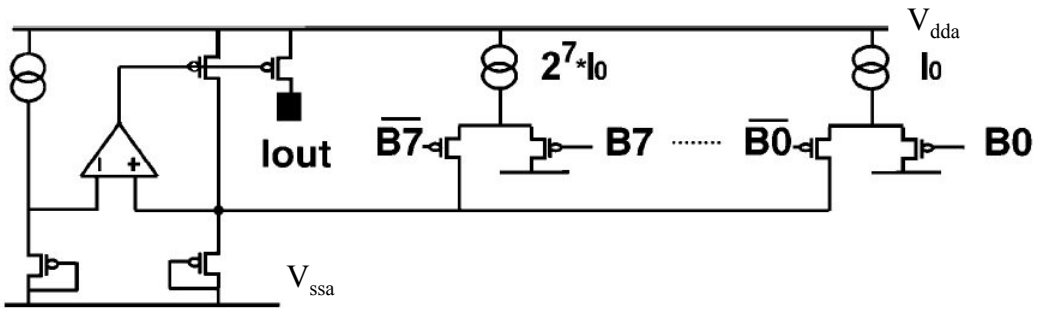

Figure 7.69: Principle of the DAC. The opamp circuit is necessary because accurate NMOS current mirrors for low current could not be made with NMOS transistors in edgeless layout [Sno01].

The main reason for this circuit topology is that the low current levels and the special layout of the NMOS preclude accurate NMOS current mirrors (see section 7.3.4). The bit value determines whether the current associated with this bit is switched into the output or not. The more current is sent to the operational amplifier input, the lower the output current of the DAC. To avoid nonlinearity errors in the DAC due to transistor parameter gradients on the chip, attention was paid to having the same centre of gravity for the current sources of all bits (see section 7.6). The voltage DACs are constructed from these current DACs by adding a current-to-voltage conversion stage, which was made robust against power supply and transistor parameter variations. A resistor could not be used for reasons of design portability and the lack of characterization of its radiation tolerance.

The output stage of each DAC has been tuned so that it provides a sensible nominal setting following a reset[26]. The voltages and currents produced by the DACs can be measured one at a time on one of two dedicated pads. The selection of which signals are sent to these pads is done during the DAC configuration via the JTAG interface. When a current is to be sensed on the output pad, it is scaled to 20 μA nominal to ease the off-chip measurement. By adjusting the code of the DAC, the output of the current DACs can be varied between almost zero and twice the nominal value. For the voltage DACs, the voltage can be varied over a range of almost 500 mV centred around the nominal value. Some of the biases were judged sufficiently sensitive to distribute them as currents individually to each column. The voltage bias required for this current is then generated locally at the bottom of each column.

Two biases must be supplied from outside. They are used to generate a reference voltage for the DACs on the chip. These two biases must be low-noise and adjustable between 0.8V and V_{dda} with a precision of ~10mV. In the test setup they are generated by a dedicated board, in the experiment they will be generated by the Analogue Pilot Chip (see Chapter 9 for more details).

7.5.3 Input/Output Pads

There are two sets of Input/Output wire-bond pads on the chip. One set is located along the top edge of the chip, and is foreseen for chip testing only. They will be inaccessible once a sensor is bump-bonded to the chip.

The other set of pads is located at the bottom of the chip. There are 103 pads along the bottom edge at a pitch of 120 μm, but grouped into three sections of 20, 30 and 53 pads with blocks of control circuitry in between. This row contains all the necessary connections to operate the chip and in the ALICE scenario will be the only pads wire-bonded to the support structure.

There are two further groups of wire-bond pads, 6 pads in each, located at the bottom left and right hand corners of the chip. These are connected internally to the 6 pads at each end of the bottom row. These are for use in the LHCb scenario, where the chip carrier may be pad-limited along one edge and connections would have to be made to the other edges of the carrier.

Each pad has a $100\ \mu m \times 100\ \mu m$ passivation cut for a wire-bond connection. Additionally, there is a cut for a possible bump-bond connection which is 20 μm in diameter. Every pad has diodes to protect against electrostatic discharge.

The width of the ALICE stave restricts the number of lines available for signalling, and for this reason a single-ended standard has been adopted. Gunning Transceiver Logic (GTL) [Gun92] will be used for all the digital signals to and from the chip. This has the advantage of

[26] The digital setting of each DAC after reset is 10000000.

a low signal swing (~800 mV, from 0.4V to 1.2V) which again reduces the risk of noise injection. Additionally, the output buffers on the chip have an adjustable slew-rate control which can be controlled to match the system requirements[27]. These buffers are powered with a supply separate from that of the rest of the chip. Finally, multiple bonding pads have been provided for the supply lines to reduce the inductance of the connection and limit the supply bounce during switching. The digital output pads can be set to a high impedance mode.

7.5.4 JTAG Interface

The chip is configurable via a JTAG interface**.** The majority of this interface circuitry complies with the IEEE Standard [IEEjt]. All necessary commands and configuration data are passed serially into the 6-pin JTAG port at the bottom left hand corner of the chip. The most important roles of the interface are:

- to transfer configuration data to the pixel columns (e.g. the TEST bits),
- to set the on-chip DACs, and
- to carry out boundary scans to test the connectivity of the chip on a board.

In compliance with the standard, the JTAG block on the chip consists of a controller, instruction register, boundary scan register, bypass register and data registers. In addition, there is a block to carry out a simple check of the functionality of the previous chip in the serial chain (scan check sequence). This is used only in the ALICE mode of operation.

7.6 Layout considerations

7.6.1 Hardening by design

The chip has been designed in our target 0.25 μm CMOS commercial technology, but with the HBD techniques already explained in Chapter 5.

In particular we used ELTs for n-channel transistors to avoid radiation induced edge effects, and p^+ guard rings to prevent radiation induced inter-transistor leakage currents (the design constraints due to ELTs, and our solutions, were discussed in the previous sections). Moreover, the digital section of the chip makes use of spatial redundancy techniques for all the configuration registers of the chip to improve their SEU-tolerance. The chip sensitivity to radiation has been tested, and is well within the ALICE (and LHCb) specifications, as will be shown in Chapter 8.

The use of HBD techniques results in transistor density loss. The average density loss factor of about three, mentioned in Chapter 5, seems to be confirmed by Figure 7.70. It shows

[27] The switching of the circuitry in the input pads is controlled by setting a threshold voltage to mid-range of the voltage swing. These thresholds must be generated externally and supplied to pins. The nominal threshold voltage is 800mV.

the chip component density (logarithmic scale) as a function of the time for 5 pixel chips designed at CERN. The OmegaD, Omega3 and Medipix2 [Llo02] chips, which use standard design, are on the same straight line. The ALICE1LHCb chip, which uses HBD techniques for radiation tolerance has a component density about three times smaller than the Medipix2 chip. However, there is still a factor of 10 which separates an HBD design from a design done using a radiation hard process, such as the Atlas FE-D (done in DMILL technology) [Cam01].

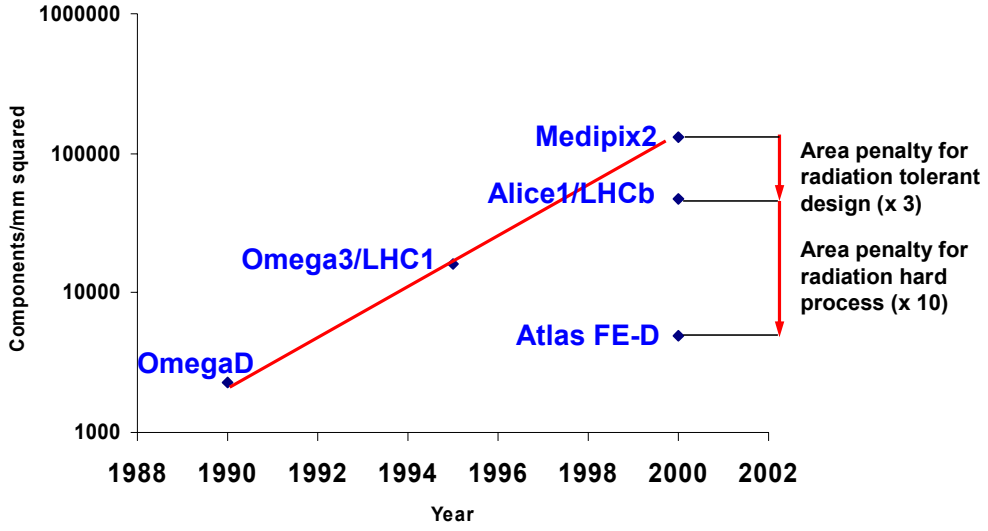

Figure 7.70: Chip component density (logarithmic scale) as a function of the time for 5 pixel chips designed at CERN. The OmegaD, Omega3 and Medipix chips use standard design, ALICE1LHCb uses HBD techniques and ATLAS FE-D uses a radiation hard process (DMILL) [Cam01].

7.6.2 Layout of the pixel cell

Figure 7.71 shows the layout of a pixel cell. The cell is 50×425 μm^2, and is divided in three sections. The digital back-end is 265 μm long, the analogue section is 125 μm long and the "head" of the pixel is 35 μm long. In this section there are the five unupsettable latches which store the chip configuration (3 for the threshold adjust, 1 for the test input and the last one for the pixel masking). These latches are not clocked during data taking, so form a "quiet" region which separates the analogue font end of a pixel (and in particular the input, which is situated at the left-hand side of the analogue section) from the digital section of its neighbour.

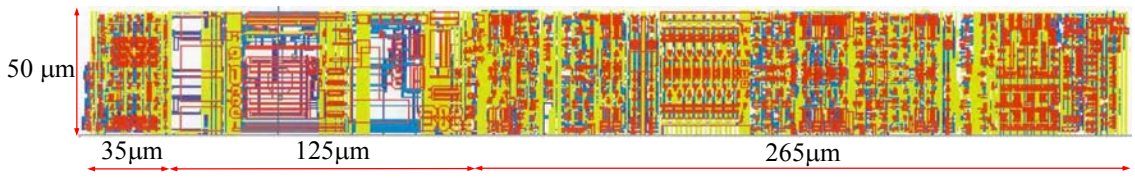

Figure 7.71: Layout of a pixel cell. The digital back-end is 265 μm long, the analogue section is 125 μm long and the "head" of the pixel is 35 μm long. The total pixel size is 50×425 μm^2.

To improve the digital-to-analogue crosstalk rejection, the analogue section in the pixel cell is separated from the digital section by means of a gap of about 5 μm and a double guard ring. The feedback transistors are also enclosed by a guard ring, which helps reduce the possibility of noise injection into the preamplifier. To cross check the circuit behaviour in more detail, the parasitic capacitances of the feedback transistors were evaluated after layout and fed back into the simulation program.

7.6.3 Signal routing and power distribution scheme

The area saving is a major objective while laying-out a pixel design, so transistors are often placed at the minimum possible distance, and all the metal layers are exploited for routing and power distribution. Our target technology offers a maximum of 6 metals, so the general rule was to use the first two metal layers for local routing and the third metal layer for bussing (bias voltages and cell interconnect). Voltage drops in the power supply lines can disturb the pixel functionality, so all the last three metal layers were used for power distribution (ground and positive power supply). The last metal has a lower resistance, so in general it is used for one of the two power lines, while the other two metal layers, shorted, for the second power line. All the sensitive area of the chip was covered with at least one metal layer, which results in a shielding effect.

The different blocks forming the analogue front-end have different sensitivities to voltage power drops. To maximize the efficiency of the power distribution, each block had its own power supply line, with a well known voltage drop. This allowed to increase the width (i.e. decrease the voltage drop) of the metal lines powering the most sensitive blocks.

For instance, the circuit is weakly sensitive to preamplifier current variations. This is also the largest current in the front-end. The voltage drop here was allowed to be larger than that in the feedback circuit and in the shaper stages, where care was taken to minimise the voltage drop.

The detrimental effect of power supply voltage drops in the digital section was underestimated, and this lead to problems of the circuit functionality for clock frequencies higher[28] than about 15-20 MHz (as explained in the next chapter).

Another technique employed to reduce analogue-to-digital crosstalk consists in separating the power supply lines powering the analogue section from the power lines powering the digital section (V_{dda} for the analogue and V_{dd} for the digital positive supply, V_{ssa} for the analogue and V_{ss} for the digital negative supply) and separating the lines carrying a high current from the lines without current. For this reason we added V_{wella} to bias the analogue well, and V_{well} for the digital well. The same thing was done for the chip analogue (V_{sub}) and digital (V_{suba}) substrate connections.

[28] The digital power supply current increases with clock frequency, and so does the voltage drop on the digital power supply lines.

7.6.4 The feedback and injection capacitances

The feedback and the injection capacitances are realised as metal-to-metal parasitic capacitors, and are laid out in a special way, shown in Figure 7.72.

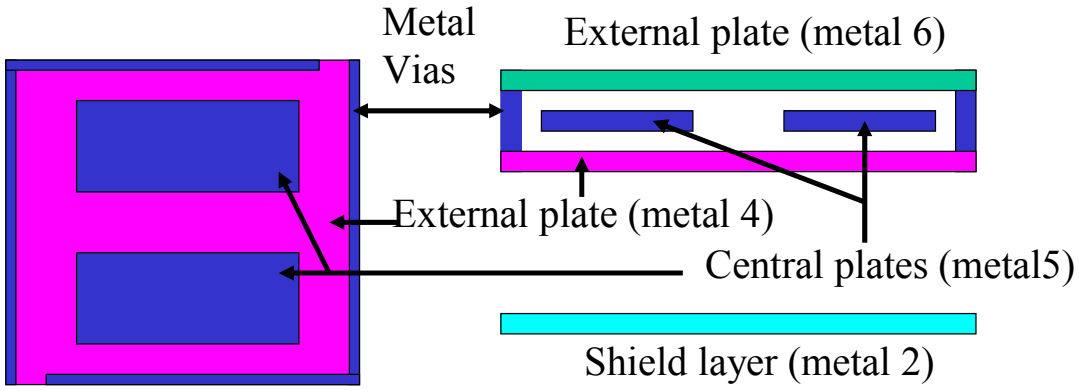

Figure 7.72: Special "sandwich structure" for the feedback and injection capacitors. Top view (left) and side cross section (right).

The "sandwich structure" used for the feedback and injection capacitances reduces the space needed for each capacitance, because it makes use of three levels of metal so that two capacitors (electrically in parallel) are stacked vertically. Moreover, this structure reduces the parasitic capacitances, especially for the preamplifier output to keep its rise time under control.

7.6.5 Transistor matching

Wherever transistor matching was a priority, common centroid structures were used to compensate for gradients in the dopant distribution along the chip. An example of common centroid structure for two matching transistors consists of splitting each device in two, and cross-interconnect the four sub-devices as shown in Figure 7.73.

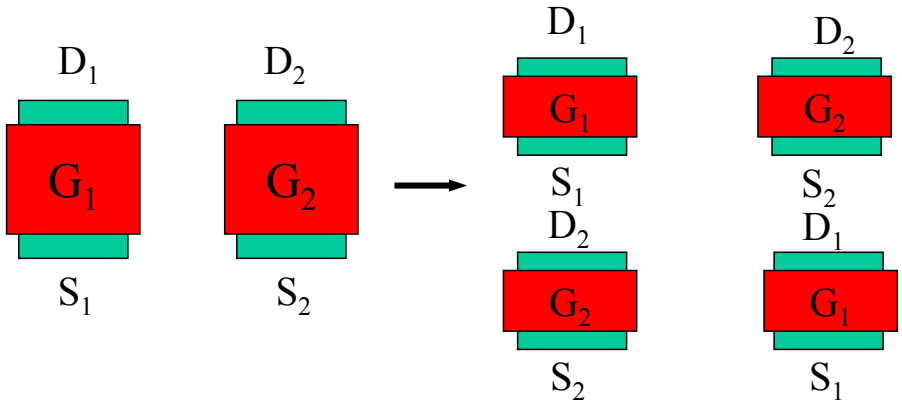

Figure 7.73: Working principle of the common centroid layout technique. Each transistor of the matching couple (left) is split in two, and the sub-transistors are cross-connected as shown in the picture (right; terminals with the same name are connected together).

7.6.6 The test structures

For the ALICE SPD 5 chips have to be bump-bonded to a single ladder detector, as explained in Chapters 3 and 4. If we suppose a chip yield of about 50% (which is reasonable given the chip size, the complex design and the stringent specifications) this would result in a ladder yield of less than 0.03% and would increase enormously the production time and cost. This means that it is extremely important to provide the bump-bonding vendor with known good dies (KGD).

If the chip is not bump-bonded to a detector, it is impossible to check its functionality, so we implemented on the chip some test features, to have the possibility to test the chips prior their bump-bonding to a detector (possibly already at wafer level), and to have the possibility to test the front-end performance, as the pixels do not have any analogue output.

7.6.6.1 The injection circuitry and the on-chip pulser

An injection capacitance is put in series with the preamplifier input of each pixel, as shown in Figure 7.33 and Figure 7.74 (right). If the pixel has been set in TEST mode the capacitance is connected, by means of CMOS switches, to the input. If the pixel is in data taking mode the capacitance is disconnected (and shorted). To simulate a charge injection from the detector a voltage pulse of known amplitude V_{in} (equal to TestHi – TestLow) is stepped onto the injection capacitance C_{inj}, which acts as a differentiator injecting into the preamplifier a charge pulse of $Q_{in} = C_{inj} \cdot V_{in}$. The capacitance is chosen to be $C_{inj} = 16$ fF, to have a simple voltage/charge conversion factor (1 mV = 100 e^-).

The voltage pulse is generated by an on-chip pulser, shown in Figure 7.74 (left). The pulser output is connected at TestLow thanks to M_{P2} if the digital signal TestPulse is high; when TestPulse switches to a logic 0, M_{P1} is switched off and M_{P2} is switched on connecting the pulser output to TestHigh. The transistors M_{P3} and M_{P4} work in the same (but complementary) way to create an inverted input pulse, which is generated only for symmetry reasons.

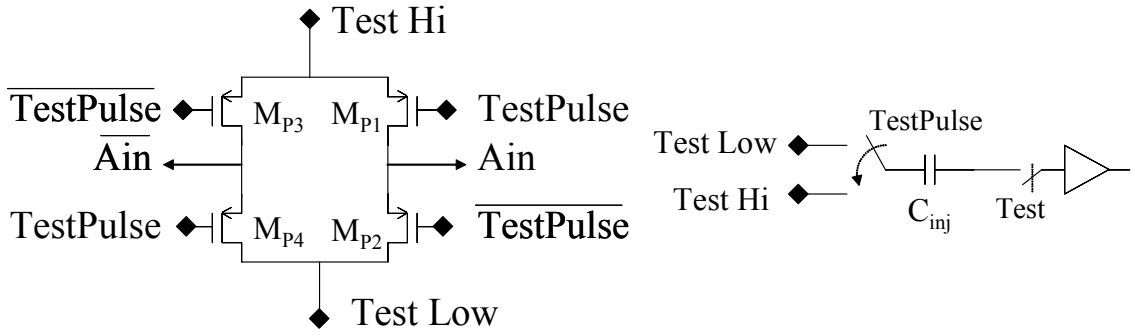

Figure 7.74: Circuit diagram of the on-chip pulser (left) and schematic representation of its working principle.

The two DC levels are provided to the chip from outside. The upper level TestHi should be at 1.6 V and TestLow should be at some level below that. Note that the switches in each pixel that determine whether the test pulse is applied to the pixel or not are PMOS transistors. Therefore, if TestLow drops below 0.8 V, there will be some saturation in the signal applied to the inputs because of increased resistance of the switch. These voltages should be as quiet as possible, and the difference between the two voltages should be controllable with a 1 mV precision.

7.6.6.2 The analogue outputs

To increase the chip testability some pixels of the top row (in column 1, 9, 17 and 25) are equipped with buffers, so that some internal nodes are made available outside the chip on some dedicated pads, as depicted in Figure 7.75.

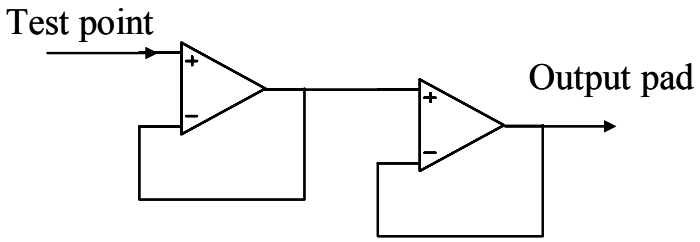

Figure 7.75: Schematic representation of the on chip buffers used to read the test points analog signals.

For the ALICE1LHCb chip the pads are available only for the chips which were not bump bonded to a detector, because the detector covers them after bump-bonding (this is done to minimise the chip dimensions). To pick up the signal from the test points we used a two-level buffer. The first buffer has limited driving capabilities but low input capacitance not to add an excessive load to the test point. The second buffer has much higher driving capabilities (and input capacitance) to drive the signal outside the chip. Figure 7.76 shows an example of signal (the second shaping stage right output) before and after the buffer stage.

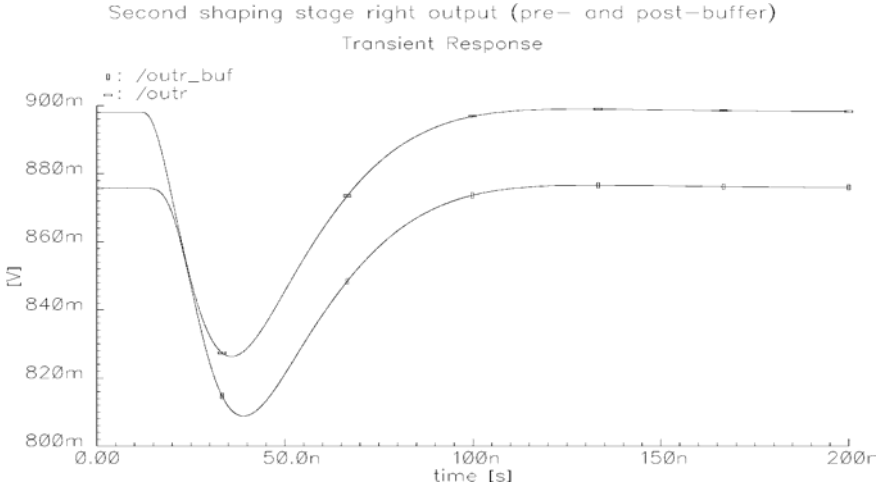

Figure 7.76: Example of signal (the second shaping stage right output) before and after the buffer stage.

The simulation has been done estimating the parasitic resistance, capacitance and inductance of the routing, of the pad and of the wire bonding. The buffer introduces a small DC offset and delay, and shifts the peaking time of the signal of some nanoseconds. The signal is also smaller of about 8% (for a 25 ns peaking time signal).

7.7 Summary

This chapter presented the design and simulations of the front-end of the ALICE1LHCb pixel chip. It makes use of a different front end scheme with respect to the more classical approach of charge integration plus pole zero cancellation and semigaussian shaping. This approach has been conceived to be used in high multiplicity environments, and to have a fast return to zero of all the blocks which compose the readout chain, whilst keeping low noise (< 200 e^-) and low power consumption (of the order of 50-100 μW/channel).

In the first section we summarized all the requirements resulting in our design choices. The second section is devoted to an in depth analysis of the choice of the front-end scheme. The charge integration plus pole zero cancellation and semigaussian shaping approach is analysed, and shown to be extremely impractical for our case. In effect pile-up can be a problem in high multiplicity environments; it would also be necessary to conceive a different scheme for the feedback stage to improve the precision of the pole-zero cancellation, and the return to zero time can be improved only by increasing the order of the shaping. Moreover, the analytical approach proposed in by Chang and Sansen in [Cha91] is only approximate for a scheme with a low-power CSA and a fast shaping time.

This lead us to try to develop a different front-end scheme, with noise performance comparable with the standard scheme, but with improved return to zero time both at the shaper output and at the preamplifier output, to be able to safely operate even in high multiplicity environments [Din03, Din01, Din00, Sno01, Sno01a, Wyl99].

The basic idea is that a system with complex poles could have a faster return to zero than a system with only real poles. Supposing that the CSA is followed by a simple stage realising a pole with time constant τ_{p2} (and with a DC gain A), and that a feedback current is fed back to the amplifier input by means of a transconductance feedback stage g_{mf}, choosing in the proper way τ_{p2}, A and g_{mf} it is possible to realize the complex conjugate poles and to chose their real and complex part. Calculations have to be carried out to investigate the feasibility of the system in terms of noise, return to zero, power consumption and area required to implement the layout. As in our scheme no feedback resistance is present in parallel with the feedback capacitor, a circuital block was conceived to absorb the detector leakage current, without interfering with the high frequency behaviour of the front-end. The scheme corrects also for the offset present at the front-end output.

Several elements point towards a possible improvement of the two poles system. First of all the shape of the output pulse presents an undershoot, which is usually not wanted. The

return to zero time can be improved and the gain has to be increased of about a factor 10, which would be better to realize in a subsequent stage to avoid having a too high gain in the front-end stage. Also for what concerns noise, the performance is comparable to the semigaussian shaping scheme, with a dominant shot noise contribution that could be reduced. For this reason we tried to improve the front-end by going to a three poles scheme.

Intuitively, the real part of the poles (in the s plane) represents the exponential decrease of the output pulse, and is dominated by the slowest pole. If all the poles have the same real part, none is dominating the time response and this should improve the return to zero capabilities of the system. Ohkawa and coworkers propose an optimum ratio of 1.1 between the real pole and the real part of the two imaginary poles. However, the assumption of having three poles with the same real part simplifies greatly the calculations, so that they can be carried out analytically and a closed form for all the formulae can be found. For these reasons we decided to analyse the configuration with three poles on a straight line. All the analysis performed for the two poles system was repeated for the three poles system, showing better performance with respect to all the circuit parameters.

A more realistic analysis is carried also out taking into account the rise time of the preamplifier. The calculated output of the four poles system for an input charge Q_{in} = 5000 e^- has a peak of 192 mV at 26 ns, while the return to zero time at 1% is 114 ns. The preamplifier also has a very fast return to zero time at 1%, of about ~65 ns. The total electronic noise is ENC_T = 117 electrons.

The detailed analysis of the implementation of all the stages which compose the front-end chip is presented, with simulation results, which follow closely the theoretical analysis. The schematic diagram of the scheme used to implement the three poles system is shown in Figure 7.18. A stage which realises a real pole τ_{p3} and a gain A_3 is added to the two poles system.

From the measurements done on the test chips it became clear that the minimum threshold and the noise performance of the chip was dominated mainly by the digital-to-analogue crosstalk , so in the new design this issue was addressed with high priority. For this reason, for example, the charge sensitive amplifier was designed with a differential input which rejects the substrate noise. Moreover it has p-channel input transistors, to be able to short the transistor bulk with the source and avoid noise injection through its gate to bulk capacitance. All the stages in the front-end chain are differential for the same reason. The system was also optimised to minimise the pixel-to-pixel capacitive charge sharing (no spurious hit due to crosstalk is produced in simulations up to an input charge of 55,000 e^- for a threshold of 2300 e^-), and to recover quickly after a huge input signal (for an input charge of 50000 e^- which is the highest charge foreseen in LHCb and a threshold of about 2300 e^- after ~200 ns the system is ready to process another pulse, even if the slow signal tail slightly offsets the pixel threshold). The impact of temperature on the front-end was simulated to be

negligible, and care was taken to have enough phase margin in the feedback loops (for the high frequency loop it is 72° and for the low frequency loop it is 82°) and to reduce the possibility of noise coupling through the power supply lines.

The discriminator, which has to compare the analogue signal produced by the front-end to a global threshold and produce a digital pulse if the signal is higher than the threshold, was designed to be low power. Moreover, it has to minimise the difference of the response delay caused by input pulses of different magnitude (time walk). The time walk at 20 ns was simulated to be better than 140 electrons.

The calculated threshold dispersion (referred at the preamplifier input) is 70 e^-, but a block was added in the discriminator to further reduce the threshold dispersion in case it would be higher than expected.

A description of the digital back-end of the chip and of the chip periphery was given in the last sections, as well as some layout considerations explaining some important strategies employed to improve chip performance.

The chip can be operated in 2 different modes, ALICE and LHCb, to meet the different requirements of the two experiments. In particular, the full digital back end is composed by the synchroniser and the delay unit selection, the delay units, the FIFO buffer and the readout logic. The first block synchronises the discriminator output with the chip clock, and feeds one of the two programmable delay lines, which has to store a hit for the duration of the trigger latency. The result of the trigger coincidence is loaded into the next-available cell of a 4-event FIFO, which acts as a multi-event buffer and de-randomiser. The contents of the FIFO cells waiting to be read out are loaded into a static flip-flop by a NEXT-EVENT-READ signal (external to the chip). The flip-flops of each column form a shift register, and data are shifted out using the system clock.

In ALICE mode, each pixel cell acts as an individual channel and the full matrix of 256×32 cells is read out. In LHCb mode, eight pixels in the vertical direction are configured as a 'super-pixel' of $400\ \mu m \times 425\ \mu m$. The logic is reconfigured accordingly. The chip periphery contains the peripheral control and JTAG logic and 42 8-bit DACs for chip biasing. Moreover, if the chip is not bump-bonded to a detector, it is impossible to check its functionality, so we implemented on the chip some test features, to have the possibility to test the chips prior their bump-bonding to a detector. An injection capacitance is put in series with the preamplifier input of each pixel to simulate a charge injection from the detector. The voltage pulse is generated by an on-chip pulser. To increase the chip testability some pixels of the top row (in column 1, 9, 17 and 25) are equipped with buffers, so that some internal nodes are made available outside the chip on some dedicated pads.

The chip is designed with HBD techniques, and this gives additional problems during the design phase, but the area penalty due to this type of approach was shown to be up to a factor 10 less than a design done using a radiation hard process, such as the Atlas FE-D chip.

Chapter 8
The ALICE1LHCb and LHCbpix1 chips: experimental results

The ALICE1LHCb chip contains a matrix of 32 columns each containing 256 readout cells, measuring 13.5×15.8 mm^2. Five chips are bump-bonded to a big detector (160 columns $\times$ 256 rows) to form the ALICE Silicon Pixel Detector (SPD) basic detecting block, *the ladder*. One chip is bump bonded to a detector (of the same dimensions, 32 columns $\times$ 256 rows) to form a *single*, the basic detecting element for the LHCb Hybrid Photon Detector (HPD). Figure 8.1 shows a photograph of the chip.

The ALICE1LHCb chip was tested extensively, both in the lab (with the test system described in Chapter 3) and in a beam of particles. The main experimental results are presented in this chapter.

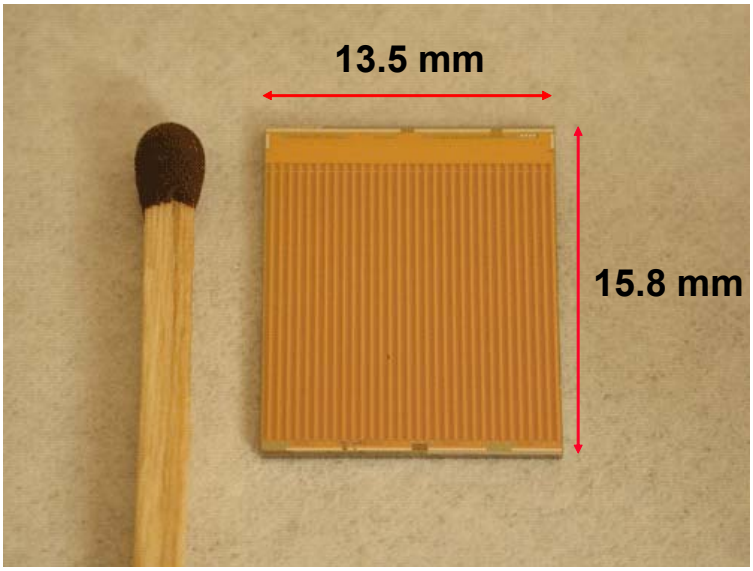

Figure 8.1: Photograph of the ALICE1LHCb chip. Its size is 13.5 mm $\times$ 15.8 mm.

8.1 The two versions of the front-end chip

The chip is meant to be used for two experiments, in the ALICE Silicon Pixel Detector and in the LHCb HPD. The first version of the chip was working well enough to meet all ALICE requirements, but this was not the case for LHCb. The major limitation was that the chip stopped to work at clock frequencies higher than 15-18 KHz, while for LHCb a 40 MHz clock is required. Moreover, the physical size of the "superpixel" described in section 7.4.4 is 425 μm × 400 μm, which is not exactly the required pixel size of 500 μm × 500 μm.

This lead to the resubmission of a second version of the chip, the LHCbpix1 chip. The channel size was increased to 65 μm × 500 μm to match the requirement of 2.5 mm × 2.5 mm at the photodetector plane and allow some overhead in coverage given the estimated image distortions due to stray magnetic fields, for a total chip area of about 16 × 21 mm^2.

Figure 8.2 shows a picture of the chip bump-bonded to a sensor and placed in the carrier.

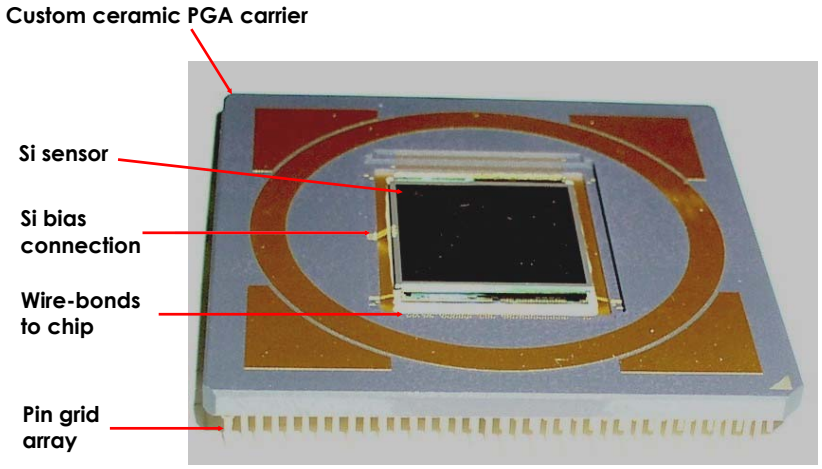

Figure 8.2: Picture of the LHCbpix1chip bump-bonded to a sensor and placed in the ceramic PGA carrier.

Another main change is a better power and signal distribution. This was not allowed in the previous chip due to the tight physical constraints imposed by the ALICE geometry. By means of these changes, functionality at 40 MHz was reached. The chance to resubmit the chip gave us the possibility to correct some other minor problems. In particular, the on-chip pulser, described in section 7.6.6, underwent a major change. The working principle is unchanged, and was explained in section 7.6.6. The redesign reduced drastically the R-C components associated with the input lines Test_low and Test_high and with the lines carrying the analogue test signal to the pixels. Moreover the pulser was replicated 32 times, under each column. The simulations show, and experimental results confirm, that a much better uniformity of the test pulse all over the chip is obtained.

A test pulser that is not uniform across the chip gives serious problems during the testing phase, because the real amount of charge injected in the pixel preamplifier is not precisely

known. This effect is shown in Figure 8.3 (all the pictures were taken with a Tektronix TDS784D scope). Four special test pixels are equipped with output buffers placed at the most important nodes of the front-end (as explained in section 7.6.6.2), so that we can monitor the pixel internal analogue signals. It can be seen that the response changes drastically from the test pixel 1, under column 1, to the test pixel 4, under column 25. The results are that, for a given pulse, less charge is injected into the front-end, the pulse gets slower and lower along the chip and a long tail is present in the output signal. This makes it very difficult to calibrate the preamplifier gain and to extract/adjust threshold and noise. For this reason, all data in the next sections which are based on the calibration of the ALICE1LHCb chip are always very approximate.

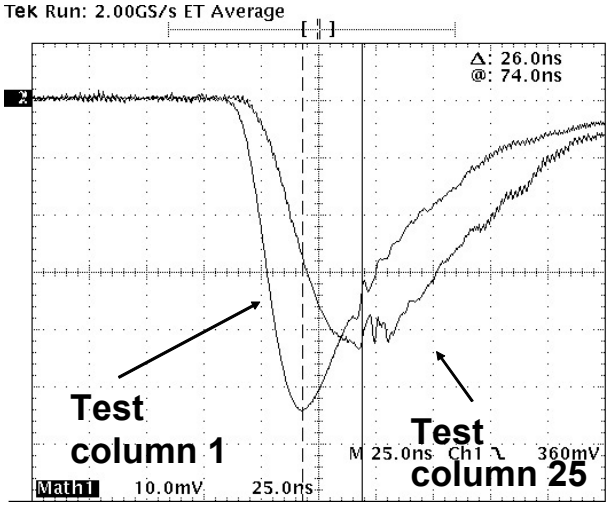

Figure 8.3: Second shaping stage differential output (ALICE1LHCb) for two test pixels (under column 1 and column 25), for the same input test pulse: the on-chip pulser injects a different amount of charge in the pixels, and with a different timing.

The same test, done on the LHCbpix1 chip gives the results shown in Figure 8.4.

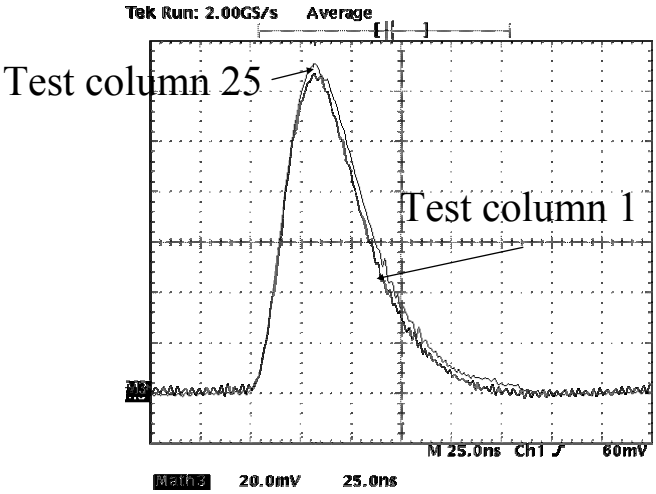

Figure 8.4: Second shaping stage differential output (LHCbpix1) for two test pixels (under column 1 and column 25), for the same input test pulse: the on-chip pulser generates much more uniform responses.

The two output pulses are almost coincident, and the slight difference is within the channel-to-channel variation. No modifications were made to the analogue front-end during the redesign of the LHCbpix1 chip. For this reason[1], many results presented in the next sections are relative to the LHCbpix1chip.

8.2 The DACs

As explained in section 7.5.2, the chip is equipped with 42 DACs for setting the bias voltages and currents. An output stage allows reading out the internal values generated by the DACs.

The behaviour of some of the DACs was not satisfactory, because their output was saturating for low (for current DACs) or high (for voltage DACs) values of the DAC code. After an in depth analysis of the problem, it was found that one of the operational amplifiers used in the DAC blocks was going out of its working region in some cases, generating the saturation effect. For most of the DACs the problem was solved (or greatly attenuated) increasing the power supply voltage of the chip, that was then raised to 1.8 V. This effect is evident in Figure 8.5, which shows the scan of the current DAC eu_vbn for different values of the power supply V_{dda}.

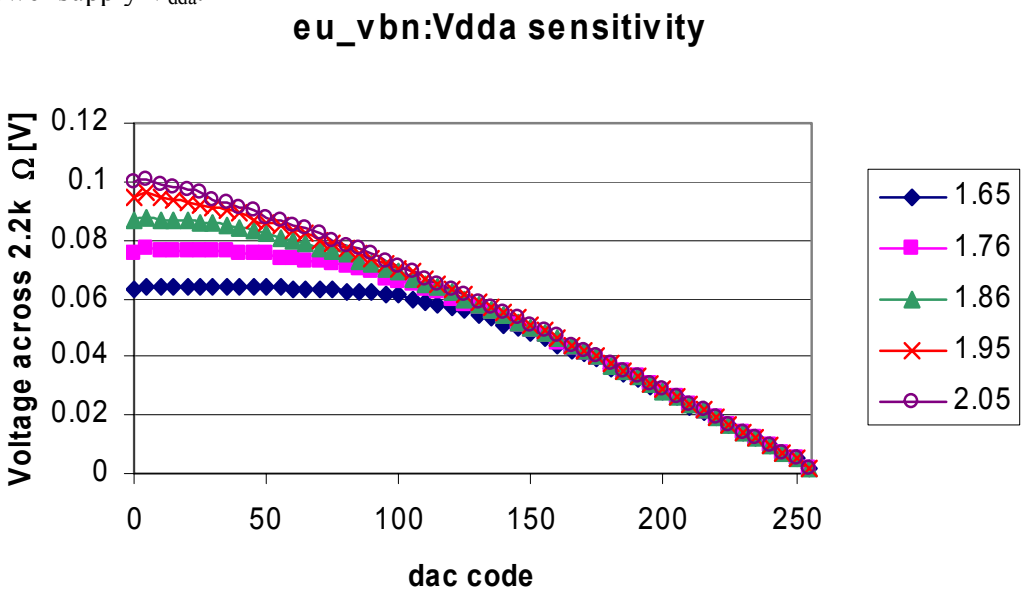

Figure 8.5: Scan of a current DAC (eu_vbn) for different values of the power supply V_{dda}: the saturation effect can be attenuated increasing V_{dda} (the current output is read across a 2.2 kΩ resistor).

However, some of the DACs did not recover after the increase of the power supply, or were still showing saturation effects. The problem was traced back not to be in the DAC

[1] And for the easier availability of the LHCb test system. Moreover, the test pads are not available in a bump-bonded ALICE assembly, while they are available in a LHCb bump bonded assembly.

itself, but in the sensing circuit, as shown in Figure 8.6. The outputs of all the DAC sensing stages are connected to the sense pad by means of a pass transistor gate. The problem is present when the DAC which has to be sensed tries to set a value on the sense pad (1 V in the example) and one of the other DACs has an output values which is lower by more than ~0.6 V (0.4 V in the example). In effect, due to the fact that the well is connected with the source, the drain to bulk diode of the circled transistor in the figure becomes positively biased and starts to draw current. This current should be provided by the output stage of the DAC under sensing, which does not have enough driving capability, so it saturates.

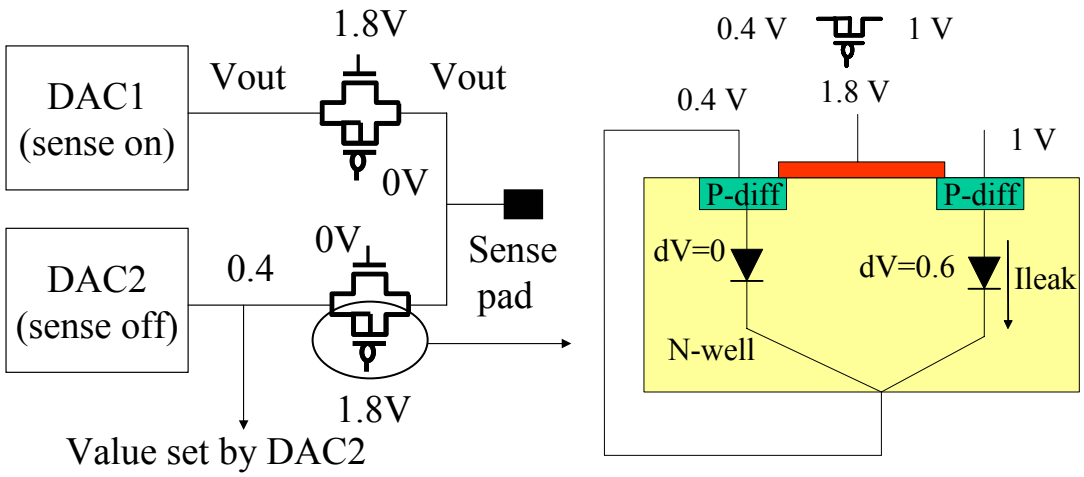

Figure 8.6: Scheme of the connection of the sensing stage of the DACs (left; only two DACs are shown) and schematic representation of the switch transistor which generates the leakage current.

The problem, which does not influence the pixel array functionality, was solved in the LHCbpix1 chip connecting the well to the positive power supply and adding a buffer at the output of each voltage DAC. Moreover, the problem can be cured with a proper DAC sensing procedure, consisting of setting all the DACs which are not scanned at their maximum output value to avoid leakage.

8.3 Calibration of the test structures

The technique used to set the global threshold in the chip was explained in section 7.3.4. To set the chip global threshold a different DAC was designed (Pre_Vth), to be more robust against temperature and process variations, and against power supply noise. The input digital code corresponds to a value (in millivolts) of the imbalance set at the output of the second shaping stage, and this corresponds to an equivalent threshold in electrons at the input of the preamplifier. A calibration of the threshold consists of extracting the threshold in electrons at the preamplifier input corresponding to each DAC code.

To perform the calibration a value of the input DAC code Pre_vth is set, and then the corresponding s-curve (explained in section 6.3.1) is plotted. From the s-curve we can extract the threshold value (and the noise). To do this we use the injection test structures, so we have to know which is the voltage step V_{in} pulsed across the injection capacitor, and the value C_{inj} of the injection capacitor, to calibrate the input charge injected in the preamplifier, $Q_{in} = C_{inj} \cdot V_{in}$. The design value of C_{inj} is 16 fF, but according to the technology design manual this value can change (in principle) by a large amount (10-20%). In practise a much better precision of the values of the capacitors in ICs can be realised with our target technology. We will then assume for the following that $V_{in} = 1$ mV corresponds to $Q_{in} = 100\ e^-$ for what concerns the tests done with the LHCbpix1 chip. Due to the pulser effect, is not so easy to estimate the voltage/charge conversion factor for the ALICE1LHCb chip; the average chip value has been estimated to be about 0.5-0.65 mV/100 e^-.

Figure 8.7 plots the calibration curve for the LHCbpix1 chip used for the lab (bare chip[2]). The data points are plotted, as well as a linear fit (solid curve).

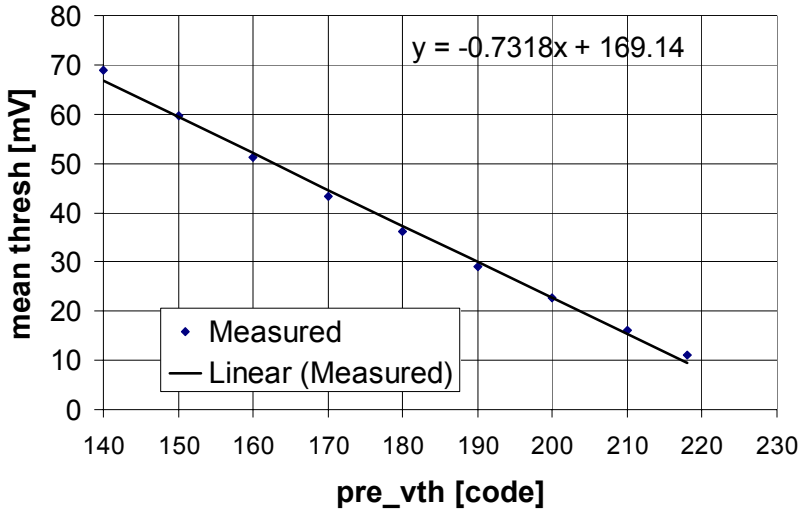

Figure 8.7: Calibration curve for the LHCbpix1 chip under test. The data points are plotted, as well as a linear fit (solid curve). The chip mean threshold is plotted as a function of the code of the DAC which sets the global threshold (Pre_vth).

8.4 The bare chip

The software and the hardware used to test the chip are based on the ALICE On-detector pixel Pilot System (OPS) described in Chapter 3, with modifications to adapt them to the LHCb scenario. In particular the carrier board has been modified to accept a pin array

[2] With bare chip we mean a chip which is not bump-bonded to a detector.

package, which will be the final packaging of the LHCb HPD. A picture of the first version of the carrier board, without the socket for the pin array package is shown in Figure 8.8.

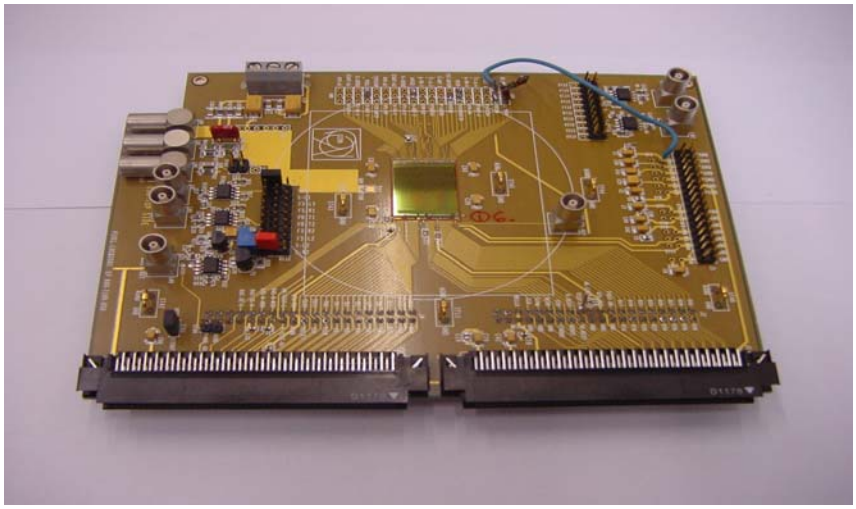

Figure 8.8: Modified carrier board for the LHCb test system (prototype version without the pin array).

8.4.1 The front-end transient analysis

The next pictures present the test outputs of the pixel under column 1 (LHCbpix1). All of them were taken using the test features of a bare (not bump-bonded) chip. The scope was in averaging mode and the background noise was subtracted if possible. This was done pulsing the chip with a test pulse of amplitude 0 V, and subtracting this output from the output obtained while pulsing with a normal input pulse. The reason for this is that the injection logic creates coherent noise which is not present when the chip is in data-taking mode, while sensing real particles. Moreover additional noise is generated by the test setup, and is partially removed with this technique.

Figure 8.9 shows the output of the preamplifier for an input charge Q_{in} = 5000 e^-; the peaking time is about 17 ns and the peak is at ~36 mV. The peaking time of this stage seems much higher than what was found in Chapter 7 (about 7-8 ns), and the peak value much lower. However, several things should be pointed out. The first is that the pulse generated by the on-chip pulser degrades the peaking time of the system. The second is that to carry the analogue signals off-chip some test buffers are used to pick up the signal at the test points. This has the double effect of loading the test point (care was taken to reduce the input capacitance of the buffer, but the output of the preamplifier is a sensitive node) and adding a shaping effect (the effect of the output buffer was shown in section 7.6.6.2). In any case, the full system is not very sensitive to the variations in the rise time of the preamplifier (as can be deduced by table 7.4), so that the other stages will show much less degradation in the peaking time and in the peak value. The last point concerns the test board. An opamp is added on the

test board to drive the oscilloscope input, and this also generates a shaping effect on fast signals. Let us point out here that all the simulations presented in Chapter 7 were done supposing a detector capacitance of 100 fF, while a rough estimation of the parasitic capacitance of a bare readout channel is ~25 fF.

The effect of a lower detector capacitance is not major, and can be summarized in a reduction of the peaking time of a few nanoseconds in all the stage outputs. The gain is increased of a few percent at the preamplifier output, almost unchanged at the first shaping stage differential output and reduced of a few percent at the second shaping stage differential output.

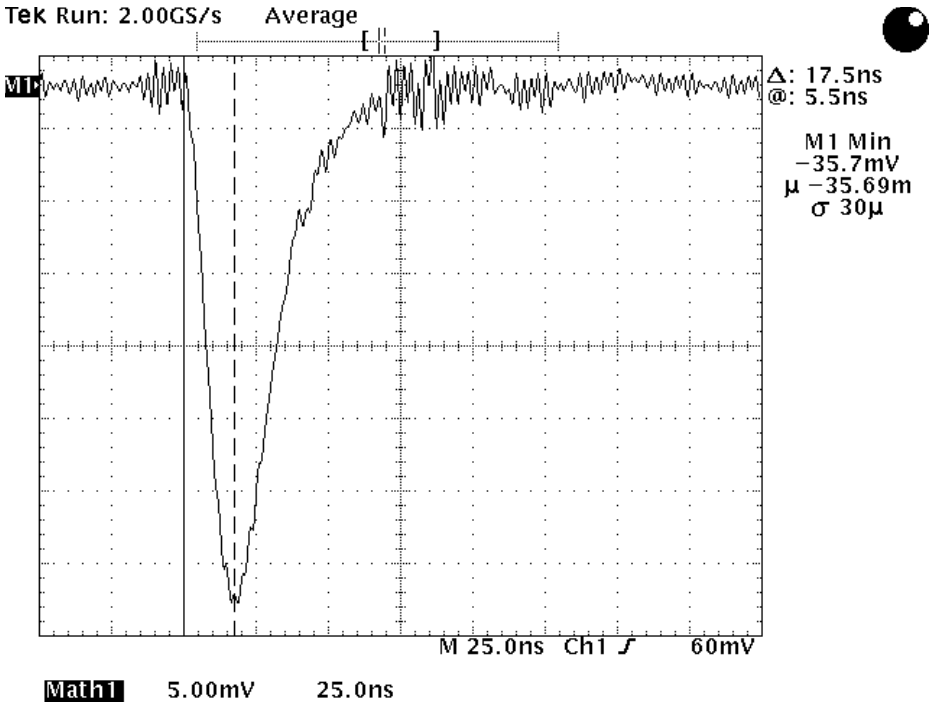

Figure 8.9: Output of the preamplifier of the test pixel under column 1 for an input signal Qin = 5000 e^-.

Figure 8.10 shows the output of the first shaping stage, left branch (left) and right branch (right) for an input signal of Q_{in} = 5000 e^-.

If we compare these pictures with Figure 8.11, that shows the same outputs but without background noise subtraction, we can see that the improvement is remarkable.

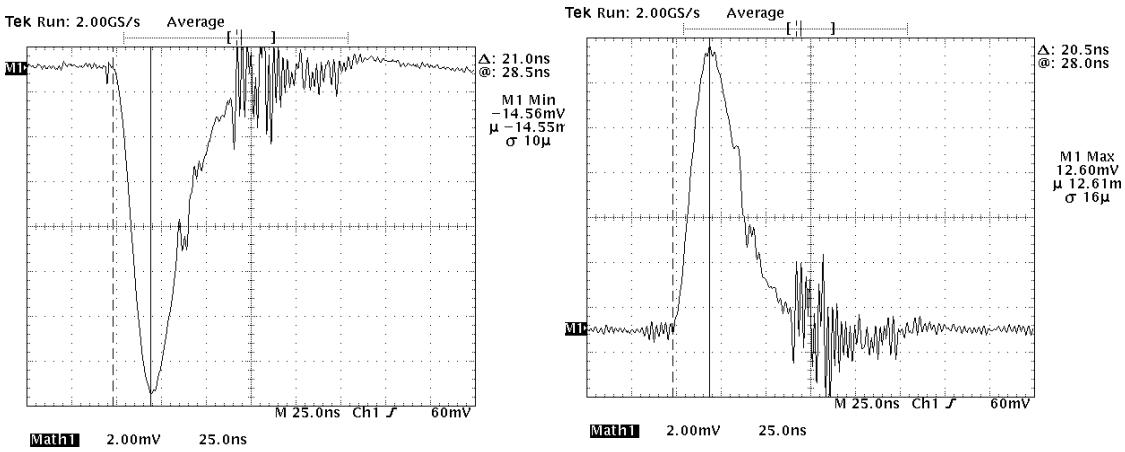

Figure 8.10: Output of the first shaping stage, left branch (left) and right branch (right) for an input signal Q_{in} = 5000 e^-.

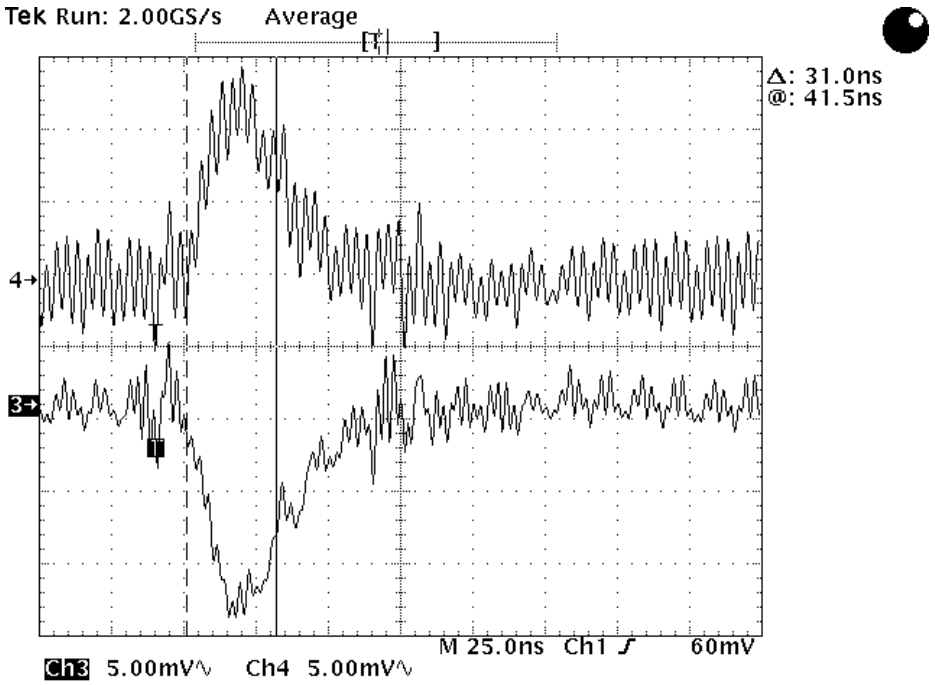

Figure 8.11: Output of the first shaping stage, left branch (left) and right branch (right) for an input signal Q_{in} = 5000 e^-, without background noise subtraction.

Figure 8.12 shows the differential output of the first shaping stage for an input signal Q_{in} = 5000 e^-. The peaking time is about 21 ns and the peak is about 28 mV.

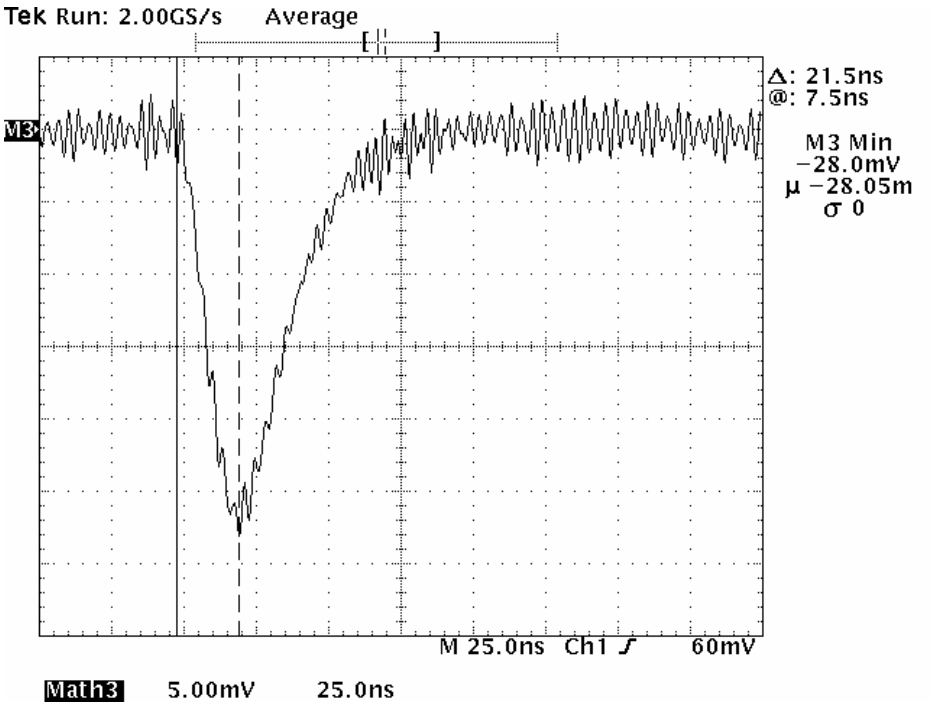

Figure 8.12: Differential output of the first shaping stage for an input signal Q_{in} = 5000 e^-. The peaking time is about 21 ns and the peak is about 28 mV.

Figure 8.13 shows the outputs of the second shaping stage for the same input signal, left branch (left) and right branch (right).

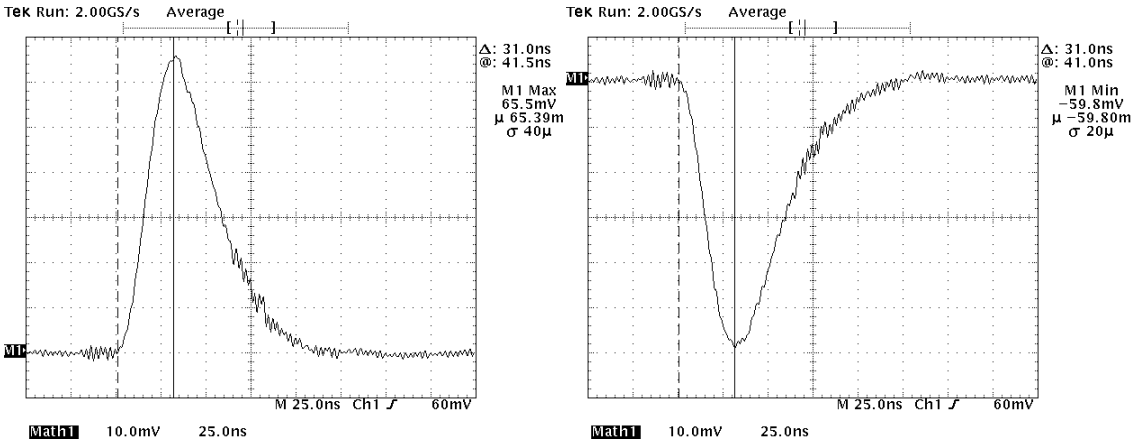

Figure 8.13: Outputs of the second shaping stage, left branch (left) and right branch (right) for an input signal Q_{in} = 5000 e^-.

Figure 8.14 shows the differential output of the second shaping stage for an input signal Q_{in} = 5000 e^-. The peaking time is about 31 ns and the peak is about 125 mV. This is consistent with the simulated values of 144 mV for the peak and ~26 ns for the peaking time, if taking into account the effect of the pulser and of the buffers.

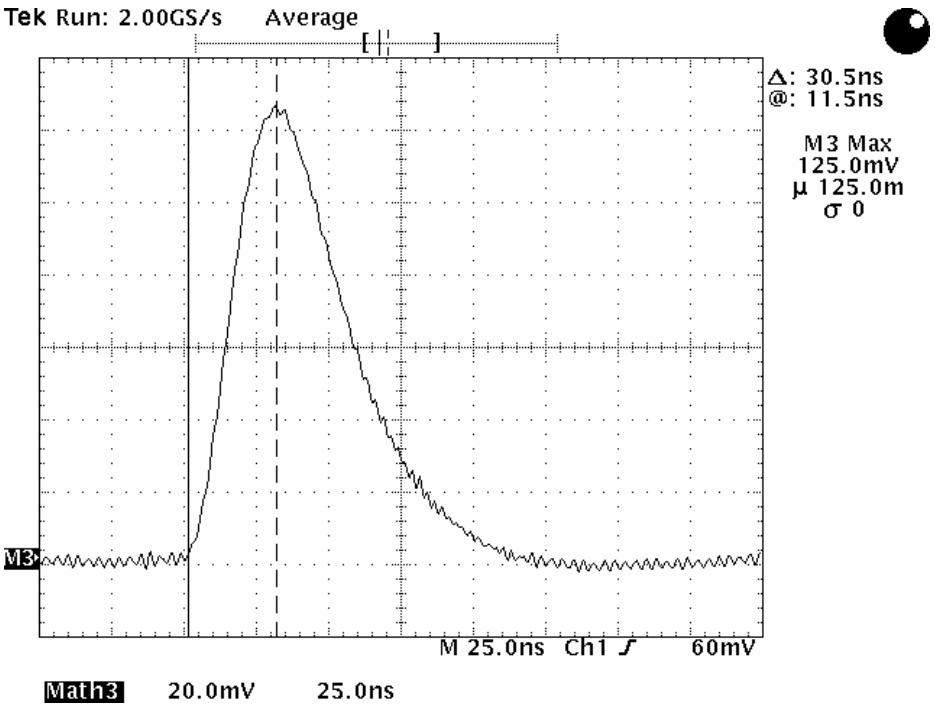

Figure 8.14: Differential output of the second shaping stage for an input signal of Q_{in} = 5000 e^-. The peaking time is about 31 ns and the peak is about 125 mV.

Figure 8.15 shows the discriminator output pulse for an input signal of Q_{in} = 5000 e^-; the time width is 90 ns.

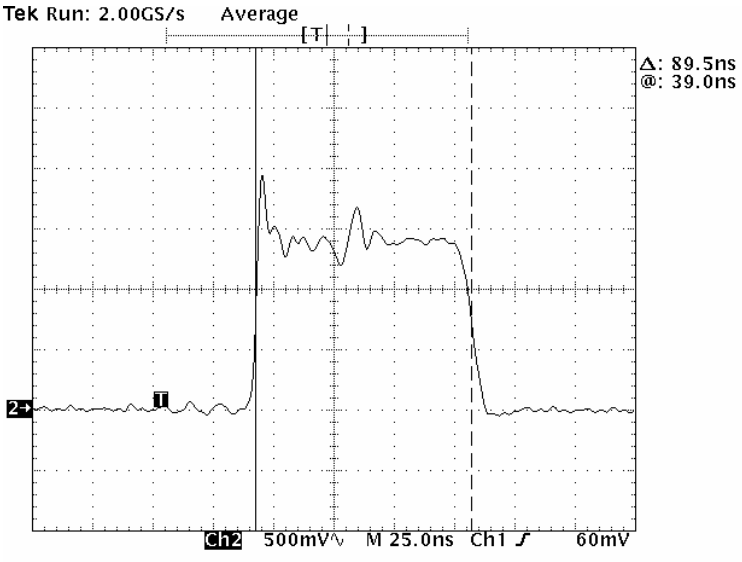

Figure 8.15: Discriminator output pulse for an input signal of Q_{in} = 5000 e^-. The time width is 90 ns.

The second shaping stage differential output for the typical ALICE signal (17000 e^-) is shown in Figure 8.16. The peaking time is about 30 ns and the peak about 260 mV. The small

undershoot is due to the nonlinearity of the system. This is again reasonably consistent with the simulated peaking time at about 25.5 ns and a peak of about 302 mV.

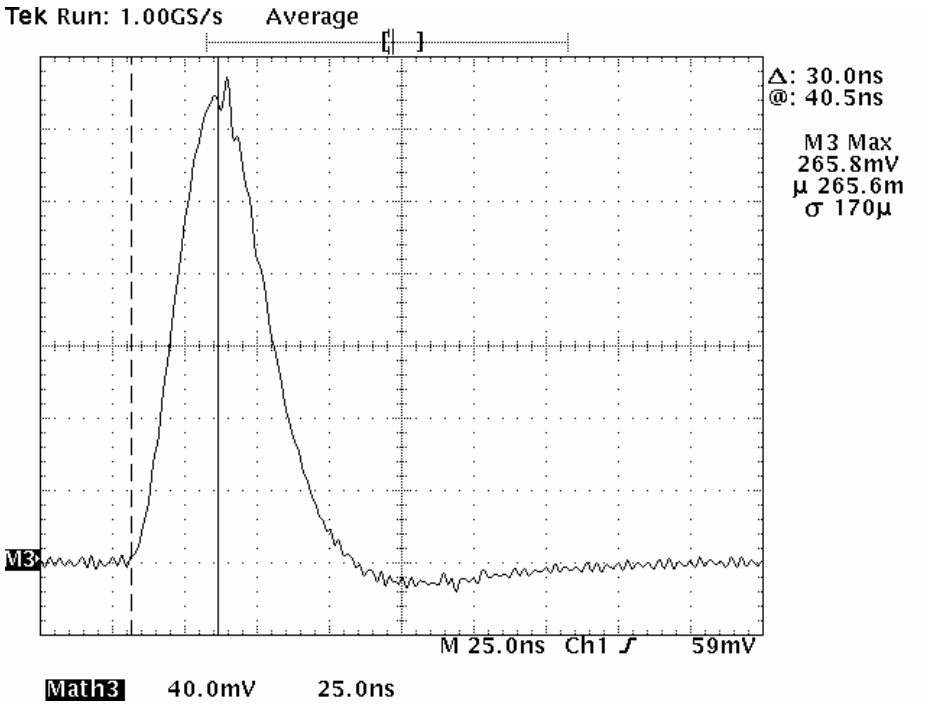

Figure 8.16: Second shaping stage differential output for the typical ALICE signal (17000 e^-). The peaking time is about 30 ns and the peak about 260 mV.

8.4.2 Electrical crosstalk

The phenomenon of crosstalk in pixel detectors, already explained in section 7.3.7.4, is mainly due to capacitive cross-coupling of the sensor cells. However, electrical coupling through the power lines or through the substrate can induce a hit in neighbouring pixels. To check if this was the case, the pixel in row 1 column 1 has been pulsed with an increasing input voltage (LHCbpix1). For a threshold of about 1800 e^- no hit in the neighbouring pixels is detected up to the maximum injected voltage of 1.6 V (corresponding to about 160,000 e^-).

8.4.3 Noise measurements

If we extract, with the procedure of the s-curve, the noise of each pixel in the chip matrix, we can plot the distribution of the chip noise. This distribution can be fitted with a Gaussian to extract the mean value. The noise distribution and the Gaussian fit for the chip under test are presented in Figure 8.17. The mean is 1.355 mV and the dispersion 0.13 mV, which corresponds to 135 e^- rms and 13 e^- rms.

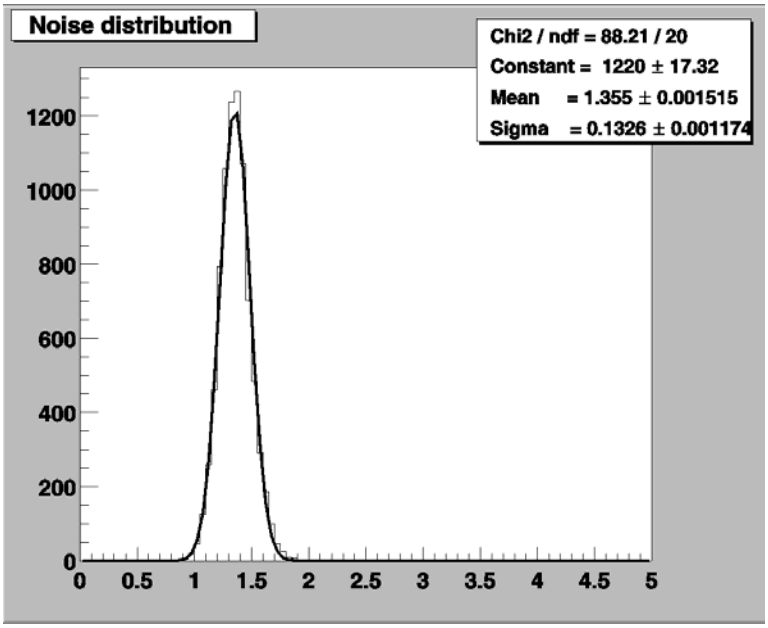

Figure 8.17: Noise distribution and Gaussian fit for the chip under test; the mean is 1.355 mV and the dispersion 0.13 mV (1 mV = 100 e^-, Pre_Vth = 217).

Figure 8.18 shows the chip noise map. It is a two-dimensional plot of the chip noise, where the noise in each pixel is represented by a different colour. The noise is uniform in the chip, and no systematic effect is visible.

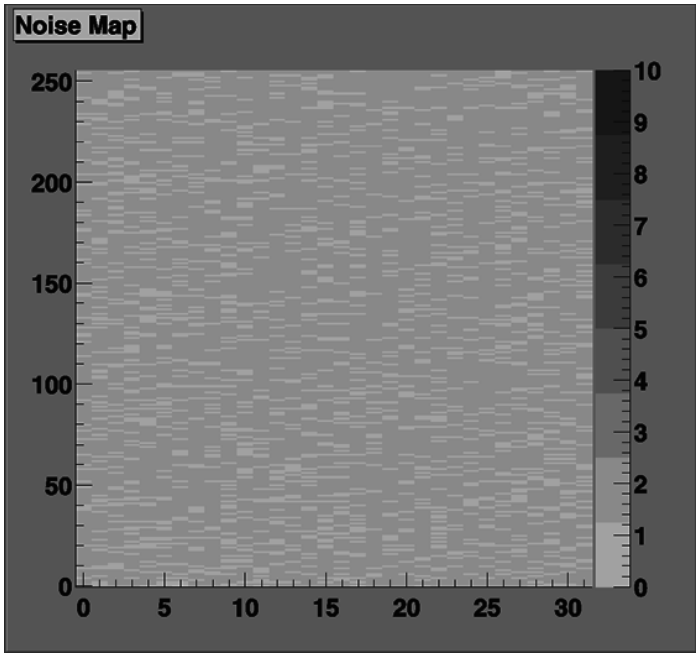

Figure 8.18: Chip noise map. The noise in each pixel is represented by a different colour. The noise is uniform in the chip, and no systematic effect is visible.

8.4.4 Recovery after a huge signal

Figure 8.19 shows how the system behaves to a huge input pulse (second shaping stage differential output, Q_{in} = ~100,000 e^-). It can be noticed that the system is back to zero in less than 300 ns, and that the discriminator output is back to zero in less than 150 ns. The behaviour is perfectly consistent with simulations (see section 7.3.7.5) and within the specifications for both experiments.

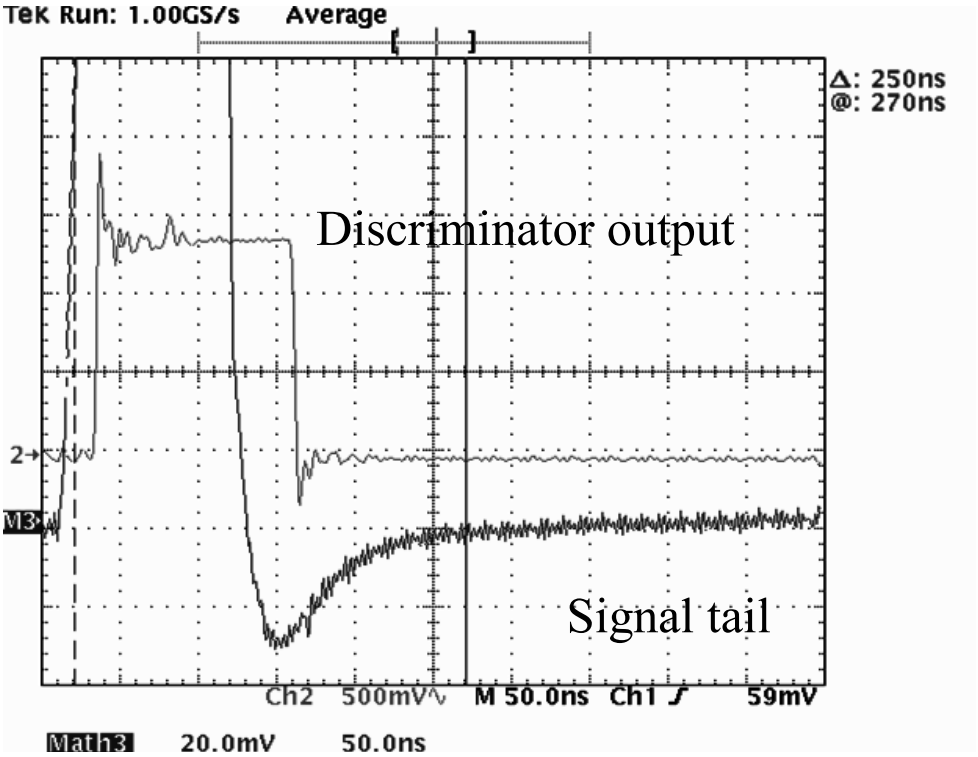

Figure 8.19: Second shaping stage differential output, Q_{in} = ~100,000 e^-

8.4.5 Threshold measurements

8.4.5.1 Threshold distribution and threshold spread

We can measure the average chip threshold (for a given DAC threshold setting) similarly to what was done for the noise. With the procedure of the s-curve we can extract the threshold of each pixel in the chip matrix, we can plot the distribution of the average noise in the chip and then fit it with a Gaussian distribution to extract the mean value and the dispersion σ. The threshold distribution and the Gaussian fit for the chip under test are presented in Figure 8.20. The mean is 11.25 mV and the dispersion 0.91 mV, which correspond to 1125 e^- and 91 e^-.

Figure 8.21 shows the chip threshold map. Again, it is a two-dimensional plot of the chip threshold, where the threshold in each pixel is represented by a different colour. No relevant systematic effect is present.

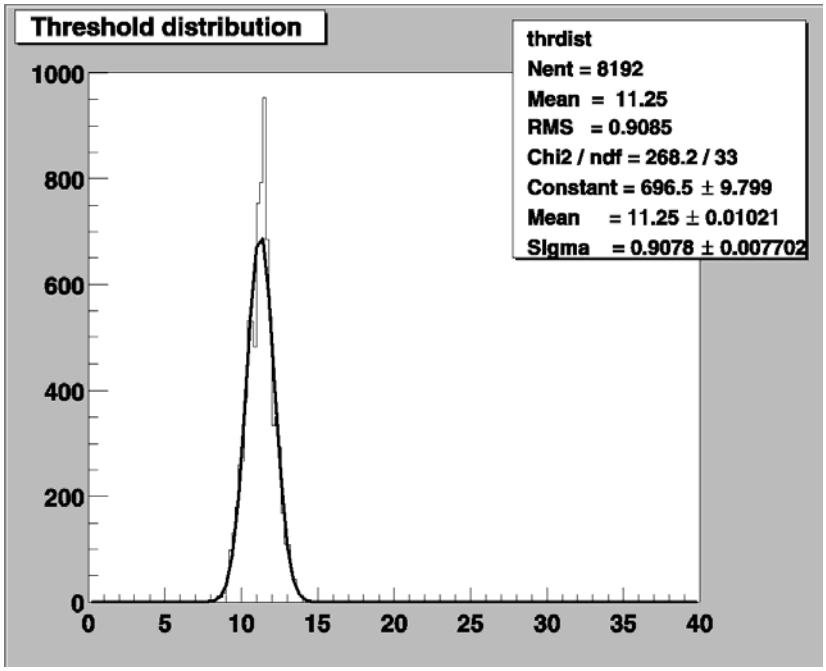

Figure 8.20: Threshold distribution and Gaussian fit for the chip under test; the mean is 11.25 mV and the dispersion 0.91 mV (1 mV = 100 e^-).

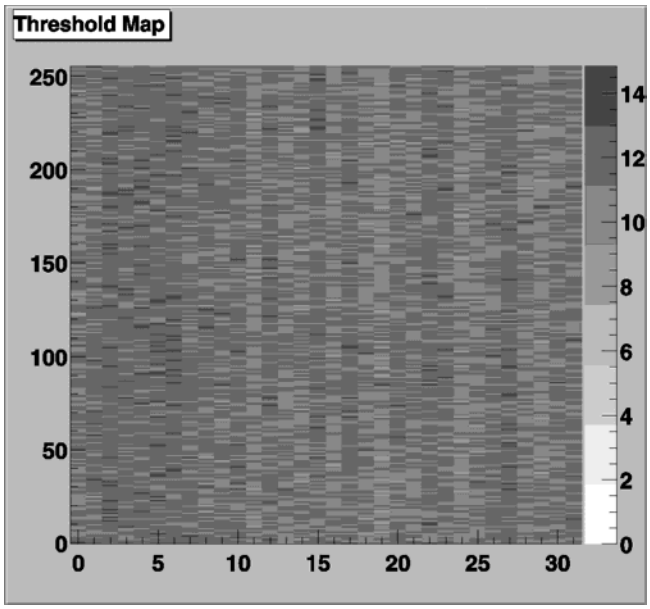

Figure 8.21: Chip threshold map. The threshold in each pixel is represented by a different colour. No relevant systematic effect is present.

8.4.5.2 Effect of the mismatch of the DAC Pre_Vth on the threshold

The calibration for three LHCbpix1 chips (clocked at 10 MHz) is shown in Figure 8.22.

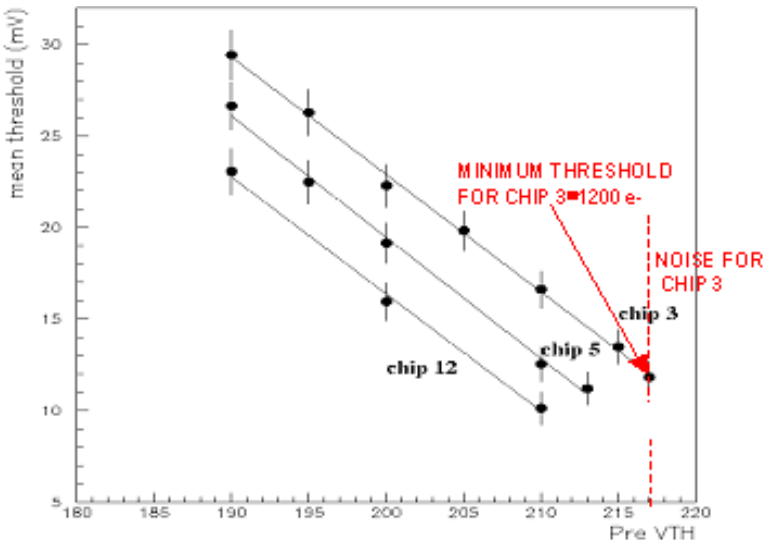

Figure 8.22: Relation between the Pre_Vth DAC code (used to set the global threshold) and the input mean threshold (in mV, 1 mV = 100 e^-) [Cas03].

The calibration curves are slightly different; this effect is due both to the offset at the output of the front-end, and to the mismatch in the output curves of the Pre_Vth DACs for the three chips. An example of this effect is shown in Figure 8.23.

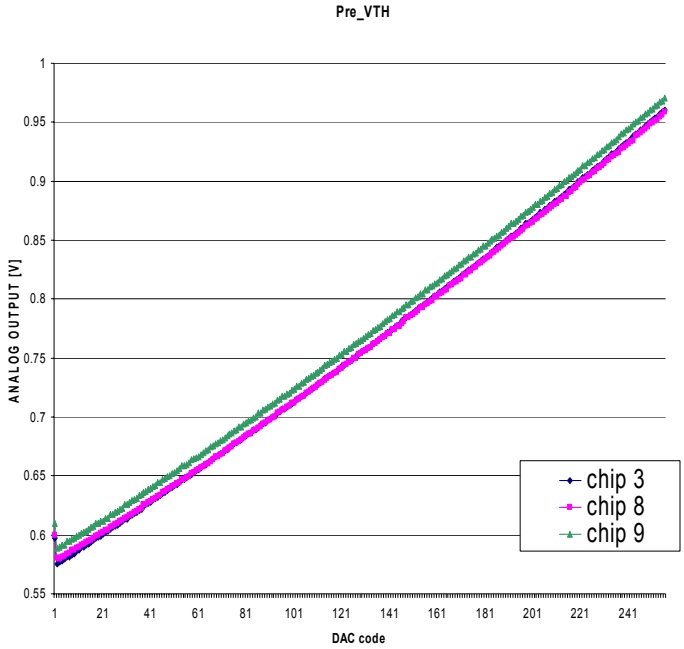

Figure 8.23: Example of Pre_Vth DAC (used to set the global threshold) scan for three different LHCbpix1 chips. The different behaviour is one of the causes of the different calibration curves shown in Figure 8.22 (the three points at code = 0 are an artefact of the test system) [Cas03].

8.4.5.3 Effect of the clock frequency on the minimum threshold

Plotting the curves from the same chips and clocking the system at 40 MHz reduces the minimum threshold for all the chips by 200-300 e^-, while the noise does not change significantly. A possible explanation for this phenomenon is that the front-end has an AC response which attenuates a signal at 40 MHz 20 times more than a signal at 10 MHz (see Figure 7.48), so the digital crosstalk is reduced by the front-end shaping. The noise measurement is not affected because the digital noise is a coherent signal, while the s-curve method extracts the pixel incoherent (white) noise.

8.4.6 Threshold adjust

As we have explained in section 7.3.8, the global threshold $V_{thglobal}$ of the chip is set by the Pre_Vth DAC, but the threshold of each pixel can be fine-tuned by means of a three-bit register (TH) and a DAC (Dis_biasth). The DAC sets the minimum threshold step ΔV_{th}, so that changing the binary code TH by 1 corresponds to a threshold shift $-\Delta V_{th}$.

Supposing a linear system we can write:

$$\text{Pixel Threshold} = V_{thglobal} - \Delta V_{th} \cdot TH \qquad (8.1)$$

The algorithm used to minimise the threshold dispersion, i.e. to determine the value of VTH for each pixel which minimises the standard deviation of the distribution of the pixel thresholds on the chip, was written by S. Jolly [Jol02] and is based on the assumption that the default value of the DAC Dis_biasth is optimal, so that it does not need to be further optimised. This assumption comes from the fact that the lower end of the chip threshold distribution for TH = 0 coincides with the upper end of the distribution for TH = 7. This means that the adjustment range is equal to the distribution width.

A first simple algorithm based on the previous assumptions is the following:

- Take the TH = 0 threshold distribution and find its width, w;
- Divide the width by 8: w/8 is the distance that the threshold of a pixel will move if its TH is increased by 1;
- For each pixel, examine its TH = 0 value and find out how far it is from the "target value";
- Divide this distance by w/8 to obtain the pixel's new value of TH.

The results of this algorithm are shown in Figure 8.24. Data written in the picture are relative to the final distribution, which has a sigma of 43 electrons rms (the starting value for the two distributions was $\sigma = 92\ e^-$ rms and $\mu = 1324\ e^-$ for TH = 7; $\sigma = 104\ e^-$ rms and $\mu = 1886\ e^-$ for TH = 0. A more sophisticated algorithm was attempted to reduce the dispersion of the final distribution. In this case the pixel threshold for every pixel and every value of TH was calculated, and the value of TH closest to the "target value" was picked up.

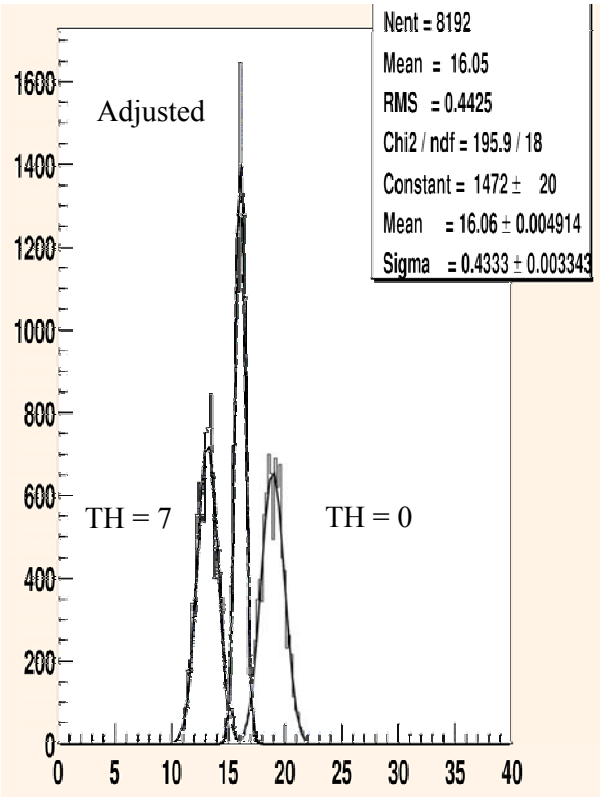

Figure 8.24: Results of the threshold adjust using the simple algorithm. Data written in the picture are relative to the final distribution, which has a sigma of 43 electrons rms.

This approach has the advantage of being completely independent of TH linearity (which is not very good at pixel level). The disadvantage is that it requires eight times as much data (i.e. the threshold scans for all values of TH instead of just TH = 0). Moreover, with this algorithm the final value of the threshold dispersion is 41 e^-, so the improvement is marginal. The reason why only a factor 2.5 of reduction in the threshold dispersion can be obtained is due to the value of ΔV_{th} which is 80 e^-, so that a final dispersion better than $\Delta V_{th} / 2 = 40$ e^- is not possible. This value was supposed to be optimal, but a new algorithm is under investigation, which would try to optimise also ΔV_{th}, allowing a finer step for the threshold adjust and so a narrower distribution.

It should be pointed out that the unadjasted dispersion of 100 e^- rms already meets the specifications of both experiments.

8.4.7 Wafer probing

A first test for the selection of the Known Good Dies (KGDs) is done directly at wafer level, with a semiautomatic procedure. A Karl-Suss PA200 probe station (shown in Figure 8.25) equipped with an 8” (200 mm) chuck is used to probe wafers received from the foundry. On this prober, if equipped with the correct needle card and chuck, it is also possible to test single chips, assemblies and ladders.

Each wafer contains 86 ALICE1LHCb chips (15.8 × 13.5 mm^2 each). A standard testing procedure has been established for the selection of ALICE KGDs, which classifies tested chips in three groups [Rie02].

- Class III chips

These are chips which show a major defect. The chips which are dead, which do not have digital output, which fail a DAC scan or which have a JTAG or a configuration registers error are placed in this class. These chips are discarded.

- Class II chips

These are chip with minor defects. They work, but have either some missing columns, or too many defective (or excessively noisy) pixels, or a minimum threshold which is too high, or have excessively low or high power supply currents (either digital or analogue). These chips can be used for testing purposes, but are not suitable for bump-bonding.

- Class I chips.

These are the chips which meet all the requirements, and can be bump-bonded to a detector: DACs, JTAG, digital out of the chip are functional; the minimum mean threshold is less 30 mV (this should correspond for ALICE to about 2000 e^-), there are less than 1% of defective (or excessively noisy) pixels, the analogue current is less than 350 mA and the digital current is less than 250 mA.

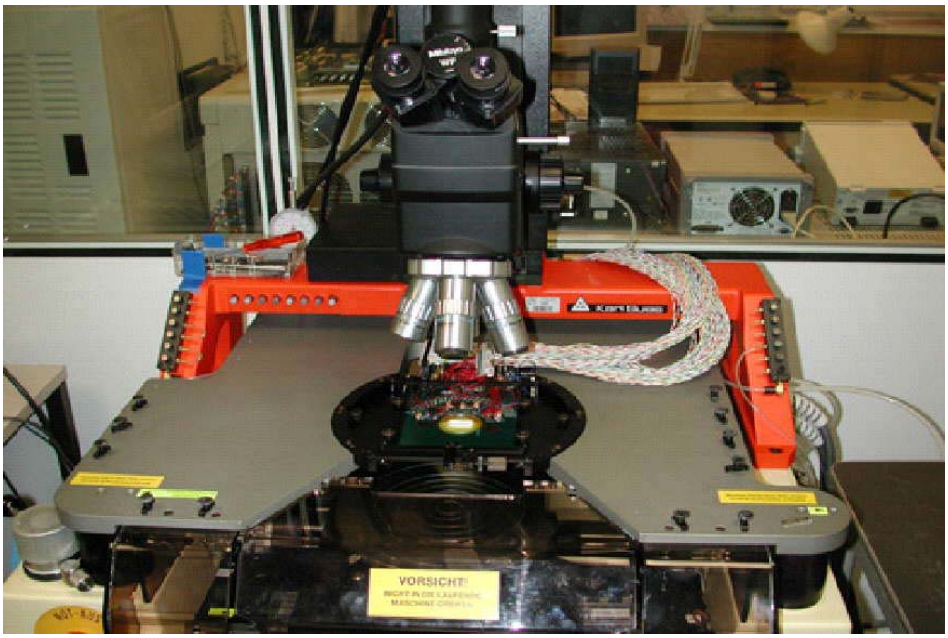

Figure 8.25: Photograph of the Karl-Suss PA200 probe station used to test chips at wafer level, but which can also be used to test single chips, assemblies and ladders.

The yield of Class 1 chips changes from wafer to wafer and lot to lot from 35% to 75%, with an average yield of about 50-55%. On the 4 wafers delivered in 2001 we found an average of 55% of class I chips, 15% of class II chips and 30% of class III chips. Up to now more than 16 wafers were tested, both 750 μm thick and 300 μm thick. The test procedure used by LHCb to select KGDs is very similar; results on 7 wafers with 71 chips each give 43% of class I chips, 15% of class II chips and 42% of class III chips [New03a].

A picture of a chip wafer with an example of the distribution of class I, II and III chips is shown in Figure 8.26.

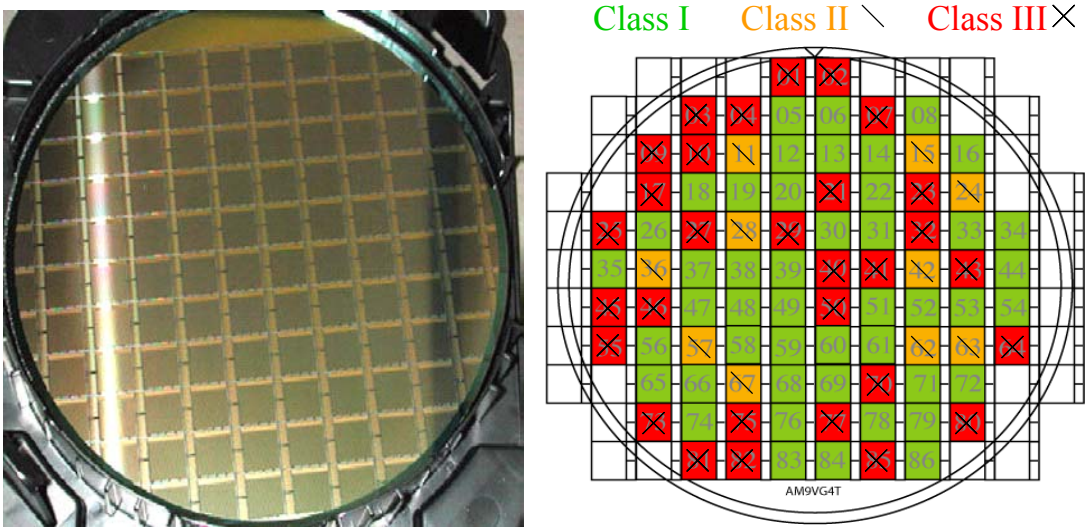

Figure 8.26: Picture of a chip wafer (left) with an example of the distribution of class I, II and III chips [Rie02].

8.4.8 X-ray irradiations

Some tests on the LHCbpix1 chips were carried out at the CERN/MIC group X-ray irradiation system [Cas03].

The system [Web10] is composed of an X-ray machine (Seifert RP149), a 6/8 inches wafer prober (Karl Suss PA200), and a CCD camera with monitor outside the irradiation cabinet. The pressure inside the cabinet can be kept slightly higher than the outside, so that the irradiation cabinet is a clean area. Moreover, the injected air is dried to lower the condensation temperature, below -15°C. The standard chuck of the wafer prober can be replaced by a 6 inches thermal chuck (Digit Concept DCT600) to perform high or low temperature measurements or irradiations. All instruments are controlled by a PC (with Labview) either via GPIB or via RS232 interfaces.

The irradiations were done with one of the two tubes available, the Tungsten (peak at 10 keV). The dose rate was chosen (and calibrated) in the available range (between about 10 and 800 rd/s) to be 10120 rd/min. The maximum area that can be irradiated with the X-ray beam is not enough to irradiate the whole chip, so the irradiation procedure was carried out for two different areas (of about 6.5 mm × 6.5 mm, as shown in Figure 8.27): one including only the array of pixels and another one with the DACs. The irradiation was done in steps of 10, 50, 200, 500, 1000 and 10000 krd. Data were recorded at the end of each irradiation step.

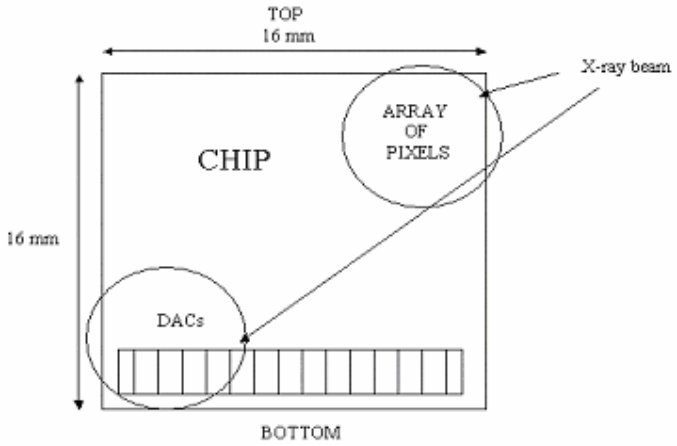

Figure 8.27: The irradiation procedure was carried out for two different areas in the chip, one including only the array of pixels and another one with some DACs.

Table 8.1 shows the measurement results for the irradiation of the pixel array only. After each irradiation and before the data acquisition, a check was made to find out the minimum threshold voltage. For all the total doses, with the exception of the 10 Mrd one, the minimum threshold voltage corresponds to a value of the DAC which sets the global threshold Pre_VTH = 216. In the case of the 10 Mrd dose the threshold voltage had to be increased setting Pre_VTH = 213.

TOTAL DOSE	Pre_VTH (minimum threshold voltage)	THRESHOLD DISTRIBUTION (Full Matrix)		NOISE DISTRIBUTION (Full Matrix)		THRESHOLD DISTRIBUTION (Irradiated pixels)		NOISE DISTRIBUTION (Irradiated pixels)	
		MEAN	RMS	MEAN	RMS	MEAN	RMS	MEAN	RMS
10 krd	216	9.52	0.886	1.284	0.144	9.45	0.8839	1.3	0.1394
50 krd	216	9.52	0.8915	1.286	0.1456	9.383	0.8927	1.302	0.1378
200 krd	216	9.575	0.9159	1.29	0.1438	9.304	0.9143	1.301	0.1398
500 krd	216	9.609	0.9334	1.282	0.145	9.219	0.9464	1.296	0.1427
1 Mrd	216	9.667	0.9453	1.286	0.1609	9.198	0.9392	1.296	0.1674
10 Mrd	213	11.44	1.008	1.278	0.1632	10.65	0.9435	1.287	0.1768

Table 8.1: Measurement results for the irradiation of the pixel array only, analyzing data for the full matrix (columns 3 to 6) and for the irradiated pixels only (columns 7 to 10). The pre-irradiation value of Pre_VTH is 216 [Cas03].

Columns 3 to 6 analyze data for the full matrix. Although all mean thresholds are around 9.5 and 9.6 mV we can observe a slight increasing trend in the values which is not significant

up to a total dose of 1 Mrd. With a total dose of 10 Mrd the minimum threshold voltage in the discriminator had to be increased to Pre_VTH = 213 because some pixels became noisy. As a consequence, the mean threshold increased from 9.67 at 1 Mrd to 11.4 at 10 Mrd, which means less than 150 e^-. For the full irradiation the threshold spread and the noise are practically unchanged.

Columns 7 to 10 analyse data for the irradiated pixels only (i.e. pixels included in a square area from column 15 to 31 and from row 0 to 125). Up to 1 Mrd there is only a little shift of the threshold towards lower values (~ 25 e^-); at 10 Mrd the need to decrease Pre_VTH increases the mean threshold to 10.65 mV (i.e. it increased by 120 e^-).

Table 8.2 shows the results of the irradiation of the DACs; analyzing data for the full matrix (columns 3 to 6) and for the irradiated pixels only (columns 7 to 10). For the last two rows the value of Pre_Vth could be increased to 227. The chip threshold is more sensitive to an irradiation of the DAC area, and even more sensitive to a "mismatched" irradiation of DACs and pixels separately. However, once again the threshold distribution shifts of some hundred electrons and can be almost recovered changing the Pre_vth DAC code. The impact on the electronic noise is negligible.

TOTAL DOSE	Pre_VTH (minimum threshold voltage)	THRESHOLD DISTRIBUTION (Full matrix)		NOISE DISTRIBUTION (Full matrix)		THRESHOLD DISTRIBUTION (Irradiated pixels)		NOISE DISTRIBUTION (Irradiated pixels)	
		MEAN	RMS	MEAN	RMS	MEAN	RMS	MEAN	RMS
10 krd	213	11.67	1.007	1.261	0.2214	11.95	0.8516	1.251	0.1364
50 krd	213	12.98	1.066	1.26	0.1901	13.2	0.909	1.285	0.1836
200 krd	213	15.46	1.2	1.273	0.1828	15.54	1.042	1.27	0.1379
500 krd	213	18.08	1.307	1.284	0.1797	17.96	1.184	1.288	0.1877
1 Mrd	227	19.86	1.447	1.293	0.1761	11.06	0.9211	1.279	0.1716
10 Mrd	227	14.2	1.118	1.332	0.2239	13.63	0.9946	1.286	0.1926

Table 8.2: Measurement results for the irradiation of the DACs, analyzing data for the full matrix (columns 3 to 6) and for the irradiated pixels only (columns 7 to 10) [Cas03].

After irradiation the chip was introduced in an oven in order to see if the irradiated areas recovered their normal behaviour after annealing. Two sets of measurements were taken: after an annealing of 24 hours at room temperature and after 1 week of annealing at 100 °C. The results of the two measurements are almost identical. Table 8.3 shows the chip parameters after the second annealing. The Pre_Vth DAC was monitored as well; after the second

annealing it does not recover to the pre-irradiation state, but only to a state which is in between 1 Mrd and 10 Mrd of radiation dose.

ANALYSED AREA	THRESHOLD DISTRIBUTION		NOISE DISTRIBUTION	
	MEAN	RMS	MEAN	RMS
FULL MATRIX	9.983	0.9775	1.286	0.1509
ARRAY OF PIXELS	9.864	0.9677	1.289	0.1327
DACs	9.727	0.916	1.264	0.1556

Table 8.3: Measurement results for the annealing of the DACs, analyzing data for the full matrix, for the pixels in the top right corner only and for the pixels in the DACs region only (Pre_Vth = 232) [Cas03].

Figure 8.28 shows the evolution of the power supply currents with the X-ray dose (pixel array irradiation; DACs irradiation results are identical); no significant change in the currents is observed up to the maximum irradiation dose (10 Mrd). This confirms the effectiveness of using HBD techniques to avoid radiation induced parasitic currents.

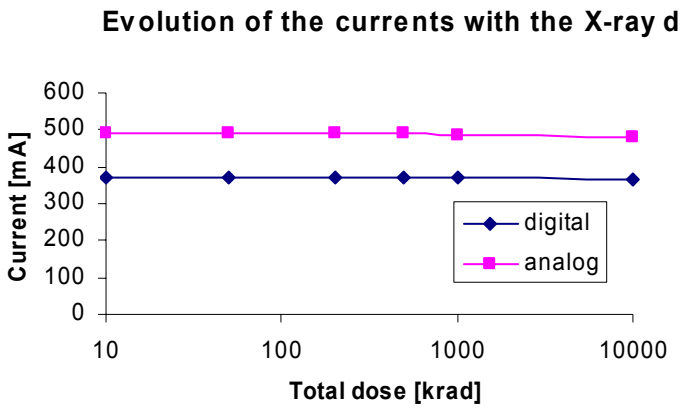

Figure 8.28: Evolution of the power supply currents with the X-ray dose (pixel array irradiation; DACs irradiation results are identical); no change is observed up to the maximum irradiation dose (10 Mrd) [Cas03].

Similar results were found irradiating an ALICE1LHCb chip [Hun01].

Let us point out that the irradiation is carried out at radiation levels 40 times higher than the ALICE foreseen doses. The effects of irradiation up to 500 krd are minor, and very often negligible.

8.4.9 Single event effects sensitivity

A hadron with an energy of several GeV does not deposit enough charge through direct ionisation to create a Single Event Effect (SEE). However, it may interact elastically and inelastically with the atoms in the pixel chip. The recoils and fragments will deposit a large amount of charge in the chip, and may therefore lead to SEEs. Both SEGRs and SELs are generally not observed for circuits designed in 0.25 μm with HBD techniques. For SEGR the electric fields are not high enough, while the implementation of guard rings prevents SEL to occur. During measurements on the ALICE1LHCb chip neither SEGRs nor SELs were observed. SEUs do however occur, so we are interested in measuring the cross-section for a SEU in the memory cells of the chip. The cross-section is determined in two different ways. Firstly heavy ions with a Linear Energy Transfer (LET) between 6 and 120 $MeVmg^{-1}cm^2$ were used, since these deposit a large amount of charge and the probability for SEU is large. From these results the SEU cross-section for other hadrons can be calculated. Secondly the measurements were repeated with 60 MeV protons. The measurements were done at the CRC Facility, Louvain-la-Neuve [Hun01].

8.4.9.1 Measurements with ions

In order to vary the value of the LET, different ions can be chosen (Xe^{26+}, Ar^{8+}, Ne^{4+}, Kr^{17+}). Additionally the chip can be tilted with respect to the propagation direction of the ions to increase the path length of the ion through the sensitive part of the memory cells and therefore increase the amount of deposited charge in this region.

To determine the number of SEUs the chip is loaded with a test pattern. After irradiation for a certain amount of time (several seconds or minutes) the memory cells are read-out and compared with the loaded test pattern. All differences are attributed to SEUs. No errors were detected repeating the test procedure prior to irradiation. The results of the measurements are shown in Figure 8.29.

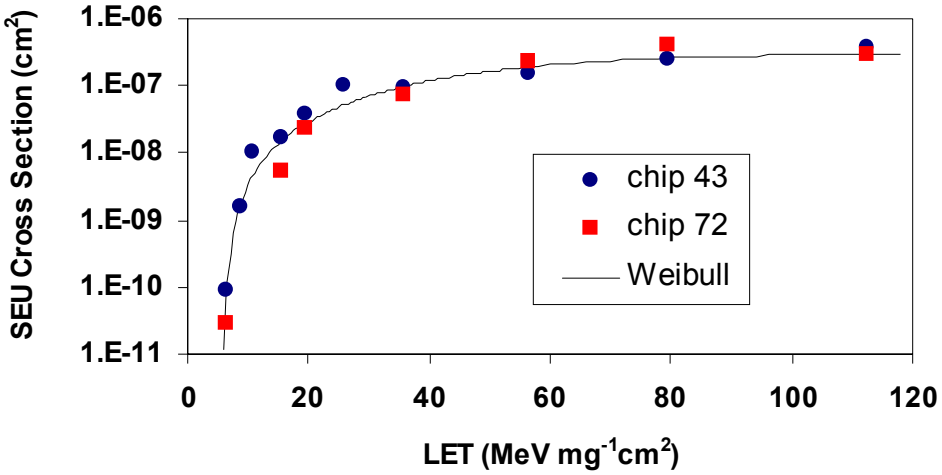

Figure 8.29: SEU cross-section as a function of the LET for 2 different pixel chips. The solid curve represents the Weibull equation and is used for the interpretation of the data [Hun01].

For two pixel chips the SEU cross-section was measured. The results for the two chips are in good agreement. When the LET is larger than 6.3 $MeVmg^{-1}cm^2$, enough charge is deposited to create a SEU. Increasing the LET increases the SEU cross-section. At high values of the LET (> 20 $MeVmg^{-1}cm^2$) the cross-section increases slowly with increasing LET due to the fact that enough charge is available for a SEU, while only the probability to deposit the charge in the sensitive region of a memory cell remains. The solid curve in Figure 8.29 represents a Weibull fit [Huh00], and can be used to estimate the SEU cross-section for hadrons irradiation. The SEU cross-section for protons with an energy of 60 MeV was estimated to be 9×10^{-16} cm^2.

8.4.9.2 Measurements with 60 MeV protons

The SEU cross-section measurements were repeated with 60 MeV protons in order to confirm the heavy ion results. The chip was irradiated for seven hours, for a total fluence of $6.4 \cdot 10^{12}$ cm^{-2}. The results of the measurements are summarised in Table 8.4. In total 84 SEUs were found, over the 41296 memory cells that were irradiated. The SEU cross-section for 60 MeV protons is then 3×10^{-16} cm^2. This result is in good agreement with the result for 60 MeV protons as calculated from the heavy ion data (9×10^{-16} cm^2).

In order to obtain an estimate for the number of SEUs in the entire ALICE SPD, the calculations which were performed for the CMS experiment are scaled with the ALICE expected particle flux and with the 60 MeV protons SEU cross-section. The neutron flux in central Pb-Pb collisions at ALICE is estimated to 6.4×10^4 $cm^{-2}s^{-1}$ for the first pixel detector layer. The hadron flux originating from the Pb-Pb interactions was simulated using GEANT and is 2×10^5 $cm^{-2}s^{-1}$. For the entire ALICE pixel detector these particle fluxes would result in an upset rate of less than 1 bit of one DAC every 10 hours. It can therefore be concluded that SEU are not a threat for continuous operation of the ALICE SPD.

Fluence (cm^{-2})	# SEUs	# irradiated cells	Cross-section (cm^2)
$6.4\ 10^{12}$	84	41296	$3.2\ 10^{-16}$

Table 8.4: SEU cross-section measurements with 60 MeV protons [Hun01].

8.4.10 Temperature sensitivity

An LHCbPIX1 chip was heated at different temperatures, ranging from 30°C to 90°C. The chip parameters were measured and the real temperature was monitored by means of a temperature sensitive resistor (PT100) stuck on the PCB just under the chip. The minimum Pre_VTH before starting the heating up was 215 and all the measurements were done at this value.

TEMPERATURE		THRESHOLD DISTRIBUTION		NOISE DISTRIBUTION	
CHIP [°C]	OVEN[°C]	MEAN	RMS	MEAN	RMS
33.1	30	12.78	1.805	1.295	0.1161
53.8	50	14.62	1.935	1.326	0.1193
73.6	70	16.28	2.069	1.423	0.1117
94.2	90	17.86	2.282	1.393	0.134

Table 8.5: Effects of the temperature on an LHCbPIX1 chip; the real temperature was measured close to the chip by means of a PT100 [Cas03].

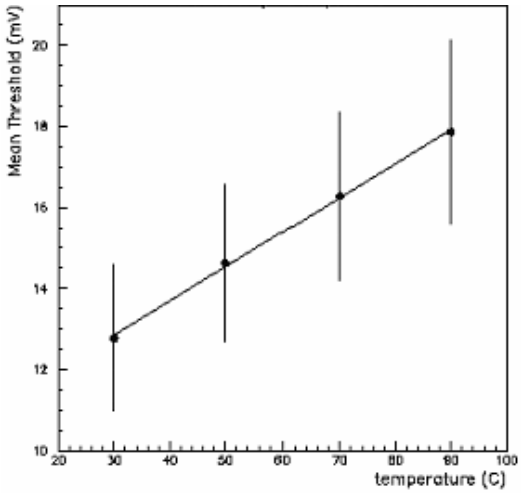

Figure 8.30: Evolution of the chip mean threshold with (oven) temperature [Cas03].

Table 8.5 shows the effects of the temperature on mean threshold and noise of an LHCbPIX1 chip; Figure 8.30 shows the evolution of the chip mean threshold with the oven temperature. The mean threshold increases linearly with temperature (with a slope of about 9 e^- / °C) as well as the threshold dispersion. Also noise changes slightly with temperature, but the effect is too small to be precisely measured.

The threshold increase with temperature is due to the change of the output of the Pre_VTH dac, as shown in Figure 8.31.

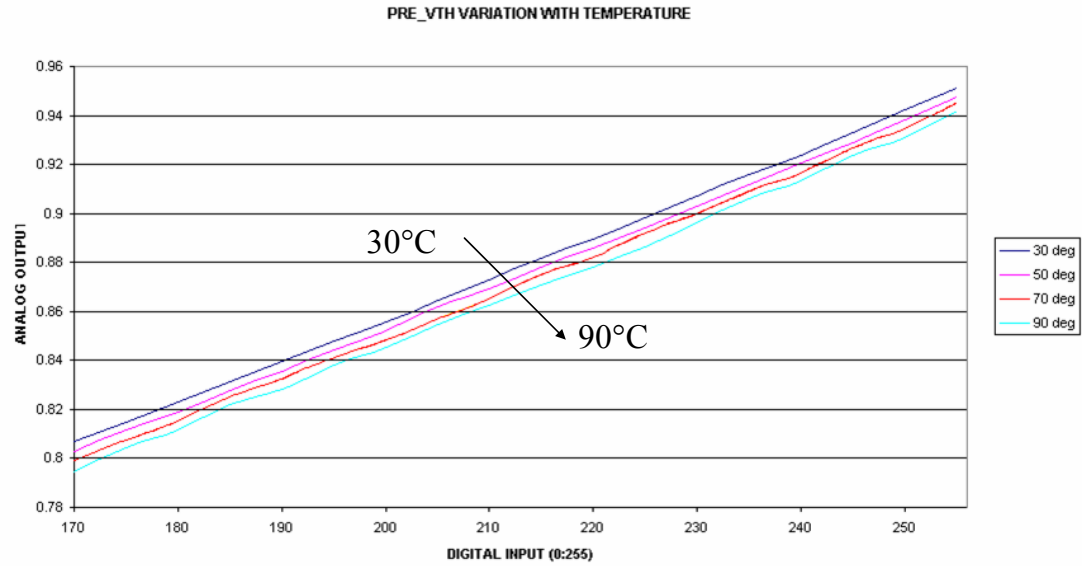

Figure 8.31: Pre_VTH (chip global threshold setting) DAC scan at different temperatures. For a given code the DAC analogue output decreases linearly with temperature [Cas03] (curves from top to bottom: 30°C, 50°C, 70°C, 90°C).

CHIP TEMP. [°C]	Pre_VTH	MEAN THRESH. [e-]	NOISE [e-]
22	220	940	130
50	223	940	140
90	227	1030	150

Table 8.6: For a given temperature it is possible to find an appropriate Pre_VTH (chip global threshold setting) DAC code which shifts the mean threshold back to around the original value at 30 °C.

In effect it is possible, for a given temperature, to find an appropriate Pre_VTH DAC code which shifts the mean threshold back to around the original value at 30 °C, as shown in Table 8.6.

8.5 The assembly

As we have seen in Chapters 3 and 4, readout chips are bump-bonded to a small detector to form a "single", or to a detector which can host five chips to form a "ladder". Tests on both structures were performed extensively, and will be reported in this section.

Measurements presented in sections 8.5.1 to 8.4.5 are done with an LHCbpix1 chip; the sensor was biased at 50 V and the total leakage current was in the order of 470 nA.

A calibration curve of the Pre_Vth DAC of the assembly used for lab tests in the region of interest is shown in Figure 8.32 .

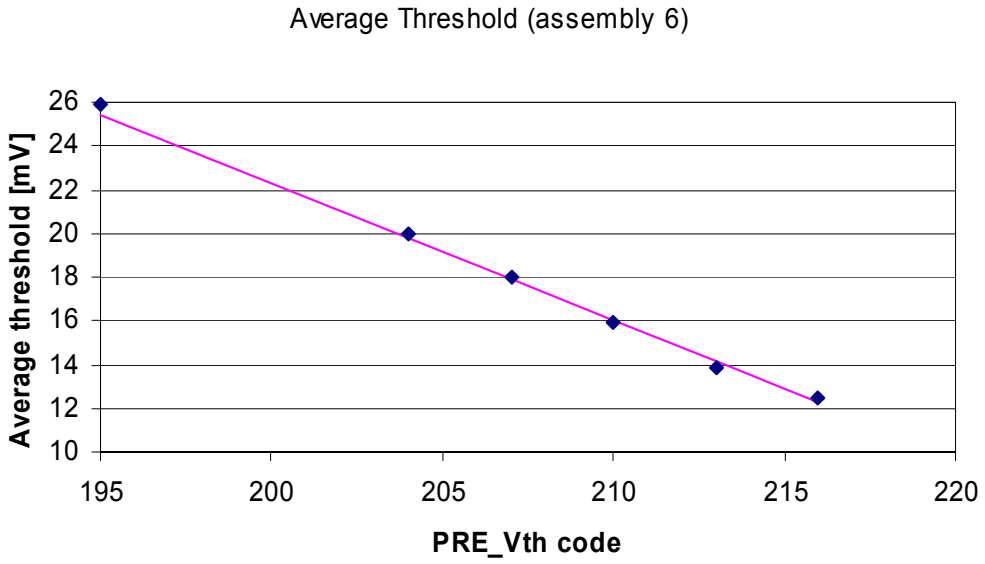

Figure 8.32: Calibration curve of the Pre_Vth DAC (chip global setting) of the assembly used for lab tests in the region of interest. The data points are plotted, as well as a linear fit (solid curve).

8.5.1 Transient analysis

Figure 8.33 shows the differential output of the second shaping stage (LHCbpix1 chip) for an input signal of Q_{in} = 5000 e^- (left) and Q_{in} = 16000 e^- (right).

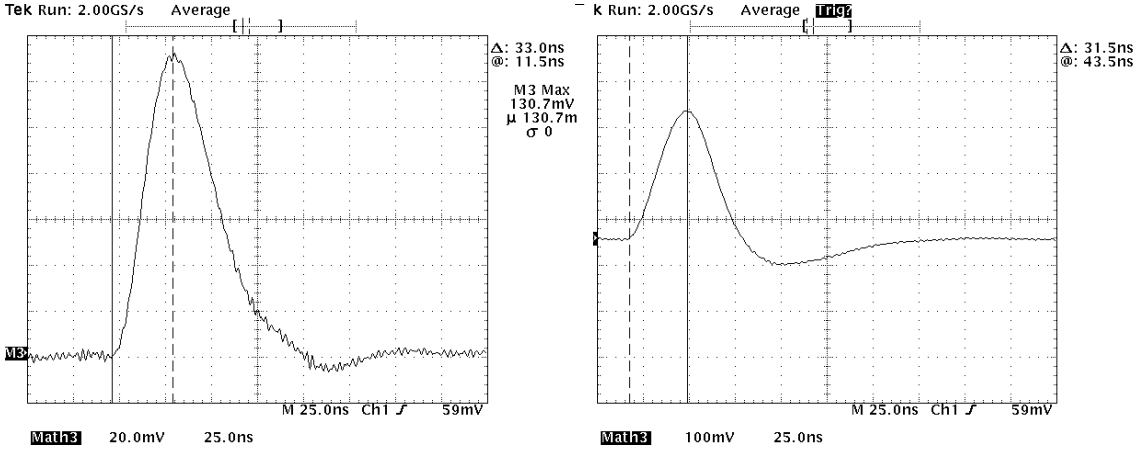

Figure 8.33: Differential output of the second shaping stage (LHCbpix1 chip) for an input signal of Q_{in} = 5000 e^- (left) and Q_{in} = 16000 e^- (right). The peaking time is about 33 ns and the peak is about 131 mV for Q_{in} = 5000 e^-. For Q_{in} = 16000 e^- the peaking time is about 31.5 ns and the peak is about 278 mV.

The peaking time is about 33 ns and the peak is about 131 mV for Q_{in} = 5000 e^-. For Q_{in} = 16000 e^- the peaking time is about 31.5 ns and the peak is about 278 mV.

A very small undershoot is present for Q_{in} = 5000 e^-, while for Q_{in} = 16000 e^- it is more pronounced (~54 mV) and is increased with respect to the un-bonded chip. The effect is due to the sensor capacitance. It should also be pointed out that, due to the increased size of the LHCb pixel (65 μm × 500 μm instead of 50 μm × 425 μm) the detector capacitance is higher than the ALICE pixel. This is consistent with simulations, as the simulated effect of higher detector capacitance is the increase of the undershoot, which increases more for bigger signals.

8.5.2 Threshold and noise measurements

The threshold distribution and the Gaussian fit for the bump-bonded chip under test are presented in Figure 8.34. The mean is 13.28 mV and the dispersion 0.91 mV, this corresponds to 1328 e^- (below the specification target value of ~2000 e^-) and 91 e^-. The threshold dispersion is very close to the calculated value (~70 e^-) and again well within the design specifications (<200 e^-).

The noise distribution and the Gaussian fit for the bump-bonded chip under test are presented in Figure 8.35. The mean is 1.72 mV and the dispersion 0.17 mV, which corresponds to 172 e^- and 17 e^-. The increase of the noise is due to the higher capacitance in parallel with the front-end input. Due to the intrinsic imprecision of noise calculations, and to

the increased detector capacitance of the LHCbpix1 chip, this value is consistent with the calculated value of ~120 e^- (see section 7.2.4.2).

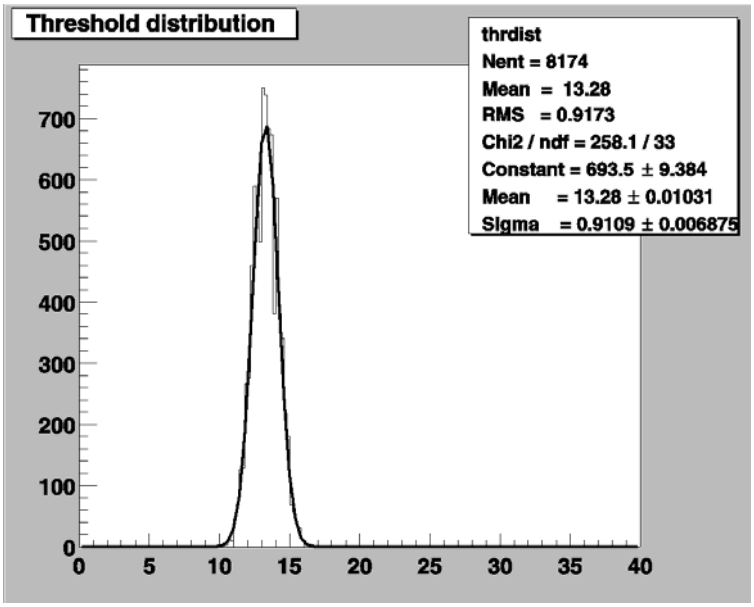

Figure 8.34: Threshold distribution and Gaussian fit for the bump-bonded chip under test; the mean is 13.28 mV and the dispersion 0.91 mV (1328 e^- and 91 e^-).

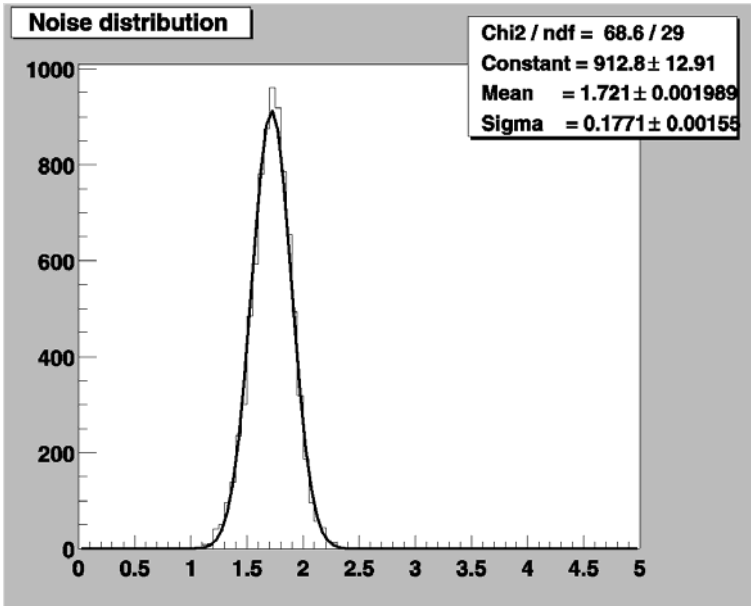

Figure 8.35: Noise distribution and Gaussian fit for the bump-bonded chip under test; the mean is 1.72 mV and the dispersion 0.17 mV (172 e^- and 17 e^-).

8.5.3 Timewalk

To measure the timewalk it is necessary to measure the threshold of the pixel under test. The threshold was extracted with the s-curve method (Figure 8.36 shows the s-curve of the

pixel under test) where the pixel was pulsed 1000 times for each value of the input charge to increase statistics. The threshold is V_{th} = 1280 e^-, the noise 160 e^-.

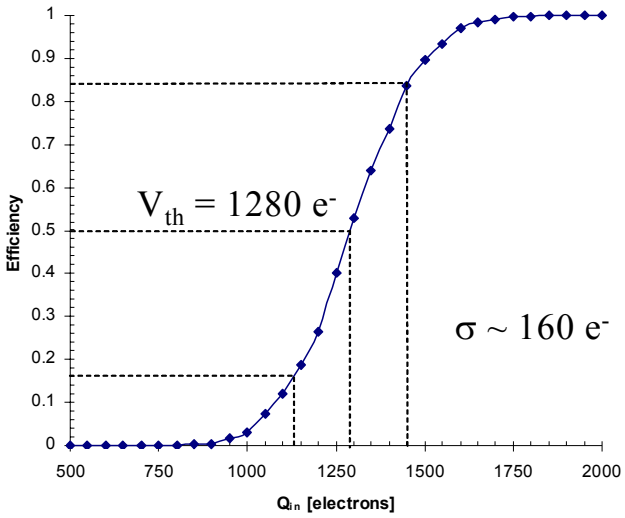

Figure 8.36: S-curve of the pixel under test. The pixel was pulsed 1000 times for each value of the input charge to increase statistic. The threshold is V_{th} = 1280 e^-; the noise 160 e^-.

Figure 8.37 plots the timewalk curve (measured as explained in section 6.3.2) referred to the delay with respect to a 50,000 e^- pulse. The pixel under test is a test pixel[3] because only in test pixels it is possible to measure the discriminator delay using the analogue test outputs.

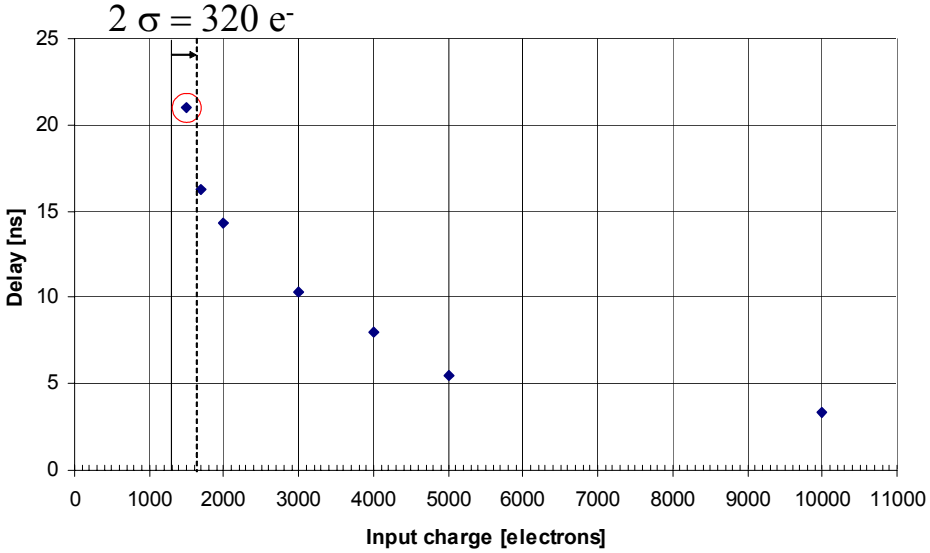

Figure 8.37: Time walk curve; the delay is referred to a 50,000 e^- pulse. The measurement of the circled point in the figure is extremely imprecise, because it is at less than 2 σ_{noise} from the threshold.

[3] The measurements refer to the test pixel (0,1).

Injecting input charges close to the pixel threshold results in a jittering of the discriminator output. The discriminator delay was estimated putting the oscilloscope in "envelope" mode and measuring the mid point of the envelope created by the discriminator edges. For this reason, the measurement of the circled point in the figure is extremely imprecise, because it is at less than 2 σ_{noise} from the threshold. This makes the estimation of a precise value for the timewalk very difficult; only an upper limit of ~250 e^- can be extrapolated from the picture.

8.5.4 Capacitive charge sharing

The pixel in row 1 column 1 (1,1) has been pulsed 10,000 times for increasing values of the input voltage. The threshold was set at about 1,800 e^- (Pre_Vth = 207). Pixel (1,1) has been chosen to have the discriminator output of a neighbouring pixel in the column available (the test pixel in row 0 column 1 (0,1)).

No signal is induced in the neighbouring pixel up to an input charge of about 55,000 e^-, which is the same value which was found with circuit simulations. At 56,000 e^- pixel (0,1) is hit at least once after 10,000 triggers on pixel (1,1). To have a signal systematically appearing in (0,1) as an almost stable output on the discriminator output checked with the scope, a signal higher than 65,000 e^- is needed. At 90,000 e^- also the neighbouring pixels in the row, (1,0) and (1,2) start to appear. This means that the ALICE and LHCb input signals (which will be in a range from 2500 e^- to 34000 e^-) will not produce spurious hits due to capacitive charge sharing.

8.5.5 Source tests

^{55}Fe-source measurements were carried out to study the response to very low signals close to the minimum threshold of the chip (~1000 electrons rms, the γ-rays from the ^{55}Fe-source create only ~1600 electron-hole pairs in 300 μm silicon) [Rie03].

Figure 8.38 shows a hit map (normalized to 1) of such a measurement after a preliminary threshold adjust. The missing pixels at the bottom left corner correspond to the position of the wire and glue drop of the backplane connection of the detector.

The pixels missing in column 10 are also insensitive to the test pulse. The missing pixels of the right-most column (columns 32) become sensitive after applying a correct reset procedure to the chip which was not done prior to this measurement. This measurement shows that the chip is also sensitive to very low input charges, and that can discriminate a signal which is only 1.6 times higher than the threshold.

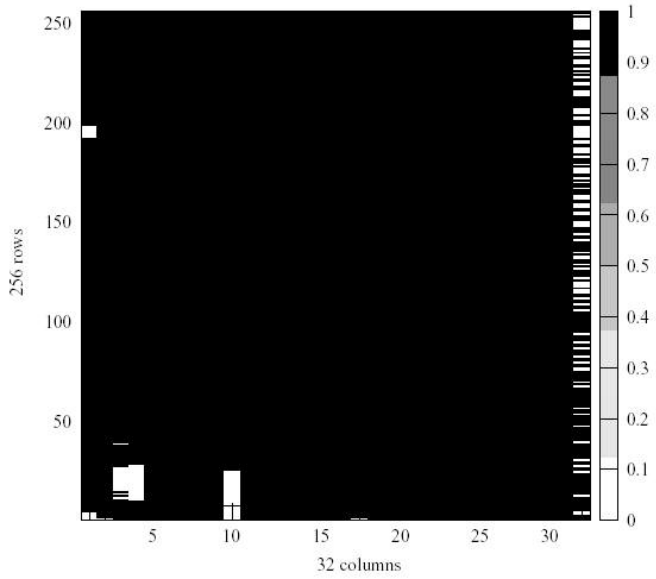

Figure 8.38: Hit map of assembly VTT 12 measured with an ^{55}Fe-source. The pixels missing in column 32 are due to an inadequate reset of the chip applied after power-up. The pixels missing in the bottom right corner are due to the wire and glue drop which provide the backplane connection for the detector. The measurements were carried out at 80 V bias [Rie03].

8.5.6 X-ray irradiation

A "single" assembly was irradiated [Rie03a] with the same setup described in section 8.4.8, with doses up to ~10 Mrd (SiO_2). This is equivalent to a maximum dose of ~17.8 Mrd (Si) for the sensor and ~1 Mrd (SiO_2) for the chip (corrected taking into account the sensor thickness). Figure 8.39 shows the total sensor leakage current as a function of the bias voltage for different irradiation doses.

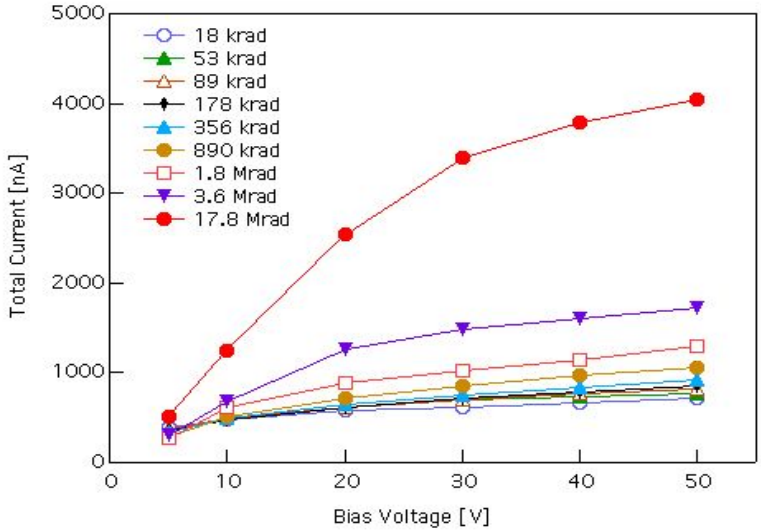

Figure 8.39: Total sensor leakage current as a function of the bias voltage for different irradiation doses [Rie03a].

The mean noise and its dispersion do not change throughout the irradiation. The threshold distribution and its sigma shift of about 20% at the maximum dose.

8.5.7 Test beam results

8.5.7.1 Test beam with 150 GeV/c pions at CERN, 2001

Several assemblies were tested during two test beam periods in 2001, in the H4 beam line at CERN using 150 GeV/c pions [Hun01, Rie03]. Figure 8.40 shows a schematic representation of the test beam setup.

A beam spot of approximately 2×2 mm^2 was selected using four scintillators in coincidence. One large scintillator (scintillator 1, ~5×5 cm^2; ~5 mm thick) was placed 10 m upstream, while two small scintillators (each 2 mm wide, 2 mm thick and 20 mm long) were placed orthogonally directly in front of the first assembly. One large scintillator was placed directly after the last assembly. The total beam size was also measured using only scintillator 1 in the trigger and determined to be ~10×5 mm^2 (3 σ).

In the first test beam period only one assembly could be mounted in between the scintillators. In the second test beam period three assemblies were mounted to study tracking. The first and the last plane were used as reference planes, while in the middle position several assemblies were tested. In total three assemblies bump bonded by AMS and three assemblies bump bonded by VTT were tested in the test beams. The assemblies were composed by 300 μm thick sensors and 750 μm thick chips.

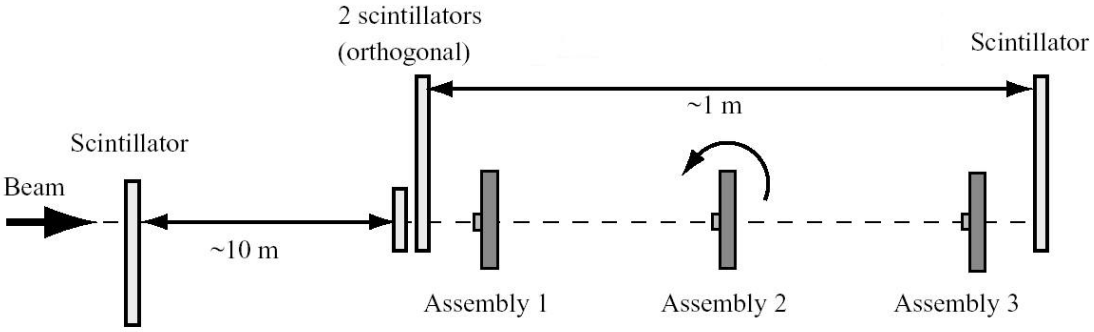

Figure 8.40: Schematic representation of the test beam setup.

The results given in this section concern mainly the first configuration. The efficiency was determined using the information from the scintillator trigger[4]. The precision of this measurement is in the order of 1–2%.

Measurements were taken with two different beam configurations. For an extraction time of 5 s the beam intensity was changed from ~10^5 particles per spill to ~10^6. The measurements at higher intensity allowed the checking of the chip performance in the case of

[4] An efficiency of 100% means that for each trigger there was a hit in the pixel assembly.

simultaneous reading from and writing to the chip. Taking into account the beam spot size of 10×5 mm^2 the high intensity setting leads to a flux of about 4×10^6 /(cm^2·s). The particle fluxes expected in ALICE are in the order of 2×10^5 /(cm^2·s).

From the image of the beam spot on the chip we extracted the beam spot dimensions. In the x-direction the beam is ~5 columns wide, corresponding to 2.1 mm; in the y direction the beam is ~40 rows wide, corresponding to 2 mm. This is consistent with the beam spot selected by the scintillators.

The strobe is generated by the scintillator trigger signal and its duration could be varied for testing purposes from 100 to 200 ns. At the CERN SPS the particle bunches arrive randomly with respect to the 10 MHz clock of the chip. The strobe width was set at 120 ns, and the trigger delay was changed in steps of 8 ns. The efficiency as a function of the strobe delay is plotted in Figure 8.41. Two different threshold settings were used, namely 200 and 215 which correspond to about 2300 and 1200 electrons, respectively (the working range for ALICE will be between 3000 and 1200 electrons). The shape of the curve is as expected, and indicates a plateau of 20 ns. With a strobe width of 100 ns there would be no plateau. There is no significant difference between the results for the different thresholds.

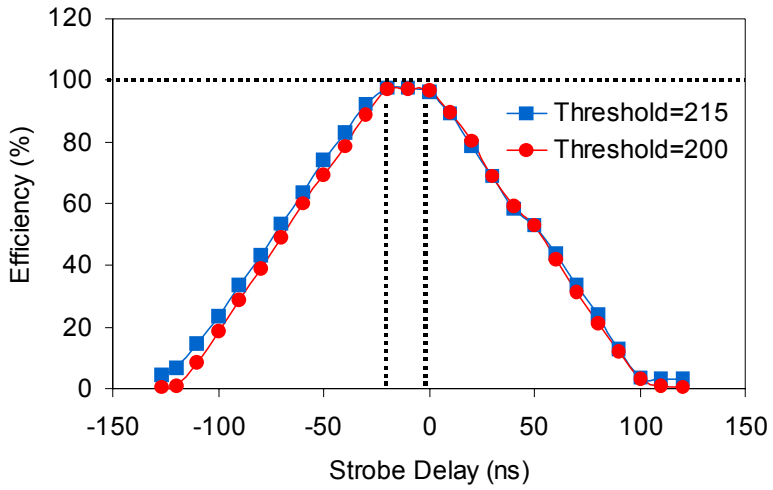

Figure 8.41: Efficiency as a function of strobe delay.

Figure 8.42 plots the efficiency as a function of the detector bias voltage. These results show that the detector can be operated with full efficiency over a large voltage range. The curve with a threshold setting of 175 (approximately 3800-4000 electrons) is shown for illustration purposes only as the chip will be operated at lower threshold. In the threshold region of interest for ALICE a stable plateau in efficiency of about 99% is observed for all assemblies.

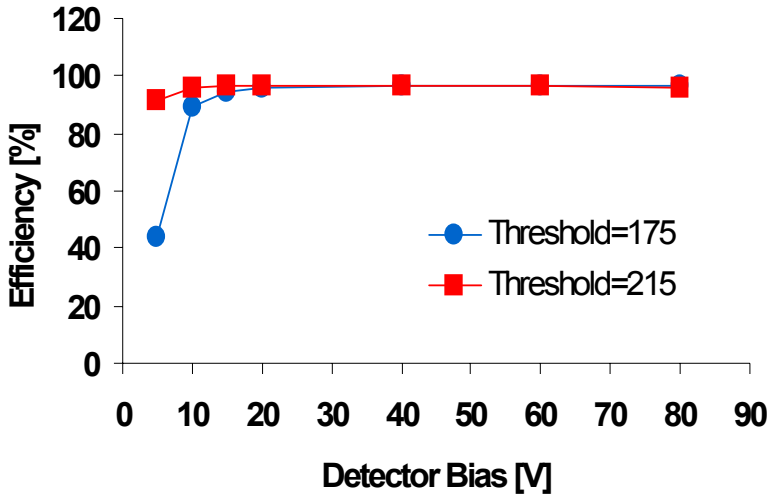

Figure 8.42: Efficiency as a function of detector bias voltage. The curve with a threshold setting of 175 (approximately 3800-4000 electrons) is shown for illustration purposes only; the chip will be operated at a threshold setting around 215, corresponding to about 1200 electrons.

The cluster size and efficiency studied as a function of the incident angle of the particles are plotted respectively in Figure 8.43 and Figure 8.44 for different thresholds. The angle between the beam and the assembly could be varied using a remote controlled stepping motor connected with an x-y table.

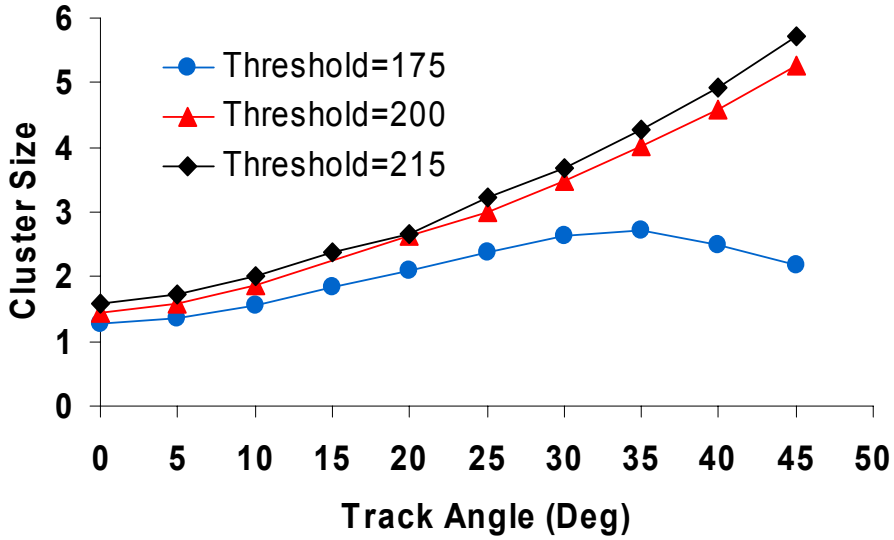

Figure 8.43: Cluster size as a function of the angle of the assembly. At zero degrees the substrate is perpendicular to the particle beam.

Figure 8.43 shows an increase of the cluster size with increasing assembly angle. For very high thresholds the cluster size decreases as a result of the decreasing charge deposition per cell when more cells are traversed. There is again no significant difference between the results with a threshold setting of 2300 and 1200 electrons.

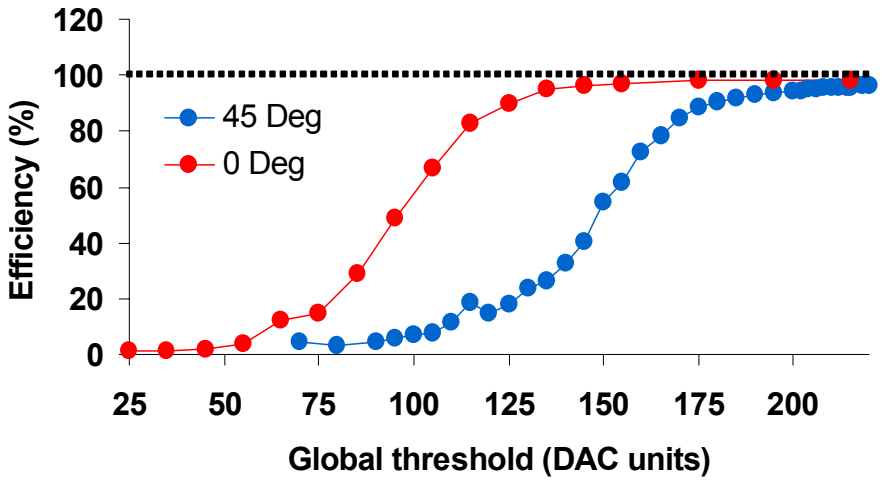

Figure 8.44: Efficiency as a function of the threshold setting for 0 and 45 degrees.

8.5.7.2 Test beam with 350 GeV/c protons at CERN, 2002

A test beam was carried out in July 2002 in the H4 beam line at CERN using 350 GeV/c protons [Rie02, Nil03]. The setup is very similar to that shown in Figure 8.40. The difference is in the implementation of the three stages of ALICE pixel detectors aligned along the beam axis. The first and the last stage were used as reference planes for tracking. Each reference plane consisted of two single chip assemblies mounted one behind the other. The distance between these two single assemblies along the beam axis is ~2 cm. Both singles were read out together via one DAQ chain similar to the readout scheme of a pixel bus (explained in Chapter3). Therefore a reference plane is referred to as mini-bus. The two reference planes, each equipped with one mini-bus, provided 4 space points for tracking. The central plane was mounted on an x-y table which allowed also to rotate the plane with respect to the beam axis. The individual pixel thresholds were not adjusted for all the measurements.

During the 2 weeks of test beam 2 single assemblies were tested in the centre position of the setup. One assembly consisted of a 200 μm thick detector bump bonded to a 750 μm thick chip (VTT49) and the other single assembly consisted of a 300 μm detector on a 750 μm thick chip (AMS76). For both assemblies timing scans, threshold scans, bias scans and measurements at divergent angles were carried out. The measured online efficiencies were 99.6% (VTT49) and 98.8% (AMS76). An example of offline efficiency[5] as a function of the chip global threshold is shown in Figure 8.45.

One ALICE ladder (VTT2-2001 composed by a 300 μm thick sensor and a 750 μm thick chips) mounted on a prototype bus was also tested in the centre position of the setup. The efficiency of the ladder was determined to be better than 99%. The same set of measurements carried out for the single assemblies was repeated for each chip on the ladder.

[5] The offline analysis was performed by P. Nilsson and is summarized in [Nil03].

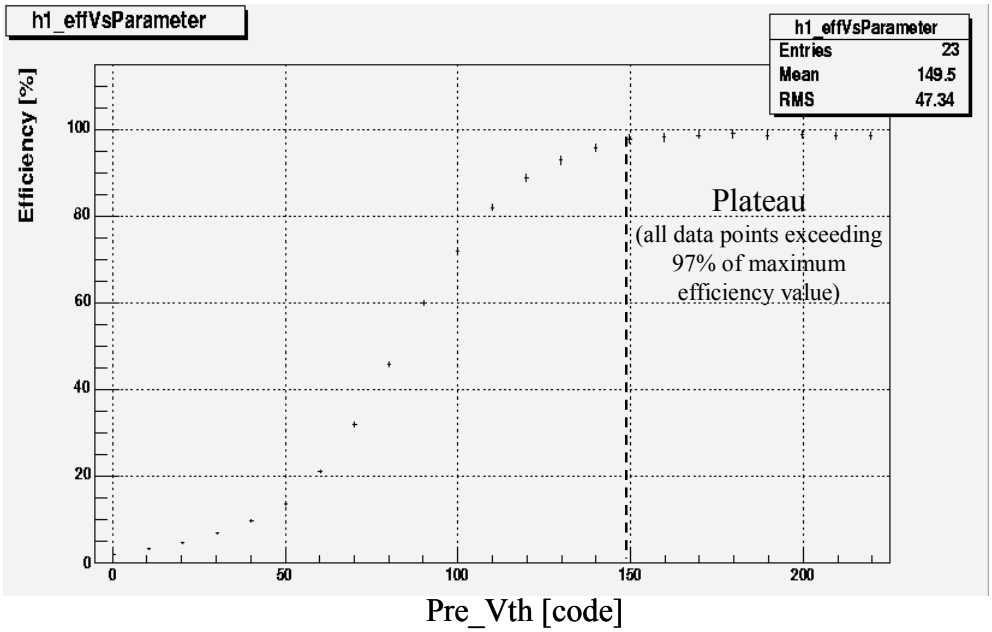

Figure 8.45: Offline efficiency as a function of the chip global threshold DAC setting (Pre_Vth); Chip 0, average plateau efficiency: (98.6 ± 0.3)%; maximum efficiency: (99.1 ± 0.9)% [Nil03].

Additional threshold scans were taken by positioning the beam spot between two chips. In the region between two chips the pixel cells are elongated (625 μm instead of 425 μm) to fully cover the gap between chips. No difference in efficiency was observed for the inter-chip regions compared to the centre of one chip.

Figure 8.46 shows the efficiency as a function of the bias voltage measured on chip 3 of the ladder.

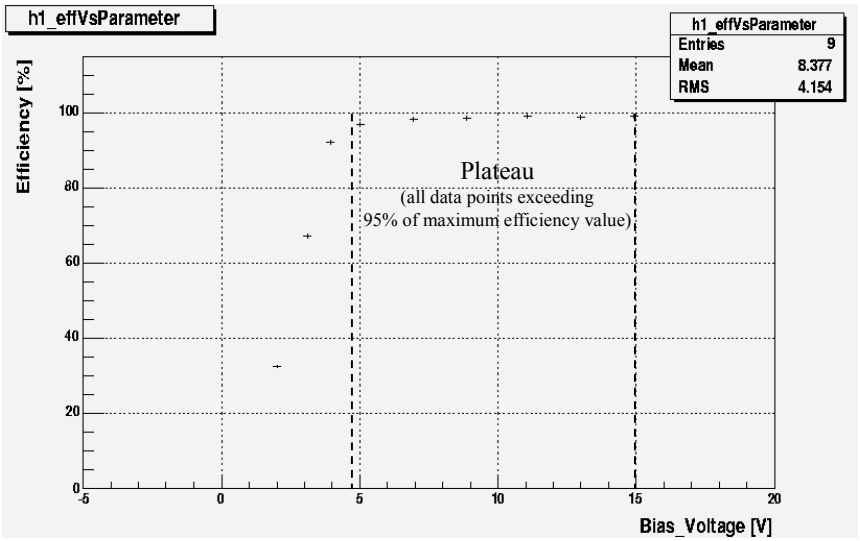

Figure 8.46: Efficiency as a function of the sensor bias voltage. A wide plateau in efficiency above 99% is present above depletion voltage. Global threshold DAC setting Pre_Vth = 200 (~2300 e^-), angle = 30°; average plateau efficiency: (98.5 ± 0.3)%; Maximum efficiency: (99.1 ± 0.6)% [Nil03].

The depletion voltage is 21 V. The measurement was repeated at three different thresholds; a wide plateau in efficiency above 98% is observed for all three threshold settings above depletion voltage.

The efficiency as a function of the strobe delay, measured in the same way as in the test beam in 2001 is plotted in Figure 8.47. A wide plateau of about 50 ns with an average efficiency of 98.4 % is present.

Figure 8.48 plots the average cluster size as a function of the chip global threshold. The cluster size increases for lower values of the threshold (higher values of Pre_Vth) due to the fact that the pixels become more sensitive to geometrical and capacitive charge sharing.

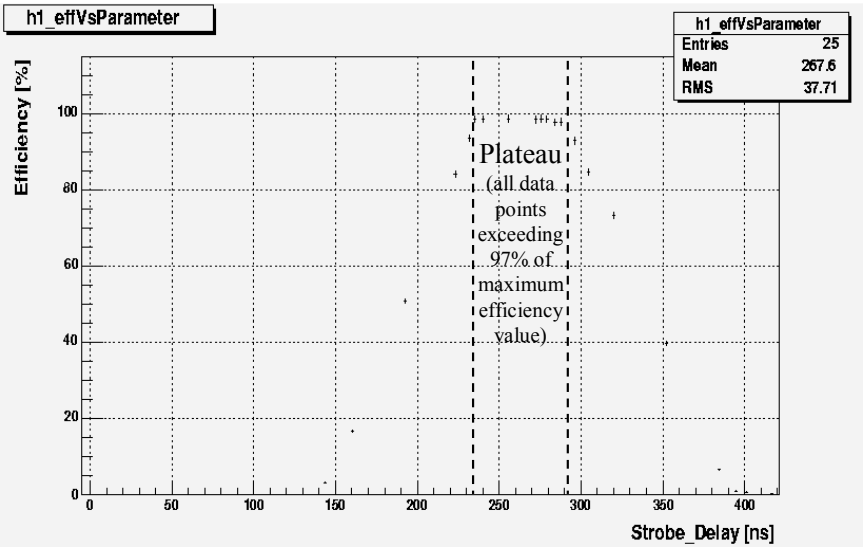

Figure 8.47: Efficiency as a function of the strobe delay; average efficiency at plateau: (98.4 ± 0.3) %; plateau width: (52 ± 4) ns [Nil03].

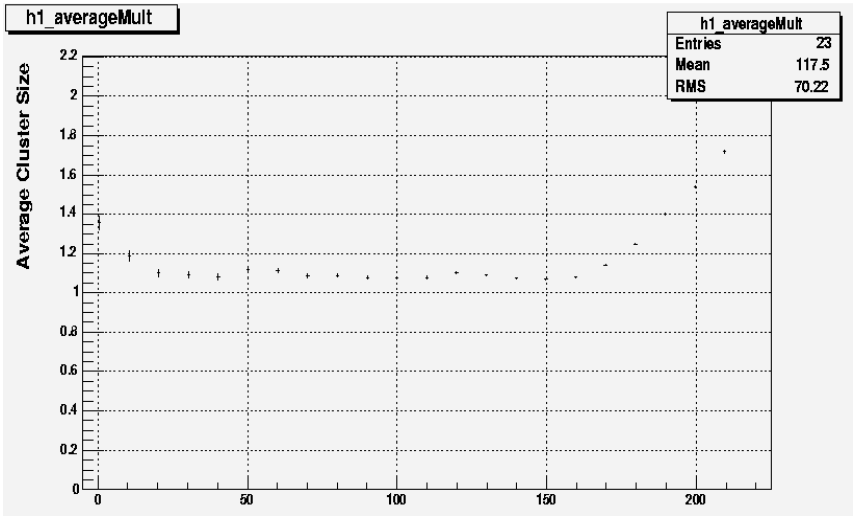

Figure 8.48: Average cluster size as a function of the chip global threshold, chip 1 [Nil03].

Figure 8.49 plots the efficiency as a function of the angle between the chip and the beam. The efficiency is always close to 100%, and the average efficiency is (99.1 ± 0.3) %.

The tracking resolution was calculated for all the measurements in both x and y directions. For the tracking resolution in the x direction (the long pixel dimension) σ_x, the average value is about 130-135 μm, close to the theoretical value of 425 μ/$\sqrt{12}$ = 123 μ. This value increases for junction pixels that are 625 μm long; on average it is 180-185 μm, very close to the theoretical value of 181 μm. For the tracking resolution in the y direction (the short pixel dimension) σ_y, the average value is about 14-15 μm, close to the theoretical value of 14.4 μm.

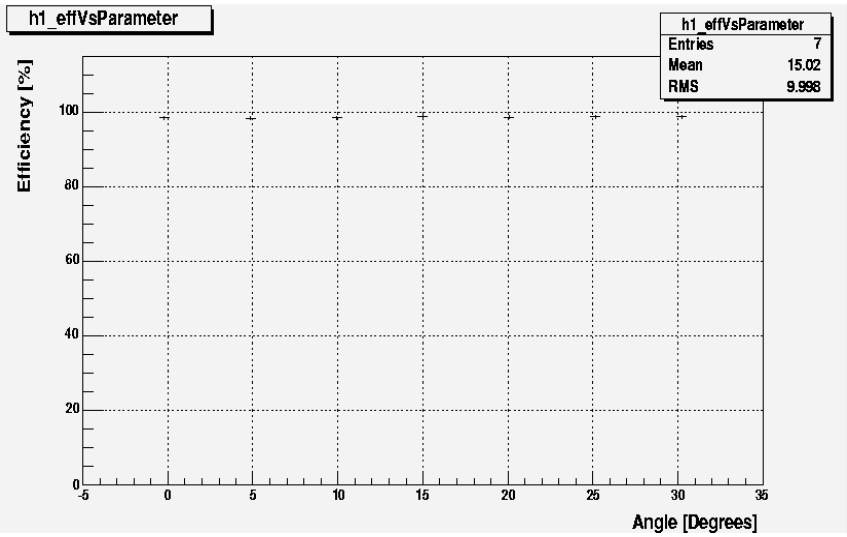

Figure 8.49: Efficiency as a function of the chip/beam angle. VTT49thin; global threshold DAC setting Pre_Vth = 210. Average efficiency: (99.1 ± 0.3) % [Nil03].

8.5.7.3 Test beam with 158 GeV/c Indium and 120 GeV/c protons at CERN, 2003

A test beam has been also carried out in October-November 2003 with two different beam settings. The chips were irradiated for 15 days with a 158 GeV/A Indium beam (10^4 ions/spill), which generates high multiplicity events. The chips were also irradiated for three days with a 120 GeV proton beam (10 protons/spill). Figure 8.50 plots the indium beam spot (colour plot, left and 3D plot, right) focussed on a single assembly (sensor bias: 32V; 10^4 ions/spill).

Up to 4 pixel planes were used for tracking (i.e. five assemblies and a "half stave"; up to 122,880 channels read out), and some tests were also carried out with a full OPS used as the last plane. Online results show the full functionality of the system.

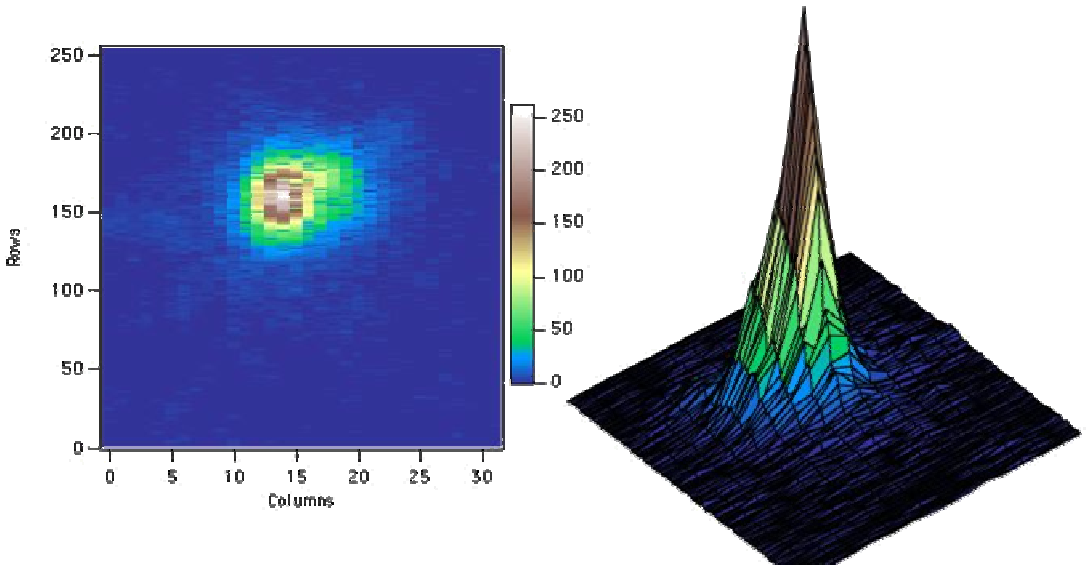

Figure 8.50: Plot of the Indium beam spot (colour plot, left and 3D plot, right) focussed on a single assembly (sensor bias: 32V; 10^4 ions/spill).

The LHCb group performed also a test beam in August-September 2003, and was able to observe the first air Cherenkov rings from pions and electrons, demonstrating the functionality of the chip also in a complete HPD.

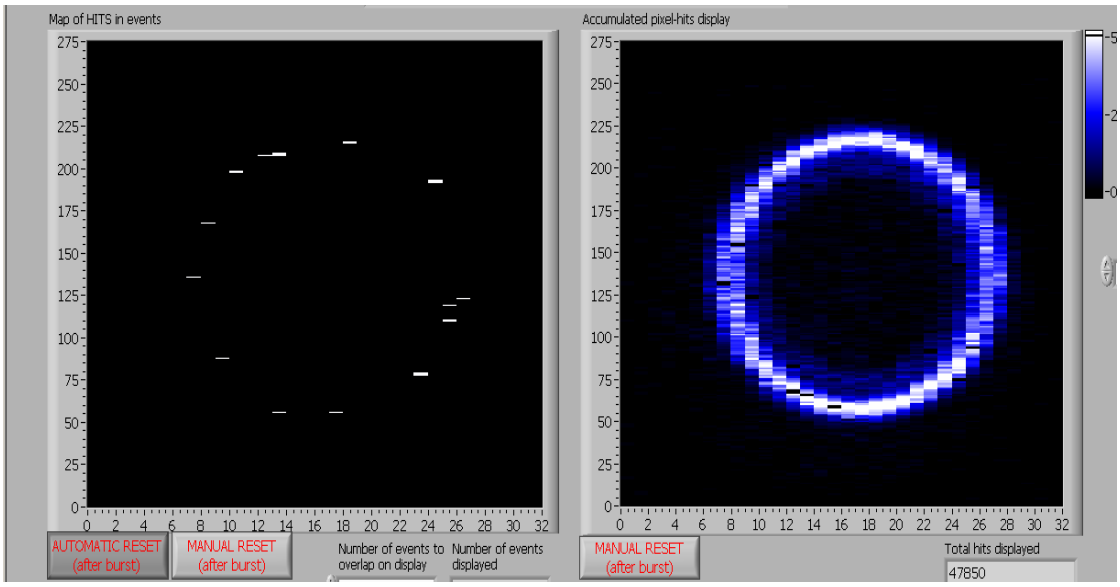

Figure 8.51: Cherenkov Rings (air radiator) obtained during LHCb testbeam in August-September 2003 (1 trigger, left; accumulated triggers, right).

8.6 Summary

The ALICE1LHCb chip is a matrix of 32 columns each containing 256 readout cells, measuring 13.5 × 15.8 mm^2. Five chips are bump-bonded to a big detector (160 columns × 256 rows) to form the ALICE SPD basic detecting block, *the ladder*. One chip is bump bonded to a detector (of the same dimensions, 32 columns × 256 rows) to form a *single*, the basic detection element for the LHCb Hybrid Photo Detector.

The first version of the chip had a limitation in the maximum achievable clock frequency (15-18 MHz). This lead to the resubmission of a second version of the chip (LHCbpix1) for LHCb which needs a 40 MHz clock.

The chips were extensively tested, and the main experimental results were presented in this chapter.

A problem with the **DACs** was solved mainly with increasing the power supply from 1.6V to 1.8V and with a proper sequencing of the testing operations for ALICE. It was solved in LHCbpix1 design.

Due to problems in uniformity of the ALICE1LHCb on-chip test pulser, the **calibration** described in section 8.3 leads to the conclusion that V_{in} = 1 mV corresponds to Q_{in} = 100 e^- for what concerns the tests done with the pulser for the LHCbpix1 chip. On the contrary, it is not so easy to estimate the voltage/charge conversion factor for the ALICE1LHCb chip; the average chip value has been roughly estimated to be about 0.5-0.65 mV/100 e^-.

Results of the **front-end of bare chips** (i.e. not bump-bonded to a detector) are shown in section 8.4 (LHCbpix1). The differential output of the second shaping stage for an input signal of Q_{in} = 5000 e^- has a peaking time of about 31 ns and a peak of about 125 mV. The front-end can recover in about 300 ns after an input pulse of about 100,000 e^-, and even at such a high input signal no electrical crosstalk is present. The chip **noise** is about 130-140 e^- rms with a dispersion of 13 e^- rms. The minimum threshold is at about 1100-1200 e^-, and the threshold dispersion less than 100 e^-. A first algorithm allowed to reduce this spread to about 43 e^- using the pixel to pixel threshold adjust.

A procedure to test chips on a **wafer probing** machine was settled, which allows the selection of good chips to be bump-bonded to a sensor.

An LHCbpix1 **chip was irradiated** with 10 keV X-rays, up to a dose of 10 Mrd (SiO_2). No significant change in the currents is observed up to the maximum irradiation dose (10 Mrd). This confirms the effectiveness of using HBD techniques to avoid radiation induced parasitic currents. A slight change in the value of the pixel threshold is induced (some hundred electrons) which can be almost recovered to the pre-irradiation value acting on the Pre_Vth DAC.

An LHCbpix1 **chip was heated** at different temperatures ranging from 30°C to 90°C; the threshold increases by ~600 e^- but can again be recovered acting on the Pre_Vth DAC.

An ALICE1LHCb **chip was irradiated** at the CRC Facility, in Louvain-la-Neuve with heavy ions with an LET between 6 and 120 $MeVmg^{-1}cm^2$. This allowed to plot the **SEU** cross-section as a function of the LET (Figure 8.29).The chip was irradiated also with 60 MeV protons, so we could calculate the SEU cross-section for 60 MeV protons, which is 3×10^{-16} cm^2. This would result in ALICE to an upset rate of less than 1 bit every 10 hours. Neither **SEGRs** nor **SELs** were observed.

Results on the **assemblies** (LHCbpix1 chip, LHCb standard sensor) are presented in section 8.5. The differential output of the second shaping stage for an input signal of Q_{in} = 5000 e^- has a peaking time of about 33 ns and the peak is about 131 mV. The minimum threshold of an assembly is slightly higher than for a bare chip (~1300 e^-), but the threshold dispersion stays below 100 e^-. The noise increases to ~170 e^- (due to the capacitance of the sensor pixel cell; the dispersion is 17 e^-). The **timewalk** at 20 ns was estimated, but due to the difficulty of the measurement, only an upper limit of ~250 e^- could be extrapolated. Also capacitive **charge sharing** was measured; no signal is induced in a neighbouring pixel up to an input charge of about 55,000 e^-.

Some ALICE1LHCb assemblies were irradiated with a ^{55}Fe-source, showing the sensitivity of the chip at very low input signals, and with 10 keV X-rays, up to a dose of 10 Mrd (SiO_2). This is equivalent to 17.8 Mrd (Si) in the sensor and 1 Mrd (SiO_2) in the chip. The mean noise and its dispersion do not change throughout the irradiation. The threshold distribution and its sigma shift of about 20% at the maximum irradiation dose.

Several ALICE1LHCb assemblies were tested during four **test beam periods** in the H4 beam line at CERN using 150 GeV/c pions, 350 GeV/c protons, 158 GeV/c Indium and 120 GeV/c protons. Efficiency measurements, timing scans, threshold scans, bias scans and measurements at divergent angles were carried out. Results were presented in section 8.5.7, and show the full functionality of the system for ALICE. The LHCb group performed also a test beam in August-September 2003, and was able to observe air Cherenkov rings from pions and electrons, demonstrating the functionality of the chip also in a complete HPD.

Chapter 9
An auxiliary chip for the ALICE and LHCb pixel chips: the Analogue Pilot Chip

The ALICE1LHCb Pixel Chip requires a number of external auxiliary features. The digital control is performed by the Digital Pilot Chip (DPC, section 3.2.3.1). The DPC can control via a JTAG interface all the chips on the half stave, and the data readout from the Pixel Chips. However, the Pixel Chip needs also some external analogue biases. Moreover, the possibility to readout the on-chip DACs once the chip has been mounted on the On-detector pixel Pilot System (OPS) is useful, as well as the possibility to measure some other analogue parameters of the OPS (i.e. voltage supplies and temperature). These auxiliary analogue functions are performed by the Analogue Pilot Chip (APC) [Ane02, Klu02]. The chip was designed to serve also the second version of ALICE1LHCb, the LHCbpix1 chip, which will be used in the LHCb RICH detector.

A prototype of the chip has been designed and tested, and a second slightly modified version has been redesigned. This chapter describes the various blocks which are in the chip, and presents measurements which show the full functionality of the final version of the APC.

9.1 Chip building blocks

9.1.1 General description

The APC is a mixed-mode IC, designed with Hardening By Design (HBD) techniques to improve radiation tolerance, containing the following blocks:

- Six 8-bit DACs, providing reference voltages to the ALICE1LHCB or LHCBpix1 chip;
- A 16-input analogue multiplexer followed by a 10-bit ADC;
- A band-gap reference, which provides a reference voltage to the other on-chip reference circuits, independent from temperature and power supply variations;
- A reference circuit which provides the necessary references to the six DACs;
- A reference circuit which provides the necessary references to the ADC;

- Four current sources for temperature monitoring (2 for ALICE and 2 for LHCb). The current sources are designed to be independent from temperature and power supply variations;
- A current sensing stage, which has the aim of reading out the Pixel Chips current DAC outputs, avoiding changing the bias condition of the DAC output stage;
- A JTAG-controlled digital block, providing all the necessary digital signals to the other blocks.

The total power consumption is around 50 mW. A block diagram of the system is shown in Figure 9.1.

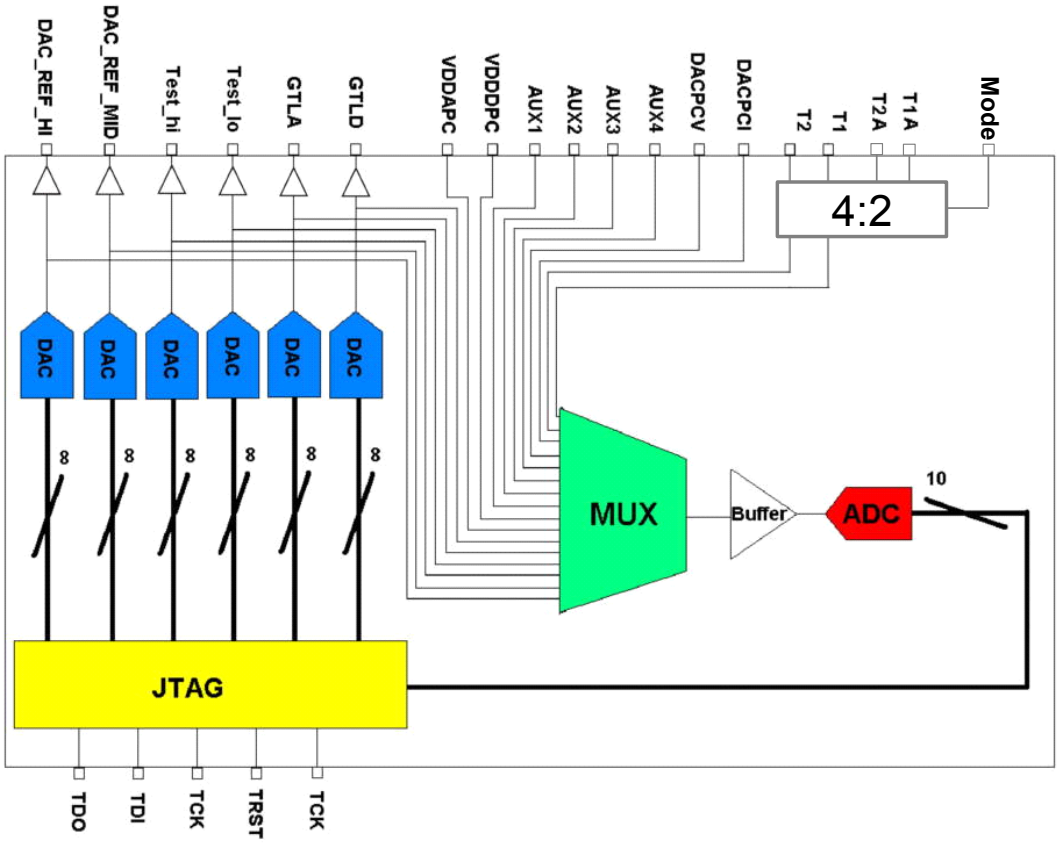

Figure 9.1: System block diagram of the APC chip.

9.1.2 The bandgap cell

Bandgap references provide precise voltages, with reduced sensitivity to temperature and process parameters. They are based on a physical quantity, for example the bandgap voltage of a semiconductor.

The working principle of the bandgap cell integrated in the APC is depicted in Figure 9.2 (left). It is based on the fact that the voltage on the two diodes D1 and on D2, biased with different current densities J1 and J2, can be subtracted resulting in a ΔV which is linearly

dependent on the absolute temperature T but with an extremely small coefficient (~ 0.2 mV/K).

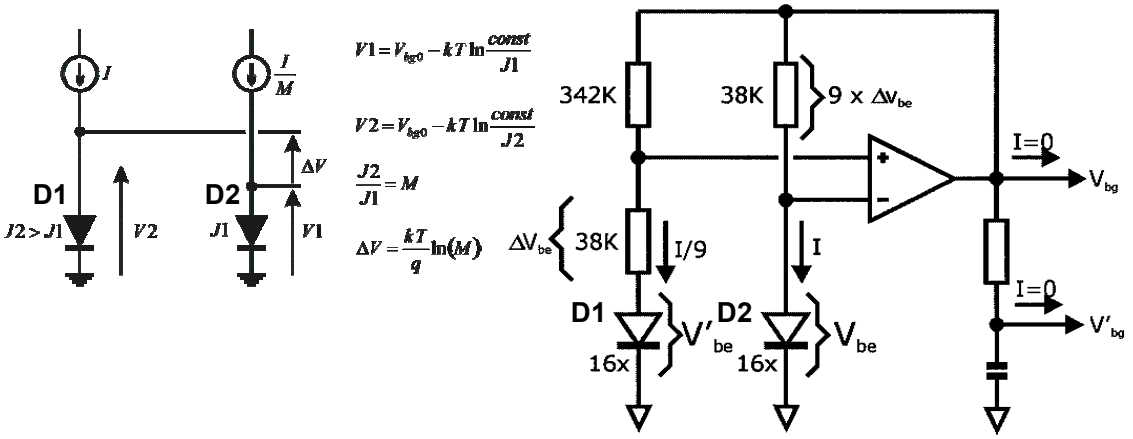

Figure 9.2: Working principle (left) and schematic circuit diagram (right) of the bandgap cell.

A schematic diagram of the bandgap cell used in the APC is shown in Figure 9.2 [Kui73]. This cell was designed by MEAD for CERN, and is available as an analogue macro-cell for the CERN designers [Web07, Mor02].

As illustrated in Figure 9.2 the band gap cell has two outputs. Both outputs provide the same voltage (bandgap voltage = 1.16 V). The output marked as V'_{bg} has an additional RC filter that should be used when a high Power Supply Rejection Ratio (PSRR) is required at high frequencies. Both outputs have no current driving capability.

The main characteristics of this cell are reported in Table 9.1.

Nominal operating voltage	2.5V
Power consumption	62.5 μW @ 2.5V
Nominal bandgap voltage	1.16V @ 0° C
Voltage temperature sensitivity	- 0.22 mV/° C (for -20° C < T < 70° C)
Power supply sensitivity	ΔV_{out} < 1 mV for V_{dd} from 1.2 V to 2.5 V.
Cell size	400 μm × 275 μm

Table 9.1: Main characteristics of the bandgap cell [Web07, Mor02].

9.1.3 The reference circuits

Several reference circuits for the ADC and the 6 on-chip DACs have been implemented. All the references are derived from the precise voltage generated by the band-gap reference block. Instead of being sent straight to the corresponding analogue blocks, the reference voltages are sent to an external pad. This was done to allow decoupling of the reference

circuits with an external capacitor and to keep the possibility (on the prototype version only) of providing the reference voltages to the analogue blocks externally. Moreover this allows the measurements of the reference voltages.

Figure 9.3 shows a schematic representation of the working principle of the current and voltage references [Rin98, Lee99].

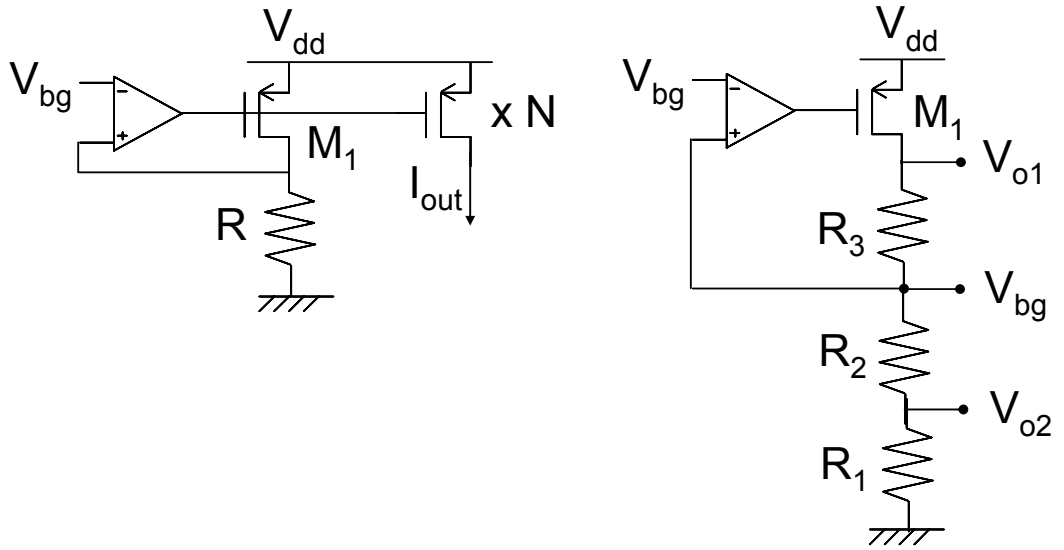

Figure 9.3: Schematic representation of the working principle of the current (left) and voltage (right) references.

The operational amplifier has virtual ground at its inputs, so that V_{bg}, applied at is negative input, is present also on its positive input. This forces a reference current V_{bg} / R in the MOS transistor M_1 (Figure 9.3, left). The current can be, if needed, multiplied (or demultiplied) using multiple copies (or scaled copies) of M_1. A drawback of this scheme is that the imprecision of the generated current depends on the imprecision of the absolute value of R, which can be in the order of 20-30% in integrated circuits. For this reason we decided to use an external resistor which can be chosen with a high precision (0.1%) and with a low temperature dependence (in the order of ten parts per million).

The working principle for the voltage references is very similar (Figure 9.3, right). Again, virtual ground at the operational amplifier input forces a current $V_{bg} / (R_1 + R_2)$ in M_1 and R_3. The output V_{o1} is used to produce reference voltages higher than V_{bg}, and is given by equation (9.1); the output V_{o2} is used to produce reference voltages lower than V_{bg} and is given by equation (9.2) .

$$V_{o1} = \frac{V_{bg}}{R_1 + R_2}(R_1 + R_2 + R_3) \tag{9.1}$$

$$V_{o2} = \frac{V_{bg}}{R_1 + R_2} R_1 \tag{9.2}$$

9.1.4 The DACs

We have seen in Chapter 7 that there are 42 8-bit DACs included in the Pixel Chip. Their role is to provide biases (voltages and currents) to the analogue and digital circuitry within the pixel cells. All are configurable via the JTAG interface. They need to be biased externally with two precise and stable voltage references, which are generated by two of the 8-bit DACs of the APC. These two outputs are sent to the two output pads vo_DRHI and vo_DRMID.

A test-input can be given to the pre-amplifier of the pixel cell in the Pixel Chip using a voltage step applied across a capacitor. The size of the step is controlled by two DC levels applied to the chip and is triggered by a logic pulse generated externally. The two DC levels are generated by two of the DACs of the APC, and sent to the two output pads vo_TESTLOW and vo_TESTHI.

The input/output pads of the Pixel Chip are 'GTL logic', and they need a reference voltage to operate properly. Two separate references have to be generated for the Analogue and for the Digital section of the Pixel Chip. These reference voltages are generated by two identical 8-bit DACs of the APC, and are sent to the two output pads vo_GTLA and vo_GTLD.

The DACs design is a modified version of the design used for the Pixel Chip itself (voltage output version) described is section 7.5.2. This revised version has to stand a higher power supply voltage (2.5V), and the reference voltages for these DACs are provided on-chip. The 6 8-bit DACs of the APC need 3 bias voltages. These 3 voltages, called vbias, DAC_REF_VDD and DAC_REF_MID, are generated on chip by some reference circuits (explained in section 9.1.3).

The DACs output stage is not low impedance, so a voltage buffer is connected at the output of each DAC to drive the DAC signal off the chip. The DACs output (before the buffers) are also brought to a pad, where a big decoupling capacitor (100 nF) has to be connected.

9.1.5 The ADC

The APC contains a 10-bit successive approximation ADC, used for the conversion of 16 DC or slow-varying signals. The ADC is based on a charge redistribution scheme, whose key elements are a binary weighted digital to analogue converter and a voltage comparator. It was designed in our target 0.25 μm CMOS technology in collaboration between CERN and the Italian National Institute for Nuclear Physics (INFN) Sezione di Torino [Riv01]. The circuit was developed to serve as a building block in multichannel data acquisition systems for High Energy Physics applications. Therefore the key features of the design are medium resolution (10 bits), very low power consumption (1 mW) and high modularity. The issue of radiation tolerance was addressed using HBD techniques, in particular ELTs and guard rings.

The circuit operates from a 2.5 V power supply, and in the original design had a maximum full scale range of 1.6 V. Tests on a first version of the chip have shown full functionality up to a clock frequency of 20 MHz, while few codes were lost above this frequency near the MSB transition. The problem has been solved in the final version with improved layout, which shows a resolution of 10 bits up to a clock frequency of 30 MHz. No degradation in the performance has been observed after the exposure of the circuit at a total ionising dose of 10 Mrd (SiO_2).

Power supply	2.5 V
Number of bits	10
Dynamic range	0.513 V – 1.925 V (in the APC)
Maximum clock speed	30 MHz
Maximum sampling frequency	~2.5 Msamples/sec
Power consumption	1 mW

Table 9.2: Main characteristics of the ADC cell.

In the APC the input range of the ADC is between the two on-chip references Vref0 and Vref1 (0.513 and 1.925; see Table 9.3). In our case, the ADC is clocked at 10 MHz. The clock and all the control signals are provided by the digital part (section 9.1.8).

A 16-to-1 multiplexer and a buffer have been put in front of the ADC, to be able to read 16 different signals: the 6 on-chip DACs, the digital and analogue power supply of the Pixel Chip, the output of the Pixel Chip voltage and current DACs and of the two temperature sensing elements. Four channels are free for other possible signals.

9.1.6 The current sensing stage

When a current of a Pixel Chip current DAC is to be sensed, it is scaled to 20 μA and is sent to a dedicated pad. To readout this current while avoiding changing the bias condition of the DAC output stage, we implemented on the APC a sensing stage, whose schematic representation is shown in Figure 9.4.

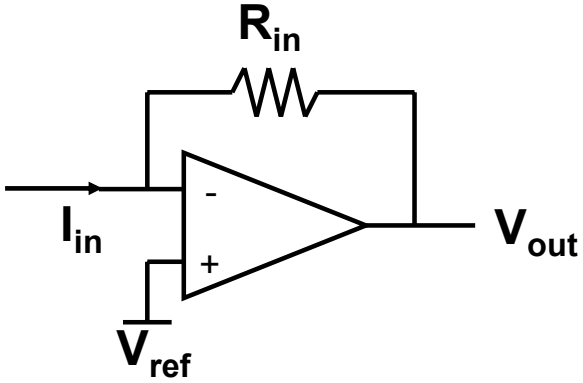

Figure 9.4: Schematic representation of the Pixel Chip current DACs sensing stage.

With this configuration the DAC output stage always sees the same voltage V_{ref}, regardless of the current which has to be sensed. The voltage $V_{out} = V_{ref} – R_{in} I_{in}$ is then sent to an input of the ADC.

9.1.7 The current sources for temperature measurement

Four current sources are implemented on chip, 2 for ALICE and 2 for LHCb. For ALICE, a group of 5 temperature sensitive resistors (PT1000) in series will be used to monitor the temperature of one ladder in the stave, while the second input will be connected to a thermistor of about 5 kΩ to sense the MCM Pilot temperature. For LHCb, there will be one PT1000 sensor with its current source for each chip. The nominal value of the current for LHCb is 1.3 mA, for ALICE is 260 μA. These values can vary with the process variations, so the system will need to be calibrated. The currents have been chosen to give a voltage within the ADC input range for temperatures up to 80 °C (and for any possible variation of the process). The voltages read on the four temperature sensitive resistors are sent to a demultiplying 4:2 block, which sends the two ALICE or the two LHCb sensed voltage to the ADC (inputs T1 and T2) depending on the logic state of the MODE signal. Figure 9.5 shows a schematic representation of the current sources circuit block for temperature measurement.

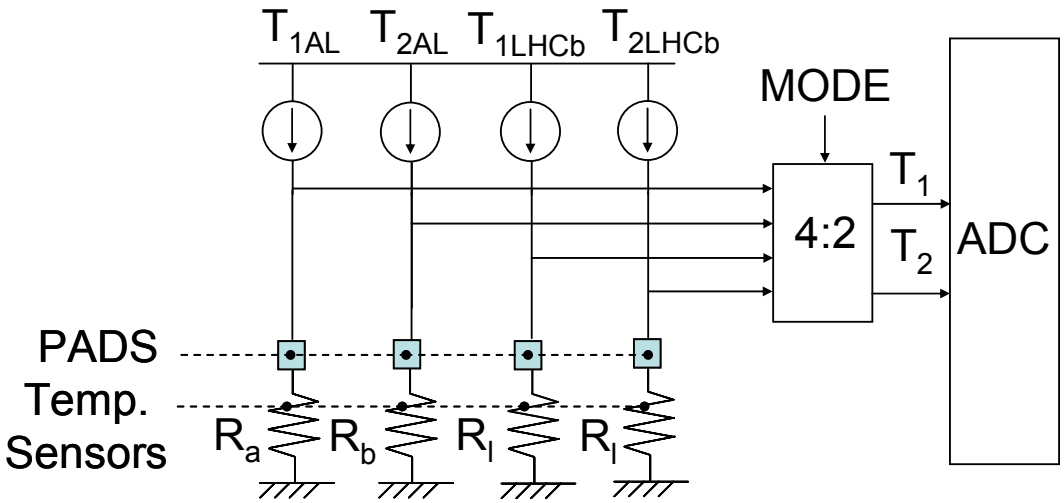

Figure 9.5: Schematic representation of the current sources circuit block for temperature measurements; the resistor R_a is a series of 5 PT1000 to monitor the temperature of one ladder (used only in ALICE mode); the resistor R_b is a thermistor of about 5 kΩ to sense the MCM Pilot temperature (used only in ALICE mode), the resistors R_l are single PT1000 (used only in LHCb mode).

9.1.8 The digital control section

The functionality of the digital core [Klu02] can be divided in four groups: programming of the six 8 bit DACs via the JTAG interface; control of the AD-conversion sequence of 16 analogue values and storage of the digital values in internal registers; read-out of the internal registers via the JTAG interface. It also allows the transmission of two analogue values on a serial output.

The digital core (about 8000 transistors in total) has been designed and synthesized using VERILOG. Digital simulations and post-layout simulations have been performed. All registers and the ADC controller have been triplicated in order to reduce the influence of single event upset errors. The maximum simulated clock frequency of the circuit exceeds 40 MHz. The maximum simulated clock skew is 110 ps. Tests performed on the chip show that the digital control section is fully functional.

9.1.8.1 Programming of the DACs

The digital core allows programming a 48 bit internal register via the JTAG interface. Each group of eight bits is connected directly to one of the six 8-bit DACs. No additional control signals are required by the DACs.

9.1.8.2 Conversion of the 16 analogue values

The APC contains one ADC and a 16 channel analogue multiplexer. Once the 'start conversion' instruction is sent to the JTAG controller, the measurement sequence is started. The controller selects the first input (input 0). After the multiplexer output has settled the AD conversion is started and consequently the ADC output is stored in the first register. This sequence is repeated automatically for all 16 analogue inputs.

9.1.8.3 Reading the measured analogue values

90 μs after the 'start conversion' command has been sent the stored values can be read via the JTAG port. The 16 10-bit registers are read out via the JTAG interface.

9.1.8.4 Serial temperature transmission

It is planned to connect one temperature sensor line to each temperature input, input0 and input1. A special signal, 'auto_temp', allows the automatic AD-conversion of the two inputs 0 and 1 only and serialization of the result. The serializer output is available as an output pad of the chip.

9.1.9 Chip layout

The final chip has an area of 4×3 mm^2; its layout is shown in Figure 9.6. The six DACs on the left-hand side of the chip and the digital control block on the right-hand side can be easily recognised.

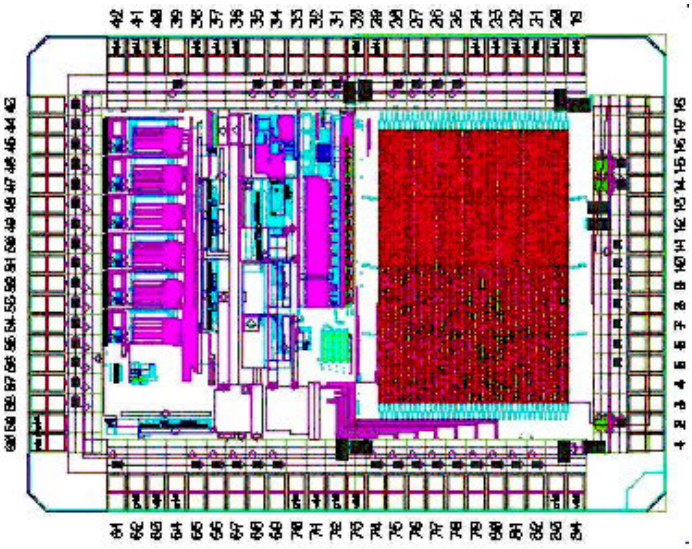

Figure 9.6: Layout of the Analogue Pilot Chip. The chip area is 4×3 mm^2.

9.2 APC measurement results

The final version of the chip (Figure 9.7 shows a photograph of the APC wire bonded to the test card), was sent to fabrication in June and received back in September 2003. This section presents the preliminary measurements done on 10 APCs; all of the 10 tested chips were fully functional (this gives an indication of the high production yield that should be expected for this chip, close to 100%).

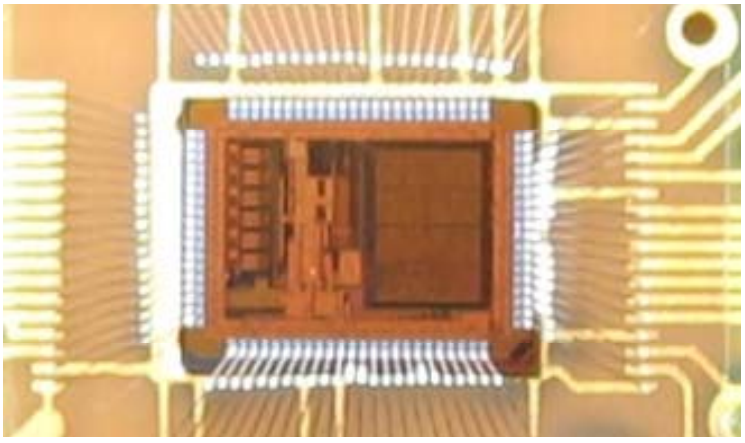

Figure 9.7: Photograph of the Analogue Pilot Chip wire bonded to the test card.

The APC was tested on a dedicated board, shown in Figure 9.8, which was plugged into an IC tester (MicroLEX systems INTEGRATest). Dedicated Labview software had to be developed to perform the required tests.

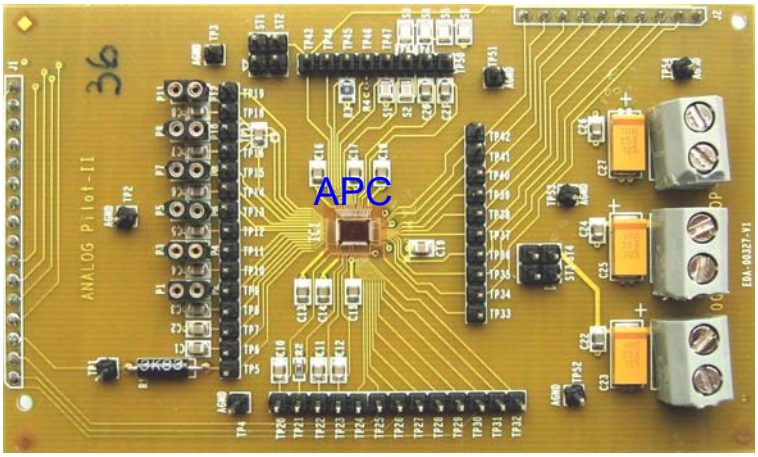

Figure 9.8: Photograph of the test board of the Analogue Pilot Chip, which is visible in the centre.

9.2.1 The bandgap cell and the reference circuits

Table 9.3 shows the measurements of the six on-chip references.

	Vref0 [V]	Vref1 [V]	Vbias [V]	DAC_ REF_V DD [V]	DAC_ REF_ MID [V]	Vout_ BG [V]
Chip 46	0.51	1.9042	1.664	0.9903	0.5561	1.1468
Chip 47	0.51	1.9035	1.664	0.9945	0.5581	1.149
Chip 36	0.5093	1.9068	1.6608	0.9894	0.5551	1.1463
Chip 24	0.5105	1.9123	1.6638	0.9931	0.5577	1.1501
Chip 35	0.5106	1.9049	1.6619	0.9888	0.5551	1.148
Chip 34	0.5098	1.908	1.6624	0.9897	0.5556	1.1454
Chip 45	0.5111	1.9084	1.6642	0.9924	0.5571	1.148
Chip 25	0.5115	1.9121	1.6636	0.9945	0.5581	1.1506
Chip 15	0.5113	1.9045	1.661	0.9941	0.5581	1.1492
Chip 26	0.5108	1.9121	1.6634	0.9951	0.5583	1.149
Nominal Value	0.513	1.925	1.655	1	0.56	1.153
AVERAGE [V]	0.5105	1.908	1.663	0.9922	0.5569	1.148
MAX - MIN [mV]	2.2	8.8	3.4	6.3	3.2	5.2
STDEV [mV]	0.71	3.48	1.29	2.42	1.32	1.67
(MAX-MIN) / AVERAGE [%]	0.43	0.46	0.20	0.63	0.57	0.45

Table 9.3: Measurements of the on-chip references done on 10 APCs. For nominal value we mean the value simulated with typical process parameters.

The six columns are:

- Vref0: lower input reference for the ADC;
- Vref1: higher input reference for the ADC;
- Vbias: reference to generate the basic current in the DACs. The nominal value is simulated with Rext = 50 kΩ (R in Figure 9.3).
- DAC_REF_VDD_out: higher input reference for the DACs;
- DAC_REF_MID_out: lower input reference for the DACs;
- Vout_BG: bandgap output; the nominal value for the bandgap is calculated at 30 °C.

The table shows also, in the last rows: the simulated (nominal) value, the average, the maximum spread, the standard deviation and the relative maximum spread.

It can be seen that both the bandgap and all the references are extremely uniform. The maximum spread is less than 10 mV and the relative maximum spread less than 0.7% for all the references. Moreover, no variation is observed changing the power supply voltage (less than 1 mV for all the references for power supplies ranging from 2.25 V to 2.75V).

To check the temperature sensitivity of the references, the chip was heated from 20 °C to 70 °C. The shift of all the references is less than -10 mV, excepted Vbias which shifts of - 15 mV [New03].

9.2.2 The DACs

Figure 9.9 shows a DAC scan for all of the six 8-bit DACs present on a chip. Each curve represents a scan on a different chip. The measurements are done connecting the buffered output of the DAC to a multimeter. The behaviour of all the DACs is very similar, and very close to the simulated behaviour. The nonlinearity of DRHI for high values of its output voltage was predicted and could not be avoided (in effect the DACs were originally designed to stand a power supply voltage of 1.6 V).

Table 9.4 presents the DACs spread from chip to chip for a given code. The absolute spread is between 10 and 20 mV, and the relative spread of about 1%, for all the DACs except DRHI. For this DAC the absolute spread (47 mV) and the relative spread (1.8%) are higher than for the other DACs. This effect has not been explained yet and is still under investigation, but does not affect the overall functionality of the chip and the possibility to use it as it is in the two experiments, without a further redesign.

If a precision in the millivolt level is required at the DAC output, a calibration database will have to be foreseen, where the voltage output of each DAC for each input code is stored. The amount of data needing to be stored is negligible if compared with the calibration data that will be stored in the database to characterize the ALICE1LHCb chip.

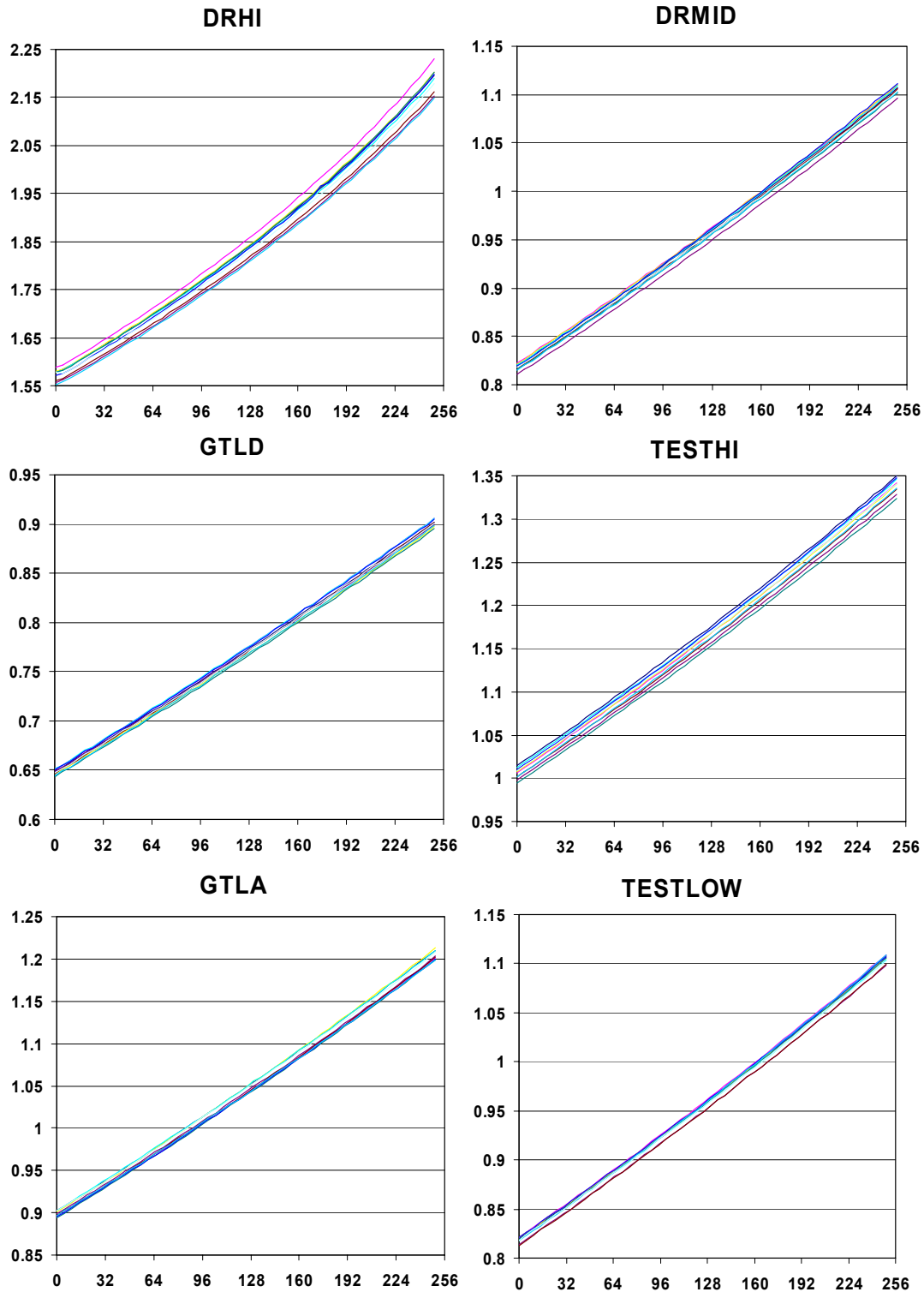

Figure 9.9: DAC scan for all of the six 8-bit DACs present on the chip; each curve represents a scan on a different chip. The measurements are done with a multimeter.

Code 115						
Chip #	DRHI [V]	DRMID [V]	GTLA [V]	GTLD [V]	TESTHI [V]	TESTLOW [V]
15	1.811	0.945	1.036	0.76	1.159	0.946
24	1.826	0.947	1.029	0.762	1.151	0.947
25	1.813	0.947	1.036	0.756	1.151	0.945
26	1.807	0.946	1.036	0.763	1.156	0.943
34	1.783	0.935	1.031	0.759	1.141	0.938
35	1.788	0.942	1.029	0.758	1.145	0.938
36	1.811	0.94	1.027	0.754	1.137	0.944
45	1.805	0.946	1.028	0.762	1.155	0.946
46	1.779	0.941	1.028	0.756	1.146	0.945
47	1.799	0.943	1.035	0.758	1.152	0.943
AVERAGE [V]	**1.802**	**0.943**	**1.031**	**0.7588**	**1.149**	**0.9435**
MAX-MIN [mV]	**47**	**12**	**9**	**9**	**22**	**9**
STDEV [mV]	**14.86**	**3.824**	**3.808**	**2.975**	**6.945**	**3.171**
MAX-MIN / AVERAGE [%]	**2.608**	**1.272**	**0.8725**	**1.186**	**1.914**	**0.9539**

Table 9.4: DACs spread from chip to chip (for DAC input code 115).

To check the driving capability of the buffers which drive the DACs outputs to the pads, test resistors were connected between the buffer output and ground. Figure 9.10 shows the measurements done on DRHI for three values of the load resistor, 10 kΩ, 30 kΩ and 50 kΩ (which result in 180 μA, 60 μA and 36 μA output current at 1.8 V). The curves are perfectly superimposed, showing that the buffer is capable of driving the load resistor without changes in the output curve. The DACs were also tested changing the chip temperature from 20 °C to 70 °C; for all of them the shift is less than 15 mV [New03].

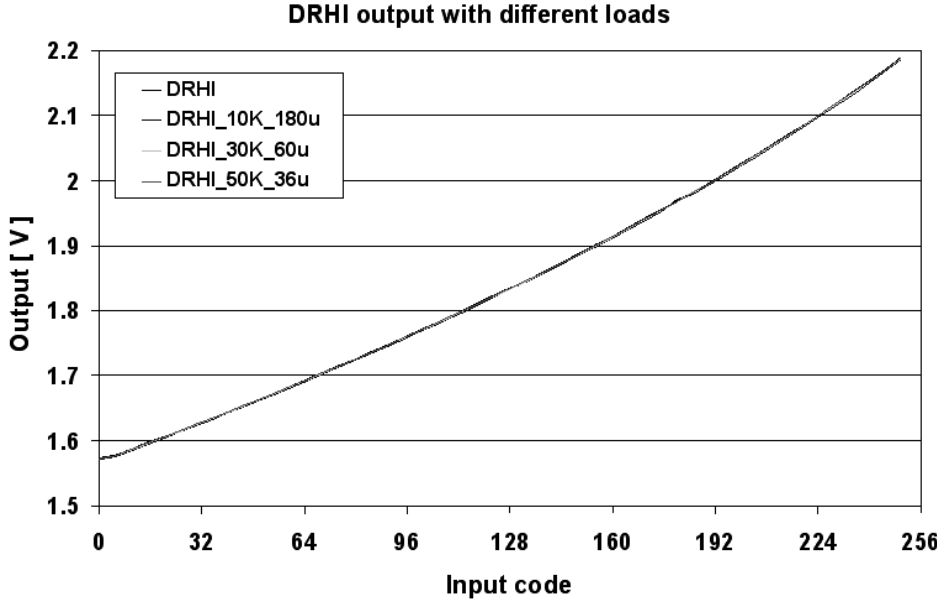

Figure 9.10: DAC scan of DRHI for three values of a test load resistor, 10 kΩ, 30 kΩ and 50 kΩ (which result in 180 μA, 60 μA and 36 μA output current at 1.8 V).

9.2.3 The ADC

The on chip ADC is fully functional. Figure 9.11shows two DAC scans done on DRHI and DRMID, both with the multimeter and with the internal ADC. The figure shows that the curves are coincident, except a small offset (in the order of some millivolts), which is due to the two buffers (one at the DAC output and one at the ADC input). The saturation effect seen for DRHI measured with the ADC is due to the fact that the signal exits the ADC dynamic range (at ~1.9 V).

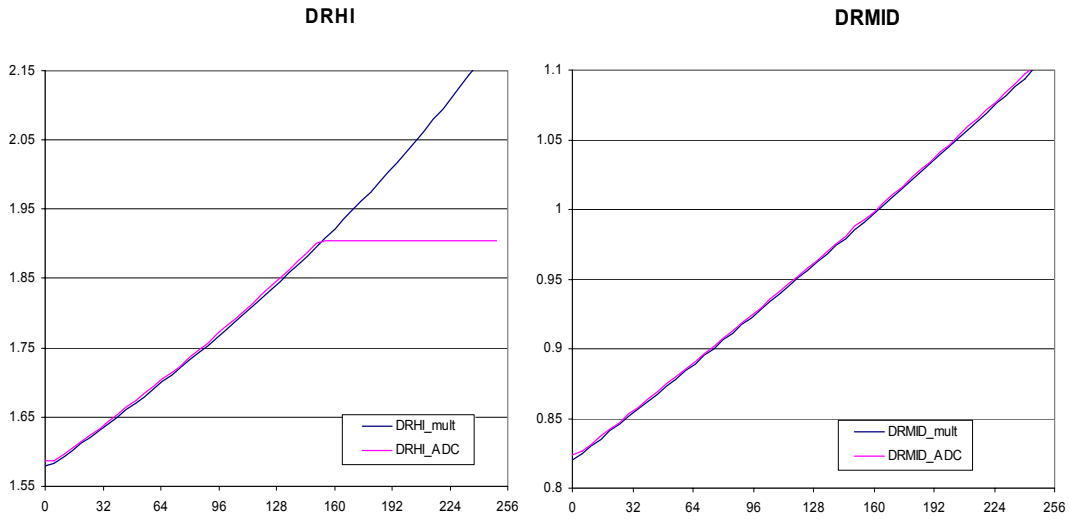

Figure 9.11: Two DAC scans done on DRHI (left) and DRMID (right), both with the multimeter and with the internal ADC. The saturation effect seen for DRHI measured with the ADC is due to the fact that the signal exits the ADC dynamic range (at 1.9 V).

Table 9.5 shows the differences between the DAC output values read with the internal ADC and with an external voltmeter. Each number is the average of the differences found for each code (scan).

Chip #	DRHI [mV]	DRMID [mV]	GTLA [mV]	GTLD [mV]	TESTHI [mV]	TESTLOW [mV]
15	-3.742	-1.765	-1.274	0.1961	-2.137	-2.020
24	-1.867	1.667	2.431	2.980	1.706	0.3922
25	-3.258	1.098	-0.7847	1.2158	-0.9216	-2.157
26	-4.5	-2.176	-1.490	-0.3529	-2	-2.6471
34	-5.588	-3.686	-2.765	-1.961	-4.412	-4.549
35	-4.697	-1.451	-1.529	-0.823	-2.941	-3.510
36	-4.968	1.0885E-14	0.01961	-0.4706	-2.020	-1.098
45	-1.875	1.6471	1.588	2.392	0.6078	-1.078
46	-3.323	-0.5098	0.05882	0.3529	-1	-2.863

Table 9.5: Differences between the DAC output values read with the internal ADC and with an external voltmeter. Each number is the average of the differences found for each code (scan).

By means of the ADC it could be possible to perform an auto-calibration of the DACs. To set an output value of the DAC the nominal code is set, and the DAC output read back with the ADC. The input code of the DAC is then changed, and the measurement procedure repeated, until the desired output value is reached. Table 9.5 shows that this can be done with an imprecision of some millivolts (up to 6, in case of DRHI), so that if a better precision is required a calibration database has to be foreseen in any case.

9.2.4 The current sensing stage

To test the current sensing stage a resistor R_{out} was connected in between a voltage source and the stage input as shown in Figure 9.12.

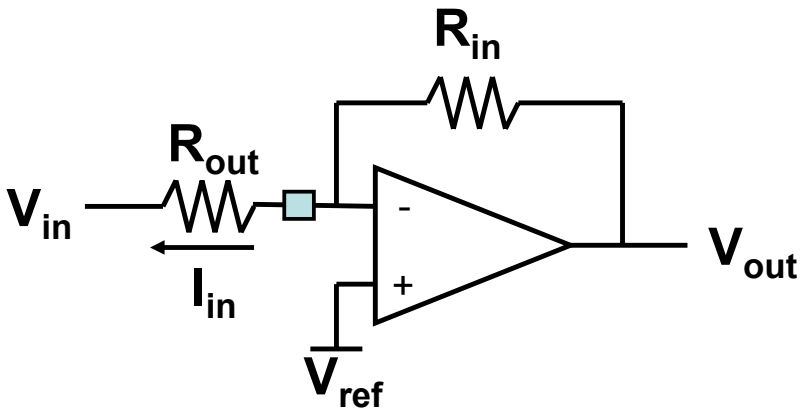

Figure 9.12: Schematic representation of the testing circuit used for the current sensing stage; the resistor R_{out} is connected in between a voltage source and the stage input to generate the input current.

The equation which gives the output voltage as a function of the input voltage is presented in equation (9.3), where V_{off} is the input offset of the operational amplifier.

$$V_{out} = (V_{ref} - V_{off})(1 + \frac{R_{out}}{R_{in}}) - V_{in}\frac{R_{in}}{R_{out}} \qquad (9.3)$$

Figure 9.13 shows the output voltage of the sensing stage as a function of the input current, both measured curve and calculated one (with $V_{off} = 3.2$ mV, $R_{in} = 8976\ \Omega$, $R_{out} = 3823\ \Omega$). The measured curve is taken with the on-chip ADC. The two curves are practically coincident.

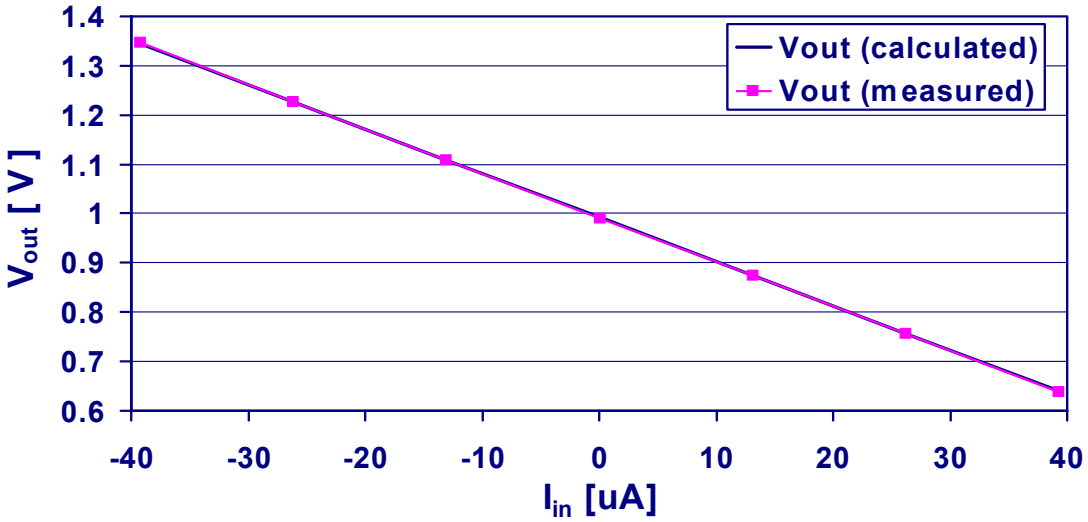

Figure 9.13: Output voltage of the current sensing stage as a function of the input current, both measured curve and calculated one (with Voff = 3.2 mV, Rin = 8976 Ω, Rout = 3823). Negative currents are sunk by the sensing circuit. The measured curve is taken with the on-chip ADC.

The precision of the current sensing stage is given by the absolute precision of the resistor R_{in}. This is an integrated resistor, whose value can vary (according to the technology design manual) up to 20-30%. For this reason a test resistor (identical to R_{in}) was designed and its terminals were accessible on two pads. The value of this resistor was measured for the 10 test chips, and the results are reported in Figure 9.14.

The results are much better than the design manual specifications. The average of the measured values differs from the design value by only 5%, and the standard deviation is 62 Ω (this means that the standard deviation divided by the average is 0.7%). It is important to point out that these are not resistors measured on the same chip, so we are measuring a chip-to-chip variation.

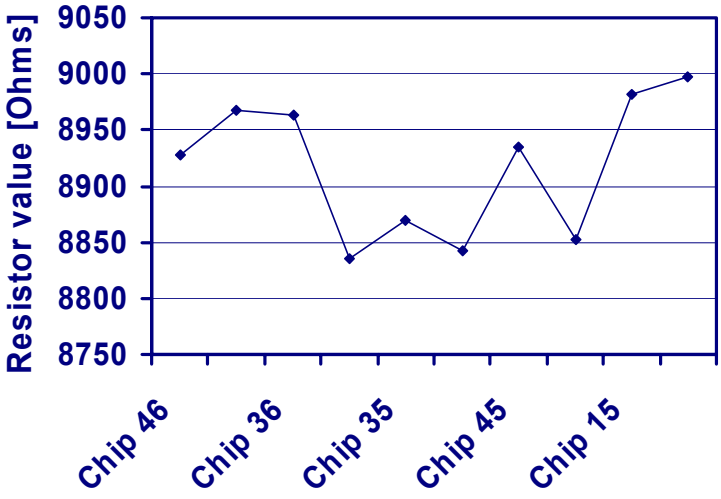

Figure 9.14: Measured value of a resistor identical to R_{in} for the 10 test chips.

9.2.5 The current sources for temperature measurement

The current sources output has been measured. To simulate different readings from a temperature sensitive resistor, resistances of known values were connected on the temperature sensing pins, and the voltage read back. Figure 9.15 shows the results for the LHCb current source (1.3 mA nominal value). The value of the current sources is almost insensitive to the load resistor, as expected. The average value of the current is 1.245 mA (this is -4.2% with respect to the nominal value), with a standard deviation of 4.5 μA [New03].

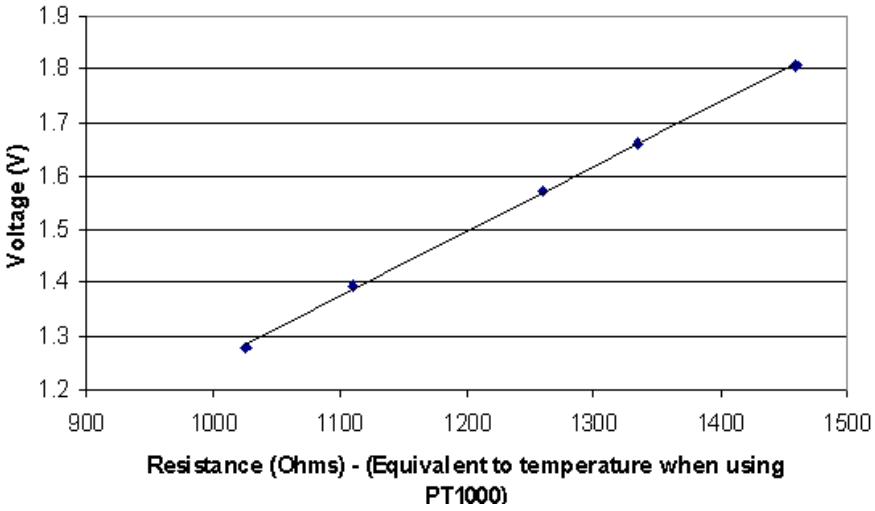

Figure 9.15: To simulate different readings from a PT1000, resistors of known resistances were connected on the temperature sensing pins, and the voltage read back [New03].

9.3 Summary

The ALICE1LHCb and LHCbpix1 Pixel Chips require a number of external auxiliary features: the digital control is demanded to the DPC, but the Pixel Chip needs also some external analogue biases. Moreover, the possibility to readout the on-chip DACs once the chip has been mounted on the On-detector pixel Pilot System (OPS) is useful, as well as the possibility to measure some other analogue parameters of the OPS (i.e. voltage supplies and temperature).These auxiliary analogue functions are performed by the Analogue Pilot Chip (APC).

The Analogue Pilot is a mixed-mode IC with a total power consumption of about 50 mW, designed with Hardening By Design (HBD) techniques to improve radiation tolerance. It contains several blocks needed to implement the required analogue features.

A prototype of the chip has been designed and tested, and a second slightly modified version has been redesigned and sent to fabrication. Tests done on the various blocks show the full functionality of the final version of the APC.

The **on-chip bandgap** cell has to provide a precise reference voltage, with reduced sensitivity to temperature and process parameters, that is used to generate the other on-chip references. Several reference circuits for the ADC and the 6 on-chip DACs have been implemented. Measurements on the bandgap and on the reference outputs show relative errors below 1%, and good temperature stability (less than 10-15 mV shift for a ΔT of 50 °C).

The **six 8-bit DACs** have to provide the reference voltages to the ALICE1LHCB or LHCBpix1 chip. The DACs design is a modified version of the design used for the Pixel Chip itself (voltage output version); this revised version has to stand a higher power supply voltage (2.5V), and the reference voltages for these DACs are provided on-chip. Moreover, a voltage buffer is connected at the output of each DAC to drive the DAC signal off the chip.

Measurements show that they are fully functional; a chip-to-chip spread (for a given code) up to some tens of millivolts is present. If a precision of the order of the millivolt is needed, this will require a calibration database, where the voltage output of each DAC for each input code is stored. The temperature sensitivity of the DACs is less than -15 mV for a ΔT of 50 °C.

The APC contains a **10-bit successive approximation ADC**, with very low power consumption (1 mW) and high modularity. The maximum sampling frequency is ~2.5 Msample/sec (for a clock of 30 MHz). In the APC it is used for the conversion of 16 DC or slow-varying signals, so it is underclocked, and the input range of the ADC is set between the 0.513 and 1.925. A 16-to-1 multiplexer and a buffer have been put in front of the ADC, to be able to read 16 different signals.

Measurements done on the ADC show that it is fully functional, and can precisely read back the input values, apart from a very small offset (in the order of some millivolts) introduced by an input buffer. This means that a self-calibration of the on chip DACs using the ADC can be performed, but that it can not precise at the millivolt level.

The **current sensing stage** has the aim of reading out the Pixel Chip current DAC outputs, avoiding changing the bias condition of the DAC output stage. Measurements show that the stage is capable to read both incoming and outgoing currents without an appreciable error.

The chip contains four **current sources for temperature monitoring**, designed to be independent from temperature and power supply variations. Their functionality was tested simulating the behaviour of a temperature sensitive resistor with resistances of known value; they are precise within some percent and are insensitive to the load resistor, as expected.

The **JTAG-controlled digital block** which has to provide all the necessary digital signals to the other blocks is fully functional.

All the blocks in the chip are working within the specifications, and the chip can be used in this version to equip the MCM of the OPS.

Conclusions

Silicon detectors offer compactness, high spatial resolution, high speed and direct and efficient conversion of the radiation energy. For this reason, in high energy physics they progressively replaced gas chambers, photographic films and other types of arrays, especially for what concerns tracking detectors. On the other hand, they need local readout electronics, which has to resist to the radiation. The luminosity at the CERN experiments has continuously increased with time, and so did the radiation levels.

A possible solution is provided by qualified radiation hard processes, normally only used for small production volumes and therefore expensive and prone to process stability, yield and availability problems. A completely different approach consists of using a standard commercial CMOS technology, hardened with special design techniques (Hardening By Design, HBD), studied in detail by the CERN RD49 project (for a summary see [Ane00]). The use of standard technology provides much lower cost and the possibility to always use the most advanced technologies.

This thesis deals with the design of the first very large pixel detector readout chip using HBD techniques and the development of the full pixel detector system for the ALICE experiment. The readout chip is also used for the RICH detector of the LHCb experiment, and the requirement to satisfy both applications lead to very stringent specifications with respect to power, speed, noise, channel-to-channel uniformity, and area. The chip contains 8192 pixel readout channels, about 13 million transistors, externally controllable biasing and testing circuitry.

Several test chips were designed, produced and tested prior to the completion of the full chip. A first test chip (ALICE1test) implemented in 0.5 μm resisted to 6-17 kGy depending on the radiation type, and demonstrated the technique, but showed the lack of density of this technology. A second test chip (ALICE2test) was implemented in 0.25 μm CMOS and withstood 260-480 kGy, total ionizing doses even beyond the requirements for LHC, but did not satisfy the very stringent front end requirements for the ALICE and LHCb experiments combined. This led to the development of the actual front end on the final chip, capable of dealing with the high rates in LHCb. The front-end implemented has a pole constellation with two complex conjugate poles and real pole, all with the same real part. This gives, in particular, a preamplifier with a very fast return to zero time (<100 ns @ 1%), which avoids pile-up in high multiplicity environments.

A first run yielded chips which could only be clocked at 15-18 MHz, after correction a second run yielded chips functional at a 40 MHz clock rate.

The chip was irradiated with 10 keV X-rays, up to a dose of 100 kGy (SiO_2), and did not show significant degradation. Irradiations with protons and heavy ions demonstrated a very good tolerance to Single Event Effects.

In conclusion, test results showed full functionality of the chip satisfying all specifications, and currently the chips have been produced and are being assembled for use in the experiments. The main results are summarized in the following table.

Parameter	Specification	Measured
Gain	28.8μV/e^- (sim. @ Q_{in}=5000e^-)	25μV/e^- (@ Q_{in}=5000e^-)
Power consumption (8192 channels)	< 1 W/chip	< 1 W/chip
Noise	< 200 e^- rms	~170 e^- rms
Threshold uniformity (σ)	200 e^-	better than 150 e^- (unadjusted)
Maximum occupancy (@40MHz)	8%	>8%
Peaking time	25 ns	~28 ns
Return to zero time	150-200 ns	<150 ns (preamp:<100ns)
Clock frequency	10 MHz and 40 MHz	~45 MHz
Radiation tolerance	> 5 kGy	>100 kGy
Individual pixel threshold adjust, mask and test	Yes	Yes
Fast recovery for large signals	Yes	<300ns @ Q_{in}=100ke^-

Table 1: Specifications and measurements of the main chip parameters (Q_{in} is the input charge).

In addition, the silicon detector to which the readout chip is connected using flip-chip soldering techniques, was designed. Electrical tests (e.g. leakage current and depletion voltage) were carried out before and after irradiation, and demonstrated that the detector performs within specifications for a radiation dose three to six times beyond the expected for the ALICE experiment.

An auxiliary chip (the APC) which generates analogue biases and measures slow-varying analogue signals was also designed and tested, and is fully functional.

This project marks the change for radiation detector readout electronics from dedicated radiation hard technologies to standard commercial deep-submicron CMOS technologies using HBD techniques. The results in this thesis demonstrate that it was possible to design a complex mixed-mode 13-million transistor chip which satisfies two experiments with different specifications, and has demonstrated full functionality when integrated in the full system. The success of this HBD approach, developed primarily in the MIC group at CERN, was such that currently most electronics for LHC with stringent radiation tolerance requirements is being produced using this approach.

Contribution of this work

The pixel detector project has been a team effort, in which I participated. Its success and many of the results published here have been obtained through the efforts of many people. To clarify the contribution of this thesis work in this large project, the tasks in which I played a significant role or for which I was responsible entirely are listed below:

Sensor (Chapter 4)

Design of:

- The singles (LHCb, NA60, ALICE)
- The ladders for ALICE (5 chips)
- The ladders for RIKEN (4 chips)
- Data analysis of the irradiation tests

ALICE1TEST (Chapter 6)

- Design of the discriminator
- Layout of several chip blocks (feedback and injection capacitances, pads…)

ALICE1LHCb (Chapter 7)

Design of:

- The front-end (preamplifier, shapers, feedback stages)
- The discriminator

LHCbPix1 (Chapter 7, 8)

- Design of the DACs
- Full chip layout and verification

Tests of both chips (Chapter 8)

- The front-end and discriminator
- The DACs
- Source tests
- Test beams

Analogue Pilot Chip, prototype and final version (Chapter 9)

Design of:

- The DACs
- The references
- Full chip layout and verification
- Tests

I also participated closely with many other tasks:

SPD system (Chapter 3)

- Development and testing of the test system hardware
- Development and testing of the test system software
- Development of the MCM module
- I presented the work of the SPD group at the LECC 2002 conference

Sensor (Chapter 4)

- Other sensor tests
- Concept of a software model to simulate the detector response (not presented in this thesis)

ALICE1TEST (Chapter 6)

- Full chip layout
- I presented the work done by the MIC design group in the general "ALICE Italia" annual meeting in Bologna (Italy) in 2000

ALICE1LHCb and LHCbPix1 (Chapter 8)

- Other testing phases (from IC testing to testbeams)

Other R&D chips (not presented in this thesis)

- Concept and design of the MACROPAD chip. It is a hybrid pixel chip for testing the feasibility of using amorphous silicon on ASIC for High Energy Physics (no test data available while writing the thesis)
- Concept of the future MEDIPIX3 chip. It is a hybrid pixel chip for x-ray imaging.

Appendix I
Noise modelling in MOS transistors

Noise sets a lower limit to the accuracy of any measurements and to the amplitude of signal that can be processed electronically. There are several noise sources in MOS transistors; the most important are:

- thermal noise in the drain-source channel,
- flicker noise in the drain-source channel,
- gate resistance thermal noise,
- bulk resistance thermal noise,
- shot noise due to the leakage current through the SiO_2 gate,
- shot noise associated with the leakage current of the drain (source) reverse biased diodes,
- thermal noise due to the source, drain and gate contact resistance.

In most cases only the first two items are important.

I.1 Thermal noise

Every resistive element generates thermal noise which is caused by random motion of charge carriers in it. A noisy resistor can be modelled as a noiseless resistance R with a series noise voltage generator of a power spectral density given by:

$$\frac{\overline{v}_R^2}{\Delta f} = 4\,k\,T\,R \qquad \text{(I.1)}$$

where k is the Boltzman constant and T is absolute temperature.

I.1.1 Channel thermal noise

The MOS transistor has a resistive channel between the drain and the source, and the random thermal motion of carriers in the channel results in thermal noise. This noise depends on bias conditions, transistor dimensions and properties of a given technology.

For the first hand noise calculation the thermal channel noise in strong inversion in the saturation region can be expressed as (taking into account also g_{mb} [Enz01]):

$$S_{Id} = 4\,k\,T\,(g_m+g_{mb}) \tag{I.2}$$

where S_{Id} is the drain current spectral density of the MOS input device, g_m is the gate-to-source device transconductance and g_{mb} is the bulk-to-source device transconductance. If we assume zero source impedance, we can find the noise voltage spectral density referred to the input:

$$S_{Vd} = 4\,k\,T\,n\frac{2}{3}\frac{1}{g_m} \tag{I.3}$$

where n is the slope factor defined in section 5.2.2.

It is then necessary to use a simple model which could be employed going from strong to weak inversion, and that can provide reasonably precise results even for simple symbolic calculations. This is why we will use for our calculations the EKV model, which is able to model all the transistor parameters even in moderate inversion, with good continuity of the model parameters in this transition region. A description of this model is presented in [Enz95, Enz01].

In weak inversion the channel thermal noise (as power spectral density of the drain current) can be expressed as [Tsi99]:

$$S_{Id} = 2q\,I_D \tag{I.4}$$

and in weak inversion the transconductance g_m is:

$$g_m = \frac{I_D}{n\,U_T} \tag{I.5}$$

where U_T is the thermal voltage $U_T = kT/q$. We can derive again the input referred channel thermal noise dividing by g_m^2:

$$S_{Vd} = 4\,K\,T\,n\frac{1}{2}\frac{1}{g_m} \tag{I.6}$$

Comparing equations (I.6) and (I.3) we can compact the two formulas in one:

$$S_{Vd} = 4\,K\,T\,n\,\frac{\gamma}{g_m} \tag{I.7}$$

where γ varies from ½ to ⅔ from weak to strong inversion for an ideal device.

For a real device measured noise is often higher than what can be predicted using equation (I.7). This is usually taken into account adding the *noise excess factor* Γ in equation (I.8), so that γ varies from ½ Γ to ⅔ Γ from weak to strong inversion for a real device.

Several functions have been proposed to interpolate γ from weak to strong inversion. In particular Enz proposes to use the following [Enz01]:

$$\gamma = \frac{\frac{1}{2} + \frac{2}{3} i_f}{1 + i_f} \Gamma \tag{I.8}$$

where i_f is the inversion factor defined in section 5.5.3.3.

For our target technology results from measurements of noise in n- and p-channel devices in all working regions are reported in Chapter 5. In particular, the white noise excess factor as a function of the gate length for devices in weak (w.i.), moderate (m.i.) and strong (s.i.) inversion are reported in Figure 5.27. It can be seen that neglecting the coefficient Γ can lead to errors as large as a factor 3.5 (for short n-channel devices in strong inversion).

I.1.2 Bulk resistance thermal noise

The three dimensional distributed substrate resistance R_B introduces a noise component which, referred to the input, can be expressed as:

$$\frac{\overline{v}_{Rb}^2}{\Delta f} = \frac{4\,k\,T\,R_B\,g_{mb}^2}{g_m^2} \tag{I.9}$$

I.1.3 Gate resistance thermal noise

The polysilicon gate resistance R_g introduces an additional thermal noise source; the resistance of the interconnections can also be included in this noise source. This noise component is directly present at the input and is expressed by:

$$\frac{\overline{v}_{Rg}^2}{\Delta f} = 4\,k\,T\,R_g \tag{I.10}$$

The resistance R_G can be calculated from layout inspection, as for example is shown in [Cha91 (pp. 22-25)].

I.2 Flicker noise

Another source of noise in MOS devices is the flicker noise, also called 1/f noise. There are two main theories which explain the physical origin of the flicker noise. According to the *McWorther model* [Mcw56], noise is caused by the random trapping and detrapping of the mobile carriers in the channel and within the gate oxide. According to that theory the flicker noise can be modelled by a voltage noise generator in series with the transistor gate with the voltage noise power spectral density given by the formula [Cha91, p.20]:

$$\frac{\overline{v}_{1/f}^2}{\Delta f} = \frac{K_f}{C_{ox}^2 W L f^{\alpha}} \tag{I.11}$$

where K_f (also called K_a) is a technology dependent constant (but independent of bias condition); α is a constant close to unity (varies in the narrow range of 0.7 to 1.2) and f is the f frequency. The formula is valid in all transistor inversion regions.

In the other approach, called *mobility fluctuation model* [Hoo81], the flicker noise is attributed to the mobility fluctuation. The electrons (holes) are scattered by phonons of lattice vibrations and the density of phonons fluctuates with a 1/f spectrum.

I.3 Shot noise

Every reverse biased junction generates shot noise which is caused by the random passage of carriers across the junction. If the current is composed by a series of random independent pulses with average value I_D then the resulting noise current has a spectral density given by [Gra84, p636]:

$$\frac{\overline{i}_R^2}{\Delta f} = 2\,qI_D \tag{I.12}$$

I.3.1 Gate shot noise

The current flowing through the SiO_2 gate generates a shot noise given by:

$$\frac{\overline{i}_g^2}{\Delta f} = 2\,qI_G \tag{I.13}$$

I.3.2 Bulk shot noise

The noise associated with the leakage current of the drain (source) reverse biased diodes, generates a shot noise given by:

$$\frac{\bar{i}_g^2}{\Delta f} = 2\,qI_B \tag{I.14}$$

Both gate and bulk shot noise have a contribution to the total noise which can usually be neglected.

I.3.3 External shot noise sources

An external source of shot noise which is very important in our case is the noise generated by the leakage current I_o of the detector, which is connected in parallel with the transistor input. Its noise contribution is given by:

$$\frac{\bar{i}_o^2}{\Delta f} = 2\,qI_o \tag{I.15}$$

All the noise contributions which generate physical noise current in parallel with the gate of the MOS transistor can be added in squares (if they are independent) in a noise component $\frac{\bar{i}_{par}^2}{\Delta f}$; this is also called "parallel noise"

I.4 Total input referred noise

In general, the noise performance of any two port network can be represented by two equivalent input noise generators i_i^2 and v_i^2. A simplified small signal model with noise sources of a MOS transistor is shown in Figure I.1 (top), and its equivalent representation in terms of equivalent input noise generators is shown in Figure I.1 (bottom).

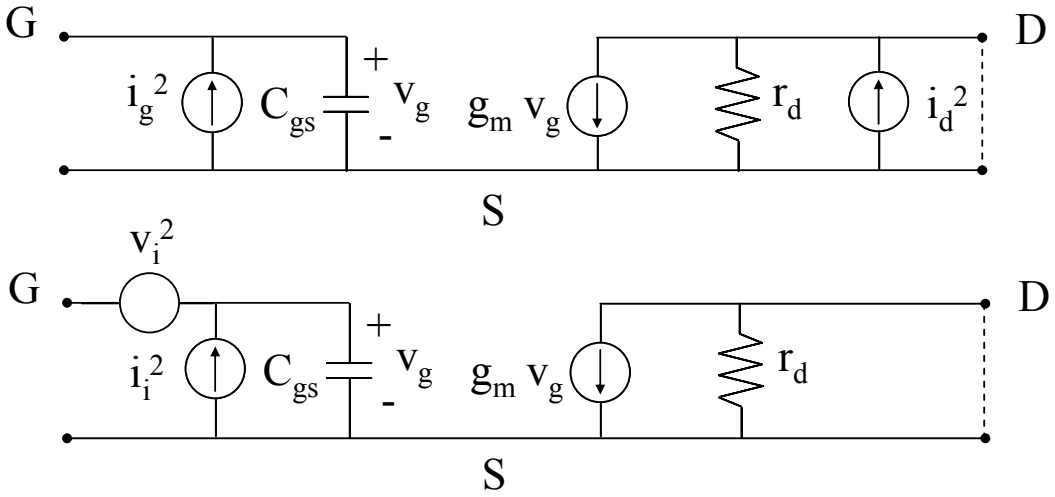

Figure I.1: MOS model with noise sources (top) and its equivalent representation with a voltage and a current input noise sources (bottom).

It can be shown that [Cha91]:

$$\frac{\overline{i}_i^2}{\Delta f} = \frac{\overline{v}_i^2}{\Delta f}\left|j\,\omega\,C_{in}\right|^2 + \frac{\overline{i}_{par}^2}{\Delta f} \tag{I.16}$$

while the expression for v_i^2 is the following, given by the sum in squares of all the noise contributions presented in the previous sections:

$$\frac{\overline{v}_i^2}{\Delta f} = \frac{4\,k\,T\,n\,\gamma}{g_m} + \frac{K_f}{C_{ox}^2\,W\,L\,f^{\alpha}} + \frac{4\,k\,T\,R_B\,g_{mb}^2}{g_m^2} + 4\,k\,T\,R_g \tag{I.17}$$

For the calculations carried out in Chapter 7 only the first two contributions are taken into account.

Appendix II
Most important abbreviations and acronyms

ALICE	A Large Ion Collider Experiment
APC	Analogue Pilot Chip
ATLAS	A Toroidal LHC ApparatuS
CAD	Computer Aided Design
CMOS	Complementary Metal Oxide Semiconductor
CMRR	Common Mode Rejection Ratio
CMS	Compact Muon Solenoid
COTS	Components Off The Shelf
CSA	Charge Sensitive Amplifier
DAC	Digital to Analogue Converter
DICE	Dual Interlock storage Cell
DPC	Digital Pilot Chip
ECC	Error Control Codes
ECT	Error Correction Technique
EDAC	Error Detection And Correction
EEPROM	Electrically Erasable Programmable Read Only Memory
FET	Field Effect Transistor
FMD	Forward Multiplicity Detector
FOX	Field Oxide
GOL	Gigabit Optical Link
GTL	Gunning Transceiver Logic
HBD	Hardening By Design
HEP	High Energy Physics
HPD	Hybrid Photon Detector
IC	Integrated Circuit
ISR	Intersecting Storage Rings
ITS	Inner Tracking System
KGD	Known Good Die
LEP collider	Large Electron Positron
LET	Linear Energy Transfer
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty experiment
LOCOS	LOCal Oxidation of Silicon
LVS	Layout Versus Schematic
MBU	Multiple Bit Upset
MCM	Multi Chip Module
MIP	Minimum Ionising Particle
MOS	Metal Oxide Semiconductor
MPC	Multiwire Proportional Chamber
NIEL	Non-Ionising Energy Loss
NMOS	Negative (n-type) Metal Oxide Semiconductor
NSREC	Nuclear and Space Radiation Effects Conference
OPS	On-detector pixel Pilot System

OTA	Operational Transconductance Amplifier
PBL	Polysilicon Buffered Locos
PCB	Printed Circuit Board
PHOS	Photon Spectrometer
PID	Particle IDentification
PMOS	Positive (p-type) Metal Oxide Semiconductor
PS	Proton Synchrotron
PSRR	Power Supply Rejection Ratio
PTS	Pixel Test System
RAM	Random Access Memory
RPC	Resistive Plate Chambers
SEB-SEBO	Single Event Burn Out
SEE	Single Event Effects
SEFI	Single Event Function interruption
SEGR	Single Event Gate Rupture
SEL	Single Event Latchup
SEM	Scanning Electron Microscope
SER	Soft Error Rate
SES	Single Event Snapback
SEU	Single Event Upset
SET	Single Event Transient
SG	Semi Gaussian
SHE	Single Hard Error
SMD	Surface Mount Devices
SOI	Silicon On Insulator
SPS	Super Proton Synchrotron
SPD	Silicon Pixel Detector
SRAM	Static Random Access Memory
STI	Shallow Trench Isolation
TDR	Technical Design Report
TID	Total Ionising Dose
TMR	Triple Modular Redundancy
TOF	Time Of Flight
TOTEM	Total Cross Section, Elastic Scattering and Diffraction Dissociation at the LHC
TPC	Time Projection Chamber
VLSI	Very Large Scale Integration
ZDC	Zero-Degree Calorimeter

Bibliography

[Add01] Canberra Semiconductor, N.V., Lammerdries 25, 2250 Olen, Belgium; http://www.canberra.com/ ;

[Add02] VTT Electronics, Kaitoväylä 1, P.O.Box 1101, FIN-90571 VTT, Oulu, Finland; http://www.vtt.fi/indexe.htm ;

[Add03] Alenia Marconi Systems (AMS) Via Tiburtina km 12,400, 00131 ROMA, Italy; http://www.amsjv.com/ ;

[Ale96] D.R. Alexander, "Design issues for radiation tolerant microcircuits for space", *1996 IEEE Nuclear and radiation effects conference short course*, Indiana Wells, CA, July 1996;

[Ale99] M. Alemi et al., "Performance of a hybrid photon detector prototype with electrostatic cross-focussing and integrated silicon pixel readout for Cherenkov ring detection", 6th International Conference on Advanced Technology and Particle Physics, Como, Italy, October 1998; published in: *Nuclear Physics B - Proceedings Supplements*, Volume 78, Issues 1-3, August 1999, Pages 360-365;

[ALI99] ALICE Technical Design Report of the Inner Tracking System (ITS), CERN/LHCC 99-12, ALICE TDR 4, 18 June 1999;

[All95] M. Allenspach et al., "SEGR in power MOSFET: prediction of breakdown and evaluation of oxide thickness dependence", *IEEE Transactions on Nuclear Science*, vol. NS-42, no. 6, December 1995, pp. 2611-2616;

[Ane97] G. Anelli, M. Campbell, F.Faccio et al., "Total dose behavior of submicron and deep submicron CMOS technologies", *Proceedings of the Third Workshop on Electronics for LHC Experiments* (LEB97, London, England, September 1997), pp. 139-143;

[Ane99] G. Anelli, M. Campbell, M. Delmastro et al., "Radiation Tolerant VLSI Circuits in Standard Deep Submicron CMOS Technologies for the LHC Experiments: Practical Design Aspects", presented at the IEEE Nuclear and Space Radiation Effects Conference (NSREC '99, Norfolk, Virginia, USA, July 1999), IEEE Transactions on Nuclear Science, vol. 46, no. 6, December 1999, pp. 1690-1696;

[Ane00] G. Anelli, "Conception et caractérisation de circuit intègres résistants aux radiations pour les détecteurs de particules du LHC en technologies CMS submicroniques profondes", Doctoral thesis, Institut National Polytechnique de Grenoble, December 2000;

[Ane01] G. Anelli, F. Faccio, S. Florian and P. Jarron, "Noise characterization of a 0.25 μm CMOS technology for the LHC experiments", *Nuclear Instruments and Methods,* A Vol. 457, Issues 1-2, January 2001, pp. 361-368;

[Ane02] G. Anelli, R. Dinapoli, A. Kluge: "CERN_Analog_pilot: Chip description and pad list", Internal document (CERN), May 2002;

[Ang92] F. Anghinolfi et al., "A 1006 element hybrid silicon pixel detector with strobed binary readout", *IEEE Transactions on Nuclear Science*, Vol. 39 (1992) p.654;

[Bar93] E. Barberis, J.G. Boissevain, N. Cartiglia et al., "Temperature effects on radiation damage to silicon detectors", *Nuclear Instruments and Methods*, A Volume 326, pp. 373-380, 1993;

[Baz97] M.P. Baze and S.P. Buchner, "Attenuation of Single Events Induced Pulses in CMOS combinatorial logic", *IEEE Transactions on Nuclear Science*, Vol. 44 no.6, pp.2217-2223, December 1997;

[Bei88] B. A. Beitman, "N-channel MOSFET breakdown characteristics and modeling for p-well technologies", *IEEE Transactions on Electron Devices*, vol. 35, no. 11, November 1988, pp. 1935-1941;

[Ben85] J.M. Benedetto, H.E. Boesch, Jr. et al. "Hole Removal in Thin-Gate Mosfets by Tunneling", *IEEE Transactions on Nuclear Science*, vol. 32, no. 6, December 1985, pp. 3916-3920;

[Bib98] J.Bibby, T.Gys, "Support document for LHC-b rich front-end electronics", LHCB/98-036, CERN February 1998;

[Bis01] D.Bisello, J. Wyss, A. Candelori, A. Kaminsky, D. Pantano, "Radiation damage of standard and oxygenated silicon diodes irradiated by 16-MeV and 27-MeV protons", *IEEE Transactions on Nuclear Science*, Vol. 48, No. 4, August 2001, pp. 1020-1027;

[Bla66] J.L. Blankenship and C.H. Nowlin, "New concepts in nuclear pulse amplifier design", *IEEE Transactions on Nuclear Science*, vol. 13, no. 3, 1966, pp. 495-507;

[Boe76] H.E. Boesch and J.M. McGarrity, "Charge Yield and Dose Effects in MOS Capacitors at 80 K", *IEEE Transactions on Nuclear Science*, vol. 23, no. 6, December 1976, pp. 1520-1525;

[Boe85] H.R. Boesch, Jr. and F.B. McLean, "Hole transport and trapping in field oxides", *IEEE Transactions on Nuclear Science*, vol. 32, no. 6, December 1985, pp. 3940-3945;

[Boe86] H.E. Boesch et al., "Saturation of Threshold Voltage Shift in MOSFET's at High Total Dose", *IEEE Transactions on Nuclear Science*, vol. NS-33, no. 6, December 1986, pp. 1191-1197;

[Bri96] C. Brisset, V. Ferlet-Carvoi et al., “Two-dimensional simulation of total dose effects on NMOSFETS with lateral parasitic transistor”, *IEEE Transactions on Nuclear Science,* vol.43, no.6, pp. 2651-2658, December 1996;

[Bur01] M. Burns, P. Chochula, R. Dinapoli et al., “The ALICE Pixel Detector Readout Chip Test System”, *Proceedings of the 7th workshop on electronics for LHC experiments*, (LEB 2001, Stockholm, Sweden, 10-14 Sept. 2001), pp.71-75, CERN-2001-005 ; CERN-LHCC-2001-034;

[Cal96] T. Calin, M. Nicolaidis, R. Velazco, "Upset Hardened Memory Design for Submicron CMOS Technology", *IEEE Transactions on Nuclear Science*, vol.43, no. 6, December 1996, pp. 2874-2878;

[Cam90] M. Campbell et al., "A 10 MHz CMOS front-end for direct readout of pixel detectors", *Nuclear Instruments and Methods,* Vol. A 290 (1990) p.149;

[Cam94] M. Campbell et al., “Development of a pixel readout chip compatible with large area coverage”, *Nuclear Instruments and Methods,* A Vol. 342, no.1, p.52, March 1994;

[Cam99] M. Campbell, G. Anelli, E. Cantatore, R. Dinapoli et al., “A pixel readout chip for 10-30 Mrad in standard 0.25 μm CMOS”, 1998 IEEE Nuclear Science Symposium and Medical Imaging Conference, Toronto, Ontario, Canada, 8 - 14 Nov 1998 - IEEE, Piscataway, NJ, 1999. - pp.827-31, published in *IEEE Transactions on Nuclear Science,* Vol.NS-46, n.3, June 1999, pp. 156-160;

[Cam01] M. Campbell, G. Anelli, E. Cantatore et al., “An introduction to deep submicron CMOS for vertex applications”, *Nuclear Instruments and Methods*, A Volume 473, Issues 1-2, 1 November 2001, Pages 140-145;

[Can98] E. Cantatore et al., “Uniformity of response of the LHC1 large area pixel detector system”, *Nuclear Instruments and Methods*, Volume 409, Issues 1-3, 21 May 1998, pp. 119-128;

[Can01] A. Candelori, R. Rando, D. Bisello et al., “Low- and high-energy proton irradiation of standard and oxygenated silicon diodes”, *IEEE Transactions on Nuclear Science,* Vol.NS-48, n.6, December 2001, pp. 2270-2277;

[Can03] A. Candelori, R. Rando, D. Bisello et al., “New evidence of dominant processing effects in standard and oxygenated silicon diodes after neutron irradiation”, Nuclear Instruments and Methods A, Vol. 512 (2003), pp. 52–59;

[Cas03] M. E. Castro Carballo, “Silicon pixel design and pixel chip testing for the RICH of the LHCb experiment at CERN”, final project, “University of Santiago de Compostela, Faculty of Physics, February 2003;

[Cha90] Z.Y. Chang and W.M.C. Sansen, "Low-noise, low-distorsion CMOS AM wide-band amplifiers matching a capacitive source", *IEEE Journal of solid state circuits*, Vol25, no.3, June1990;

[Cha91] Z.Y. Chang and W.M.C. Sansen, "Low-noise wide-band amplifiers in bipolar and CMOS technologies", Kluwer Academic Publishers, 1991;

[Cho03] P. Chochula, M. Burns, R. Dinapoli et al.: "The ALICE silicon pixel detector", Quark Matter 2002, 16th Int. Conf. on Ultra-Relativistic Nucleus-Nucleus Collisions, published in: *Nuclear Physics* A Volume 715, 10 March 2003, Pages 849-852;

[Chr99] J. Christiansen, "Requirements to the L0 front-end electronics", LHCb Note 99-029 Electronics (CERN), 1999;

[Cor00] F. Corsi, R. Dinapoli, P. Lamanna, C. Marzocca, "Design and Characterization of a DAC for the Slow Control of the Pixel Chip", *Proceedings of the sixth Workshop Electronics for LHC Experiment* (LEB00, Cracow, Poland, September 2000), p.560-564;

[Dam83] C.J.S. Damerell et al., "First measurements of efficiency and precision of CCD detectors for high energy physics", *Nuclear Instruments and Methods,* A Vol. 213 (1983), p.201-215;

[Deb02] W. de Boer et al., "Measurements with a CMOS pixel sensor in magnetic fields", *Nuclear Instruments and Methods,* A Vol. 487 (2002) pp.163-169;

[Del92] P. Delpierre et al.: "Development of silicon micropattern (pixel) detectors", *Nuclear Instruments and Methods,* A Vol. 315 (1992) pp.133-138;

[Des03] M. Despeisse, G. Anelli, S. Commichauet al., "Characterization of 13 μm and 30 μm thick hydrogenated amorphous silicon diodes deposited over CMOS integrated circuits for particle detection application", to be published in the Proceedings of the 9th Pisa Meeting on Advanced Detectors, La Biodola, Isola d'Elba, May 2003 as a special *Nuclear Instruments and Methods,* A issue, in press, uncorrected proof (already available online);

[Die82] S.E. Diehl et al., "Error analysis and prevention of cosmic ion induced errors in CMOS RAM", *IEEE Transactions on Nuclear Science*, vol.29, no. 6, December 1982, pp. 2032-2039;

[Din77a] A.G.F. Dingwall and R. E. Stricker, "C2L: A New High-Speed High-Density Bulk CMOS Technology", *IEEE Journal of Solid-State Circuits*, vol. 12, August 1977, pp. 344-349;

[Din77b] A.G.F. Dingwall, R. E. Stricker and J. O. Sinniger, "A High Speed Bulk CMOS C2L Microprocessor", *IEEE Journal of Solid-State Circuits*, vol. 12, October 1977, pp. 457-462;

[Din00] R. Dinapoli, M. Campbell, W. Snoeys, K.Wyllie et al., "An analog front-end in standard 0.25μm CMOS for silicon pixel detectors in ALICE and LHCb", *6th Workshop on Electronics for LHC Experiments*, Cracow, Poland, 11 - 15 Sep 2000

- CERN, Geneva, 25 Oct 2000. [CERN-2000-010 ; CERN-LHCC-2000-041 - pp.110];

[Din01] R. Dinapoli, M. Campbell, W. Snoeys, K.Wyllie et al., "A front-end for silicon pixel detectors in ALICE and LHCb"; *Nuclear Instruments and Methods in Physics Research,* Section A, Volume 461, Issue 1-3, April 2001, Pages 492-495;

[Din02] R. Dinapoli, A. Kluge, M. Morel et al., "The Alice Silicon Pixel Detector Readout System – Moving towards system integration", *8th Workshop on Electronics for LHC Experiments*, Colmar, France, 9 - 13 Sep 2002, pp.164-169;

[Din03] R. Dinapoli, "A readout front-end for pixel detectors in the ALICE experiment at CERN, with dosimetry applications", proceedings of the JDOCS 2003, Montpellier, March 2003;

[Din04] R. Dinapoli, "Challenges and experiments at the LHC", http://dinapoli.home.cern.ch/dinapoli/Tesi/addon1.pdf

[Din04a] R. Dinapoli, "Semiconductor detectors: basic properties", http://dinapoli.home.cern.ch/dinapoli/Tesi/addon2.pdf

[Duf92] C. Dufour, et al., "Heavy Ion induced Single Hard Error on sub micronic memories", *IEEE Transactions on Nuclear Science*, vol. NS-39, no. 6, December 1992, pp. 1693-1697;

[Enz95] C. Enz, F. Krummenacher, E. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications", Analog Integrated Circuits and Signal Processing, 8, 83-114, 1995;

[Enz01] C. Enz, E. Vittoz, "Advanced engineering course on low-voltage analog CMOS IC design", part I, Lausanne, Switzerland, July 2001;

[Fac98] F. Faccio, G. Anelli, M. Campbell et al., "Total dose and single event effects (SEE) in a 0.25 μm CMOS technology", *Proceedings of the 4th Workshop on Electronics for LHC Experiments* (LEB98, Rome, Italy, Sept. 1998), pp. 105-113;

[Fac98a] F. Faccio, C. Detcheverry, M. Huhtinen, "First evaluation of the Single Event Upset (SEU) risk for electronics in the CMS experiment", CERN CMS note 1998/054, 1998;

[Fac99] F. Faccio, K. Kloukinas et al. "SEU effects in registers and in a Dual-Ported Static RAM designed in a 0.25 μm CMOS technology for applications in the LHC", *Proceedings of the Fifth Workshop on Electronics for LHC Experiments* (LEB99, Snowmass, Colorado, USA, September 1999), pp. 571-575;

[Fac99a] F. Faccio, A. Marchioro, T. Calin, M. Nicolaidis, R. Velazco et al., "Single Event Effects in Static and Dynamic Registers in a 0.25 μm CMOS Technology", *IEEE Transactions on Nuclear Science*, vol. 46, December 1999, pp. 1434-1439;

[For98] R. Forty and O. Schneider, "RICH Pattern Recognition", LHCb Note 98-040 RICH, CERN 1998;

[Gat84] E.Gatti and P. Rehak, “Silicon drift chambers — an application of a novel charge transport scheme”, *Nuclear Instruments and Methods,* A vol. 225, n.3, September 1984, pp. 608-614;

[Gir98] A. Giraldo, "Evaluation of Deep Submicron Technologies with Radiation Tolerant Layout for Electronics in LHC Environments", Ph.D. Thesis at the University of Padova, Italy, December 1998;

[Gir00] A. Giraldo, A. Paccagnella, A. Minzoni, "Aspect ratio calculation in n-channel MOSFETs with a gate-enclosed layout", *Solid-State Electronics*, vol. 44, 1st June 2000, pp. 981-989;

[Gor02] Y.Gornushkin et al., “Test results of monolithic active pixel sensors for charged particle tracking”, *Nuclear Instruments and Methods*, A Vol. 478 (2002) pp.311-315;

[Gra84] P.R. Gray, R.G. Meyer, “Analysis and design of analog integrated circuits, second edition”; John Wiley & Sons, New York 1984;

[Gra99] T. Grassi, “Development of the digital read-out system for the CERN Alice pixel detector”, University of Padova – Department of Electronics and Computer Engineering (DEI), Doctoral Thesis, December 31, 1999;

[Gri82] P. Grignoux and R.L. Geiger, "Modeling of MOS Transistors with Nonrectangular-Gate Geometries", *IEEE Transactions on Electron Devices*, vol. 29, no. 8, August 1982, pp. 1261-1269;

[Gun92] B. Gunning et al., “A CMOS Low Voltage-Swing Transmission-Line Transceiver”, *ISSCC Dig. Of Tech. Papers*, February 1992, p.58;

[Gys95] T. Gys, C. D'Ambrosio, H. Leutz et al.: “A new position-sensitive photon detector based on an imaging silicon pixel array (ISPA-tube*)”, Nuclear Instruments and Methods* Sec. A,Volume 355, Issues 2-3, 15 February 1995, Pages 386-389;

[Gys00] T.Gys, “The use of Pixel Hybrid Photon Detectors in the RICH counters of LHCb”, LHCb Note 2000-064 RICH, CERN 2000;

[Hat68] K. Hatch, “Aspects of resolution degradation in a nuclear pulser analyzer system”, *IEEE Transactions on Nuclear Science*, vol. 15, no. 1, 1968, pp. 303-314;

[Hat85] H. Hatano and K. Doi, “Radiation-tolerant high-performance CMOS VLSI circuit design”, *IEEE Transactions on Nuclear Science*, vol. 32, no. 6, December 1985, pp. 4031-4035;

[Hat86] H. Hatano and S. Takatsuka, “Total dose radiation-hardened latch-up free CMOS structures for radiation-tolerant VLSI designs”, *IEEE Transactions on Nuclear Science*, vol. 33, no. 6, December 1986, pp. 1505-1509.

[Hei80] E.Heijne et al., “A silicon surface barrier microstrip detector designed for high energy physics”, *Nuclear Instruments and Methods,* A vol. 178 (1980) pp.331-343;

[Hei88] E.Heijne et al.: "The silicon micropattern detector: a dream?", *Nuclear Instruments and Methods,* A vol. 73 (1988) pp.615-619;

[Hei89] E.Heijne, P.Jarron, "Development of silicon pixel detectors: an introduction", *Nuclear Instruments and Methods,* A Vol. 275 (1989) pp.467-471;

[Hei94] E.Heijne et al.: "Development of silicon micropattern pixel detectors", *Nuclear Instruments and Methods,* A Vol. 384 (1994) pp.399-408;

[Hei96] E. Heijne et al., "LHC1: a semiconductor pixel detector readout chip with internal, tunable delay providing a binary pattern of selected events", *Nuclear Instruments and Methods*, A Vol. 383, no.1, pp. 55-63, December 1996;

[Hei99] W.F. Heidergott, "System Level mitigation strategies", *1999 IEEE Transactions on Nuclear and Radiation Effects Conference Short Course,* Norfolk, VA, July 1999;

[Hoh87] J. H. Hohl and K. F. Galloway, "Analytical model for Single Event Burn Out of power Mosfet", *IEEE Transactions on Nuclear Science*, vol. 34, no. 6, December 1987, pp. 1275-1280;

[Hol96] P. Holl, P. Rehak, F. Ceretto, U. Faschingbauer, J. P. Wurm, A. Castoldi and E. Gatti, "A 55 cm^2 cylindrical silicon drift detector", *Nuclear Instruments and Methods* A vol.377, Issues 2-3 , August 1996, pp. 367-374;

[Hoo81] F.N. Hooge, T.G. Kleinpenning, L.K. Vandamme, "Experimental verification studies on 1/f noise", *Reports on Progress in Physics*, vol. 44, 1981, pp. 479-532

[How01] J.W. Howard et al., "Total dose and single event effects testing on the Intel Pentium III (P3) and AMD K7 microprocessors", *2001 IEEE Radiation Effects Data Workshop record*, IEEE cat. no. 01TH8588 pp. 38-47, July 2001;

[Huh00] M. Huhtinen and F. Faccio, "Computational method to estimate Single Event Upset rates in an accelerator environment", *Nuclear Instruments and Methods* sec. A vol. 450 (2000) pp.155-172;

[Hun01] J.J. van Hunen, M. Campbell, R. Dinapoli et al., "Irradiation and SPS Beam Tests of the Alice1LHCb Pixel Chip", *Proceedings of the 7th workshop on electronics for LHC experiments*, (LEB 2001, Stockholm, Sweden, 10-14 Sept. 2001), pp. 85-89, CERN-2001-005 ; CERN-LHCC-2001-034;

[IEEjt] IEEE Std 1149.1. Test Access Port and Boundary Scan Architecture;

[Jar03] P. Jarron, G. Anelli, S. Commichau et al., "A new concept of Monolithic Silicon Pixel Detectors: Hydrogenated Amorphous Silicon on ASIC", to be published in the Proceedings of the 9th Pisa Meeting on Advanced Detectors, La Biodola, Isola d'Elba, May 2003 as a special *Nuclear Instruments and Methods,* A issue, in press, uncorrected proof (already available online);

[Joh96] A.H. Johnston, “The influence of VLSI technology evolution on radiation induced latchup in space systems”, *IEEE Transactions on Nuclear Science*, vol.43, no.2, April 1996, pp.505-521;

[Joh98] A.H. Johnston, "Radiation Effects in Advanced Microelectronics Technologies", *IEEE Transactions on Nuclear Science*, vol. 45, no. 3, June1998, pp. 1339-1354;

[Jol02] S. Jolly, “Minimising the LHCBPIX1 Threshold Distribution”, RICH meeting 9/12/2002; CERN 2002 http://agenda.cern.ch/askArchive.php?base=agenda&categ=a021898&id=a021898s1t2/transparencies ;

[Klo98] K. Kloukinas, F. Faccio, A. Marchioro and P. Moreira, “Development of a radiation Tolerant 2.0 V standard cell library using a commercial deep submicron CMOS technology for the LHC experiments”, *Proceedings of the Fourth Workshop on Electronics for LHC Experiments* (LEB98, Rome, Italy, September 1998), pp. 574-580;

[Klu00] A. Kluge, “Raw data format of one SPD sector”, to be submitted as an ALICE internal note (CERN);

[Klu01] A. Kluge, P. Chocuka, R. Dinapoli et al., “The ALICE on-detector pixel PILOT system-OPS”, *Proceedings of the 7th Workshop on Electronics for LHC Experiments* (LEB 2001, Stockholm, Sweden, 10-14 September 2001), pp. 95-100, CERN-2001-005; CERN-LHCC-2001-034;

[Klu02] A. Kluge, “Specification of the digital control part of the analog pilot chip”, Internal document (CERN), April 2002;

[Kno89] Glenn F. Knoll, “Radiation Detection and Measurement”, Second edition, John Wiley & SonsGlenn, 1989;

[Kon68] M. Konrad, “Detector pulse shaping for high resolution spectroscopy”, *IEEE Transactions on Nuclear Science*, vol.15, no.1, 1968, pp.268-282;

[Kru91] F. Krummenacher, “Pixel detectors with local intelligence: an IC designer point of view,” *Nuclear Instruments and Methods*, A Vol. 305 (1991) pp.527-532;

[Kui73] K. E. Kuijk, "A precision reference voltage source", *IEEE Journal of solid-state circuits*, vol. Sc-8, no. 3, June 1973, pp: 222-226;

[Lac99] R.C. Lacoe, J.V. Osborn et al., “Total dose radiation tolerance of a Chartered 0.35-µm CMOS process”, *1999 IEEE Radiation Effects Data Workshop record*, IEEE cat. no. 99TH8463 pp. 82-86, July 1999;

[Lac00] R.C. Lacoe, J.V. Osborn et al., “Application of hardness-by-design methodology to radiation-tolerant ASIC technologies”, *IEEE Transactions on Nuclear Science*, vol.47, no.6, pp.2334-2341, July 2000;

[Lac01] R.C. Lacoe. “Performance impact of Radiation-Hardness-By-Design”, *Pres. 2001 Hardness-By-Design Workshop*, Albuquerque, NM, July 2001;

[Lac03] R.C. Lacoe, "CMOS scaling, design principles and hardening-by-design methodologies", 2003 IEEE NSREC short course, section II, Monterey CA, July 2003;

[Lak94] K. R. Laker and W. M. C. Sansen, Design of Analog Integrated Circuits and Systems, McGraw-Hill International Editions, 1994;

[Lee99] B.S. Lee, "Technical review of low dropout voltage regulator operation and performance", Texas Instrument Application Report, August 1999;

[LHC98] LHCb Technical Proposal, CERN/LHCC 98-4; CERN February 1998;

[Llo02] X.Llopart, M. Campbell, R. Dinapoli, D. San Segundo, E. Pernigotti, "Medipix2, a 64k pixel read out chip with 55 μm square elements working in single photon counting mode" *IEEE Nuclear Science Symposium and Medical Imaging Conference,* Vol. 3, 2001, pp. 1484-1488;

[Ma89] F.B. McLean et al., "Electron-hole generation, transport and trapping in SiO_2", in *Ionizing radiation effects in MOS devices & circuits*, chapter 1, T.P. Ma and V.Dressendorfen; J. Wiley & Sons, New York, 1989;

[Man02] L. Manchanda, G. R. Weber et al., "A new method to fabricate thin oxynitride/oxide gate dielectric for deep submicron devices", Microelectronic Engineering, Volume 22, Issues 1-4 , August 1993 , Pages 69-72;

[Mar98] A. Marchioro, "Deep Submicron Technologies for HEP", *Proceedings of the Fourth Workshop on Electronics for LHC Experiments* (LEB98, Rome, Italy, September 1998), pp. 40-46;

[Mas01] L.W. Massengill, J.R. Schwank et al., "Hevy-Ion-Induced Breakdown in Ultra-Thin Gate Oxides and High-k Dielectrics", *IEEE Transactions on Nuclear Science*, vol.48, no.6, pp. 1904-1912, December 2001;

[Mav02] D.G. Mavis and P.H. Eaton, "Soft Error Rate mitigation techniques for modern microcircuits", Proc. of 2002 Intl. Rel. Phys. Symp., pp.216-225, April 2002;

[Mcl76] F.B. McLean, G.A. Ausmand,l Jr. et al., "Application of stochastic hopping transport to hole conduction in amorphous SiO_2", J. Appl. Phys., vol. 47, no.4, pp. 1529-1532, April 1976;

[Mcl77] F.B. McLean and G.A. Ausmand,l Jr., "Simple approximate solutions to continuous time random-walk transport", *Phys. Rev. B*, vol.15, no.2, pp.1052-1061, January 1997;

[Mcl89] F.B. McLean et al., "Electron-hole generation, transport and trapping in SiO_2", in *Ionizing radiation effects in MOS devices & circuits*, T.P. Ma and V. Dressendorfen, chapter 3, J. Wiley & Sons, New York, 1989;

[Mik00] B. Mikulec, "Single photon detection with semiconductor pixel arrays for medical imaging applications", PhD Thesis, University of Vienna, Austria, June 2000;

[Mil71] J. Millman, C.C. Halkias, “Integrated electronics: analog and digital circuits and systems”; McGraw-Hill international book company (1971);

[Mol02] M. Moll, E. Fretwurst, M. Kuhnke, G. Lindstrom, “Relation between microscopic defects and macroscopic changes in silicon detector properties after hadron irradiation”, Nuclear Instruments and Methods B, vol. 186 (2002) pp.100–110;

[Mor00] P. Moreira et al., “G-Link and Gigabit Ethernet Compliant Serializer for LHC Data Transmission”, NSS 2000 p.9/6-9 vol.2, 2000;

[Mor02] P. Moreira, “Bandgap Cell – Irradiation Tests”, Internal document (CERN), 2002;

[Mor03] M. Morel, “Status on the developments of the stave components”, EP/ED meeting, CERN, 4 September 2003, http://morel.home.cern.ch/morel/presentation.htm ;

[Mul86] R.S Muller and T.I. Kamins, “Device electronics for Integrated circuits”, J. Wiley & Sons, New York, 2nd edition, 1986;

[Mus91] O. Musseau, “Effets des ions lourds énergétiques sur les circuits intégrées. Application au cas de circuits MOS, MOS sur isolant et GaAs”, Thèse d’état, France 1991;

[Na6I] NA6i proposal: CERN/SPSC 2000-10/SPSC/P316;

[Nap82] L. S. Napoli, R. K. Smeltzer, R. Donnelly and J. Yeh, "A Radiation Hardened 256*4 Bulk CMOS RAM", RCA Review, vol. 43, September 1982, pp. 458-463;

[Neu94] K. Neumeier, H. P. Bruemmer, "Radiation hard LOCOS field oxides", *IEEE Transactions on Nuclear Science*, vol. 41, no. 3, June 1994, pp. 572-576;

[New03] Charlotte Newby, “Analogue Pilot Tests”, RICH meeting (CERN) 25th February 2003; http://cnewby.home.cern.ch/cnewby/analogue_pilot_tests.htm

[New03a] Charlotte Newby, “Wafer probing”, RICH meeting (CERN) 25th February 2003; http://cnewby.home.cern.ch/cnewby/Wafer_work.htm

[Nil03] Paul Nilsson, “Test Beam Analysis”, SPD General Meeting, CERN, October 22, 2003;

[Now70] C.H. Nowlin, “Pulse shaping for nuclear pulse amplifiers”, *IEEE Transaction on Nuclear Science*, Volume NS-17, Issue 1, February 1970, Pages 226-41;

[Now03] N. Nowlin, J. Bailey, B. Turfler and D. Alexander, “A total-dose hardening-by-design approach for high-speed mixed-signal CMOS integrated circuits”, Intl. Jour. High Speed Elec. Sys., in press;

[Oco02] P. O’Connor, G. De Geronimo, “Prospects for charge sensitive amplifiers in scaled CMOS”, *Nuclear Instruments and Methods*, vol.480, no.2, March 2002, pp.731-725;

[Ohk76] S. Ohkawa, M. Yoshizawa and K. Husimi, “Direct synthesis of the Gaussian filter for nuclear pulse amplifiers”, *Nuclear Instruments and Methods*, Volume 138, Issue 1, 1 October 1976, Pages 85-92;

[Osb98] J.V. Osborn, R.C. Lacoe, D.C. Mayer et al., "Single Event Latchup Characteristics of Three Commercial CMOS Processes", *IEEE Transactions on Nuclear Science*, vol. 45, no. 3, June 1998, pp. 1458-1463;

[Osb98b] J.V. Osborn, R.C. Lacoe, D.C. Mayer and G. Yabiku, "Total Dose Hardness of Three Commercial CMOS Microelectronics Foundries", Proc. 7th NASA Symp. on VLSI design, pp. 4.2.1-4.2.14, Albuquerqe, NM, 1998;

[Pas03] B. Pastircak, "ALICE radiation levels", ALICE Technical Forum, CERN, June 2003;

[Pel89] M. J. M. Pelgrom, A. C. J. Duinmaijer and A. P. G. Welbers, "Matching Properties of MOS Transistors", *IEEE Journal of Solid-State Circuits*, vol. 24, October 1989, pp. 1433-1440;

[Pen96] F.X. Pengg, "Monolithic silicon pixel detectors in SOI technology", Doctoral thesis, Johannes Kepler University, Linz, 1996;

[Rad68] V. Radeka, "Optimum signal-processing for pulse-amplitude spectrometry in the presence of high-rate effects and noise", *IEEE Transactions on Nuclear Science*, vol. 15, no. 3, 1968, pp. 455-470;

[Rad88] V. Radeka, "Low-noise techniques in detectors"; *Ann. Rev. Nucl. Part. Sci.*, 1988. 38. 217. 77; pp. 217-277, 1988;

[Reh86] P. Rehak, Jack Walton et al., "Progress in semiconductor drift detectors", *Nuclear Instruments and Methods* A, Vol. 248, Issues 2-3 , August 1986, pp. 367-378;

[Rie98] P. Riedler: "Radiation damage effects and performance of silicon strip detectors using LHC readout electronics", doctoral thesis, University of Vienna, 1997;

[Rie02] P. Riedler, M. Campbell, R. Dinapoli, M.Morel et al., "The ALICE Silicon Pixel Detector", Proceedings of the International Workshop on Semiconductor Pixel Detectors for Particles and X-Rays (PIXEL2002 Carmel, California, USA September 9-12, 2002); edited by Stewart C. Loken; eConf C020909 (2002); http://www.slac.stanford.edu/econf/C020909/;

[Rie03] P. Riedler, G. Anelli, R. Caliandro, R. Dinapoli, et al., "First results from the ALICE silicon pixel detector prototype", Proceedings of the VERTEX 2001, Brunnen, Switzerland; published in *Nuclear Instruments and Methods* A, vol. 501, n. 1, 21 March 2003, pp. 111–118;

[Rie03a] P. Riedler, R. Dinapoli et al.,"The ALICE Silicon Pixel Detector", presented at VERTEX 2003, Low Wood Lake Windermere Cumbria,UK, 14-19 September 2003, to be published as a special issue of Nuclear Instruments and Methods A;

[Rie03b] P. Riedler, M, Cinausero et al., "ALICE Silicon Pixel Detector Sensor Production Readiness Review", CERN, June 30, 2003;

[Rin98] G.A. Rincon-Mora and P.E. Allen, “A low-voltage, low quiescent current, low drop-out regulator”, IEEE Journal of Solid-State Circuits, vol.33, No.1, January 1998;

[Riv99] A. Rivetti, G. Anelli, F. Anghinolfi et al., “Analog Design in Deep Submicron CMOS Processes for LHC”, *Proceedings of the Fifth Workshop on Electronics for LHC Experiments* (LEB99, Snowmass, Colorado, USA, September 1999), pp. 157-161;

[Riv01] A. Rivetti et al., “A Low-Power 10-bit ADC in a 0.25-mm CMOS: Design Considerations and Test Results”, *IEEE Trans. Nucl. Sci.*, vol. 48, no. 4, pp. 1225-1228, August 2001;

[Roc99] P. Roche: “Etude de l'alea logique (SEU) induit par une particule ionisante dans des mémoires SRAM développes en technologies CMOS submicroniques”, Thèse d’état, Université Montpellier II, Montpellier November 1999;

[Rub98] György Rubin, Pierre Vande Vyvre, “The ALICE Detector Data Link Project”, *Proceedings of the Fourth Workshop on Electronics for LHC Experiments*, INFN Rome, 21-25 September 1998, CERN/LHCC/98-36.

[Ruz99] A. Ruzin, G. Casse, M. Glaser et al., “Comparison of radiation damage in silicon induced by proton and neutron irradiation“, *IEEE Transactions on Nuclear Science*, vol. 46, no. 5, October 1999, pp. 1310-1313;

[Sak84] N.S. Saks, M.G. Ancona and J.A. Modolo, “Radiation Effects in MOS Capacitors with Very Thin Oxides at 80 K”, *IEEE Transactions on Nuclear Science*, vol. 31, no. 6, December 1984, pp. 1249-1255;

[Sak86] N.S. Saks, M.G. Ancona and J.A. Modolo, “Generation of Interface States by Ionizing Radiation in Very Thin MOS Oxides”, *IEEE Transactions on Nuclear Science*, vol. 33, no. 6, December 1986, pp. 1185-1190;

[Sch02] J. Schwank, “Total dose effects in MOS devices”, 2002 IEEE NSREC Short Course, Section III, July 2002;

[Seg77] Emilio Segrè, “Nuclei and particles”, Second edition, Addison-Wesley Publishing Company, Inc., 1977;

[Sex85] F.W. Sexton and J.R. Schwank, “Correlation of radiation effects in transistors and integrated circuits”, *IEEE Transactions on Nuclear Science,* vol.32, no.6, December 1985, pp. 3795-3981;

[Sex97] F.W. Sexton, D.M. Fleetwood, M.R. Shaneyfelt et al., “Single Event Gate Rupture in Thin Gate Oxides”, *IEEE Transactions on Nuclear Science*, vol. 44, no. 6, December 1997, pp. 2345-2352;

[Sha98] M.R. Shaneyfelt, P. E. Dodd, B. L. Draper and R. S. Flores, "Challenges in Hardening Technologies Using Shallow-Trench Isolation", *IEEE Transactions on Nuclear Science*, vol. 45, December 1998, pp. 2584-2592;

[Sna03] W.P. Snapp, “Mixed-signal RHBD strategies at Boeing”, *Pres. 2002 DARPA HBD Workshop*, Washington DC, February 2003;

[Sno92] W. Snoeys, “A new integrated pixel detector for high energy physics”, doctoral thesis at Stanford University, California (1992);

[Sno98] W. Snoeys, M. Campbell, E. Cantatore, R. Dinapoli, et al., “Radiation tolerance beyond 10 Mrad for a pixel readout chip in standard submicron CMOS”, *Proceedings of the 4th Workshop on Electronics for LHC Experiments*, Rome, Italy, 21-25 Sep 1998, pp.114-118 -CERN, Geneva, 1998 [CERN-LHCC-98-036];

[Sno00] W. Snoeys, F. Faccio et al., "Layout techniques to enhance the radiation tolerance of standard CMOS technologies demonstrated on a pixel detector readout chip”, 8th European Symposium on Semiconductor Detectors, Schloss Elmau, June 14-17, 1998, published in *NIM*, A Vol. 439, Jan. 2000, pp. 349-360;

[Sno00a] W. Snoeys, G. Anelli, M. Campbell et al., “Integrated Circuits for Particle Physics Experiments”; Proceedings of the International Solid State Circuits Conference, 2000, pp.156-160; *IEEE JSC*, Vol. 35, No. 12, December 2000, pp.2018-2030;

[Sno01] W. Snoeys, M. Campbell, E. Cantatore, R. Dinapoli et al., “Pixel readout electronics development for the ALICE pixel vertex and LHCb RICH detector”, *Nuclear Instruments and Methods*, A Vol. 465 (2001), pp. 176–189;

[Sno01a] W. Snoeys, M. Campbell, E. Cantatore, R. Dinapoli et al., “Pixel readout chips in deep submicron CMOS for ALICE and LHCb tolerant to 10 Mrad and beyond”, *Nuclear Instruments and Methods*, A Vol. 466 (2001), pp. 366–375;

[Sto98] P. A. Stolk, F. P. Widdershoven and D. B. M. Klaassen, "Modeling Statistical Dopant Fluctuations in MOS Transistors", *IEEE Transactions on Electron Devices*, vol. 45, no. 9, September 1998, pp. 1960-1971;

[Sze81] S.M.Sze, “Physics of semiconductor devices”, J. Wiley&Sons, sec. edition (1981);

[Tsi99] Y. Tsividis, “Operation and Modeling of The MOS Transistor”, Second Edition, McGraw-Hill International Editions, 1999;

[Tsu61] M. Tsukuda, “Pulse analyzing system for a gridded ionisation chamber”, *Nuclear Instruments and Methods*, A Vol. 14 (1961), pp. 241–251;

[Tui97] H. P. Tuinhout, A. H. Montree et al., "Effects of Gate Depletion and Boron Penetration on Matching of Deep Submicron CMOS Transistors", *Technical Digest of the IEEE International Electron Device Meeting* 1997, pp. 631-634;

[Tui01] H. Tuinhout, M. Vertgart and M. Pelgrom, “Matching of MOS transistors”, Advanced Engeneering Course on Low-Voltage, Low-Power analog CMOS IC design, Lausanne, Switzerland, July 2001;

[Tur01] R. Turchetta et al., “A monolithic active pixel sensor for charged particle tracking and imaging using standard VLSI CMOS technology”, *Nuclear Instruments and Methods,* A Vol. 458 (2001) pp. 677-689;

[Wal92] R. Walker et al., "A Two-Chip 1.5 GBd Serial Link Interface", *IEEE Journal of solid state circuits*, Vol. 27, No. 12, December 1992, Agilent Technologies, Low Cost Gigabit Rate Transmit/ Receive Chip Set with TTL I/Os, HDMP-1022, HDMP-1024, http://www.semiconductor.agilent.com , 5966-1183E(11/99);

[Web00] CERN home pages: www.cern.ch and public.web.cern.ch/public/ (general public);

[Web01] Fermi National Accelerator Laboratory (Fermilab) home page: www.fnal.gov;

[Web02] Brookhaven National Laboratory (BNL) home page: http://www.bnl.gov/world ;

[Web03] Deutsches Elektronen-Synchrotron (DESY) home page: http://www.desy.de/html/home/index.html ;

[Web04] CERN Grey Book; http://greybook.cern.ch/ ;

[Web05] Lawrence Berkley National Laboratory, "The particle adventure, the fundamentals of matter and force", http://particleadventure.org/particleadventure/ ;

[Web06] Stanford University, The theory behind particle physics, http://www2.slac.stanford.edu/vvc/theory.html ;

[Web07] CERN Deep Submicron Technology available cells, "Bandgap Cell", P. Moreira: http://deepsub.web.cern.ch/deepsub/docs/common_cells/BandgapDescription.html

[Web08] Agilent Technologies, Agilent HDMP-1032, HDMP-1034, Transmitter/Receiver Chip Set, http://www.semiconductor.agilent.com , 5968-5909E(2/00);

[Web09] Röhm GmbH & Co. KG, http://www.roehm.com/en/communication.html ;

[Web10] CERN/MIC group X-ray irrad. system: proj-xraymic.web.cern.ch/proj-XrayMIC/ ;

[Web11] ALICE home page: http://alice.web.cern.ch/Alice/AliceNew/; presentation of the experiment: alice.web.cern.ch/Alice/AliceNew/public/presentation/ ;

[Wei96] P.Weilhammer et al., "Si Pad Detectors", *NIM* A Vol. 383 (1996), pp.89-97;

[Win89] P.S. Winoukur: "Radiation induced interface traps", in *Ionizing radiation effects in MOS devices & circuits*, chapter 4, T.P. Ma and V. Dressendorfen; J. Wiley & Sons, New York, 1989;

[Wyl97] K.H. Wyllie: "The design and development of radiation tolerant microstip detectors for tracking at the future large hadron collider", doctoral thesis, University of Cambridge, RAL-TH-1998-013, 1997;

[Wyl99] K.H. Wyllie, W. Snoeys, M. Campbell, R. Dinapoli et al. "A pixel readout chip for tracking at ALICE and particle identification at LHCb", *5th Conference on Electronics for LHC Experiments*, Snowmass, CO, USA, 20 - 24 Sep 1999, pp.93-97 - CERN, Geneva, 1999 [CERN-99-09 ; CERN-LHCC-99-033];

[Wyl00] K.H. Wyllie for the LHCb RICH group, "Level-0 Electronics for the LHCb RICH", LHCb technical note, LHCb2000-075 RICH, CERN August 2000;

[Wyl01] K.H. Wyllie for the ALICE pixel design group, "The ALICE1LHCb chip manual", draft 4, 3/07/01, http://kwyllie.home.cern.ch/kwyllie/pixel_doc/ALICE1LHCB_Documentation.pdf ;

Acknowledgements

First of all, I would like to thank my CERN supervisors, Pierre Jarron (for the MIC group) and Giorgio Stefanini (for the ALICE SPD group), as well as Mike Letheren (the MIC Group Leader), for the possibility they gave me to work at CERN in such a stimulating and interesting environment, and for their constant support.

Thanks to Jean Gasiot, and to Henry Luquet and Jean-Claude Vaissiere, for having offered me the possibility of working on a Ph.D. Thanks "Chef" for having accepted to be my thesis director, and thanks also for having so often trusted my judgement.

Thanks to André Touboul and Raoul Velazco for their kindness and availability, for having read my thesis and for their useful comments. Thanks also for having accepted to be in the Jury. The defence of the thesis was source of great satisfaction for me: it was a pleasure to discuss my work with such a Jury (and public…).

I would like to specially thank Walter Snoeys. He supervised a big part of my work, and from the long discussions we had and his comments I learnt really a lot. In particular, his comments on my thesis and on the defence of the thesis were fundamental for the final quality of both. Moreover, he was so kind to accept to be in the Jury. Thank you Walter, it was a great pleasure to work (and to share the office) with you.

Several other people revised parts of my manuscript: thanks to Giovanni Anelli, Francis Anghinolfi, Rocco Caliandro, Petra Riedler, Alexander Kluge, Matthieu Despeisse. But most of all, thanks to Ken Wyllie: he revised *all* of it, and I am pretty sure he risked severe neurological damages because of that (narcolepsy, most probably)… Thanks to you Ken, I really appreciated.

To work in the CERN ALICE SPD group has been very nice, thanks to all of you: Alex, Fabio, Giorgio, Michel, Mike, Paul, Peter, Petra, and thanks to the people of the LHCb group, Elena, Maria, Thierry, Yeroen…

Many of the colleagues of the MIC group were also good friends: I was lucky to work in such a nice environment. Thanks, together with the colleagues already mentioned, to Angelo, Bettina, Danielle, Domenico, Erik, Eugenio, Federico, Giovanni (both of them), Jean-Claude, Lukas, Michael, Nico, Paul, Raffa, Robert, Vincent, Yan... And thanks Xavi, it was great to have you as an office mate for two years! Thanks to Bert Van Konigsveld, too, for the zillion

of everyday problems that he solved for me (well, to be honest I also appreciated a lot his walnuts, figs, blueberries and so on).

Merci aussi à tout le group du CEM II pour leur accueil e support, et en particulier à Jean Roch, Laurent et Matthieu. Mais surtout à Paule et Florence (DRED) pour leur aide.

Grazie a tutta la famiglia (mamma Tata e papa' Marco; Giangio, Lorella e Marco; Giggione e Alessia; Mako e Silvia, nonche' la nonna e l'enorme moltitudine di zie zii cugine e cugini di ogni grado, con relativi compagni/compagne e pargoli che hanno allietato le feste comandate) per il sostegno e l'amore di sempre. Un grazie speciale a mia madre anche per essersi sobbarcata quasi 2000 km per venire a vedere la mia seduta di dottorato, pur sapendo che non ci avrebbe capito granche' visto che non parla francese e che nella lotta col videoregistratore e' sempre il videoregistratore che vince (insomma la tecnologia non e' esattamente il suo forte...).

Grazie a Stefania per il suo amore, e per avermelo dimostrato venendo sin qui; per essermi stata vicino e avermi sostenuto nei momenti difficili, nonche' per essersi sorbita le tonnellate di "lamentazioni" tipiche di qualunque studente di dottorato che stia scrivendo la tesi.

Eight hours a day in front of a screen, drawing green, red and yellow rectangles with the final aim to study Quark Gluon Plasma can drive you quite far from the "real world"… The need to use your hands to produce something which does not need to be taped-out and sent to a foundry helps to keep in touch with reality, I think. Maybe this is why I ended up with the strange hobby of home cheesemaking: I spent hundred of the typical "Geneva week-ends" (i.e. rainy-cold-foggy-windy) trying to make mozzarella at home... Quindi grazie a Rosa (e alla sua famiglia), prima responsabile di questo strano hobby, per essere stata una maestra cosi' paziente, e per la sua sempre calorosa ospitalita'. Et merci à Edith de m'avoir fourni le (merveilleux et indispensable) lait nécessaire pour toutes mes expériences laitières. Just for the record: as many people both in Montpellier and at CERN already know thanks to my Ph.D. parties, not only I managed to make mozzarella at home, but also several other kinds of cheese.

Vorrei in conclusione ringraziare uno strano gruppo di persone, formatosi per caso a partire dal comune interesse per la cucina. Delle persone speciali che, nonostante la distanza, sono state enormemente presenti in questo periodo. Grazie a tutti, siete stati davvero un grande sostegno. Se sono arrivato a scrivere la parola "fine" su questa tesi e' anche merito vostro.

TITRE an anglais:
A radiation tolerant pixel detector system for the ALICE and LHCb experiments at CERN

RESUME en anglais:

Due to the high radiation levels, standard electronics cannot be used for experiments at the CERN future accelerator (LHC), in particular for what concerns tracking detectors, which stand very close to the interaction point. A possible solution is provided by qualified radiation hard processes, normally only used for small production volumes and therefore expensive and prone to process stability, yield and availability problems. A completely different approach consists of using a standard commercial CMOS technology, hardened with special design techniques (Hardening By Design, HBD). This thesis deals with the design of the first very large pixel detector readout chip using HBD techniques and the development of the full pixel detector system for the ALICE experiment. The readout chip is also used for the RICH detector of the LHCb experiment, and the requirement to satisfy both applications lead to very stringent specifications with respect to power, speed, noise, channel-to-channel uniformity and area. The chip contains 8192 pixel readout channels, about 13 million transistors, externally controllable biasing and testing circuitry.
Several test chips were designed, produced and tested prior to the completion of the full chip. They were functional and could stand a radiation dose much higher than the dose foreseen for the ALICE experiment, but did not satisfy the very stringent front-end requirements for the ALICE and LHCb experiments combined. This led to the development of the actual front end on the final chip, capable of dealing with the high rates in LHCb. The front-end implemented has a pole constellation with two complex conjugate poles and real pole, all with the same real part. This gives, in particular, a preamplifier with a very fast return to zero time (<100 ns @ 1%), which avoids pile-up in high multiplicity environments. A first run yielded chips which could only be clocked at 15-18 MHz, after correction a second run yielded chips functional at a 40 MHz clock rate. The chip was irradiated with 10 keV X-rays, up to a dose of 100 kGy (SiO_2), and did not show significant degradation. Irradiations with protons and heavy ions demonstrated a very good tolerance to Single Event Effects. In conclusion, test results showed full functionality of the chip satisfying all specifications, and currently the chips have been produced and are being assembled for use in the experiments.
In addition, the silicon detector to which the readout chip is connected using flip-chip bump bonding soldering techniques was designed. Electrical tests (e.g. leakage current and depletion voltage) were carried out before and after irradiation, and demonstrated that the detector performs within specifications for a radiation dose three to six times beyond the expected for the ALICE experiment.
An auxiliary chip (the APC) which generates analogue biases and measures slow-varying analogue signals was also designed and tested, and is fully functional.

RESUME en français :
Etant donné le haut niveau de radiations, une électronique standard ne peut être utilisée pour les expériences prévues au CERN au sein du futur accélérateur (LHC), en particulier pour ce qui concerne les détecteurs de trace, qui sont placés très proches du point d'interaction. Une possible solution consiste en l'utilisation d'un procédé durci qualifié sous rayonnement**,** d'ordinaire utilisé uniquement pour de faibles volumes de production ; mais ceci résulte en un prix élevé et un procédé sujet à des problèmes de stabilité, de rendement et de disponibilité. Une approche totalement différente consiste à utiliser une technologie commerciale standard CMOS, durcie par des techniques spécifiques de dessin (Hardening By Design, HBD). Cette thèse traite du dessin du premier très large circuit intégré pour la lecture d'un détecteur à pixels, en utilisant les techniques HBD, et du développement du système global du détecteur à pixels pour l'expérience ALICE. La puce électronique de lecture est aussi utilisée dans le détecteur RICH de l'expérience LHCb. Pour satisfaire aux deux applications, le cahier des charges comporte des spécifications strictes par rapport à la consommation de puissance, à la vitesse, au bruit, à l'uniformité entre les canaux et à la surface. La puce contient 8192 canaux de lecture, environ 13 millions de transistors, une polarisation contrôlable depuis l'extérieur, et une circuiterie de tests.

Plusieurs puces de test ont été dessinées, produites, et testées avant le dessin de la puce entière. Elles étaient fonctionnelles et ont tenu une dose d'irradiation totale bien plus élevée que la dose prévue pour l'expérience ALICE, mais ne satisfaisaient pas le strict cahier des charges imposé par les deux expériences combinées ALICE et LHCb. Ceci a mené au développement de la version finale de l'électronique d'entrée, capable de faire face aux taux d'occupation élevés du canal. L'électronique d'entrée implémentée a une constellation de pôles, avec deux pôles complexes conjugués et un pôle réel, les trois pôles présentant une partie réelle identique. Ceci résulte principalement en un préamplificateur présentant un retour à l'origine très rapide (<100 ns @ 1%), permettant d'éviter les effets d'accumulation liés aux environnements de haute multiplicité. Une première version de la puce ne pouvait être utilisée que pour des fréquences d'horloge inférieures à 15-18 MHz. Après correction, une seconde version de la puce a démontré une parfaite fonctionnalité pour une fréquence d'horloge de 40 MHz. La puce a été irradiée avec des rayons X de 10 keV, jusqu'à une dose totale de 100 kGy (SiO_2), et n'a montré aucune dégradation notoire. Des irradiations avec des protons et des ions lourds ont démontré une excellente tolérance aux effets singuliers. En conclusion, les résultats des tests ont montré une parfaite fonctionnalité de la puce, satisfaisant toutes les spécifications. Les puces ont été produites et sont en cours de montage dans les différentes expériences. De plus, le détecteur silicium connecté à la puce de lecture en utilisant la technique d'assemblage dite « flip-chip » a été dessiné. Des tests électriques (par exemple la caractérisation des courants de fuite et des tensions de désertion), ont été réalisés avant et après irradiation, et ont démontré que le détecteur fonctionne conformément au cahier des charges, pour une dose allant jusqu'à 6 fois la dose prévue pour l'expérience ALICE. Une puce auxiliaire (APC), qui génère des polarisations analogiques et mesure des signaux analogiques variant lentement, a aussi été dessinée et testée, et est entièrement fonctionnelle.

TITRE ET RESUME an anglais: En dernière page

SPECIALITE: Électronique et Micro-optoélectronique

MOTS-CLES: Circuits intégrés, Durcissement aux radiations, Détecteurs de particules, Technologies CMOS, Large Hadron Collider, Transistors à géométrie fermée, ALICE, LHCb, électronique d'entrée, pixels

---

INTITULE ET ADRESSE DU LABORATOIRE DE RATTACHEMENT:
CERN, laboratoire Européen pour la Physique des Particules, CH-1211 Genève 23, Suisse

---