Radiation hardened design of CMOS pixel sensor for the micro-vertex detector of the CBM experiment

Yue Zhao

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE PHYSIQUE ET CHIMIE-PHYSIQUE

Institut Pluridisciplinaire Hubert CURIEN (IPHC) - UMR7178

THÈSE présentée par :

Yue ZHAO

Soutenue le : **Février 18 2021**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Physique

Radiation Hardened Design of CMOS Pixel Sensor for the Micro-Vertex Detector of the CBM Experiment

THÈSE dirigée par :

M. HU Yann	Professeur, Université de Strasbourg
M. BAUDOT Jérôme	Professeur, Université de Strasbourg

RAPPORTEURS :

M. BARBERO Marlon	Professeur, Université d'Aix Marseille
M. GOIFFON Vincent	Professeur, ISAE - SUPAERO

AUTRES MEMBRES DU JURY :

Mme. ORSINI Fabienne	Directrice de recherches, Synchrotron SOLEIL
M. UHRING Wilfried	Professeur, Université de Strasbourg

Contents

Contents I
Résumé VII
Chapter 1. The MVD for the CBM experiment 1
1.1 The CBM experiment 1
1.1.1 The SIS Machine 2
1.1.2 The Physics Program 3
1.1.3 The subdetectors of CBM Experiment 5
1.2 The Micro Vertex Detector (MVD) 6
1.2.1 Requirements 6
1.2.2 Geometries 8
1.2.3 The predicted particle environment for MVD 9
1.3 The introduction of Sensors for MVD 11
1.3.1 Introduction into semiconductors 11
1.3.2 Introduction to silicon sensors 13
1.3.3 The Pixel Sensors for MVD 14
1.4 The MIMOSIS for CBM MVD 19
1.4.1 Design of the Pixel 20
1.4.2 Buffer structure 21
1.4.3 The development schedules 23
1.5 The strategy and focus of this work 23
1.6 Bibliography 25

Chapter 2. The Radiation Effects on CMOS Sensors.....27
2.1 Basic Radiation Mechanism on Silicon Sensors and I.C. 28
2.1.1 Electrons and positrons.....28
2.1.2 Protons and ions.....31
2.1.3 Energy loss, linear energy transfer and stopping power ..32
2.2 Single event effects.....34
2.2.1 Mechanism.....34
2.2.2 SEE terminology and classification.....37
2.2.3 SEE Response in Digital Circuits.....38
2.3 Total ionizing dose effects.....42
2.3.1 Mechanism.....42
2.3.2 Impact of TID Deficiencies on Bulk MOS Structures 45
2.3.3 Impact of TID Deficiencies on STI MOS Structures.....48
2.4 Displacement damage.....49
2.5 Summary.....52
2.6 Bibliography 53
Chapter 3. Radiation Harden Techniques for MVD against SEE and TID 59
3.1 Introduction to System Reliability.....59
3.2 Triple modular redundancy (TMR) 61
3.2.1 Topologies of TMR in digital circuits 61
3.2.2 The trade-off on reliability and cost 63
3.3 Error correction code (ECC).....66
3.3.1 Mechanism and Codeword distance 66
3.3.2 Hamming code.....67
3.3.3 CRC 68
3.3.4 Discussion.....70

3.4 Refreshing and reloading 70
3.5 Re-entrant gate layout and guard ring 72
3.6 Reverse body biasing 74
3.7 Silicon-on-insulator (SOI) 75
3.8 Summary 76
3.9 Bibliography 78
Chapter 4. The Radiation-Harden Design of The Frontend Readout System 80
4.1 Design principle of the readout system 81
4.2 Evaluation 86
4.2.1 The cross-section and hit density of MVD 86
4.2.2 The error rate in control flows 87
4.2.3 The error rate in data flows 88
4.2.4 The error rate in global controller 88
4.3 TMR Deployment 89
4.3.1 Design Flow with TMR 89
4.3.2 Detailed Deployment 90
4.4 ECC Deployment 95
4.5 Verification 95
4.5.1 Modelling SEE pulse 96
4.5.2 Simulation with SEE pulse injected 99
4.6 TMR Optimization 101
4.7 Summary 103
4.8 Bibliography 104
Chapter 5. The Radiation Harden Design of the Phase-Locked Loop 105
5.1 Architecture of PLL 105

5.1.1 Topology 106
5.1.2 Noise Model 107
5.2 Circuit design 108
5.2.1 VCO 108
5.2.2 Charge Pump 110
5.2.3 Phase/Frequency Detector and Lock State Detector 111
5.2.4 Loop Parameter 113
5.2.5 Divider 114
5.2.6 Layout 115
5.3 Corrections and Improvements 115
5.3.1 The deep N-well connection 115
5.3.2 New lock detection structure 116
5.3.3 Dual Modular Redundancy on PFD 117
5.3.4 Charge Pump balance 118
5.3.5 Re-entrant gate layout for bias 119
5.3.6 Compact layout 120
5.4 Test method and results 121
5.4.1 Test board, system 122
5.4.2 Static test results 123
5.4.3 Jitter performance 124
5.4.4 Dynamic test results 126
5.5 The improvements 128
5.6 Summary 131
5.7 Bibliography 132
Chapter 6. Preparation for the Radiation-Tolerance Evaluation of MIMOSIS-1 133
6.1 SEE Cross-section Calibration 133

6.1.1 Characterization Method 134
6.1.2 Chip design 135
6.1.3 Test Design 137
6.2 SEE Beam Test Method of the Readout System 140
6.2.1 The testability design of MIMOSIS-1 140
6.2.2 The in-Time Test Method of MIMOSIS-1 142
6.3 TID and SEE Beam Test Method of the PLL 143
6.3.1 TID Test Method 143
6.3.2 SEE Test Method 144
6.4 Summary 145
6.5 Bibliography 146
Chapter 7. Conclusions and Perspectives 147
7.1 Conclusions 147
7.2 Perspectives 151
7.2.1 SEE in smaller-scale process 151
7.2.2 SEE evaluation by simulation 155
7.3 Bibliography 159

Résumé

1. Introduction

Il est prévu que l'expérience sur la matière baryonique comprimée (CBM) démarre vers 2024 auprès du complexe d'accélérateurs de particules FAIR (Facility for Antiproton and Ion Research in Europe) sur le site de GSI (Gesellschaft für Schwer Ionenforschung, Darmstadt). Son programme de recherche va consister à explorer le diagramme de phase de l'interaction nucléaire forte QCD (Quantum Chromo Dynamic) dans la région des densités baryoniques élevées en utilisant des collisions noyau-noyau à cible fixe de haute énergie ou proton-noyau. Il comprendra l'étude de l'équation d'état de la matière nucléaire à des densités élevées et la recherche du début du déconfinement et de la transition de phase chirale. Comme le montre la Figure 1, le détecteur de micro vertex (MVD) est le premier des instruments CBM, juste après la cible. Sa tâche sera de reconstruire avec une excellente résolution l'origine des différentes particules secondaires produites lors de la collision. L'instrument lui-même est composé de quatre plans de détection consécutifs, chacun équipés de capteurs mesurant la position grâce à une pixellisation d'environ 27×32 μm^2. Pour répondre aux exigences de l'analyse physique, le MVD doit fonctionner dans le vide et présenter un très faible budget de matière d'environ 0,3% de longueur de radiation par plan.

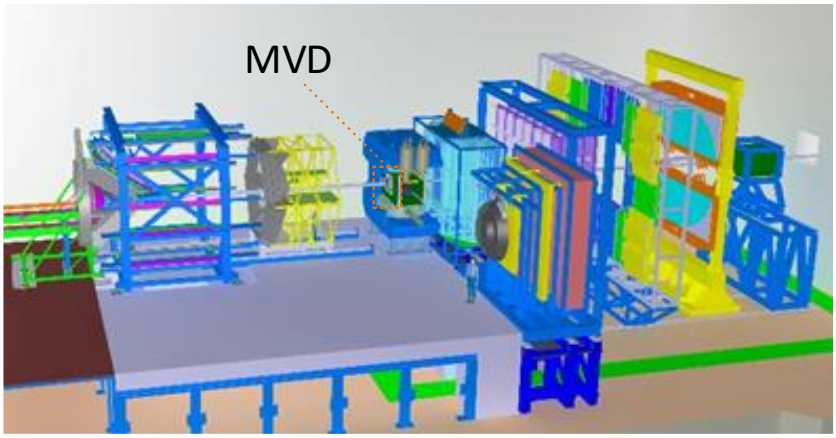


Figure 1: Modèles 3D de l'expérience CBM, avec l'emplacement du MVD. Le faisceau arrive de la gauche.

Le capteur MIMOSIS-1 est un capteur à pixels actifs monolithique en technologie CMOS (MAPS) conçu par l'IPHC en tant que premier capteur prototype de taille réelle pour le MVD. Au-delà des fonctionnalités utiles à la physique, les conditions environnementales de fonctionnement du MVD ont dû être prises en compte dans la conception du circuit. Les particules hautement ionisantes générées par les collisions, telles que les ions or ou carbone et les protons, peuvent introduire différents effets de rayonnement dans les capteurs MVD : événements singuliers (SEE), effets de la dose ionisante totale (TID) et de la dose de dommages par déplacement. Mon objectif de thèse était de comprendre l'impact des effets SEE et TID sur les circuits de lecture du capteur, d'élaborer une stratégie de conception pour en atténuer l'impact et enfin de mettre en œuvre cette conception spécifique.

Un événement singulier est une perturbation électrique qui peut engendrer une modification du fonctionnement normal d'un circuit. La cause est liée au passage d'un seul ion à travers ou à proximité d'un nœud sensible dans un circuit. Les particules chargées perdent de l'énergie et génèrent une série de paires électrons-trous lorsqu'elles traversent le matériau semi-conducteur. Cet effet est utilisé pour détecter la particule initiale avec un nœud de collection spécifique (diode). Dans les autres micro-circuits nécessaires aux fonctionnalités du capteur, de telles paires électrons-trous déplacées par le champ électrique seront également collectées par les électrodes de circuit. Comme le montre la Figure 2, ces

charges collectées provoquent une courte impulsion de courant dans les micro-circuits qui peuvent provoquer une erreur (par exemple logique) dans leur fonctionnement. Les SEE tels que les événements provoquant un upset (SEU), l'apparition d'un courant transitoire (SET), un vérrouillage de courant ou latchup (SEL), une interruption fonctionnelle (SEFI)… peuvent entraîner des erreurs fonctionnelles temporaires ou permanentes du circuit.

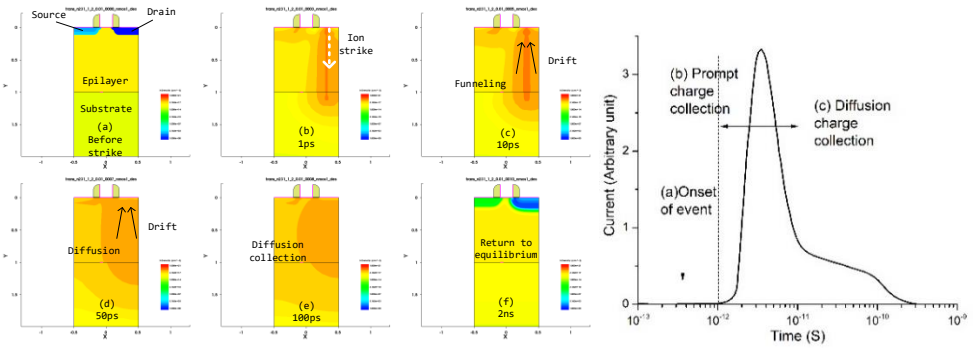


Figure 2 La simulation TCAD de SEE

L'effet de dose ionisante totale est dû à la collecte de charges ionisées par le piégeage profond des trous près de la couche isolante et du substrat, c'est-à-dire proche des interfaces Si/SiO2 dans le MOSFET. L'effet TID peut entraîner un décalage de la tension de seuil et une augmentation du courant de fuite, ce qui peut provoquer un dysfonctionnement du circuit analogique et une augmentation de la consommation statique.

2. Travail Doctoral

La première partie de mon travail a consisté à étudier les différentes techniques de durcissement des circuits intégrés contre les radiations. Trois méthodes peuvent être employées au niveau de la conception du système : Triple redondance modulaire (TMR), Code de correction d'erreur (ECC) et rafraîchissement et rechargement des registres. L'utilisation de transistors MOS fermés et la polarisation inverse du bulk sont des techniques de durcissement applicables au niveau de la conception de circuits. Enfin, la technologie SOI (Silicon-on-Insulator) est un procédé de fabrication qui présente une bonne tolérance aux effets d'événements singuliers. Les principes, les méthodes d'implémentations, les avantages et les inconvénients de ces différentes techniques seront discutés en détail durant

la thèse. Nous résumons ici les résultats. Le TMR est adapté aux blocs critiques et de petite taille. L'ECC est généralement utilisé pour le flux de données. Le rafraîchissement et le rechargement sont mieux adaptés aux registres de configuration et de monitoring pour protéger le flux opérationnel d'un système. Ces trois méthodes sont efficaces pour se protéger contre les SEE. L'utilisation de MOS fermés et la polarisation inverse du bulk conviennent davantage pour diminuer l'effet de dose TID. La technologie SOI peut éliminer les SEE en principe mais aurait nécessité de changer la technologie de détection déjà validée pour MIMOSIS-1 et pourrait présenté d'autres effets indésirables.

En se focalisant sur l'impact des SEE et l'effet de dose TID, deux modules de circuits critiques dans MIMOSIS-1 ont été identifiés. Le premier est le front-end gérant la lecture des données de mémoires tampons. Afin de répondre aux spécifications en termes de vitesse de lecture et de compression élevée des données, une structure à 3 couches de mémoires tampons a été proposée pour le système de lecture de MIMOSIS-1. Cette architecture permet de modifier le processus de lecture en passant d'un flux de données en rafale rapide en un flux de données séquentiel lent. Chaque couche de mémoire tampon a sa mémoire et sa logique de contrôle. L'utilisation de TMR et ECC est bien adapté pour le front-end de lecture et est décrit en détail dans la thèse. Il y a un compromis à trouver au niveau des paramètres de conception entre la fiabilité requise dans l'environnement radiatif, la surface occupée et les budgets de puissance dissipée du circuit.

Afin de pallier maladroit de cette conception, nous proposons dans cette thèse une méthode d'optimisation de conception pour l'implémentation de TMR basée sur la simulation multi-niveaux. Pour estimer la fiabilité du circuit face aux SEE, deux grandeurs sont nécessaires : le taux de particules hautement ionisantes et la section efficace SEE du circuit. Les deux quantités consequence de grandes incertitudes puisque l'accélérateur et le capteur sont en développement. Par conséquent, l'évaluation est d'abord basée sur la section efficace connue d'un autre capteur à pixels CMOS similaire (ALPIDE) utilisé dans l'expérience ALICE, puis sur le modèle proposé dans la thèse. Les paramètres d'impulsion SEE correspondant à

différentes pertes d'énergie ionisante peuvent être obtenus grâce à une simulation TCAD basée sur les paramètres du processus de fabrication. L'outil de génération SEE peut ensuite injecter plusieurs impulsions SEE aléatoires basées sur le modèle SEE dans le circuit, alors que le système de lecture est en simulation. Grâce à l'étude statistique des résultats de la simulation, la stabilité du module subissant des SEE peut être évaluée. Avec une conception optimisée de la Figure 3, le système peut toujours fonctionner après 10^7 SEE, ce qui correspond à plus de cinq ans de fonctionnement du MVD. En même temps, le taux d'erreurs avec la conception durcie est réduit de 4 à 5 ordres de grandeur. De plus, le taux d'erreur de sortie de données équivalent dans l'environnement de travail est inférieur à 10^{-9}.

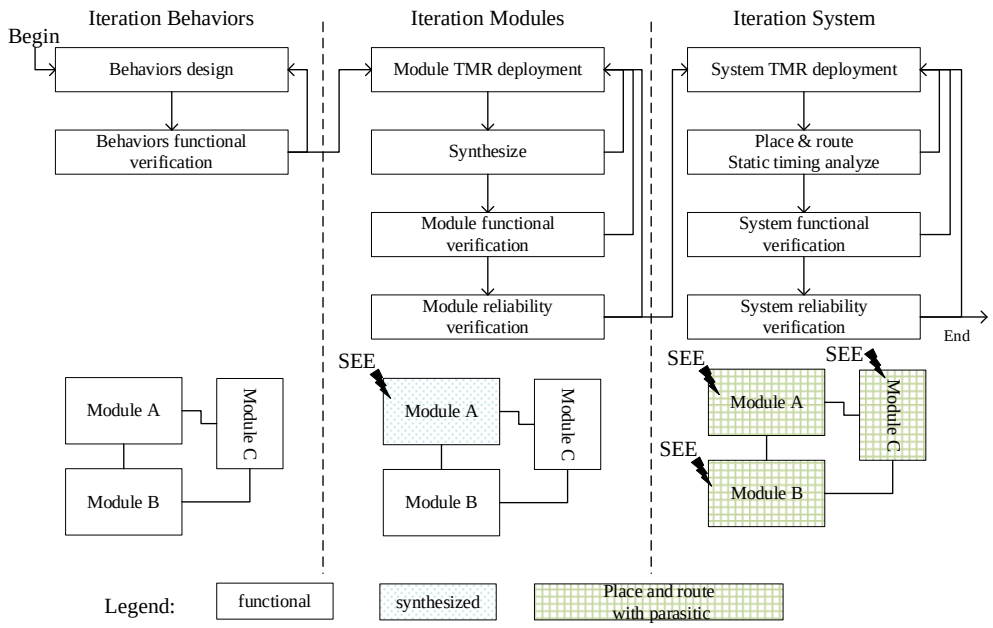


Figure 3. Le processus de conception et de vérification avec le contrôle de fiabilité SEE.

Le second module critique est le circuit de génération de fréquence, principalement la boucle à verrouillage de phase (PLL). Le pilotage de MIMOSIS-1 est basé sur une logique numérique synchrone pour laquelle des horloges stables sont essentielles. Si la PLL reçoit une impulsion en sortie inattendue, cela peut amener à une période instable ou une perte de verrouillage et le circuit numérique qu'elle pilote aura de graves problèmes de fonctionnement. Dans le circuit réalisé dans ce travail, le TMR, les

transistors MOS fermés et la polarisation inverse du bulk sont implémentés. Là encore, la conception du circuit durci minimise la surface occupée et la dissipation de puissance.

Avec une conception optimisée avec des techniques de durcissement aux radiations, le prototype PLL-1 + a une petite surface (0,115 mm^2), une faible consommation de puissance (3,6 mW), un temps de verrouillage court (5us) et une fonction de détection de verrouillage. Il a un niveau de gigue total à 2% UI dans le test indépendant et 8% UI dans le test d'intégration à forte activité.

La dernière partie de cette thèse est consacrée aux méthodes de vérification du durcissement aux radiations. Afin de valider les résultats de la simulation, des tests en faisceaux sont prévus. Cependant, il est difficile de calculer directement la section efficace de la puce fonctionnelle de grande taille (MIMOSIS-1). Une autre puce d'évaluation a été conçue pour calibrer la section efficace du processus de fabrication. Les méthodes de test au SEE statique et dynamique pour la puce d'évaluation, MIMOSIS-1 et PLL-1+ ont toutes été conçues. Le test TID pour le circuit PLL-1+ est également abordé.

3. Conclusion

En résumé, dans cette thèse, nous avons étudié les techniques de durcissement des circuits adaptées à l’environnement radiatif exigeant du MVD de l’expérience CBM et nous avons ensuite présenté une stratégie globale de durcissement au niveau de la conception où des contraintes supplémentaires (comme la taille des microcircuits et le budget de puissance) sont à satisfaire. La méthode est basée à la fois sur des simulations spécifiques et une évaluation du niveau critique de chaque fonctionnalité dans le circuit global.

La stratégie a été implémentée sur le premier prototype de capteur de taille finale pour le détecteur Micro Vertex de l'expérience CBM. La simulation et les résultats des tests montrent qu'il est possible de concevoir un détecteur avec des contraintes fortes pour répondre aux besoins de tolérance aux rayonnements. Les circuits numériques et analogiques fonctionnent correctement sous les effets SEE et TID grâce à la simulation.

Le procédé décrit dans cette thèse peut également être étendu à d'autres conceptions de circuits intégrés durcis aux radiations.

Chapter 1. The MVD for the CBM experiment

This initial chapter introduces the background for the physics of the CBM experiment and the requirements on its detector, which is the source of this research work. Firstly, the physics motivation is summarized and emphasized by the exploration of the phase diagram of strongly interacting matter utilizing heavy-ion collisions at an energy scale below 5 GeV. Secondly, the chapter depicts the general concept of the CBM detector with a focus on the Micro Vertex Detector (MVD), which can achieve an excellent vertexing and tracking to reconstruct the secondary vertices and to measure precisely the momenta of tracks. With the physical requirements, high-precision vertex detector with outstanding performance in terms of flavor tagging and track reconstruction is presented. Finally, the chapter terminates with an introduction of the Monolithic Active Pixel Sensors (MAPS) that is proposed to equip the vertex detector, i.e., MIMOSIS. As the main theme of this thesis is the radiation hardness design of the MVD, the presentation will mainly focus on the context of the MVD and the necessity of radiation hardness.

1.1 The CBM experiment

Darmstadt's GSI Helmholtz Center will be adequately reinforced for many years to come. The new facility in construction, Future International Antiproton and Ion Research Facility (FAIR), will offer unparalleled research chances in nuclear, hadronic, atomic and plasma physics, which is a multipurpose accelerator facility, based on that will supply primary beams (protons up to 90 GeV, Uranium up to 35 AGeV, nuclei with Z/A from 0.5 up to 45 AGeV) and secondary beams (rare isotopes and

antiprotons) with high intensity and quality. The facility layout depicted in figure 1.1 consists of a double-ring synchrotron, rings for accumulation, cooling, and storage of primary and secondary beams, as well as dedicated detector arrangements. The research programs incorporate studies of nuclear matter at extreme densities, studies of nuclei far from stability, hadronic physics with antiproton beams, the investigation of plasmas induced by ion and laser beams, as well as atomic physics [1][2].

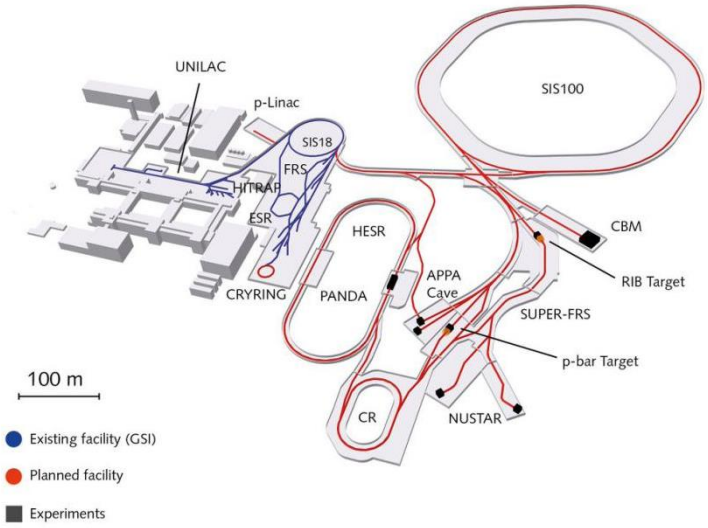


Figure 1.1 The planned FAIR facility.

1.1.1 The SIS Machine

In this updated version, the SIS-18 synchrotron will be accompanied by the new SIS-100 and SIS-300 ring accelerators, both of which operate along in the same underground tunnel with a circumference of 1084 m and magnetic rigidities of 100 and 300 Tm, respectively. The floor tunnel is 17 m deep below the surface. The SIS-100 and SIS-300 accelerators can deliver beam energies of up to 11 AGeV and 35 AGeV for Au beams, respectively. Therefore, SIS-300 is anticipated to reach into the region of the highest baryonic densities. However, the great superiority of the new facility doesn't consist in the energy, but in the beam quality and intensity. Both accelerators are prospective to offer 10^{11} ions per second and 10^{13} protons per second [3].

The 300 Tm machine will operate at bending fields of 6 T at ramp rates of up to 1T/s. Through utilizing high charge states (e.g., U92+), heavy ion beams will be accelerated in SIS300 to energies of $35 - 45\ GeV/u$, with a maximum intensity of 2×10^{10} ions per cycle and allows to have long spills of up to $100\ s$ while maintaining a sufficient average intensity. Moreover, SIS300 can be operated as a stretcher, and both primary and secondary (radioactive and antiproton) beams are then injected, cooled, and stored in a system of storage rings with internal targets and in-ring experiments. Based on the phase-space cooling techniques applied in the storage rings, the future program will broadly capitalize on the little emittance and low momentum spread of beams in high precision experiments.

The FAIR synchrotrons will compress heavy-ion and proton beams down to very short bunch lengths demanded for the production and subsequent storage, as well as efficient cooling of exotic nuclei (~60 ns) and antiprotons (~25 ns). With the double-synchrotron facility, continuous beams with high average intensities of up to 3×10^{11} ions per second are offered at energies of 1 GeV/u for heavy ions, either directly from the SIS100 or by transferring to the 300 Tm ring and slowly extracting from it. The SIS300 will supply ion beams of maximum energies around 45 GeV/u for Ne^{10+} beams and close to 35 GeV/u for fully stripped U^{92+} beams, respectively.

1.1.2 The Physics Program

Nowadays, the properties of the protons and neutrons are widely known, and the Standard Model theory and Quantum Chromo Dynamics (QCD) depict their fundamental interactions well. However, the current knowledge is still insufficient to forecast precisely the collective properties of multibody systems formed by these particles, especially in extreme conditions, for instance, in terms of temperature and pressure. The characteristics of this system currently require more in-depth theoretical and experimental research.

Of great concern is the existence of different stages of nuclear matter, among which the best known is the so-called liquid phase that is realized

in the ground state of the nuclei. At a given temperature, the binding force between nucleons is not enough to maintain the stability of the nucleus, and protons and neutrons escape into a vacuum. Due to the comparability to the evaporation of a liquid, the process is explained as a phase change to a nuclear gas.

The CBM experiment parses the phase transition forecasted by QCD, which is much higher than the temperature demanded to evaporate the nucleus. This is a phase transition from a hadron matter to the so-called Quark-Gluon Plasma (QGP). If more and more energy is diverted into the hadron gas, its temperature will go up to the point where thermal excitation of nuclear resonance happens. Despite extreme conditions, the quarks of a single hadron are limited.

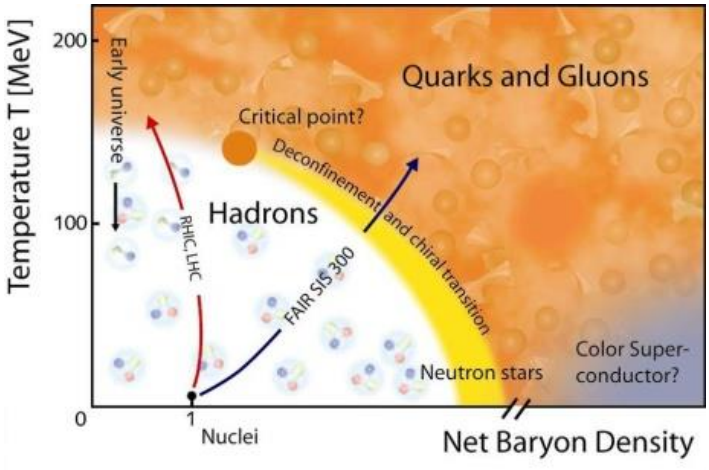


Figure 1.2 The phase diagram of hadronic matter as a function of the temperature (T) and the baryonic density (μ_b) [4].

The discrepancy between the hadronic matter and QGP is defined by the emergence of a deconfinement of quarks. It is predicted that such deconfinement will result in dramatically high particle density, which will bridge the distance between the quarks of adjacent hadrons compared with the quarks that form the hadrons. In these circumstances, the “potentials” of hadrons merge, allowing the quarks to move freely among the conventional “potentials” [5].

In QGP, the distance between disparate hadrons becomes shorter due to high compression or strong particle generation towing to high temperature,

which leads to the potential occurrence of single hadrons. In the potential of merging, the quarks of adjacent hadrons also attract quarks. In consequence, the barrier is weakened in the neighboring direction. At some point, the barrier between hadrons falls below the particle generation limit. As a result, a confined quark may now leave its partners and move towards the neighboring hadron. If the hadron density is high enough, the quark can be regarded as a quasi-plasmid-free equal to the conduction electrons in the metal. So far, this matter has reached the state of QGP.

1.1.3 The subdetectors of CBM Experiment

The CBM detector is devised with a fixed-target setup. A variety of subsystems is used to cover a wide range of possible heavy-ion studies at FAIR. The detector acceptance covers polar angles of 2.5°- 25° with full azimuthal coverage. It is necessary for the electronic components to be radiation hard to sustain the expected radiation dose. Besides, the sensors and support materials need to be constructed as thin as they can be to lessen multiple scattering. Moreover, fast operating sensors are required to support the highest beam intensities.

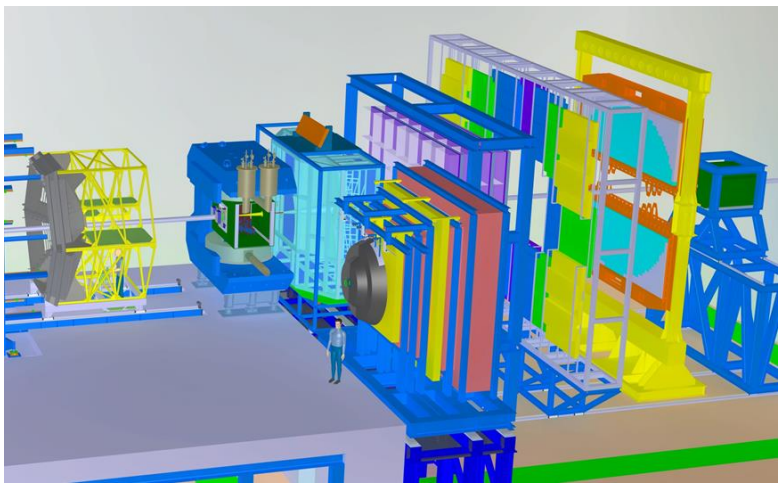


Figure 1.3 Structure of CBM detectors

The nucleus of the experiment will be a silicon tracking and vertex detection system installed in a large acceptance dipole magnet. The Micro-Vertex Detector (MVD) is required to determine high-precision secondary vertices for *D* meson recognition. The MVD consists of two layers of ultra-

thin and highly granular Monolithic Active Silicon Pixel Sensors (MAPS), located near the target.

The Silicon Tracking System (STS) is composed of a low material quality silicon microstrip detector and may be replenished by one or two layers of hybrid-pixel detectors to supply precise spatial point measurements. STS can execute track reconstruction over a wide range of momentum from about 100 MeV to over 10 GeV with a momentum resolution of about 1%.

The Ring Imaging Cherenkov (RICH) detector (for momenta < 8-10 GeV/c) together with Transition Radiation Detectors (TRD) (for momenta > 1.5 GeV/c) will be utilized for the survey of electrons. Muons will be measured by using an active hadron absorber system composed of an iron layer and a Muon Tracking Chamber (MuCh). For muon measurements, MuCh will take place in the RICH position. Charged hadron identification will be executed by a time-of-flight (TOF) detector. The Electromagnetic Calorimeter (ECAL) will offer information about photons and neutral particles. The Projectile Spectator Detector (PSD) is required for the measurement of the collision centrality and the orientation of the reaction plane.

1.2 The Micro Vertex Detector (MVD)

To illustrate the potential physics mechanisms of newly observed processes, the authentication of impact particles will play a crucial role. The Micro Vertex Detector (MVD) is the indispensable tool to achieve very high-performance particles tagging by reconstructing displaced vertices, and also plays a critical role in the track reconstruction, particularly for low momentum particles that do not reach the central tracker or barely penetrate its sensitive volume.

1.2.1 Requirements

The MVD requirements are mainly motivated by two competing sources of constraints: the physics goals and the running conditions.

The open charm physics program at CBM dominantly relies on D-mesons, which will be reconstructed by way of their hadronic decay

channels $D^0 \rightarrow K^- + \pi^+$and $D^0 \rightarrow K^- + \pi^+ + \pi^+$. Among the addressed D-mesons, D^0 is deemed to be the most difficult to reconstruct. The lifespan of the D−mesons is $c\tau = 123.0\ \mu m$ for D^0 and $c\tau = 311.8\ \mu m$ for $D^{\pm}$. Since the lifespan of charm mesons does not allow them to reach a detector, they must be reconstructed by detecting their daughter particles. Consequently, the design guidelines for the Mini Vertex Detector (MVD) and CBM's STS are based on the requirements determined by the observed object. The relatively low production multiplicity of open charm mesons raises additional requirements, that is an intense beam is required to produce a sufficient number of those particles in a reasonable amount of time. On the target, CBM's estimated collision speed is 10^7 collisions per second and approximately 10% of these hits can be regarded as central hits. The CBM physics goals dictate an unprecedented spatial three-dimensional point resolution and a very reduced material budget. At the same time, high granularity and fast readout compete and are inclined to increase the power dissipation, which will result in an augmented material budget in turn.

Table 1.1 The specifications of the MVD for CBM

Requirement	Goal
Spatial resolution	$5\mu m$
Material budget	$0.3\% - 0.5\% X_0$
Detection efficiency	99.5%
Hit rate (peak)	$7.5 \times 10^7 hits/cm^2$
Time resolution	$50\mu s$
Ionizing radiation hardness (Without safety factor)	$50KGy$
Non-ionizing radiation hardness	$10^{14} n_{eq}/cm^2$

For purpose of detecting individual particles and their attenuation topologies, it is necessary to reconstruct the particle tracks passing through various sites of the silicon detector. Superconducting dipole magnets with a magnetic field of up to 1T will bend the tracks of the charged particles so that their momentum can also be reconstructed. An isolation envelope with

dimensions of $1400 \times 2000 \times 1100\, mm^3$ is needed to separate the tracking stations from the magnet, and the MVD will be placed in a vacuum vessel, but the vacuum environment is not conductive to the heat dissipation of electronic components. If it were not for the help of air convection, components would only dissipate heat through radiation and PCB conduction. Hence, designing additional heat dissipation structures is compulsory. However, this requirement goes against the low mass quality. A compromise solution is needed to balance these requirements.

1.2.2 Geometries

As shown in Figure 1.4, the MVD for CBM is composed of 4 stations of detectors and both sides of a station layer will attach detectors, distributed around the middle hole, where the beam will pass through. The front stations deploy fewer detectors than the rear stations, which means the front stations have a higher dose rate.

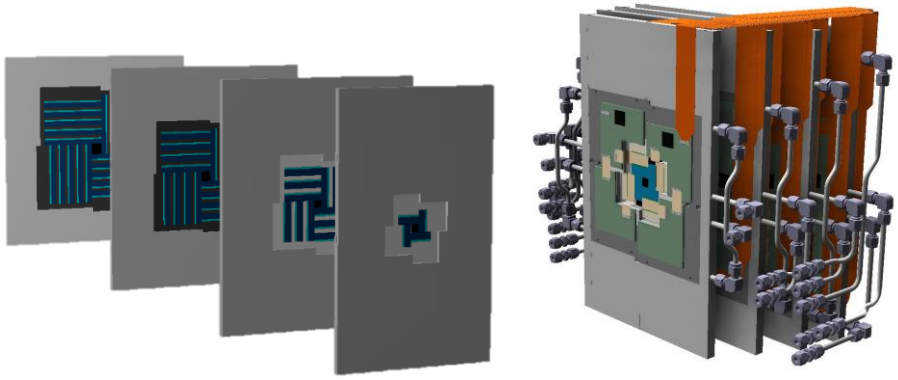


Figure 1.4 The geometry of MVD for CBM.

The design of station support is dictated by the cooling requirements and the material budget. As the equipment will be placed in a vacuum, it is necessary for the layer support to be designed as a part of the cooling system. As a result, in a vacuum, heat must be removed by heat conducting or liquid cooling. As shown in Figure 1.5, the support of the layer station is composed of carbon fiber material. Between two layers of carbon fiber material, there is a layer of “micro-pipes”, which is mainly realized by

liquid cooling and some separator walls, and the latter is utilized to ensure mechanical stability. [6]

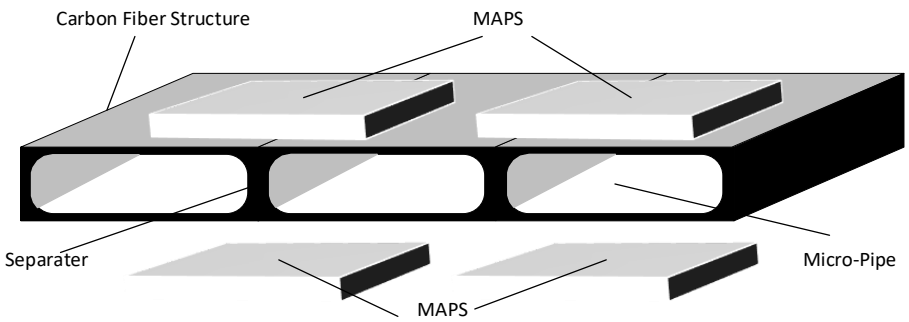


Figure 1.5 Cross-section of the proposed vertex detector station. Only some MAPS chips are shown for clarity. In reality, they are intended to match as close as possible.

1.2.3 The predicted particle environment for MVD

Figure 1.6 indicates the average number of hits in an MVD detector station as a function of its distance from the target, which comes from a detailed GEANT simulation study of Au + Au collisions at disparate incident energies. The dashed line illustrates the collision result of only the particles formed in the nuclear collision, and the solid line comprises delta electrons that shot down from the target by an electron beam [7].

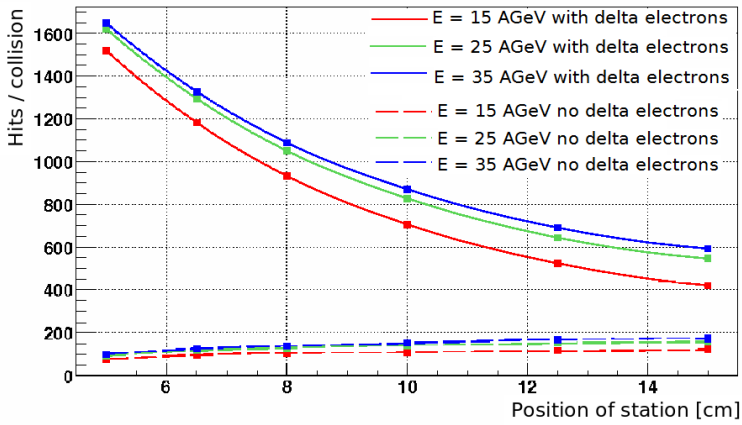


Figure 1.6 The average number of hits on one MVD detector station per Au+Au collision as a function of the position of this station. Two cases are shown: with (full lines) and without (dashed lines) the inclusion of the contribution of delta electrons (see text). Three different incident energies were studied: 15, 25 and 35 AGeV [7]

From Figure 1.6, it can be known that if incremental electrons are deleted from the analysis, the average number of hits per collision will be less than 200 and will be completely independent of the station's position (dashed line). On the other hand, when delta electrons are not excluded, for a station located 5 cm (solid line), the average number of hits per collision prevails, reaching 1600 hits/collisions, which means that the number declines as the distance from the target increases. On account of these results, it can be summarized that at such a high hit density (up to 3.5 hits/mm^2/collision), it is very fundamental to use a silicon pixel detector with high granularity.

The strong particle flux expected on the MVD will induce a considerable radiation dose, particularly for sites close to the target [6]. The expected ionizing and non-ionizing radiation doses in CBM experiments have been assessed through system simulation[6][8], which are normalized to a typical CBM operating year, assuming that the effective persistence of the operating year (i.e., the target beam) is about $5 \times 10^6 s$ (about two months). Assuming a nominal beam intensity of 10^9 beams per second and an interaction possibility in the target of 1%, it can be anticipated that approximately 5×10^{13} collisions will happen during this period.

The non-ionizing doses provoked by different particles are calculated by means of the NIEL factor, which normalizes the radiation damage caused by 1 MeV neutrons [9]. For the ionizing dose, it is assumed that all charged particles are minimum ionizing particles. It has been found that at 25 AGeV, for each Au + Au collision, the ionizing and non-ionizing radiation doses hitting the MVD detector station can reach 0.7 μGay and 30.2 n_{eq}/cm^2, respectively.

In order to meet the requirements for MVD radiation hardness, the numbers must be normalized according to the beam intensity. CBM will operate in two main phases: the first starts with the measurement of the SIS-100 synchrotron and the second begins with SIS-300, comprising the open charm measurement in the A+A collision. Currently, as for the second stage of SIS-300, it is anticipated that CBM-MVD will be used to conduct two stages of open charm measurement. The requirements for the detector

will be lower in the first phase and will be upgraded in the second. Due to diverse technical limitations, which will be discussed later in this thesis, it appears to be reasonable to operate the first generation MVD at a collision rate of about 10^5 collisions/sec. In the circumstances, the initial version of CBM-MVD on the SIS-300 should be able to withstand $10^{13}\ n_{eq}/cm^2$, and each CBM runs for approximately 30KGray [10].

1.3 The introduction of Sensors for MVD

With the continuing advance of semiconductor technology, silicon detector technology is persistently improved. In high-energy physics applications, silicon detectors show the advantages of high speed, small size, and high integration. Back in the early 1960s, silicon was used in radiation detection applications to replace gas detectors, which marked a revolution in particle physics experiments [11], because the moderate bandgap energy of silicon (~ 1.2 eV) is neither too high to generate a large number of charge carriers by ionized particles, nor too low to prevent large leakage currents from generating electron-hole pairs. Complementary Metal Oxide Semiconductor (CMOS) is a production process based on ideas emerging in the 1960s and prevailing in today's electronics market. The foundation is the Metal-Oxide Semiconductor Field Effect Transistor (MOSFET), which is a primary element in all modern electronic devices.

1.3.1 Introduction into semiconductors

Pure silicon with negligible impurities, i.e., intrinsic silicon, has the same number of electrons and holes. The attributes of silicon can be adjusted by doping other additional elements in the intrinsic material. Donors (e.g., As, P, or Sb) are introduced into intrinsic silicon, resulting in electrons increasing to form n-type silicon. Likewise, acceptors (e.g., Al, Ga, or Sn) cause holes to form p-type silicon, and both n-type and p-type materials are called extrinsic semiconductors. External doping introduces additional energy levels into the material. In consequence, by adjusting the doping type and concentration, the electrical nature of silicon can be intentionally changed.

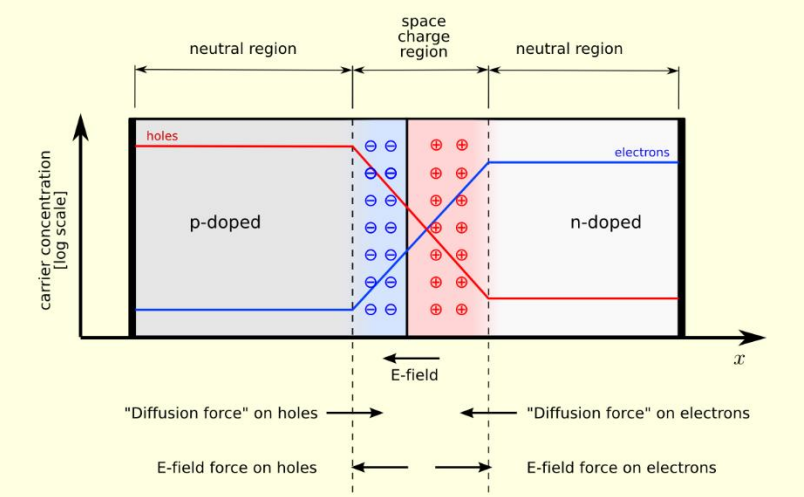


Figure 1.7. PN junction with zero bias thermal equilibrium. The electron and hole concentrations are indicated by blue and red lines, respectively. The gray area is electrically neutral. Bright red areas are positive, and bright blue areas are negative. The electric field is shown at the bottom. Electrostatic forces act on electrons and holes and their diffusion orientation [12].

As the most crucial structure in a semiconductor device, a p-n junction is constructed by joining two opposite doping types extrinsic semiconductors together. Diodes are easy to implement by using p-n junctions. As a two-terminal device, the feature of the diode is that in the case of an external voltage bias is applied, the current can only be conducted in one direction. For the sake of diode characterization, the analysis begins with the diode reaching thermal equilibrium. In the state of thermal equilibrium, the distribution of electrons and holes is consistent. Nevertheless, once the two semiconductors are connected, the discrepancy in concentration will bring electrons to diffuse from n-type silicon to p-type silicon, while holes will diffuse in opposite directions. The ionized donor retained in n-type silicon and the ionized acceptor in p-type silicon will have the consequence the building of an electric field, which will counteract the diffusion of electrons and holes. When the diffusion is compensated by the electric field, the diode is in dynamic thermal equilibrium and a neutral region without moving carriers is shaped as a depletion region. The depletion region of the diode can assemble the free charge generated very quickly. So, the reverse-biased diodes are implemented as charge collection nodes for silicon detectors.

1.3.2 Introduction to silicon sensors

When passing through matter, charged particles lose energy by interacting with the electromagnetic fields inside the atoms and scattering with electrons and nuclei, which is described quantitatively in [13]. As the valence electrons are loosely combined with the silicon atom, the impacting particle generates an electron-hole pair cloud along its path. The charge carriers diffuse around until they lose the residual energy and recombine. If they were collected before the recombination, impacting particles could be detected. In a silicon detector, electron-hole pairs formed by incident charged particles are collected by an electrode. The travel of these charged carriers is controlled by the built-in electric field (drift) and the distribution (diffusion) of non-uniform charge carriers. Both mechanisms affect charge collection.

The principle of silicon detectors leans upon collecting the charges from incident particles, so fast signal response and high charge collection efficiency are fundamental. In the depletion region, drift is the elementary mechanism in the motion of charge carriers, which is more rapid and has less possibility of charge trapping compared with diffusion. Therefore, full depleted silicon detectors are favored for particle detection. Nevertheless, the use of fully depleted structures is commonly restricted by operational circumstances. The width W of the depletion region is estimated as

$$W = \sqrt{\frac{2\varepsilon_S}{q}\left(\frac{1}{N_A}+\frac{1}{N_D}\right)V_b} \tag{1.1}$$

In the equation, ε_S is the relative dielectric permittivity of the semiconductor, N_A and N_D is the doping spatial densities and q is the electron charge. As described in Eq. (1.1), the width W of the depletion region is proportional to the square root of the reverse-biased voltage V_b. For the sake of full depletion of the silicon, an external high voltage is required, but this high voltage is normally unavailable in the standard CMOS process owing to the limitation in breakdown voltage. Furthermore, the additional high voltage distributed to each sensor complicates the design of detectors. As a result, partial depletion is ordinary in silicon

detectors, such as in CMOS pixel sensors (CPS). In partially depleted silicon, drift is merely available in the vicinity of the anode of a diode. As the impurities present in the silicon bulk could trap charge carriers, the quantity of the collected charge depends on the carrier's lifespan. The longer charge collection time, the more loss of charge carriers. The different mechanisms in charge carrier's collection will affect charge collection time. In fully depleted detectors, charge collection time depends on the strength of the electric field, while in partially depleted ones, diffusion constant and doping concentration contribute mainly to charge collection time. Typically, a partially depleted n-type silicon with a thickness of 300 μm and 10 $k\Omega \cdot cm$ resistance value has a typical charge collection time of approximately 30–90 ns, three times longer than that of the fully depleted structure.

1.3.3 The Pixel Sensors for MVD

With the advance of physics and the progress of various High Energy Physics (HEP) experiments, various types of silicon detectors have been proposed and established, including strip detectors, hybrid pixel detectors, DEPleted Field Effect Transistor (DEPFET) detectors, and Monolithic Active Pixel.

- **The strip detectors**

The strip detectors are the first detector devices using the lithographic capabilities of microelectronics, which can be found in all high energy physics experiments of the last two decades. As shown in figure 1.8, the detector utilizes the arrangement of strip implants acting as charge collecting electrodes, placed on a low doped fully depleted silicon wafer; these implants create a one-dimensional array of diodes. A position-sensitive detector is built by connecting each of the metalized strips to a charge sensitive amplifier, and two-dimensional position measurements can be implemented by applying additional strip-like doping on the wafer backside (double-sided technology).

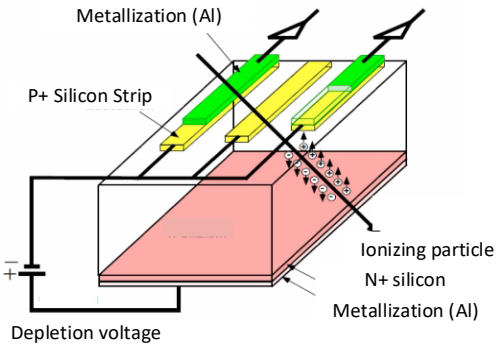


Figure 1.8 The principal of the silicon strip detector.

The structure of the silicon strip detector is simple and easy to implement, has also clear issues. Since the silicon strip detector can only tell one dimension of information at a time, the high particle fluences ambiguities create difficulties for the track reconstruction. In consequence, deriving the point resolution from just one coordinate is not enough information to reconstruct a secondary vertex.

- **Hybrid Pixel Sensor (HPS)**

Meanwhile, pixel detectors allow track reconstruction at high particle rates without ambiguities. The Hybrid Pixel Sensor (HPS) is a classical idea for HEP as shown in figure 1.9, where the readout chip is mounted directly on top of the pixels by bump-bonding. Now that each pixel has its own readout amplifier, as there are two wafers, we can choose the appropriate process for sensor and readout chip separately. Due to the parallel working of the readout channels, the readout speed is very fast, and the separated wafers help to gain a high radiation tolerance performance.

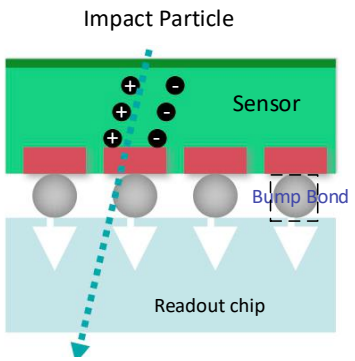


Figure 1.9 The Principal of Hybrid Pixel Sensor

The defect of HPS is that the pixel area is defined by the size of the readout chip. Generally, the readout part is larger than the sensor part, which limits the improvement of spatial resolution. Another shortcoming is that because there are two chips, the material budget is high, and it is the same with the power consumption.

- **DEPFET Sensor**

A DEPFET Sensor is composed of a field-effect transistor (FET) placed on the uppermost part of a high resistivity n-type completely depleted silicon bulk. As shown in figure 1.10, deep n-doping implantation, the "internal gate," situated under the transistor channel (around 1 μm below), generates a minimum of potential for the electrons. [14]

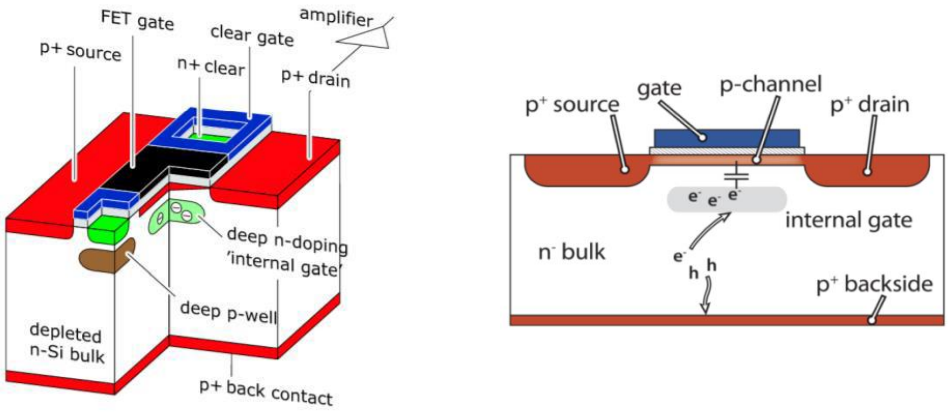


Figure 1.10 Sketch of a DEPFET pixel and the internal gate location.

The electrons generated by the incident particles are collected by the internal gate, and the transistor channel current is regulated by the internal gate potential changes. When the electrons on the internal gate need to be reset, a positive voltage pulse will be applied to the transparent gate.

The DEPFET has a higher amplification factor because of collecting the charges by its gate instead of the active part (source or drain). The internal gate also results in the disadvantage of weak radiation hardness, which is hard to manufacture in an industrial process. The reset of the internal gate requires a high power, which sets the limitation for high integration.

- **Monolithic Active Pixel Sensor (MAPS)**

With the continual advance of deep-well technology and double well technology, it is feasible to integrate signal sensing and readout electronics on a homogeneous substrate. In 2000, researchers of the LEPSI and IReS came up with a Monolithic Active Pixel Sensor (MAPS) for charged particle tracking and imaging, which have offered an attractive balance among granularity, material budget, radiation tolerance, and readout speed[15] [16].

As shown in figure 1.11, the first generation of MAPS has been made of a photodiode by the P-N junction created between the N-well and the epitaxial layer (P-type). The doping level of the P well (P + type) and the P substrate (P ++ type) is higher than that of the epitaxial layer (P-type), bringing about two potential barriers. The first potential barrier is situated in the boundary region between the epitaxial layer and the P-well, and the second, in the boundary area between the epitaxial layer and the P substrate.

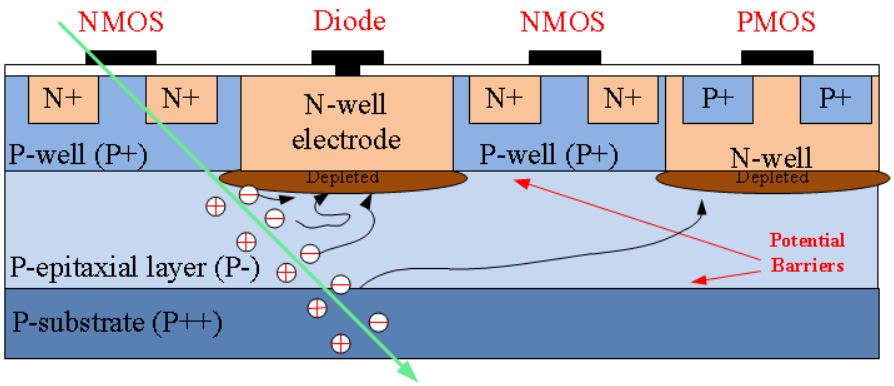


Figure 1.11: The principal of the first-generation MAPS structure.

When charged particles cross the MAPS detector, the energy emitted by the particles will be converted to e-/hole pairs, normally generating about 80 electron-hole pairs per micrometer. The electrons generated in the epitaxial layer thermally diffuse and converge at the N-well collector electrode. Owing to the high level of P-type doping, more electrons produced in the substrate will recombine, and only a few electrons will be transferred to the epitaxial layer and collected by the collection electrode.

In the first-generation MAPS structure, only NMOS transistors could be used in pixels, because of N-wells that would host PMOS transistors

would compete for charge collection with N-well / P-epi diodes. To avert this, the solution is to use a deep P-well in the P substrate to segregate the N-well from the epitaxial layer or substrate in the second-generation structure, as shown in Figure 1.12. The deeper P-well prohibits the N-well from collecting electrons, so a full CMOS process can be used in pixels [15][17][18].

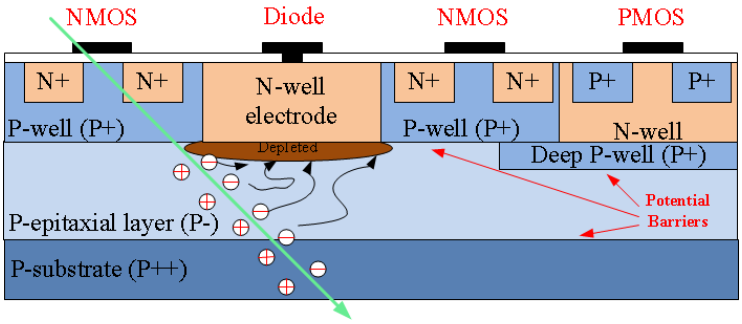


Figure 1.12: The principal of the second-generation MAPS structure.

The MAPS structure was used for the first time in the vertex detector upgrade of the STAR (Solenoidal Tracker at RHIC) experiment at RHIC (Relativistic Heavy Ion Collider). The ALPIDE utilized in the ALICE experiment is a monolithic pixel sensor that is manufactured in the 180 nm CMOS imaging sensor process of TowerJazz, which achieves detection efficiency above 99%, a power consumption less than 40 mW/cm^2, and a dimensional resolution of around 5 μm. These characteristics meet the excessive requirements of the ALICE experiment, even after neutron irradiation to 1.7×10^{13} n_{eq}/cm^2 at 1 MeV [19].

Table 1.2: Advantages and disadvantages of various detector technologies

Technology	Advantages	Disadvantages
Strip detector	Large area possible Fast speed	Expensive cost
Hybrid pixel	Low detector capacitance Low leakage current large signal-to-noise ratio High radiation-hard High Readout Speed	Large power consumption Expensive cost
DEPFET	Integrated amplification Low power consumption Low noise Fast readout speed	Only small devices Not radiation-hard High power at reset No industry standard
MAPS	Standard IC process Integrated amplification Low noise Low power consumption High spatial resolution	Moderate Readout Speed Moderate radiation tolerance

- **Conclusions**

In consideration of the low material budget, low power consumption, and high spatial resolution requirements mentioned above, a MAPS-based detector is more appropriate for CBM-MVD application scenarios than other types of detectors. The MAPS process is based on the standard IC process with only one thin wafer used, which makes it easy to manufacture and has very low material quality. With the limitation of lower power consumption, it still has high spatial resolution and moderate readout speed, as well as certain radiation resistance under the current process, but it still needs to be optimized.

1.4 The MIMOSIS in MVD for CBM

The MIMOSIS is now being devised as the CMOS pixel sensor in the MVD for the CBM experiment. The elementary design of the pixel's matrix is based on the ALPIDE sensor [14], which has been developed for the ALICE ITS upgrade. Nevertheless, the rate capability and the radiation

tolerance of MIMOSIS have to be drastically beyond the ones of ALPIDE, which have introduced the need for a dedicated CPS design [20].

1.4.1 Design of the Pixel

MIMOSIS leans upon the TOWER/JAZZ 180 nm quad-well technology, which allows to integrate a complete amplifier-shaper-discriminator chain into each pixel. The pixel matrix will comprise 1024 columns of 504 pixels with a size of 26.88 × 30.24 μm^2, and the block diagram of its layout is shown in Figure 1.13. The DC-coupled version of the pixels is inspired by ALPIDE's related pixels. Nevertheless, different from ALPIDE, which is optimized for triggering, MIMOSIS will run in a constant readout according to CBM requirements. Hence, the trigger logic is substituted by a different type of sample-and-hold logic, which will hold the potential signal for a frame time. This adjustable frame time is controlled by the global shutter signal, which is 5 μs by default. Once the global shutter reaches the pixel, the result of the sample-and-hold circuit will shift to the output buffer and clear the circuit. In the general case where the timeout threshold of a pixel goes beyond the frame time, digital edge detection will be used to avert double counting, which is necessary because, compared with earlier MIMOSA sensors, the readout rate of the MIMOSIS is extremely high. [21] This mechanism helps to minimize the data rate.

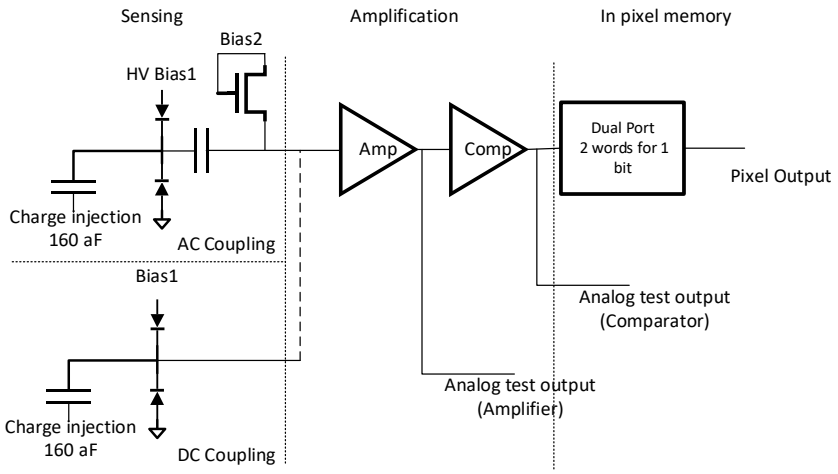


Figure 1.13 Block diagram of the pixel of MIMOSIS.

Besides the DC version, it is also considered to equip MIMOSIS with the AC-coupled pixel version, which is shown in Fig.1.13. Both pixels are analogical, but the transistor of the amplifier on the pixel of the AC pixel can be preserved from depletion voltage by a capacitor. As a result, the depletion voltage of the collection diode is not constrained by the voltage that can be accepted on the gate of the transistor and can be incremented up to +40V while keeping the active medium at the mass potential. AC pixels are compatible with the modified 180 nm TOWER/JAZZ process, which is optimized for enhanced depletion [22] and is anticipated to significantly fortify the sensor's radiation tolerance.

1.4.2 Buffer structure

As shown in figure 1.14., the buffer structure of MIMOSIS must meet the requirement to read any record hits without reducing the data rate under the help of a trigger system. The reading of pixels is performed by a priority encoder, which works for two columns at a reading frequency of $20\ MHz$. The time resolution of the system is subject to the frame length, and the latter is expected to be 5 μs. Other values can be chosen within some restrictions at the cost of reducing the hit rate capability. The maximum number of hits recorded from the double column is consistent with the number of clock cycles within the frame time. If the number of hits surpasses this number, the unread hits will be abandoned.

The data of eight double columns will be supplied to a so-called region buffer by way of a data compression unit that identifies groups of up to four adjacent emission pixels in the reading order, which corresponds to a cluster found within the double-column, but it is not applicable to identify clusters spanning more than one double-column. The region buffer is devised as a double buffer, where one element applies a 16bit × 20 MHz bus to read the priority-encoded data from the double column, and the second buffer writes the data of the antecedent frame to the next level. The region buffer adds a dedicated "region" header, which decreases the number of bits required to encode the number of matching columns. The depth of the region buffer is 100 words, consistent with the theoretical maximum output of a single double-column, but it is usually shared among

these columns. This sharing symbolizes the first average of fluctuations in occupancy rates incurred by Poisson fluctuations.

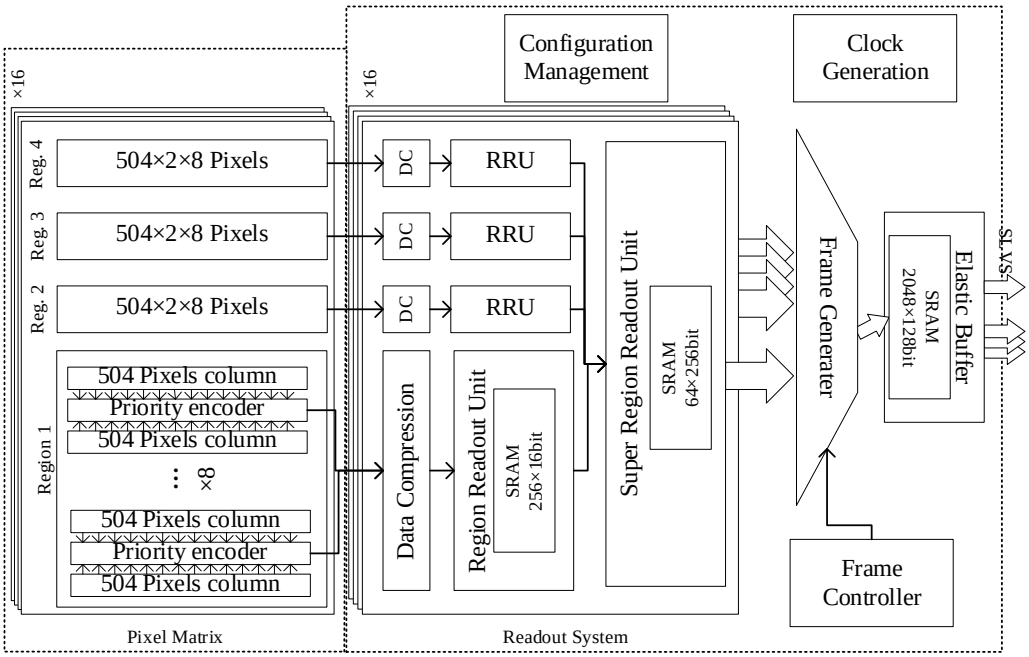


Fig. 1.14. MIMOSIS functional organization, with the three layers buffer (RRU, SRRU, and Elastic Buffer).

The 16 super regions converge data from the four regions through a 32-bit × 40 MHz bus with priority encoded with a depth of 256 words, whose purpose is to load balance and prepare data for transmission by way of the 256bit × 40 MHz main bus. The bus will transmit data to a central elastic buffer through the idle word removal. The intention of idle words removal is to eliminate possible idle words, which are created to fill messages with a width of more than 16 bits in the bus. Up to 3200 words can be processed every 5 μs, and the data is thereby written to the elastic buffer with an input bandwidth of 3200 words per 5 μs and depth of 16384 words, which is enough to store the peak data rate corresponding to 5 frames. The elastic buffer constitutes the key element of time and space load balance, which transmits data to up to 8 adjustable 320 Mbps output buffers. In the case of low data load, the number of output buffers can be decreased to save power and material budget for routing.

1.4.3 The development schedules

The final MIMOSIS sensor will be built through a dedicated R&D program, which is currently regarded to comprise three prototype submissions.

The first prototype, MIMOSIS-0, was presented in 2018, and both electrical function and radiation tolerance are tested. The sensor comprises 16 double columns, each with 2×504 pixels, half of which are realized by AC and the other half by DC coupling. A priority encoder for pixel readout and a DAC needed for pixel control is carried out. A dedicated number of pixels can be read through an exclusive analog readout line, which allows direct observation of the analog signal in the extensive chain within the pixel. The test results indicate that both AC and DC coupling pixels sense with low noise, allowing the threshold to be set to <150 e, and the noise can be further decreased by applying a slight reverse bias of -1 V to the p-well structure. The pulse shape and length of the amplification chain within the pixel were studied, and it was summarized that it is conceptually adequate to achieve a time resolution of ~1 μs combined with a lead time of ~10 μs, which is quicker than the ideal frame time 5 μs.

The second prototype, MIMOSIS-1, was submitted with in 2020 and is now under testing. MIMOSIS-1 is the first reticle size prototype with full 1024×504 pixels. In this thesis, we will mainly discuss the design of radiation hard structures in these prototypes.

The final prototype is MIMOSIS-2, which will mainly correct the errors in MIMOSIS-1. Moreover, new modifications could be brought to this prototype if the radiation resistance does not meet the requirements of MVD.

1.5 The strategy and focus of this work

The main target of this thesis is to find an appropriate radiation harden design method for MVD. The backdrop and motive of the thesis have been shown in the first chapter, which explains the objective, requirements and detector composition of the CBM experiment, the principle, structure,

manufacture processing, working environment and the urgent radiation requirements of the MVD.

In the second chapter, the mechanism of radiation and its effect on CMOS detector (single event effects, total ionizing dose, and displacement damage) will be introduced in detail, and then the principles and trade-offs of the main methods of radiation harden design will be described in Chapter 3. Afterwards, there will be two part of the detailed design. The first part (chapter 4) is the data buffer in the readout circuit, which symbolized the design method of the digital circuit. The second part (chapter 5) is the phase-locked loop, which is used to generate the clock for the system and represents the analog part. The sixth chapter is the test design and results of the previous two parts, including electrical performance test and radiation tolerance test. Finally, there is the summary of this thesis and the prospect of the future work direction.

1.6 Bibliography

[1] P. Senger, *The CBM experiment at FAIR*, *J. Phys. Conf. Ser.* **50**, (2006) 357.

[2] *GSI - "The Machine"*https://www.gsi.de/en/researchaccelerators/fair/the_machine.htm (26 February 2020).

[3] P. Spiller, *The Superconducting SIS100 Synchrotron for High Intensity Proton and Heavy ION Beams*, in *HB 2008 - 42nd ICFA Adv. Beam Dyn. Work. High-Intensity, High-Brightness Hadron Beams*, (2008).

[4] *Phasendiagram.jpg (460×310)*https://fias.uni-frankfurt.de/~hossi/Bilder/BR/Plotl/Phasendiagram.jpg (27 February 2020).

[5] G. Rosner, *Future Facility: FAIR at GSI*, *Nucl. Phys. B - Proc. Suppl.* **167**, Elsevier, (2007) 77.

[6] M. Deveaux, *Development of fast and radiation hard Monolithic Active Pixel Sensors (MAPS) optimized for the D 0 detection with the CBM - vertex detector*, *Building*, (2007).

[7] C. A. Dritsa, *Design of the Micro Vertex Detector of the CBM experiment : Development of a detector response model and feasibility studies of open charm measurement*, (2011) 207.

[8] D. Bertini, *Radiation dose simulations*, *CBM Collab. Meet.*, (2009).

[9] A. Akkerman, J. Barak, M. B. Chadwick, J. Levinson, M. Murat and Y. Lifshitz, *Updated NIEL calculations for estimating the damage induced by particles and γ-rays in Si and GaAs*, *Radiat. Phys. Chem.* **62**, Pergamon, (2001) 301.

[10] M. Deveaux, S. Amar-Youcef, N. Bialas, D. Doering, M. Domachowski, H. Düring, I. Fröhlich, T. Galatyuk, J. Michael, et al., *Status of the Micro Vertex Detector of the Compressed Baryonic Matter Experiment*.

[11] G. Lutz and G. Lutz, *Introduction*, in *Semicond. Radiat. Detect.*, Springer Berlin Heidelberg, (2007).

[12] W. Shockley, *The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors*, *Bell Syst. Tech. J.* **28**, (1949) 435.

[13] M. Tanabashi, K. Hagiwara, K. Hikasa, K. Nakamura, Y. Sumino, F. Takahashi, J. Tanaka, K. Agashe, G. Aielli, et al., *Review of Particle Physics*, in *Phys. Rev. D* **98**, American Physical Society, (2018).

[14] M. Boronat, *DEPFET pixel detector for future e-e+ experiments*, *Nucl. Part. Phys. Proc.* **273**–**275**, Elsevier B.V., (2016) 982.

[15] R. Turchetta, J. D. Berst, B. Casadei, G. Claus, C. Colledani, W. Dulinski, Y. Hu, D. Husson, J. P. Le Normand, et al., *Monolithic Active Pixel Sensor for charged particle tracking and imaging using standard VLSI CMOS technology*, *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **458**, Elsevier Science B.V., (2001) 677.

[16] C. Hu-Guo, J. Baudot, G. Bertolone, A. Besson, A. S. Brogna, C. Colledani, G. Claus, R. De Masi, Y. Degerli, et al., *CMOS pixel sensor development: a fast read-out architecture with integrated zero suppression*, *J. Instrum.* **4**, (2009) P04012.

[17] W. Snoeys, *CMOS monolithic active pixel sensors for high energy physics*, *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **765**, Elsevier, (2014) 167.

[18] C. Cavicchioli, P. L. Chalmet, P. Giubilato, H. Hillemanns, A. Junique, T. Kugathasan, M. Mager, C. A. Marin Tobon, P. Martinengo, et al., *Design and characterization of novel monolithic pixel sensors for the ALICE ITS upgrade*, *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **765**, Elsevier, (2014) 177.

[19] G. Aglieri Rinella, *The ALPIDE pixel sensor chip for the upgrade of the ALICE Inner Tracking System*, *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **845**, Elsevier, (2017) 583.

[20] M. Deveaux, B. Arnoldi-Meadows, G. Bertolone, G. Claus, A. Dorokhov, M. Goffe, A. Himmi, K. Jaaskelainen, P. Klaus, et al., *Observations on MIMOSIS-0, the first*

dedicated CPS prototype for the CBM MVD, Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip. **958**, (2020).

[21] M. Deveaux, B. Arnoldi-Meadows, G. Bertolone, G. Claus, A. Dorokhov, M. Goffe, A. Himmi, K. Jaaskelainen, P. Klaus, et al., *Observations on MIMOSIS-0, the first dedicated CPS prototype for the CBM MVD, Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **958**, (2020) 1.

[22] W. Snoeys, G. Aglieri Rinella, H. Hillemanns, T. Kugathasan, M. Mager, L. Musa, P. Riedler, F. Reidt, J. Van Hoorne, et al., *A process modification for CMOS monolithic active pixel sensors for enhanced depletion, timing performance and radiation tolerance*, *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **871**, Elsevier B.V., (2017) 90.

Chapter 2. The Radiation Effects on CMOS Sensors

In 1947, J. Bardeen, W. Brattain, and W. Shockley invented the semiconductor transistor, taking the lead in a new era in semiconductor microelectronics. In 1959, J.S. Kilby invented the first solid-state integrated circuit, which turned out to be a milestone in the history of microelectronics. From then on, the microelectronics industry has experienced tremendous progress over the past 40 years, particularly in the development of the performance of products (i.e., integrated circuits), while considerably reducing the manufacturing costs through fundamental integrated functions. The persistent decrease in the silicon surface occupied by these essential components retains the speed of integration at speed stipulated by the famous "Moore's Law," which states that the number of transistors per integrated circuit will duplicate every 18 to 24 months.

Dating back to the early 1960s, research done at the U.S. Naval Research Laboratory (NRL) has shown that MOS devices were sensitive to radiation. Prior to this, there were thoughts that MOS transistors (most-carrier devices) were not as sensitive to radiation as bipolar transistors. At that time, MOS devices were used in the first reconnaissance satellite around the world by NRL due to their high-impedance input and low current characteristics. By using γ-ray radiation from the Cobalt-60 source mainly for the passive oxide layer of the bipolar transistor, NRL began the study of the surface effect mechanism incurred by radiation.

As the dimension of the MOSFET decreases, it has been found that the sensitivity of integrated circuits to radiation from natural spaces or ground environments will be enhanced dramatically [1][2][3]. Nowadays, for processes that extremely scaled down, natural radiation is one of the reasons for the highest malfunction rate among all equipment and circuit reliability issues in the field of nanoelectronics [4][5].

2.1 Basic Radiation Mechanism on Silicon Sensors and Integrated Circuits (IC)

In order to understand the radiation effect caused by the collision of incident particles with the silicon semiconductor, we need first to grasp its physical process. The interaction between a particle and a substance is usually depicted by a single collision between the incident particle and one in the matter. The incident particles are mainly considered to be electrons-positrons, protons, and ions, and the reference material can be an intrinsic semiconductor material, i.e., silicon.

2.1.1 Electrons and positrons

- **Elastic collisions**

Elastic collisions, caused by the Coulomb interaction of electrons with the nuclear field, have for consequence changes in the direction of electron and positron motion, but are not responsible for energy loss. Thus, this part will not cause ionization.

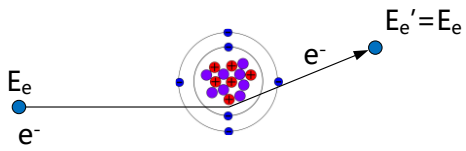


Figure 2.1 Diagram of electronic elastic collision. The electrons are scattered by the nucleus, leaving their energy almost unchanged.

- **Inelastic collisions**

Inelastic collisions and the emission of braking radiation will lead to the loss of energy from electrons (and positrons) in the matter, and this is the

primary energy loss mechanism for low to medium energy impact particles (up to 10 MeV). Inelastic scattering is the consequence of the Coulomb interaction between incident electrons and atomic electrons. Some of the energy and momentum of the incident electron is transferred into the target system, and the final state of the interaction can contain not only single-electron excitation or atomic ionization (accompanied by the generation of electron-hole pairs), but also many atoms in the solid [6].

Bethe [7] first obtained the quantum mechanical computation of inelastic collision based on the Born approximation. Special calculations for the particular cases of incident electrons and positrons will respectively bring about the theory of electron energy loss and the special formulas of Møller and Bhabha [8] for inelastic electron-positron scattering. The energy loss of the electron/positron can be expressed as

$$-\frac{dE}{dx} = K\frac{Z}{A}\frac{1}{\beta^2}\left[\ln\left(\frac{\tau^2(\tau+2)}{2(I/m_e c^2)}\right) + F(\tau) - \delta - 2\frac{C}{Z}\right] \tag{2.7}$$

In the formula, τ is the kinetic energy of the particle in units of $m_e c^2$, δ is the density correction, C gives the shell correction, Z and A are the number of protons and nucleus, β is the speed in unit of c, and $F(\tau)$ is different for electrons and positrons. For a detailed treatment of the energy loss of electrons and positrons, see also [9][10]. Section 2.1.3 incorporates energy loss formulae for protons and ions, besides the pertinent definitions of stopping power.

The energy loss caused by ionization can also be contrasted with the radiative component of electron energy loss so that the relative significance of the two energy loss mechanisms can be assessed based on the electron energy.

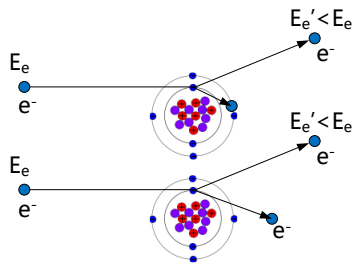


Figure 2.2 Diagram of inelastic electron collisions leading to atomic excitation or ionization.

- **Braking radiation (Bremsstrahlung)**

Electrons and positrons that pass-through substance undergoes deceleration and emit radiation when they interact with the electrostatic field of the atom. This effect is called bremsstrahlung or braking radiation, as shown in Figure 2.3.

At medium energies, without screening, the energy loss can be approximated as [10]

$$-\frac{dE}{dx_{Brem}} = 4Nr_e^2\alpha Z^2\left(\ln\frac{2E}{m_ec^2} - \frac{1}{3} - f(Z)\right) \tag{2.8}$$

and at high energies, for complete screening, as

$$-\frac{dE}{dx_{Brem}} = 4Nr_e^2\alpha Z^2\frac{E}{m}\left(\ln\frac{183}{Z^{1/3}} + \frac{1}{18} - f(Z)\right) \tag{2.9}$$

In the formula, If the logarithmic term is neglected, the energy loss rate can be approximately equivalent to the energy of the incident electron:

$$-\frac{dE}{dx_{Brem}} = \frac{E}{X_o} \tag{2.10}$$

In the formula, X_o is the radiation length and represents the distance over which radiative emission decreases the initial projectile energy by a factor $1/e$. The dependence on $1/m$ makes braking radiation emission much more significant for electron than that for more massive particles such as muons, protons or ions, while the dependence on Z^2 makes it extremely significant for high-Z materials, with important consequences for shielding strategies.

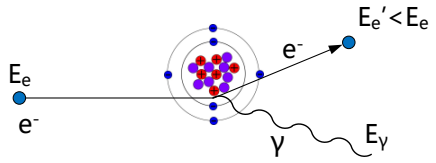


Figure 2.3 Diagram of the electron-induced radiation emission process. The electron emits a photon by the deceleration induced by the atomic electrostatic field.

2.1.2 Protons and ions

Finally, the effects for protons and ions passing through matter can be summarized as particle deceleration (energy loss) and deflection. The main interactions that lead to these effects are elastic collisions with atomic nuclei and inelastic collisions with atomic electrons in the matter.

- **Elastic collisions**

The elastic scattering of protons and ions in matter occurs because of the interaction with the nuclei of the matter shielded by the electron cloud. This process is the same as the elastic scattering of electrons mentioned above, and the main discrepancy is the non-negligibility of the mass of the projectile. Although the total kinetic energy is retained, the kinetic energy has been partially transferred from the projectile to the target core.

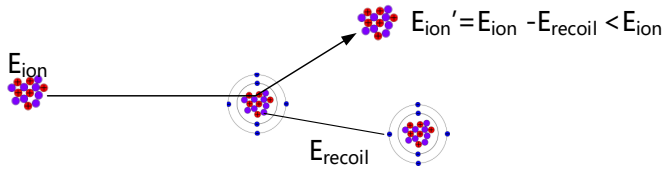


Figure 2.6 Diagram of the elastic electromagnetic interaction of an incident ion with a target atom, screened by the electron cloud, with the transfer of part of the energy to the target nucleus.

- **Inelastic collisions**

Ion electromagnetic inelastic collisions are interactions of the incident ions with the field of the atomic electrons in the matter, the result of which is the excitation or the ionization of the target atom.

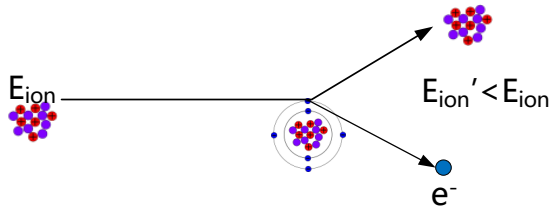


Figure 2.7 Diagram of inelastic electromagnetic interaction of incident ions with electrons of the target atoms, inducing atomic excitation or ionization.

2.1.3 Energy loss, linear energy transfer and stopping power

Stopping power is the average energy loss per unit path-length as a charged particle passes through a material, for which the typical units are MeV/cm or $KeV/\mu m$. What can also be ignored is the dependence of the stopping power on material density with reference to mass stopping power, providing the average energy loss per unit areal mass, for instance, $MeV\ cm^2/mg$.

Generally due to the inelastic interaction of the projectiles with the field of the electrons in the target nucleus, a major part of the energy loss of electrons and ions in a material is widely known as electronic stopping power or linear energy transfer (LET). The quantity of energy transferred to the electrons in each collision for ions is a small fraction of the projectile energy, but the number of collisions in media of average density is high, giving rise to a significant loss of kinetic energy. Collisions with a finite transfer of energy are described as soft, while less frequently hard collisions occur, which cause atomic excitation with the ejection of energetic, fast electrons (often referred to as δ-rays). The part of the electronic stopping power that does not contain the production of δ-rays above a given threshold Δ is denominated as restricted stopping power.

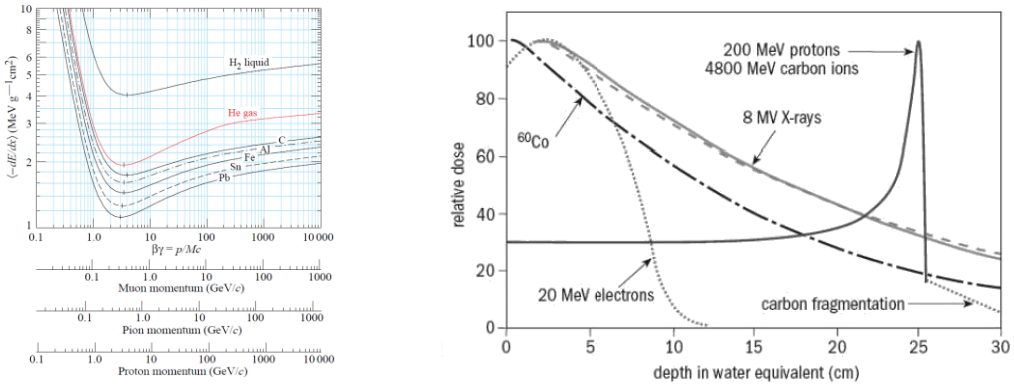
Energy loss by collisions imparting energy to the atomic nuclei, contrary to the above-mentioned electrons, is often named nuclear stopping (and the name often causes misunderstandings, as no nuclear/hadronic interaction is involved). Nuclear stopping is much smaller than electronic stopping, generally making up less than 0.1% of the total energy loss.

Bethe and Bloch et al. [11] first gave an accurate quantum mechanical presentation of the energy loss phenomena. The formula below can approximate the resulting (electronic) stopping power

$$-\frac{dE}{dx} = Kz^2\frac{Z}{A}\frac{1}{\beta^2}\left[\ln\left(\frac{2m_e\gamma^2v^2W_{\max}}{I^2}\right) - 2\beta^2 - \delta - 2\frac{C}{Z} + zL_1 + z^2L_2\right] \quad (2.12)$$

In the formula, W_{max} is the maximum energy transfer in a single collision, and I is the mean excitation potential of the target material. Two correction terms are generally deemed as the density correction δ and the shell effect

given by C/Z . The extra terms $zL_1 + z^2L_2$ represent the Barkas and the Bloch corrections.



(a) (b)

Figure 2.8 (a) Mean energy loss rate in various materials [12]. *Extra axes have been added to convert $\beta\gamma$ to kinetic energy for muons, pions and protons. (b) Qualitative comparison of depth dependence of deposited dose for protons, carbon ions, gammas, x-rays* [13].

Figure 2.8 (a) indicates the mean energy loss rate in diverse materials from [12], as a function of $\beta\gamma$. Extra axes have been appended specifically for muons, pions and protons, and particle loses energy at a minimum rate at approximately $\beta\gamma = 3{\sim}4$ whichever the material. The energy loss rate goes up at higher energies on account of the relativistic extension of the transversal electric field ($\ln(\beta\gamma)^2$, with saturation owing to the density effect, the δ correction), while at low energies, when particles slowdown in the materials, it augments as $1/\beta^2$. The latter behavior is the source of what is called Bragg peak, i.e., an increasing energy loss rate along the proton or heavy ion path when the particles are close to stopping. These characteristics are increasingly applied in cancer therapy for helping to maximize the dose on target while sparing healthy tissues, especially in contrast to an electron or gamma irradiation, as shown in Figure 2.8b. In modern microelectronics, the Bragg peak can trigger single event phenomena. Muons, on account of their relatively big mass compared with the mass of electrons, behave in an analogous way to protons in energy loss mechanisms, although they afterwards experience dissimilar processes when at rest.

2.2 Single event effects

Single Event Effects (SEE) are the consequence of the interaction of high-energy particles (such as protons, neutrons, alpha particles, or heavy ions) in a microelectronic device or circuit. A SEE may intervene in the operation of the device/circuit (for instance, inverting or flipping the data state of memory cells, latches, flip-flops, etc.), or ultimately do harm to the circuit (for example, gate oxide cracks, destructive latch-up event). For more than four decades, people have widely understood the mechanism, and have characterized and modelled the interaction of protons or heavy ions in semiconductor devices to produce SEE [14]. For the most recent ~10nm technologies, the influence of other atmospheric particles generated in nuclear cascade showers on circuits has been verified [15, 16].

2.2.1 Mechanism

The physical mechanisms relevant to the production of SEE in microelectronic devices schematically comprise three main consecutive steps: (a) the charge deposition by the energetic particle striking the sensitive region, (b) the transportation of the released charge into the device and (c) the charge convergence in the sensitive region of the device, which are illustrated in Figure 2.9 in the case of the passage of a high-energy ion through a reverse-biased n+/p junction.

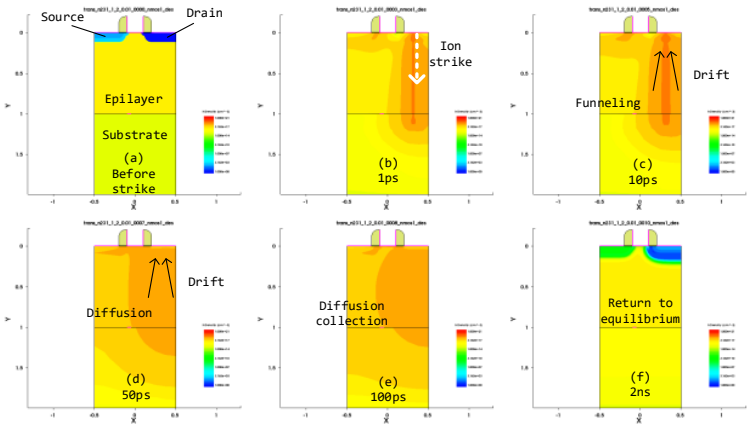


Figure 2.9 TCAD simulation results of ion strike with LET = 1 MeV/cm on NMOS.

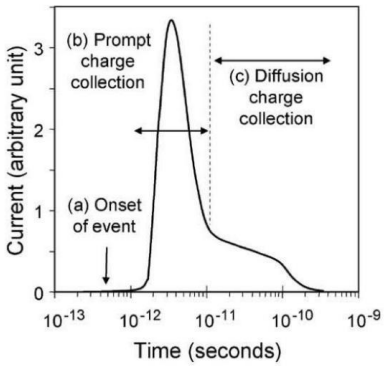


Figure 2.10: The resultant current pulse caused by the passage of a high-ionization particle.

- **Charge generation**

When a high-energy charged particle crosses the device, the charge is deposited along the particle path through the following mechanisms: direct ionization through interaction with the material or indirect ionization through secondary particles induced by reaction with atomic nuclei. Direct ionization by heavy ions of the HEP experimental environment is particularly important, where the heavy ions interact with the target material basically by inelastic interactions and transmit a large amount of energy to the electrons of the struck atoms. These electrons generate a cascade of secondary electrons which thermalize and create electron-hole pairs along the particle path [Fig. 2.9 (b)].

In a semiconductor or insulator, a large quantity of the deposited energy is thus converted into electron-hole pairs, the remaining energy transformed into heat and a very small quantity into atomic displacements. It has been shown experimentally that the energy required for the creation of an electron-hole pair is related to the material bandgap. In a microelectronics silicon substrate, one electron-hole pair comes into being for every 3.6 eV of energy lost by the ion, other particles, such as the neutrons of the terrestrial environment, do not interact directly with the atomic electronics of the target material, and do not ionize the matter on their passage. These particles should nevertheless not be neglected because of their ability to produce SEE on account of their possibility of nuclear reaction with the atoms of materials that constitute the microelectronic

devices. This mechanism is denominated as indirect ionization, where the charged products produced by nuclear reactions can deposit energy along their trajectory in the same way as direct ionization. As this process also leads to a line of electron-hole pairs created along the secondary emitted particles, the same general models and concepts can be utilized.

- **Charge transport**

When a column of charge carriers is generated in the semiconductor by an ionizing particle, they are rapidly transported and collected by elementary structures (e.g., p-n junctions). The transportation of charge leans upon two main mechanisms [Figs. 2.9 (c) to (e)]: charge drift in regions with an electric field and charge diffusion in neutral zones. The stored charges can also be recombined with other mobile carriers that exist in the lattice.

- **Charge collection**

The charges transmitted in the device bring about a parasitic current transient [Fig. 2.10], which can cause disturbances in the device and associated circuits. The devices that are the most sensitive to ionizing particle strikes are usually devices containing reversely biased p-n junctions, since the strong electric field existing in the depletion region of the p-n junction allows a very efficient collection of the stored charge. The effects of ionizing radiation are disparate according to the intensity of the current transient, as well as the number of circuit nodes impacted. If sufficiently important, the current can give rise to permanent damage on gate insulators (gate rupture, SEGR) or the latch-up (SEL) of the device. In normal low-power circuits, the transient current may commonly result in only an eventual change of the logical state (cell upset, SEU).

Many papers discuss the process of charge absorption in SEE. Based on the SEE mechanism, charges Q generated on average by high-ionization particles are computed by the following equation:

$$Q = \eta \cdot q_e \cdot \rho \cdot l \cdot LET/E_{eh} \tag{2.13}$$

In the formula, η is the charge absorption efficiency of an electronic node, q_e is the electron charge, i.e., $1.6 \times 10^{-19} C/e^-$, ρ is the density of target

material, which, for silicon, is $2330mg/cm^3$, l is the travel path of the particle, LET is the linear energy transfer of the particle and E_{eh} is the ionization energy to produce an electron-hole pair, 3.6 eV in silicon. The charge absorption efficiency η is influenced by many spatial parameters, such as the impact angle and position of the particle as well as the layout of the circuit [17]. In Eq. 2.13, it is supposed that the LET of the particles is constant, which is completely accurate for high energy ionizing particles.

2.2.2 SEE terminology and classification

As defined by the JEDEC standard JESD89A [18], JESD57 [19], and ESCC25100 [20], Single Event Effects represent any measurable or observable changes prompted by high-energy single-particle impacts in the condition or performance of microelectronic devices, components, subsystems or systems (digital or analog). These changes are primarily separated into two parts, which are temporary recoverable soft errors and permanent irreversible hard errors.

Soft errors, including Single-Event Upset (SEU), Single Event Transient (SET), and Single-Event Functional Interrupt (SEFI), have for consequence an incorrect output signal from the latch or storage unit, which can be modified by performing one or more standard functions of the device containing the latch or storage unit. For multi-unit, there are also Multiple-Cell Upset (MCU) and Multiple-Bit Upset (MBU).

In memory, the soft error rate can representatively be expressed in terms of Failures in Time (FIT) (the number of failures per 10^9 hours of operation). A practical equation is used during an accelerated or real-time test, which is, for instance, $Single\ Error\ Rate,$ SER $= \frac{N_r}{AF \times \Sigma_r} \times 10^9 (FIT/M\ Bit)$, where N_r is the number of bit flips observed at the time T_r, Σ_r is the number of $MBit \times h$ accumulated at time T_r and AF is the acceleration factor with respect to a given reference.

Hard errors are irreversible variations in operation, normally perpetual damage to one or more elements of the device or circuit (e.g., the gate oxide rupture, destructive latch-up events).

A Single-Event Latch-up (SEL) is a self-sustaining high current state in a device, which is prompted by a single high-energy particle passing through the sensitive area of the device structure and resulting in the loss of device functionality, and meanwhile, can bring about permanent damage to the device. If the device is not permanently damaged, it is necessary to restart the device (turn off then on) to get back to normal operation. An instance of SEL in CMOS devices is that when a single particle strike leads a parasitic thyristor structure (p-n p-n), thereby resulting in a short to ground.

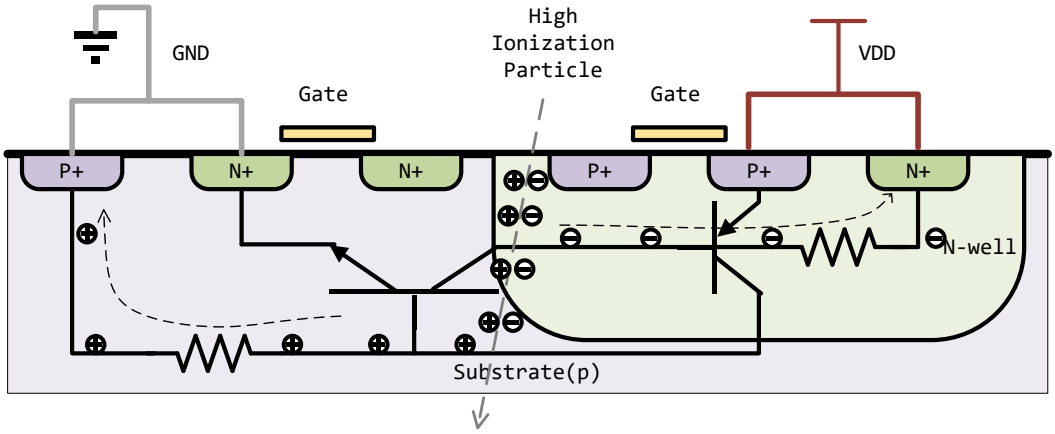


Figure 2.11 The mechanism of SEL

The Single Event Gate Rupture (SEGR) is the thorough or partial damage of the dielectric gate material on account of an avalanche breakdown, which can easily affect the MOSFETs with high current consumption (such as switching MOS). First of all, the gate oxide layer is weakened by the passage of ions through the gate oxide layer. Furthermore, the charge ionized along the ion track in silicon will enhance the oxide electric field instantaneously [21].

2.2.3 SEE Response in Digital Circuits

As the feature size of CMOS keeps decreasing, it is well established that Single Event Transient (SET) has turned out to be a significant error mechanism of great concern for digital circuit designers. With the advance of CMOS scaling, higher operating frequency, lower power supply voltage, and lower noise tolerance make the sensitivity of the circuit to SET higher.

Digital Single-Event Transients (DSET) are voltage or current transients brought by charges ionizing by energetic particles and collected by the circuit nodes [22]. Even if this transient does not result in SEU in the stroked circuit, it may transmit to subsequent circuits and be stored as corrupted data when it hits the latch or storage element. Different from SRAM cells, in combinatorial logic nodes, only when the charge propagates through the circuit and is locked in the static cell, will a SET with sufficient charge emerge as a "persistent" error [23]. DSET must meet many conditions to give rise to errors in storage elements [24]:

(1) Ion impact must generate a transient that can transmit through the circuit.

(2) There must be an open logical passage through which the DSET can transmit to reach the lock or storage element.

(3) The DSET must have adequate amplitude and duration to change the latch/memory state.

(4) In synchronous logic, the DSET must reach the latch during a clock pulse that enables the latch. As a result, the possibility of capturing SET goes up as the clock frequency increases.

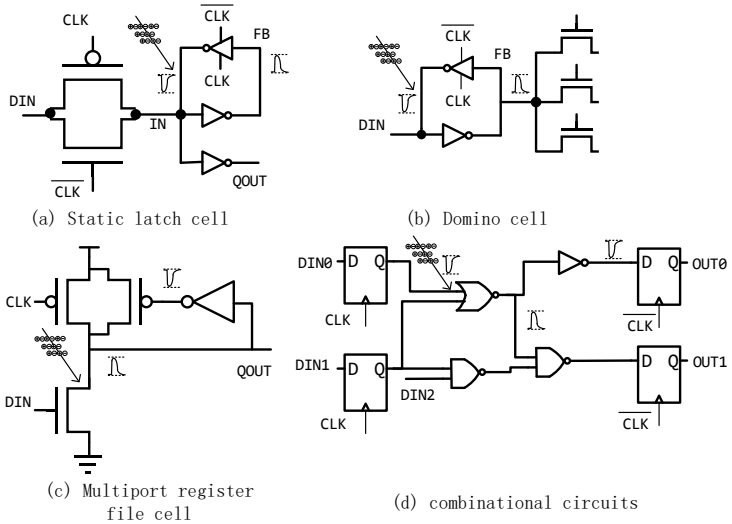


Figure 2.12: Illustration of typical sequential logic (latch, domino, register) and combinational circuits (random logic block).[23]

Digital circuits consist of sequential elements (such as latches, flip-flops, register cells) and combinational logic (such as NAND and NOR gates).

The following illustration briefly describes the effects of transients prompted by a single event in these two types of circuits.

- **Sequential logic**

Representative sequential elements in the core logic are a latch [Fig. 2.12(a)], comprising a domino cell [Fig. 2.12(b)] or a register file cell [Fig. 2.12(c)]. State changes can happen in core logic similarly to the way if it happened in memory elements. In sequential logic (like in SRAM), the soft error rate is independent of the clock frequency of the circuit [25]. For instance, the latch state can be flipped by the charge stored by a particle strike on a circuit node in spite of the state of the clock signal.

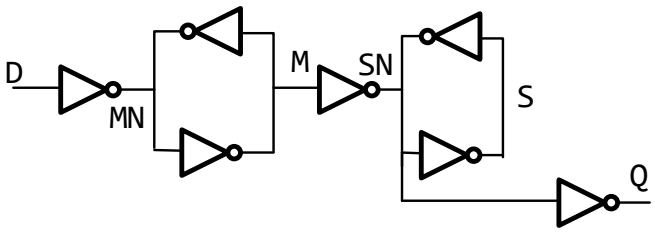


Figure 2.13: Simplified schematic of the flip-flop circuit. The sensitive nodes are labeled as M.N., M, S.N., and S [26].

Flip-flop circuits (Fig.2.13) are other representative sequential logic circuits, which, with technology scaling, have turned to be more susceptible to soft errors, mainly because of the reduction in supply voltage and of their node capacitances. The simplified schematics of Fig. 2.13 indicate that flip-flops circuits are analogous to SRAM cells, since both apply feedback loops of cross-coupled inverter-pairs. The critical charge (Q_{crit}) and the collection efficiency (Q_S) determine the soft error sensitivity of this class of circuits. In an SRAM cell, Q_{crit} is primarily the same for the two storage nodes as the cell is symmetrical. In flip-flops, the differently sized inverters have disparate fan-outs, which makes the flip-flop circuit asymmetric compared with the SRAM cell. As a consequence, the individual storage nodes in a flip-flop have different critical charges than in SRAMs, whose SER sensitivity can vary according by several orders of magnitude[27].

- **Combinational logic**

Any node in the combinational circuit can be affected by an SEU and lead to a voltage transient which can transmit through the combinational stages [Fig. 2.12 (d)] and bring about an error if latched by a successive element, such as a memory cell. In combinational logic, a certain number of transients will not be latched, and even when latched, some of these data will not be detected as errors for the software operation. A transient error in the logic circuit may not be intercepted in the memory circuit since it may be masked by one of the three phenomena as follows [28], [29]:

(i) Logical masking [30] emerges when a particle beats a portion of the combinational logic that cannot affect the output owing to a subsequent gate whose consequence is totally determined by its other input values. For instance, if the strike occurs on an input to a NAND (NOR) gate (as illustrated in Fig. 2.14(a)), but one of the other inputs is in the controlling state (e.g., 0(1) for a NAND (NOR) gate), the strike will be fully masked, and the output will be unvaried (i.e., the particle strike will not give rise to a soft error).

(ii) Electrical masking arises for transients with bandwidths higher than the cutoff frequency of the CMOS circuit and these transients will thereafter be attenuated [31]. The pulse amplitude may decrease, the rise and fall times increase, and, finally, the pulse may vanish [as shown in Fig. 2.14(b)]. On the other hand, since most logic gates are nonlinear circuits with a massive voltage gain, low-frequency pulses with enough initial amplitude will be amplified [23].

(iii) Temporal masking (or locking window masking) emerges when the pulse produced by the impact of particles reaches the latch, but not at the time when the latch captures the clock transition of its input value [29], which is explained in Fig. 2.14(c): when the transient transmits towards a sequential element (a latch in Fig. 2.14(c)), the interference in node DIN may be outside the latching window [32]. In consequence, the error will not be locked, and no software error will appear.

On account of the above-mentioned reasons, masking affects the soft error rate in combinational logic and renders the effects considerably lower than they could be. In addition to these masking mechanisms, two key-factors influence the soft error rate in combinational logic: the clock

frequency and the SET pulse width. With increasing clock frequency, there are more latching clock edges to acquire a pulse, and thereby the error rate rises. The pulse width is a key parameter that decides both the distance that the SET will pass through the combinational chain and the possibility that the SET be latched in a memory element as erroneous data. The wider the SET pulse width, the greater possibility it has of arriving at the latching edge of the clock. If the transient grows longer than the period of the clock, then every induced transient will be latched [33]. The SET pulse width and amplitude are subject to both process and circuit parameters (substrate and epitaxial layer doping, circuit capacitance, etc.) [34].

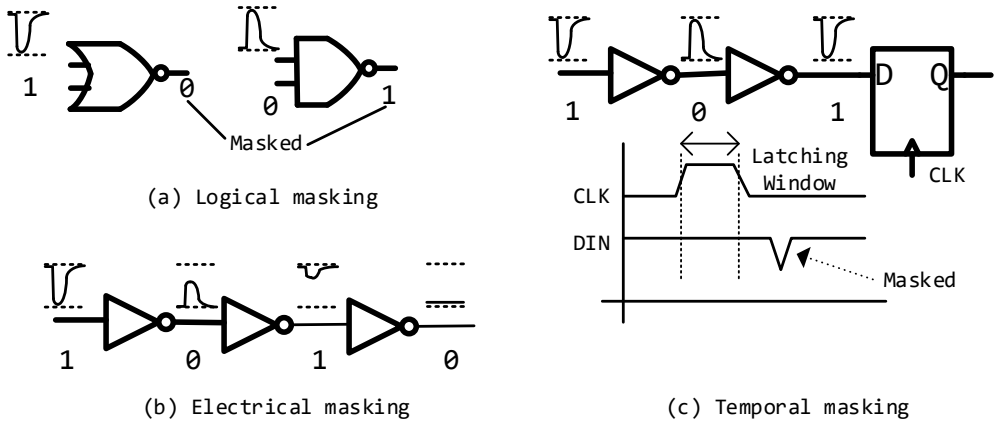


Figure 2.14: Illustration of the masking phenomena in combinational logic [23].

2.3 Total Ionizing Dose effects

Accumulated long-term ionizing damage caused by protons and electrons can make devices have threshold shifts, increasing device leakage (i.e., power consumption), timing changes, and decreasing functionality, etc. This long-term cumulative radiation effect is generally related to the total dose of ionizing radiation, and is therefore called Total Ionizing Dose (TID) effects.

2.3.1 Mechanism

The physical processes that lead from the initial storage of energy by ionizing radiation to the production of ionization deficiencies in the dielectric of a Metal-Oxide-Semiconductor (MOS) structure are: 1) the

generation of electric-hole pairs, 2) the rapid recombination of a fraction of the generated electric-hole pairs, 3) the transportation of free carriers remaining in the oxide, and either 4a) the formation of trapped charge by way of hole trapping in defect precursor sites or 4b) the formation of interface traps through reactions mostly involving hydrogen, which are all outlined graphically in Fig 2.15 [35] [36].

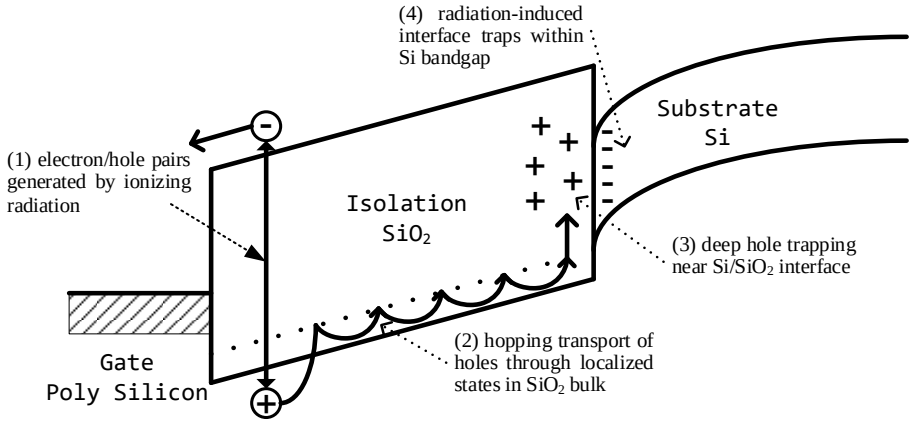


Fig. 2.15. Illustration of the main processes in TID damage [35].

In electric-hole pair generation (process 1), a fraction of the kinetic energy of the incident particle is lost to electric-hole pairs creation. The mean energy, E_p, required to produce an electron-hole pair in a material is related to the bandgap of the target material. The number of electric-hole pairs caused by a given dose is thereby a function of E_p as well as the material density. The relationships between ionization energy, material density, and generated carriers are given in Table 2.1 for three materials: GaAs, Si, and SiO_2 [37]. The electric-hole pair density per rad, denoted as κ_g (column 3 in Table 2.1), is acquired by using the following conversion equation:

$$\kappa_g[\frac{\mathrm{N_{ehp}}}{\mathrm{cm^3 rad}}] = 100[\frac{\mathrm{erg}}{\mathrm{g}}][\frac{1}{\mathrm{rad}}] \cdot \frac{1}{1.6\times10^{-12}}[\frac{\mathrm{eV}}{\mathrm{erg}}] \cdot \frac{1}{E_p}[\frac{\mathrm{N_{ehp}}}{\mathrm{eV}}] \cdot \rho[\frac{\mathrm{g}}{\mathrm{cm^3}}] \quad (2.13)$$

Once formed, a fraction of the electric-hole pairs is eliminated through either columnar or geminate recombination (process 2) [38]. The electric-hole pairs that keep clear of this initial recombination process divided by the total number of electric-hole pairs produced is the fractional charge yield, f_y. Fig 2.16 plots f_y as a function of the radiation type. The figure

also demonstrates that electron-hole pair recombination is a function of the electric field within the material [39].

Table 2.1 Relationship between ionization energy, material density, and generated carriers [37].

Material	Mean Ep (eV)	Density (g/cm^2)	Pair density per rad κ_g ($N_{ehp}/cm^3/rad$)
GaAs	~4.8	5.32	$\sim 7 \times 10^{13}$
Si	3.6	2.328	4×10^{13}
SiO_2	17	2.2	8.1×10^{13}

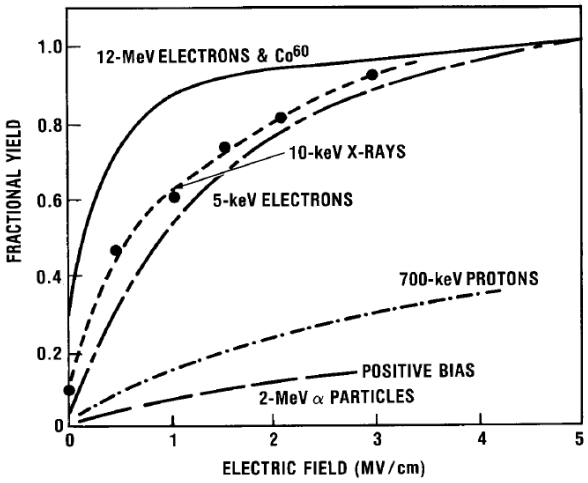


Fig. 2.16. The fractional yield of holes generated in SiO_2 as a function of the electric field in the material [38].

The fractional yield of electric-hole pairs goes up monotonically as the local electric field rises. It is usually believed that electrons, which have much higher mobility than holes in oxides, are quickly swept out of a SiO2 dielectric. The surviving holes will experience polaron hopping transport through shallow traps in the SiO2 (process 3). A fraction of these transported holes may sink into deep traps in the oxide bulk or near the Si/SiO2 interface, thus forming trapped positive charge (process 4a) [40]. The hole trapping efficiency (f_{ot}) is also a function of the electric field in the oxide because of the field dependence of the hole trapping cross-section [37]. The trapped hole defect may, depending on its proximity to the interface, exchange charge with the underlying Si via electron tunneling [41]. Reactions between holes and hydrogen-containing deficiencies or

dopant complexes can also give rise to the formation of the second type of ionization defect: the interface trap (process 4b) [42].

2.3.2 Impact of TID Deficiencies on Bulk MOS Structures

Oxide trapped charge is representatively net positive because of the capture of holes in neutral oxygen vacancies and the subsequent formation of oxygen vacancy deficiencies, or E' centres [43]. There are primarily two types of E' deficiencies: E_δ', and E_γ'. The E_δ' center is a "dimer" vacancy, which forms a relatively shallow trap for holes in the oxide, and mostly has energies located in the SiO2 bandgap within 1.0 eV of the oxide valence band [44]. The shallow trap level of the E_δ' makes it a good candidate for the defect type, which is responsible for hole transport through SiO2 (process 3) [44]. The E_γ' center is a remarkably deeper trap than the E_δ' defect, residing at energy levels greater than 3 eV above the oxide valence band[45]. While E_γ' centers may be situated throughout the oxide, most are found near the Si/SiO_2 interface [46].

Both types of E' centres can exchange charge with the adjacent Si layer and the ability of the E' defect to "communicate" with the Si is a strong function of its proximity to the interface [47]. E' centres that readily capture carriers from or emit carriers to the adjacent Si are often called border traps or switching states. They are generally situated within 3 nm of the Si/SiO_2 interface and can exchange charge through electron tunneling on time scales of microseconds to seconds [41]. E_γ' centers situated at distances greater than 3 nm from the interface (i.e., in the oxide bulk) may capture and emit carriers, but with low probability for the occurrence of this process. Thus, E_γ' deficiencies in the oxide bulk are usually treated as fixed (i.e., bias independent) positive oxide charge (N_{ot}). Removal or compensation of N_{ot} may require heightened temperature and/or biased annealing over relatively long periods of time. A schematic diagram of the location of border traps (switching states) and oxide trapped charge (fixed states) in the MOS system is shown in Fig 2.17. The formation of Not in an oxide can be expressed as [48]:

$$\Delta N_{ot} = D\kappa_g f_y f_{ot} t_{ox} \tag{2.14}$$

In the formula, D is the total ionizing dose deposited, and as Eq. (2.14) shows, N_{ot} is proportional to the thickness of the oxide. It should be noticed that for extremely thin gate oxides, N_{ot} buildup shows a sub-linear dependence on tox on account of the existence of tunneling processes that enhance charge compensation and annealing.

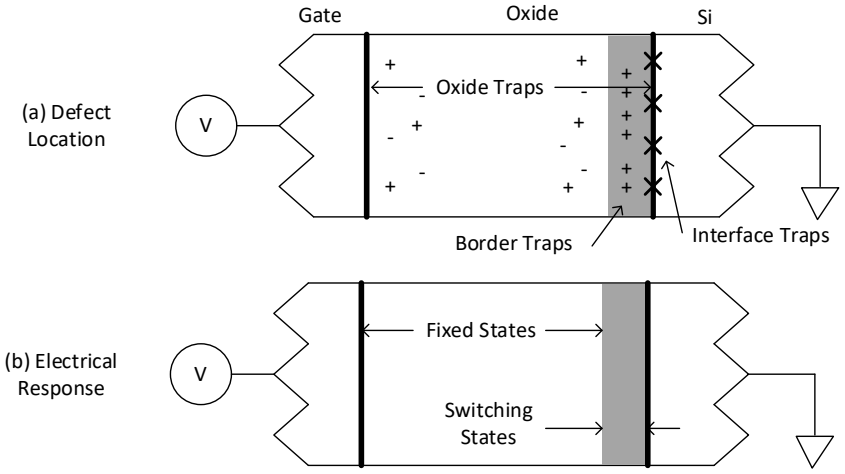


Fig. 2.17. Location and stability of trapped charge in SiO_2 [34].

Fixed oxide trapped charge can have a remarkable impact on the DC parameters of CMOS devices and integrated circuits. One of the most well-studied effects is the negative shift in the DC drain current vs. gate-to-source voltage (V_{gs}) for both n- channel and p-channel MOSFETS. That is illustrated in Fig. 2.18, showing that for a fixed I_D, the radiation-induced formation of N_{ot} shifts the V_{gs} bias point more negative (i.e., by V_{ot}) when the TID increases. In n-channel MOSFETs, this shift brings about a reduction in threshold voltage and an increment in off-state and drive currents. In p-channel MOSFETs, V_t increases negatively, while off-state and drive currents are reduced.

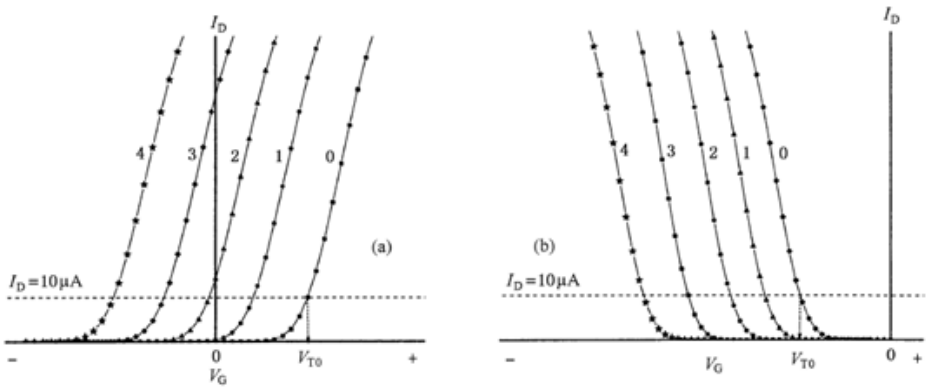


Fig. 2.18. Illustration of the effect of fixed oxide trapped charge on (a)n-MOS and (b) p-MOS devices. The 0~4 is the I-V of MOS when the dose increase. (0 pre-irradiation, 4- max dose.)

Radiation-induced DC voltage shifts can be computed by using the following equation:

$$\Delta V_{ot} = -\frac{t_{ox}}{k_{ox}\varepsilon_0} q\Delta N_{ot} \tag{2.17}$$

In the formula, k_{ox} is the dielectric constant of SiO_2 and ε_0 is the dielectric constant of free space[49]. Considering the equations (2.14) and (2.17), the theory forecasts that negative threshold voltage shifts prompted by fixed oxide trapped charge formation are proportional to the square of oxide thickness, i.e.,

$$-\Delta V_t(N_{ot}) = -\Delta V_{ot} \propto t_{ox}^2 \tag{2.18}$$

This theoretical relationship has been confirmed via countless experiments[48], which in Eq. (2.18) points out that as the gate oxides of advanced CMOS technologies are scaled to thinner dimensions, the threat of shifts in DC parameters prompted by N_{ot}. The formation in the gate oxide is declined [50]. Instead, hole trapping in the thicker shallow trench isolation dielectrics is now a greater radiation threat in modern CMOS technologies, as typical STI trenches are much thicker than gate oxides. For advanced CMOS technologies, the STI thicknesses range from 300 nm to 450 nm [51]. The influence of fixed positive oxide trapped charge formation in shallow trench isolation (STI) structures will be discussed minutely below.

2.3.3 Impact of TID Deficiencies on STI MOS Structures

In the 1980s, Saks and Ancona forecasted that technology scaling would decline the MOSFET's susceptibility to radiation-induced damage in gate oxides [52] primarily due to the fact that, to first order, defect formation in gate oxides scales with oxide thickness tox. Radiation tolerance in thin gate oxides is further fortified by the elevated possibility of positive charge eradication or compensation by tunneling electrons from the adjacent materials [14], and inherent gate oxide hardness is proved at the 0.25 um technology node [53]. At this node, the oxide thickness is representatively less than 6 nm, which is less than twice the approximate distance for high possibility electron tunneling (i.e., ~ 3 nm) [14]. Gate oxide hardness trends have continued to be observed in subsequent smaller technologies, e.g., 0.18 um (tox = 3.2 nm) and 0.13 um (tox ~ 2 nm). Defect formation in thicker isolation oxides is representatively the major factor of radiation-induced degradation in the DC parameters of modern CMOS devices and integrated circuits. A representative cross-section of a modern bulk n-channel MOSFET is shown in Fig. 2.19, which illustrates that the shallow trench isolation (STI) structures enclosing the active device are much thicker (by more than two orders of magnitude) than the gate oxide.

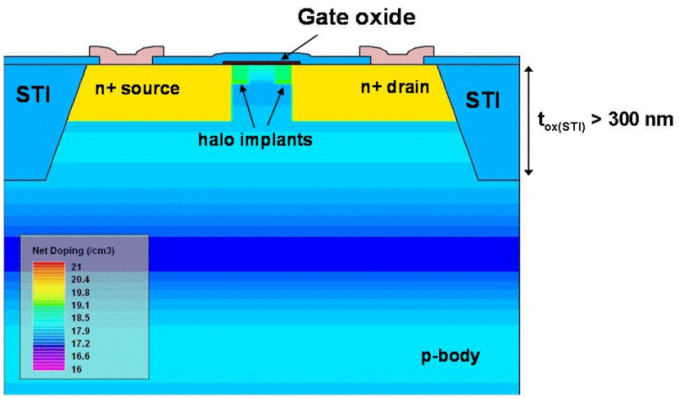


Fig. 2.19. Representative cross-section of a bulk n-channel MOSFET in modern CMOS technology.

Fig. 2.20 illustrates the cross-sectional diagrams of n-channel MOSFETs with (a) LOCOS isolation and (b) STI, where the radiation-induced charge formation is plotted with the "+" symbol. In the case of

LOCOS isolation, charge formation occurs at the bird's beak region and along the base of the field oxide (extending from drain to source where the gate overlaps the thick field oxide). In STI oxides, the radiation-induced charge will be enhanced near the trench corner and along the base of the field oxide. In most deep-submicron technologies, LOCOS isolation has been substituted with STI.

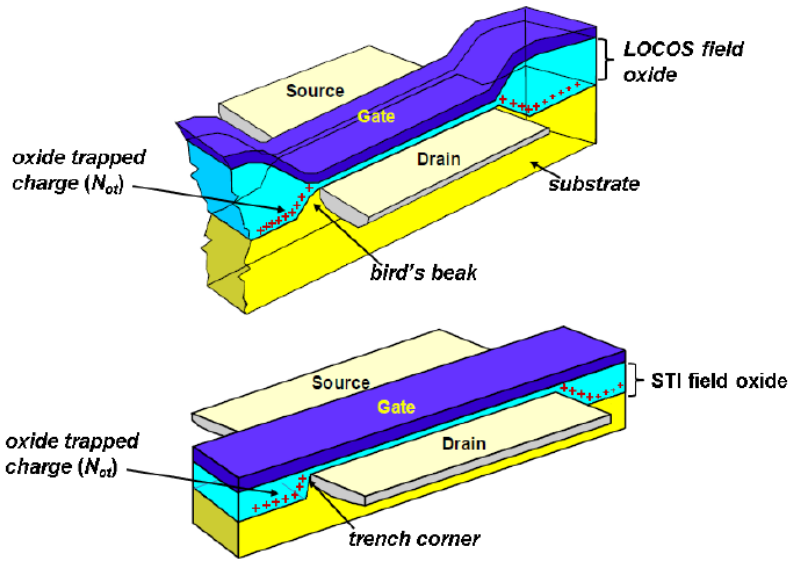


Fig. 2.20 Cross-sectional diagrams of n-channel MOSFETs with (a) LOCOS isolation and (b) STI [54].

2.4 Displacement damage

Highly energized particles (protons, electrons, neutrons, and heavy ions) can destroy semiconductor materials by displacing atoms as the particle moves through the material. In silicon, Frenkel deficiencies (interstitial silicon and vacancy pairs) are shaped when incident particles collide with silicon nuclei or when primary recoil atoms impact other atoms in the lattice [55]. The vacancies and interstitials generated by these collisions introduce allowable energy states in the bandgap of semiconductor material. The existence of these states can vary the electrical performance of electronic devices, resulting in various radiation responses. For minority carrier devices such as diodes and bipolar transistors, the primary influence of the radiation-induced energy levels is the production of bulk traps (N_t) in the material, which decrease carrier lifetime and thus augment carrier

recombination in a forward-biased junction (e.g., the base-emitter junction in active operation), resulting in the increment of the base current and the possible reduction of collector current. Both alterations to current result in a reduction in current gain. A linear relationship between incident particle fluence and bipolar current gain was postulated by Messenger and Spratt [56]:

$$\frac{1}{\beta(\Phi)} = \frac{1}{\beta_{pre}} + K\Phi \tag{2.17}$$

In the formula, $\beta_{pre}$is the current gain prior to radiation exposure, $\beta(\Phi)$is the current gain after irradiation, Φ is the particle fluence (i.e., the number of considered particles), and K is a composite displacement damage factor with the units of cm^2. Experimental studies, whether in the past or at present, utilize the damage factor as an analytical tool for describing displacement damage effect. For instance, there is a linear relationship between reciprocal gain and particle fluency in the experimental data as shown in Fig. 2.21, which, originally presented by Summers et al. in 1987, demonstrate the dependence of the damage factor, K (the slope of the curves), on the type and energy of the incident particle [57]. As the figure shows, the damage factor is larger for 16.8 MeV He ions than either 40.0 MeV He ions or 4.3 MeV deuterons.

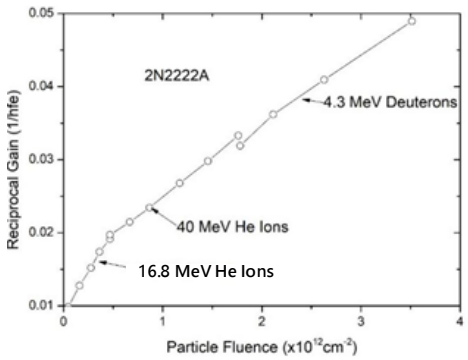


Fig. 2.21. Reciprocal gain vs. particle fluence for sequential irradiations on the 2N2222A NPN BJT with 16.8 and 40.0 helium ions and 4.3 MeV deuterons [57].

The displacement damage brought by different particles can be correlated analytically on the basis of Non-Ionizing Energy Loss (NIEL). NIEL, generally represented as the variable S, is computed as the energy

loss per unit mass to lattice displacement as a particle moves through the target material [55], which has the same units as LET (i.e., $MeV\ cm^2\ mg^{-1}$) and, as with LET, is a nonlinear function of the type and energy of the incident particle type. In Fig. 2.22, the non-ionizing energy loss is plotted vs. energy for three kinds of incident particles: electrons, protons, and neutrons [58]. The curves have been computed using the analytical model developed by Burke [59]. The calculation of NIEL demands approximations for the fluxionary cross-section for atomic displacements (dσ/dΩ) and the average energies of recoil atoms (T), corrected for Lindhard energy partitioning [60]. NIEL is computed by assessing the integral:

$$S = \frac{N}{A}\int_{\theta_{min}}^{\pi} \frac{d\sigma(\theta)}{d\Omega} T(\theta) d\Omega \tag{2.18}$$

In the formula, N is Avogadro's number, A is the atomic mass, θ is the scattering angle, $d\Omega$ is the fluxionary solid angle centered about θ, and θ_{min} is the scattering angle for which the recoil energy equals the threshold for atomic displacement [61].

The model calculations of non-ionizing energy loss in silicon offer values of $S_p = 1.94 \times 10^{-3}\ MeV\ cm^2\ g^{-1}$ for 200 MeV protons and $S_n = 2.04 \times 10^{-3}\ MeV\ cm^2\ g^{-1}$ for 1 MeV neutrons [60]. Once given these values, the NIEL ratio of $200\ MeV$ protons to $1\ MeV$ neutrons (S_p/S_n) is 0.95. The secondary electrons produced by $10\ keV$ x-rays have energies below $0.01\ MeV$. As Fig. 2.22 manifests, the x-ray induced NIEL in silicon is many orders of magnitude smaller than that of either protons or neutrons, hence the displacement damage produced by 10 keV x-rays is regarded negligible in solid-state materials, even in materials used in advanced non-volatile memory.

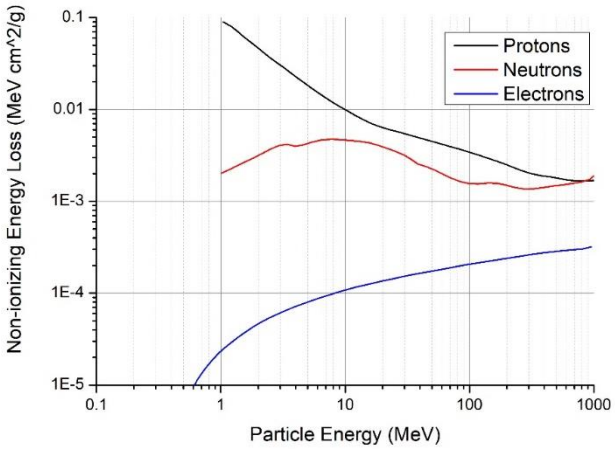


Fig. 2.22. Calculation of non-ionizing energy deposition in silicon by electrons, protons, and neutrons [58].

2.5 Summary

In this chapter, the introduction starts with the basic physical processes and explains the main physical processes that occur when high ionization particles enter semiconductor materials. The SEE is mainly concerned with memory or sequential logic, which has a very limited effect on combinational logic. Nevertheless, the SET on the clock tree can still cause a critical error, which, together with SEGR, can make permanent damage to the circuit. The TID is an accumulation process which can cause the leakage current increment and threshold shift especially on NMOS. The displacement damage effect mainly affects the high current or sensing devices, and is not considered as the main radiation risk in the condition of MVD in CBM.

2.6 Bibliography

[1] P. E. Dodd and L. W. Massengill, *Basic mechanisms and modeling of single-event upset in digital microelectronics*, *IEEE Trans. Nucl. Sci.* **50 III**, (2003) 583.

[2] D. K. Nichols, J. R. Coss, L. S. (Ted) Smith, B. Rax, M. Huebner and K. Watson, *Full temperature single event upset characterization of two microprocessor technologies*, *IEEE Trans. Nucl. Sci.* **35**, (1988) 1619.

[3] M. Huhtinen and F. Faccio, *Computational method to estimate Single Event Upset rates in an accelerator environment*, *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **450**, (2000) 155.

[4] ITRS, *International Technology Roadmap for Semiconductors 2012 Update*, *Overview*, *Rep. - Itrs*, (2012) 70.

[5] D. Munteanu and J.-L. Autr, *Investigation of Sensitivity to Heavy-Ion Irradiation of Junctionless Double-Gate MOSFETs by 3-D Numerical Simulation*, in *Comput. Numer. Simulations*, InTech, (2014).

[6] Giovanni Santin, *Section III: Space Radiation Transport Models*, *2006 IEEE Nucl. Sp. Radiat. Eff. Conf. Short Course* **3**, (2006).

[7] A. C. T. Chan, *Distorted Wave Born Approximation For Inelastic Atomic Collision by*, in *Proc. Sci.* **370**, Sissa Medialab Srl, (2007).

[8] J. D. Ullman, H. Frauenfelder, H. J. Lipkin and A. Rossi, *Determination of electron and positron helicity with Møller and Bhabha scattering*, *Phys. Rev.* **122**, (1961) 536.

[9] R. C. Fernow, *Introduction to Experimental Particle Physics*, in *Introd. to Exp. Part. Phys.*, Cambridge University Press, (1986).

[10] W. R. Leo, *Techniques for Nuclear and Particle Physics Experiments*, *Tech. Nucl. Part. Phys. Exp.*, (1994).

[11] H. Bichsel, *Straggling in thin silicon detectors*, *Rev. Mod. Phys.* **60**, American Physical Society, (1988) 663.

[12] M. Tanabashi, K. Hagiwara, K. Hikasa, K. Nakamura, Y. Sumino, F. Takahashi, J. Tanaka, K. Agashe, G. Aielli, et al., *Review of Particle Physics*, *Phys. Rev. D* **98**, American Physical Society, (2018) 30001.

[13] U. Amaldi and G. Kraft, *Particle accelerators take up the fight against cancer*, in *Cern Cour.* **18**, (2006).

[14] R. D. Schrimpf and D. M. Fleetwood, *Radiation Effects and Soft Errors in Integrated Circuits and Electronic Devices*, in *World Sci.* **34**, WORLD SCIENTIFIC, (2004).

[15] D. F. Heidel, P. W. Marshall, K. A. Label, J. R. Schwank, K. P. Rodbell, M. C. Hakey, M. D. Berg, P. E. Dodd, M. R. Friendlich, et al., *Low Energy Proton Single-Event-Upset Test Results on 65 Nm SOI SRAM*, in *IEEE Trans. Nucl. Sci.* **55**, (2008).

[16] E. H. Cannon, M. Cabanas-Holmen, J. Wert, T. Amort, R. Brees, J. Koehn, B. Meaker and E. Normand, *Heavy Ion, High-Energy, and Low-Energy Proton SEE Sensitivity of 90-Nm RHBD SRAMs*, in *IEEE Trans. Nucl. Sci.* **57**, (2010).

[17] D. Munteanu and J. L. Autran, *Modeling and simulation of single-event effects in digital devices and ICs*, *IEEE Trans. Nucl. Sci.* **55**, (2008) 1854.

[18] JEDEC, *Measurement and Reporting of Alpha Particle and Terrestrial Cosmic Ray Induced Soft Error in Semiconductor Devices*, *JEDEC Stand. JESD89A*, (2006) 1.

[19] E. I. Association and others, *Test procedures for the measurement of single-event effects in semiconductor devices from heavy ion irradiation*, *EIA/JEDEC Stand.*, (1996).

[20] ECSS, *Single Event Effects Test Method and Guidelines*, in *Ecss*, (2002).

[21] J.-M. Lauenstein, *Single-Event Gate Rupture in Power MOSFETs: A New Radiation Hardness Assurance Approach*, (2011).

[22] V. Ferlet-Cavrois, L. W. Massengill and P. Gouker, *Single event transients in digital CMOS - A review*, *IEEE Trans. Nucl. Sci.* **60**, (2013) 1767.

[23] J. M. Benedetto, P. H. Eaton, D. G. Mavis, M. Gadlage and T. Turflinger, *Digital single event transient trends with technology node scaling*, *IEEE Trans. Nucl. Sci.* **53**, (2006) 3462.

[24] P. E. Dodd, *Physics-based simulation of single-event effects*, in *IEEE Trans. Device Mater. Reliab.* **5**, (2005).

[25] M. J. Gadlage, P. H. Eaton, J. M. Benedetto and T. L. Turflinger, *Comparison of heavy ion and proton induced combinatorial and sequential logic error rates in a deep submicron process*, *IEEE Trans. Nucl. Sci.* **52**, (2005) 2120.

[26] P. Roche and G. Gasiot, *Impacts of front-end and middle-end process modifications on terrestrial soft error rate*, in *IEEE Trans. Device Mater. Reliab.* **5**, (2005).

[27] T. Heijmen, B. Kruseman, R. Van Veen and M. Meijer, *Technology Scaling of Critical Charges in Storage Circuits Based on Cross-Coupled Inverter-Pairs*, in *IEEE Int. Reliab. Phys. Symp. Proc.* **2004-Janua**, Institute of Electrical and Electronics Engineers Inc., (2004).

[28] T. Karnik, P. Hazucha and J. Patel, *Characterization of soft errors caused by single event upsets in CMOS processes*, *IEEE Trans. Dependable Secur. Comput.* **1**, Institute of Electrical and Electronics Engineers Inc., (2004) 128.

[29] P. Shivakumar, M. Kistler, S. W. Keckler, D. Burger and L. Alvisi, *Modeling the effect of technology trends on the soft error rate of combinational logic*, *Proc. 2002 Int. Conf. Dependable Syst. Networks*, (2002) 389.

[30] P. Shivakumar, M. Kistler, S. W. Keckler, D. Burger and L. Alvisi, *Modeling the Effect of Technology Trends on the Soft Error Rate of Combinational Logic*, in *Proc. 2002 Int. Conf. Dependable Syst. Networks*, (2002).

[31] P. Dahlgren and P. Liden, *Switch-Level Algorithm for Simulation of Transients in Combinational Logic*, in *Proc. - Annu. Int. Conf. Fault-Tolerant Comput.*, IEEE, (1995).

[32] P. Liden, P. Dahlgren, R. Johansson and J. Karlsson, *On Latching Probability of Particle Induced Transients in Combinational Networks*, in *Dig. Pap. - Int. Symp. Fault-Tolerant Comput.*, Publ by IEEE, (1994).

[33] M. J. Gadlage, R. D. Schrimpf, J. M. Benedetto, P. H. Eaton, D. G. Mavis, M. Sibley, K. Avery and T. L. Turflinger, *Single Event Transient Pulsewidths in Digital Microcircuits*, in *IEEE Trans. Nucl. Sci.* **51**, (2004).

[34] J. M. Benedetto, P. H. Eaton, D. G. Mavis, M. Gadlage and T. Turflinger, *Digital Single Event Transient Trends with Technology Node Scaling*, in *IEEE Trans. Nucl. Sci.* **53**, (2006).

[35] T. R. Oldham, *Basic Mechanisms of TID and DDD Response in MOS and Bipolar Microelectronics*, in *Proc. NSREC Tutor. Short Course*, (2011).

[36] H. J. Barnaby, *Total-ionizing-dose effects in modern CMOS technologies*, *IEEE Trans. Nucl. Sci.* **53**, (2006) 3103.

[37] J. R. Schwank, *Basic mechanisms of radiation effects in the natural space radiation environment*, *U. S. Dep. Energy* **1**, (1994) 3.

[38] K. L. Hughes, D. M. Fleetwood and J. R. Schwank, *Charge yield for Cobalt-60 and 10-keV x-ray irradiations of MOS devices*, *IEEE Trans. Nucl. Sci.* **38**, (1991) 1187.

[39] M. R. Shaneyfelt, J. R. Schwank, D. M. Fleetwood, P. S. Winokur, K. L. Hughes, F. W. Sexton and K. L. Hughes, *Field Dependence of Interface-Trap Buildup in Polysilicon and Metal Gate MOS Devices, IEEE Trans. Nucl. Sci.* **37**, (1990) 1632.

[40] D. M. Fleetwood, *Total ionizing dose effects in MOS and low-dose-rate-sensitive linear-bipolar devices, IEEE Trans. Nucl. Sci.* **60**, (2013) 1706.

[41] D. M. Fleetwood, *Fast and slow border traps in mos devices, IEEE Trans. Nucl. Sci.* **43**, (1996) 779.

[42] L. Tsetseris, R. D. Schrimpf, D. M. Fleetwood, R. L. Pease and S. T. Pantelides, *Common origin for enhanced low-dose-rate sensitivity and bias temperature instability under negative bias, IEEE Trans. Nucl. Sci.* **52**, (2005) 2265.

[43] J. R. Chavez, S. P. Karna, K. Vanheusden, C. P. Brothers, R. D. Pugh and B. K. Singaraju, *Microscopic structure of the e'δ center in amorphous sio2: A first principles quantum mechanical investigation, IEEE Trans. Nucl. Sci.*, (1997).

[44] C. J. Nicklaw, Z. Y. Lu, D. M. Fleetwood, R. D. Schrimpf and S. T. Pantelides, *The Structure, Properties, and Dynamics of Oxygen Vacancies in Amorphous SiO2*, in *IEEE Trans. Nucl. Sci.*, (2002).

[45] D. B. M. Klaassen, *A unified mobility model for device simulation-I. Model equations and concentration dependence, Solid State Electron.*, (1992).

[46] R. Waser, R. Dittmann, C. Staikov and K. Szot, *Redox-based resistive switching memories nanoionic mechanisms, prospects, and challenges*, in *Adv. Mater.*, (2009).

[47] D. M. Fleetwood, *"Border Traps" in MOS Devices*, *IEEE Trans. Nucl. Sci.*, (1992).

[48] R. C. Lacoe, J. V. Osborn, D. C. Mayer, S. Brown and J. Gambles, *Total-Dose Tolerance of the Commercial Taiwan Semiconductor Manufacturing Company (TSMC) 0.35-Mm CMOS Process*, in *IEEE Radiat. Eff. Data Work.*, (2001).

[49] R. Muller and T. Kamins, *Device Electronics For Integrated Circuits*, (1977).

[50] F. T. Brady, *A scaleable, radiation hardened shallow trench isolation1*, *IEEE Trans. Nucl. Sci.*, (1999).

[51] J. Wong, J. Weil, C. Whiting, C. Yu, B. Porth, J. Hart, G. Matteson, H. Mochiki and M. Hagihara, *A Hardmask STI Etch Process for 0.13um Logic Technology and Beyond*, in *IEEE Int. Symp. Semicond. Manuf. Conf. Proc.*, (2001).

[52] N. S. Saks and M. G. Ancona, *Generation of interface states by ionizing radiation at 80k measured by charge pumping and subthreshold slope techniques*, *IEEE Trans. Nucl. Sci.*, (1987).

[53] J. V. Osborn, *Total dose hardness of three commercial cmos microelectronics foundries*, *IEEE Trans. Nucl. Sci.*, (1998).

[54] M. R. Shaneyfelt, P. E. Dodd, B. L. Draper and R. S. Flores Sandia, *Challenges in hardening technologies using shallow-trench isolation*, *IEEE Trans. Nucl. Sci.*, (1998).

[55] T. P. Ma and P. V. Dressendorfer, *Basic mechanisms for the new millennium*, (1998) Medium: ED; Size: 119 p.

[56] G. C. Messenger and J. P. Spratt, *The Effects of Neutron Irradiation on Germanium and Silicon*, in *Proc. IRE* **46**, (1958).

[57] G. P. Summers, C. J. Dale, E. A. Burke, E. A. Wolicki, P. W. Marshall and M. A. Gehlhausen, *Correlation of particle-induced displacement damage in silicon*, *IEEE Trans. Nucl. Sci.*, (1987).

[58] C. J. Dale, P. W. Marshall, E. A. Burke, G. P. Summers and E. A. Wolicki, *High energy electron induced displacement damage in silicon*, *IEEE Trans. Nucl. Sci.*, (1988).

[59] E. A. Burke, *Energy dependence of proton-induced displacement damage in silicon*, *IEEE Trans. Nucl. Sci.*, (1986).

[60] G. P. Summers, E. A. Burke, P. Shapiro, S. R. Messenger and R. J. Walters, *Damage correlations in semiconductors exposed to gamma, electron and proton radiations*, *IEEE Trans. Nucl. Sci.* **40**, (1993) 1372.

[61] S. R. Messenger, *Nonionizing energy loss (NIEL) for heavy ions*, *IEEE Trans. Nucl. Sci.*, (1999).

Chapter 3. Radiation Hardening Techniques for MVD against SEE and TID

As described previously in Chapter 1, the MVD deployment conditions and particle environment in CBM, the usage of MAPS for the MVD stations, which are fabricated in a CMOS process, known to be affected by SEE and TID. In this chapter, several radiation hardening techniques for CMOS processes and appropriate for MVD sensors will be introduced. Since such methods mostly intend to minimize system failures, possibly induced by SEE, elements of system reliability theory will be introduced first.

3.1 Introduction to System Reliability

In general, the possibility that a system performs a specified function under specified conditions within a specified time defines the product reliability. The system can be a small part, a larger component, or a complete equipment.

For purpose of simplifying the system working station, the state of a component is defined as a binary function:

$$X_i(t) = \begin{cases} 1, & when\ component\ is\ up \\ 0, & when\ component\ is\ down \end{cases} \tag{3.1}$$

Then let $X(t) = (X_1(t), \dots, X_n(t))$ be the component state vector at time t. The system lifetime T is defined as

$$T = \inf\,[t: \phi(X(t)) = 0] \tag{3.2}$$

In the formula, ϕ is a function that can tell the system is up or down base on the component state vector X. The survival possibility function of a system defines the reliability and is expressed as

$$R(t) = P(T > t) \tag{3.3}$$

In the formula, t is time, T is the lifetime of the system. The cumulative distribution function of fault is the opposite $F(t)$ of the system survival possibility.

$$F(t) = P(T \leq t) = 1 - R(t) \tag{3.4}$$

The fault possibility density function $f(t)$ is the derivative of $F(t)$.

$$f(t) = F'(t) = -R(t) \tag{3.5}$$

The fault rate $\lambda(t, \Delta t)$ is defined as

$$\lambda(t, \Delta t) = \frac{P(t < T \leqslant t + \Delta t | T > t)}{\Delta t} \tag{3.6}$$

When $\Delta t \to 0$,

$$\begin{aligned} \lambda(t) &= \lim_{\Delta t \to 0} \frac{P(t<T\leqslant t+\Delta t)}{P(T>t)\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{F(t+\Delta t)-F(t)}{\Delta t R(t)} \\ &= \frac{F\prime(t)}{R(t)} = \frac{f(t)}{R(t)} = -\frac{R'(t)}{R(t)} \end{aligned} \tag{3.7}$$

The differential equation (3.7) can be solved with

$$R(t) = \mathrm{e}^{-\int_0^t \lambda(\tau) d\tau} \tag{3.8}$$

If $\lambda(t)$ is a constant value λ, the solution becomes $R(t) = \mathrm{e}^{-\lambda t}$.

Now, we have to consider the reliability of a combination of n components with corresponding reliabilities $R_1(t), \ldots, R_n(t)$. Generally speaking, in a complex system, the system reliability is based on the system structure function ψ, given by

$$R(t) = \psi(R_1(t), \ldots, R_n(t)) \tag{3.9}$$

For the reliability of a system where components are serially chained as depicted in figure 3.1 (a), the function ψ is simply the product of reliabilities:

$$R(t) = \prod_{i=1}^{n} R_i(t) \tag{3.10}$$

When the components of the system are connected in parallel as in figure 3.1 (b), the function becomes:

$$R(t) = 1 - \prod_{i=1}^{n} (1 - R_i(t)) \tag{3.11}$$

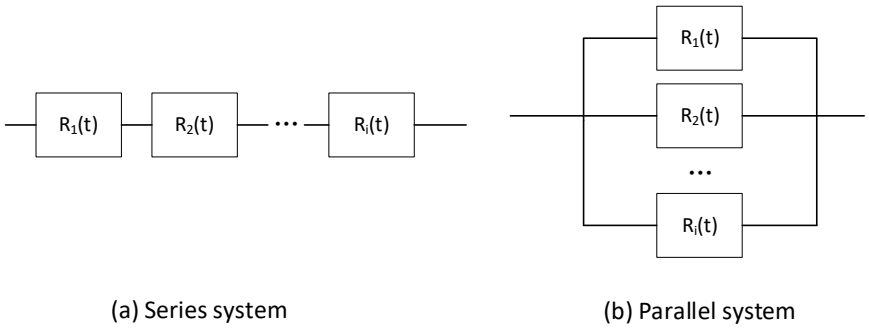


Fig. 3.1. The system reliability model for components connected (a) in series and (b) in parallel.

3.2 Triple modular redundancy (TMR)

In the 1950s', the components in computers featured very low reliability. In order to ensure that the information can be stored correctly, the intuitive idea to save a few more copies was followed and redundancy became a classic and effective error mitigation mechanism. In these cases, the triple modular redundancy (TMR) can provide more reliability at an acceptable cost [1]. It is now widely used in space satellite systems [2].

3.2.1 Topologies of TMR in digital circuits

The register-transmit level (RTL) model as the essential element of the digital circuit shown in Figure 3.2 (a) has two levels of logic-register block. As per the idea of TMR, the first TMR block is composed of 3 replicated components and a voter and the latter has three inputs, A, B, and C, and the logic for its output is given by $Out = AB \cup BC \cup AC$. From this equation, it is clear that the TMR system, as shown in Figure 3.2 (b), will

provide the wrong result only if there are 2 or 3 components in error, or if the voter fails. Duplication errors may stem from SEU in a register or SET in the logic. For purpose of preventing the voter from becoming a fatal part of the circuit, the voter can also be tripled, as shown in Figure 3.2 (c). However, if the output is in conjunction with redundant and non-redundant parts, the voter should be a single line to acquire the merged signal. In spite of voter redundancy, the clock signal may still introduce wrong pulses caused by SET and the duplication of the clock needs not only to separate the input of each register but also to introduce a Δt latency on each clock, as shown in Figure 3.2(f).

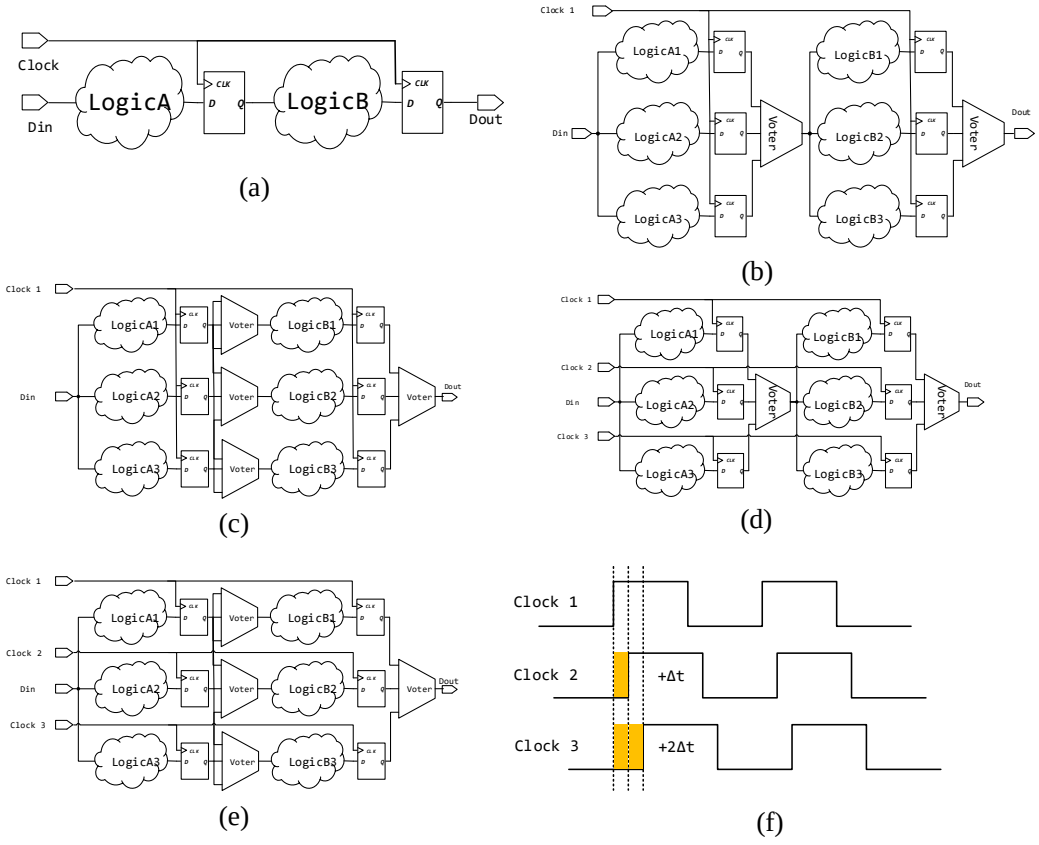


Figure 3.2 The different topology diagrams of TMR: (a) underlying logics without redundancy, (b) duplicated logic, (c) duplicated voter, (d) duplicated clock, (e) duplicate both voter and clock and (f) the schedule of the triplicated clock.

Since TMR brings higher reliability with additional micro-circuits, it has a cost in terms of power and area for the overall circuit. The redundancy ratio ρ_r is defined as

$$\rho_r = \frac{N_{additional\ part}}{N_{original\ part}} \tag{3.12}$$

In TMR, $\rho_r > 200\%$ (2 more copies and voters). If the clock is also tripled, the speed of the circuit will be slower by the effect of latency Δt. With the constantly increasing scale and speed of digital integrated circuits in MAPS, the cost of the implementation of TMR for the whole digital part of the chip has become unacceptable.

3.2.2 The trade-off on reliability and cost

Before the trade-off in the TMR implementation is discussed, a simple SEE error distribution model in digital circuits is come up with here. There are two assumptions to simplify the SEE fault model. The first is that the position of the incident particles and the energy distribution are uniform. That is to say, for a unit area and a particle with a specified LET, there is an incidence possibility $P(LET)$. The second is that the response to the same impact particles of disparate MOSFETs is the same, which is rough because larger MOSFETs will collect more charge during the SEE process. Nevertheless, if it is considered that in the synthesis of digital circuits, automation design tools will balance the drive conductance (size) $1/R$ and load capacity (fan-out) C of the logic unit, the larger MOSFETs have more capacitors to buffer the charge, and a larger drive current to release the charge. With these two assumptions, the logic elements have the same fault rate, i.e., SEE cross-section, which is defined as faults bits in-unit area and unit time ($as\ \lambda(t)$), $bit/mm^2/s$.

If the components A have a reliability R_A and fault rate λ_A, the expected reliability of the TMR system with a single voter is

$$R_{TMR-sv} = (\binom{3}{2}R_A^2(1-R_A) + R_A^3)\cdot R_V = (3R_A^2 - 2R_A^3)\cdot R_V$$

$$= R_A^2(3-2R_A)\cdot R_V = (3-2e^{-\lambda_A t})e^{-2\lambda_A t}\cdot e^{-\lambda_v t}. \tag{3.13}$$

With 3 voters, it is

$$R_{TMR-sv} = (\binom{3}{2} R_A^2(1-R_A) + R_A^3) \cdot (\binom{3}{2} R_V^2(1-R_V) + R_V^3)$$
$$= (3R_A^2 - 2R_A^3) \cdot (3R_V^2 - 2R_V^3) \quad (3.14)$$

If there is a functional logic circuit that includes M logic cells with unique reliability R_l, which form N processing block (block i has n_i logic cells, $\Sigma n_i = M$), the reliability of this logic block is

$$R_{logic} = R_l^{\sum n_i} = R_l^M = e^{-\lambda_l t M} \quad . \quad (3.15)$$

If there are k blocks out of N processing block (the block i in set K) have deployed the single-voter TMR, the reliability is

$$R_{logic-sv} = R_V^k R_l^{M-\Sigma n_i, i\in K} \prod_{i\in K}^K (3R_l^{2n_i} - 2R_l^{3n_i}) \quad (3.16)$$

Now the quality (improvement) factor of TMR is defined as

$$\eta = \ln\left(\frac{R_{logic-sv}}{R_{logic}}\right) = \ln\left(R_V^{\mathrm{k}} \frac{\prod_{i\in K}^K \left(3R_l^{2n_i} - 2R_l^{3n_i}\right)}{R_l^{\Sigma n_i, i\in K}}\right)$$
$$= \ln\left(e^{-\lambda_v kt} \cdot e^{-\lambda_l t \Sigma n_i, n_i \in K} \cdot \prod_{i\in K}^K \left(3 - 2e^{-\lambda_l t n_i}\right)\right)$$
$$= -\lambda_v kt - \lambda_{\,l} t \sum_{i\in K}^K n_i + \sum_{i\in K}^K ln\left(3 - 2e^{-\lambda_l t n_i}\right) \quad (3.17)$$

There are two parts in the equation for the maximization of the η. The first part, drawback of introducing the voters, is $\lambda_v kt$, which means fewer numbers and more reliable voters. The second part is $\sum_{i\in K}^K (ln\left(3 - 2e^{-\lambda_l t n_i}\right) - \lambda_l t\, n_i)$. When $\Lambda_{\mathrm{i}} = \lambda_l t\, n_i$ $(\Lambda > 0)$, the function $f(\Lambda_{\mathrm{i}}) = \ln(3 - 2e^{-\Lambda_{\mathrm{i}}}) - \Lambda_{\mathrm{i}}$ has a maximum point at $\Lambda_{\mathrm{i}} = \ln\left(\frac{4}{3}\right) \approx 0.288$. Considered with λ_l and t, the n_i needs to be designed to an appropriate value. Meanwhile, there is a sum for the k groups, and the k needs to augment here, which conflicts with the previous part.

In addition, we assume that the quantity of circuit witch needs TMR is fixed, and each block have the same size, i.e. $n_i = n_A (i \in K), S_{tmr} = n_A \cdot k$. Now the quality factor is

$$\eta = -\lambda_{\mathrm{v}} t \frac{S_{tmr}}{n_A} - \lambda_{\,l} t \mathrm{S}_{\mathrm{tmr}} + \frac{S_{tmr}}{n_A} ln\left(3 - 2e^{-\lambda_l t n_A}\right)$$

$$= \frac{S_{tmr}(ln\left(3-2e^{-\lambda_l t n_A}\right)-\lambda_{\mathrm{v}} t)}{n_A} - \lambda_{\,l} t \mathrm{S}_{\mathrm{tmr}} \quad (3.18)$$

In the expression, the $S_{tmr}, t, \lambda_{\mathrm{v}}, \lambda_{\,l}$ are given as the parameters and there is a simple need to consider merely $\frac{ln\left(3-2e^{-\lambda_l t n_A}\right)-\lambda_{\mathrm{v}} t}{n_A}$. It is rather complex to find the peak value for this function by an algebraic method. But with a plot tool, we can find that it has a jump breakpoint because of $\lambda_{\mathrm{v}} t$, and it keep decreasing after the breakpoint, which means if the voter has a lower error rate than the components, i.e. $\lambda_{\mathrm{v}} \ll \lambda_l$, the smaller the block (n_A), the higher the reliability (η).

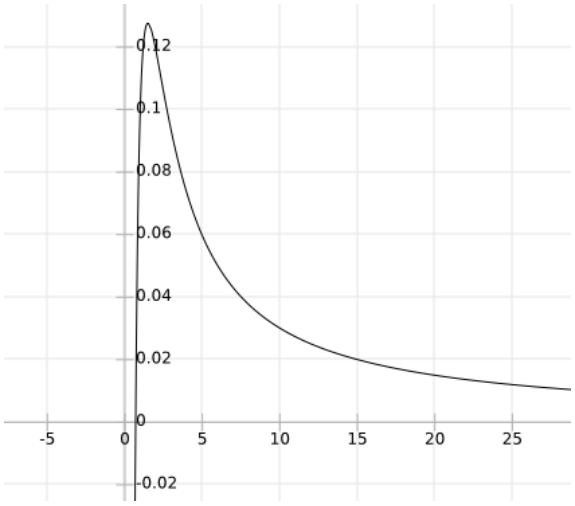


Figure 3.3 The plot of $f(n_A) = \frac{ln(3-2e^{-n_A})-1}{n_A}$. There is a jump breakpoint at x=0.

As per the previous discussion, in the circuit design, two parts, *K* and n_i need to focus on optimization, which represent the number of voters and their distribution in the circuit. Too many voters will decline the reliability and further elevate the complexity of the system. The module being tripled needs to be appropriate because the large module has bad reliability, and a small one needs more voters. As per the actual situation in MIMOSIS, more TMR design solutions will be introduced in detail in Chapter 4.

3.3 Error Correction Code (ECC)

TMR is an appropriate choice for digital circuits, which, however, will cost too many resources for large amount of data storage. It is necessary to have a proper error detection and correction method for the data array, where Error Correction Code (ECC) is designed to detect and even correct errors. It was first designed for unreliable data storage hardware and communication protocol. There are several branches based on disparate encoding methods and error detection roles.

3.3.1 Mechanism and Codeword distance

Above all, it is impossible to judge the correctness from unpredictable and unrelated data. The key is to make the data have some connection with itself with some rules. For instance, the parity check makes the count of bit '1' in the data, which is always odd, or even with an additional parity bit. If the data received does not fit the parity check, there is at least one-bit error, but the parity check cannot identify the error position.

The code distance is a critical indicator of the error detection method. It makes a certain distance between valid codewords. The minimum distance of two codes is defined as

$$d := \min_{\substack{m_1, m_2 \in \Sigma^k \\ m_1 \neq m_2}} \Delta[C(m_1), C(m_2)] \tag{3.19}$$

In the formula, Σ is the alphabet of the code, k is the message length, n is the block length, and $\Delta(C_1, C_2)$ is the Hamming distance of two code, that is to say, the number of positions in which C_1 and C_2 differ. Since any code has to be injective, any two codewords will be disparate in at least one position, so the distance of any code is at least 1 [3]. The hamming distance limit of binary block codes is shown in figure 3.4.

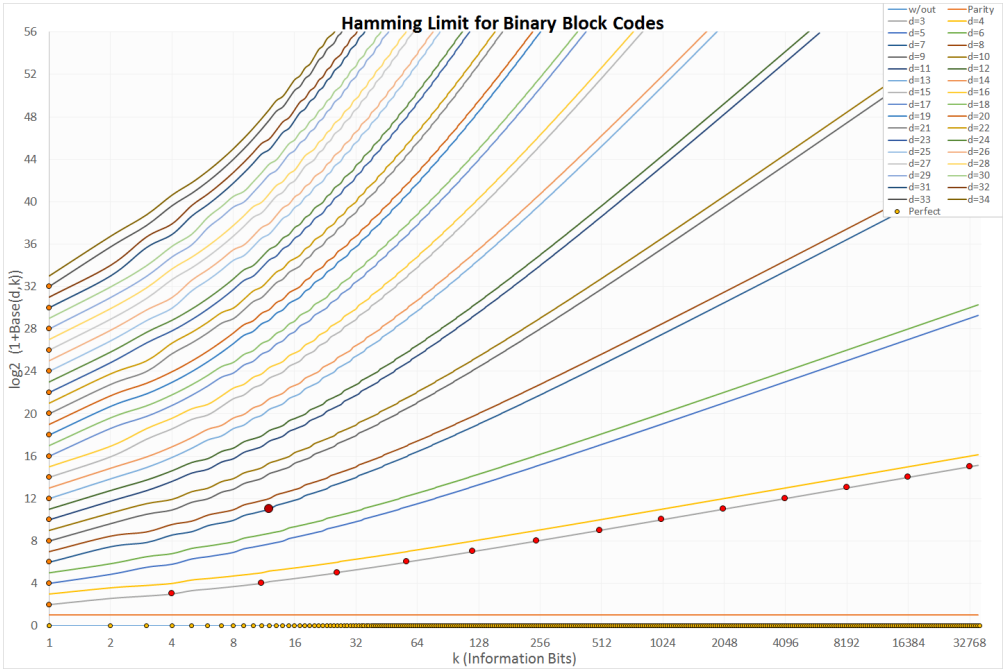


Fig. 3.4. The hamming limit of binary block codes [4].

A code with distance d allows the receiver to detect up to $d-1$ transmission errors since changing $d-1$ positions of a codeword can never accidentally yield another codeword. Moreover, if no more than $(d-1)/2$ transmission errors occur, the receiver can uniquely decode the received word to a codeword, which is because every received word has at most one codeword at a distance $(d-1)/2$. If more than $(d-1)/2$ transmission errors occur, the receiver cannot uniquely decode the received word in general owing to the possibility of several codewords. One way for the receiver to cope with this situation is to use list decoding, in which the decoder outputs a list of all codewords in a certain radius.

3.3.2 Hamming code

In the field of telecommunications, the Hamming code is a linear error correction code, which was invented by Richard Wesley Hamming in 1950s. In contrast, in addition to not correcting errors, a simple parity check code can only detect an odd number of errors. The Hamming code is a complete code, which can achieve the highest code rate among the codes with the same packet length and the minimum distance of 3. [4]

In mathematical terms, Hamming code is a binary linear code. For all integers r $\geqslant$ 2, there is a packet length $n = 2^r - 1, k = 2^r - r - 1$ encoding. As a result, the code rate of the Hamming code is $R = \frac{k}{n} = \frac{1-r}{2^r-1}$, which is the highest for a code with a minimum distance of 3 and a packet length of $2^r - 1$. The parity check matrix of the Hamming code is constructed by listing all non-zero column vectors of length r.

In order to encode the m bits data into a Hamming code, there will be need for k additional bits which will fit $m + k - 1 \leq 2^k$. The parity bit $P_i(i\ in\ [0, k-1])$ is situated in 2^i position of the code content and the data fill the left places in the chain. The parity bit P_i controls the data with the position binary code bit i equals 1, that is $P\ \&\ 2^i = 1$. For instance, 15-bit data need 5 parity bits. The encoding process is shown in figure 3.5. The value of parity bits is the parity checksum of the controlled data. If the data bit D7 at position 11 upsets, the parity bits P0, P1, P3 will not fit, and then we can get the position 11 from the binary code 1011.

No.		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
content		P 0	P1	D1	P2	D2	D3	D4	P3	D5	D6	D7	D8	D9	D10	D11	P4	D12	D13	D14	D15
parity bit coverage	P 0	■		■		■		■		■		■		■		■		■		■	
	P1		■	■			■	■			■	■			■	■			■	■	
	P 2				■	■	■	■					■	■	■	■					■
	P 3								■	■	■	■	■	■	■	■					
	P 4																■	■	■	■	■

Fig. 3.5. The encoding process of hamming code.

3.3.3 CRC

Cyclic Redundancy Check (CRC), first published by W. Wesley Peterson in 1961 [5], is a hash function that forms short fixed-digit check codes based on data such as network packets or computer files, which is mainly utilized to detect or verify errors that may occur after data transmission or storage. The produced number is computed before transmission or storage and appended to the data, and then the receiver checks to determine whether the data has changed. It is widely utilized because this function is easy to use with binary computer hardware, easy

to perform mathematical analysis, and is particularly good at detecting errors prompted by transmission channel interference.

CRC is a kind of checksum, which is the remainder acquired by dividing the two-byte data stream by binary division (no carry, use XOR instead of subtraction). The divisor is the binary representation of the information data stream that needs to be checked, which is a predefined binary number of lengths $(n + 1)$, usually expressed in terms of polynomial coefficients. Before dividing, it's necessary to add n zeros after the information data. CRC is a polynomial ring based on a finite field GF (2) (i.e., congruence divided by 2). In general, the formula is:

$$M(x) \cdot x^{\{n\}} = Q(x) \cdot K(x) - R(x) \tag{3.20}$$

In the equation, $M(x)$ is the original information polynomial, $K(x)$ is the key polynomial of n order, $M(x) \cdot x^{\{n\}}$ means that the original information is followed by n zeros, and $R(x)$ is the remainder polynomial, which is the CRC checksum. In communication, the sender appends n-bit R (replaces the originally added 0) after the original information data M and then sends it out. After receiving M and R, the receiver checks whether $M(x) \cdot x^{\{n\}} + R(x)$ can be $K(x)$ divisible, and if so, the receiver thinks the information is correct.

The error detection capability of CRC depends on the order of the key polynomial and the specific key polynomial utilized. The error polynomial $E(x)$ is the XOR result of the received message codeword and the correct message codeword. The CRC algorithm cannot detect errors if and only if the error polynomial can be divisible by the CRC polynomial. Since the calculation of CRC is based on division, no polynomial can detect errors in a set of all zero data or zeros that were lost in the past. Whereas, this problem can be solved in the light of a variant of CRC. All errors with only one data bit can be perceived by any polynomial with at least two non-zero coefficients. The error polynomial is $x^{\{k\}}$, and $x^{\{k\}}$ can only be separated by the polynomial $x^{\{i\}}$ of $i \leq k$. CRC can detect all double-bit errors whose separation distance is less than the polynomial order. In the circumstances, the bit error polynomial is $E(x) = x^i + x^k = x^k \cdot (x^{i-k} +$

1), $i > k$. For CRC polynomials, as shown in table 3.1, there are many standard polynomials available, which, furthermore, can also be designed as other polynomials, with reference to the research of [6].

Table 3.1 the standard CRC polynomials

Name	Polynomial	Binary
CRC-1	$x + 1$	0x1
CRC-8	$x^8 + x^7 + x^6 + x^4 + x^2 + 1$	0xD5
CRC-16-CCITT	$x^{16} + x^{12} + x^5 + 1$	0x1021
CRC-32-IEEE802.3	$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$	0xEDB88320

3.3.4 Discussion

TMR can fix errors in real-time, but the cost is at least 3 times that of components. The Hamming code can correct the one-bit error, but it needs to be designed with a proportional data word size. If the word is short, many parity bits are required. On the other hand, long data words make the check polynomial very long, which means that it is difficult to implement in hardware. The CRC can accept various length of input data with a fixed cost of hardware.

3.4 Refreshing and reloading

The CMOS devices, such as microcontroller, FPGA, CPU, SRAM, are easy to be influenced by SEE. Nevertheless, most of the SEEs on CMOS devices are temporary effects, which can be recovered by a reset or reload, but the EEPROM or Flash EPROM devices are not susceptible to SEE. As shown in figure 3.6, the EEPROM or Flash EPROM use a floating gate to restore data without power supply. When it is powered off, the device has very limited capacity to collect the charge. EEPROM and Flash EPROM memory cells will not be changed by heavy ions and protons in read mode, but single event transients (SET) in peripheral circuits can cause read errors on one or more bits of the data word.

As a result, there is a design plan to use ROMs to store data backup. When the system is unstable on account of an error, a specific circuit will detect this malfunction and read out the data in the ROM and write it into the system. This process is called refresh or reload.

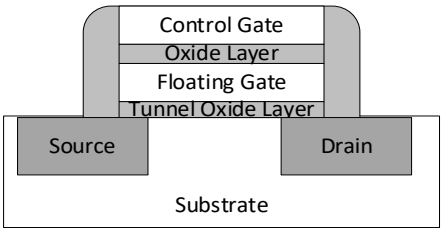


Figure 3.6. The Structure of EPROM and Flash Memory.

In most of the microcontroller, there is a peripheral called watchdog, which is a timer that can trigger an interrupt or soft reset when time up. As shown in figure 3.7, if the program set the watchdog periodically, the watchdog will not trigger. Whereas, if the program crash, the watchdog will be triggered, and the processor will jump to run the interruption handle program to recover the system.

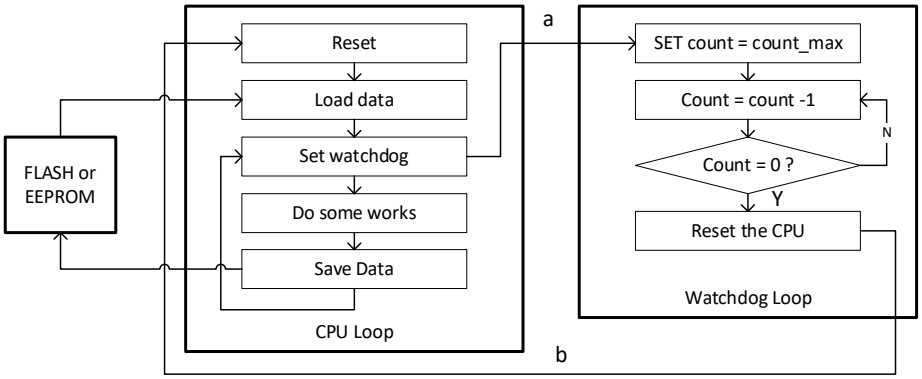


Figure 3.7 The working flow of watchdog. If the CPU does not set the watchdog counter regularly as (a), the watchdog counter will go to 0, which will reset the CPU (or trigger a non-maskable interrupt).

The same approach can be implemented on FPGA or ASIC design. Both external radiation protection watchdog and integrated TMR watchdog are available. Watchdogs can apply not only counters, but also data patents or

checksums as identification standard. Some commercial FPGA even integrated FLASH or reload circuit in the package.

The refreshing or reloading approach can handle the temporary errors with a system reset or interrupt, which needs fewer components but a long reconfiguration time. The reset also makes the system lose the current working condition. If the error occurs frequently, the system becomes unusable.

3.5 Re-entrant gate layout and guard ring

The effects of ionization damage in bulk CMOS ICs fabricated in modern commercial processes can be mitigated by utilizing a particular layout and circuit design methods that suppress intra- and inter-device leakage. As shown in Fig. 3.8, a typical layout of a fringeless field effect transistor, i.e., a re-entrant gate (also called as enclosed gate) layout transistor is asymmetrical. The transistors' characteristics are subject to whether the drain D or the source S is inside the gate, surrounded by the gate G. The effective aspect ratio of this transistor can simply be approximated by the midline, that is to say, the geometric mean of the gate width $W_{midline}$ divided by the common length L_{common} of the shortest distance from the source to the drain. The $W_{midline}$ and L_{common} representations of re-entrant gate layout transistors are shown in figure 3.8, where this approximation is not very accurate.

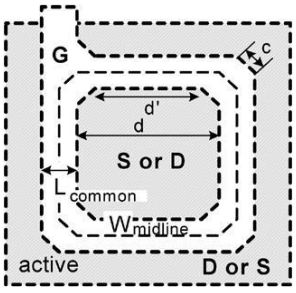


Figure 3.8. A typical layout of a re-entrant gate layout transistor.

In order to determine the effective W/L of the gate, some groups [7] have developed an approximate formula for the surrounding gate layout transistor based on device simulation:

$$\frac{W}{L}=\frac{8\alpha}{\ln\frac{d'}{d'-2\alpha L}}+\frac{4K(1-\alpha)}{\sqrt{\alpha^2+2\alpha+5\ln\frac{1}{\alpha}}}+\frac{3(d-d')}{2L} \tag{3.21}$$

In the formula, d' is the effective length minus the 45° corner side, where $d'=d-c\cdot\sqrt{2}$, c is the side length at the 45° angle, d is the side length of the middle gate (or source) square. This formula has two fitting parameters K and α, and manifests good agreement when measuring I-V data of a 0.25μm process. For $L>0.5\mu m$, the value of K should be 4. When the short channel effect is considered, the value of K should be 3.5. The typical value of α is 0.05.

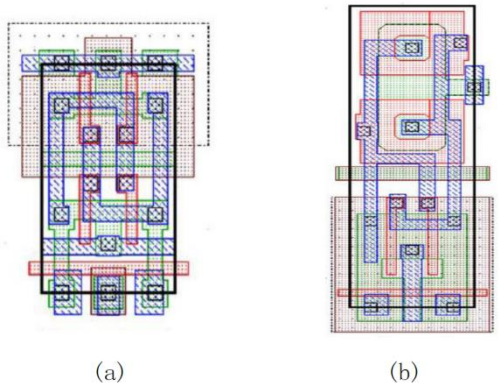


Figure 3.9. 6T memory circuit layout. (a) Standard 6T layout and (b) 6T re-entrant gate reinforcement layout

Figure 3.9. illustrates a typical Radiation Hardening by Design (RHBD) layout for a 6T SRAM bit cell. As the figure illustrates, the nFET pull-down devices are designed with reentrant gates. It should be noted that for most modern batch technologies, the accumulation of deficiencies in shallow trench isolation (STI) only reverses moderately doped p-type regions (for instance, the body of an nFET or p-substrate below STI). As a consequence, it is assumed that the pFET (pull-up and access transistor) in this SRAM design is not impacted by radiation damage and does not require a recessed gate. The layout in Figure 3.9 also has a p+ guard ring that separates the two pull-down nFETs from the n-well surrounding p-channel transistor. The addition of the guard ring augments the effective doping and the reversal potential below STI. Thus, it is less likely to

accumulate deficiencies during irradiation, which, thereby, will not cause the leakage between the n-well and the nFET drain/source diffusion layer.

One of the primary drawbacks to RHBD using reentrant gate layout technique is elevated area usage. The figure 3.9 (a) shows a layout of a standard non-RHBD SRAM cell that uses single stripe gates for all the transistors, no guard band, and nFET access devices. It can be readily observed that the area of the unhardened layout is considerably smaller than the hardened cell. Additional limitations to the hardened SRAM cell described above are the larger gate capacitances of the pull-down nFETs and the use of pFETs for access transistors, both of which decrease read/write access times. Besides, there are further limitations to design the hardened nFET circuits. As shown in the eq. 3.21, the ratio of W/L is computed by d, L and c besides W. Non-intuitive parameter settings and model inaccuracy make the circuit design more difficult.

3.6 Reverse body biasing

As described in the section 2.3, the TID may augment the device leakage current. In the deep sub-micron digital circuit design, the quiescent current (leakage current) accounts for an increasing proportion, which can be reduced by using power gating (PG) and reverse body biasing (RBB) methods [8]. The power gating declines the leakage current by cutting off the power supply of devices, which does not meet the need here. As shown in figure 3.10, the reverse body biasing uses a separated body connection besides the source (V_b). If $V_b < V_s$, the threshold voltage will augment and the leakage current will decrease, which also has a corrective effect on the offset of the threshold [9].

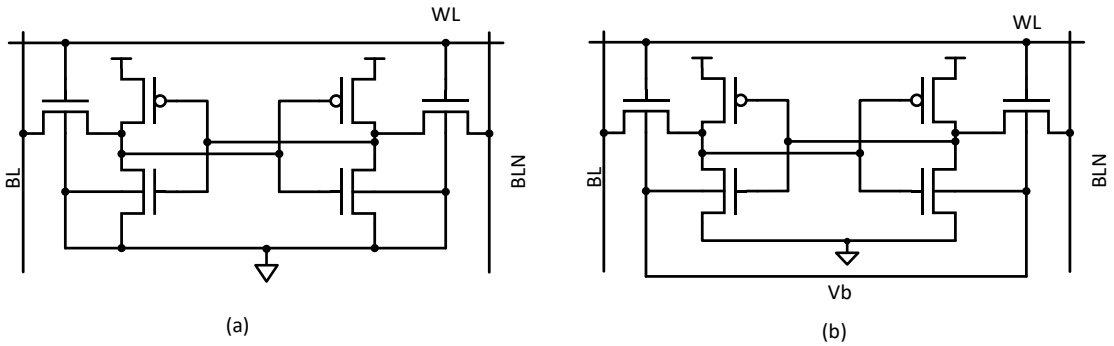
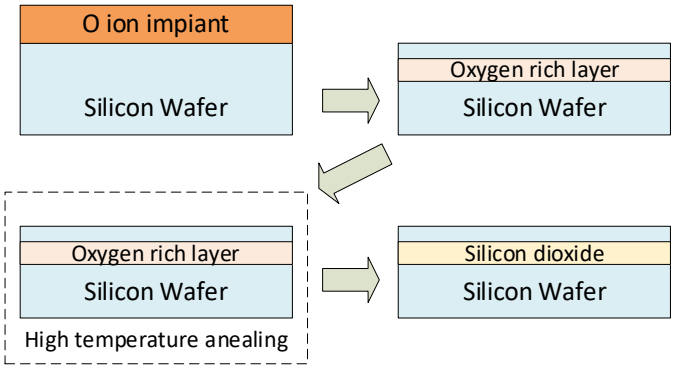


Fig 3.10 The 6T SRAM cell schematic (a) without RBB, (b) with RBB.

The reverse body bias is beneficial for hardened digital circuits, especially circuits with large leakage currents such as SRAM. Meanwhile, the body bias incurs a body effect in the MOSFET response model, which will lead to a large shift of the threshold, so it is crucial for analog circuit design. For digital circuits, the RBB also bring about the degradation in access times and noise margins as well as elevated voltage supply levels. Another disadvantage of RBB is that since the TID is a cumulative effect, the circuit is constantly changing while suffering the radiation. However, with appropriate modeling, it may be possible to identify lower reverse biases that achieve the same benefits.

3.7 Silicon-On-Insulator (SOI)

In semiconductor manufacturing, Silicon-On-Insulator (SOI) technology is to manufacture silicon semiconductor devices in a silicon-insulator-silicon layered substrate to decrease parasitic capacitance within the device, thereby improving the performance. [10]



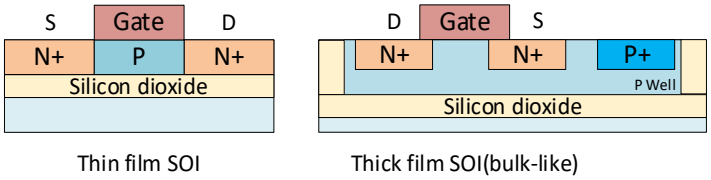


Figure 3.11. The process of manufacture the SOI wafer.

In radiation harden design, the utilization of SOI is often regarded as an attractive alternative CMOS technology used in harsh radiation environment applications. Most hardened processes are developed in SOI so as to address the primary radiation threats in deep sub-micron technologies, i.e., single event effects and soft errors [11]. For a thin film SOI process, as shown in figure 3.11, there is almost no epitaxial layer so that when SEE is produced, the active area collects very limited charges [12]. Concerning the TID effects, SOI offers few advantages over bulk. While full dielectric isolation virtually eliminates inter-device leakage, SOI transistors, particularly fully depleted n-FETs, exhibit fortified susceptibility to intra-device leakage and single transistor degradation [13]. Defect formation in the thick buried oxide can invert the backside interface creating leakage paths between drain and source, decrease threshold voltage through back-to-front gate coupling [14], and exacerbate gate-induced-drain-leakage (GIDL) [15].

3.8 Summary

In this chapter, beginning with the theory of system reliability, a series of radiation hardening methods are introduced. Certain methods do not conflict with each other and can be deployed at the same time, such as the TMR, the ECC as well as refreshing and reloading, which are from the system design level. On the other hand, the re-entrant gate layout and reverse body biasing are from the view of analog circuit designer, and the SOI is the processing way. The principles, implementation methods, advantages and disadvantages of these methods are discussed in detail. The TMR is fit for small but critical part and the ECC is usually utilized for data flow, while the refreshing and reloading are utilized for monitoring a

system working flow. These three methods are mainly utilized to resist SEEs. The re-entrant gate layout and reverse body biasing are utilized to reduce the TID effect.

In the design of MVD sensors, the radiation-tolerance performance is not the only design parameter, but material budget, power dissipation and detection performance are also important figures of merit. However, improving simultaneously all these design parameters result in conflicts. Especially, implementing all radiation-hardening techniques will enlarge the material and power budget. The key is to balance these design parameters, and the specific optimization reached for the MIMOSIS sensor is described in the next chapter.

3.9 Bibliography

[1] R. E. Lyons and W. Vanderkulk, *The Use of Triple-Modular Redundancy to Improve Computer Reliability*, *IBM J. Res. Dev.* **6**, IBM, (2010) 200.

[2] ECSS, *Space Product Assurance - Techniques for radiation effects mitigation in ASICs and FPGAs handbook*, *Structure*, (2016) 1.

[3] W. C. Huffman and V. Pless, *Fundamentals of Error-Correcting Codes*, in *Fundam. Error-Correcting Codes* **215**–**216**, Cambridge University Press, (2003).

[4] *HammingLimit - Block code - Wikipedia*https://en.wikipedia.org/wiki/Block_code#/media/File:HammingLimit.png (12 August 2020).

[5] W. W. Peterson and D. T. Brown, *Cyclic Codes for Error Detection*, in *Proc. IRE* **49**, (1961).

[6] Philip Koopman, *Best CRC Polynomials*, *Carnegie Mellon Univ.*, (2015)https://users.ece.cmu.edu/~koopman/crc/index.html (9 July 2020).

[7] A. Giraldo, A. Paccagnella and A. Minzoni, *Aspect ratio calculation in n-channel MOSFETs with a gate-enclosed layout*, *Solid. State. Electron.* **44**, Elsevier Science Ltd, (2000) 981.

[8] H. Xu, R. Vemuri and W. Ben Jone, *Run-Time Active Leakage Reduction by Power Gating and Reverse Body Biasing: An Energy View*, in *26th IEEE Int. Conf. Comput. Des. 2008, ICCD*, (2008).

[9] L. T. Clark, K. C. Mohr, K. E. Holbert, X. Yao, J. Knudsen and H. Shah, *Optimizing Radiation Hard by Design SRAM Cells*, in *IEEE Trans. Nucl. Sci.* **54**, (2007).

[10] G. K. Celler and S. Cristoloveanu, *Frontiers of silicon-on-insulator*, in *J. Appl. Phys.* **93**, American Institute of PhysicsAIP, (2003).

[11] P. R. Fleming, B. D. Olson, W. T. Holman, B. L. Bhuva and L. W. Massengill, *Design technique for mitigation of soft errors in differential switched-capacitor circuits*, *IEEE Trans. Circuits Syst. II Express Briefs* **55**, (2008) 838.

[12] Y. Zhang, Y. Zhao and R. Zhao, *Model to predict the number of transistors in an asymmetrical priority Address-Encoder and Reset-Decoder readout circuit for monolithic active pixel sensors for high-energy physics*, *Nucl. Instruments Methods*

Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip. **904**, Elsevier B.V., (2018) 171.

[13] P. Paillet, M. Gaillardin, V. Ferlet-Cavrois, A. Torres, O. Faynot, C. Jahan, L. Tosti and S. Cristoloveanu, *Total ionizing dose effects on deca-nanometer fully depleted SOI devices*, *IEEE Trans. Nucl. Sci.* **52**, (2005) 2345.

[14] O. Flament, A. Torres and V. Ferlet-Cavrois, *Bias Dependence of FD Transistor Response to Total Dose Irradiation*, *IEEE Trans. Nucl. Sci.* **50**, (2003) 2316.

[15] P. C. Adell, H. J. Barnaby, R. D. Schrimpf and B. Vermeire, *Band-to-Band Tunneling (BBT) Induced Leakage Current Enhancement in Irradiated Fully Depleted SOI Devices*, in *IEEE Trans. Nucl. Sci.* **54**, (2007).

Chapter 4. Radiation-Hardened Design of the MIMOSIS Frontend Readout System

Schematically, any pixel sensors can be divided in two main functional parts: the pixel array and the frontend readout system. As introduced in Chapter 1, the expected performances and operational environment of the MVD in the CBM experiment require the MIMOSIS sensor to feature a complex readout system. Hence, this digital part of the chip is rather extended and possibly prone to radiation effects. In this chapter, we concentrate on the frontend readout system which consists mostly of the digital part of the sensor. In the summary of Chapter 3, it has been explained that the primary contradiction is the needs of radiation reliability and design limitations. To solve this problem, this chapter proposes a method for radiation hardening design through optimization, which is shown in Figure 4.1. Firstly, the original design without radiation hardening technology is introduced, and then the error rate of each module is evaluated by calculation. Afterwards, TMR and ECC are deployed on

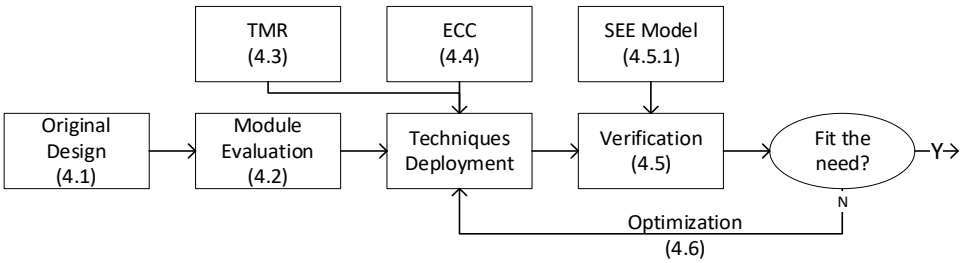


Figure 4.1 The design flow of the methodology presented in this chapter.

the module. Using the SEE model shown in Section 4.5.1, verification can be performed through multi-level simulation. If the simulation results cannot meet the requirements, the design for deployment of techniques will be revised. Through this design and verification cycle, an optimized design can be acquired.

4.1 Design principle of the readout system

As shown in Figure 1.14, the frontend readout system buffers the output of the pixel array, which outputs a 16-bit encoded data word each time, and the data format is shown in Figure 4.2, consisting of the pixel address in the column, the column number and the compression code. The 3-bit compression code presents the state of the next 3 pixels. With the help of data compression, the data word can present four adjacent pixel hits.

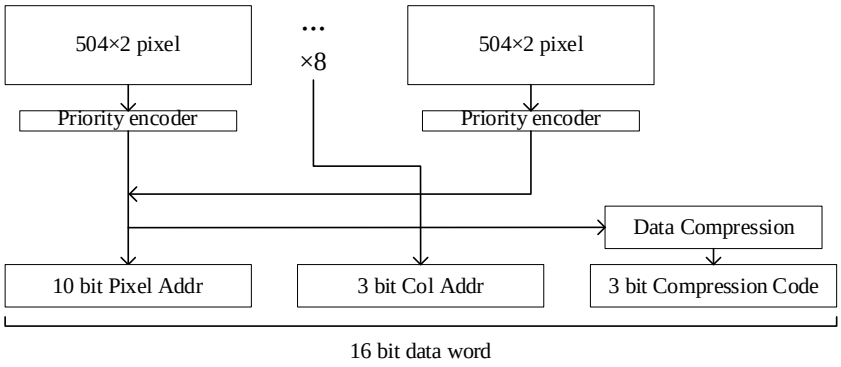


Figure 4.2 The data word format of the pixel array. Including pixel address, column address and data compression code.

The 16-bit data word is first buffered in Region Readout Unit (RRU), which, as shown in Figure 4.3, has a $128 \times 32bits$ SRAM that can buffer 256 words. In the RRU, the 6-bit region head code (64 RRUs in MIMOSIS-1) is added in front of the data word set, and the RRU try to save all the data from the global $Frame\_Start$ signal until the $Frame\_End$ signal. But, it will overwrite the older data if it overflows. The RRU output is controlled by a token ring, which, when the token is acquired, reads and outputs the buffered data words and hands over the token when the buffer

is empty. The write process working in parallel in the pixel array, under $20MHz$ clock, while the read process works with the $40MHz$ clock.

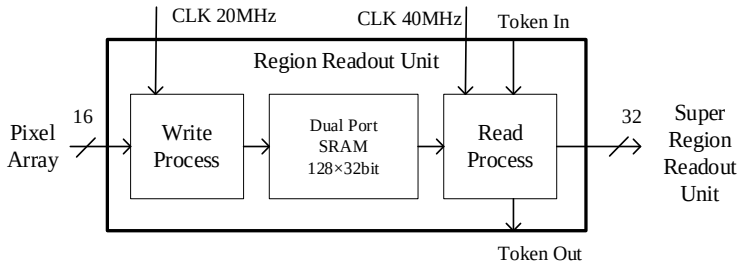


Figure 4.3 The data processing block diagram of the Region Readout Unit.

As shown in Figure 4.4, the Super Region Readout Unit (SRRU) is combined with the data from four RRUs. Each RAM word contains 8 RRU words and the RAM capacity of SRRU equal to 4 RRUs'. Both the write and read process use the $40MHz$ clock in SRRU: the former uses the last token of the RRU as the finished signal and the latter uses the token ring between the SRRUs to control the output.

The RRU and SRRU have the same architecture, but disparate memory size and clock domain. The RRU is designed to match the readout speed of the pixel array, while SRRU aims to buffer and merge RRU data to acquire higher readout speed and avoid overflow. In the detector, pixels fire as clusters, which means some columns have lots of events, but others do not. Meanwhile, the hit events are not coming continuously but bristly and multi-layer buffer helps to slow down the data rate.

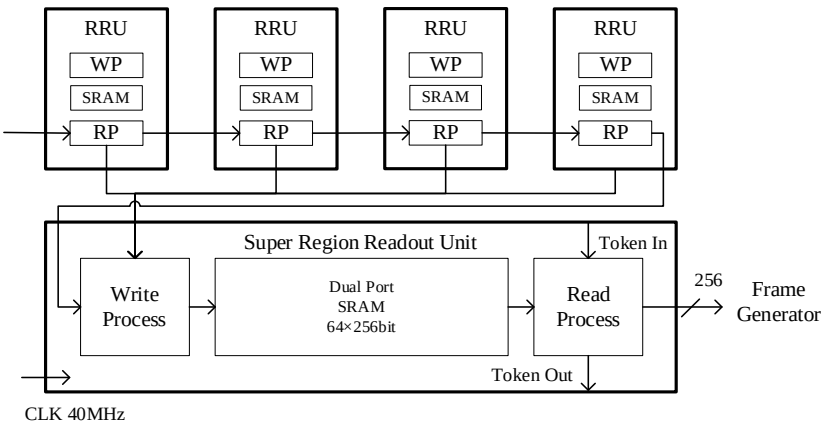


Figure 4.4 The data processing block diagram of the Super Region Readout Unit.

After the SRRU, as shown in Figure 4.5, there is a multiplexer to route the 16 SRRUs outputs into one. The outputs of SRRUs' are controlled by a token ring. So, in normal condition, there are only one SRRU that has output. Then, the frame generator adds "frame start" and "frame end" label including the frame number to the data word set so as to identify a "frame", which is the time base of the detector. In order to further improve the data readout efficiency, the idle remover will combine the SRRU word with the idle words added in the SRRUs. The frame generator buffers the SRRU output with $40MHz$ but processing the data with $80MHz$.

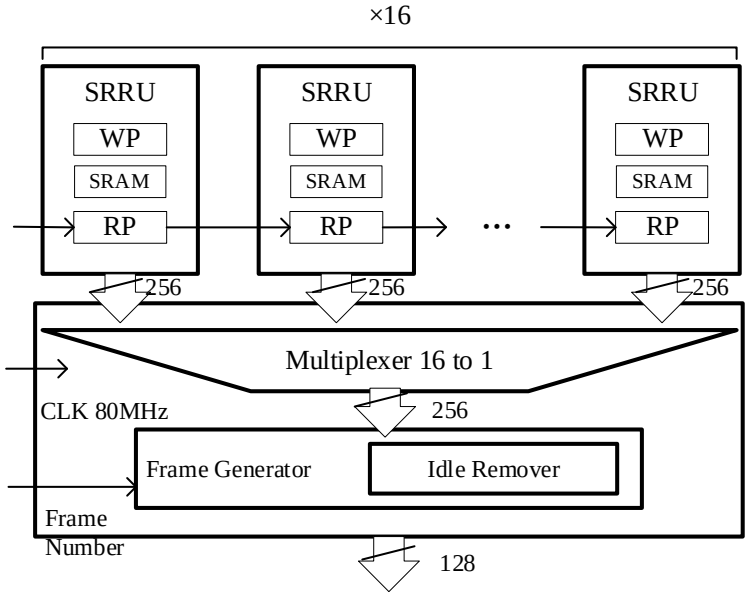


Figure 4.5. The block diagram of the Multiplexer and Frame Generator.

The last buffer of the readout system is the Elastic Buffer, which has a large SRAM FIFO with the same size of all 16 SRRUs' as shown in Figure 4.6. The FIFO buffers output of the Frame Generator and feeds the Data Serializers that are working on $320MHz$ clock and have only one-bit output connected with SLVS Driver. Since the Data Serializers have very high-power consumption, the Serializers is designed to be able to switch off to save power when the data rate is low. On the back stations of MVD in CBM, the hit rate is much lower than the front stations. Hence, the sensors on station 2 or 3 can switch off some outputs to decline power consumption and heat dissipation budget.

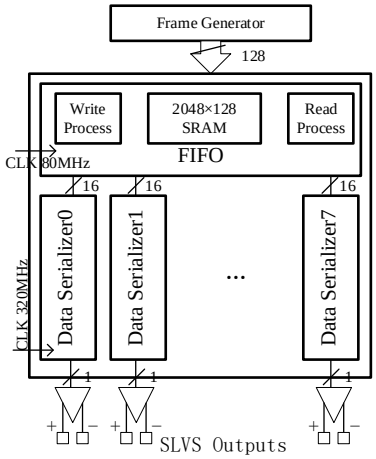


Figure 4.6. The data processing block diagram of the Elastic Buffer and Output Data Serializers.

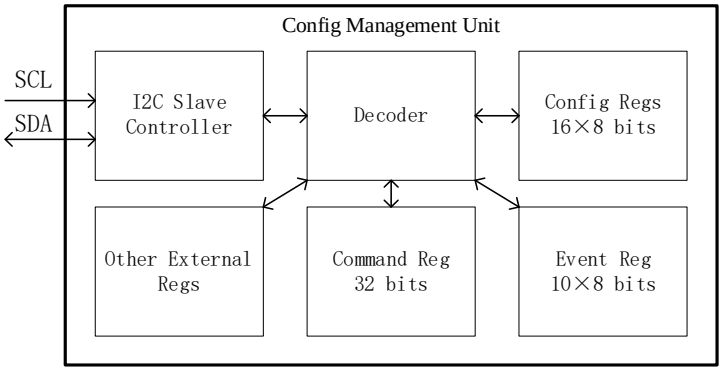


Figure 4.7. The diagram of Configuration Management Unit.

The configuration registers of the sensor chip are managed by the Configuration Management Unit (CMU), which as shown in Figure 4.7, allows testers to use I^2C protocol to exchange data with configuration registers, event monitor registers, command registers and other external

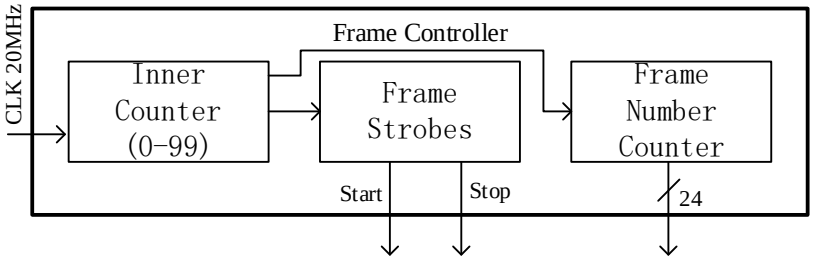


Figure 4.8. The diagram of frame controller.

peripheral registers. In normal working situations, the configuration is loaded before the detection and barely reloaded during the acquisitions.

Besides, there are global controllers that serve as the frame controller and the clock divider. As shown in Figure 4.8, the frame controller contains an internal counter with a maximum count of 100 and the counter runs on a $20MHz$ clock with a period of $50ns$. For 100 counts, the frame duration is 5 μs. A 24-bit accumulator is utilized to generate frame number.

In the clock divider, as shown in Figure 4.9, there are 4 stages of divider registers, each of which divides the input frequency by two. The output of the divider constitutes the four clock domains of the system.

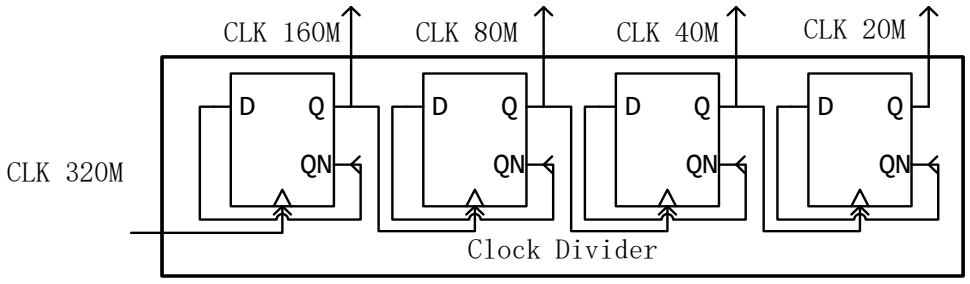


Figure 4.9. The diagram of clock divider. The input clock is coming from the PLL.

In the readout system, there are 768K bits of SRAM and each stage has its buffer and control logic, where the registers and SRAMs are sensitive to SEU. The readout system has 4 clock domains and the clock trees are sensitive to SET. The TID effect may cause the static power of the SRAM to augment and performance to reduce. According to the methods introduced in Chapter 3, TMR is adequate for registers while clock trees and ECC is suitable for SRAM. But with the limitation of the experiment environment, the deployment of TMR and ECC needs optimization, which is on two parts: use less material and ensure the reliability. The first step is to assess the reliability of the original system.

4.2 Evaluation

Before the design, there is need to assess the reliability of the system by calculations based on the cross-section, which can help to roughly estimate the error-rate of each part.

4.2.1 The cross-section and hit density of MVD

For estimating the circuit reliability against SEE, two ingredients are indispensable: the rate of highly ionizing particles and the SEE cross-section of the circuit. Both quantities bear large uncertainties since both the accelerator and the sensor are in development. Consequently, the verification is based on the known cross-section from another similar CMOS pixel sensor (ALPIDE) utilized in the ALICE experiment. The ALICE experiment of CERN's LHC has some similarities with the CBM experiment. The Internal Tracking System (ITS) also uses the TowerJazz 0.18 μm MAPS process. In Pb-Pb collisions, the event rate is 50 kHz (up to 10 MHz Au-Au collisions in CBM) [1]. But higher is the energy of ALICE, which uses ring collider detectors, while the CBM uses fixed targets and planar detectors, which may result in more energetic particles to strike the detectors. The SEE test results of ITS in ALICE is shown in Table 4-1 the SEE test results of ITS in ALICE.[2]

Table 4-1 the SEE test results of ITS in ALICE

Typical SEU cross section	Pixel mask bit (Registers)	$\sigma_{Reg} = 10 \times 10^{-14} cm^2/bit$
	Region Memory (SRAM)	$\sigma_{SRAM} = 7 \times 10^{-14} cm^2/bit$
TID Threshold shift	<40mV for NMOS L=0.18um @ 20Mrad <120mV for NMOS L=0.35um @ 20Mrad	
NIEL	$>10^{13}$ 1MeV n_{eq}/cm^2	

The maximum hit density of ITS layer-0 is $1.5 MHz/cm^2$ [3], in which the high energy hit density ($Energy > 20 MeV$) is $770 KHz/cm^2$, i.e., ~5%

of the total impact particles. Scaling to the CBM experiment, the peak hit density is expected to be $7 \times 10^7 Hz/cm^2$. Even though the geometry and collision energy are very disparate between ALICE and CBM, the high energy hit density is foreseen as $\phi = 3.5 MHz/cm^2$. The SEU rates of registers and SRAMs are computed as:

$$\rho_{Reg} = \sigma_{Reg}\Phi = 3.5 \times 10^{-7} Hz/bit$$

$$\rho_{SRAM} = \sigma_{SRAM}\Phi = 2.45 \times 10^{-7} Hz/bit \quad (4.1)$$

4.2.2 The error rate in control flows

In RRU, SRRU, Frame Generator and Elastic Buffer, there are total registers $N_{Reg} = 7725$, which contains N_{DReg} =4736 data buffer registers and $N_{CReg} = 2989$ control registers. In station layer 0, there are about N_{chips} =20 chips. If assuming all the control registers are crucial, the mean time to malfunctions (MTTF) of SEU will be:

$$T_{mttf-SEU} = \frac{1}{N_{chips} N_{Creg} \cdot \rho_{Reg}} \approx 48\text{s} . \quad (4.2)$$

This means, only considering SEU on control registers, on average in each 48 seconds, the system will have a crucial error and needs to be reset, and MVD with this error rate is not useable for CBM experiment. If the TMR is deployed on all the control registers, the TMR will correct all the single bit error in a clock period. There are 4 clock domains in read out system, but we use the maximum $T = \frac{1}{20MHz} = 50ns$ to have calculation margin. So, the single bit toggle possibility is

$$P_{se} = \rho_{Reg} T = 1.75 \times 10^{-14}/bit . \quad (4.3)$$

As introduced in section 3.2, for the TMR, the output errors can only be prompted by the voter or at least 2 error input. The voter error is a SET error which is much rare compare with the SEU. The possibility of these errors can be ignored, so the double error rate per triplet is

$$\frac{P_{de}}{T} \cong \frac{3(1-p_{se})p_{se}^2}{T} \cong 1.84 \times 10^{-20} s^{-1} . \quad (4.4)$$

And the MTTF of SEU will be:

$$T_{mttf-SEU} = \frac{1}{N_{chips} \cdot N_{Creg} \cdot \frac{P_{de}}{T}} = 9.09 \times 10^{14} s \ . \quad (4.5)$$

The result seems to be good for reliability, but the additional registers and voters will take up 50% more space and 100% more power consumption, and besides, the routing and time costs are not regarded here. The calculation indicates that either the original or the full TMR design is not adequate for the requirement of CBM, so we need to find a compromise between the two.

4.2.3 The error rate in data flows

For the data registers in readout system, the error rate is computed as

$$P_{e-Dreg} = \frac{N_{Dreg} \cdot \rho_{Reg}}{R_{data}} = \frac{4736 \times 3.5 \times 10^{-7} s^{-1}}{2.56 Gbps} = 6.5 \times 10^{-15} \quad (4.6)$$

In the formula, R_{data} is the read-out data rate of the system, and then the error rate in SRAM is

$$P_{e-SRAM} = \frac{N_{SRAM} \cdot \rho_{SRAM} \cdot \eta_{full}}{R_{data}} = 7.53 \times 10^{-11} \quad (4.7)$$

In the equation, $N_{SRAM} = 768 Kbit$ is the total SRAM size of the system, η_{full} is the SRAM buffer fulfill level (take 100% as margin) and R_{data} is the read-out data rate of the system. The results indicate that the data registers and SRAMs have large margin from the data error rate requirement. The harden method is unnecessary for the SRAM and data registers.

4.2.4 The error rate in global controller

The clock divider has only 4 registers. The error rate of clock divider is

$$T_{mttf-SEU} = \frac{1}{N_{chips} N_{reg} \cdot \rho_{Reg}} = 1.79 \times 10^{4} s \approx 4.96 hours \ . \quad (4.8)$$

One error in five hours is acceptable for data path but not for clock because an error on clock may cause the load circuit out of the synchronization which needs a reset to recover the circuit. Meanwhile, the TMR

deployment on the clock divider does not cost a lot, and a full triplicated (both registers and voters) diagram is thus appropriate for clock divider.

The CMU has 240 config registers and 190 control registers. The error rate of configuration registers CMU is

$$T_{mttf-SEU} = \frac{1}{N_{chips} N_{ConfigReg} \cdot \rho_{Reg}} = 298s \, . \tag{4.8}$$

And for control registers, the error rate is

$$T_{mttf-SEU} = \frac{1}{N_{chips} N_{ControlReg} \cdot \rho_{Reg}} = 376s \, . \tag{4.9}$$

Although the error rate is moderate, the configuration registers may not be refreshed for a long time, which will cause errors to accumulate. There are two ways to fix the error. One is to reload the configuration periodically. The refresh does not need external circuits in the controller but need to add a refresh schedule on the master control blocks. At the same time, refreshing will occupy the I^2C bus for a long time, during which, other registers on the chip cannot be accessed. The other way is to use ECC to refresh the data internally through a slow clock. The hamming code can fix one-bit amount and a data word which is suitable for the CMU.

4.3 TMR Deployment

4.3.1 Design Flow with TMR

As shown in Figure 4.10, the TMR Generation (TMRG) tools designed by CERN [4] can be applied to deploy TMR and demonstrate that the circuit with TMR will function. The tool is conveniently based on the verified behaviour or RTL netlist. With user's assuming of TMR deployment list, the triplication tool (tmrg) can produce the aim netlist automatically. After the logic synthesis, the placement generation (plag) tool can help to produce the place and route (P&R) constraints to avoid the registers in TMR group placed too close with each other. With the test-bench generation (tbg) tool and SEE stimulus generation (seeg) tool, the simulator can demonstrate the response of the hardened circuit to the SEE strike.

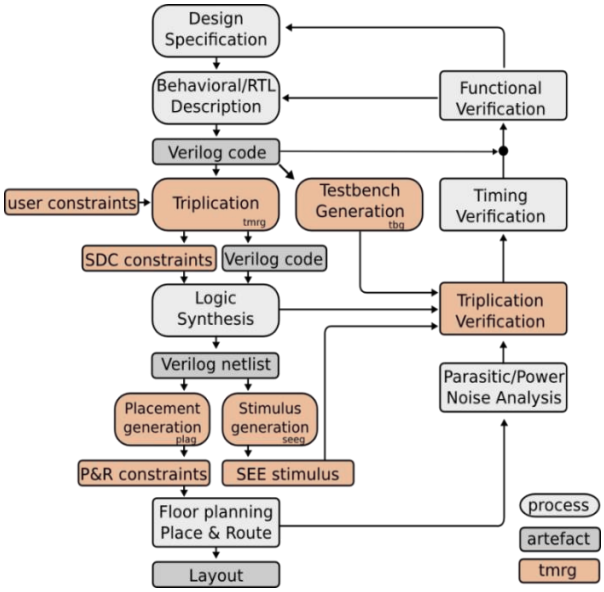


Figure 4.10. The design flow of digital back-end with TMR deployment [4].

4.3.2 Detailed Deployment

Based on the calculation in Section 4.2, the data buffer registers and SRAMs are not sensitive to SEU, while the control registers are quite the opposite. According to the severity of the error after suffering from the SEU and the size of the control area, the registers are sorted by criticality, and the result is shown in Figure 4.11.

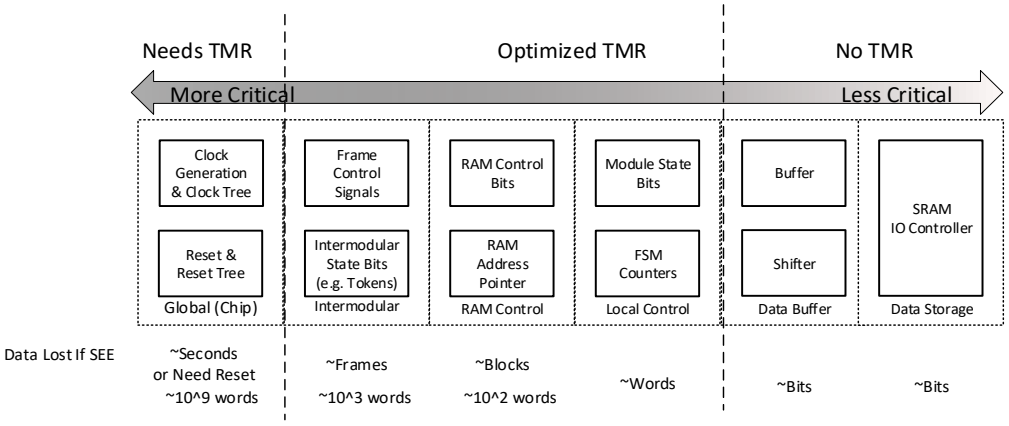


Figure 4.11. The sort of the critical level over SEE of the blocks in readout system. Registers and SRAMs are mainly concerned. Also the clock and reset circuits are involved.

First, if the clock generation or reset register fails, the system may have serious timing problems or reset problems, which will incur the entire chip fail to work during a few seconds. These modules require highest level protection, and the clock divider is devised as in Figure 4.12. Each stage has triple registers and 4 voters, among which, three are for the feedback loop of divider, and the other is for the output. The feedback voter cannot be omitted because the feedback loop will keep the error state perpetually.

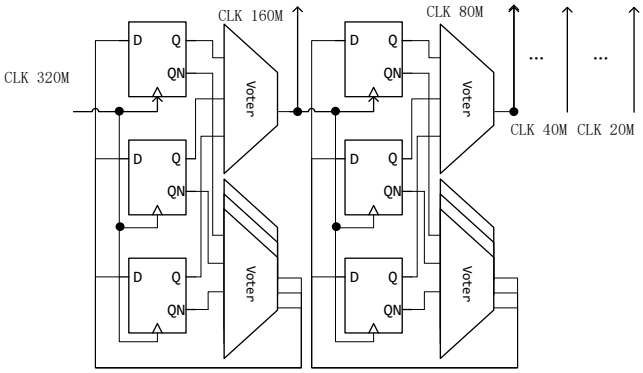


Figure 4.12. The schematic of clock divider with TMR deployment.

In addition, the “feedback loop registers” are common in digital design. The accumulator is composed of plus-one logic and registers, which form a feedback loop as shown in Figure 4.13 (a). For instance, the circular shift register has the structure shown in Figure 4.13 (b). For the "feedback loop register", if TMR needs to be deployed on the register, the voter should be

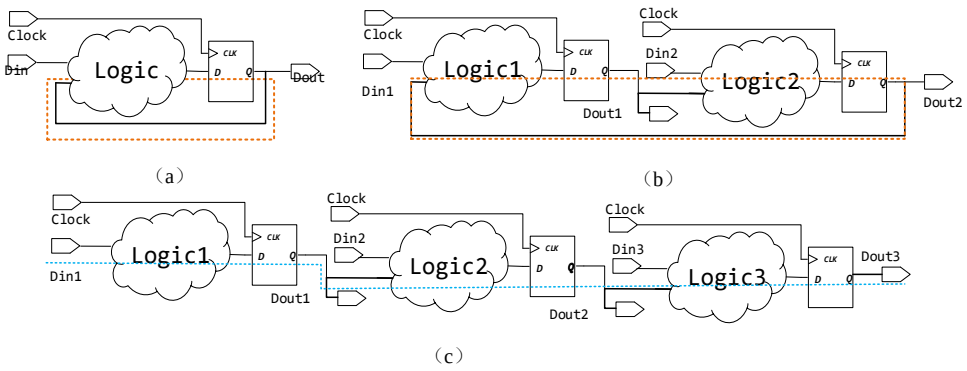


Figure 4.13. The RTL level schematic of “feedback loop registers” and flow processing registers. Part (a) and (b) have a feedback path in the data path. Part (c) is a flow processing data path.

situated inside the loop (as shown in (a)). If not (as (b)), the error will be transmitting inside the loop and will not be rectified by the voter. The single-bit error accumulating inside the loop may then result in a double-bit error.

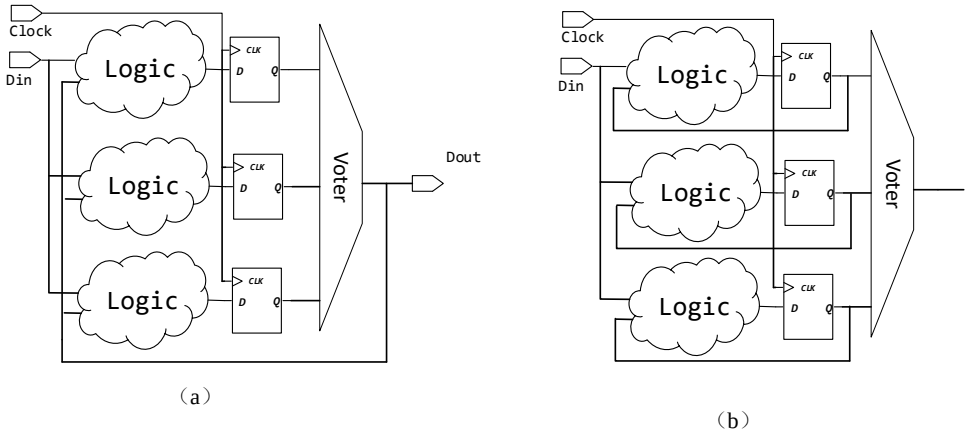


Figure 4.14*. The TMR schematic for "feedback loop registers". Part (a) and (b) has same function and components but (b) will keep the error in the loop.*

The intermodular signals like frame control strobes and tokens can cause a whole frame data lost. The frame control strobes, i.e. the $FrameStart$ and $FrameEnd$ signal, are shared amount the readout

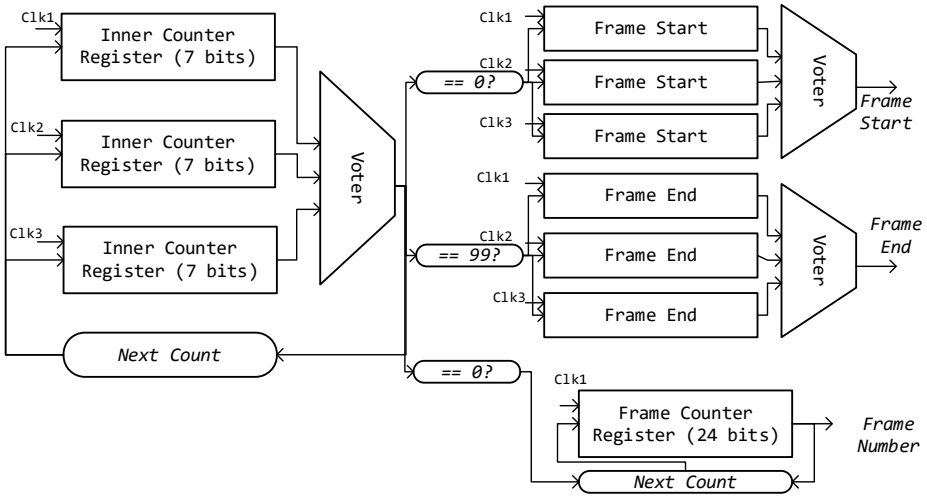


Figure 4.15. The schematic of Frame Controller. The registers are tripled and summed up with a single voter. The registers are driven by different clock domains. The Frame Counter and logics are not tripled.

system, to which most of the memory controllers are relevant to. As the design is shown in Figure 4.15, the internal counter and the strobe registers are important but the frame number is not hardened. If error occurs on the frame number counter, it is possible to correct it by software processing.

The clock tree is a balanced inverter or buffer H-tree caused by the place and route tool. The design purpose of the clock tree is to equalize the delay of the clock to each register. Even though the combinational logic is less likely to be impacted by the SEE, the clock tree is a special case because the clock signal directly controls the sampling time of registers. If there is a SET on the clock signal, it will let registers sample a wrong value, and the Propagation Induced Pulse Broadening (PIPB) effect [5] [6] may cause further errors. As described in Section 3.2, the clock can be tripled to drive each register in the TMR group.

The "next count" logics in Figure 4.15 are not tripled. If SET occurred in these combinational logics, there is a chance to cause an error. An ideal way to harden is to introduce 3 disparate phase clocks for different registers in the TMR group. As shown in Figure 4.16, these clocks have tiny phase delay Δt with each other, which can filter the SET pulse shorter than Δt, but the delay also results in a timing problem in the sequential synthesis. The extra $2\Delta t$ must be added to the total delay constraint, making the circuit run more slowly than originally devised. The larger Δt brings higher reliability but slower circuit, which is a parameter that needs to be optimized.

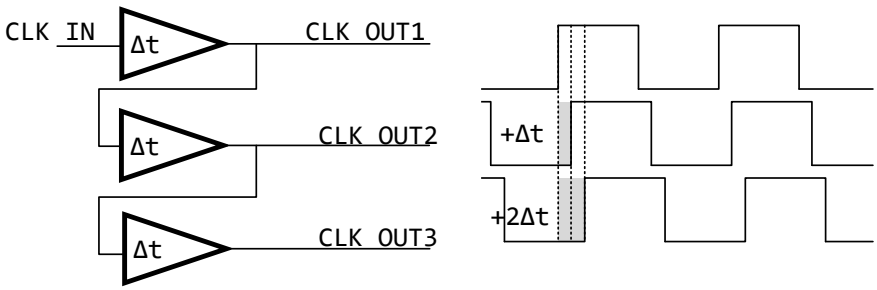


Figure 4.16. The phase delayed clock module. Each buffer has a Δt delay.

The following parts are the Finite State Machines (FSM) in memory controllers, i.e. the Reading and Writing Process. In the FSM, the state jumping is based on state counter registers, which are set to working states (A) by the start signal and reset to idle state (I). Assuming there are 4 working states in the FSM (A, B, C and D), the normal schedule of the FSM is shown in Figure 4.17 (a). If the SEE impact at the 'B' process, the content state register will be unpredictable (N) as Figure 4.17 (b), and the following processes will be lost. Fortunately, the state register can be recovered by the end signal and the recovery time of the status register is less than one processing period. If the status register is reinforced by TMR, it can be recovered quickly on the next clock edge. The correction time can be shortened to 1 clock cycle without an error output. Normally, the shorter the recovery time, the lower the error rate. Nevertheless, if the error is not crucial, the error rate may be acceptable, which is another parameter that needs to be optimized.

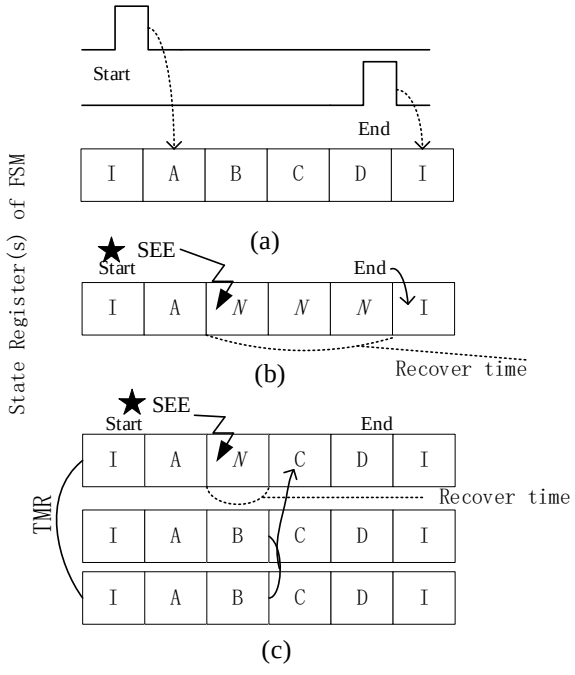


Figure 4.17. The recovery time of the State Registers in common condition and TMR condition.

Finally, there are counter's and pointer's registers as local control registers, which are applied to point out the read or write location in the

memory with the same length of the memory address. The control area of these registers is only in the block. It can be regarded that the error of the register will bring about the loss of the entire block of data at most. The balancing between error rate and size is the final optimization here.

4.4 ECC Deployment

On the basis of calculations, hamming code needs to be applied to reinforce the CMU, which has 30×8 bits registers. Each 8-bit word requires a Hamming code correction block as shown in Figure 4.18. This structure is transparent for register reading and writing and has two clock inputs: one is the normal write clock and the other is the refresh clock. The refresh clock is at a very low frequency ($\sim 100KHz$) and is disabled when the write operation is enabled. The refresh clock controls the loop, which decodes and corrects the data, then re-encodes and stores it in the register. The refresh operation can correct the error on time to prevent the double bit errors in the code word.

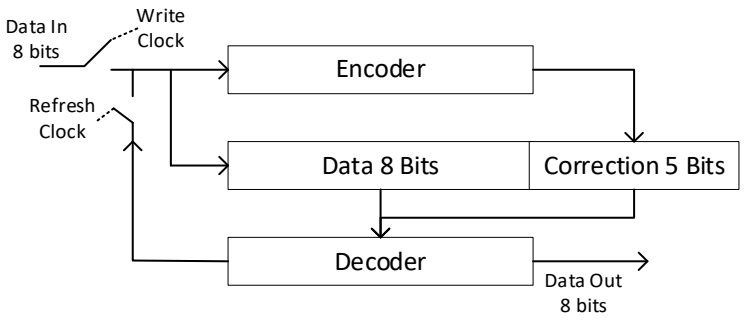


Figure 4.18. Hamming code correction protected register schematic.

4.5 Verification

In the section 4.3, the parameters that need to be optimized are determined, but a benchmark is needed to guide the optimization. Calculation is no longer a reliable way to accurately assess the error rate of a dynamic system because the calculation uses too many averages and assumptions. The static analysis is not appropriate for the random characteristics of SEE, and too many margins are applied during the

calculation which is not conducive to the optimization. In consequence, we use Monte Carlo simulation as the benchmark to be closer to the actual situation.

For the benchmark, we need to find out the input of the simulation, i.e. the SEE response of the circuit. So, we need modelling the SEE in simulation first.

4.5.1 Modelling SEE pulse

The SEE particle simulation is come up with in Synopsis Technology Computer-Aided Design (TCAD) toolset. As shown in Figure 4.19, the N-MOSFET is described as TCAD 3-D mesh network model and the rest of the components are described in the SPICE (Simulation Program with Integrated Circuit Emphasis) netlist. The driver of this circuit is a fixed input inverter, i.e., the Vout terminal will stay high. The loads are N inverters with the same dimension as for the driver.

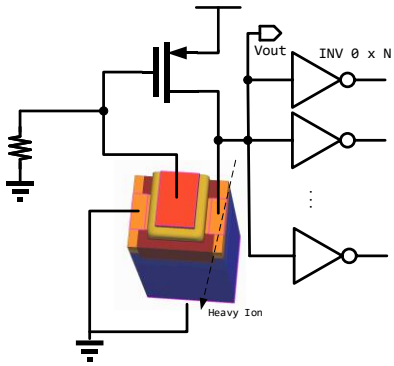


Figure 4.19. The schematic of TCAD simulation for modeling the pulse of the SEE in NMOS FET.

Table 4-2 The energy, LET and range of impact particles into silicon

Material	Energy (MeV)	LET $(MeV \cdot cm^2/mg)$	Range (μm)
H	100	5.857×10^{-3}	4.162×10^{4}
C	100	1.455	181.52
Au	100	50.51	14.72
Au	1000	93.64	60.08

The impact particles of MVD in the CBM experiment are mainly protons, carbon and gold. With the help of SRIM (Stopping and Range of Ions in Matter) [7] software, the energy, stopping power (i.e. LET) and range of impact particles into silicon are listed in the Table 4-2. As the epitaxy layer thickness of the process is about $25\mu m$, all the listed particle will be passing through.

The Heavy Ion model, based on the LET of the particle, is defined by the TCAD simulation tool. In the simulation, the heavy ion stroked the drain part of the N-MOSFET with a LET from $1\ to\ 500\ MeV \cdot cm^2/mg$, and the voltage observed at terminal $Vout$ is shown in Figure 4.20. As LET augments, the voltage drop pulse width becomes longer, but when LET augments exponentially, the pulse width only augments linearly, which is related to the ability of the active area to absorb charge. When the $LET > 10MeV \cdot cm^2/mg$, the lowest voltage of $Vout$ is close to 0, which means the active area is fully saturated. Thus, the field is too weak to get the charge carriers (electrons and holes) separated. The charges are combined with each other, which makes the pulse width shorter than expected. Besides, the impact angle and the position also affect the absorption of the ionized charges [8], but for this simulation condition, only the case of right-angle incidence is regarded.

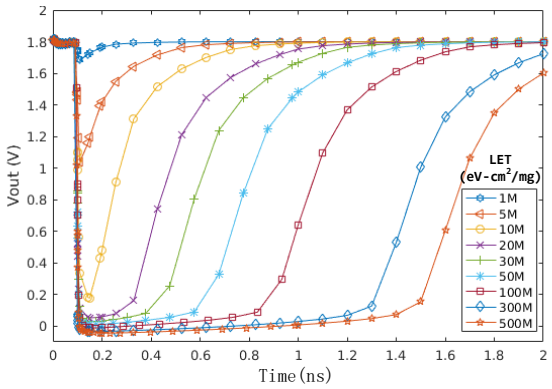


Figure 4.20. Voltage shift at Vout vs time under different LET (Fanout = 1)

Considering the different fan-out (load) of the circuit, as shown in the Figure 4.21, the pulse widths have small changes. In the low LET part, the

pulse width becomes shorter when the load augments, but in the high LET part, the pulse width augments, which is because the load capacitor helps to share the charge and maintain the level of Vout when the charge is low, and when the charge is high, it prolongs the level recovery time because more charge is stored. The result indicates that a high fan-out circuit has a higher SET/SEU threshold but once the LET is high, the high load may become a drawback of the circuit performance.

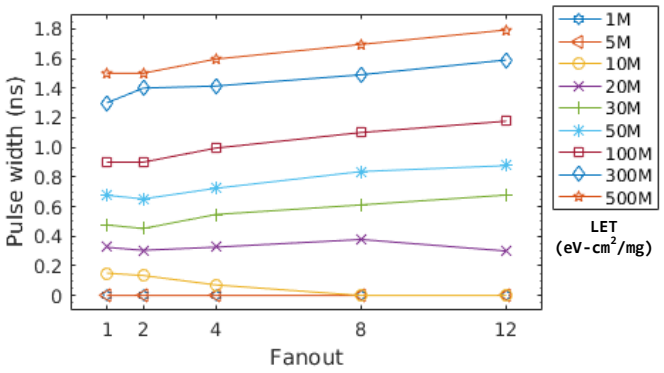


Figure 4.21. Pulse width vs fanout under different LET. The threshold of pulse is 0.9V.

With these simulation results, the particle p_i cross-section $\Theta(p_i)$ of a register or logic standard cell can be assessed by the equations:

$$\Theta_c(p_i) = \int D(p_i, E) \cdot \Theta_{\mathrm{LET}}(f_{LET}(p_i, E))\, dE$$

$$\Theta_c(LET) = P_e\left(f_{pw}(LET)\right) \cdot \rho_{sensitive} \qquad (4.10)$$

In the formula, p_i is a specific impact particle, $D(p_i, E)$ is the energy distribution function of p_i, i.e. $\int_0^{+\infty} D(p_i, E)dE = 1$, $f_{LET}(p_i, E)$ is the corresponding function of particle energy and LET as shown in Table 4-2, $\Theta_c(LET)$ is the LET cross-section, ρ_{active} is the active (sensitive) area part of the cell, $f_{pw}(LET)$ is the function shown in Figure 4.18 and $P_e(pw)$ is a threshold function, showing whether the pulse is enough to cause SET/SEU in the cell, and $\rho_{sensitive}$ is the sensitive area (mostly the drain

area) ratio of the cell. So, the number of SEEs N in a time period T corresponding to the particles' fluence $\Phi(\mathrm{p_i})$ on a M-cells circuit is

$$N = \sum_{p_i \in P}^{P} (\Phi(\mathrm{p_i}) \cdot \sum_{c_j \in M}^{M} \Theta_{\mathrm{C_j}}(p_i) \cdot S_{c_j}) \cdot T \qquad (4.11)$$

The $D(E)$ is available for some specific applications, such as aerospace, nuclear power plants, etc. Meanwhile, unfortunately, this distribution is not available for High Energy Physical (HEP) experiments. The detector is built to get the particle distribution. However, at this point, we assume that the particles are distributed equally at the energy range and take the high energy hit density as $\phi = 3.5 \times 10^6\ Hz/cm^2$. With these parameters, the SEE pulse for the SET and SEU is modeled.

4.5.2 Simulation with SEE pulse injected

With the model described in last section, the SEE pulse is produced with a serie of random numbers shown in

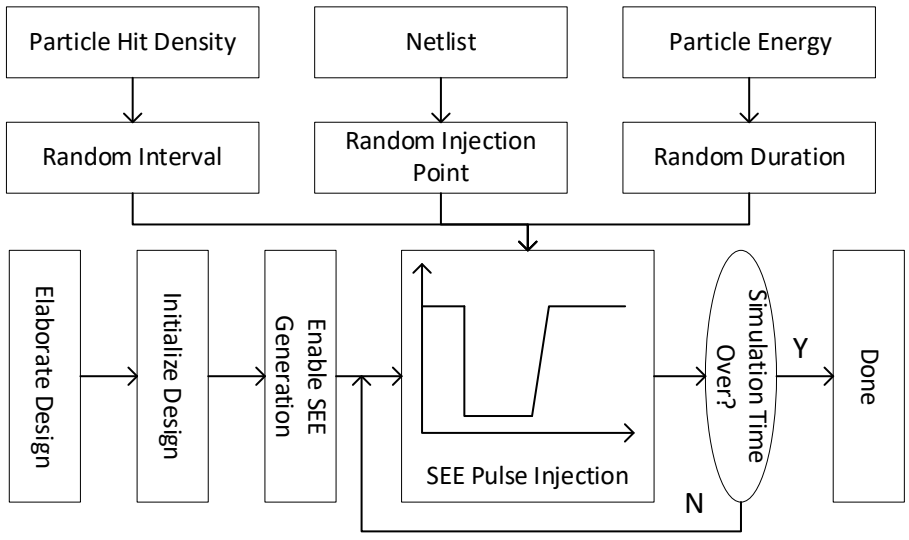


Figure 4.22. In order to produce a SEE pulse in the circuit, there are some parameters to be determined: the interval between two SEE pulse T_I, the duration of SEE pulses T_D and the injection point L_I of the circuit. The generation of T_I that needs to satisfy the expectation of T_I is $ET_I = \phi \frac{\Sigma\, \Theta(P)\, N_{cell}}{\Sigma N_{cell}}$, in which ϕ is the high energy hit density mentioned before, the $\frac{\Sigma\, \Theta(P)\, N_{cell}}{\Sigma N_{cell}}$ is the weighted average of the cell count N_{cell}, for each

standard cell, and $\Theta(P)$ is computed from the cell layout by Eq. (4.10). However, the actual error rate is low compared with the simulation time scale ($\sim 10^{-3} Hz$, considerd with the calculations in Eq. 4.1). To accelerate the simulation, the T_I is shrunk by a factor $K = 10^8$, which makes the error rate as $\sim 10^5 Hz$. The duration of SEE pulse T_D is produced from the particle energy with $T_D = E_p \cdot f_{LET}(P \cdot E) \cdot f_{pw}(LET, fan\text{out})$, where E_p is the random particle type and energy produced based on table 4.2, $f_{LET}(P \cdot E)$ is the LET of the particle energy and $f_{LET,fan-pw}$ is the function shown in Figure 4.21. The $f_{LET}(P \cdot E)$ and $f_{pw}(LET, fan\text{out})$ is implemented by lookup table and interpolation. The generation of the injection point L_I is based on the netlist of the circuit, and the possibility of each cell is $P_{I-cell} = \frac{S_{cell}}{S_{module}}$.

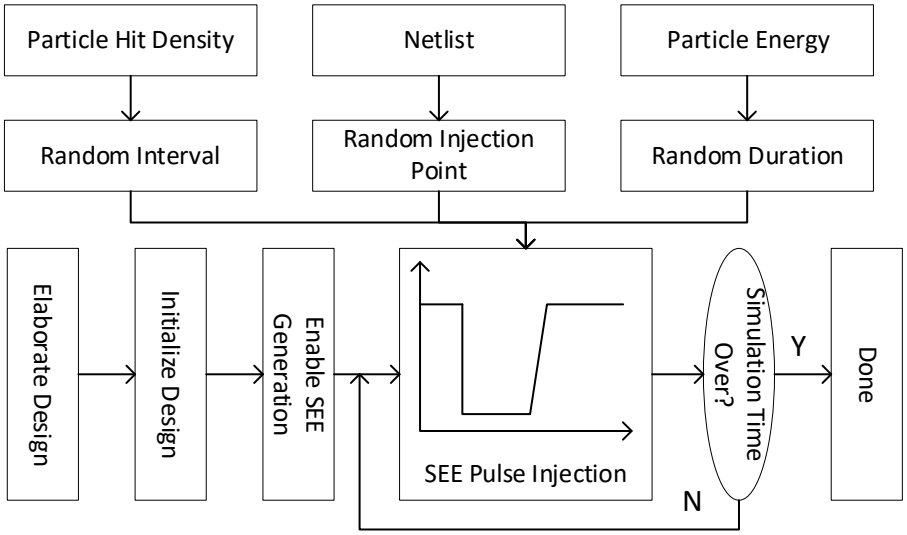


Figure 4.22. The simulation flow with SEE pulse generation.

As shown in Figure 4.23, pseudo-random numbers are formed through algorithms based on random number seeds. The sequence of random number generation is fixed for the specified seed, and the random one, key to repeat the same simulation conditions, is specified before the simulation to produce a certain list of pulses by the SEE generator. A test pattern recorded fired pixels is fed to the input of the pixel matrix emulator which imitates the pixel array and feed data to the readout system in the

simulation, and the deserializer decodes the serial data out of the readout system and record in an output data file. Another data file is generated using the same system, but will disable SEE generation, which makes the reference version. The data files are compared frame by frame to find out the differences including bit errors, data lost, data redundancy and out of synchronization. The evaluation will be performed multiple times get enough statistics for the readout system. The resulting database records the random seed, the data errors, and the SEE generation list to help identify the error position and perform again the simulation.

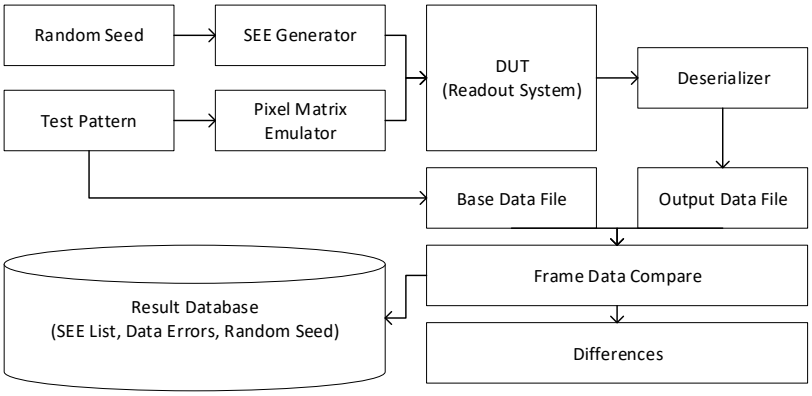


Figure 4.23. The evaluation data path of the readout system.

4.6 TMR Optimization

With verification, the design parameters mentioned in the design part can be optimized and Figure 4.24 shows the TMR optimization iterations in the design and verification flow. After the behaviour level iteration, the module is synthesized block by block with TMR deployment and will be checked by functional and reliability simulations. In the SEE injection simulation, the module in verification is described in RTL while others are in functional. If the results are not satisfying, the module will be re-designed under the guidance of the results. Based on the theory of

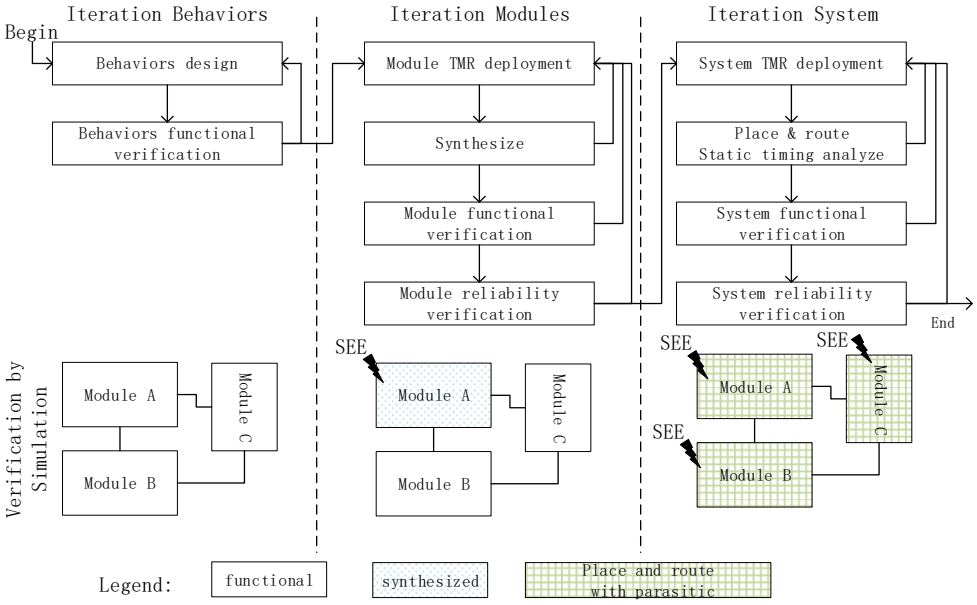


Figure 4.24. The design and verification flow with TMR optimization

reliability in Section 3.1, the module's reliability should be 1-2 orders of magnitude higher than the system requirements. After all the modules are verified, it comes to the system iteration which consider the global clock tree and reset tree, the place and route, as well as the timing analyzation. In the system level verification, the parasitic after place and route is included, in other words, the simulation is running with cell and wire delay. The clock tree is called into being and the simulation will check the reliability of all the components.

After the optimization, the module and the system are verified in a long duration simulation ($100s$, $20M$ frames), the results of which are shown in Table 4-3. Since the SEE generation is accelerated by $K = 10^8$, the simulation is produced up to 10^7 SEE pulses. The area is raised by 15.85%, which is acceptable, and the MTTF of the circuit is measured as the number of SEEs that occurred before the system has failed. We have found that without radiation hardened techniques, the system is susceptible to SEE, but with hardened ones, the system can still run after 10^7 SEEs, which corresponds to more than 5 years of operation. The error rate is measured by the number of data errors after the SEE has occurred. In the original design, the system is very susceptible to SEE and is prone to malfunction, and the error rate is too high to be robustly quantified measured. Meanwhile, the error rate in the TMR deployment design is quite low, and the equivalent error rate in the working environment is lower than 10^{-9}.

Table 4-3 Comparisons before and after TMR deployment.

		Super Region Readout Unit	Elastic Buffer	Readout System
Area (mm^2)	Before	1.75	6.37	43.47
	After	2.20	7.93	50.56
MTTF (SEEs to failure) (Min/Mid/Max)	Before	3/13/48	5/16/78	12/18/32
	After	$>10^7$	$>10^7$	$>10^7$
Error rate (data errors/SEE) (Min/Mid/Max)	Before	>1000	>1000	>1000
	After	0.015/0.085/0.185	0.048/0.055/0.059	0.050/0.135/0.219

4.7 Summary

In this chapter, the deployment and optimization flow of radiation hardening techniques, mainly the TMR and the ECC, are introduced in the design of readout system in MIMOSIS. The TMR deployment strategy is a low-cost way to design high reliability circuits, while the SEE model and verification way supply us with a guide to the optimal design that meets the requirements. The circuit designed under this optimization is fabricated, and its reliability will be tested under particle irradiation. The result of this chapter indicates that a complete system that is not sensitive to SEE can be

devised and optimized for a specific particle environment through the iteration of TMR design and verification flow.

4.8 Bibliography

[1] J. Schambach, J. Alme, M. Bonor a, P. Giubilato, R. Hannigan, H. Hillemanns, M. Lupi, S. V. Nesbø, A. Rehman, et al., *A Radiation-Tolerant Readout System for the ALICE Inner Tracking System Upgrade*, *2018 IEEE Nucl. Sci. Symp. Med. Imaging Conf. NSS/MIC 2018 - Proc.* **9**, (2018) 8.

[2] H. Hillemanns, I. Aimo, C. Bedda, C. Cavicchioli, A. Collu, P. Giubilato, A. Junique, T. Kugathasan, P. La Rocca, et al., *Radiation Hardness and Detector Performance of New 180nm CMOS MAPS Prototype Test Structures Developed for the Upgrade of the ALICE Inner Tracking System*, in *IEEE Nucl. Sci. Symp. Conf. Rec.*, Institute of Electrical and Electronics Engineers Inc., (2013).

[3] B. Abelev, J. Adam, D. Adamová, M. M. Aggarwal, G. Aglieri Rinella, M. Agnello, A. Agostinelli, N. Agrawal, Z. Ahammed, et al., *Technical design report for the upgrade of the ALICE inner tracking system*, *J. Phys. G Nucl. Part. Phys.* **41**, (2014).

[4] S. Kulis, *Single Event Effects Mitigation with TMRG Tool*, in *J. Instrum.* **12**, IOP Publishing, (2017).

[5] L. Sterpone, N. Battezzati, F. L. Kastensmidt and R. Chipana, *An analytical model of the propagation induced pulse broadening (PIPB) effects on single event transient in flash-based FPGAs*, *IEEE Trans. Nucl. Sci.* **58**, (2011) 2333.

[6] R. Chipana, F. L. Kastensmidt, J. Tonfat and R. Reis, *SET susceptibility estimation of clock tree networks from layout extraction*, *LATW 2012 - 13th IEEE Lat. Am. Test Work.*, (2012).

[7] J. F. Ziegler, M. D. Ziegler and J. P. Biersack, *SRIM - The stopping and range of ions in matter (2010)*, *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms* **268**, North-Holland, (2010) 1818.

[8] P. E. Dodd and L. W. Massengill, *Basic mechanisms and modeling of single-event upset in digital microelectronics*, *IEEE Trans. Nucl. Sci.* **50 III**, (2003) 583.

Chapter 5. The Radiation Harden Design of the Phase-Locked Loop

As described in Section 4.1, the readout system functionality requires a 320MHz clock which is provided by the Phase-Locked Loop (PLL) block. Of course, this PLL is a critical element of the sensor and requires protection against radiation effects. The design flow of PLL is reviewed in this chapter, where three PLL versions have been designed: the first is PLL-0, a functional prototype, the second is PLL-1, which corrects some initial version's errors and improves the layout for integration into MIMOSIS-1, and the third is PLL-1+, which has the same architecture as PLL-1 but implements radiation hardening techniques. The design of the PLL is primarily focused on the balance between performance, area and power rather than the implement of the TMR in Chapter 4. The test results of the three PLLs are presented afterwards.

5.1 Architecture of PLL

The PLL is a closed-loop feedback system that produces an output signal's phase, which is related to the input signal phase. There are several disparate types of PLLs: Analog PLL (APLL), All-Digital PLL (ADPLL), and Charge-Pump PLL (CPPLL), whose common idea is to implement a variable frequency oscillator and a phase detector in the feedback loop. The oscillator forms a periodic signal, and the phase detector compares the phase of the signal with the phase of the periodic input signal and adjusts the oscillator to maintain the phase match. Keeping the input and output

phases in sync also means keeping the input and output frequencies the same. Consequently, in addition to the synchronization signal, the phase-locked loop can also track the input frequency or call into being frequencies by multiplying the input frequency.

5.1.1 Topology

The MIMOSIS needs a low jitter 320MHz main clock, which is not convenient to import from the pad. The reasonable way is to use the input clock source as 40 MHz caused by the external active crystal oscillator and a PLL to produce the 320MHz primary clock with reference to the 40 MHz clock.

As per the deployment environment of MIMOSIS, the area and power consumption are mainly concerned. The ADPLL is area efficient but with large power consumption [1][2], while the APLL is both area and power-efficient but not robust enough for the radiation condition. The CPPLL is robust and straightforward, and can achieve high efficiency in the area and power consumption through design. It has a broad frequency and phase capture range, which maintains the lock state after the radiation. It has a digital output with a full CMOS swing range, which is appropriate for MIMOSIS. Nevertheless, the phase/frequency detector of CPPLL may cause a dead zone problem [3], because the lock time and noise level of CPPLL are higher than that of APLL.

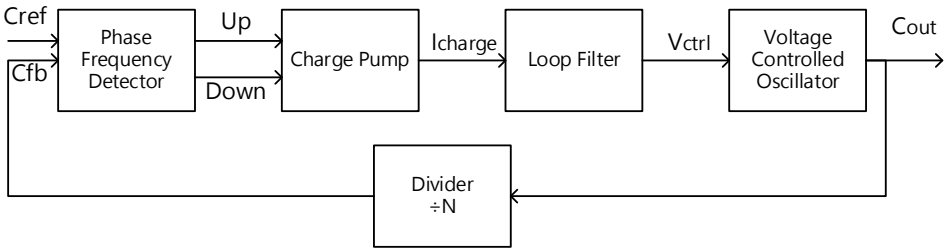


Figure 5.1 The block diagram of a typical CPPLL

As shown in Figure 5.1, the phase/frequency detector (PFD) compares the reference clock (C_{ref}) and the feedback clock (C_{fb}), which outputs the up and $down$ signal. The charge pump produces the charge current pulse with the up signal or the discharge current pulse with the $down$ signal.

The capacitors and resistors in the loop filter convert the current signal into a voltage control signal (V_{ctrl}), and the output clock is generated by a Voltage Controlled Oscillator (VCO). By change of its control voltages (V_{ctrl}), the VCO can change the output clock frequency (f_{Cout}). The ratio between f_{Cout} and V_{ctrl} is defined as $K_{VCO} = \Delta f_{Cout}/\Delta V_{ctrl}$.With an N-divider, the frequency of the feedback clock is $1/N$ of the output clock. Since the C_{fb} has the same frequency as the C_{ref}, the $f_{Cout} = N f_{Cref}$.

5.1.2 Noise Model

The phase-domain model of CPPLL is shown in Figure 5.2, and the forward gain of the CPPLL is

$$G_{fwd}(s) = \frac{K_{det}K_{vco}H(s)}{s} \tag{5.1}$$

In the formula, $K_{det}$is the max current of the charge pump (I_{cp}), and $H(s)$ is the transfer function of the loop filter.

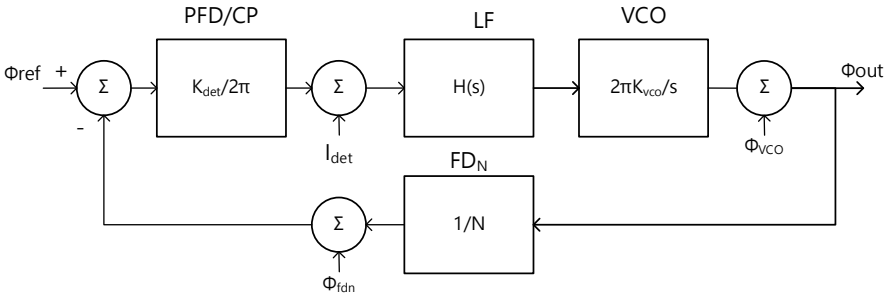


Figure 5.2 The phase-domain model of CPPLL.

And with the feedback path, the open-loop gain transfer function is

$$G_O(s) = \frac{K_{det}K_{vco}H(s)}{Ns} \tag{5.2}$$

Then the transfer functions from the various noise sources to the output are

$$G_{\text{ref}} = \frac{G_{\text{fwd}}}{1+G_O} = \frac{NG_{\text{fwd}}}{N+G_{\text{fwd}}} \tag{5.3}$$

$$G_{\text{vco}} = \frac{1}{1+G_O} = \frac{N}{N+G_{\text{fwd}}} \tag{5.4}$$

$$G_{\text{fdn}} = \frac{\phi_{\text{out}}}{\phi_{\text{fdn}}} = -G_{\text{ref}} \tag{5.5}$$

$$G_{\mathrm{det}} = \frac{\phi_{\mathrm{out}}}{i_{\mathrm{det}}} = \frac{2\pi G_{\mathrm{ref}}}{K_{\mathrm{det}}}. \tag{5.6}$$

These transfer functions allow specific overall characteristics of phase noise in PLLs to be identified. When $s \to \infty$, $G_{\mathrm{fwd}} \to 0$ because of the VCO and the low-pass loop filter, $G_{\mathrm{ref}}, G_{\mathrm{det}}, G_{\mathrm{fdn}} \to 0$ and $G_{\mathrm{vco}} \to 1$. At high frequencies, the noise of the PLL mainly comes from the VCO because the low-pass LF prevents any feedback at high frequencies. When $s \to 0$, $G_{\mathrm{fwd}} \to \infty$ because of the $1/s$ term from the VCO. As a result, at DC, $G_{\mathrm{ref}}, G_{\mathrm{det}}, G_{\mathrm{fdn}} \to N$ and $G_{\mathrm{VCO}} \to 1$. On the other hand, at low frequencies, the PLL's noise is contributed by the OSC, PFD/CP, and FDN, and the gain of the loop diminishes the noise from the VCO [4].

5.2 Circuit design

As mentioned in Chapter 1, the working conditions of MIMOSIS impose many restrictions on the design. For the PLL, the main problem is that all the components, including the bandgap reference voltage, bias, and the large capacitors in the loop filter, need to be implemented in the circuit. These limitations make it necessary for circuit design to consider the realization of functions and performance, to consider area, power consumption, and reliability after irradiation.

5.2.1 VCO

The circuit design of a CPPLL depends on the performance of the VCO, the main feature of which is the frequency gain K_{VCO}. A higher K_{VCO} brings about faster VCO response and wider output frequency but a worse output phase noise. For this design, the target frequency is fixed, and a lower K_{VCO} helps to control the phase noise.

As shown in Figure 5.3, a differential ring oscillator structure is picked for the VCO considering the need of low power consumption and area effectiveness. In the single stage of VCO, M2, M3, M6, and M7 basically consist of the delay part. For instance, in A/NB's rising edge, the NB node is driven by M7. Loads of M7 are the C_{db} of $M7$ and M3 as well as the $C_{gb}$of M6 and M3 in the next stage. To keep the node in balance, M2/M3 and M6/M7 have to be of the same size. Since the transistors work in a

linear region when oscillating, the current drive fits $I_d \propto (W/L)^2$, and the main load part $C_{gb} \propto W \cdot L$. Then the delay fits $\tau \propto C_{load}/I_d \propto L^3/W$. M4 and M5 are devised to control the charging current and thus the delay length. In order to control the K_{VCO} in a lower value, the size of M4 and M5 is much smaller than M6/M7. The M1 is applied to control the drain current of the oscillator, which helps control the oscillator's power and adjust the center frequency when considering the patristic parameters in layout.

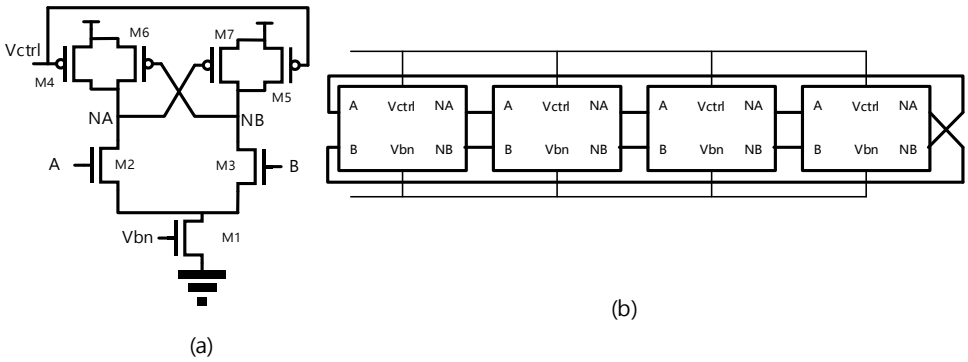


Figure 5.3 The schematic of (a) a single stage of VCO and (b) the loop of 4 stages.

Considering the manufacturing uncertainty, it is necessary to confirm that the target frequency is in the VCO frequency range under each process variation. As shown in Figure 5.4, the V_{tune} for $320MHz$ is between $0.8 - 1V$, which is in the middle of $1.8V$ power supply range. The K_{VCO} is around $250 - 350MHz/V$.

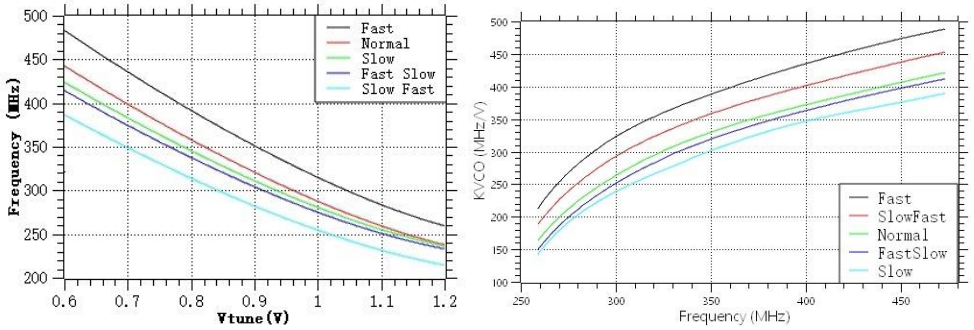


Figure 5.4 Corner simulation of output frequency and K_{VCO} versus V_{tune}.

5.2.2 Charge Pump

The charge pump is essentially a switched current source, which converts the front-stage phase detector's output signal into a current signal. The current signal will produce a voltage on the loop filter capacitor to control the VCO. The charge pump, whose output current accuracy and matching performance have an essential impact on the noise of the PLL, is the core component of the phase-locked loop, and has non-ideal effects such as clock feedthrough, charge sharing of the MOS switch, instability of the current source, and mismatch of the charging and discharging current sources, all of which will affect the noise performance of the phase-locked loop.

The schematic of the charge pump implemented in the PLL-0 is shown in Figure 5.5, where the MP8 and MN8 are the charge and the discharge path of the charge pump. The switch transistors, MP3 and MN3, are switched place with the driver to reduce the clock feedthrough from the switch to I_{out}, and MP7 and MN7 have the same size as MP3 and MN3 to cancel the switch gate's charge sharing effect. The current mirrors in the charge pump are fully cascoded structure to reduce the charge and discharge current mismatch. For the control signals, transmit gates are deployed to match the inverters' delay.

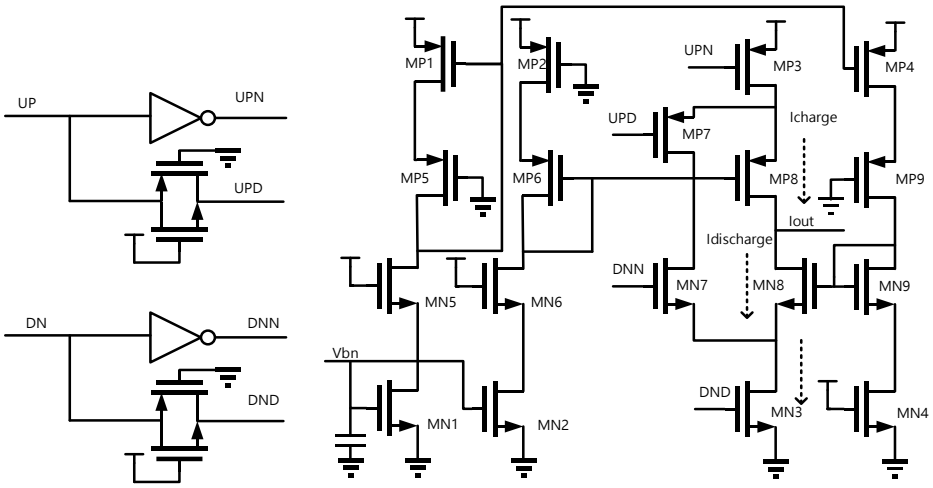


Figure 5.5 The schematic of the charge pump

The simulation result of the mismatch of the charge pump is shown in Figure 5.6. Under disparate process corners and control voltage, the current mismatch is under $400nA$ corresponding to $10\mu A$ charge current, and the

shifting of current mismatch versus V_{tune} is on account of the channel length modulation effect of MP8 and MN8.

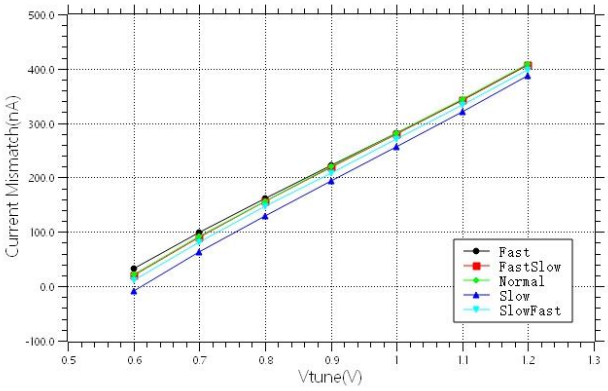


Figure 5.6 Current mismatch of the current pump. $I_{mis} = I_{discharge} - I_{charge}$.

5.2.3 Phase/Frequency Detector (PFD) and Lock State Detector

In the charge pump PLL, the reference clock and the feedback clock compare rising edges through two registers. When both rising edges are sampled, the register is reset by the NAND gate. For instance, as shown in Figure 5.7, if the reference clock is faster than the feedback clock, the $Down$ (DN) signal will open the discharge path and make the VCO enter a higher frequency (PMOS controlled), and vice versa. There are also two AND gates to stop the control signal when reset is low.

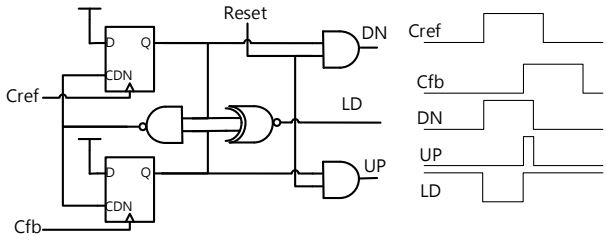


Figure 5.7 The schematic of PFD and its timing diagram.

The PFD registers are not standard cells but optimized dynamic D flip-flops for faster working speed. As the topology and size of transistors shown in Figure 5.8, the locked loop of the register is composed of two end-to-end inverters, one of which is 25 times the size of the other. The clock's rising edge is detected by the inverter (MP1 and MN1) delay. The

inverter produces a pulse current path from MN4 to the ground which will change the lock loop state. This optimized register has a very low delay (~60ps) from the clock rising edge to the output compared with the standard cell (~180ps).

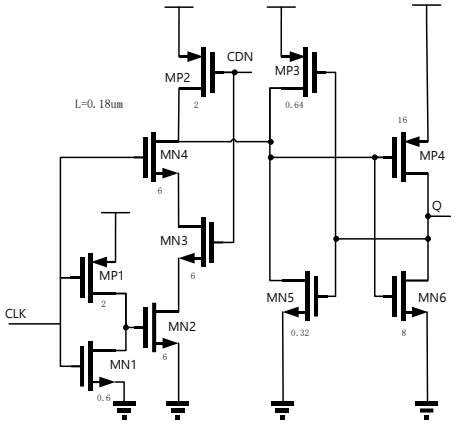


Figure 5.8 The schematic of register optimized for PFD.

A further function of PFD in this design is the lock state detection. In the CPPLL, the phase error is represented by the pulse width difference between UP and DN. With an XNOR gate for the UP and DN, the LD signal can be applied to detect the lock state of the PLL. The schematic shown in Figure 5.9 presents an unbalanced charge pump with a large charge current and small discharge current. If the duty cycle of LD is lower than the ratio of $I_{discharge}/I_{charge}$, the voltage on C1 (V_{C1}) will decline. If the $V_{C1}$is lower than V_{cmp}, the comparator will give a high value for the $Lock$ signal. Similarly, if the LD goes to a high duty cycle, the V_{C1} will rise sharply, which will cause the $Lock$ signal drop, i.e., loss of lock. The lock state detection is devised to turn off the clock's output when the PLL has not reached or loses lock state, thereby preventing subsequent load circuits from losing synchronization.

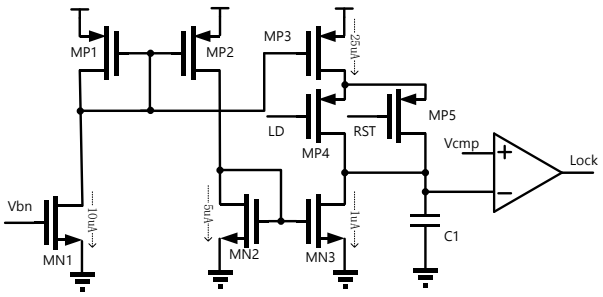


Figure 5.9. The Lock Detector circuit. The check limit of duty cycle of LD is fixed as 4%.

5.2.4 Loop Parameter

The loop filter's function is to convert the charge pump's current signal into a voltage signal and control the frequency of the voltage-controlled oscillator's output signal. Besides, the loop filter constructs the PLL system's dynamic performance, thus determining the system's stability and locks time. A typical second order low pass loop filter for CPPLL is shown in Figure 5.10, whose transfer function is

$$H(s) = \frac{RC_2 s+1}{RC_1C_2 s+C_1 s+C_2 s}. \tag{5.7}$$

Based on Eq. 5.2, the close loop gain transfer function is

$$H_C(s) = \frac{NH_O(s)}{1+H_o(s)} = \frac{\omega_n^2\left(1+\frac{s}{\omega_z}\right)}{\left(s^2+2s\zeta\omega_n+\omega_n^2\right)} \tag{5.8}$$

In the formula, ω_n is the undamped natural frequency in rad/s, ω_z is the stabilizing zero points at $1/RC_2$ (rad/s), and ζ is the damping factor. The second order CPPLL has two pole point (P_1, P_2) and a zero point (Z). If $\zeta < 1$, the complex poles result in damped oscillation before lock, and if $\zeta > 1$, Z and P_1 are canceled, the PLL will approach the lock state directly. The undamped natural frequency (ω_{n}), phase margin (Φ_m) and damping factor (ζ) are computed by

$$\left\| \frac{K_{vco}\cdot I_{cp}}{N\cdot s}\cdot H(s) \right\|_{s=\omega_{\mathrm{n}}} = 1, \tag{5.9}$$

$$\Phi_M = 180 + \mathrm{atan}(\omega_n RC_2) - \mathrm{atan}(\omega_n RC_1), \tag{5.10}$$

The value of C1, C2, and R can be solved with a given ω_{n}, Φ_M and ζ from the following equations:

$$\begin{aligned} T_1 &= \frac{\sec(\Phi_M)-\tan(\Phi_M)}{\omega_{\mathrm{n}}} \\ T_2 &= \frac{\zeta}{\omega_{\mathrm{n}}^2 T_1} \end{aligned} \tag{5.11}$$

$$C_1 = \frac{K_{vco}I_{cp}}{\omega_{\mathrm{n}}^2 N}\frac{T_1}{T_2}\sqrt{\frac{1+\omega_{\mathrm{n}}^2T_2^2}{1+\omega_{\mathrm{n}}^2T_1^2}} \tag{5.12}$$

$$C_2 = C_1(\frac{T_2}{T1} - 1) \tag{5.13}$$

$$R = \frac{T_2}{C_2} \tag{5.14}$$

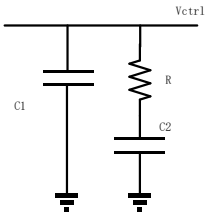


Figure 5.10. The second-order loop filter for CPPLL.

To maintain the stability of the CPPLL, the undamped natural frequency ω_n should fit $\omega_n < F_{ref}/20$, and a lower ω_n brings about a lower -3 dB loop bandwidth, which means it has better control with the low-frequency noise, i.e., lower jitter. Meanwhile, the T_2 increases a lot and so does the value of C_2. The phase margin Φ_m needs to be at least $45°$ to keep the loop stable because lower phase margin is appropriate for slightly better spurs, especially for higher-order filters. In this design, $\omega_n = 1MHz$, $\Phi_M = 60°, \zeta = 0.7, K_{VCO} = 350MHz/V$ and $I_{CP} = 10uA$. Then, we can get $C_1 = 3.07pF$, $C_2 = 26.87pF$, and $R = 15.47K\Omega$.

5.2.5 Divider

The clock divider in the PLL is shown in Figure 5.11, and its structure is similar to the design in Figure 4.12 but with disparate voters. The voters in this divider is specially devised to decline the possibility of SET in the voters, and the capacitor of node $outb$ is larger than that in the standard cells.

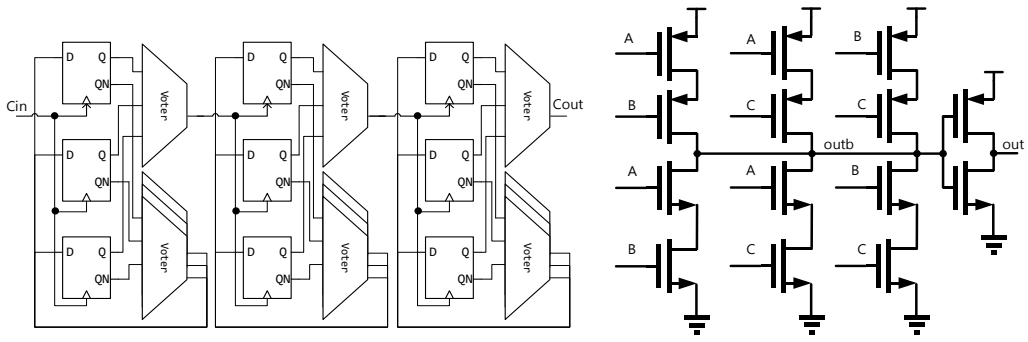


Figure 5.11 Frequency divider and voter

5.2.6 Layout

As shown in Figure 5.12, the layout of the CPPLL is compact and all integrated. The dimension of the core region is $650 \times 170\ um^2$, and only the core region is needed to be implemented in MIMOSIS sensor.

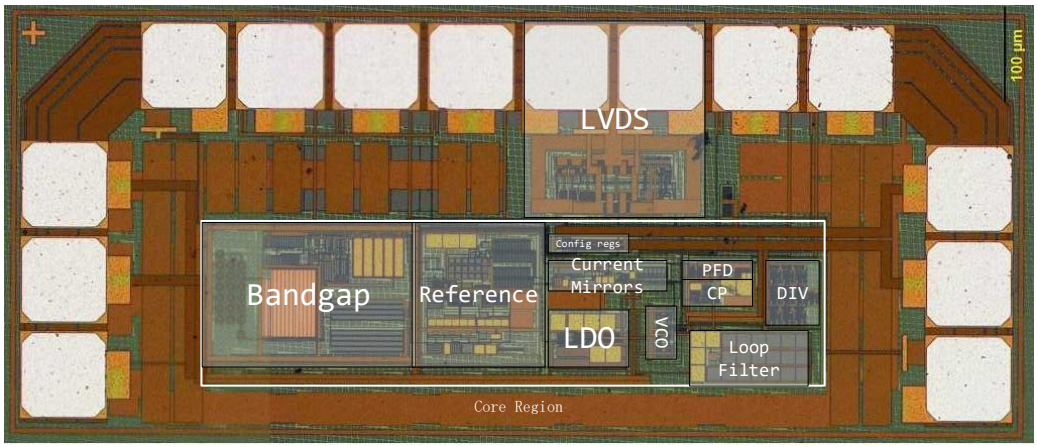


Figure 5.12 Layout of PLL-0.

5.3 Corrections and Improvements

After testing, PLL-0 can work as devised in advance, but there are some problems, which are corrected in the new version, and new radiation harden designs are then introduced.

5.3.1 The deep N-well connection

The first problem is that there is soft connection between disparate N-well contact. As presented in Chapter 1, the MIMOSIS uses the modified process with a deep N-well, leading to a short path from an N+ doping region (well contact NW1 in Figure 5.13) through N-well and deep N-well to another N+ doping region (well contact NW2). If the deep N-well is separated, the connection is broken (i.e., well contact NW3). In the PLL-0, the loop filter uses the N-well PMOS capacitor to gain a larger capacitance. The negative port of the capacitor, i.e., the N-well contact, is connected to the ground while the other is connected to the power. After FIB (Focused Ion Beam) editing, the negative port of the capacitor is disconnected with ground but softly connected with the power, and the soft connection

reverses the capacitor in the loop filter, causing the MOS capacitor's value to become 1/3 of the original. In PLL-1, this error is corrected by separated every block with a distance.

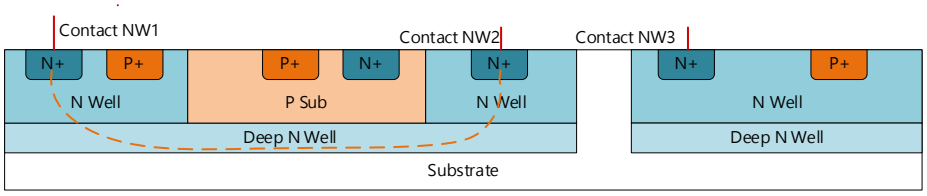


Figure 5.13. The two N-well contact (NW1 and NW2) soft connected by a deep N-well. Separate the deep N-well to avoid.

5.3.2 New lock detection structure

The lock-detect circuit in Figure 5.9 is not well functional when the input clock or power supply is not ideal, and there are two main reasons for this problem. On the one hand, the ratio of charging and discharging as well as the comparator's input are fixed, making the circuit incompatible with more noise. On the other hand, on account of the dead zone problem of the PFD, the LD signal in Figure 5.9 does not always remain at 0. There will be some short pulses even after the PLL is locked. The voltage pulses will cause the MP4 to have a charge sharing phenomenon, thereby making the voltage of $C1$ unstable.

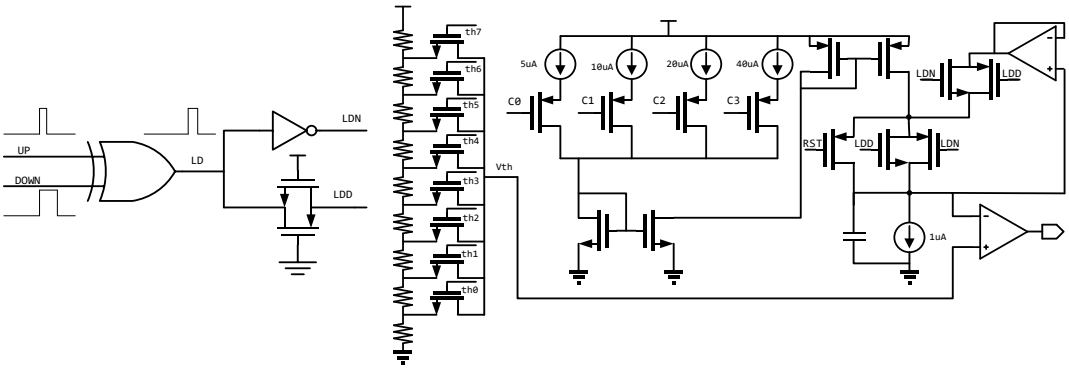


Figure 5.14 The new lock detection circuit schematic.

The structure shown in Figure 5.14 is deployed in PLL-1 so as to improve the lock detection circuit's stability. The compared threshold voltage (V_{th}) is selected by enabling (th7 ~ th0), and C3 ~ C0 controls the charge current from 75mA to 0mA (disable). With intention to decrease

the LD switch's charge sharing, complementary switches are applied to enhance the conductivity and a buffer to keep the voltage across the switches equal. The additional current mirrors between the current source and the switch also decrease the switches' capacitor load.

5.3.3 Dual Modular Redundancy on PFD

In the PLL-0, the digital elements are PFD, divider, and VCO, among which the divider has been devised as TMR hardened, the VCO uses a differential structure that is not sensitive to SEE, and the PFD is additionally hardened in the PLL-1+.

The new PFD registers are hardened with Dual Modular Redundancy (DMR), similar to TMR, but it does not require a voter but a reset. As shown in Figure 5.15, the register for reference and feedback clock is doubled. If SEU reverses one of these registers to 1, the XNOR gate (Xnor1 and Xnor2) will call into being the reset through AND gate (And5). Moreover, the AND gate filters the SEU, so it is safe to keep the DN and UP at a low value while PLL goes to the locking state. The low-value cutoff of the charge and discharge current in the current pump brings about the fact that the PLL loses the correction of only the current clock period and prolonged lock time. The DMR structure prevents the DN or UP from going to a high value because of the SEU in registers.

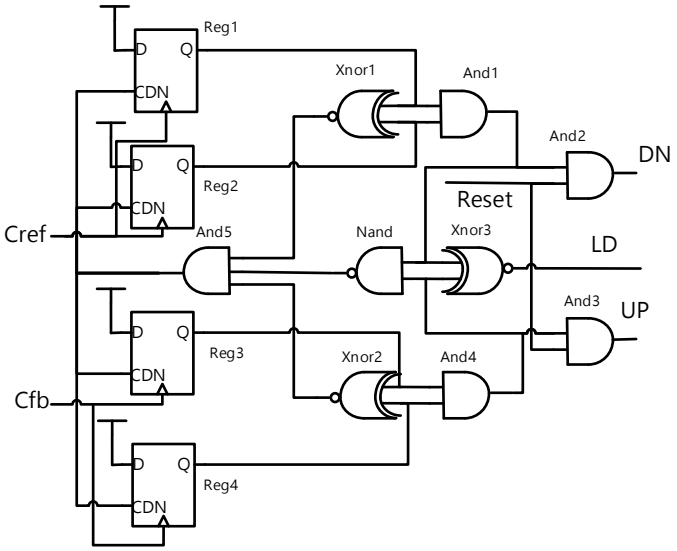


Figure 5.15. The schematic of PFD with DMR implemented.

5.3.4 Charge Pump balance

When the phase-locked loop is locked, the output control signal is a control signal with the same width, which means that two current sources will be opened together for a period of time. This period requires the two currents of the charge pump to be highly matched; otherwise, it will cause many jitters. The change in output voltage (V_{tune}) makes the channel length modulation effect of current sources MP8 and MN8 noticeable (4%).

In the new design, the dashed frame circuit introduces a negative feedback structure to decrease the channel length modulation effect when the output voltage changes, and when the charge pump is charged, the output voltage will augment. Owing to the channel length modulation effect, the output current will decline. Nevertheless, for MP11, the increment of the output voltage will decrease the branch's current, so the gate of the MP8 will reduce to make the charging current larger, and the sum current remains a fixed value. The MP8, MP10, and MP11 form negative feedback help to keep the charging current constant.

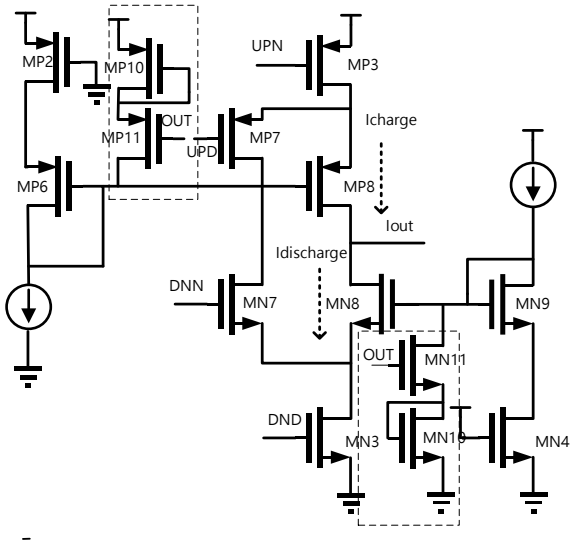


Figure 5.16. The additional feedback path of the charge pump. The dashed frame circuit helps offset the channel length modulation effect, resulting in better current matching.

The simulation results are shown in Figure 5.17, where the mismatch between charge and discharge current is less than $170nA/10uA$ under disparate corners, much smaller than the results in Figure 5.6. The other

method may achieve a lower current mismatch but with a complicated structure. Moreover, less than a 2% error in this design has made the charge pump no longer a performance bottleneck.

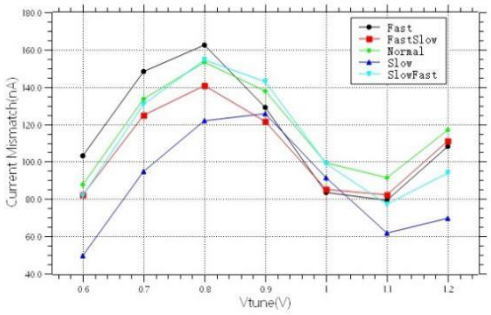


Figure 5.17. The simulation *result of the current mismatch.*

5.3.5 Re-entrant gate layout for bias

In Section 3.5, the re-entrant gate layout (RGL) is introduced as a fringeless field-effect transistor that can mitigate the TID effects in bulk ICs, but with several obvious drawbacks. First of all, it is hard to compute the equivalent W/L ratio for the FET with RGL accurately, and Eq. (3.21) gives the right approach. Nevertheless, like Calibre, the EDA verification tools have their approach to identify MOSFETs, which conflicts with the calculation before, thereby bringing many difficulties to the verification work. On the other hand, the efficient W is at least four times of L, as per Eq. (3.21), which is feasible for digital circuit design, but very restrictive for analog circuit design.

Considering the limitations of RGL, in the analog circuit, only the NMOS of the bias circuit uses RGL (the NMOS in the dashed frame in Figure 5.18). Owing to the size limitation of RGL, MN1, MN2, and MN3 are in the conventional NMOS layout. Since MN1 is in a closed control loop with an amplifier, the threshold drift has little effect, while MN2 and MN3 have a large L but a small W, and the current mirror is in a cascode state, so these two FETs can also be in a conventional layout.

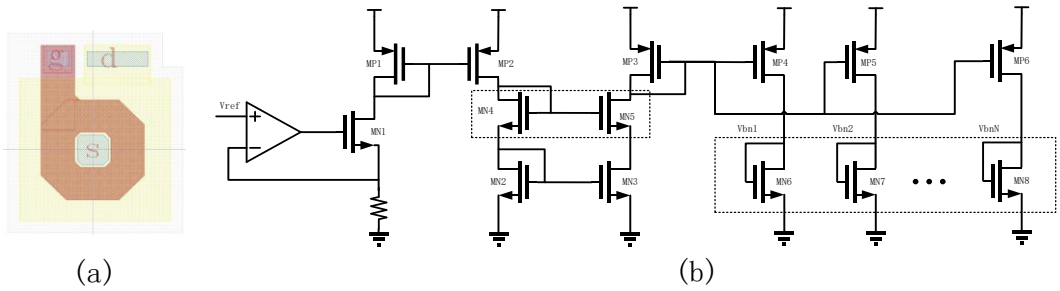


Figure 5.18 The (a) NMOS in re-entrant gate layout and the PLL's schematic of the bias generation circuit.

5.3.6 Compact layout

When MIMOSIS-1 was manufactured, the MIM (Metal-Insulator-Metal) capacitor was removed because of process changes that removed the CTM (Capacitor Top Metal) layer of the MIM capacitor. A MIM capacitor, generally composed of the top two metal layers and a particular CTM metal layer in the middle, is equivalent to a parallel plate capacitor. The dielectric layer in the middle of the CTM and bottom layer metal is relatively thin, and the formed capacitance density is high. As it is situated on the top layer, it has small parasitic and high accuracy. Without MIM capacitors, the amplifiers and the loop filter can only use the MOS capacitors. MOS transistors' gate capacitance can achieve a higher capacitance density, but the capacitance value will vary with the gate voltage, with prominent nonlinearity. Moreover, the MIM capacitor can be placed over the other components to save the area, but the MOS capacitor cannot, which gives rise to a larger area of the circuit.

In this condition, with the design and optimization, the layout of PLL-1 and PLL-1+ is shown in Figure 5.19, in which the core region of PLL-1 is implemented in MIMOSIS-1. The PLL-1 has the same structure as PLL-0 but with new lock detection and RGL bias. The PLL-1+ uses more new structures likes PFD with DMR and a balanced charge pump based on PLL-1 and has the same core size. The core size is $310 \times 350um^2$, which is smaller than PLL-0 even with more circuits and without MIM capacitors.

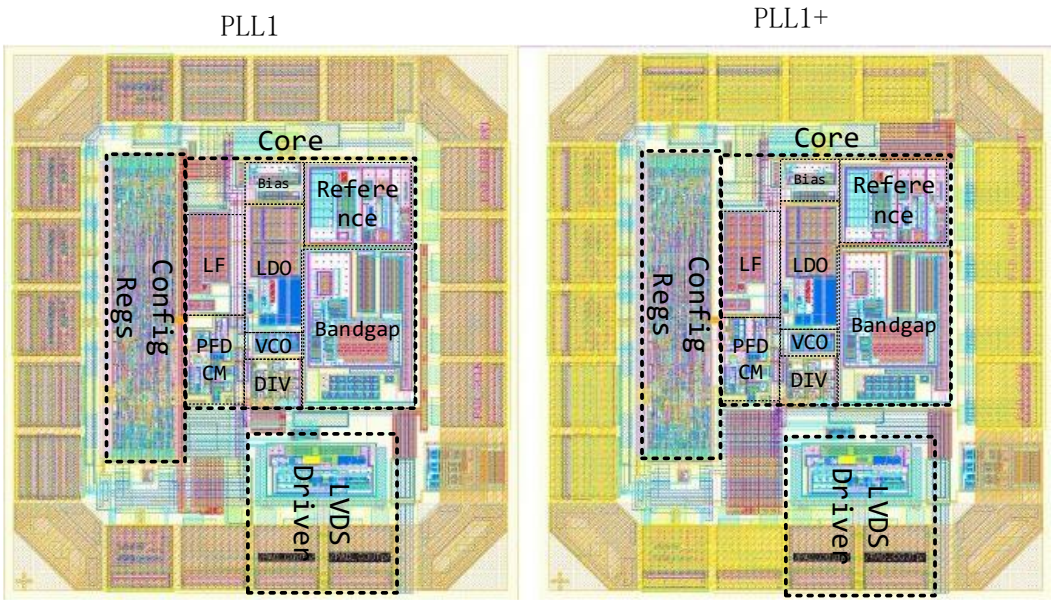


Figure 5.19 the layout of PLL-1 and PLL-1+. The core region of PLL-1 is implemented in MIMOSIS-1.

5.4 Test method and results

After completing the chip design, tape-out, and packaging, the PLL's working status and performance level have been systematically tested in a laboratory environment, and the test content basically contains performance tests such as input clock and power supply range, power consumption, output clock noise, and lock time.

Because of the soft connection mentioned in Section 5.3.1 of PLL-0, the circuit needs the FIB edition to cut the deep N-well connection to the ground. Thus, only two chips were tested. In the second submission, there are PLL-1 and PLL-1+ on the same wafer with MIMOSIS-1. 3 PLL-1 and 1 PLL-1+ chip have been tested.

5.4.1 Test board, system

The block diagram of the PLL test board and test system is shown in Figure 5.20. Since the PLL integrates all components, the test board contains a power regulator, clock input, and level conversion. The PLL's output clock directly outputs the LVDS signal to the oscilloscope differential probe through the SMA interface. The bias current of the

LVDS driver can be tuned externally by a potentiometer, and the LVDS clock input introduces the clock input. Through the crossbar switch, it is possible to switch between single-line input and LVDS input.

Outside the board, the clock is produced by a SI5344 clock generation board, and the typical RMS jitter of its clock output is $90fs$. The CP2112 USB-I^2C converter board is applied to config the PLL through I^2C bus. The oscilloscope is a LeCroy 640Zi with active differential amplifier probe D620, which has 6GHz bandwidth.

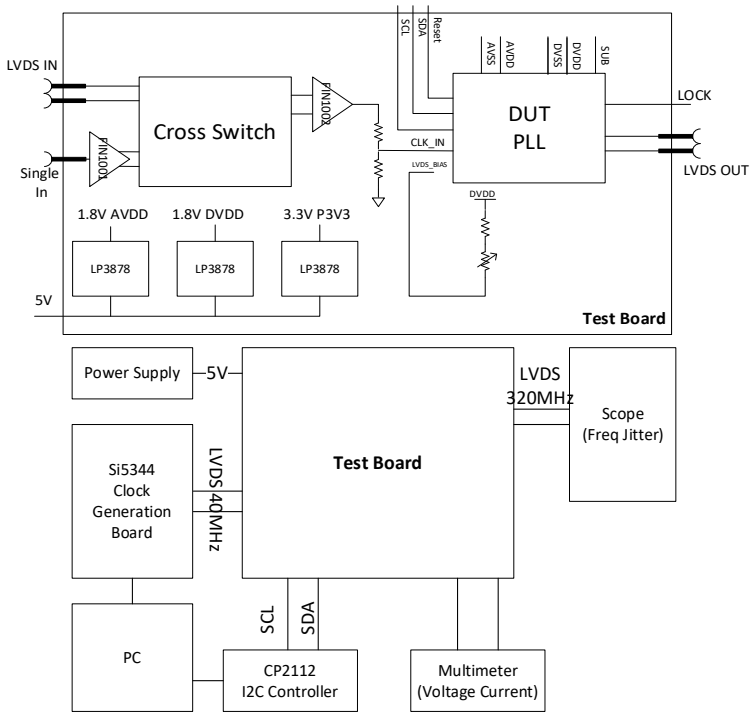


Figure 5.20. The test board and test system of PLL.

5.4.2 Static test results

The typical power consumption, the input frequency, and power range test results are shown in Table 5-1. The typical power is tested under the condition $V_{in} = 1.8V, F_{in} = 40MHz$ and LVDS driver off. The power consumption of PLL1 and PLL1+ is higher, basically because PLL-0 uses a shift register as a configuration, while PLL-1 uses an I2C controller and deploys TMR, which consumes more current. The PLL-1+ has slightly higher power than the PLL-1, which is caused by the DMR on PFD.

Table 5-1 The static test results of PLL

	Power (mW)	Input Frequency (Min-Max) ($MHz\ V_{in} = 1.8V$)		Power Supply (Min-Max) ($V, F_{in} = 40MHz$)	
PLL-0	4.75	18	70	<1.4	>2.0
PLL-1	5.31	25	80	<1.4	>2.0
PLL-1+	5.35	25	80	<1.4	>2.0

5.4.3 Jitter performance

Jitter is the characterization of the clock signal's random change in the time domain, while the phase noise is the representation in the frequency domain, both of which represent the accuracy of the clock signal. Unlike most analog signals such as voltage or current, the most crucial feature of a clock signal is its threshold transition time value, and the jitter is the random variation of these transitions in time value. As the jitter in timing domain, phase noise characterizes the energy of undesired frequency components in the clock signal. Thermal and other noise processes cause jitter and phase noise in the circuit.

Absolute jitter (J_{ab}), also known as the edge to edge jitter, is the absolute jitter of a clock signal, which refers to the discrete sequence of time that its transition moment deviates from its ideal value.

$$J_{ab}(k) = t_k - t_{ideal} = kT + \tau_k - kT = \tau_k \tag{5.15}$$

Period jitter (J_p) refers to sampling each cycle of the clock signal and calculating the offset from the ideal signal cycle. Period jitter is the first order difference from absolute jitter.

$$J_p(k) = t_{k+1} - t_k - T = \tau_{k+1} - \tau_k \tag{5.16}$$

Under most conditions, it is hard to get an ideal signal, so it is impossible to determine the ideal signal's edge moment, and the absolute jitter cannot be computed. But, for period jitter, after sampling enough periods, the average value of the sampling period can be applied to replace the period of the ideal signal for calculation.

As a form of random noise, period jitter usually obeys Gaussian distribution. Therefore, the RMS value of jitter is usually applied to characterize the accuracy of the clock signal.

$$J_p\ RMS = \sqrt{\frac{1}{\mathrm{N}}\sum_{k}^{k\leq N} J_p(k)^2} \tag{5.17}$$

Peak-to-peak period jitter ($J_p\ P-P$) is the difference between the maximum clock period and the minimum clock period of all single clock periods in the observation window, with a usual value of 1,000 or 10,000 cycles. This feature is useful for ensuring the setup and hold time of flip-flops in digital systems.

A crest factor is computed for peak-to-peak jitter given an RMS jitter value and a tolerable Bit Error Rate (BER). Once computed, the resulting crest factor is applied to convert the RMS jitter value and peak-to-peak value.

In the measurement of jitter, the dual-Dirac jitter model applied in "MJSQ" (Methodologies for Jitter and Signal Quality Specification)[6] takes the total jitter (T_j) of a clock as the sum of with random jitter (R_j) and deterministic jitter (D_j). The T_j is corresponding to the J_p described above with R_j weighted by a multiplier α that is determined from the bit error ratio.

$$\mathrm{T_j} = \alpha(\mathrm{BER}) * \mathrm{R_j} + \mathrm{D_j}(\delta\delta) \tag{5.18}$$

In the formula, α = 14.07 for a typical BER value of 10^{-12}, and $\mathrm{D_j}(\delta\delta)$ is the "model-dependent" deterministic jitter, which is determined from fitting to the dual-Dirac model.

Figure 5.21 shows the jitter over a range of input-clock frequency test results of PLL-0, PLL-1, and PLL-1+. The total jitter (BER = 10^{-12}) is lower than $0.04UI$ when the input clock is around 40MHz, and when the input clock rises, the total jitter increases considerably, an effect caused mainly by two reasons. The first is that the higher input clock makes the VCO's working area offset, which results in the loop filter not being able to remove in-band noise well and is reflected in the R_j contribution. On the other hand, because the working bandwidth of the LVDS driver is only

about $400MHz$, when the frequency rises further, it will cause waveform distortion and introduce D_j contribution.

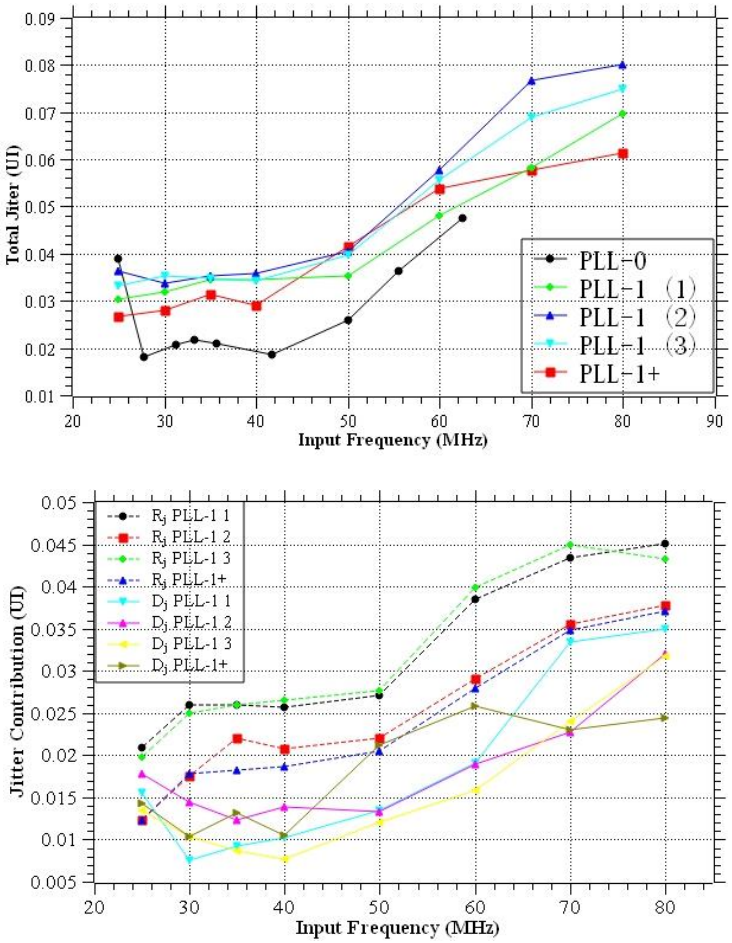


Figure 5.21. The total jitter and the jitter contribution of R_j and D_j over T_j. The value of R_j is multiplied with factor $\alpha(BER) = 14.07$.

5.4.4 Dynamic test results

The dynamic test of PLL contains the lock time and the jitter performance with large power loads, and the lock time result of the various conditions is shown in Table 5-2, where the PLL can lock in $15\mu s$ in less than all conditions. The fast locking feature is on behalf of the high loop bandwidth and the appropriate damping factor. Moreover, with the help of the reset circuit, the lock time after the reset of PLL is $1.5\mu s$ (~60 input clock period), which is very fast for a CPPLL. The reset circuit after the charge pump will pre-charge the capacitor in the loop filter to make the VCO run at $\sim 300MHz$ after the reset. If the bandgap is also reset, the lock

time will be much longer because the feedback path of the bandgap needs at least $7\mu s$ to keep balance. If the clock input is reconnected at the PLL in the working state, the lock time is less than $15\ \mu s$. The reason for the longer lock time is because when the PLL has no clock input, the loop will try to match at a shallow frequency ($\sim 160MHz$), and keep the capacitor voltage in the loop filter away from the locked state, which will take a long time to recharge.

Table 5-2 The lock time test results of PLL

	Lock time after reset PLL	Lock time after reset Bandgap	Lock time from the clock input
PLL-1	$\sim 1.5\mu s$	$\sim 8.5\mu s$	$< 15\mu s$

The jitter performance with large power loads (power noise) is shown in Table 5-3. The test is on the MIMOSIS-1 chip instead of the prototype chip, and there are three levels of activities of pixel matrix and sequencer. In level 0, only PLL is working, i.e., the best performance. In level 1, the readout system is active while the pixel matrix is disabled. In level-2, based on level-1, a global reset is made for the pixel matrix. In level 3, a overall pixel flushing signal is sent periodically between the two resets. In order to compare the PLL performance, a clock generation board SI5344EB by Silicon Labs is utilized to replace the PLL clock and works in condition level 3.

Table 5-3 The jitter performance of PLL with activities

	Level-0	Level-1	Level-2	Level-3	Comparison
$T_j(ps)$	154 (5%)	355(11%)	559(18%)	797(26%)	242(7%)
$R_j(ps)$	9.08	22.2	37.4	55.9	12.8

The jitter performance during heavy activity is not ideal because the low-dropout (LDO) regulator for VCO does not work correctly. The capacitor in the LDO, which maintains the feedback loop's stability, is not well devised with the MOS capacitor, resulting in the oscillation of the control loop. The correction will be carried out in future designs.

Another way to correct the LDO stability problem is to give a higher bias current. With an external bias tuner, the LDO's performance is much better than the test result before, as shown in Figure 5.22. With a higher bias current, the LDO feedback loop's bandwidth is more extensive, which results in better loop stability. With 50% more power consumption, the jitter performance in activity level-3 is comparable to the external clock.

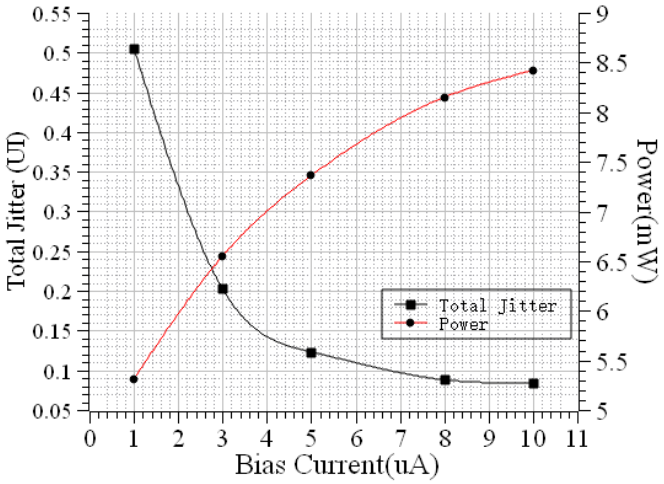


Figure 5.22. The performance of the PLL-1 with external bias. In power consumption, the configuration circuit belongs to the core part, also take into account.

5.5 The improvements

The test result of the MIMOSIS-1 shows that there are still area and power consumption budget available on the chip for the PLL. As presented in the last section, the PLL does not have a good performance in the heavy-activity test, and the main problem is that the LDO is not correctly working as designed. A new schematic of the LDO is devised for the VCO as shown in Figure 5.23, where the LDO has two feedback loops. The first loop is devised with 2 stage amplifiers to call into being the output voltage based on the resistors' bandgap reference, and the second loop's primary purpose is to maintain stability with a low current consumption and capacitance load (the VCO). The first loop basically produces the power supply rejection (PSR) gain. The PSR of the LDO is devised as $80dB@10KHz$ and $40dB@1MHz$.

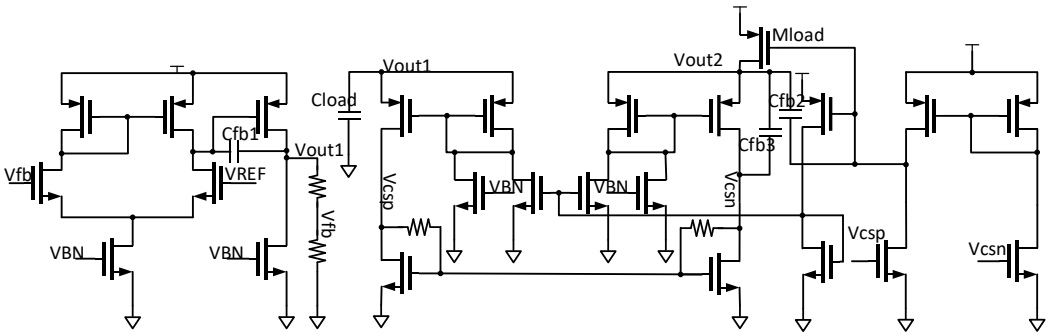


Figure 5.23 The schematic of LDO. $V_{out2} = V_{out1} = 1.3 \times V_{REF} = 1.5V$.

The feedback capacitor of the first loop (C_{fb1}) is replaced with a 5-layer metal-finger capacitor shown in Figure 5.24, which is much stable but four times larger than the MOS capacitor. The other capacitors ($C_{fb2}, C_{fb3}\ and\ C_{load}$) are implemented with MOS capacitors.

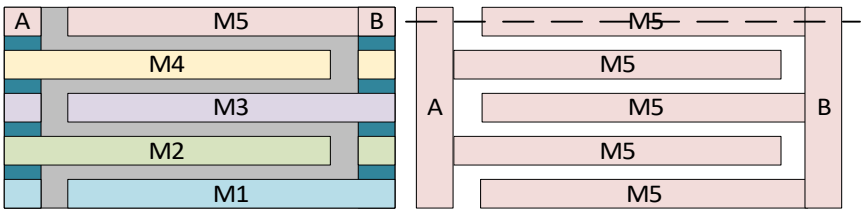


Figure 5.24. The Cross-section and the top view of the M5-M1 metal finger capacitor. The A and B terminals have the capacitance in the same layer or the adjacent layers.

With a larger area, the capacitors in the loop filter can be larger than before. Hence, the bandwidth of the control loop can be lower. A new structure is devised for the next version. Based on the PLL-1+, the architecture of PLL-2 is shown as Figure 5.25. When the area is larger, the loop filter can implement more capacitors so that the frequency division ratio (N) can be augmented, and the input clock frequency (F_{ref}) as well as power consumption can be reduced. With the LDO, the VCO is devised only for a 1.5V power supply, with which, the gain of the VCO (K_{vco}) is devised much lower ($\approx 100MHz/V\ @320MHz$) than before by deploying a proportion V_{tune} level shifter. All the above modifications have a positive contribution to the jitter performance.

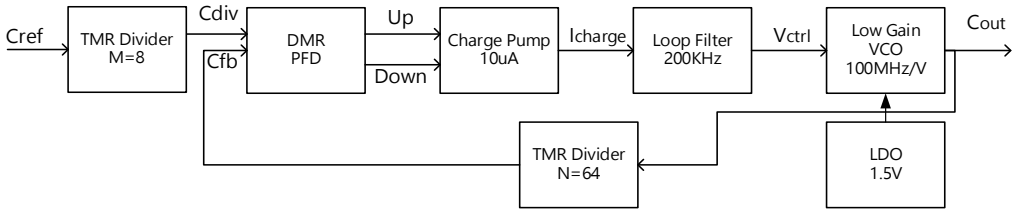


Figure 5.25. The block diagram of PLL-2.

Moreover, a large capacitor can make the loop bandwidth narrower, thereby suppressing noise's influence on the final output clock, and Figure 5.26 indicates the spectrum of phase noise with the clock output of PLL-1. The total jitter of the clock based on the phase noise is computed with the equation [6]:

$$T_j = \frac{\sqrt{2\int_0^{2F_o} PN(f)df}}{2\pi F_o} \text{ (s)} \tag{5.19}$$

In the formula, F_o is the output frequency, and $PN(f)$ is the phase noise power shown in Figure 5.26.

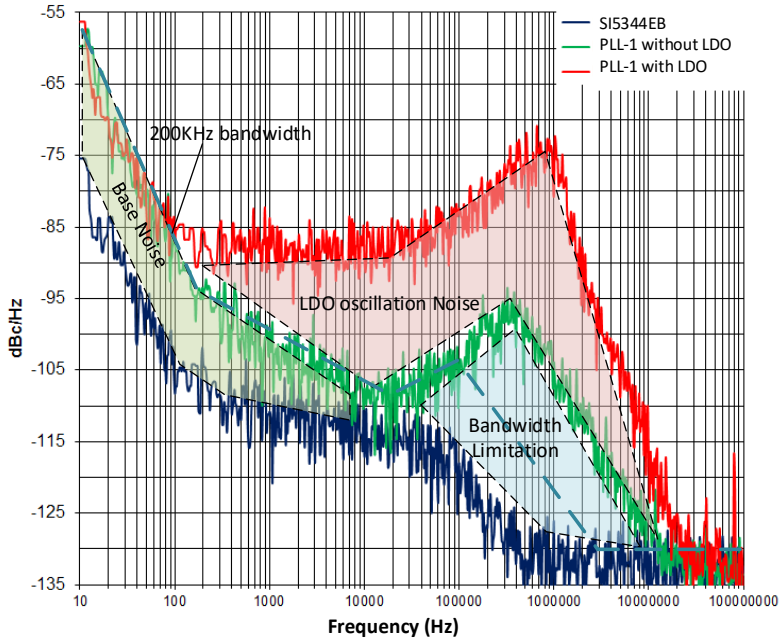


Figure 5.26. the spectrum of phase noise power of PLL-1 with/without LDO compared with a clock generation board SI 5344EB.

If we take the resulting plot without LDO, we can come to the conclusion that phase noise's main contribution is the bandwidth limitation of the PLL, where the bandwidth is $1MHz$, corresponding to the $-3dB$ knee point in the Figure 5.26. If we augment the capacitors in the loop filter to reduce the bandwidth to $200KHz$, the phase noise power will be an order of magnitude lower (the blue dash line).

5.6 Summary

In this chapter, the design and test flow of PLL in MIMOSIS is presented, and the test results indicate that the design meets the needs of MIMOSIS with low power consumption and a small area with full integration. Radiation hardness techniques are pertinent in the PLL. In the PFD and divider, the redundancy is performed against the SEU, and in the charge pump, a full cascode structure and a feedback mechanism without amplifiers are presented to keep the charge and discharge current balance to counter the TID effect's threshold shift effect. Moreover, in the bias circuit, the RGL is utilized for NMOS to reduce the TID effect's charge collection.

Bibliography

[1] X. Chen, J. Yang and L. X. Shi, *A fast locking all-digital phase-locked loop via feed-forward compensation technique, IEEE Trans. Very Large Scale Integr. Syst.* **19**, (2011) 857.

[2] P. L. Chen, C. C. Chung, J. N. Yang and C. Y. Lee, *A clock generator with cascaded dynamic frequency counting loops for wide multiplication range applications, IEEE J. Solid-State Circuits* **41**, (2006) 1275.

[3] L. Zhang, Z. Wang, T. Zhang and X. Peng, *An Improved Fast Acquisition Phase Frequency Detector for High Speed Phase-Locked Loops*, in *AIP Conf. Proc.* **1955**, (2018).

[4] K. S. Kundert, *Predicting the Phase Noise and Jitter of PLL-Based Frequency Synthesizers*, in *Phase-Locking High-Performance Syst. From Devices to Archit.*, (2003).

[5] W. Corp, *Fiber Channel - Methodologies for Jitter and Signal Quality Specification - MJSQ*, (2001).

[6] W. Kester, *MT-008 TUTORIAL Converting Oscillator Phase Noise to Time Jitter*.

Chapter 6. Preparation for the Radiation-Tolerance Evaluation of MIMOSIS-1

Previous chapters have explained how the frontend readout system and phase lock loop of MIMOSIS-1 are hardened against SEE and TID by design. The MIMOSIS-1, PLL-1/PLL-1+, and an evaluation chip were fabricated on the same wafer during the spring of 2020. Because of the delay in the submission of the MIMOSIS-1 chip and initial testing, there is no time left during the course of this work to demonstrate the success of this specific design experimentally. Nevertheless, the preparation of the planned irradiation tests has been fully completed.

In this chapter, we first explain the experiment measurement method of SEE cross-section $\Theta(LET)$ presented in Eq. (4.10). An evaluation chip is devised for the measurement, which can measure both the SEU error rate of registers and SRAMs and SET pulse width in inverters. The $\Theta(LET)$ is the base assumption for optimization and is made from the simulation results, which will help to guide the design of next-generation or under similar conditions.

Then the testability design for SEE is introduced and the in-time SEE test method is shown. Finally, the TID and SEE beam test is devised for the PLL, which is mainly composed of analog parts.

6.1 SEE Cross-section Calibration

Experimentally, the device cross-section $\Theta(P)$ or $\Theta(LET)$ is computed as the ratio of the number of upsets observed to the particle fluence

irradiating the device. The experiment has to be repeated for various beam species and energy, allowing to assess the cross-section at various LET.

6.1.1 Characterization Method

To measure the cross-section of a device, in experiments, a variety of high-energy heavy ions or protons are usually applied to irradiate semiconductor devices. Each particle has specific energy in the test, corresponding to a specific LET for the particular semiconductor material. Afterwards, a procedure is required to identify, and count errors caused in the irradiation beam's device to compute the devices' error rate. It is hard for a complex functional circuit to determine the error location without a high testability design, which means transient signals, and stored data are highly available during the test. Meanwhile, the circuit to gain high testability also has an error rate. Hence, a particular circuit is usually devised to facilitate the statistical error rate.

The SEU is more straightforward to measure than SET because errors will always remain in the register or memory until the next overwrite. A shifter link is usually applied to test registers and the read-write scan loop for the RAM test. There are several test patterns for SEU: all 0, all 1, and checkerboard bit 0x55 or 0xAA, with which the SEU cross-section may vary. When the circuit stores different values, especially for registers, the sensitive nodes that can collect charges in the circuit are also different [1][2].

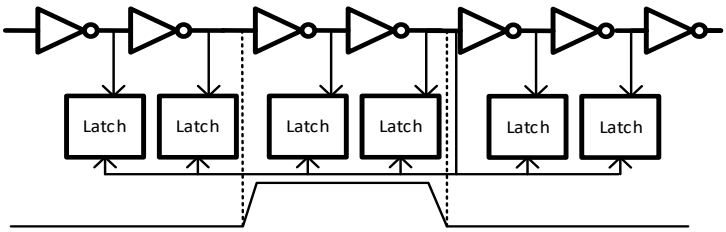


Figure 6.1. On-chip characterization of the SET pulse width.

The characterization of SET is more difficult because, as introduced in Section 2.2.3, the pulse is transient in the circuit and might be masked by subsequent circuits in the treatment chain. The on-chip characterization

method shown in Figure 6.1 can capture the pulse when SET is produced. The Nth stage output can supply a hold signal for latches to freeze the data and the SET pulse [3].

As shown in Figure 6.2, a typical cross-section versus LET plot has a threshold at lower energy, and the error rate will then augment until it reaches the knee point and stabilizes at the saturation value. The Weibull distribution can reproduce such behavior of the SEE cross-section of SEE. The accumulated distribution function of a Weibull fit for the SEE rate in function of the LET is

$$F(LET;\lambda,k)=\begin{cases}1-e^{-((LET)/\lambda)^{k}} & LET\geq 0\\ 0 & LET<0\end{cases} \tag{6.1}$$

In the formula, $1>k>0$ is the shape parameter, and $\lambda>0$ is the scale parameter of the distribution.

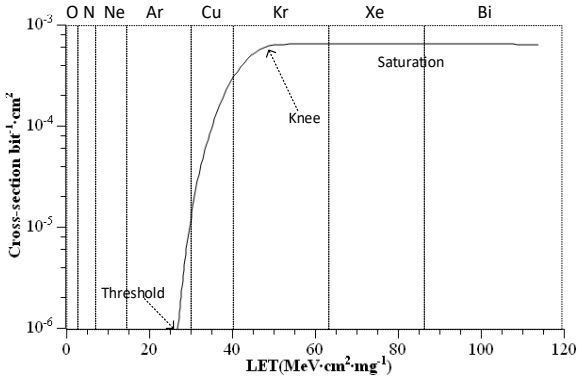


Figure 6.2. A typical cross-section V.S. LET plot. Heavy ions with the corresponding LET are marked at the top of the figure.

In order to solve the two unknown parameters in Eq. (6.1), at least 2 data points around the threshold are needed. Whereas, since the threshold position might not be known, it is necessary to measure the SEE rate over a relatively wide range of LETs.

6.1.2 Chip design

A SEE evaluation chip, devised to assess the SEE rate using the process used to fabricate the CBM experiment sensors, is controlled through the I^2C protocol, as the block diagram in Figure 6.3 illustrates. The chip

implements three memory blocks offered by the manufacturer, several registers and invertor test chains. In order to save the I/O pad for the

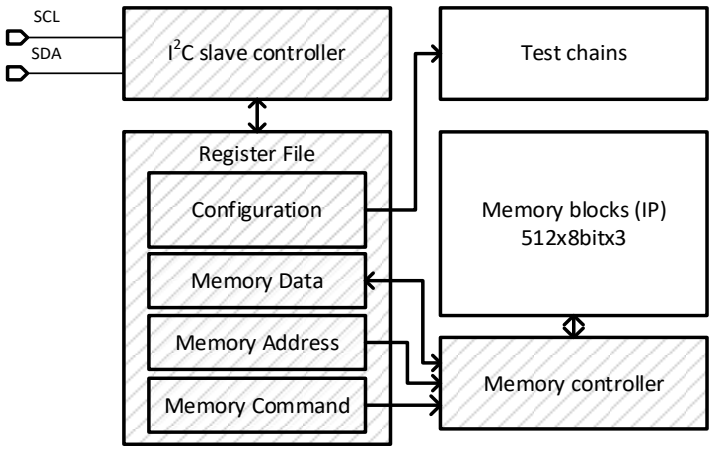


Figure 6.3. The block diagram of the SEE evaluation chip. The blocks with slash shadow are implemented with TMR.

memory test, a memory controller and a series of memory control registers are implemented in the evaluation chip, and the memory controller fully controls the 3 blocks of SRAM. The memory address, data, and command registers can be accessed from the I2C controller. The I2C slave controller, the register file, and the memory controller are all accomplished with TMR to avoid the SEE on these testability circuits.

The test chains are shown in Figure 6.4, where there are 4 inverter chains and 1 register chain performed in the evaluation chip. The inverter chains are devised to observe SET in various sizes of inverters, and the register chain is utilized to characterize SEU in static mode and measure the SET pulse width in dynamic mode. The inverters between the two registers' clock (block 1 in Figure 6.4) have a longer delay than that on the data line (block 3 in Figure 6.4).

It is worth noting that the unbalanced pull-up and pull-down current and the inverters' threshold voltage will primarily affect the transmit signal's duty cycle, which is called the pulse width adjustment effect (PWAE). [4] To minimize this effect, the inverters in Figure 6.4 are specifically devised to maintain the duty cycle shifting under $\pm 300ps$ after more than 10^4 levels of inverters. The evaluation chip is fabricated besides the PLL-

1/PLL-1+ using the modified TowerJazz 180nm process, with an area of $1500 \times 1200 \mu m^2$.

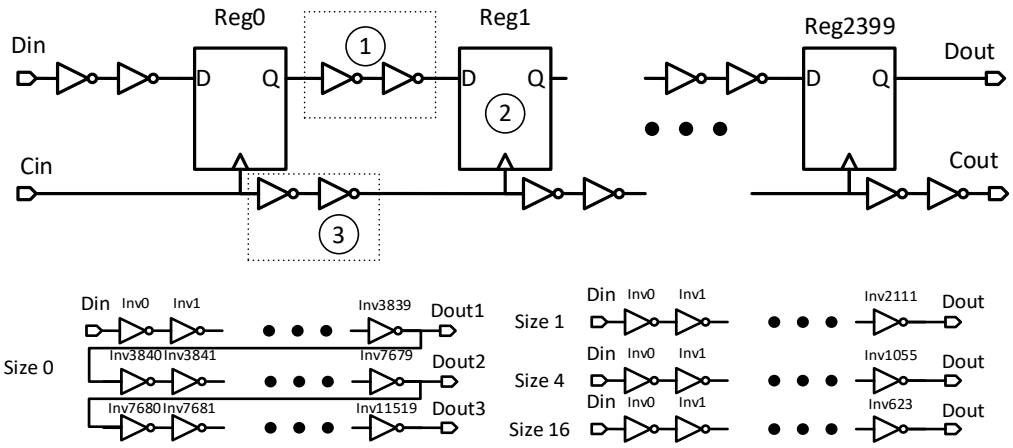


Figure 6.4. The test chains of the SEE evaluation chip.

6.1.3 Test Design

As presented in Figure 6.2, to extract the parameters in Eq. (6.1), the error rate needs to be assessed over a range of LETs obtained from various ion species and energies. Typically, Oxygen (2.19 $MeV\ cm^2 mg^{-1}$, same unit after), Neon (2.8) and argon (8.9) are utilized for the low-LET testing in Silicon, and heavier ions such as xenon (54.7) and gold (88.4) deliver much more ionization for the high-LET testing.

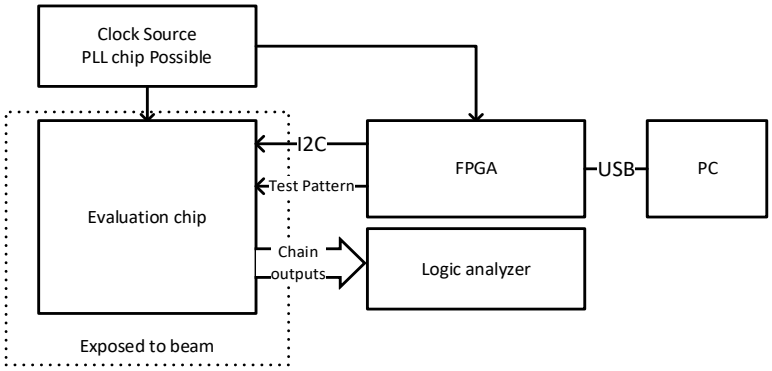


Figure 6.5. The test structure of the evaluation chip.

The test system is devised as depicted in Figure 6.5, where the clock source can be a commercial clock chip or the PLL-1 devised in Chapter 5. As the two chips are presented on the same wafer die, it is possible to bond

the clock output pad of PLL-1 with the evaluation chip's clock input pad. An FPGA with a fast I/O interface (~300MHz) is deployed to program the configuration by way of I^2C bus and call into being the test pattern for the memory and registers. The logic analyzer monitors the output of the chains and figures out the errors by recognizing the patterns.

- **Static test for SEU**

For the SEU test of registers and memory, the static test schedule is as follow:

1. Write the pattern into the registers or the memory with beam off;
2. Turn off the clock of the control and turn on the beam for a certain period;
3. Read out the result, count the upsets, and rewrite the test pattern (all 0 or all 1);
4. Augment or decline the irradiation duration if the upset count is too low or too high, and cumulate enough data for robust error rate estimation;
5. The cross-section is given by $\Theta(LET) = N_E/N_{bits}/\Phi_p$, where N_E is the error count, N_{bits} is the number of test bits and Φ_p is the fluence of the impacting particles;
6. Switch to another particle and/or energy and restart at step 1.

We can extract the parameters from a maximum likelihood fit of the Weibull distribution function to the SEE rate evolution with LET through this series of measurements. Then it is possible to find the cross-section threshold and saturation value of the register and SRAM.

- **Static test for SET**

On account of the uncertainty of the semiconductor manufacturing process, there will always be some manufacturing deviations in the device's inverter chain. In design, with Monte Carlo simulation, the PWAE of the inverters is lower than $300\ ps@\pm 3\sigma$ for 10^4 inverters, i.e., $< 30fs/cell$, so the pulse width can be observed directly at the data out terminal. To

minimize the effect of the pulse width adjustment of inverters, a calibration flow is devised as follow:

1. Feed the inverter chain with a full-scale sine wave with frequency (F_{in});
2. Observe the signal at data output 0-2 and record the duty cycle as $D_0\ D_1\ D_2$;
3. And the expectation of the PWAE of the inverter chain is $E(PWI)= (D_2 - D_0\ )/2/F_{in}$.

- **Dynamic test for SET**

The manufacturing deviations brings indeterminacy to the static measurement of the SET pulse width, which, however, can be utilized to measure the SET pulse width with dynamic mode. Based on the test method shown in Figure 6.1, the register chain can be utilized to measure the SET pulse width. As shown in Figure 6.6, when the register chain is running at a low frequency (~1MHz), the data pattern is transmitting in the register chain fast. If the chain is fed with pattern 0x55 (0 and 1 alternatively), the system output value is monitored with the logic analyzer, which can find out the timing when the data are disparate from the test pattern. As shown in Figure 6.7, if the SEE occurs on data inverters and registers (block 1 and 2 in Figure 6.4), the single-bit data error will be found in the output data. However, if the SET occurs on clock inverters (block 3 in Figure 6.5), a series of registers will take an additional sample that shifts the data by an additional place. Besides, the propagation of the SET pulse is influenced by the clock inverters, and the inverters will cut down/augment the pulse width and finally mask the pulse on the clock line or output the pulse at the $COUT$ terminal. So, the data shift length can be measured on the output and the pulse width can be computed as follow:

If the pulse is observed at $COUT$:

$$PW_{SET} = PW_{COUT} + N_D \cdot PWI \tag{6.2}$$

In the formula, PW_{SET} is the pulse width of SET, PW_{COUT} is the pulse width observed at COUT, N_D is the number of not sequenced data, and

PWI is the pulse width modulation parameter of inverters calibrated in the last section.

Alternatively, if the pulse cannot be observed at $COUT$:

$$PW_{SET} \leq T_{setup-DFF} + N_D \cdot PWI \tag{6.3}$$

In the formula, $T_{setup-DFF}$ is the setup requirement time for D flip-flop.

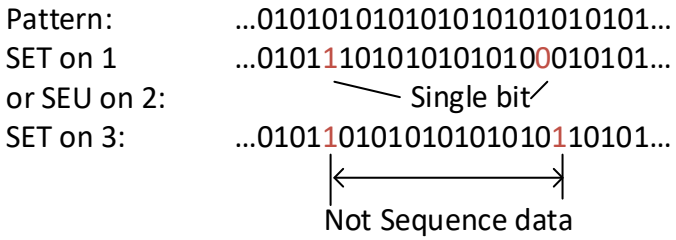


Figure 6.6. The error data may find in the monitoring.

6.2 SEE Beam Test Method of the Readout System

In Section 4.5.2, the readout system is simulated with SEE pulses injected, but the simulation is primarily devised to guide the optimization of TMR deployment, so the actual reliability has to be assessed from measurements in a beam. In this section, a beam test is devised to demonstrate the design method in Chapter 4 and provides another approach to measure the cross-section.

6.2.1 The testability design of MIMOSIS-1

It is arduous to identify the correctness by the random hit data of the pixel matrix. To make the readout system's output predictable, a further Data Generation for Multi-Frame Pattern Emulation (DGMFPE) module is devised. Figure 6.7 indicates that the DGMFPE module, situated between the pixel matrix and the region readout buffer, generates a specific data pattern when enabled instead of random hit data of the pixel matrix, and comprises 64 memories with 8 words of 8 bits utilized with the Multi-Frame Emulation mode, which is enabled with a global enable signal. Each

memory is coupled to a region, while each word is associated with a frame. These registers are available through I2C configuration control.

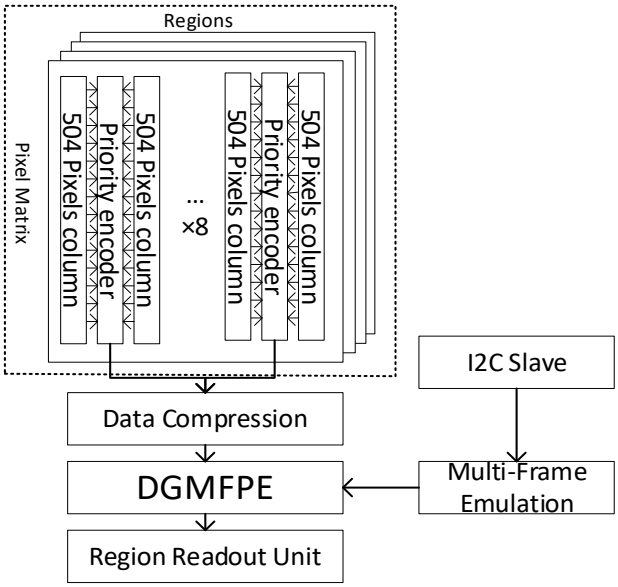


Figure 6.7. The DGMFPE module in the readout system.

The DGMFPE module sends 16-bits words at the input of the region buffers to replace the pixel response, and Table 6-1 gives the scheme of the sent word. The MSB (Most Significant Bit) is fixed to 0 to ensure no dedicated words in the data frame. Bits 14 to 12 are the truncated region number LSBs (Least Significant Bits), and bits 11-8 are the truncated frame counter LSBs. Moreover, bits 7 to 0 are the emulated pixel counter, which is incremented each time a pixel is read (every 50 ns). So, the DGMFPE module can also call into being the test pattern for the SEU test of the memory.

Table 6-1: Data word send in Multi-Frame Emulation mode

Bit	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Name	Fixed	Region Number			Frame Counter				Emulated Pixel Counter							
Value	0	0 to 7			0 to 15				0 to N							

For the registers' test, the MIMOSIS-1 has the configuration register in each pixel called the mask register, which controls whether the pixel is

available (unmasked) or not (masked). To save the pixel circuit area of the pixel, the mask registers in pixels are performed with a custom latch with the same store structure as registers, which form a register chain similar to Figure 6.4 and are accessed through the I^2C. With its high availability, it is an ideal material to test the cross-section of the registers.

6.2.2 The in-Time Test Method of MIMOSIS-1

The primary objective of the test is to identify the reliability of the MIMOSIS-1 readout system, and a test system is devised to monitor the MIMOSIS-1 readout system's performance, as shown in Figure 6.8. An FPGA contains an I^2C controller, a readout system in the behavior model, and a frame data comparator, and the frame data comparator synchronizes DUT's (Design Under Test) output with the behavioral model and figures out the frames' difference. The PC is connected to the FPGA through the USB interface to collect the results and statistical data so as to analyze them later.

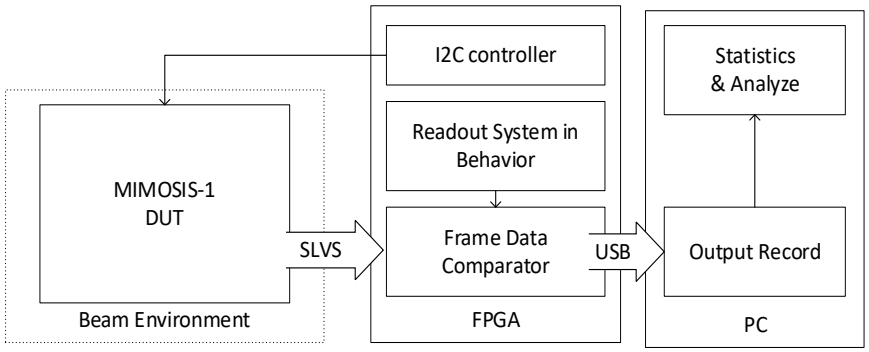


Figure 6.8. The beam test block diagram of MIMOSIS-1

The test is similar to the SEU test in Section 6.1.3, which follows these steps:

1. Configure the MIMOSIS-1 to use the DGMFPE module as the data source instead of the pixel matrix, same with the behavior model;
2. Start the two blocks at the same time and set the DUT in the beam environment;

3. Record the frame data comparator's data output with P.C. and figure out and count the single-bit errors, word errors, and frame errors.

Moreover, referring to the static test method for SEU in Section 6.1.3, the mask registers can be utilized to characterize the cross-section of the latches.

6.3 TID and SEE Beam Test Method of the PLL

As introduced in Chapter 5, the PLL is a mixed-signal component. On the one hand, the analog parts, like the charge pump, bandgap, and bias, are sensitive to TID. On the other hand, the digital parts like VCO, divider, and PFD, are sensitive to SEE. We will successively describe methods to assess the PLL radiation tolerance performance under the beam test.

6.3.1 TID Test Method

For the TID test of the devices, there are several standards for aerospace applications [5][6][7]. Generally, the ionizing dose is built up through irradiation by high energy γ from a Cobalt 60 source or by electrons from an accelerator beam. Different from SEE evaluation, real-time monitoring is not fundamental for measuring the effect of TID, but the circuit needs to be biased, i.e., powered up, to ensure that charges produced by ionization will move and accumulate as in operational conditions. Moreover, the TID test primarily focuses on the degradation of system performances and the augment of power consumption generated by bias and leakage current.

The annealing (i.e., compensation) of radiation-induced damage in SiO2 is a long-term process strongly subject to temperature and applies electric field. The basic mechanisms for electron compensation are tunneling an electron from the Si substrate and compensation by thermal excitation of an electron from the valence band [8]. Annealing will redistribute the fixed charge in the oxide, thereby restoring part of the device's original characteristics. Thus, annealing behaviors have crucial implications with total dose assessment of electronic devices.

The procedure follows the steps below:

1. Record initial performance (jitter for the PLL circuit) and power consumption in standard conditions before any irradiation;
2. Irradiate the tested circuit with a specific incremental dose with power ON;
3. Test the functionality, performance, and power consumption;
4. Repeat steps 2 and 3 until the total ionizing dose targeted for the test is reached;
5. Bias annealing and testing at room temperature for 24, 48 and 168 hours;

6.3.2 SEE Test Method

A SEE beam test is devised to demonstrate the reliability of the PLL-1+ devised in Chapter 5, and as shown in Figure 6.9, the board will carry 2 PLL chips with the same clock source. As a result, the outputs of the two PLL chips have the same frequency and phase. With a signal processing board comparing the difference of the two chips' outputs, any non-expected phase, which is the error caused by SEE, can be found. Counting such errors allows the statistical inference of the SEE cross-section for the PLL-1+ device.

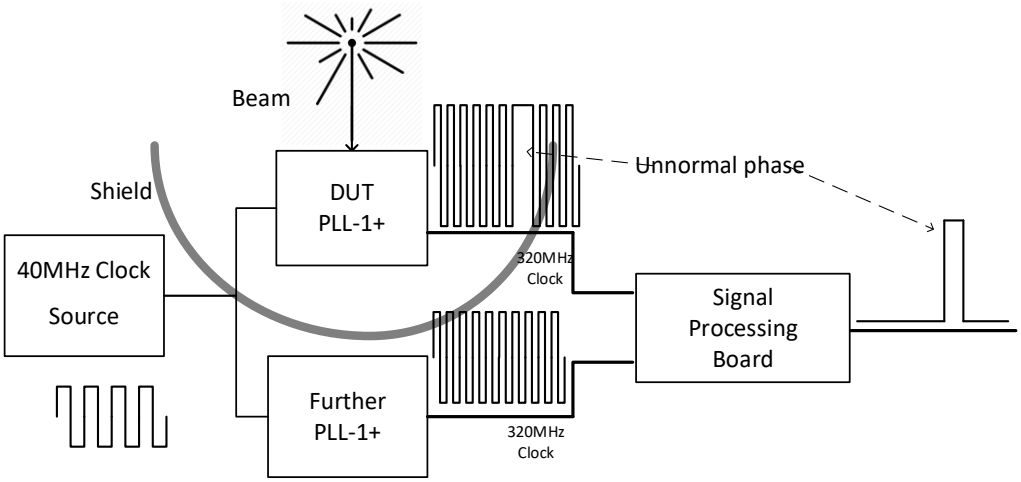


Figure 6.9. The block diagram of the SEE test of the PLL-1+.

6.4 Summary

In this chapter, an evaluation chip is introduced to characterize the SEU and SET cross-section of the standard cell library, the test method of SEU and SET of the evaluation chip are presented, and the reliability test of the MIMOSIS-1 readout system is devised. In addition, for the TID and SEE test of the PLL, static and dynamic test methods are devised. These test results can help us demonstrate the SEE model, confirm the radiation hardening design's effectiveness, and further guide the subsequent designs.

6.5 Bibliography

[1] C. H. Chen, P. Knag and Z. Zhang, Characterization of heavy-ion-induced single-event effects in 65 nm bulk CMOS ASIC test chips, IEEE Trans. Nucl. Sci. **61**, IEEE, (2014) 2694.

[2] A. Samaras, P. Pourrouquet, N. Sukhaseum, L. Gouyet, B. Vandevelde, N. Chatry, R. Ecoffet, F. Bezerra and E. Lorfevre, Experimental characterization and simulation of electron-induced SEU in 45-nm CMOS technology, IEEE Trans. Nucl. Sci. **61**, IEEE, (2014) 3055.

[3] B. Narasimham, V. Ramachandran, B. L. Bhuva, R. D. Schrimpf, A. F. Witulski, W. T. Holman, L. W. Massengill, J. D. Black, W. H. Robinson, et al., On-chip characterization of single-event transient pulsewidths, IEEE Trans. Device Mater. Reliab. **6**, (2006) 542.

[4] A. Balasubramanian, B. Narasimham, B. L. Bhuva, L. W. Massengill, P. H. Eaton, M. Sibley and D. Mavis, Implications of total dose on Single-Event Transient (SET) pulse width measurement techniques, IEEE Trans. Nucl. Sci. **55**, (2008) 3336.

[5] European Space Agency, ESCC Basic Specification No. 22900, Total Dose Steady-State Irradiation Test Method., (2016) 1.

[6] ASTM, F1892-12: ASTM Standard Guide for Ionizing Radiation (Total Dose) Effects Testing of Semiconductor Devices, (2012).

[7] D. of Defense, Test Method Standard, in MIL-STD-883j, Test Method Stand. Microcircuits **49**, (2013).

[8] J. R. Schwank, M. R. Shaneyfelt, D. M. Fleetwood, J. A. Felix, P. E. Dodd, P. Paillet and V. Ferlet-Cavrois, Radiation effects in MOS oxides, IEEE Trans. Nucl. Sci. **55**, (2008) 1833.

Chapter 7. Conclusions and Perspectives

7.1 Conclusions

The CBM experiment is expected to start around 2024 at the accelerator complex SIS-100 of the FAIR in the GSI, Darmstadt, whose research program is to explore the strong nuclear interaction phase diagram, driven by Quantum Chromo Dynamic (QCD), in the region of high baryon densities using high-energy fixed-target nucleus-nucleus or proton-nucleus collisions, including the study of the equation-of-state of nuclear matter at high densities and the search for the onset of deconfinement and chiral phase transition. The Micro Vertex Detector is the first of the CBM instruments, right after the target, whose task is to reconstruct with exquisite resolution the origin of the various secondary particles produced in the collision. The instrument itself is composed of four consecutive detection planes, each paved by sensors measuring the position thanks to pixelization of about $27 \times 30\ \mu m^2$. To match the requirements for physics analysis, the MVD has to operate in a vacuum and feature a very light material budget of about 0.3% of radiation length per plane.

The MIMOSIS-1 sensor, a Monolithic Active Pixel Sensor in the CMOS technology, is now being devised by IPHC as the full-size prototype sensor for the MVD. Beyond the functionalities used for physics, environmental conditions of operation for the MVD have to be taken into account in the design of the circuit. The highly ionizing particles generated by collisions, such as gold or carbon ions and protons, may indeed introduce various radiation effects into the MVD sensors: Single-Event

Effects (SEE), Total Ionization Dose (TID) effects, and displacement damages. In this thesis, the objective is to understand the impact of radiation effects on the sensor readout circuitry, elaborate a design strategy to mitigate this impact, and finally implement this specific design.

A SEE is an electrical disturbance that disrupts the normal operation of a circuit, and the cause is the passage of a single ion through or near a sensitive node in a circuit. The charged particles lose energy and generate a series of electron-hole pairs when passing through the semiconductor material, which is used for detecting the initial particle, with a specific collection node (diodes). In the other micro-circuits needed for the sensor functionalities, such electron-hole pairs driven by the electric field will also be collected by circuit terminals, resulting in a short current pulse in micro-circuits which may cause an error in their operation (e.g. logic). SEEs may result in temporary or permanent circuit functional errors such as Single-Event Upset (SEU), Single-Event Transient (SET), Single-Event Latch-up (SEL), Single-Event Functional Interrupt (SEFI), etc.

The TID effect is on account of the collection of ionized charges by the deep hole trapping near the insulating layer and the substrate, i.e. Si/SiO_2 interfaces, in the MOSFET. The TID effect can give rise to threshold voltage shifting and leakage current increment, which will provoke analog circuit dysfunctionality and make static power increase.

The first part of my work is to investigate the various techniques to harden integrated circuits against radiations, where three methods can be employed at the system design level: Triple Modular Redundancy (TMR); Error Correction Code (ECC); refreshing and reloading of registers. The re-entrant gate layout and reverse body biasing techniques, they address TID effects, very different from the TMR, ECC, refresh and reload technique. Finally, the Silicon-on-insulator (SOI) is a manufacturing processing approach. The principles, implementation methods, advantages, and disadvantages of these various methods are discussed in detail, and we summarize here our findings. The TMR is fit for a small but crucial part and the ECC is commonly utilized for data flow. Refreshing and reloading are better suited for configuration registers or monitoring a system working

flow. These three methods are well adapted to protect against SEEs, and the re-entrant gate layout and reverse body biasing are more suited to alleviate the TID effect. The SOI can eliminate SEE in principle but would require changing the sensing technology already validated for MIMOSIS-1 and will bring some undesirable side effects.

Focusing on SEE and TID impact, two crucial circuit modules in MIMOSIS-1 are identified, the first of which is the readout data buffer frontend. In order to meet the needs of the readout speed and high data compression specifications, a 3-layer-buffer structure is come up with as the readout system of MIMOSIS-1. This architecture converts the reading process from a fast burst data stream to a slow sequential data stream, and each layer of the buffer has its memory and control logic. The deployment of TMR and ECC is well suited for the readout frontend and is described in detail, so, it is necessary to balance the design parameters between the reliability in the radiation environment with the area and power budgets of the circuit.

In order to alleviate the contradiction in this design, this thesis proposes a design optimization approach for TMR deployment based on multi-level simulation. For estimating the circuit reliability against SEE, two ingredients are indispensable: the rate of highly ionizing particles and the SEE cross-section of the circuit. Both quantities bear large uncertainties since both the accelerator and the sensor are in development. Consequently, the evaluation is firstly based on the known cross-section from another similar CMOS pixel sensor (ALPIDE) applied in the ALICE experiment and then based on the model presented in the thesis. With a TCAD simulation based on the manufacture process parameters, the SEE pulse parameters, corresponding to various ionizing energy losses, can be acquired. The SEE generation tool can then inject several random SEE pulses based on the SEE model into the circuit, while the readout system is in simulation. Through the statistics of the simulation results, the stability of the module suffering the SEE can be acquired. With optimized design, the system can still run after 10^7 SEEs, which corresponds to more than five years of MVD operation. Meanwhile, the error rate with the

hardened design is reduced by 4 to 5 orders of magnitude. Moreover, the equivalent error rate in the working environment is lower than 10^{-9}.

The second crucial module is the frequency generation circuit, primarily the phase-locked loop (PLL), and MIMOSIS-1 steering is based on a synchronous digital logic, for which stable clocks are crucial. If the PLL has an unexpected output pulse, unstable period, or locks loss, the digital circuit drives will have serious functional issues. In the design produced in this work, TMR, re-entrant gate layout, and reverse body biasing are deployed. Again, the produced radiation-hard design is suitable within a small area and low power consumption.

With optimized design with radiation harden techniques, the prototype, PLL-1+, has a small area ($0.115mm^2$), low power consumption ($3.6mW$), short lock time ($5\mu s$), and lock detection function, with a total jitter level of $2\%UI$ in the independent test and $8\%UI$ in the heavy activity integration test.

The final part of this thesis is composed of the verification methods of the radiation hardening. In order to cross-check the simulation results, beam tests are foreseen, but it is difficult to compute the cross-section with the large functional chip (MIMOSIS-1) directly. A further evaluation chip is devised to calibrate the cross-section of the manufacturing process, and the static and dynamic SEE test methods for the evaluation chip, MIMOSIS-1, and PLL-1+ are all devised. Besides, the TID test for PLL-1+ is introduced.

In summary, this thesis has studied the requirement and environment of the MVD in CBM experiment, presented the harden techniques suit for the MVD based on the mechanism of the radiation effects, and then presented a global strategy for radiation-hardening by design where additional constraints (like microcircuit sizes and power budget) are stringent. The method is based on both specific simulations and assessment of the crucial level of each functionality in the overall circuit.

The strategy is deployed on the first full-size prototype sensor for the Micro Vertex Detector of the CBM experiment. Simulation and test results manifest that it is possible to design a detector with strict restrictions to

meet the needs of radiation tolerance, and both the digital and analog circuits work properly under SEE and TID effects through simulation. The method in this thesis can also be extended to other radiation-hardened integrated circuit designs.

7.2 Perspectives

On the basis of the research in this thesis, there are further steps that can be taken for SEE hardening in smaller-scale processes and simulation methods.

7.2.1 SEE in smaller-scale process

As of today, 5nm process chips have begun mass production. Although the size of CMOS processes continues to shrink under the view of "More Moore's Law", many factors can directly affect their sensitivity to SEE, such as shrinking device sizes, increasing operating frequencies, and lowering of the crucial charge required to cause SEE.

The combination of a reduction in device feature size and an increment in circuit integration (i.e., the number of transistors per unit area) has important implications for soft errors in [1]:

- the reduction of the per-bit cross-section presented to an incident ionizing particle;
- the reduction of the energy deposition volumes traversed by the particle;
- the increment of the particle region of influence in the circuit plan.

It should be noted that, for bulk technologies, the scaling of planar dimensions has not been accompanied by like scaling of the vertical dimensions in the FEOL (Front End of Line) processing such as wells or epitaxial depths. Consequently, the efficiency of energy transfer from an incident ionizing particle track to circuit nodes has scaled at a rate closer to the feature size squared rather than feature size cubed [1].

Figure 7.1 illustrates the increment in the area of influence particles in the circuit, which is one of the most spectacular effects observed in recent technology. As the influence of a single event is not "point-like" but has a certain radial expansion (resulting in the radial charge distribution shown

below), the point of effect can intersect more or less diverse parts of the circuit, related to the specific technological node. In Figure 7.1, the influence point of alpha particles intersects with up to 6 memory cells in a 45 nm SRAM, while the previous influence is limited to a single cell at the 130 nm node. This pure geometric effect is partly (and only partly owing to) the cause of the multicellular upset (MCU) observed in all recent technologies (usually below the 130/90 nm technology node).

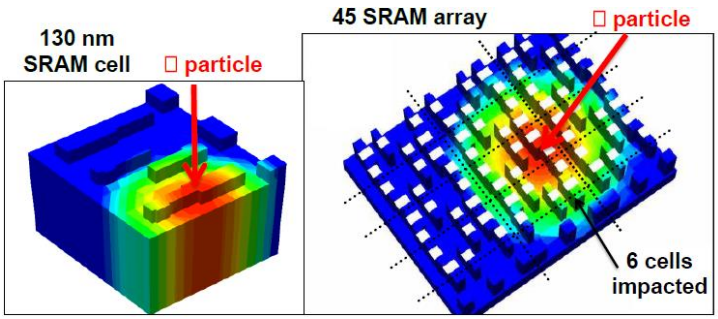


Figure 7.1. Illustration of the domain of influence of a single alpha-particle striking an SRAM in both 130 nm and 45 nm technologies. Only a single cell is impacted in 130nm whereas a cluster of 6 adjacent cells is impacted.[2]

The concept of critical charge (Q_{crit}) is a first-order measure introduced to quantify the sensitivity of static memory from one logical state to another, which is consistent with the smallest quantity of charge that can reverse the data bit deposited in the memory cell. It has an exponential relationship with the Soft circuit Error Rate (SER), as shown by the analysis model developed by [3] :

$$SER = K \times A \times Fe \times \exp(-Q_{crit}/Q_S) \qquad (7.1)$$

In the formula, K is a scaling factor, F_e is the particle flux ($cm^{-2} \times s^{-1}$), A is the area of the circuit sensitive to particle strikes (cm^2), Q_{crit} is the crucial charge and Q_S is the charge collection efficiency of the device (same unit as Q_{crit}).

The two key parameters of SER are the critical charge (Q_{crit}) of the SRAM cell and the charge collection efficiency (Q_S) of the circuit, both of which are determined by process technology, and Q_{crit} also depends on the characteristics of the circuit, especially the power supply voltage and the effective capacitance of the drain node. Q_{crit} and Q_S are essentially

independent, but both drop off as the feature size decline. Equation 7.1 emphasizes that changes in the ratio Q_{crit}/Q_S will have a very large impact on the final SER, which is also proportional to the area of the sensitive area of the device, thus decreasing in proportion to the square of the device size.

There are papers [4][5] that propose a simple process with decreased differences in full SPICE simulation (⩽10 and below improved analysis process). This last model takes into account the dynamic behavior of the cell and proves a simple technique to decouple the nonlinearly coupled storage nodes. The crucial value model developed in this way is composed of NMOS and PMOS transistor parameters, which is consistent with the results of commercial 90nm CMOS process SPICE simulation, and the maximum error is less than 5%.

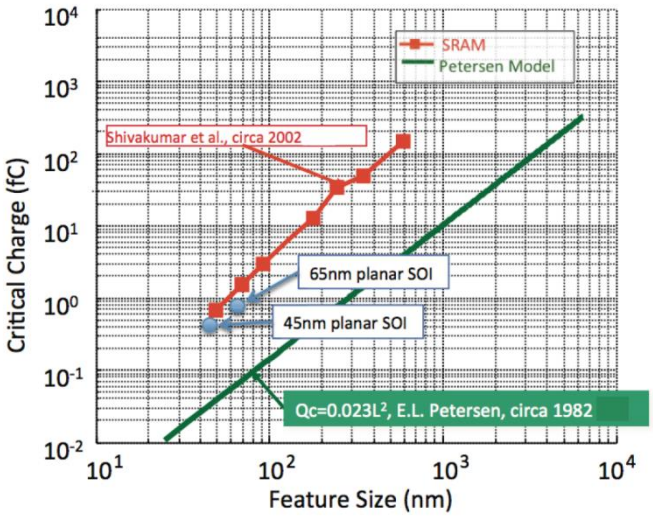


Figure 7.2. Critical charge scaling as a function of feature size. SRAM data from [6] *are shown in red. The 1980's scaling model* [7] *is based on a second power dependence on feature size is shown in green.*

Meanwhile, as the crucial charge reduce along with the scaling of feature size, not only the high energy particle source but also the natural radiation at ground level may cause soft errors.[8]

Low-energy protons are usually produced during the scattering of high-energy protons or neutrons, primarily in the space radiation environment [9], whose impact on modern electronics is an important issue for space and terrestrial applications. Although the shielding material can easily

absorb low-energy protons, the scattering of high-energy protons will produce a large number of low-energy protons, which affects electronic devices. In addition, as secondary low-energy protons are produced in the exfoliation reaction of high-energy neutrons with silicon and other materials existing in modern semiconductor devices, it is also necessary to consider the SER prompted by low-energy protons in the ground radiation environment [10].

Atmospheric muons are an important part of the ground natural radiation background. Muons are leptons, but come from the "hard" components in the cosmic ray cascade of the atmosphere, and are the products of attenuation of charged ions (charged muons $\mu+$ and μ -) through weak interactions, and these particles can easily penetrate the atmosphere, and constitute the most abundant high-energy charged particles at sea level.

The impact of muons on electronics has been the object of only a few studies until recently. We can cite the pioneering work of Ziegler and Lanford [11] in the 1980s, and some experimental research results using artificial muon beam storage, as well as the experimental research results of sub-beam storage [12][13][14]. Recently, Sierwski et al. [15][16] have combined measurement and numerical simulation in combination with measurement and numerical simulation, and conducted a study on the influence of low energy (<3MeV) and atmospheric positive muons on advanced processes. The impact on the process has carried out the first major work on this topic, which has verified and quantified the effect of muon direct ionization on disparate bulk SRAMs at different process nodes (65, 55, 45, and 40nm). The data presented in Figure 7.3 indicates the possibility of abnormality of the M20B surface muon beam using TRIUMF when subjected SRAM at nominal biases to four different μ+ radiation.

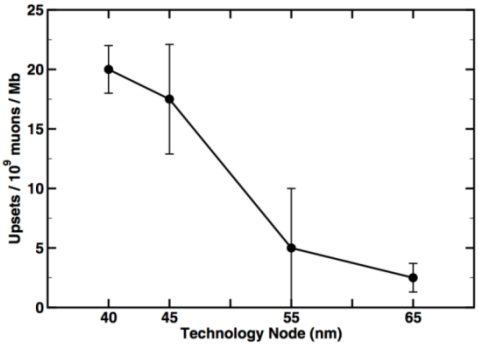


Figure 7.3. Experimental muon-induced single event upset possibility for different SRAMs under test operated at nominal supply voltage and irradiated using a 21.6 MeV/c μ+ beam. [16]

These latest studies clearly indicate that for the most advanced CMOS technology, low-energy atmospheric muons are crucial as a new radiation constraint on the ground. It is necessary to conduct further analysis in the future to study the exact ratio of positive and negative ions in particular, which will not only affect the circuit architecture but also affect the local environment (shielding) of the circuit, thereby depositing a large amount of charge in the silicon. The distribution of low-energy atmospheric muons below 1 MeV will be profoundly influenced. As a result, the current lack of experimental and theoretical knowledge relevant to atmospheric muon distribution below 1 MeV should represent a limitation and cannot accurately estimate the impact of muons on ground electronic devices.

7.2.2 SEE evaluation by simulation

In Section 4.5.1, the SEE pulse is modeled in a mixed TCAD 3-D mesh network and the SPICE netlist, which is not a very accurate way to identify the parameter of the SEE pulse. Further tools may be combined with physics simulation tools likes Geant4 [17] to gain a more accurate SEE response of the circuit.

The Geant4 is a Monte-Carlo based on particle transportation simulation tool, and the accuracy of Monte-Carlo prediction depends on the accurate description of the entire simulation chain, including the description of the radiation source, physical transmission process, geometric model, and response modeling. Some tools cover (almost) all

particle types, while other tools have better algorithms for certain types of particles.

SER-based physical simulation based entirely on Monte Carlo supplies a very powerful approach that can make the error rate prediction process more detailed than previous models and analysis calculations [18][19]. Schematically, the Monte Carlo simulation code solves the radiation problem through two main steps, namely, the interaction of radiation, the device and the subsequent charge movement as well as the change of node current and/or voltage in the device/circuit. The complete simulation chain is very complicated owing to its multi-scale and multi-physical characteristics.

- **MRED**

The Monte-Carlo Radiative Energy Deposition (MRED) software developed at Vanderbilt is a framework for treating single event effects in integrated circuits, in which the component tool describing radiation interactions and transport in the matter are a built-in, Monte Carlo, binary-collision code. The elementary principle is to comprehensively handle all forms of radiation that interact with materials and offer an interface that runs smoothly with pertinent programs that deal with other parts of the conundrum, such as the transmission of radiation-induced charges or circuit effects analysis of radiation-induced changes. As one of the software systems, probably the most important single design decision is the attribute of the linguistic interaction among various independent constituents [18]. In MRED, Python is the chosen language for this interaction and for upset rate predictions, particle transport and energy deposition use the Geant4 physics and a device geometrical model that contains the FEOL and BEOL structures. The use of Geant4 libraries allows to model the spatial and temporal distribution of charge that is caused by a particle. Afterwards, the energy deposition is tracked in the sensitive volumes, modeled as a set of Rectangular Parallelepiped (RPP) boxes.

- **TIARA-G4**

In recent years, TIARA-G4 [20] has been jointly developed at Aix-Marseille University (IM2NP Lab) and STMicroelectronics (Central R&D Center, Crolles), which is a general Monte Carlo simulation code written in C++, completely based on Geant4 toolkit, applied to model Geant4 particles (including neutrons, protons, muons, alpha particles, and heavy ions) and the interaction of various architectures and electronic circuits. TIARA stands for the tool suite applied for radiation reliability assessment, whose primary objective is to embed the latest knowledge and methods of SER evaluation into a unique simulation platform. The initial version of TIARA [21][22] is a standalone C++ native code that is dynamically linked to the IC CAD stream through coupling with the SPICE solver. The code has been developed so as to make it simple to add a new radiation environment, physical model, or new circuit architecture. On the one hand, this first version can handle the transmission and energy deposition of charged particles (heavy ions and alpha particles) without the demand for nuclear codes like Geant4. Only the SRIM table was applied as an input file to compute the flow of particles in the silicon and simplified to a single layer with a simplified BEOL structure. On the other hand, as for neutrons, a separate database is used, which is compiled with a specific Geant4 application, to call into being nuclear events in the simulation stream that are caused by the interaction of incident neutrons with the circuit.

The new version of TIARA is denominated as TIARA-G4 since it has been completely rewritten in C++ and compiled into a complete Geant4 application using Geant4 classes and libraries. Nuclear events are no longer provided from the database but directly produced by Geant4 in the simulation code stream, which now allows us to use the virtual geometry model (VGM[23]) factory and perform computations with Geant4 and visualize with Root[24], thus considering all the complexity of the circuit in terms of materials, doping, and 3D geometry. That is to say, compared with the first version of the code, the primary enhancement of TIARA-G4 is the conversion of the code in the Geant4 application, which allows the utilization of Geant4 classes to depict the geometry and materials of the

circuit (now including the real BEOL structure) and integrate particle transmission and tracking directly in the simulation process without the demand for external databases or other files.

With these EDA and radiation effects simulation tools, the design of radiation harden integrated circuits can be more automated and intelligent.

7.3 Bibliography

[1] L. W. Massengill, B. L. Bhuva, W. T. Holman, M. L. Alles and T. D. Loveless, *Technology Scaling and Soft Error Reliability*, in *IEEE Int. Reliab. Phys. Symp. Proc.*, (2012).

[2] P. Roche, *Year-In-Review on SER*, utorial at the International Reliability Physics Symposium (IRPS), Phoenix, (2006).

[3] P. Hazucha and C. Svensson, *Impact of CMOS technology scaling on the atmospheric neutron soft error rate*, *IEEE Trans. Nucl. Sci.* **47**, IEEE, (2000) 2586.

[4] J. L. Autran, S. Serre, D. Munteanu, S. Martinie, S. Semikh, S. Sauze, S. Uznanski, G. Gasiot and P. Roche, *Real-time soft-error testing of 40nm SRAMs*, *IEEE Int. Reliab. Phys. Symp. Proc.*, IEEE, (2012) 1.

[5] S. M. Jahinuzzaman, M. Sharifkhani and M. Sachdev, *An analytical model for soft error critical charge of nanometric SRAMs*, *IEEE Trans. Very Large Scale Integr. Syst.* **17**, (2009) 1187.

[6] P. Shivakumar, M. Kistler, S. W. Keckler, D. Burger and L. Alvisi, *Modeling the effect of technology trends on the soft error rate of combinational logic*, *Proc. 2002 Int. Conf. Dependable Syst. Networks*, (2002) 389.

[7] F. L. Petersen, P. Shapiro, J. H. Adams and E. A. Burke, *Calculation of cosmic-ray induced soft upsets and scaling in VLSI devices*, *IEEE Trans. Nucl. Sci.* **29**, (1982) 2055.

[8] G. Just, J. L. Autran, S. Serre, D. Munteanu, S. Sauze, A. Regnier, J. L. Ogier, P. Roche and G. Gasiot, *Soft errors induced by natural radiation at ground level in floating gate flash memories*, *IEEE Int. Reliab. Phys. Symp. Proc.*, (2013).

[9] N. Seifert, B. Gill, J. A. Pellish, P. W. Marshall and K. A. Label, *The susceptibility of 45 and 32 nm bulk CMOS latches to low-energy protons*, *IEEE Trans. Nucl. Sci.* **58**, IEEE, (2011) 2711.

[10] K. P. Rodbell, D. F. Heidel, H. H. K. Tang, M. S. Gordon, P. Oldiges and C. L. Murray, *Low-Energy Proton-Induced Single-Event-Upsets in 65 Nm Node, Silicon-on-Insulator, Latches and Memory Cells*, in *IEEE Trans. Nucl. Sci.*, (2007).

[11] J. F. Ziegler and W. A. Lanford, *Effect of cosmic rays on computer memories*, *Science (80-.).* **206**, (1979) 776.

[12] J. F. Dicello, C. W. McCabe, J. D. Doss and M. Paciotti, *The relative efficiency of soft-error induction in 4k static rams by muons and pions*, *IEEE Trans. Nucl. Sci.*, (1983).

[13] C. J. Gelderloos, R. J. Peterson, M. E. Nelson and J. F. Ziegler, *Pion-induced soft upsets in 16 Mbit DRAM chips*, *IEEE Trans. Nucl. Sci.*, (1997).

[14] S. Duzellier, D. Falguère, M. Tverskoy, E. Ivanov, R. Dufayel and M. C. Calvet, *SEU Induced by Pions in Memories from Different Generations*, in *IEEE Trans. Nucl. Sci.*, (2001).

[15] B. D. Sierawski, M. H. Mendenhall, R. A. Reed, M. A. Clemens, R. A. Weller, R. D. Schrimpf, E. W. Blackmore, M. Trinczek, B. Hitti, et al., *Muon-Induced Single Event Upsets in Deep-Submicron Technology*, in *IEEE Trans. Nucl. Sci.*, (2010).

[16] B. D. Sierawski, R. A. Reed, M. H. Mendenhall, R. A. Weller, R. D. Schrimpf, S. J. Wen, R. Wong, N. Tam and R. C. Baumann, *Effects of Scaling on Muon-Induced Soft Errors*, in *IEEE Int. Reliab. Phys. Symp. Proc.*, (2011).

[17] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, et al., *GEANT4 - A simulation toolkit*, *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, (2003).

[18] R. A. Weller, M. H. Mendenhall, R. A. Reed, R. D. Schrimpf, K. M. Warren, B. D. Sierawski and L. W. Massengill, *Monte Carlo simulation of single event effects*, *IEEE Trans. Nucl. Sci.* **57**, (2010) 1726.

[19] R. A. Reed, R. A. Weller, A. Akkerman, J. Barak, W. Culpepper, S. Duzellier, C. Foster, M. Gaillardin, G. Hubert, et al., *Anthology of the development of radiation transport tools as applied to single event effects*, *IEEE Trans. Nucl. Sci.*, (2013).

[20] J.-L. Autran, S. Semikh, D. Munteanu, S. Serre, G. Gasiot and P. Roche, *Soft-Error Rate of Advanced SRAM Memories: Modeling and Monte Carlo Simulation*, *Numer. Simul. - From Theory to Ind.*, (2012).

[21] S. Uznanski, G. Gasiot, P. Roche, C. Tavernier and J. L. Autran, *Single event upset and multiple cell upset modeling in commercial bulk 65 nm CMOS SRAMs and flip-flops*, *Proc. Eur. Conf. Radiat. its Eff. Components Syst. RADECS* **57**, (2009) 194.

[22] S. Uznanski, G. Gasiot, P. Roche, S. Semikh and J. L. Autran, *Combining GEANT4 and TIARA for Neutron Soft Error-Rate Prediction of 65 Nm Flip-Flops*, in *IEEE Trans. Nucl. Sci.*, (2011).

[23] I. Hivnáčová and B. Viren, *The Virtual Geometry Model*, in *J. Phys. Conf. Ser.* **119**, Institute of Physics Publishing, (2008).

[24] *ROOT: analyzing petabytes of data, scientifically. - ROOT*https://root.cern.ch/ (12 October 2020).

Direction de la **recherche**
et de la **valorisation**
Collège **doctoral**
Université de Strasbourg

Déclaration sur l'honneur

Declaration of Honour

J'affirme être informé que le plagiat est une faute grave susceptible de mener à des sanctions administratives et disciplinaires pouvant aller jusqu'au renvoi de l'Université de Strasbourg et passible de poursuites devant les tribunaux de la République Française.

Je suis conscient(e) que l'absence de citation claire et transparente d'une source empruntée à un tiers (texte, idée, raisonnement ou autre création) est constitutive de plagiat.

Au vu de ce qui précède, **j'atteste sur l'honneur que le travail décrit dans mon manuscrit de thèse est un travail original et que je n'ai pas eu recours au plagiat ou à toute autre forme de fraude.**

I affirm that I am aware that plagiarism is a serious misconduct that may lead to administrative and disciplinary sanctions up to dismissal from the University of Strasbourg and liable to prosecution in the courts of the French Republic.

I am aware that the absence of a clear and transparent citation of a source borrowed from a third party (text, idea, reasoning or other creation) is constitutive of plagiarism.

In view of the foregoing, ***I hereby certify that the work described in my thesis manuscript is original work and that I have not resorted to plagiarism or any other form of fraud.***

Nom : Prénom : ZHAO Yue

Ecole doctorale : ED182 Physique et Chimie-Physique

Laboratoire : Institut Pluridisciplinaire Hubert CURIEN (IPHC)

Date : 05/06/2021

Signature :

Yue ZHAO

Radiation Hardened Design of CMOS Pixel Sensor for the Micro-Vertex Detector of the CBM Experiment

Résumé

L'expérience CBM (*compressed baryonic matter*) étudiera le diagramme de phase de la matière nucléaire à des densités extrêmes, en observant des collisions or-or ou proton-or. Le détecteur Microv Vertex (MVD), l'instrument le plus proche des collisions. est conçu à l'aide de capteurs à pixels CMOS. Ces collisions produisent de nombreuses particules très ionisantes, telles que des ions or, carbone ou des protons. Ces rayonnements peuvent générer différents types de perturbations dans les capteurs du MVD : effets d'événement unique (SEE), effets de dose d'ionisation totale (TID) et dommages par déplacement.

Cette thèse propose une méthode de durcissement aux rayonnements par conception (RHDB) qui respecte les principales contraintes des capteurs imposées par l'expérience, notamment leur taille. En identifiant les fonctionnalités critiques du circuit et avec des simulations dédiées, nous introduisons une conception HRDB de manière sélective dans deux parties numériques des capteurs du MVD et réussissons ainsi à augmenter significativement leur résistances aux SEE. Il s'agit de l'architecture de lecture des données et d'un circuit à boucle de verrouillage (PLL).

Mots-clés: capteurs à pixels CMOS, effets d'événement unique, effet de dose d'ionisation totale, durcissement aux rayonnements par conception

Résumé en anglais

The compressed baryonic matter experiment (CBM) will explore the behaviour of nuclear matter at extreme densities using Gold-Gold or proton-Gold collisions. The Micro Vertex Detector (MVD) is the CBM instrument closest to the collisions and is being designed using CMOS pixel sensors. The high-ionization particles generated by these collisions, such as gold or carbon ions and protons, may introduce radiation effects into the MVD sensors, such as: single-event effects (SEE), total ionization dose (TID) effects and displacement damages.

In this thesis, we derive a method to implement radiation-hardening-by-design (RHBD) without compromising the essential features of the CMOS circuits, essentially its size. Using dedicated simulations and considering critical functionalities, we selectively introduced RHDB in two digital parts of the MVD sensor prototype, hardening them considerably against SEE: the read-out architecture and a phase-locked loop circuit.

Keywords: CMOS Pixel Sensors, Single-Event Effects, Total Ionization Dose Effects, Radiation-Hardening-By-Design