UC Riverside
UC Riverside Electronic Theses and Dissertations

Title
Radio Frequency Switch Design with Interference Suppression and Electrostatic Discharge for 5th Generation of Mobile Network

Permalink
https://escholarship.org/uc/item/7cc3p5vs

Author
Wang, Chenkun

Publication Date
2018

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Radio Frequency Switch Design with Interference Suppression and Electrostatic Discharge for 5th Generation of Mobile Network

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Chenkun Wang

March 2018

Dissertation Committee:
Dr. Albert Wang, Chairperson
Dr. Sheldon Tan
Dr. Ming Liu

Copyright by
Chenkun Wang
2018

The Dissertation of Chenkun Wang is approved:

____

____

____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

After an amazing period of three years and six months, today is the day: writing this note of thanks is the finishing touch on my dissertation. It has been a period of intense learning for me, not only in the scientific arena, but also on a personal level. I would like to reflect on the people who have supported and helped me so much throughout this period.

Firstly, I would like to express my sincere gratitude to my advisor Prof. Albert Wang for the continuous support of my PhD study and related research, for his patience, motivation and immense knowledge. His guidance helped me in all the time of research and life in these years. I could imagine having not only a mentor but a friend in my future life.

Beside my advisor, I would like to thank the rest of my dissertation committee: Prof. Sheldon Tan and Prof. Ming Liu, for their insightful comments and encouragement, but also for the questions which incented me to widen my research from various perspectives.

It's my honor to work with my fellow labmates, Dr. Fei Lu, Dr. Rui Ma, Dr. Qi Chen, Mr. Feilong Zhang and Mr. Cheng Li, for the discussions and guidance, for the sleepless nights we were working together before the deadlines, and for all the fun we have had in the past three years. Also, I thank Dr. Gang Chen, his students and my friends in SMIC for their professional support.

I appreciate Marvell Semiconductor for providing me an opportunity to join as an internship. Many thanks to my manager Mr. Zhendong Guo, my colleagues Mr. Hui Zhao

and special thanks to my mentor Mr. Peng Liu, his immense knowledge and unselfish share is invaluable throughout my life.

Last but not the least, I would like to thank my parents and my girlfriend, Gege Yang, for supporting me spiritually throughout this “journey” and my life in general.

ABSTRACT OF THE DISSERTATION

Radio Frequency Switch Design with Interference Suppression and Electrostatic Discharge Consideration for 5th Generation of Mobile Network

by

Chenkun Wang

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, March 2018
Dr. Albert Wang, Chairperson

In the next few years, the 5th generation of mobile network employing the massive multiple input multiple output (MIMO), beam-forming, mm-wave frequency bands and carrier aggregation techniques will further increase the data rate to enrich the growing of mobile devices at the price of more complex multi-band, multi-frequency front end module (FEM). As an indispensable part of radio-frequency front end (RFFE), antenna switch circuit needs more restrict performance not only the basic insertion loss, isolation and power handling capability, but the requirements from higher data rate, low interference and better reliability of 5G application.

To achieve those additional requirements for 5G RF switch with 45nm silicon-on-insulator (SOI) CMOS technology, this dissertation presents a novel multi-bands switch array structure to analysis the interference between switches in a single chip. The comprehensive study of switch array reveals that existing noise isolation techniques are

insufficient and calls for novel in-die interference elimination. To reduce this in-die crosstalk, an above-silicon through back-end-of-line (BEOL) metal wall is developed as a practical solution with about 18.5dB (~98.6%) suppression. To reach a higher data rate with available frequency bands, millimeter wave (mm-wave) switches (28GHz/38GHz) has been demonstrated with the consideration of reliability issue of electrostatic discharge (ESD) which will introduce severe parasitic effects under this frequency level and degrade the performance of RFICs. The insertion loss and isolation together with ESD protection capability have been compared which shows the importance of ESD-RFIC co-design. Considering the necessity of accurate estimation for ESD performance, a novel methodology for both human body model (HBM) and charged device model (CDM) ESD protections using combined TCAD simulation and TLP/VFTLP measurements is depicted. To improve the accuracy of parasitic capacitance extraction, this dissertation introduces an enhanced de-embedded method with the help of HFSS simulation which reduces one third of the testchip size and gives a better reference for co-design.

Contents

List of Figures x

List of Tables xiv

Chapter 1 Introduction 1

1.1 History of wireless communication 1
1.2 Organization of the Thesis 5

Chapter 2 Background of 5G RFICs 7

2.1 What is 5G? 7
2.2 Front End Introduction 19
2.3 RF Switch for 5G Application 24
2.4 45nm silicon-on-insulator (SOI) technology 26

Chapter 3 RF Switch Array 28

3.1 SPnT RF Switch Design 29
3.2 Switch Array and Interference 36
3.3 Measurements and Discussion 37

Chapter 4 RF Flying Crosstalk Suppression 41

4.1 Interference Suppression Using Metal Wall 44
4.2 Implementation in RF Switch Circuits 47
4.3 Measurements and Discussion 49

Chapter 5 Electrostatic Discharge Protection 55

5.1 ESD Introduction 56
5.2 Human Body Model (HBM) ESD Protection 63
5.3 Charged Device Model (CDM) ESD Protection 77

Chapter 6 ESD Parasitic Effects 95

6.1 ESD Parasitic Capacitance 95
6.2 Enhanced Parasitic Extraction 103
6.3 Mm-wave Switch and ESD Protection Parasitic Impacts 108

Chapter 7 Conclusion 117

References 119

List of Figures

Figure 1 Maxwell's Equations 1
Figure 2 The development of mobile network generations (1G – 5G) 3
Figure 3 The frequency range of different technologies 4
Figure 4 5th generation mobile networks user cases...8
Figure 5 RF front end module of mobile phones (simplified) 19
Figure 6 RFFR module of Smartphone using discrete switches or Antenna Switch Modules (ASMs) 20
Figure 7 802.11 a/b/g/n/ac dual-band WiFi front end components.................................21
Figure 8 Block diagram of 5G user equipment front end wireless system architecture...22
Figure 9 FFC approved mm-wave bands for 5G research...23
Figure 10 mm-wave attenuation in rain (left) and atmosphere (right). 24
Figure 11 Recent research shows great isolation, insertion loss and return loss are hard to achieve at mm-wave frequency..25
Figure 12 Simulated doping profile of a 45nm asymmetric (AFET) NMOS device. 27
Figure 13 Basic parasitic capacitance from a simple MOSFET. 30
Figure 14 Basic model of simple NMOS switch during ON/OFF-state...........................31
Figure 15 Gate voltage changing with drain side input when AC floating. 32
Figure 16 Voltage distribution of stacked NMOS with/without FFC. 33
Figure 17 Simplified switch schematics of series-shunt, gate resistance, stacked FETs and FFC techniques..33
Figure 18 Insertion loss of high band T/Rx SPDT switch...34
Figure 19 Isolation of high band T/Rx SPDT switch. 34
Figure 20 Input and output power of high band T/Rx SPDT switch at 2GHz.................34
Figure 21 Insertion loss of low band T/Rx SP3T switch..35
Figure 22 Isolation of low band T/Rx SP3T switch. 35
Figure 23 Input and output power of low band T/Rx SP3T switch at 750MHz. 35
Figure 24 The 2x2 switch array design splits including layout, die photo and testing board photo. 36
Figure 25 Measured S21 (Insertion loss for ON-state/Isolation for OFF-state) for the individual switches at PCB level...38
Figure 26 Measured inter/inner-band interference under 0dBm input at 2GHz shows varying interference suppression related to the layout distances of different switches.....39
Figure 27 A simplified EM simulation to compare through-BEOL flying crosstalk with in-Si crosstalk in a 45nm SOI CMOS platform 42
Figure 28 EM simulation results of the two cases: In-Si and From BEOL (Flying Crosstalk) 42
Figure 29 X-section illustration of new in-BEOL metal wall flying crosstalk suppression structure applied to RF T/Rx switches in a 45nm SOI CMOS. 44
Figure 30 A die-scale photo shows the through-BEOL metal wall between the two SPDT switches on a chip...46

Figure 31 (a) A simplified SPDT switch schematic featuring series-shunt MOSFET topology. (b) Circuit features including 12-FET stacks, gate resistor (RG) and feed-forward capacitor (FFC) are used in each stack branch. (c) Summary of circuit specifications for the SPDT switches. 47
Figure 32 Layout of the SPDT splits for inter-switch interference study: (a) Original switches (No ISO), (b) switches with BI-ring, (c) Die photo for switches using the new in-BEOL metal wall (Metal Wall), also same circuits after the trench-etching but before metal-filling (Air Trench). 48
Figure 33 SEM die photos for the Switch-A/B circuit pair: (a) Switch-A/B pair after trench-etching (20μm wide, 15μm deep), (b) Zoom-in for Switch-A/B showing the trench etched, but before metal filling (Air Trench split), (c) Zoom-in photo for the trench filled by nano silver powder (Metal Wall split). 49
Figure 34 HFSS-ADS co-simulation of the four switch circuit splits at sub-6GHz frequency. 50
Figure 35 Measured and simulated key circuit specs for the SPDT switch pair show satisfactory circuit performance for the T/Rx switch circuits. (a) Insertion loss for the SPDT switches from 1710 to 2155MHz. (b) Isolation for the SPDT switches from 1710 to 2155MHz. (c) SPDT power handling capability and good signal linearity given as the 1dB-compression point (P1dB) at 2GHz. 52
Figure 36 Crosstalk comparison for the four SPDT splits by (a) HFSS-ADS co-simulation, and (b) measured (0dBm incident signal to Input of Switch-A and interference monitored at Output of Switch-B). 52
Figure 37 ESD protection design window. 55
Figure 38 The categories of ESD events with equivalent circuits and models. 57
Figure 39 2000V HBM current waveform with its characteristics. 58
Figure 40 200V MM current waveform with its characteristics. 59
Figure 41 500V CDM current waveform with its characteristics. 60
Figure 42 Transmission line pulse (TLP) test bench. 61
Figure 43 Typical protection circuit topology – double diode network. 62
Figure 44 STI Diode X-section. 63
Figure 45 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow. 64
Figure 46 Gated Diode X-section. 64
Figure 47 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow. 65
Figure 48 Diode String X-section. 65
Figure 49 GGNMOS X-section. 66
Figure 50 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow. 66
Figure 51 SCR X-section. 67
Figure 52 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow. 68
Figure 53 DTSCR X-section and equivalent circuit. 68

Figure 54 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow. 69
Figure 55 N+PW STI Diode TLP measured I-V curve under different width. 70
Figure 56 N+PW Gated diode TLP measured I-V curve under different width. 71
Figure 57 Diode string TLP measured I-V curve under different diode numbers. 72
Figure 58 GGNMOS TLP measured I-V curve under different dimensions. 73
Figure 59 SCR TLP measured I-V curve under different width. 75
Figure 60 DTSCR TLP measured I-V curve under different trigger diode numbers. 76
Figure 61 Simplified CDM tester schematic and its ESD discharging waveform. 77
Figure 62 Measured I-V curves for 2DTSCR and SCR structures by VFTLP. 79
Figure 63 Comparison of ESD source pulse waveforms from CDM standard and VFTLP tester....79
Figure 64 TCAD simulated transient I-V curve for 2DTSCR and 3DS under CDM stressing....81
Figure 65 Cross-section views of ESD discharging current distribution across DTSCR by TCAD simulation shows the two-step DTSCR triggering mechanism: (a) current starts in the diodes (left) before triggering the SCR (right), (b) more current steers into SCR at its triggering threshold, and (c) most current conducts through the SCR after DTSCR fully turns-on. 82
Figure 66 Comparison of voltage transient waveform of DTSCR by testing and TCAD simulation under VFTLP. 83
Figure 67 Comparison of TCAD simulated transient volatge waveform for 2DTSCR under VFTLP and CDM pulse. 84
Figure 68 Cross-sections for N+PW (a) and P+NW (b) STI diodes with doping concentration by TCAD simulation. 85
Figure 69 The transient voltage waveforms by TCAD simulation for STI diodes under 125V CDM stressing. 86
Figure 70 Simulated transient electrical field distribution for N+PW STI diode. 87
Figure 71 The simulated x-section for a sample N+PW gated diode. 88
Figure 72 TCAD simulated voltage waveform for gated diodes under 125V CDM pulse. 88
Figure 73 Simulated transient electrical field distribution for a sample N+PW gated diode. 89
Figure 74 VFTLP-measured transient voltage and current waveforms for sample ESD diodes. 89
Figure 75 Cross-section of a sample 3-diode string structure with doping concentration.90
Figure 76 TCAD-simulated voltage waveform for diode string structures with different diode number. 90
Figure 77 VFTLP-measured transient voltage and current waveforms for sample ESD diode string. 91
Figure 78 TCAD-simulated cross-sections for sample 2-diode STI DTSCR (a) and 2-diode gated-diode DTSCR (b) ESD protection structures. 92
Figure 79 TCAD-simulated voltage waveforms for sample DTSCR ESD protection structures under 125V CDM ESD stressing. 93

Figure 80 VFTLP-measured voltage waveforms for sample DTSCR ESD protection structures. 94
Figure 81 De-embedded method patterns for stand-alone C_{ESD} extraction. 96
Figure 82 Equivalent circuit of de-embedded method with Y parameter calculation. 97
Figure 83 Extracted capacitance of N+PW STI diode from 10MHz to 40GHz. 97
Figure 84 Extracted capacitance of N+PW gated diode from 10MHz to 40GHz. 98
Figure 85 Extracted capacitance of N+PW diode string from 10MHz to 40GHz. 99
Figure 86 Extracted capacitance of GGNMOS from 10MHz to 40GHz. 100
Figure 87 Extracted capacitance of SCR from 10MHz to 40GHz. 101
Figure 88 Extracted capacitance of STI DTSCR from 10MHz to 40GHz. 102
Figure 89 Equivalent circuit models for new de-embedding C_{ESD} extraction methods. 104
Figure 90 HFSS 3D model of the “short” metal wire. 105
Figure 91 Comparison of C_{ESD_T} with C_{ESD_C} for ESD devices shows extraction errors for C_{ESD_T}. 105
Figure 92 Comparison of tested and simulated output period for a 37-stage ring oscillator using different width ESD devices. 106
Figure 93 Comparison of S_{11} between measurement and simulation for the input buffer (Upper) and output buffer (Lower) with a 120μm gated diode ESD device. 107
Figure 94 Series-shunt (a) and series-only (b) switch topologies feature stacked MOSFETs with output ESD protection. 109
Figure 95 Measured insertion loss and isolation for 28GHz switches with different ESD protection: no dedicated ESD protection (noESD), and 5DS and SCR ESD protection. 112
Figure 96 Measured insertion loss and isolation for 38GHz switches with different ESD protection: no dedicated ESD protection (noESD), and 5DS and SCR ESD protection. 112
Figure 97 TLP tested I-V curve of 28/38GHz switch self-protection. 114
Figure 98 TLP tested I-V curve of 28GHz (a) and 38GHz (b) switch with different ESD protection. 115

List of Tables

Table 2-1 User experience of data rate, latency and mobility under different user cases 13
Table 3-1 2X2 Switch Matrix Frequency bands and function 29
Table 3-2 Optimized design parameters of high/low bands switches. 33
Table 3-3 Interference in 2X2 switch array with different layout floorplans. 40
Table 5-1 Interference in 2X2 switch array with different layout floorplans. 61
Table 5-2 Critical parameters of STI diode structures. 71
Table 5-3 Critical parameters of gated diode structures. 72
Table 5-4 Critical parameters of diode string structures. 73
Table 5-5 Critical parameters of GGNMOS structures. 74
Table 5-6 Critical parameters of SCR structures. 75
Table 5-7 Critical parameters of DTSCR structures. 76
Table 6-1 De-embedded capacitance of STI diode at 10GHz..98
Table 6-2 De-embedded capacitance of gated diode at 10GHz. 99
Table 6-3 De-embedded capacitance of diode string at 10GHz....................................100
Table 6-4 De-embedded capacitance of GGNMOS at 10GHz. 101
Table 6-5 De-embedded capacitance of SCR at 10GHz. 102
Table 6-6 De-embedded capacitance of DTSCR at 10GHz. 102
Table 6-7 A summary for return loss baud rate masks of I/O buffers. 108
Table 6-8 Switch design specifications of 28GHz series-shunt SPDT and 38GHz series-only SPDT. 110
Table 6-9 Measured mm-wave switch RF & ESD performance with different ESD protection. 113

Chapter 1 Introduction

1.1 History of wireless communication

With continuous development in technological advancements, wireless communication, such as Wi-Fi, LTE, Bluetooth, GPS, etc., has made a tremendous progress from simple to complex systems that are able to communicate across multiple networks and devices. The concepts of massive multi-input multi-output (MIMO), internet of things (IoT), automatic drive, 5th generation of mobile communication (5G), etc., are becoming closer to everyone all over the world.

Radio frequency (RF) is any of the electromagnetic wave whose frequencies are lying in the range extending from around 3kHz to 300 GHz, which include those frequencies used for communications or radar signals. Every electrical engineer knows at least a bit about James Clerk Maxwell (Figure 1); he wrote those equations that made life extra busy back in sophomore year or thereabouts. Not only did he write the electrodynamic equations that bear his name, he also collected all that was then known about electromagnetic

$$\nabla \cdot \mathbf{D} = \rho \quad (1) \quad \text{Gauss' Law}$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2) \quad \text{Gauss' Law for magnetism}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (3) \quad \text{Faraday's Law}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \quad (4) \quad \text{Ampère-Maxwell Law}$$

Figure 1 Maxwell's Equations

phenomena and invented the displacement (capacitive) current term that allowed him to derive an equation that led to the prediction of electromagnetic wave propagation. Then between 1886 and 1889 came Heinrich Hertz, who was the first to verify experimentally Maxwell's prediction that electromagnetic waves exist and propagate with a finite velocity which truly indicates the born of radio frequency wave earlier than the vacuum tube, the first transistor. [1]

At present, communication using radio frequency dominates the wireless market. From the mobile telephone limited to phones installed in cars or other vehicles in 1973 to the first handheld mobile phone by Motorola weighted 1.1Kg with 23 cm long and works as a true "brick". Then in the 1990s, the "second generation" digital mobile phone systems emerged with Global System for Mobile communication (GSM) and Code Division Multiple Access (CDMA), during that time, talk and text are the major features. As the use of 2G phones became more widespread and people began to utilize mobile phones in their daily lives, the demand for data was growing for sharing pictures, browsing webpages, etc. However, the 2G technology was nowhere near up to the job, so the industry began to work on the next generation of technology known as 3G Wideband Code Division Multiple Access (WCDMA) and CDMA Evolution Data Optimized (EVDO), that enables the data rate up to 2 Mbit/s maximum data rate indoors, and in the mid-2000s, WCDMA evolved to high-Speed Packet Access (HSPA) with higher data rate up to 14Mbps. So far, the 4G Long-term Evolution (LTE) is most commonly used, which increases the capacity and internet speed using different radio interface together with core network improvements. More than 40 bands have been used mobile communication using frequency division

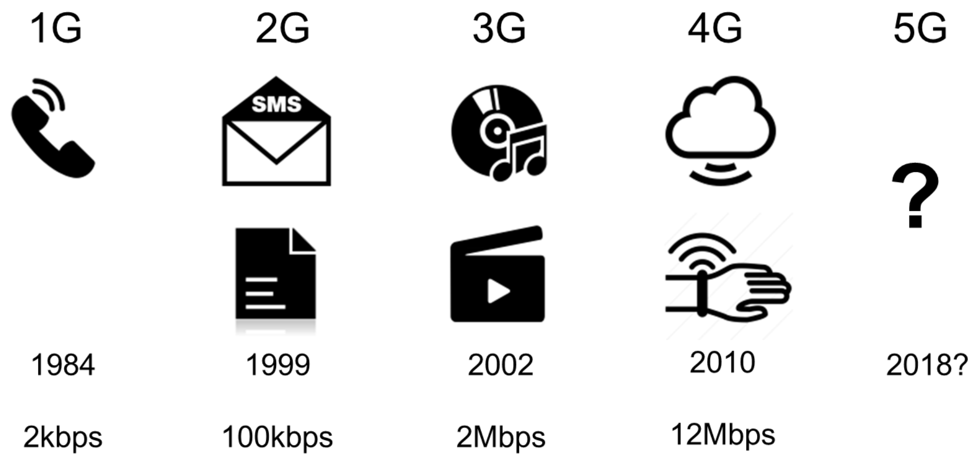


Figure 2 The development of mobile network generations (1G – 5G)

duplex (FDD) and time division duplex (TDD). But the demand from daily use of mobile internet, automotive and IoT require higher data rate and network capability which inspires the development of 5G. Though it's still in the phase of research, more and more details have been settled with the cooperation of worldwide companies and groups. With a steady stream of news trickling out – South Korea will deploy a 5G test network during the Winter Olympics, and AT&T will provide mobile 5G service in a dozen cities by the end of 2018 – it falls to 3rd Generation Partnership Project (3GPP) to deliver the technical blueprints that will make those promises a reality. Figure 2 shows the history of mobile communication from 1G to 5G.

The development of RF communication might be due to the galloping progress of integrated circuit (IC), which is a set of electronic circuits on one small flat pie of semiconductor material. ICs were made possible by experimental discoveries showing that semiconductor devices could perform the functions of vacuum tubes, and by mid-20th-centruy technology advancements in semiconductor device fabrication. Since their origins in the 1960s, the size, speed, and capacity of chips have progressed enormously, driven by

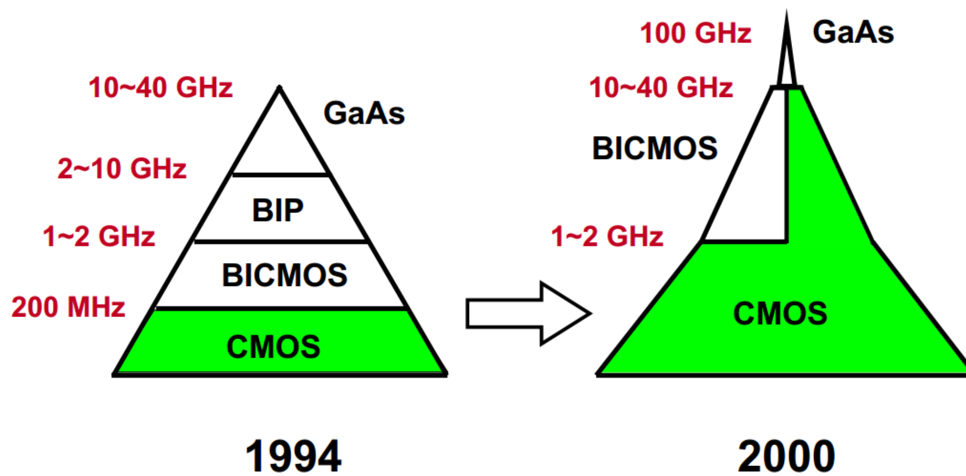


Figure 3 The frequency range of different technologies

technical advances that allow more and more transistors on chips of the same size – a modern chip may have several billion transistors in an area the size of a human fingernail. These advances, roughly following Moore's law, allow a computer chip of today to have millions of times the capacity and thousands of times the speed of the computer chips of the early 1970s. [2]

The complementary metal-oxide-semiconductor (CMOS) devices, although the bipolar transistor is historically first. Nowadays the number of CMOS transistors integrated on chips, vastly outnumber the bipolar ones. Indeed, previously CMOS devices were reserved for logic as they offer the highest density (in gates/mm^2). Most high-frequency circuitry was carried out in bipolar technology. As a result, a lot of analog functions were realized in bipolar technology. The highest-frequency circuits have been realized in exotic technologies such as GaAs and now InP technologies. The channel length of CMOS transistors shrinks continuously however. In 2004, a channel length of 0.13um is standard and 28nm CMOS technology has been widely used all over the world. 14nm even 7nm fin

field-effect-transistor (FinFET) was coined several years ago and Moore's law is continuing. This ever-decreasing channel length gives rise to ever increasing speeds. As a result, CMOS devices are capable of gain at ever higher frequencies.

Today CMOS and bipolar technologies are in competition over a wide frequency region, extending all the way to 10 and even 40 GHz, as predicted in Figure 2. BiCMOS which contains both CMOS and bipolar technologies is expensive but keep growing to fulfill the system and circuit requirements of RFICs. And the GaAs process achieves better power handling capability for high power RF applications when reaching 60 GHz which may be used as 5G band.

1.2 Organization of the Thesis

This dissertation is organized as follows. First, the history of wireless communication as an introduction is in Chapter 1. In Chapter 2, the RF front end (RFFE) module and RF switch with details about 5G requirements and applications will be introduced together with the 45nm SOI technology which is the mainly used process in this thesis. Then a switch array design as a 5G application for better understanding the importance of interference is dedicated in Chapter 3. To solve the consequent interference between RF switches, Chapter 4 reveals a solution of above-Si through Back-end-of-line (BEOL) metal wall which suppress the crosstalk a lot. Chapter 5 describes a comprehensive study of ESD consideration with different protection model, Human Body Model (HBM) and Charged Device Model (CDM). Overshoot issue induced from diode-based structures under CDM ultra-fast pulse is analyzed with TCAD simulation and very fast transmission line pulse (VFTLP) measurements. The parasitic effects induced by ESD protections severely

damage the performance of RFICs especially at high frequency, high data rate circumstances. This results a discussion about an enhance method of accurate parasitic extraction and parasitic effects to mm-wave RFICs in Chapter 6, followed by the conclusion.

Chapter 2 Background of 5G RFICs

2.1 What is 5G?

5th generation mobile networks, abbreviated 5G, are the proposed next telecommunications standards beyond the current 4G/IMT-advanced standards aiming at higher capacity, density of mobile broadband users, and supporting device-to-device, ultra-reliable, and massive machine communications. The technologies to be used in 5G are still being defined, but there are many details on which everyone agrees. Lots of companies, research institutions and even governments are cooperating and pushing forward the research and development of 5G standards and technologies. [3]–[7]

5G network will use a type of encoding called Orthogonal Frequency Division Multiplexing (OFDM), which is similar to the encoding that LTE uses. The air interface will be designed for much lower latency and greater flexibility than LTE, though.

The new networks will predominantly use very high frequencies that can transmit huge amounts of data, but only a few blocks at a time. The standard will work all the way from low frequencies to high, but it gets the most benefit over 4G at higher frequencies. 5G may also transmit data over the unlicensed frequencies currently used for Wi-Fi, without conflicting with existing Wi-Fi networks.

5G networks are much more likely to be networks of small cells, even down to the size of home routers, than to be huge towers radiating great distances. Some of that is because of the nature of the frequencies used, but a lot of that is to expand network capacity. So 5G networks need to be much smarter than previous systems, as they're juggling many more, smaller cells that can change size and shapes. But even with existing macro cells, 5G will

be able to boost capacity by four times over current systems by leveraging wider bandwidths and advanced antenna technologies. Here are some use cases and requirements mentioned in NGMN 5G white paper [4].

2.1.1 Use Cases

The user cases are more complicated in 5G technology which can be divided into 8 groups showed in Figure 4.

1) Broadband Access in Dense Areas

This family highlights the broad range of growing and new use cases of the fully connected society. The focus is service availability in densely-populated areas (e.g., multi-store buildings, dense urban city centers or events), where thousands of people per square kilometer (km^2) live and/or work. Communications are expected to be pervasive and part of everyday life. Augmented reality, multi-user interaction, three-dimensional (3D) services will be among the services which play an increasingly significant role in the 2020+

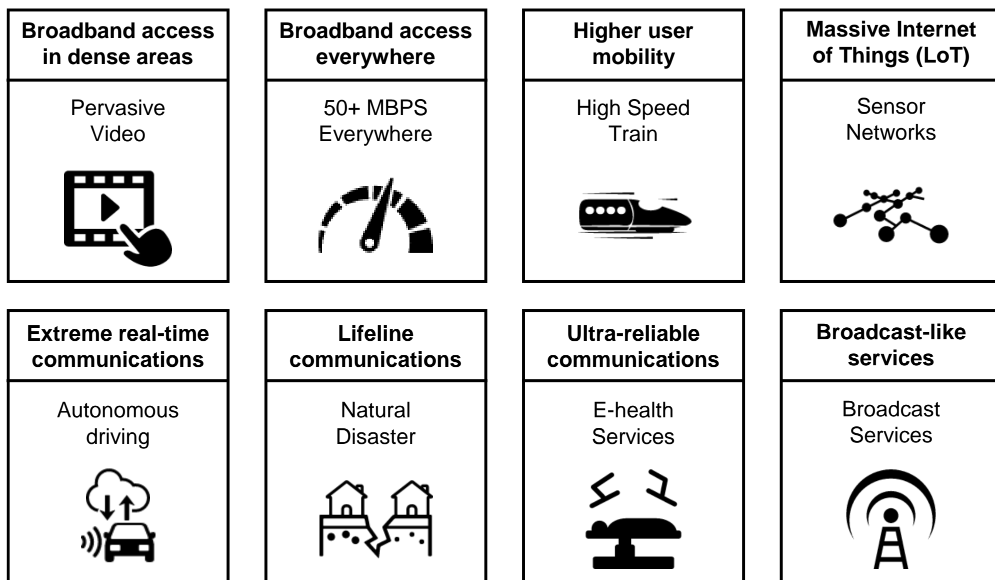


Figure 4 5th generation mobile networks user cases

timeframe. Context recognition will be an essential aspect, at the network edge (i.e. close to the user), ensuring delivery of consistent and personalized services to the customers.

This family includes the following use cases: pervasive video, smart office, operator cloud services, photo sharing in stadium and open-air gathering.

2) Broadband Access Everywhere

This family highlights the need to provide access to broadband service everywhere, including the more challenging situations in terms of coverage (from urban to suburban and rural areas). A consistent user experience with respect to throughput needs a minimum data rate guaranteed everywhere. Further development of digital inclusion of people living in scarcely populated areas and in developing countries requires the infrastructure deployment cost to be a key factor in services.

This family includes the following use cases: 50+ Mbps everywhere and ultra-low-cost networks.

3) Higher User Mobility

Beyond 2020, there will be a growing demand for mobile services in vehicles, trains and even aircrafts. While some services are the natural evolution of the existing ones (navigation, entertainment, etc.), some others represent completely new scenarios such as broadband communication services on commercial aircrafts (e.g., by a hub on board). Vehicles will demand enhanced connectivity for in-vehicle entertainment, accessing the internet, enhanced navigation through instant and real-time information, autonomous driving, safety and vehicle diagnostics. The degree of mobility required (i.e. speed) will depend upon the specific use case.

This family includes the following use cases: high speed train, remote computing, moving hot spots and 3D aircrafts connectivity.

4) Massive Internet of Things (IoT)

The vision of 2020 and beyond also includes a great deal of growing use cases with massive number of devices (e.g., sensors, actuators and cameras) with a wide range of characteristics and demands. This family will include both low-cost/long-range/low-power MTC as well as broadband MTC with some characteristics closer to human-type communication (HTC).

This family includes the following use cases: smart wearables (clothes), sensor networks and mobile video surveillance.

5) Extreme Real-time Communications

This family covers use cases which have a strong demand in terms of real-time interaction. These demands are use-case specific and, for instance, may require one or more attributes such as extremely high throughput, mobility, critical reliability, etc. For example, the autonomous driving use case that requires ultra-reliable communication may also require immediate reaction (based on real-time interaction), to prevent road accidents. Others such as remote computing, with stringent latency requirement, may need robust communication links with high availability.

Tactile internet is a typical use case in this family.

6) Lifeline Communication

Public safety and emergency services that are provided today are continuously improving. In addition to new capabilities for authority-to-citizen and citizen-to-authority communication for alerting and support, these use cases will evolve to include emerging and new applications for authority-to-authority communication, emergency prediction and disaster relief. Furthermore, there will be an expectation that the mobile network acts as a lifeline, in all situations including times of a more general emergency. Therefore, the use cases require a very high level of availability in addition to the ability to support traffic surges.

Natural disaster is one of the use cases in this family.

7) Ultra-reliable Communications

The vision of 2020 and beyond suggests not only significant growth in such areas as automotive, health and assisted living applications, but a new world in which the industries from manufacturing to agriculture rely on reliable MTC. Other applications may involve significant growth in remote operation and control that will require extreme low latency as well (e.g., enterprise services or critical infrastructure services such as Smart Grid). Many of these will have zero to low mobility.

This family includes many use cases: automated traffic control and driving, collaborative robots, eHealth, remote object manipulation, 3D drone connectivity and public safety.

8) Broadcast-like Services

While personalization of communication will lead to a reducing demand for legacy broadcast as deployed today, e.g. linear TV, the fully mobile and connected society will

nonetheless need efficient distribution of information from one source to many destinations. These services may distribute content as done today (typically only downlink), but also provide a feedback channel (uplink) for interactive services or acknowledgement information. Both, real-time or non-real time services should be possible. Furthermore, such services are well suited to accommodate vertical industries' needs. These services are characterized by having a wide distribution which can be either geo-location focused or address-space focused (many end-users).

This family includes the following use cases: news and information, local/regional/national broadcast-like services.

2.1.2 5G Requirements

The 5G vision and user cases above result lots of requirements like faster user experience, better system performance and massive devices connections.

1) User Experience

User experience requirements address the end user's experience when consuming one or more services. User experience will have to be managed in highly heterogeneous environments and under different user scenarios/contexts.

- Consistent User Experience

The 5G system should be able to deliver a consistent user experience over time for a given service everywhere the service is offered. Consistent user experience is defined by service-dependent minimum KPIs (e.g. data rate, latency) being met over the service coverage area, with a level of variation configurable by the operator. These service-dependent KPIs are for further study and evaluation, within the NGMN program and

elsewhere across the ecosystem, as appropriate. A consistent user experience across time and space depends obviously on the technology performance and capabilities, and on the operator deployment. The requirements address only the technology performance and capabilities.

- User Experienced Data Rate

Data rate requirements are expressed in terms of user experienced data rate, measured in bit/s at the application layer. The required user experienced data rate should be available in at least 95% of the locations (including at the cell-edge) for at least 95% of the time within the considered environment. The user experienced data rate requirement depends on the targeted application/use case. It is set as the minimum user experienced data rate required for the user to get a quality experience of the targeted application/use case.

Use case specific user experienced data rates up to 1 Gb/s should be supported in some specific environments, like indoor offices, while at least 50 Mb/s shall be available everywhere cost effectively. Use case specific user experienced data rate requirements are specified in Table 2-1.

Use Case	Data Rate	Latency (E2E)	Mobility
Broadband access in dense areas	DL: 300 Mbps UL: 50 Mbps	10 ms	0 - 100 km/h
Indoor ultra-high broadband access	DL: 1 Gbps UL: 500 Mbps	10 ms	Pedestrain
50+ Mbps everywhere	DL: 50 Mbps UL: 25 Mbps	10 ms	0 - 500 km/h
Airplanes Connectivity	DL: 15 Mbps UL: 7.5 Mbps	10 ms	Up to 1000 km/h
Ultra-low latency	DL: 50 Mbps UL: 25 Mbps	<1 ms	Pedestrain
Broadcast like services	DL: Up to 200 Mbps UL: Modest (e.g. 500 kbps)	<100 ms	0 - 500 km/h

Table 2-1 User experience of data rate, latency and mobility under different user cases

- Latency

When considering latency requirements, the following metrics are considered:

End to End (E2E) Latency: Measures the duration between the transmission of a small data packet from the application layer at the source node and the successful reception at the application layer at the destination node plus the equivalent time needed to carry the response back.

User Plane Latency: Measures the time it takes to transfer a small data packet from user terminal to the Layer 2 / Layer 3 interface of the 5G system destination node, plus the equivalent time needed to carry the response back.

The E2E latency is the latency perceived by the end user. It accounts for the time needed for the data packet to cross all the nodes up to the application server and back, which includes nodes of the 5G system and nodes potentially outside the 5G system. In contrast, the user plane latency is limited to the 5G system only. Both latency metrics approximately coincide when the application server is located within the 5G system. In the latter case, the latency is minimized when the application server is co- located with a radio node, e.g., the radio base station or another user terminal (for the case of device-to- device, D2D, communication). As a result, the requirements on minimum latency are expressed in terms of E2E latency.

The 5G system should be able to provide 10ms E2E latency in general and 1ms E2E latency for the use cases which require extremely low latency. Note these latency targets assume the application layer processing time is negligible to the delay introduced by

transport and switching. Use case specific E2E latency requirements are specified in Table 2-1.

The 5G system should also give the end user the perception of being always connected. The establishment of the initial access to the network (or status change from idle state to connected) should then be instantaneous from the end user perspective.

- Mobility

Mobility refers to the system's ability to provide seamless service experience to users that are moving. In addition to mobile users, the identified 5G use cases show that 5G networks will have to support an increasingly large segment of static and nomadic users/devices. 5G solutions therefore should not assume mobility support for all devices and services but rather provide mobility on demand only to those devices and services that need it. In other words, mobility on demand should be supported, ranging from very high mobility, such as high-speed trains/airplanes, to low mobility or stationary devices such as smart meters.

The mobility requirements are expressed in terms of the relative speed between the user and the network edge, at which consistent user experience should be ensured (see Consistent User Experience requirement). Use case specific mobility requirements are specified in Table 2-1.

2) System Performance

System performance requirements define the system capabilities needed to satisfy the variety and variability of users and use cases.

- Connection Density

Up to several hundred thousand simultaneous active connections per square kilometre shall be supported for massive sensor deployments. Here, active means the devices are exchanging data with the network. Note this KPI assumes a single operator in the considered area.

- Traffic Density

The 5G network should be able to serve massive number of HTC and MTC devices. In the extreme cases:

a) Data rates of several tens of Mb/s should be supported for tens of thousands of users in crowded areas, such as stadiums or open-air festivals.

b) 1Gb/s to be offered simultaneously to tens of workers in the same office floor.

Traffic Density measured in bit/s/m2 is defined as the total amount of traffic exchanged by all devices over the considered area. The KPI requirement on the minimum Traffic Volume Density / Areal Capacity for a given use case is given by the product: [required user experienced data rate] x [required connection density]. For the sake of defining this KPI, a single operator is considered in the considered area.

- Spectrum Efficiency

Spectrum efficiency should be significantly enhanced compared to 4G in order for the operators to sustain such huge traffic demands under spectrum constraints, while keeping the number of sites reasonable. Spectrum efficiency improvements should apply in both small and wide area cells, in both low and high frequency bands, in both high and low mobility scenarios.

In particular the average spectrum efficiency (measured in bit/s/Hz/cell) and the cell-edge spectrum efficiency (measured in bit/s/Hz/user) should be improved.

- Coverage

The 5G technology should allow the data rates requirements to be achieved in rural areas with only the current grid of macro sites. The coverage requirement for other environments is for further study, within the NGMN program, and elsewhere, as appropriate.

- Resource and Signaling Efficiency

Signaling efficiency should be enhanced, so that the related radio resource and energy consumption are minimized and justified by the application needs. More specifically, network function specific signaling should only be transmitted when needed. In this context, user end (UE) capability handling should also be designed for network flexibility and scalability.

For certain IoT/MTC applications, additional measures should be considered to avoid a surge by volume in case a large number of devices attempt to access the network simultaneously.

3) Device Requirements

Smart devices in the 5G era will grow in capability and complexity as both the hardware and software, and particularly the operating system will continue to evolve. They may also in some cases become active relays to other devices, or support network-controlled device-to-device communication.

- Operator Control Capabilities on Devices

5G terminals should have a high degree of programmability and configurability by the network, for example in terms of terminal capabilities, access technology used, transport protocol used and certain lower layer functions (e.g. error control schemes). This will enable efficient logical division for different services (slicing) while removing dependency on terminal type.

In particular, flexible and dynamic UE capability handling should be assured. This would allow the network or the UE to choose one of the profiles depending on quality of service (QoS) needs, radio node capability and/or radio conditions.

The 5G devices should provide the capability to operators to check the hardware and software platform configuration over the air, the capability to update the smart device's operating system over the air, and the ability to diagnose the malfunction of devices or malware in smart device plus the ability to fix the problems or update device software that affect end user experience or overall network performance.

Operators shall be able to retrieve network as well as service-related performance data (e.g., voice call drops, handover failure, network registration failure, and instantaneous throughput) from the UE in order to collect information on real-life operation and use them as an input for service experience optimization and customer care.

- Multi-Band-Multi-Mode Support in Devices

To enable true global roaming capability, smart devices should be able to support multiple bands as well as multiple modes (TDD/FDD/mixed). Note that IoT/MTC devices which are stationary may not require multiple bands/modes.

Furthermore, to achieve the high data rates, devices should be able to use multiple bands simultaneously, without impacting the single band performance or network performance. 5G terminals shall support aggregation of data flows from different technologies and carriers.

- Device Power Efficiency

Battery life shall be significantly increased: at least 3 days for a smartphone, and up to 15 years for a low-cost MTC device.

- Resource and Signaling Efficiency

At the device side, the resource and signaling efficiency requirement is even more crucial as frequent signaling has a significant impact on the battery life.

2.2 Front End Introduction

2.2.1 Radio Frequency Front End (RFFE) Module Basic

All RF signal transceivers including Wi-Fi, Bluetooth, Cellular and GPS needs a front-end module which is a generic term for all the circuitry between the antenna up to and

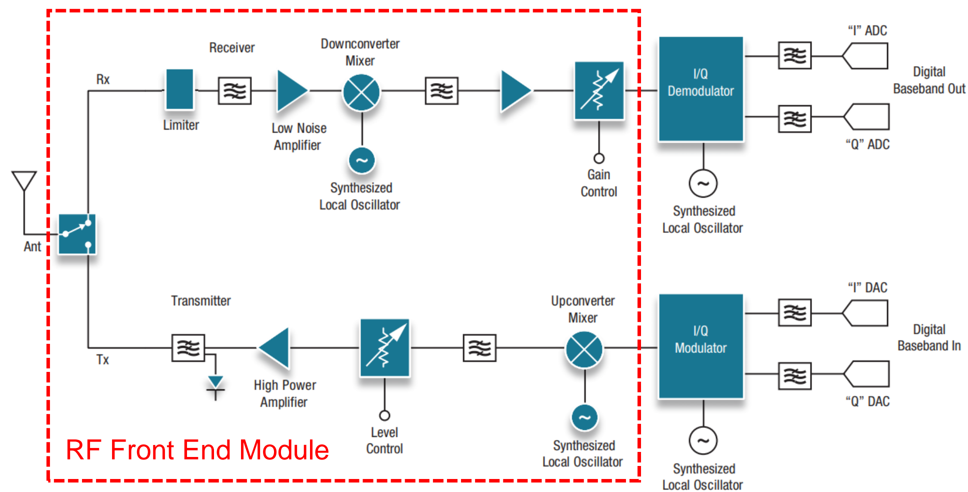


Figure 5 RF front end module of mobile phones (simplified)

sometimes including the mixer stage. It consists of all components in the receiver that process the signal at the original incoming modulated RF signal until demodulator and in the transmitter that process the modulated intermediate frequency (IF) to antenna(s). Figure 5 shows the simplified cellular transceiver including the RFFE which contains the basic components such as switch, filter, low noise amplifier, power amplifier, mixer and so on.

Reaching to 4G network, the RFFE module has been more complicate with massive bands and diversity antennas. An example of multi-bands multi-modes transceiver with diversity antenna is revealed in Figure 6 which is already widely used in normal cell phones. Even the WiFi front end also has more components due to the higher requirements

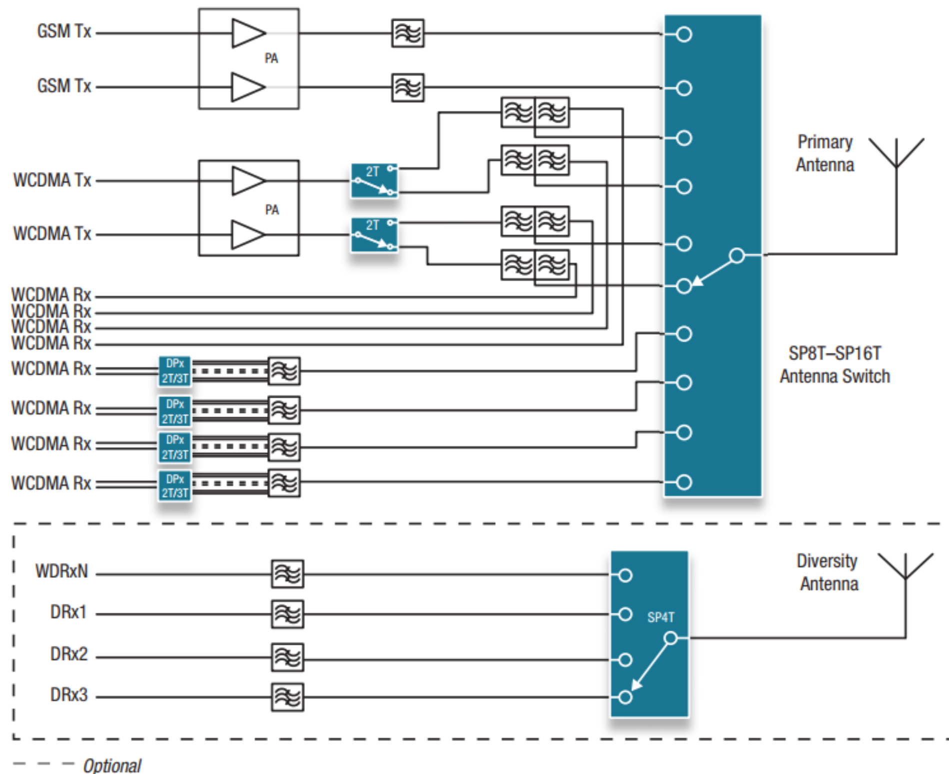


Figure 6 RFFR module of Smartphone using discrete switches or Antenna Switch Modules (ASMs)

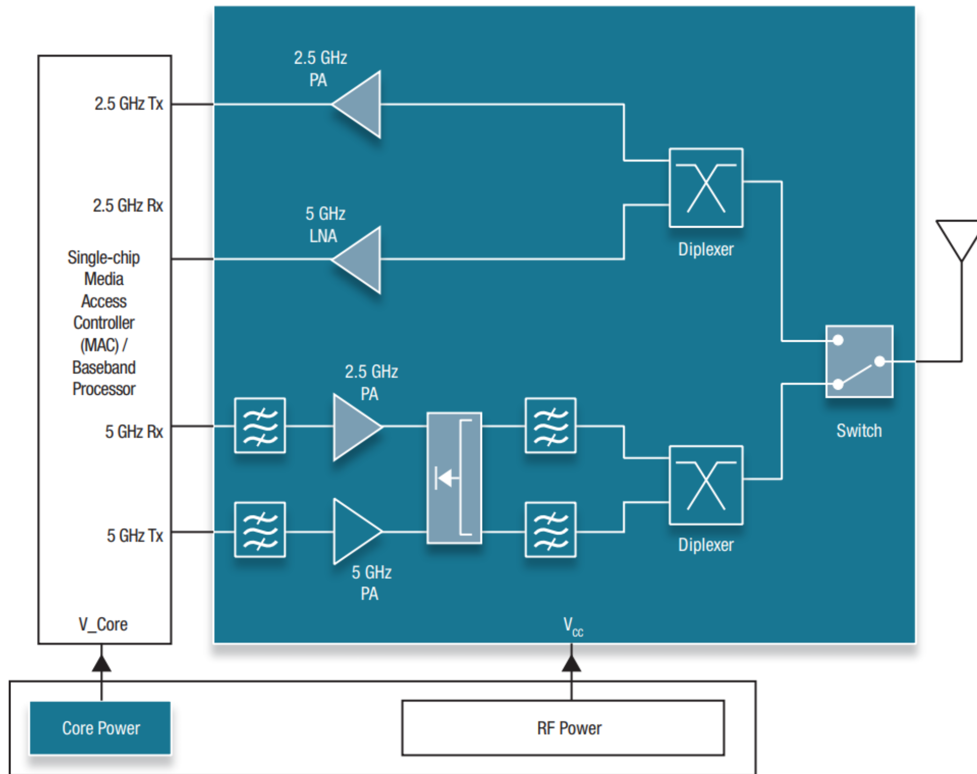


Figure 7 802.11 a/b/g/n/ac dual-band WiFi front end components

of network speed and higher frequency bands. Figure 7 indicates a front-end module of dual-band WiFi for 802.11 a/b/g/n/ac standards.

As we can see, though the basic topology of RFFE module doesn't change a lot, due to the development of process and filter technology, more complex structure covering multiple bands will be used for future application. The multi antennas front end required by MIMO application have forced more RFFE modules in a single chip.

2.2.2 5G RFFE Vision

5G mobile communications will need to accommodate huge traffic demands in the near future. Massive MIMO technology utilizing hundreds of antenna elements has drawn attention as a key antenna configuration for envisioned 5G applications. Realizing the

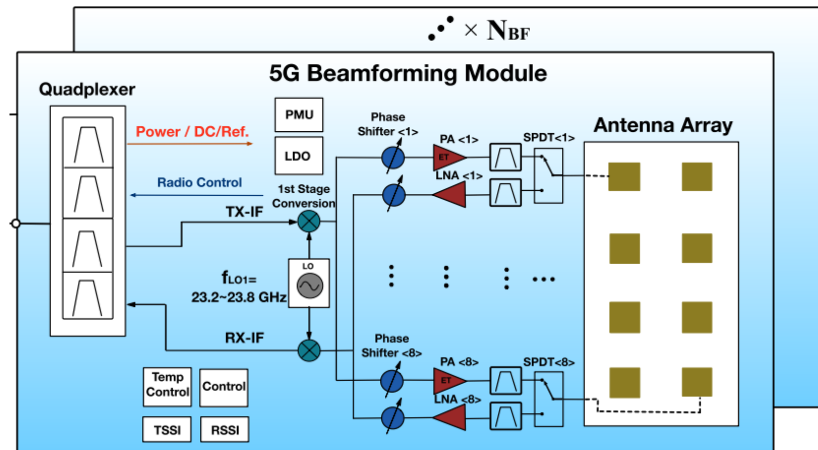


Figure 8 Block diagram of 5G user equipment front end wireless system architecture.

massive multi-input-multi-output (MIMO) concept of active phased-array antennas (APAAs) for 5G will require small-size, low-power-consumption, and highly accurate phase control over the wide-hand frequency range, which poses significant challenges for the RF front end [8]. As one of the most important block, RFFE requires more devices and terminals capability which also needs higher speed. More bands and channels are needed to achieve it which results more complicate front-end. Massive MIMO with antenna array and multi-band-multi-mode (MBMM) with a lot more branches in a single device even in a single chip are predictable. Although, the final standards of 5G is not settled, some topics can be expected such as wideband power amplifier (PA) and low noise amplifier (LNA), multi-pole-multi-threw switch, diplexer, duplexer and massive RF path integration like the RFFE 5G beamforming module in Figure 8 [9].

What's more, higher frequency should be exploited. In some researches recently [7], 28GHz and 38GHz bands may be a great choice for 5G application and the U.S. Federal Communications Commission (FCC) approved the spectrum for 5G, including the 28GHz, 37 GHz and 39 GHz bands showed in Figure 9, on July 14, 2016 [10]–[12]. From Figure

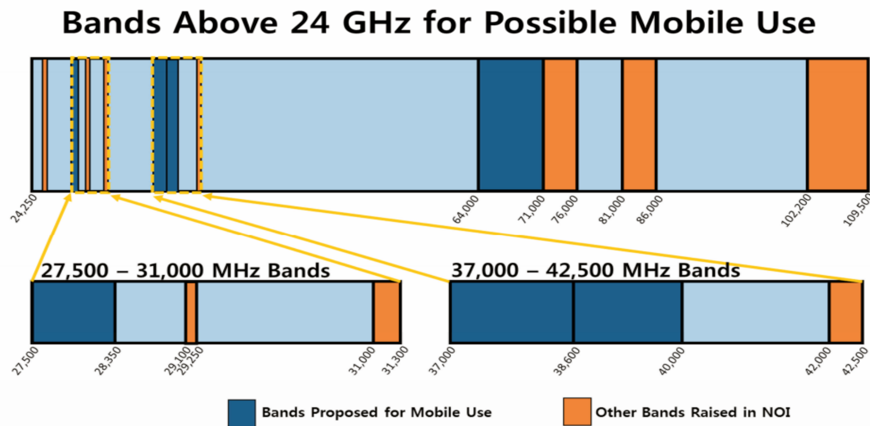


Figure 9 FFC approved mm-wave bands for 5G research.

10, we could see that the rain attenuation and atmospheric absorption are not bad as other mm-wave frequency and those may be a good choice for 5G transmission. Recent studies suggest that mm-wave frequencies could be used to augment the currently saturated 700MHz to 2.6GHz radio spectrum bands for wireless communications with a much higher data rate. 5GHz bands has been used for Wi-Fi communication already together with 2x2 MIMO routers which includes two or more antennas inside. With the help of multi-antennas structure, beam forming technology which achieves better signal strength at some specific spots gives the user end a higher data speed with same receiving device. The combination of cost-effective CMOS technology that can now operate well into the mm-wave frequency bands, and high-gain, steerable antennas at the mobile and base station, strengthens the viability of mm-wave wireless communications. Further, mm-wave carrier frequencies allow for larger bandwidth allocations, which translate directly to higher data transfer rates. Mm-wave spectrum would allow service providers to significantly expand the channel bandwidth far beyond the present 20MHz channels used by 4G customers. By increasing the RF channel bandwidth for mobile radio channels, the data capacity is greatly

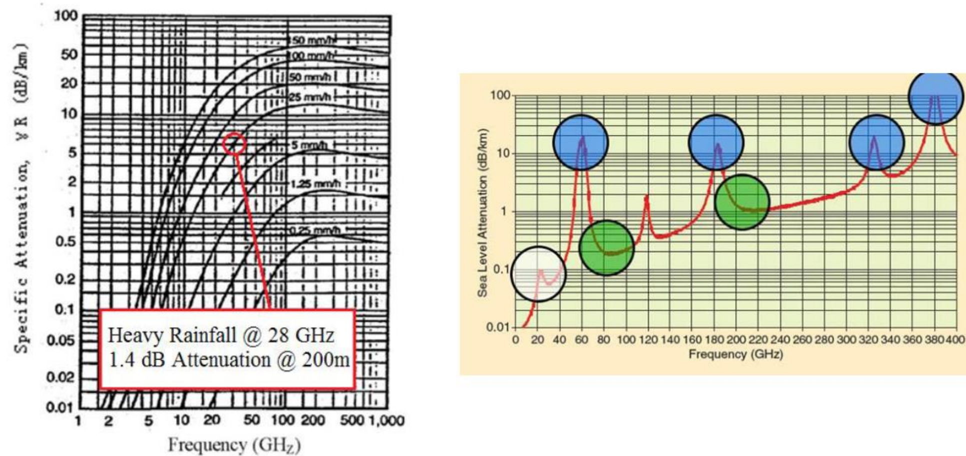


Figure 10 mm-wave attenuation in rain (left) and atmosphere (right).

increased, while the latency for digital traffic is greatly decreased, thus supporting much better internet-based access and applications that require minimal latency such as automotive. Mm-wave frequencies, due to the much smaller wavelength, may exploit polarization and new spatial processing techniques, such as massive MIMO and adaptive beamforming. The integration will be easier due to the smaller size of antennas.

As a result, for the blocks in front-end, working at mm-wave frequency and multi-blocks in a single die will be a challenge to achieve 5G requirements. How to deal with the sensitive capacitance and inductance under high frequency level and how to achieve multi-blocks without interference between them will be important. Nowadays, with the blowout of 5G developments, researches are paying more attention to solve those issues with different design topologies and different materials [13].

2.3 RF Switch for 5G Application

As we know, an RF switch is a device to route high frequency signals through transmission path. It’s an important block inside all RFFE for the use of antenna switch which change the path from antenna to different transceiver, band switch which change the

signal from one band to another and transmitter/receiver (T/Rx) switch which change the direction of signal. Different application has different requirements. For antenna switch, power handling capability is the toughest parameter with considering of ESD pulse from antenna end. But as band switch, signal pole multiple throw (SPnT) (n sometime is >10) allow a single antenna to cover all bands for different carriers. The T/Rx switch on the other hand requires higher switching speed which can handle the data switching rate of time division duplex (TDD). RF switches can be simply categorized into two equally mainstream and essential groups: electromechanical switches and solid-state switches. Electromechanical switches are based on the simple theory of electromagnetic induction. They rely on mechanical contacts as their switching mechanism. A solid-state switch is an electronic switching device based on semiconductor technology (e. g. MOSFET, PIN diode). It functions similarly to an electromechanical switch except that it has no moving parts. The advantages of solid state switch using MOSFET for RF application are evident with its high speed and compatibility with integrated circuits.

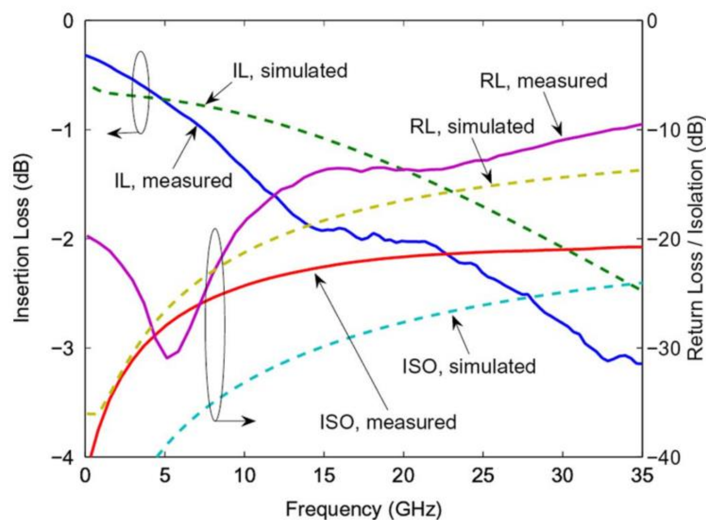


Figure 11 Recent research shows great isolation, insertion loss and return loss are hard to achieve at mm-wave frequency.

When reaching 5G technology like massive MIMO/MBMM applications, the switch design is more critical due to the interference between each branch. This directly effects the most important performance like insertion loss (IL) and isolation (Iso) which shows the ratio of output power and input power with ON (IL) and OFF (Iso) state. To handle the high power on each branch side, a more efficient structure is needed to reduce the size and cost of each switch.

What's more, higher frequency range up to mm-wave bands are also becoming a challenge to the topology of switch structures and performance. Normal series-shunt switch can only handle the frequency lower than 10GHz, poor insertion loss, return loss and isolation will be found at that mm-wave range. How to meet the same or even more specifications as 4G LTE standard under high frequency bands are still under research. Figure 11 shows a recent research about RF switch at 28GHz band [14].

2.4 45nm silicon-on-insulator (SOI) technology

Global Foundries 45nm SOI CMOS technology delivers outstanding power & performance for a broad range of applications [15]. With their 6th generation of leading-edge SOI technology, the significant transistor performance especially the lower junction capacitance and power improvement over traditional bulk technology are achieved. There are several advantages associated with the elimination of junction isolation that is no N or P wells. For the partially depleted floating body device there isn't a path for the leakage current. Additionally, the elimination of the N-well removes the N-well proximity effect from the FET models and the shallow STI reduces the stress effects. Last but not the least, with SOI, the floating body voltage is a function of applied S/D voltages and prior

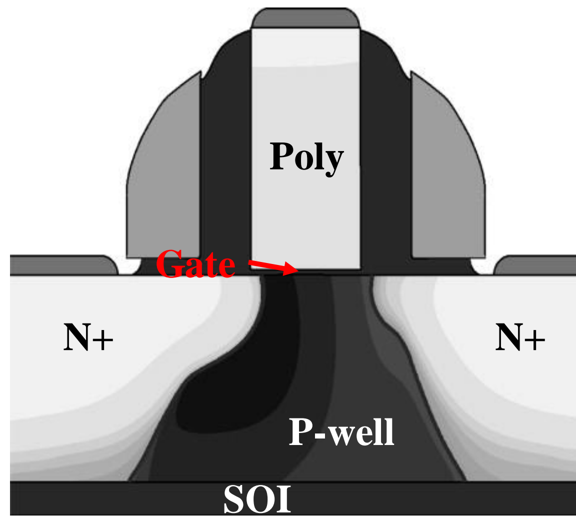


Figure 12 Simulated doping profile of a 45nm asymmetric (AFET) NMOS device.

operation, this can be used to advantage to reduce switching time and reduce the voltage drop across pass gate and source follower circuits. Compared to a bulk technology, there are several benefits of SOI such as the reduction of the S/D capacitance which improves the technologies power performance. And bulk analog circuits share the same substrate and there will be coupling through the substrate as an interference. This can be minimized by region substrate contacts and high resistivity silicon, but in SOI this coupling is virtually eliminated by the isolation provided by the BOX and shallow trench isolation (STI). A typical NMOS cross-section is showed in figure 12.

This leading-edge SOI technology gives a potential application for ultra-high frequency, low cross-talk RFIC design especially for RF switch due to its high-speed response, good power capability, reduced capacitance and great isolation techniques.

Chapter 3 RF Switch Array

Wireless revolution is around the corner. 5G cellular technology will ensure ultra-high data rates and broad coverage over hot spots at lower latency, power and costs as mentioned above [16]. And internet of things (IoT) requires massive wireless networks to link everything together each of which needs RF blocks to communicate. The corner-stone wireless technologies (e.g., 4G/5G, WiFi, GPS) involve plenty frequency bands that are aggregated together such as 4G/LTE uses 70 bands covering 450MHz to 6GHz. On another word, many-band RF chips involving aggregation, noise/interference coupling among different circuit blocks is a major challenge in RF IC designs. For example, a MIMO network uses many RF switches to select antennas, Tx/Rx circuits and frequency bands [17]. An integrated massive switch structure like an array is a key RF IC block for to wireless aggregation where inter/inner-band interferences are more severe that must be addressed using various noise isolation techniques. For RFICs design, SOI CMOS is superior to bulk CMOS due to its unique features, e.g., lower parasitic and better noise isolation [18], [19]. For RF switch design, a series-shunt topology is commonly used where the series and shunt MOSFETs dominate the insertion loss and isolation, respectively [20]. The trade-off of width and capacitance to reach better performance is frequency dependent and hard to estimate. Circuit techniques, e.g., stacked FETs [21]–[23], gate resistance [24], W/L optimization [25] and FFC [26]–[28], are used to enhance insertion loss, isolation and linearity. Other noise suppression techniques are substrate isolation and layout optimization including better layout floor planning. However, these technologies or design

<table>
<tr><th>Structure</th><th colspan="6">Bands & Frequency (MHz)</th><th>Function</th></tr>
<tr><td rowspan="3">SPDT</td><td colspan="3">Band 4</td><td colspan="3">Band 25 (2)</td><td rowspan="3">T/Rx Switch</td></tr>
<tr><td colspan="3">2110-2155</td><td colspan="3">1930-1995</td></tr>
<tr><td colspan="3">1710-1755</td><td colspan="3">1850-1915</td></tr>
<tr><td rowspan="3">SP3T</td><td colspan="2">Band 12 (17)</td><td colspan="2">Band 13</td><td colspan="2">Band 26 (5)</td><td rowspan="3">Band Switch</td></tr>
<tr><td colspan="2">729-746</td><td colspan="2">746-756</td><td colspan="2">859-894</td></tr>
<tr><td colspan="2">699-716</td><td colspan="2">777-787</td><td colspan="2">814-849</td></tr>
</table>

Table 3-1 2X2 Switch Matrix Frequency bands and function

based interference suppression techniques are not sufficient for advanced RF ICs of high complexity and integration involving large number of frequency bands.

This chapter presents a study of interferences inside an RF switch array, aiming to understand the design influences on interference characteristics. The 2x2 single-pole double/triple-throw (SP2T/SP3T) Tx/Rx band switch array, featuring a series-shunt topology with gate resistance and feed-forward capacitance (FFC) and covering low (699-894MHz) and high (1710-2155MHz) bands showed in Table 3-1, was designed and fabricated in a 45nm SOI CMOS. The inter-band and inner-band interferences were characterized, which reveals that existing noise isolation techniques, e.g., substrate isolation and layout floor planning, are insufficient for interference reduction. It therefore calls for novel in-die interference elimination techniques [29].

3.1 SPnT RF Switch Design

The series-shunt structure is the most widely used the topology for RF switch design due to its wideband balance of insertion loss and isolation. The resistance of each FET branch mainly depends on its width. The wider the transistors, the smaller the resistance but in the same time higher capacitance from source to drain. So, for the branch which is ON, a small resistance of series FET and high impedance of shunt FET are needed to

achieve smaller insertion loss. As a result, we will make the series ones wider than the shunt ones. However, for the branch which is OFF, a high impedance with small capacitance of series FET and low resistance of shunt FET which connects output and ground are required to get a better isolation at OFF output side. At this moment, the series part needs smaller width on the contrary. Now, how to balance the performance of insertion loss and isolation to meet the specifications depends on the optimization of width.

Figure 13 shows a basic structure of MOSFET including five parts of capacitance: C_{gs}, C_{gd}, C_{sb}, C_{gb} and C_{db}. The name of capacitor is based on its connection of source (s), drain (d), gate (g) and body (b). In my case, the SOI technology remove the body part replaced by BOX which is an insulator. So, the parasitic capacitance can be simplified to C_{gs} and C_{gd}. The value of each capacitors is related to the thickness of oxide, the length and width of gate and body process. And all those capacitors can pass AC signal from one side to another. The higher the frequency, the less the impedance. So, let's assume the control gate of an OFF-state transistor is connected to ground. At 2GHz RF frequency, the signal will quickly leak to ground from the drain side through the gate capacitance.

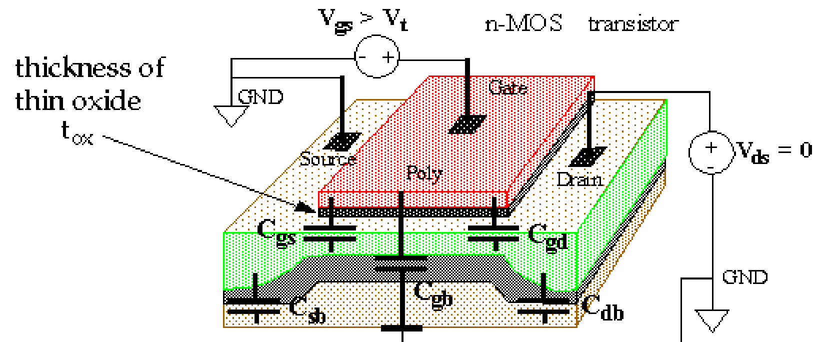


Figure 13 Basic parasitic capacitance from a simple MOSFET.

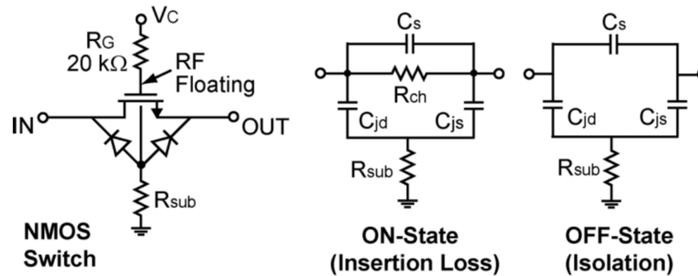


Figure 14 Basic model of simple NMOS switch during ON/OFF-state

As a result, a very basic technique called high gate resistance which can make the gate side AC floating. The big resistor can block the AC signal from drain side to DC control side which looks like a ground for AC signal. But the DC voltage can still be delivered to the gate due to the almost infinity gate resistance. With C_{gs} and C_{ds}, the AC signal voltage will almost evenly drop from drain, gate and source. On another word, the gate voltage will be in the middle of drain and source. Adding R_G allows the gate-source and gate-drain capacitance to keep the gate to channel voltage nearly constant during the signal period. But it will limit the switching speed. The simplified model is showed in Figure 14 of ON/OFF state.

Another thing is power handling capability. The output power from transmitter side from antenna can reach about 23dBm based on 4G/LTE standard. To handle this high power, all transistors must be strong enough to tolerate the voltage drop from output to ground. This not only requires a high trigger voltage FET and needs to use stacked structures to evenly distribute the voltage. All RF signal crossing the FET's source and drain are AC which means the power can jump across each stack through the gates. As an example, assuming 6V peak to peak AC signal is applied to an OFF stacked FETs. If the stack number is 6, each stack will evenly distribute 1V from its drain to source. With gate

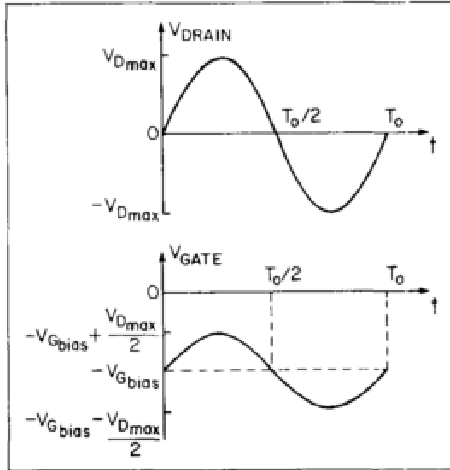


Figure 15 Gate voltage changing with drain side input when AC floating.

floating technique, the voltage from source to gate can reach about 0.5V which can easily turns-on most of the FETs with modern technology. Figure 15 shows the voltage on drain and gate when a sinusoidal AC voltage applied to a gate biased MOSFET. So, the stack number must be large enough to make sure the voltage distribution on V_{gs} of each FET must be smaller than turn-on threshold voltage (V_{th}).

The threshold voltage of our 45nm technology is about 0.225V which results $V_{peak}/2N$ must less than 0.225V. If the switch needs to handle 23dBm Tx output, the peak voltage can reach 4.525V. So, $N > \frac{4.525}{2\times 0.225} \approx 10.04$. At least 10 stacks of FETs are required and the resistance from source to drain is not that linear when V_{GS} approaching 0.225V. So the final stack number is set to 12 to make sure the P1dB is about 30dBm based on the simulation which gives enough margin for layout and fabrication.

What's more, another technique used in my multi-stacks switch design is feed forward capacitors. The voltage distribution is not evenly because the high frequency sinusoid signal can't give enough time for balancing each stack. As a result, two ends of stacked

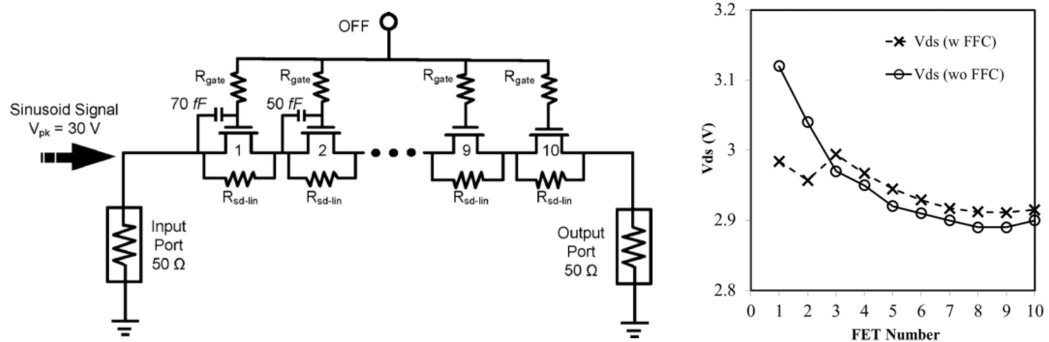


Figure 16 Voltage distribution of stacked NMOS with/without FFC.

MOSFETs will suffer more voltage swing cross them with positive and negative signal. Figure 16 depicts the voltage distribution when a positive sine voltage applied to 10 stacks of FETs with/without FFC.

Two frequency levels (low: 750MHz and high: 2GHz) are optimized based on a series of simulation and trade-off. The final design parameters of two RF T/Rx SPDT/SP3T switches are listed in Table 3-2 and the simplified schematic are showed in Figure 17.

Purpose		Stack Number	Width (um)	R_G (Ω)	FFC (fF)
High Band	Series	12	451.44	12K	20
	Shunt	12	351.12	14K	20
Low Band	Series	12	802.56	18K	20
	Shunt	12	451.44	20K	20

Table 3-2 Optimized design parameters of high/low bands switches.

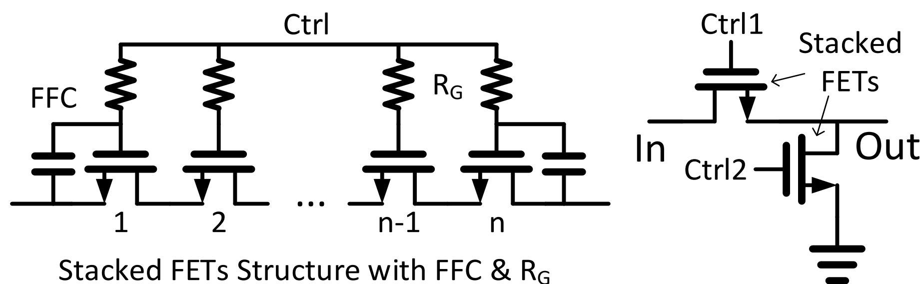


Figure 17 Simplified switch schematics of series-shunt, gate resistance, stacked FETs and FFC techniques.

The simulated performance of high band (1710MHz to 2155MHz) including insertion loss, isolation and 1dB compression point (P1dB) are showed in Figure 18-20. At 2GHz point, the insertion loss is ~1.082dB and the isolation is ~73.2dB with ~33dBm P1dB.

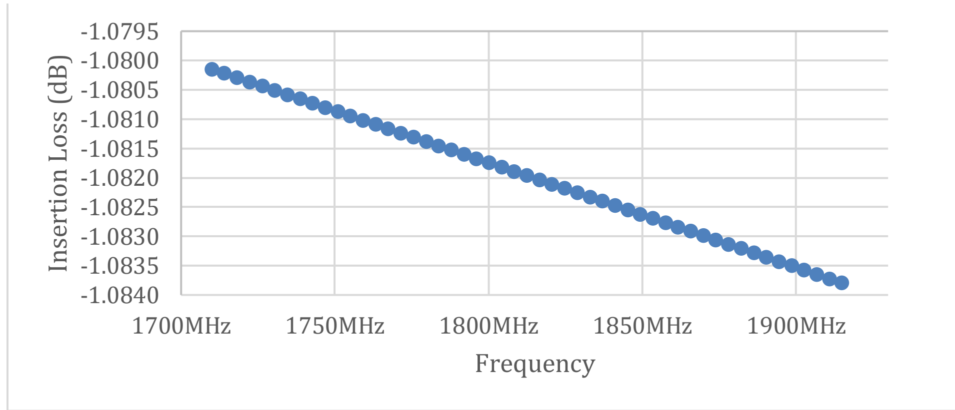


Figure 18 Insertion loss of high band T/Rx SPDT switch.

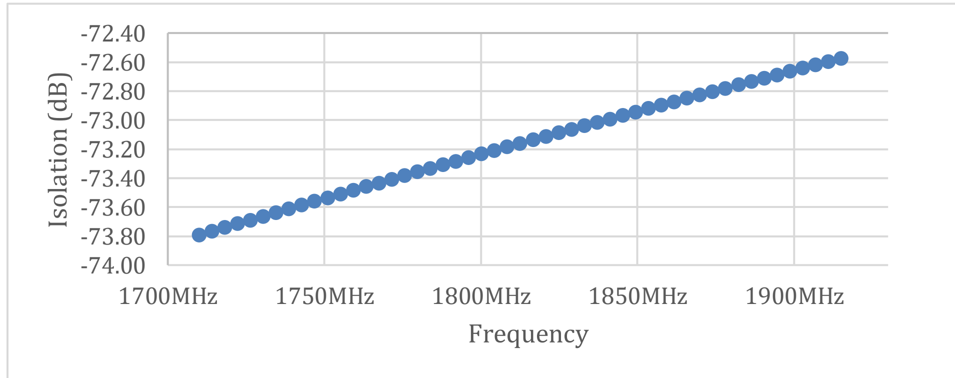


Figure 19 Isolation of high band T/Rx SPDT switch.

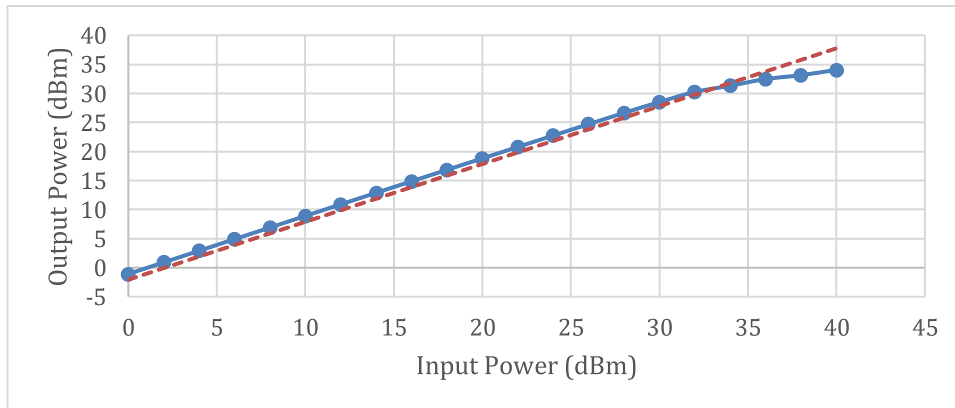


Figure 20 Input and output power of high band T/Rx SPDT switch at 2GHz.

The simulated performance of low band (699MHz to 894MHz) including insertion loss, isolation and 1dB compression point (P1dB) are showed in Figure 21-23. At 750MHz point, the insertion loss is ~0.731dB and the isolation is ~76.3dB with ~32dBm P1dB.

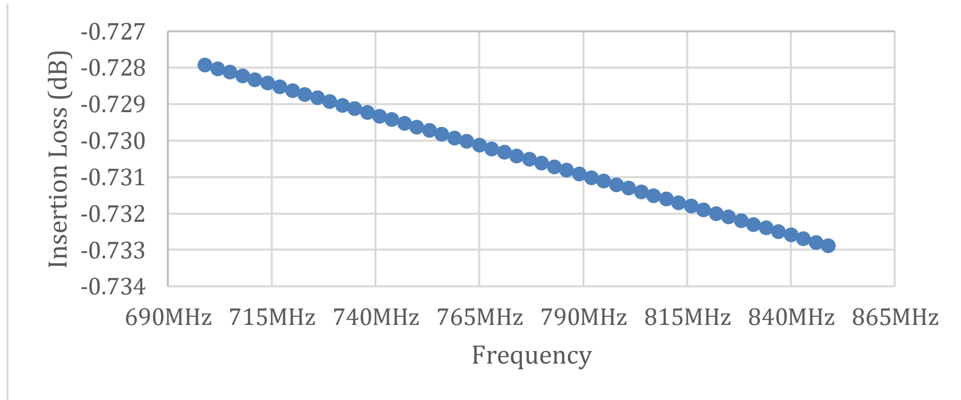


Figure 21 Insertion loss of low band T/Rx SP3T switch.

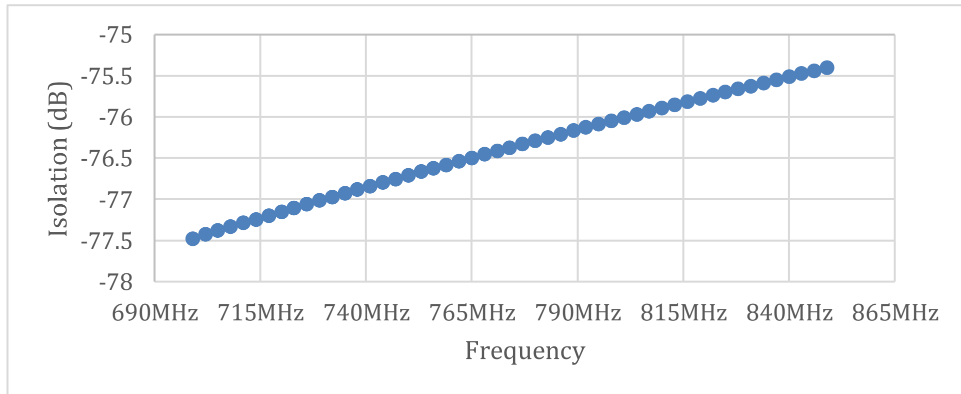


Figure 22 Isolation of low band T/Rx SP3T switch.

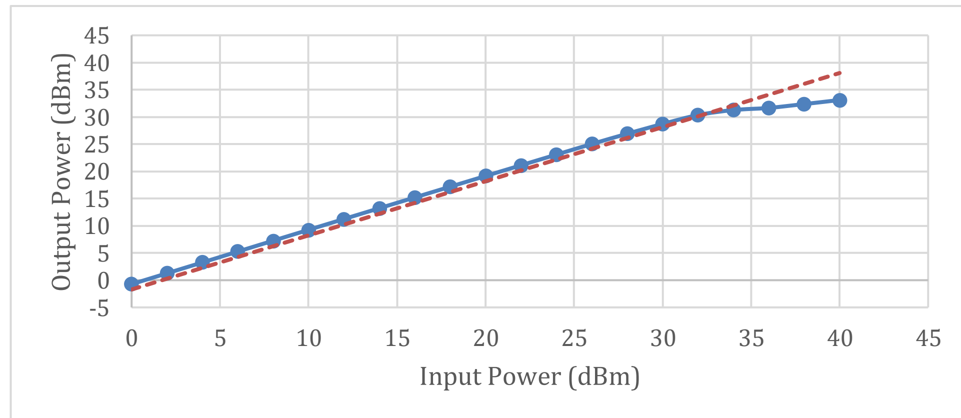


Figure 23 Input and output power of low band T/Rx SP3T switch at 750MHz.

3.2 Switch Array and Interference

To investigate the design influences (i.e., layout floor planning) on interference suppression, a 2X2 RF switch array was designed in 45nm SOI CMOS using the switches for 4G/LTE bands above. Two low band SP3T switch and two high band SPDT switch are used to demonstrate the layout induced interference for 4x4 MIMO applications. Each T/Rx switch can be used to connect the antenna and the interference from each side will be a severe damage for RF transceiver performance.

Figure 24 shows the two design splits used to analyze the layout effects of interference reduction in 2X2 switch array. H1/2 are the high band switches working at 2GHz level and

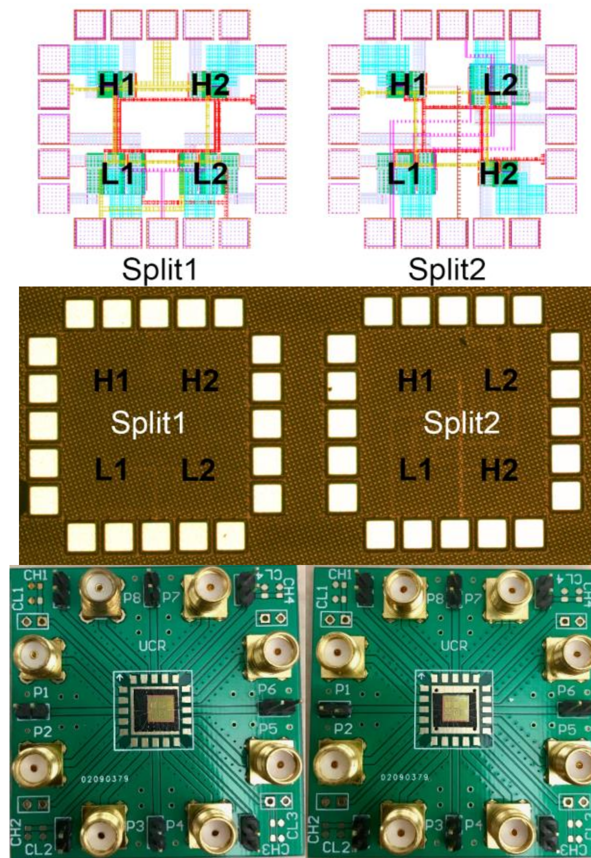


Figure 24 The 2x2 switch array design splits including layout, die photo and testing board photo.

L1/2 are the low band switches working at 750MHz level. The die photo and testing printed circuit board (PCB) are showed below. The interconnects transmission lines are optimized to 50Ohm. The die is attached to PCB using conducting tape and manually wire bonding using aluminum wire.

Interference, due to inevitable parasitic capacitance and EM crosstalk, is a major challenge to RF IC designs. Particularly, interferences between RF switches on a die can seriously affect RF IC performance. For example, in a MIMO SoC, high signal power in Tx channel will leak through switches to other branches in Rx mode who is extremely sensitive, resulting in performance degradation. Common interference suppression techniques exist. At technology level, techniques include using high-resistance substrates, guard rings, trench isolation, SOI, etc. At design level, the most commonly used technique may be careful layout floor planning where noisy elements are placed away from sensitive elements. However, the layout method is entirely limited by the chip size and complexity. In addition, entirely limited by the chip size and complexity. In addition, the distance based technique may not work in some designs. To investigate the effectiveness of the floorplan-based interference suppression technique for noise-sensitive RF ICs, a design split plan was used based on the 2X2 switch array in this work. Fig. 2 shows the design plan consisting of two layout splits, each has two high-band and two low-band switches that are placed at different locations on the dies.

3.3 Measurements and Discussion

The RF specifications including insertion loss and isolation of individual switches were measured using Agilent N5812A vector Network Analyzer. The control voltage of

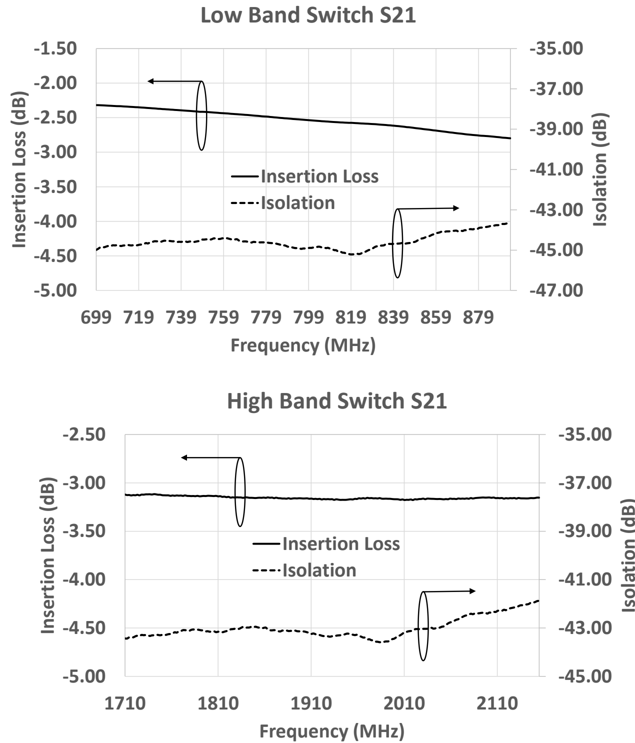


Figure 25 Measured S21 (Insertion loss for ON-state/Isolation for OFF-state) for the individual switches at PCB level

ON-state and OFF-state are 1V/0V with power supply. Figure 25 presents measured PCB-level isolation and insertion loss showing IL of 2.4dB and Iso of 44.5dB at 750MHz for a SP3T band switch, and IL of 3.2dB and isolation of 43.4dB at 2GHz for a high-band SPDT T/Rx switch. The performance is degraded because of the mismatching from PCB interconnection, wire bond and SMA connector. Based on our GSG probe testing of other designs, at least 1.5dB insertion loss degradation and about 5dB isolation improvement can be found from parasitic effects.

The interference was characterized using Agilent E8363B spectrum analyzer with 0dB input signals at 2GHz and 750MHz. To evaluate the interferences with different distance and frequency, two splits showed before were used with different layout floorplans. Figure

26 shows the measured inter-band (different frequency bands' switches) and inner-band (same band's switches) interference with 2GHz input signal. Same testing was duplicated at 750MHz. The inter-band interferences were measured using Split 1 where the output of the high-band SPDT (H2) affects the two low-band SP3T switches at different distances, i.e., L1 next to H2 for shorter distance (Near) and L2 at the diagonal for longer distance (Far). It shows that a longer distance helps to reduce interference by ~3.57dB. On the other hand, the inner-band interferences were measured for the high-band switch pairs, H1 and H2, on Split 1 and Split 2 with different distances, i.e., Near case for H1 and H2 on Split

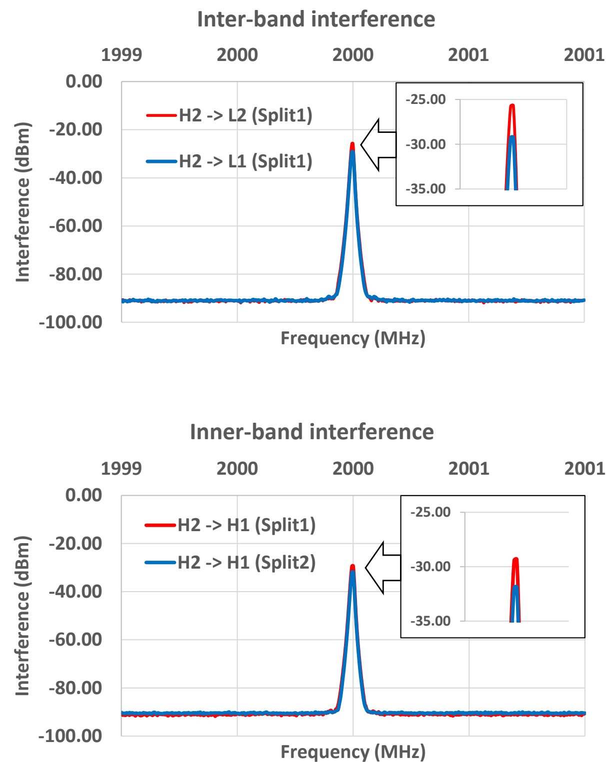


Figure 26 Measured inter/inner-band interference under 0dBm input at 2GHz shows varying interference suppression related to the layout distances of different switches

Input Signal	Distance	Inner-band	Inter-band
2GHz	Far (dBm)	-31.875	-29.251
	Close (dBm)	-29.325	-25.683
	Suppression (dB)	***2.55***	***3.568***
750MHz	Far (dBm)	-29.197	-43.244
	Close (dBm)	-25.705	-34.655
	Suppression (dB)	***3.492***	***8.589***

Table 3-3 Interference in 2X2 switch array with different layout floorplans.

1, and Far case for H1 and H2 on Split 1 and Split 2, respectively. Measurement shows an interference reduction of ~2.55dB. Table 3-3 summaries the improvement of layout floorplan technique which helps to reduce inter-band and inter-band interferences: a good reduction of 8.59dB was achieved for 750MHz signal input, while a small reduction of 2.55dB for 2GHz case. The results depict that while the layout technique (distance optimization) is useful in interference suppression in some case with low frequency, it is absolutely insufficient for the RF front end blocks like switches especially when multiple switches (switch array) in a single chip where the interferences after using both technology techniques (45nm SOI with shallow trench isolation) and design techniques (floor-planning) is still too high, unacceptable to complex multi-band RF ICs. To attack this noise isolation challenge, we devised a novel through-BEOL metal wall structure for interference isolation in next chapter that was validated at circuit level using T/Rx antenna switch.

Chapter 4 RF Flying Crosstalk Suppression

Crosstalk of noises or interferences via conductive semiconductor substrates (e.g., silicon) has been a major design challenge, particularly for noise-sensitive RF, and analog and mixed-signal (AMS) integrated circuits (ICs). While silicon-based CMOS ICs continuously benefit from aggressive scaling down at a pace of Moore's Law for more than fifty years, the global crosstalk problem has become increasingly severe for larger and more complex systems-on-a-chip (SoC) designed in advanced CMOS technologies [29]. Over years, substantial research efforts have been devoted to developing various on-chip noise reduction techniques in CMOS. As mentioned in Chapter 2, Silicon-on-insulator (SOI) technology can reduce substrate noise coupling utilizing features such as shallow trench isolation (STI), buried oxide (BOX) layer, high resistive (HR) substrate and in-substrate buried isolation ring (BI ring), making it attractive to RF and AMS ICs [27], [30]. Various in-silicon noise isolation techniques have been used in CMOS, e.g., guard ring, oxide trench, porous Si trench, decoupling by-pass capacitor, shielding ground and buried ground plane[31]–[37]. Many exotic means were also reported to reduce in-silicon crosstalk. For example, trap-rich substrates were reported to reduce crosstalk induced by the under-BOX parasitic surface conduction effect in SOI CMOS [38]–[40]. A Faraday cage made of metal-filled trench by microelectromechanical system (MEMS) process was reported that requires creating a deep local cavity in the wafer backside [41]. A poly-Si filled trench cage in SOI was reported that requires bonding a Si handle wafer [42]. A dotted Cu-filled trench cage formed by through-Si vias (TSV) that also requires wafer

backside etching to form a local cavity was reported [43], [44]. Unfortunately, these exotic crosstalk reduction techniques are impractical to foundry CMOS technologies.

Recent studies reveal that, with various in-substrate (in-Si) crosstalk-reduction techniques in use, the through-Si-substrate noise coupling problem has been greatly alleviated. Instead, the above-Si-substrate "flying crosstalk", defined as the noise coupling effect through the back-end-of-line (BEOL) stacks (above silicon) in CMOS emerges as the dominating global crosstalk problem. A simplified electromagnetic (EM) field simulation was conducted to show that the above-Si in BEOL flying crosstalk is becoming dominating in SOI CMOS. As shown in Fig. 27 for a simple EM simulation deck, two structures were simulated for crosstalk: on for Si substrate only and the other on has a full

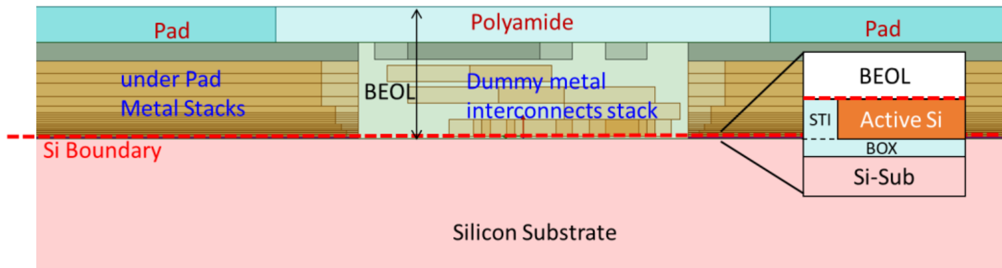


Figure 27 A simplified EM simulation to compare through-BEOL flying crosstalk with in-Si crosstalk in a 45nm SOI CMOS platform

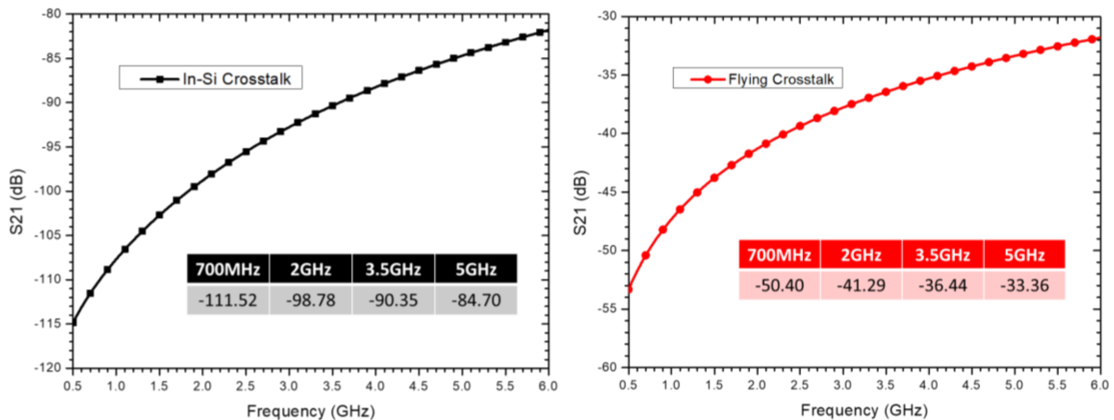


Figure 28 EM simulation results of the two cases: In-Si and From BEOL (Flying Crosstalk)

BEOL stacks as in a SOI CMOS process. The BEOL stack has a complex metal interconnect pattern with 10 Cu layers in this 45nm SOI CMOS process. EM simulation shows that the crosstalk through the BEOL stack above the Si substrate is dominating over its through-substrate counterpart as shown in Fig. 28. Specifically, at 2GHz, the in-Si crosstalk through the substrate is only about 1/500000 of the total global crosstalk, which is dominated by the through-BEOL flying crosstalk. Two factors were considered in this simulation comparison. First, while real-world ICs may have almost unlimited back-end design variations, in this EM simulation, we chose a simple case in a 45nm SOI CMOS featuring 10 metal layers and used a dummy random metal interconnect stacks to mimic complex metal interconnects layout designs. The 45nm SOI CMOS features a thin active Si layer of 80nm and a BOX layer of 145nm. The BEOL stack is about 11.3µm thick containing 10 metal interconnect layers. Second, the absolute values of the EM simulation may not be very accurate because the details of the BEOL stacks are foundry-confidential that is not available to us. Nevertheless, this simplified EM simulation shall be good for a comparison study, which roughly reflects the real-world observation that substantial crosstalk still exists even after using all kinds of conventional in-Si noise isolation structures commonly provided by a foundry.

The severity of the flying crosstalk is attributed to the massive scale of and complex metal interconnects in the BEOL stacks on IC chips, particularly for today's large and sophisticated RF/AMS SoC chips implemented in SOI CMOS where the active Si layer is typically less than 150nm and the through-substrate crosstalk can readily be suppressed using various in-Si noise reduction techniques. This above-silicon flying crosstalk problem

will particularly be a killing global interference challenge to advanced RF SoCs for fifth-generation (5G) global smartphones, which will operate in more than 40 frequency bands, and be implemented in a 7nm CMOS using reach to 18 copper metal layers to interconnect more than 5 billon transistors on one Si chip. Therefore, it is imperative to develop novel and disruptive in-die global noise isolation techniques to attack this ever-increasing above-Si flying crosstalk challenge for future RF/AMS ICs, which cannot be addressed by any existing in-Si noise isolation techniques.

4.1 Interference Suppression Using Metal Wall

Electromagnetic waves can travel through media. The coupling effect of unwanted RF noises, or interferences, through the thick BEOL stacks in CMOS ICs can be greatly exacerbated by the complex multi-layer metal interconnects on a chip, resulting in significant global flying crosstalk. We introduce a disruptive concept of in-die through-BEOL metal wall technique to suppress the RF flying crosstalk above an IC substrate, which was validated using a set of RF switch circuits implemented in a foundry 45nm SOI

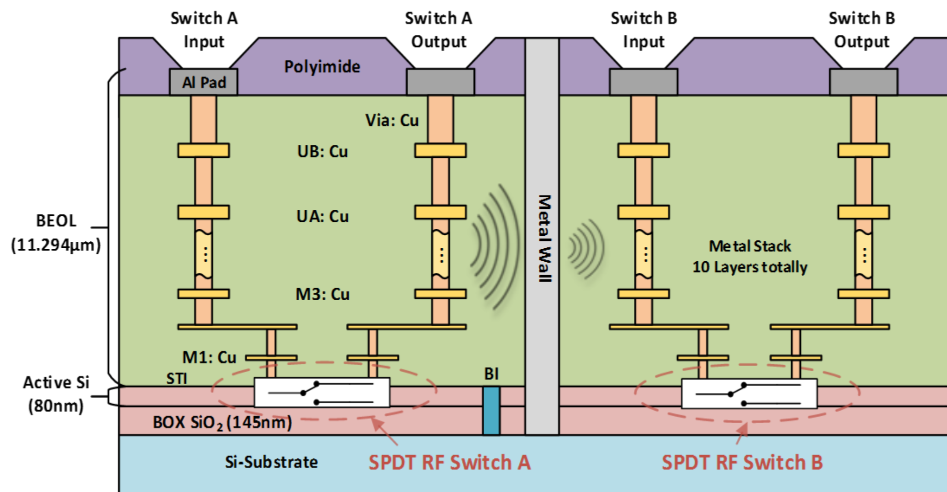


Figure 29 X-section illustration of new in-BEOL metal wall flying crosstalk suppression structure applied to RF T/Rx switches in a 45nm SOI CMOS.

CMOS technology. Fig. 29 depicts the through-BEOL metal wall structure in a SOI CMOS process. Briefly, circuit blocks are noise-isolated by a metal wall, closed or unclosed as needed, formed by selectively etching a trench into the BEOL stack that is then filled with metal. According the electromagnetic (EM) theory, the outgoing or incoming RF noises can be substantially blocked by the properly designed metal walls, hence isolating RF circuit blocks from each other on a single chip. The efficiency of flying crosstalk isolation is determined by several factors, including the type of metal and its formation process, dimensions (e.g., width), layout (e.g., single or multiple walls, and the spacing between the walls if using multiple metal walls) and shape (closed or unclosed) of metal walls. EM field simulation serves to guide the design of a metal wall to optimize the crosstalk isolation efficiency. Uniquely, the new concept was proposed with CMOS process compatibility in mind to ensure foundry manufacturability. Hence, a new process module was developed to make the through-BEOL metal walls in standard CMOS platforms by a new 3D heterogeneous integration flow.

The key novelty of this through-BEOL metal wall noise-isolation concept is that it can be readily implemented in standard CMOS processes through 3D heterogeneous integration. Hence, the fabrication processes must be CMOS-compatible without introducing many modifications to foundry CMOS processes. For concept demonstration purpose and because we cannot ask the foundry to change its SOI CMOS processes, we chose to develop a simple post-CMOS process module to make the through-BEOL metal wall structures in our cleanroom facility. For simplicity, we chose to use FIB technique for deep etching and used silver nanoparticle powder to fill the trenches to form the metal

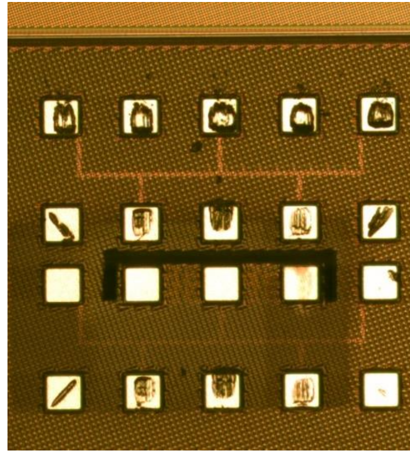


Figure 30 A die-scale photo shows the through-BEOL metal wall between the two SPDT switches on a chip.

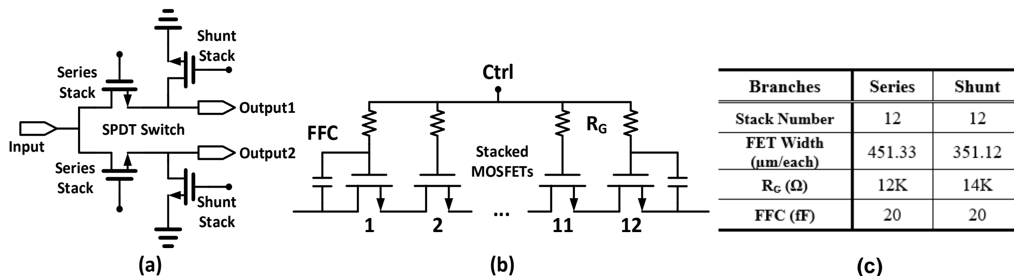
walls. Fig. 30 shows a large die photo for the IC using the inter-switch crosstalk-isolation metal wall structure. In this demonstration, a partial metal wall was used per EM simulation. In real-world designs, a closed metal wall may be used for enhanced isolation of global crosstalk suppression. For foundry production, both design and process optimization are needed for making the metal walls. The FIB etching can be replaced by high-aspect-ratio etch technique commonly used in standard CMOS processes for etching accuracy. A CMOS-friendly metal, other than silver, will likely be used. A standard trench-filling method, e.g., sputtering, may be use at a foundry to improve the quality of the metal walls. An alternative way, sometimes maybe easier to be integrated into a foundry CMOS process, is to form a stacked-via metal wall in the back end. The process follows: where the noise-isolation metal wall is needed on a die, use standard vias to construct the metal wall layer-by-layer by replacing the vias with suitable metal materials selected for making the metal wall. Since each layer of vias are fabricated using standard CMOS processes, no high-aspect-ratio etching would be needed. A new set of layout design rule may be used to make the stacked-via metal walls accordingly. The basic idea of this alternative way is to

use standard CMOS back end processes as much as possible, which could be more attractive to the foundries.

Therefore, the extra process steps needed to make the through-BEOL metal walls must be minimized and fully compatible to foundry CMOS processes. The through-BEOL crosstalk-isolation metal walls can be formed in two different ways: the first way is that, after completion of the BEOL stack in a standard CMOS process, a trench will be formed by etching through the full BEOL stack, which is then filled by a specific metal. Apparently, high-aspect-ratio etching is critical to deep trench etching and metal filling. The second way is to form a stacked-via noise-isolation metal wall, where within the wall region, selected metal vias will be converted into a dotted or continuous metal wall in a standard CMOS back-end process, hence minimizing any extra process steps needed. Either way, the through-BEOL crosstalk-isolation metal walls can be made easily in CMOS with limited process modification, hence, suitable for mass production.

4.2 Implementation in RF Switch Circuits

The new concept was validated using a set of RF transmitter/receiver (T/Rx) switch circuits implemented in a foundry 45nm SOI CMOS, featuring an 80nm thin active Si layer on a 145nm BOX layer. Fig. 31 (a) and (b) show the schematics for the single-pole-double-



Branches	Series	Shunt
Stack Number	12	12
FET Width (μm/each)	451.33	351.12
R_G (Ω)	12K	14K
FFC (fF)	20	20

Figure 31 (a) A simplified SPDT switch schematic featuring series-shunt MOSFET topology. (b) Circuit features including 12-FET stacks, gate resistor (RG) and feed-forward capacitor (FFC) are used in each stack branch. (c) Summary of circuit specifications for the SPDT switches.

threw (SPDT) T/Rx switches featuring a series-shunt MOSFET topology, high gate resistance, stacked MOSFETs and feed-forward capacitor (FFC) techniques introduced in Chapter 3. The SPDT switches were designed for RF SoCs for multi-band global smartphones using 4th–generation long-term evolution (4G/LTE) standard, covering the LTE bands-2/4/25 of 1710–2155MHz. The 4G/LTE T/Rx switch circuit was selected for demonstration because it is extremely sensitive to the inevitable noises on a chip and the inter-switch interferences due to its high Tx power. The global crosstalk in general will become more severe to RF SoCs for the incoming 5G mobile systems. To handle the Tx power of up to 23dBm with non-negative control voltage, 12 MOSFETs are used in the MOSFET stacks for both series and shunt branches. Fig. 31 (c) gives the switch design parameters optimized for low insertion loss (IL) and high isolation (Iso). The total die area of a SPDT is about 0.012mm^2 without pads. In this work, inter-switch interference is studied for two adjacent switches, depicted in Fig 32, where an incident signal is injected into the Input of Switch-A and its output signal leaks into Switch-B. Without any

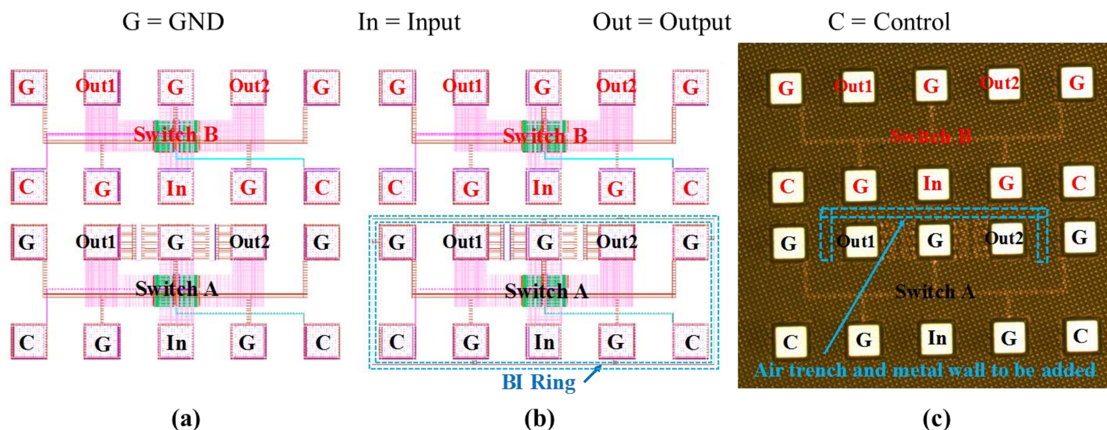
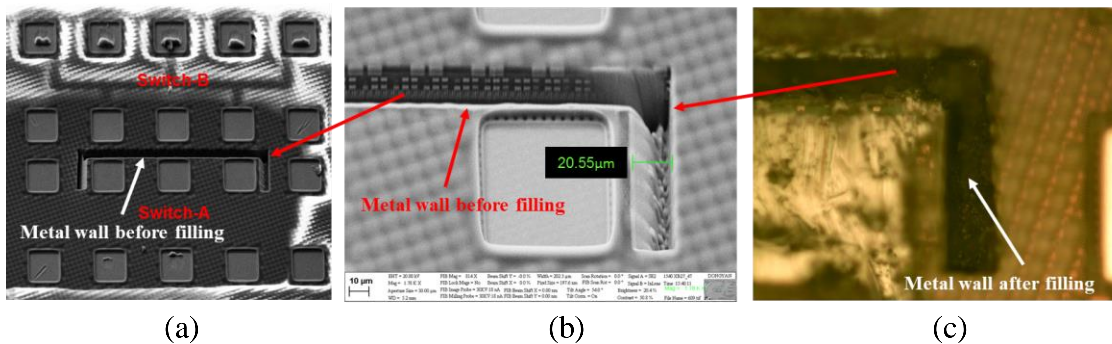


Figure 32 Layout of the SPDT splits for inter-switch interference study: (a) Original switches (No ISO), (b) switches with BI-ring, (c) Die photo for switches using the new in-BEOL metal wall (Metal Wall), also same circuits after the trench-etching but before metal-filling (Air Trench).

intentional input signal to Switch-B, the Output of Switch-B is monitored for the crosstalk coupled from Switch-A. Fig. 32 shows four design splits of Switch-A/B pairs: (1) original SPDT without noise isolation (No ISO), (2) using a foundry-provided in-Si buried isolation plug ring (BI ring) for noise reduction, (3) using an above-substrate open trench (Air Trench), and (4) using the new above-substrate metal wall (Metal Wall). The Metal Wall and Air Trench splits, both are through-BEOL structures, are same except that the latter is the unfilled trench to be used to make the Metal Wall by filling with silver nano powder in our post-fab process. The BEOL stack contains 10 Cu metal interconnect layers and is about 11.3µm thick for the 45nm SOI CMOS process used.

4.3 Measurements and Discussion

Fig. 33 shows photos of the fabricated SPDT switches featuring a new through-BEOL metal wall, including: Split-3, after-trench-etching-no-filling (Air Trench), and Split-4, after-filling (Metal Wall). The through-BEOL metal wall is 20µm wide and 15µm deep. The filling metal used was silver nano power in this demonstration. The T/Rx switch circuits were first optimized by simulation for key parameters including IL, Iso, and linearity represented by the 1dB-compression point (P1dB). Interference suppression was



(a) (b) (c)

Figure 33 SEM die photos for the Switch-A/B circuit pair: (a) Switch-A/B pair after trench-etching (20µm wide, 15µm deep), (b) Zoom-in for Switch-A/B showing the trench etched, but before metal filling (Air Trench split), (c) Zoom-in photo for the trench filled by nano silver powder (Metal Wall split).

then studied by EM co-simulation using a field simulator (HFSS) for 3D EM field simulation of the through-BEOL metal wall structure and an RF circuit simulator (ADS) for switch circuit co-simulation.

EM co-simulation was used to guide the design of SPDT switch circuits utilizing the through-BEOL metal walls for crosstalk isolation. HFSS-ADS co-simulation was conducted for the four switch circuit splits as depicted in the main text, i.e., Split-1 (No ISO), Split-2 (BI ring), Split-3 (Air Trench) and Split-4 (Metal Wall), designed in a 45nm SOI CMOS. Co-simulation was conducted for the switch circuits over a wide frequency spectrum of sub-6GHz that was adopted for the next-generation 5G mobile communications. Fig. 34 shows that the flying crosstalk is clearly dominating in the RF switch circuits designed in the 45nm SOI CMOS that cannot be suppressed by any existing in-Si noise-isolation techniques, e.g., using a BI-ring. The Flying crosstalk can be reduced

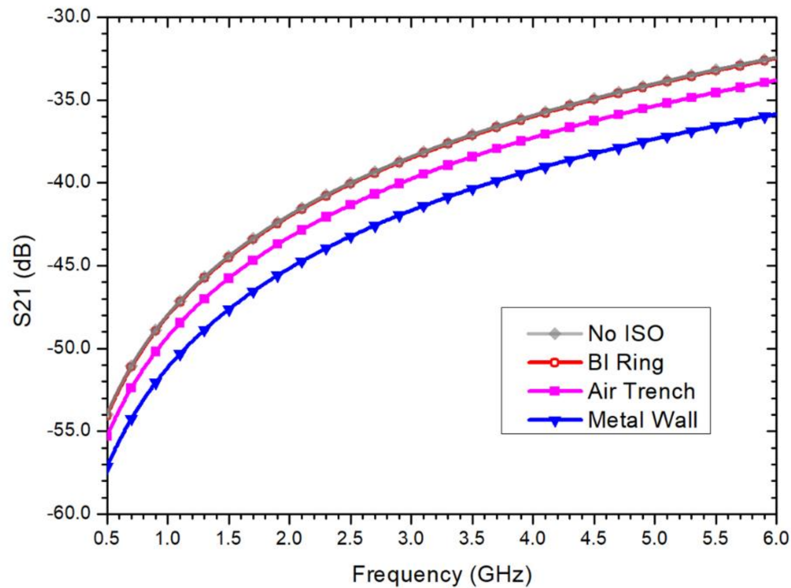


Figure 34 HFSS-ADS co-simulation of the four switch circuit splits at sub-6GHz frequency.

substantially by the Air Trench structure that, however, cannot be used in real-world ICs. Co-simulation readily shows that the proposed through-BEOL metal wall structure can suppress almost all the flying crosstalk through the BEOL stack above the Si substrate, which was confirmed experimentally. First, 3D EM field simulation using HFSS was conducted to optimize the through-BEOL metal wall structures for crosstalk isolation efficiency, which provides design guidelines for selection of metal materials and process method, as well as the dimensions (i.e., width), shapes (e.g., closed or partial wall) and layout (e.g., single or double wall, and wall-to-wall spacing). Second, the full-chip crosstalk suppression behaviors were studied by SPDT circuit simulation using ADS that includes the S-parameter data file for the metal wall from HFSS field simulation.

Comprehensive RF characterization was conducted for the switches to study chip level crosstalk reduction details. To ensure accurate RF measurement, a 50Ω load was used for impedance matching at the input and output ports, and ground-signal-ground (GSG) test patterns were designed for testing the T/Rx switches. The RF characterization was conducted using Vector Network Analyzer (Agilent E8363B), Signal Generator (Keysight N5182A) and Spectrum Analyzer (Agilent E4405B). The control voltages are 1V (ON) and 0V (OFF). S-parameters were measured to obtain insertion loss and isolation parameters of the SPDT switches. Large signal measurement was conducted to obtain P1dB values to evaluate the power handling capability and linearity. Inter-switch interference effect was monitored using a signal generator (Keysight N5182A) and a spectrum analyzer (Agilent E4405B). Fig. 35 (a) (b) shows the simulated and measured IL and Iso results from 1710MHz to 2155MHz for the original SPDT switch. Simulation

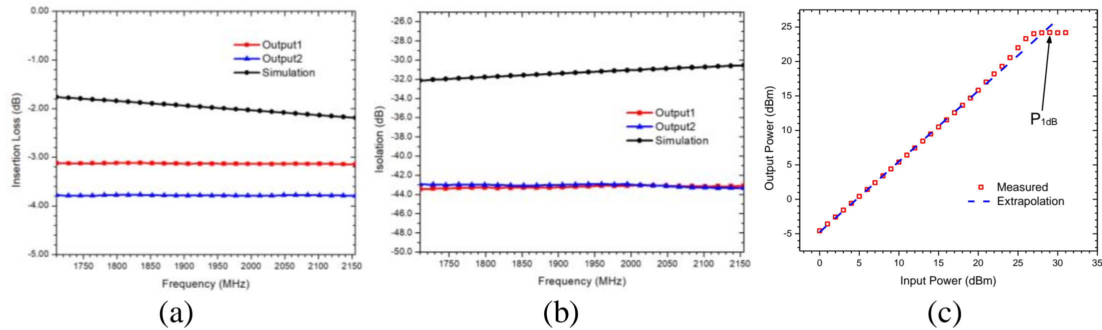


Figure 35 Measured and simulated key circuit specs for the SPDT switch pair show satisfactory circuit performance for the T/Rx switch circuits. (a) Insertion loss for the SPDT switches from 1710 to 2155MHz. (b) Isolation for the SPDT switches from 1710 to 2155MHz. (c) SPDT power handling capability and good signal linearity given as the 1dB-compression point (P1dB) at 2GHz.

shows symmetric SPDT performance with IL of ~2.04dB and Iso of ~31.03dB. Measurement shows slightly asymmetric results at 2GHz, i.e., IL~3.17dB and Iso~43.39dB for Output1, and IL~3.81dB and Iso~41.86dB for Output2. Fig. 35 (c) depicts the measured input-output curve where the extrapolated linearity is P1dB ~ 29dBm. These key specs are good for T/Rx switches in a 45nm SOI CMOS.

Fig. 36 (a) shows the concerned inter-switch interferences for the four SPDT splits by HFSS-ADS co-simulation. As expected, the foundry-provided conventional in-Si BI ring, which was designed to block the active silicon noise coupling effect, seems to be useless

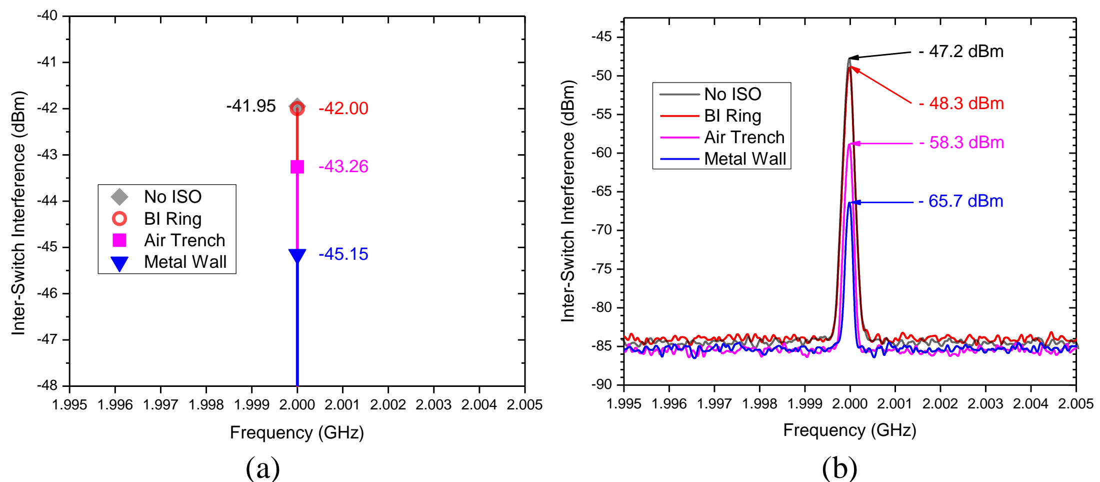


Figure 36 Crosstalk comparison for the four SPDT splits by (a) HFSS-ADS co-simulation, and (b) measured (0dBm incident signal to Input of Switch-A and interference monitored at Output of Switch-B).

for suppressing the above-Si flying noises coupling through the BEOL stack (-42.00dBm/Split-2 vs. -41.95dBm/Split-1). Since the in-BEOL flying noise coupling effect becomes dominating, the through-BEOL Air Trench can alleviate the flying crosstalk substantially (-43.26dBm/Split-3) as compared with the No ISO switch (-41.95dBm/Split-1) because the air gap reduces the capacitive coupling effect by changing the dielectric parameter. Of course, an air trench is not accepted for CMOS ICs. It is readily observed that the new Metal Wall structure can significantly suppress the in-BEOL flying crosstalk as expected (-45.15dBm/Split-4 vs. -41.95dBm/Split-1). It is worth noting though that, while the HFSS-ADS co-simulation provides design guidance, it may not be accurate because the complex BEOL stack details and materials parameters are normally unknown to IC designers due to foundry confidentiality. Fig. 36 (b) shows measured crosstalk suppression results for the four switch splits when biased in ON state. It clearly confirms that the in-Si BI ring is useless for blocking flying noises (-48.3dBm/Split-2 vs. -47.2dBm/Split-1). In contrast, it readily shows that the Metal Wall significantly suppressed the inter-switch flying interferences by 18.5dB (-65.7dBm/Split-4 vs. -47.2dBm/Split-1). It is obvious that the new through-BEOL metal wall can almost completely remove the in-BEOL flying crosstalk (~98.6%, in linear scale) as expected. Both simulation and measurement confirm that the new through-BEOL metal wall crosstalk isolation method is very efficient in blocking the dominating above-Si flying crosstalk, which cannot be addressed by conventional in-Si noise-isolation techniques.

For the purpose of a quick concept demonstration, a simple post-CMOS process module to make the through-BEOL metal walls in our cleanroom facility was developed,

which can be readily integrated into a foundry CMOS flow by 3D heterogeneous integration. Although it is successfully demonstrated that the new flying crosstalk suppression technique can reduce interference a lot, its performance is limited by many factors: For example, the high-aspect-ratio etching cannot be done well using research level FIB, which however shall be an easy process step in any industrial foundry with thinner but better metal filled walls. On the other hand, the proposed alternative stacked-via through-BEOL metal wall structure and process flow can be readily realized and optimized by a commercial IC foundry through 3D heterogeneous integration, which only requires minor modification to a standard CMOS process flow. It can be envision that this simple and fully CMOS-compatible through-BEOL metal wall noise-isolation technique is a viable solution to the emerging above-Si global flying crosstalk problem for next-generation RF/AMS ICs.

Chapter 5 Electrostatic Discharge Protection

ESD failure is a major reliability concern to integrated circuits (ICs) and electronic products (e.g., smart phones). As IC technologies continue migrating into beyond 40 nm domain, ESD protection is emerging as a key IC design challenge, particularly for high-frequency operating at multi-GHz and high-speed ICs with data rates beyond tens gigabits per second (Gbps) [45]. The ESD phenomenon originates from transfer of electrostatic charges between two objects of different electrical potential and the resulting fast and large ESD transients can damage ICs. One emerging ESD protection design challenges along with IC technology scaling-down is the ESD Design Window Shrinking Effect. As depicted in Fig. 37, an ESD Design Window is defined by the breakdown voltage (BV) and power supply voltage (VDD) of the core circuit under protection with a proper safe margin for practical IC designs [46], [47]. An ESD protection solution requires accurate design of ESD-critical parameters including the ESD triggering voltage, current (V_{t1}, I_{t1}), the ESD holding voltage and current (V_h, I_h), the ESD discharging resistance (R_{ON}), and

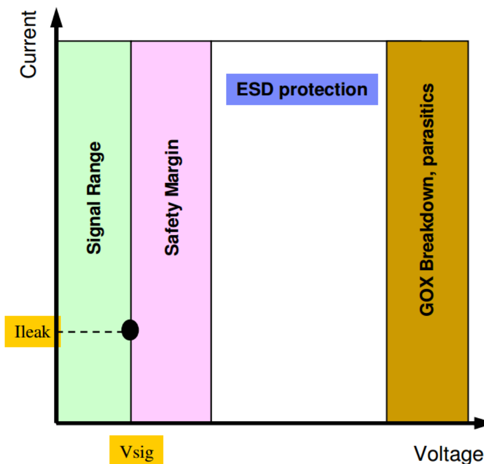


Figure 37 ESD protection design window.

the ESD failure voltage and current (V_{t2}, I_{t2}). The rationale for setting the ESD Design Window is that the triggering of an ESD protection structure must be lower than the BV of the IC node protected ($Vt1 < BV$) to prevent any ESD damage; meanwhile, to avoid latch-up effect, the holding voltage of an ESD protection structure has to be higher than the IC supply voltage (i.e., $Vh>VDD$). Unfortunately, as IC technology scaling-down continues, BV of IC devices decreases dramatically, however, IC supply voltages only drop slightly, which results in a narrower ESD design window, i.e., ESD Design Window Shrinking. As such, a quantitative ESD protection design method is required for accurate ESD protection circuit designs in advanced IC technologies, especially for mixed-signal ICs involving multiple supply voltage domains [48].

This chapter introduces the basic ESD protection strategies and design consideration. Then, more details about human body model (HBM) and charged device model (CDM) ESD protection will be explained including TCAD simulation, TLP/VFTLP testing, typical structures, overshoot issues and performance validated in 28nm bulk CMOS technology.

5.1 ESD Introduction

ESD is an important consideration in chip design and needs to be developed from initial stages of the design process. Considering ESD as a last-minute addition often causes problems of unsatisfactory reliability or degraded chip performance.

All pins, even high speed, RF or sensitive analog pins, must have ESD protection. Low ESD protection causes packaging, test fallout and product failures in the field. Months of engineering effort, failure analysis, and partitioning experiments in manufacturing are

consumed trying to understand failure of parts, only to potentially trace the problem back to ESD vulnerability.

There are two general categories of electrostatic discharge considered in the microelectronics industry: (i) the electronic die is touched by a charged human being or by a charged machine, which discharges to the ground via the electronic die and (ii) the capacitance of the die itself charges up or discharged when the die experiences a contact with the environment. The most developed models to reproduce the known ESD events and covered by standards are: (1) human body model (HBM), (2) machine model (MM) and (3) charged device model (CDM). Most chips have specified requirements and/or targets for protection levels to these models. The basic characteristics of each event are mentioned in the Figure 38.

Category	Equivalent circuit	ESD event experienced
1. a: Die touched by a human being during process or handling. **Human Body caused ESD event**	R_{HBM}, V_{HBM}, C_{HBM}, DUT	
1. b: Die touched by a machine during process or handling. **Machine caused ESD event**	L_{MM}, V_{MM}, C_{MM}, DUT	
2. Die's capacitance charges/discharges during contact with environment. **Charged device ESD event**	DUT, V_{CDM}, C_{CDM}, R_{CDM}	

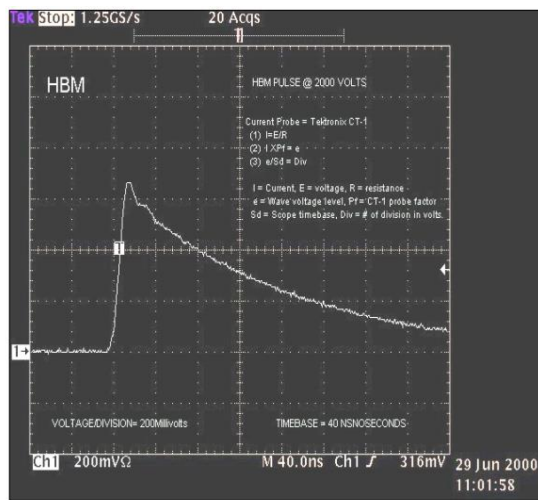
Figure 38 The categories of ESD events with equivalent circuits and models.

5.1.1 Human Body Model (HBM)

As the name implies, this model was developed to represent a human discharging across two pins in an IC. The equivalent circuit for HBM is a 1.5kOhm resistor in series with a 100pF capacitor. The capacitor is charged to specified ESD voltage, and then the switch is closed and the current flows through the Device Under Test (DUT). An HBM ESD event occurs between any two pins on a module. HBM is the longest ESD event of the three models, but it has the lowest current. Specialized ESD testers equipped with the appropriate resistor and capacitor network are used to apply an HBM stress to a module to determine its robustness. Also, transmission line pulse (TLP) tester can also be used to generate similar pulse waveform showed in Figure 39.

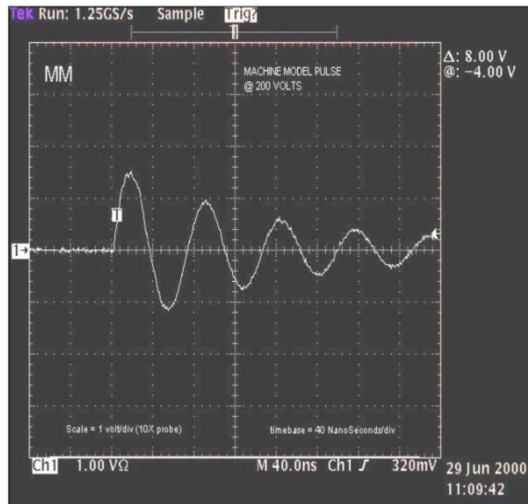
5.1.2 Machine Model (MM)

Similar to HBM, the MMESD event occurs across any two pins on a module. MM, however, simulates a tool discharging across the two pins. This event is more rapid than HBM and has a higher current level. The waveform is also quite a bit different, the MM



Characteristics:
- Relatively slow (rise time 5-10ns, pulse width ~150ns)
- Unidirectional current ~ 1-3A
- IHBM = VHBM/1.5kΩ [example: 2kV HBM = 1.3A current]
- Electrically looks like a current source

Figure 39 2000V HBM current waveform with its characteristics.



Characteristics

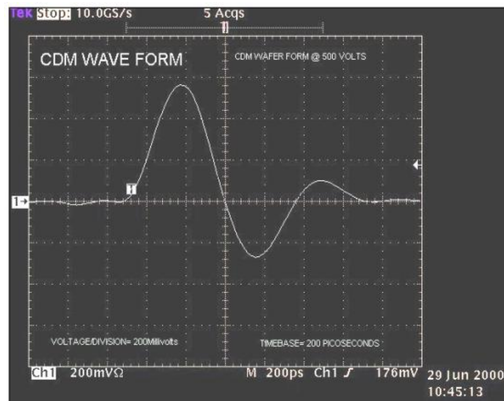
- Rapid event
- Additional series inductance (0.5-1.0uH) in equivalent ckt not shown
- High currents (~3 – 5A)
- 200V MM = 2.8-3.8A (allowed tester range per JEDEC spec)
- Bidirectional -- current oscillates into and out of the pin.

Figure 40 200V MM current waveform with its characteristics.

wave form is a bi-directional damped oscillation showed in Figure 40. The test procedure for MM is similar to HBM, but the MM is represented by a capacitor in series with an inductor as shown in Figure 38. The performance of MM protection can be estimated from the results of HBM.

5.1.3 Charged Device Model (CDM)

Unlike HBM and MM, CDM ESD events only involve a single pin on the module. Modules can develop charge through triboelectric (frictional) effects. When this occurs, unless a high resistive path is used to discharge the module, it may be exposed to a CDM ESD event. The CDM event charges a chip and then discharges to ground out of a single pin. CDM is a very high current event but occurs in a very short time interval. CDM testing is generally accomplished only on completed chips. A field plate is used to impose a specified charge on the module, and then a probe with a low resistance path to ground drops down on a single pin to discharge the module. This procedure is repeated for each pin on the module, and a functional test is employed after CDM stressing is completed to



Characteristics
- Very rapid event (rise time ~ 100s of ps, duration 1-2ns).
- High currents (1-2A / 100V CDM vs. 66mA / 100V in HBM).
- Module capacitance in the range of 1-20pF

Figure 41 500V CDM current waveform with its characteristics.

determine if the module has been damaged. The current waveform is showed in Figure 41 with detailed characteristics. Very fast transmission line pulse (VFTLP) are sometimes used to estimate the CDM protection performance.

5.1.4 Transmission Line Pulse (TLP)

TLP testing is a newer test model than the ESD models described previously and is generally not a required stress for a product qualification. The main use of TLP is to develop current-voltage relationships for semiconductor devices and circuits under ESD like conditions. Device operation under the high current and short duration of an ESD event can be significantly different than normal functionality. TLP is used extensively for analyzing discrete devices at wafer level during technology development activities, and most of the results that will be presented in the sections that follow are from TLP testing on devices in this technology. The basic concept of TLP testing is applying a square wave to the DUT and measuring the current through and voltage across the device. Different pulse widths can be used to simulate different ESD events. Generally, a 100ns pulse is used to simulate HBM, and a 30ns pulse is being evaluated for simulating MM. Both of these pulses can be delivered using most of the available TLP testers. To investigate devices

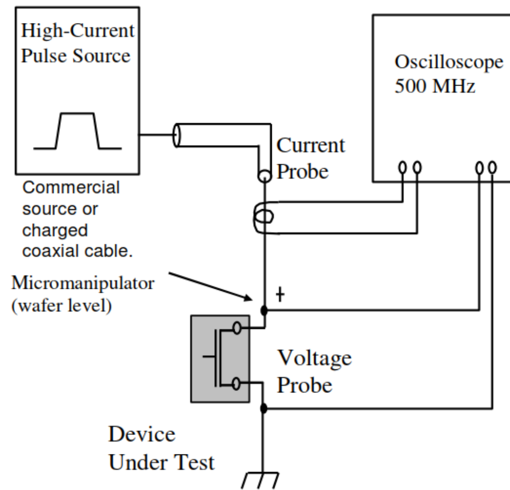


Figure 42 Transmission line pulse (TLP) test bench.

under CDM conditions, an extension to TLP often referred to as Very Fast TLP (VFTLP) is employed. The pulse width used for a VFTLP tester is typically between 1ns and 3ns, with rise and fall times on the order of 100ps – 500ps to evaluate the CDM performance. Figure 42 shows the TLP tester components including high-current pulse source and oscilloscope.

5.1.5 ESD Model Standards

ESD protection levels are generally set by product owners based on how well the environments where the chips will be handled are controlled and by the requirements of

ESD Model	Specification
HBM	JEDEC JESD22-A 114-B MIL-STD-883E Method 3015.7 ESDA STM5.1-1998
MM	JEDEC EIA/JESD22-A115-A 1997 ESDA STM5.2-1999
CDM	JEDEC JESD22-C101B.01 2004 ESDA STM5.3.1-1999
TLP	No accepted Standard Exists

Table 5-1 Interference in 2X2 switch array with different layout floorplans.

customers of the chips. The different protection standards seen in the industry are recorded in the Table 5-1 including JEDEC, MIL and ESDA standards that describe each of the ESD models.

5.1.6 ESD Protection Strategy

One of the most commonly used ESD protection strategies is to use two diodes on the pin, one to the supply bus and the other to the ground bus. Figure 43 shows the current paths in the chip for the four possible discharges between the pin and its supply and ground busses. Sometimes, the internal circuitry has some self-protection capability which can help improving ESD performance. Good ESD results are achieved when the discharge paths are confined to these routes. If the resistance of the wiring, the power clamp or the protect devices causes the signal pad or Vdd bus to rise too high in voltage the discharge path may be rerouted through I/O or internal circuitry. Therefore, the discharge path must not only be able to handle the current, it must be able to do so without allowing large voltages to develop across the parallel circuitry due to I-R effects.

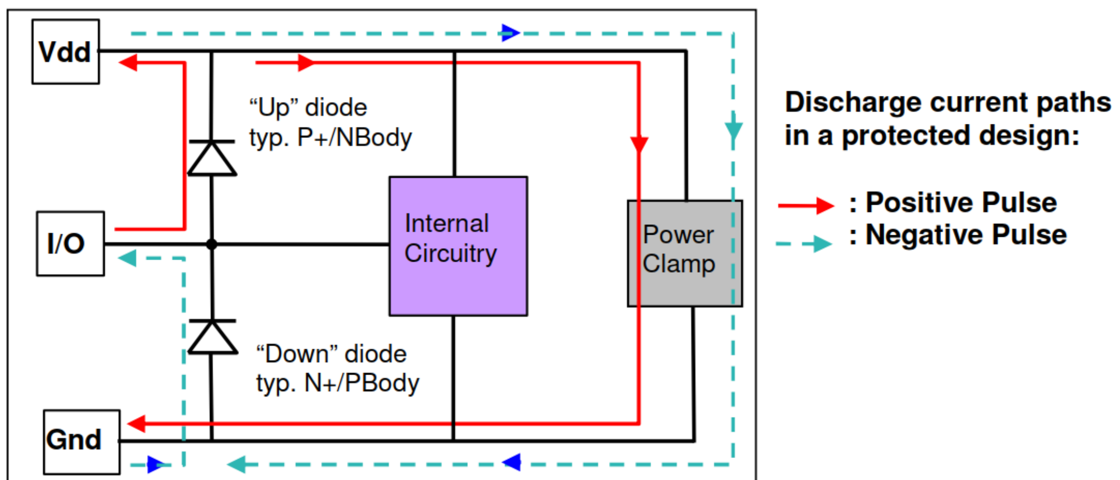


Figure 43 Typical protection circuit topology – double diode network.

5.2 Human Body Model (HBM) ESD Protection

5.2.1 Stand-alone ESD structures

Stand-alone ESD structures, including shallow trench isolation (STI) diode, gated diode, diode string, gate grounded NMOS (GGNMOS), silicon-controlled rectifier (SCR) and diode triggered SCR (DTSCR), have been characterized for HBM protection with TCAD simulation and VFTLP measurements. Critical ESD parameters, including trigger voltage (V_{t1}), holding voltage (V_h), turn-on resistor (R_{on}), breakdown current (I_{t2}) and leakage current have been extracted for 28nm process. The structures of each type of protection devices are demonstrated here with physical based model building and TCAD simulation.

a) STI Diode

STI diode is the most basic ESD protection structure with about 0.7V trigger voltage and well integration capability. Double diodes structure is widely used in lots of full-chip bi-directional protection. Figure 44 shows the cross-section of N+PW/P+NW STI diodes. The ESD current injects from the anode side to cathode.

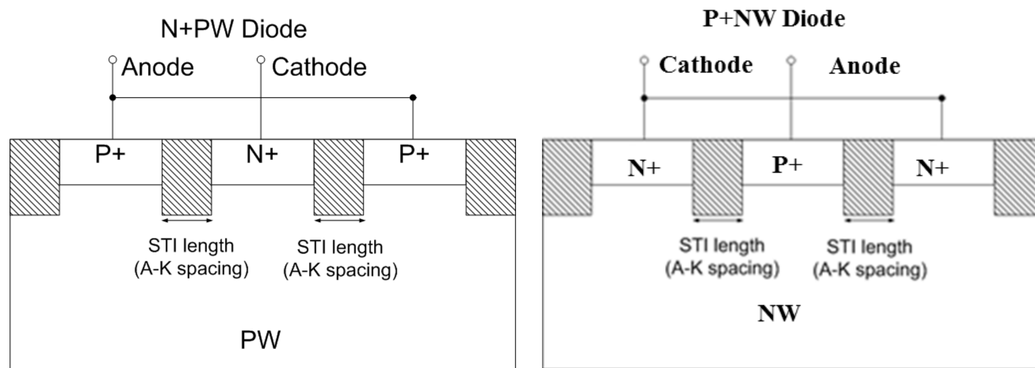


Figure 44 STI Diode X-section.

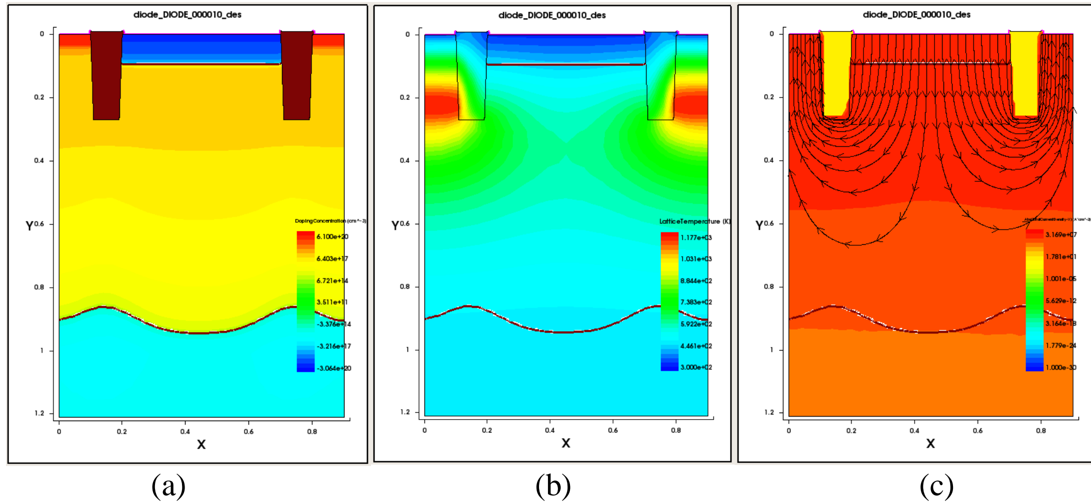


Figure 45 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow.

2D TCAD model has been built based on the fabrication process. HBM current pulse is stimulated to the structures. Figure 45 shows the doping density, temperature distribution and current flow inside STI diode structures. The current and temperature are crowded around the STI because it is the shortest path.

b) Gated Diode

Gated diode has better R_{on} by replacing STI with a gate to reduce the length of the current path. However, the parasitic capacitance will be increased because of the gate structure. Figure 46 shows the X-section of two different gated diodes.

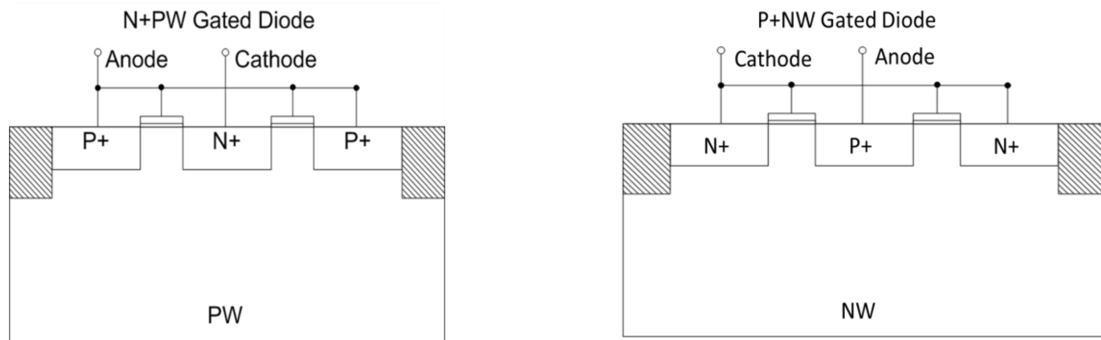
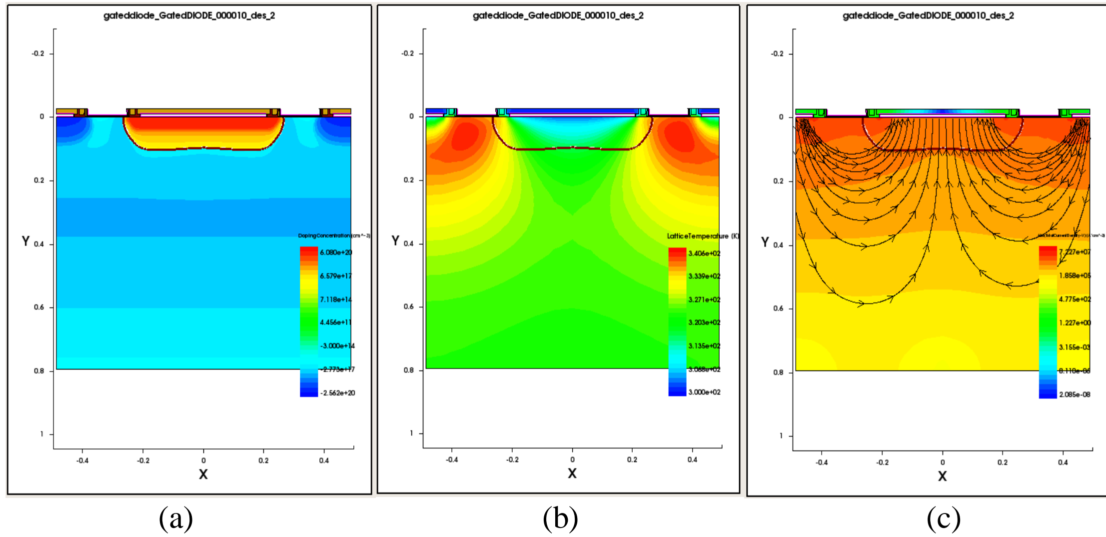


Figure 46 Gated Diode X-section.



(a) (b) (c)

Figure 47 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow.

2D TCAD model has been built based on the fabrication process. HBM current pulse is stimulate to the structures. Figure 47 shows the doping density, temperature distribution and current flow inside gated diode structures. Compared with STI diode, the temperature hot spot and current dense area are near the junction. A shorter path can be achieved without STI block.

c) Diode String

Diode string increases the trigger voltage by stacking the diodes which can match some high voltage application. The V_{t1} can be well controlled by the number of diodes.

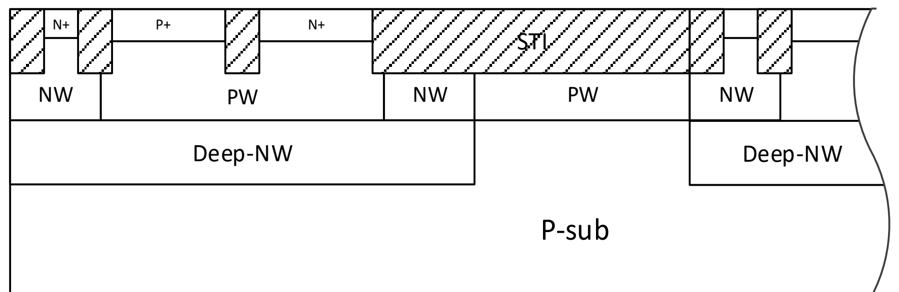


Figure 48 Diode String X-section.

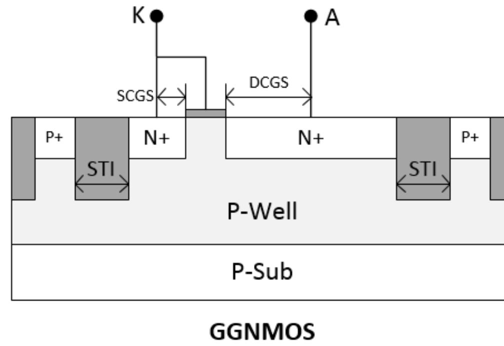


Figure 49 GGNMOS X-section.

What's more, the series connected parasitic capacitor will reduce the total capacitance. Figure 47 shows the X-section of diode string.

The diode string structure performance can be estimated by series STI diodes. Deep-NW and NW on the side are used to isolate each diode which also introduce more die area.

d) GGNMOS

GGNMOS structure is widely used as a high voltage protection due to its snapback I-V curve. The junction inside MOSFET is broken down first and then the current will go through the parasitic BJT inside. The trigger voltage is larger than the voltage drop across

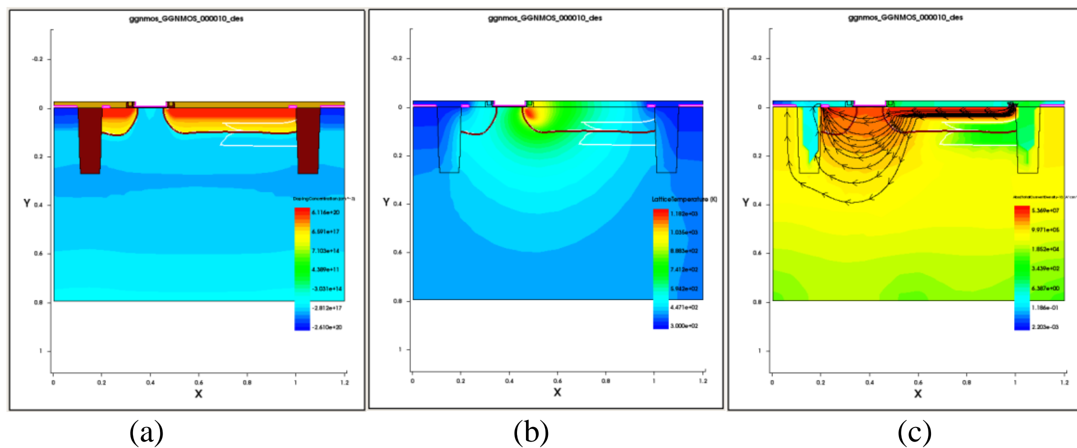


Figure 50 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow.

it when the BJT is ON. So, from the I-V curve, a snapback structure can be seen. Figure 49 shows the X-section of GGNMOS.

2D TCAD model has been built based on the fabrication process. HBM current pulse is stimulate to the structures. Figure 50 shows the doping density, temperature distribution and current flow inside GGNMOS during the HBM stress. The hotspot is at the breakdown junction and current go through under the gate directly after it triggered.

e) SCR

The SCR structure use the latch-up from two BJT inside a NPNP or PNPN structure to generate the snapback performance as a useful ESD protection device. The trigger voltage of SCR is higher than GGNMOS because the N-well/P-well junction has to be breakdown first than trigger the parasitic BJTs. Benefiting to its low capacitance, it's now used for high frequency IC design with a smaller size and same current handling capability. Figure 51 shows the cross-section of basic PNPN SCR structure.

2D TCAD model has been built based on the fabrication process. HBM current pulse is stimulate to the structures. Figure 52 shows the doping density, temperature distribution

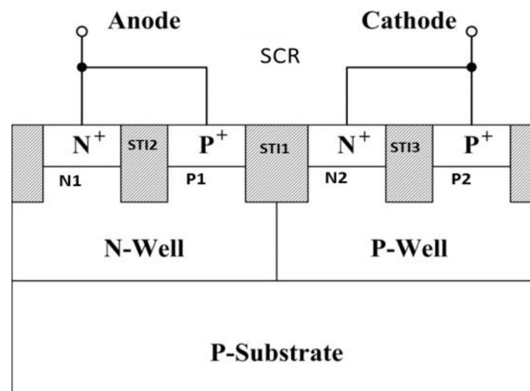


Figure 51 SCR X-section.

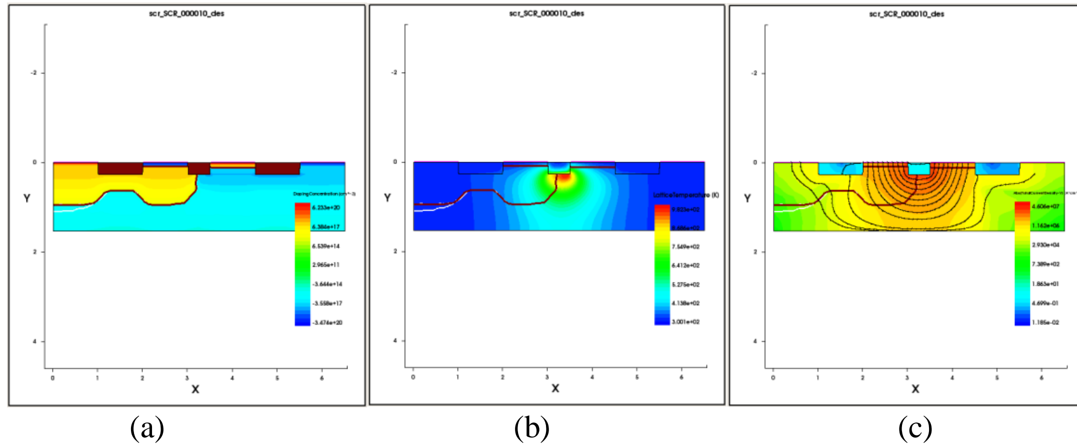


Figure 52 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow.

and current flow inside SCR. Same as GGNMOS, the spot where breakdown happens has the highest temperature and the current will go through quickly after parasitic BJTs turns on. Due to its active device performance, the current capability of SCR is better than normal diode and NMOS based structures.

f) DTSCR

The limitation of SCR is the trigger voltage, normally about 12V, which is too high to be used as low voltage modern applications. One solution to solving this issue is use the diode to help trigger the intrinsic BJT with parasitic base current. Diode string can be used

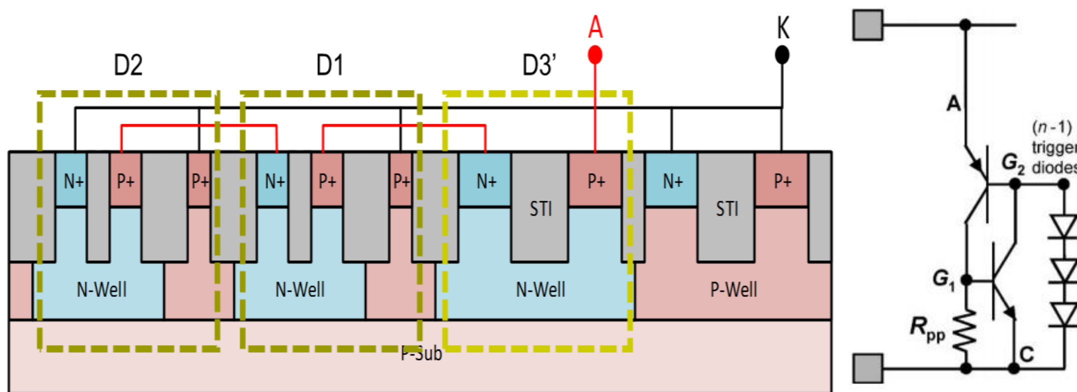
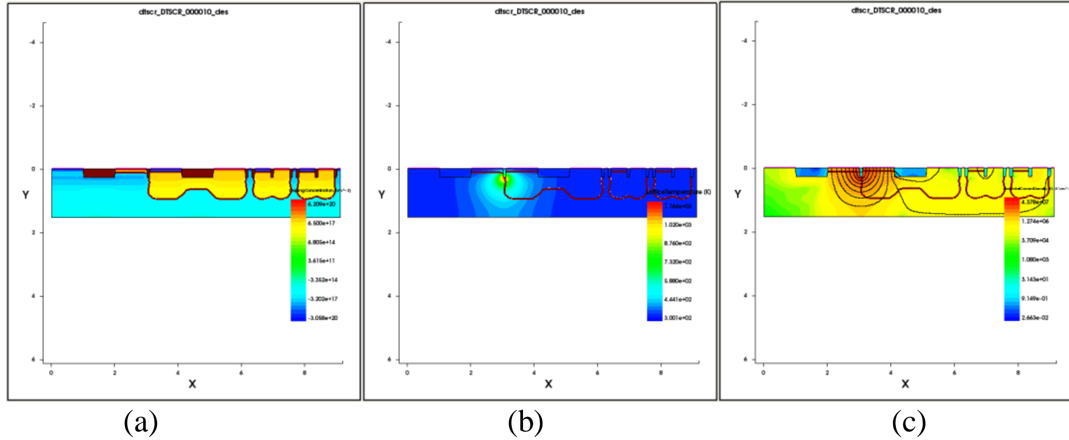


Figure 53 DTSCR X-section and equivalent circuit.



(a) (b) (c)

Figure 54 TCAD simulation results: (a)doping density (b) temperature distribution (c) current flow.

to control the trigger voltage accordingly. The cross-section with equivalent schematic are showed in Figure 53. The total trigger diode number is 4 due to an intrinsic one from SCR itself.

Since the trigger can be done by parallel diode string, the temperature dense area showed in Figure 54 is much smaller than breakdown first device and current will finally use the path of SCR instead of using triggering diodes.

5.2.2 TLP measurements

100ns TLP is used to characterize the HBM ESD performance of different structures. Each structure has different splits to analysis the size corresponding to its protection capability. If same protection level is achieved, the smaller size and lower parasitic capacitance will be better. The key parameters including trigger voltage (V_{t1}), breakdown current (I_{t2}), turn-on resistance (R_{on}), holding current (I_h) and voltage (V_h) of snapback devices are extracted which will be a good reference for 28nm technology.

a) STI Diode

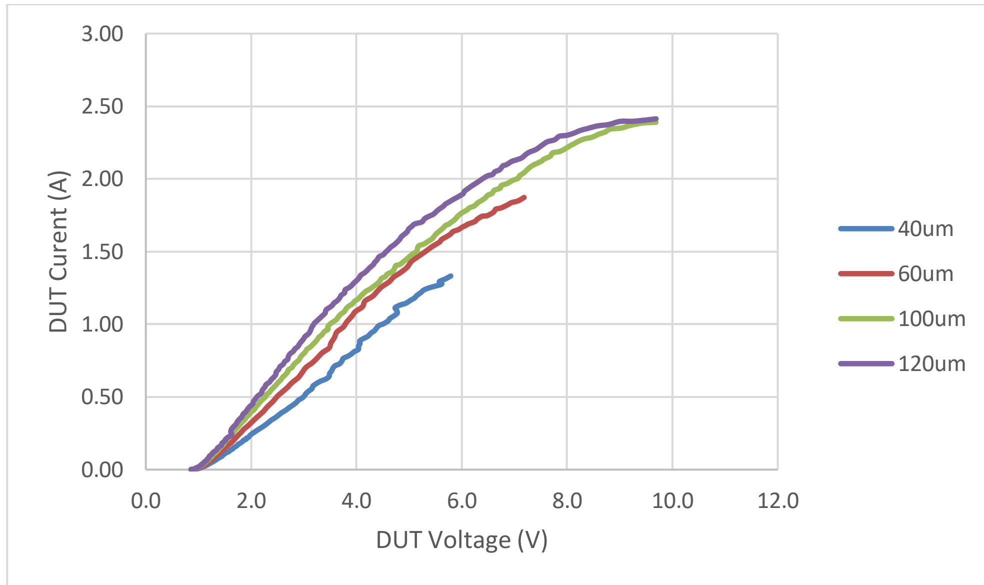


Figure 55 N+PW STI Diode TLP measured I-V curve under different width.

Figure 55 shows the TLP measured I-V curve of N+PW STI diode with different width as an example. The trigger voltage mainly depends on the junction turn-on voltage which is about 1V under quick transient pulse. But the current handling capability is increased with the width increasing.

Same measurements were applied to P+NW STI diode. All critical parameters are collected based on the following rules:

1. Trigger voltage is defined with the voltage when current reach 0.01A.
2. Breakdown voltage and current are located at the point when tested leakage current, after the pulse, is much higher than normal which means it's thermal broken.
3. The turn-on resistance is defined as the slope after device triggered.

Table 5-2 summarized the parameters of all STI diode structures. It shows the trigger voltages are almost constant which is manly depended on the doping level of junction. The turn-on resistance is decreased with the width increasing because current can go through a

Cellname	Vt1(V)	Vt2(V)	It2(A)	Ron(Ohm)
N+PW _40	1.04	5.79	1.33	3.42
N+PW _60	1.02	7.18	1.87	3.02
N+PW _100	0.96	10.91	2.41	2.74
N+PW _120	0.94	11.16	2.41	2.47
P+NW _40	1.03	6.14	1.57	3.20
P+NW _60	0.98	8.47	2.16	3.13
P+NW _100	0.95	10.46	2.43	2.90
P+NW _120	0.94	11.07	2.45	2.85

Table 5-2 Critical parameters of STI diode structures.

wider device with same distance. In the meanwhile, the peak current which shows the protection capability is also positively related to the width. There is no big different between P+NW and N+PW junction whose doping difference are same.

b) Gated Diode

Figure 56 shows the TLP measured I-V curve of N+PW Gated diode with different width as an example. Same as STI diode, the trigger voltage mainly depends on the junction

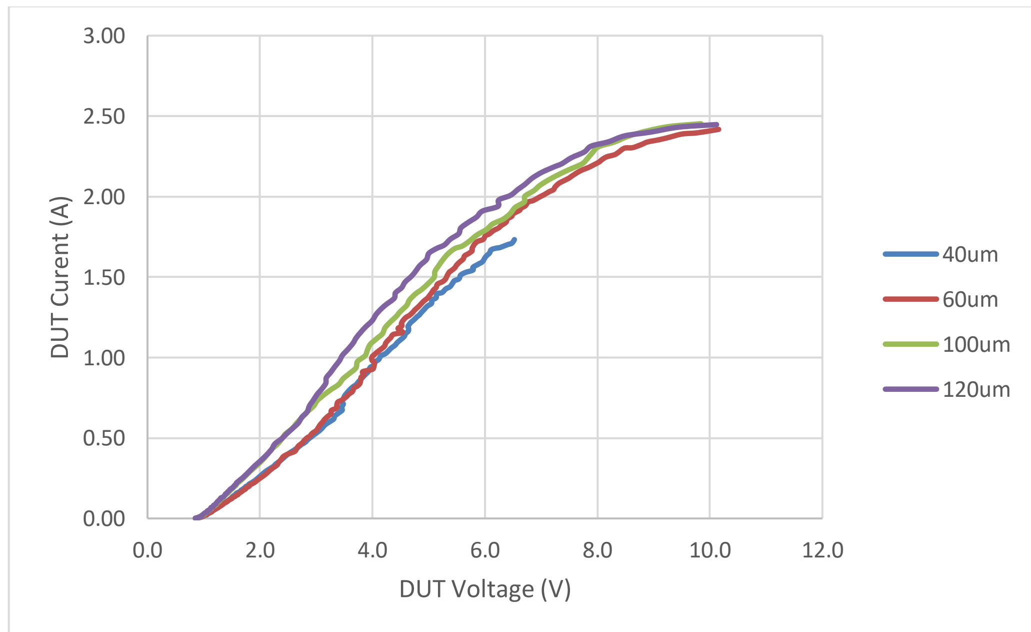


Figure 56 N+PW Gated diode TLP measured I-V curve under different width.

Cellname	Vt1(V)	Vt2(V)	It2(A)	Ron(Ohm)
Gated_N+PW _40	0.98	6.52	1.73	3.03
Gated_N+PW _60	0.96	10.16	2.42	2.82
Gated_N+PW _100	0.94	9.84	2.45	2.79
Gated_N+PW _120	0.93	10.12	2.45	2.53
Gated_P+NW _40	0.97	6.69	1.78	2.94
Gated_P+NW _60	0.97	9.78	2.38	2.79
Gated_P+NW _100	0.93	9.77	2.42	2.71
Gated_P+NW _120	0.93	10.56	2.44	2.69

Table 5-3 Critical parameters of gated diode structures.

turn-on voltage which is about 1V under quick transient pulse. In this case, some of the devices, due to the metal interconnects fails before the real device thermal breakdown, it is not obvious that the peak current doesn't increase when current reach about 2.4A.

Same measurements were applied to P+NW gated diode. All critical parameters are extracted using same definition as STI diodes showed in Table 5-3. The performances of gated diode structures are very similar to STI diode. But due to the shorter path of current flow by removing the STI, the turn-on resistance is smaller which gives less voltage with same ESD current.

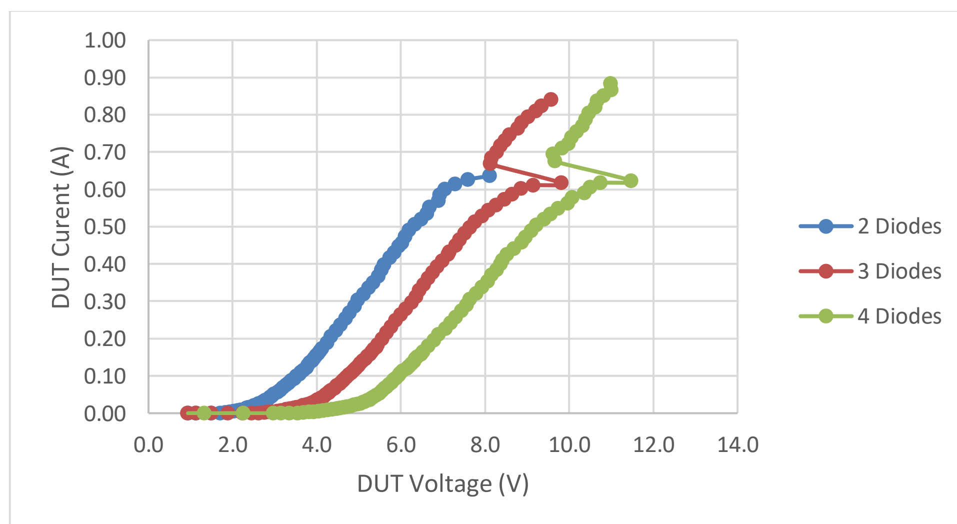


Figure 57 Diode string TLP measured I-V curve under different diode numbers.

Cellname	Vt1(V)	Vt2(V)	It2(A)	Ron(Ohm)
Diode_string _2	2.26	8.21	0.64	7.98
Diode_string _3	3.36	9.87	0.63	8.45
Diode_string _4	4.39	11.57	0.63	9.64

Table 5-4 Critical parameters of diode string structures.

c) Diode String

Figure 57 shows the TLP measured I-V curve of N+PW diode string with different diode number. The trigger voltage can be well controlled by diode number which is about 1V each. Comparing with three curves, it is not hard to find out that the turn-on voltage is almost same because the width of each series diode is equal. One of the diode in diode string will finally thermal breakdown first which gives almost same current handling capability.

Same extraction is used to define the critical parameters listed in Table 5-4. Since we use 0.01A current to define the trigger voltage, the V_{t1} of each diode is a little bit higher because series diode increases the turn-on resistance in the same time. The width of each diode is a little bit small which is hard to cover even 1kV HBM pulse in this case.

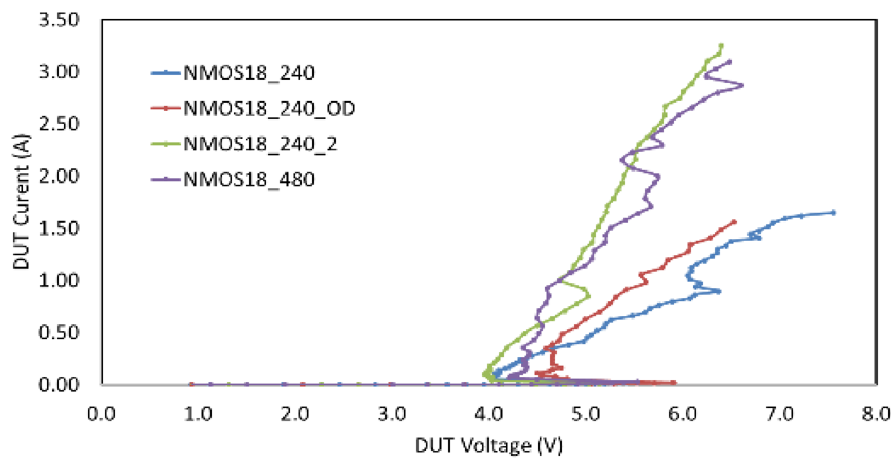


Figure 58 GGNMOS TLP measured I-V curve under different dimensions.

Cellname	Vt1(V)	Vh (V)	Ih (mA)	Vt2(V)	It2(A)	Ron(Ω)
GGNMOS18_240	5.52	4.03	65.3	7.55	1.65	2.22
GGNMOS18_240_OD	5.90	4.50	112.0	6.52	1.56	1.40
GGNMOS18_240_2	5.14	3.96	99.1	6.40	3.25	0.77
GGNMOS18_480	5.53	4.21	65.0	6.48	3.09	0.75

Table 5-5 Critical parameters of GGNMOS structures.

d) GGNMOS

TLP tested I-V curves of GGNMOS with different dimensions are depicted in Figure 58. The total width of these GGNMOS are 240um or 480um with one of which is two 240um NMOS in parallel. The overdrive (OD) device is designed for high voltage application which have a thicker gate oxide. It is obvious that the snapback curve can be found in all structures due to the parasitic BJT inside.

Table 5-5 summarizes the critical parameters defined before and two additional factors including holding voltage and holding current are also extracted. Those parameters are located at the point with the lowest voltage after triggered. The trigger voltage of GGNMOS is mainly depended on the reverse breakdown of N+PW junction. Then the substrate current accumulates the voltage to turn on the parasitic BJT resulting a deep snapback at an enough current accumulated point. The turn-on resistance of snapback devices is very small but also related to the width of MOSFET.

e) SCR

A deeper snapback structure called SCR is widely used for many high-power applications these days due to its low R_{ON}. The TLP I-V curve with different width are showed in Figure 59. The trigger voltage is very high for SCR structure due to its low doping junction which needs to reverse breakdown first before triggering two parasitic

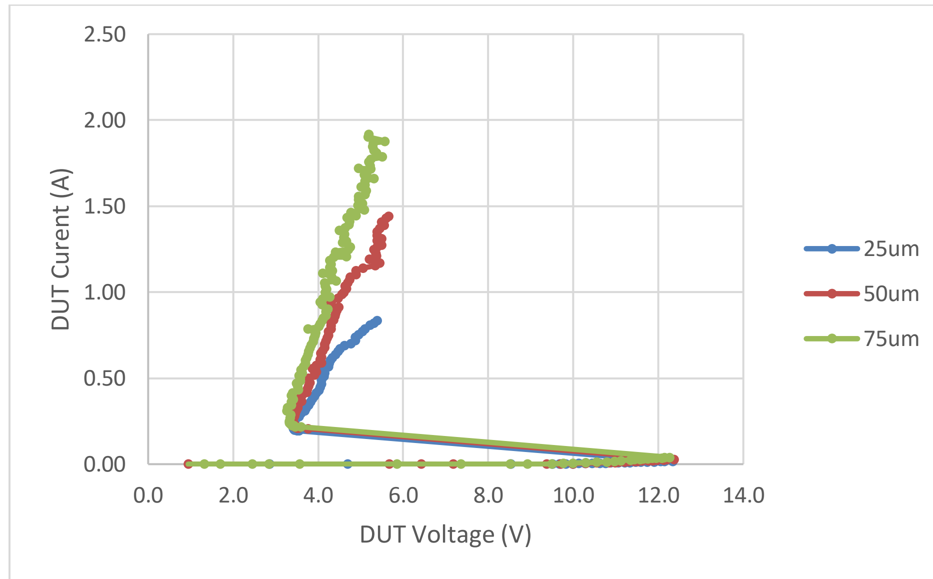


Figure 59 SCR TLP measured I-V curve under different width.

Cellname	Vt1(V)	Vh (V)	Ih (A)	Vt2(V)	It2(A)	Ron(Ω)
SCR _25	12.33	3.41	0.21	5.38	0.84	2.96
SCR _50	12.36	3.41	0.22	5.65	1.44	1.92
SCR _75	12.26	3.32	0.24	5.19	1.92	1.33

Table 5-6 Critical parameters of SCR structures.

BJTs. After the SCR triggered, the resistance inside is very small and the ESD current will pass through quickly.

The corresponding critical parameters are listed in Table 5-6 which shows a compatible performance except the trigger voltage is too high for mordern process. With 2kV protection level, SCR has the smallest area. However, the latch-up issue needs to be paid more attention which results continuous triggered even after ESD pulse finished.

f) DTSCR

Diode triggered SCR is used to compensate the disadvantage of high trigger voltage by using diode string in parallel. The diode number can limit the trigger voltage by intentional inducing current to substrate. In the meantime, it has a larger size but the current

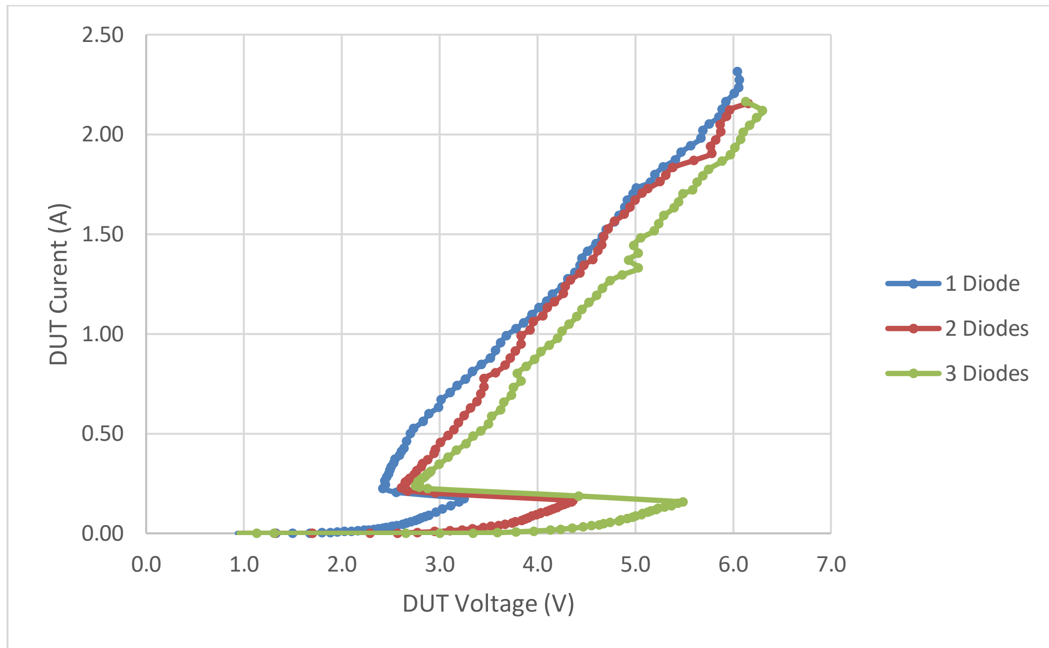


Figure 60 DTSCR TLP measured I-V curve under different trigger diode numbers.

Cellname	Vt1(V)	Vh (V)	Ih (A)	Vt2(V)	It2(A)	Ron(Ω)
DTSCR _1D	3.26	2.42	0.23	6.05	2.31	1.88
DTSCR _2D	4.37	2.61	0.23	6.16	2.16	1.77
DTSCR _3D	5.49	2.76	0.24	6.13	2.16	1.86

Table 5-7 Critical parameters of DTSCR structures.

mainly go through the triggered SCR with a low resistivity. Figure 60 shows the measured I-V curve under TLP testing. The triggering contribution can be observed with the trigger voltage drops rapidly. But the turn-on resistance is maintained because the triggered SCR has much smaller resistance which pass more current.

Critical parameters are listed in Table 5-7 which confirms that the trigger diodes only effect the V_{t1} and increase the area a little bit of course. These V_{t1} values are more useful to low supply, low gate breakdown voltage processes.

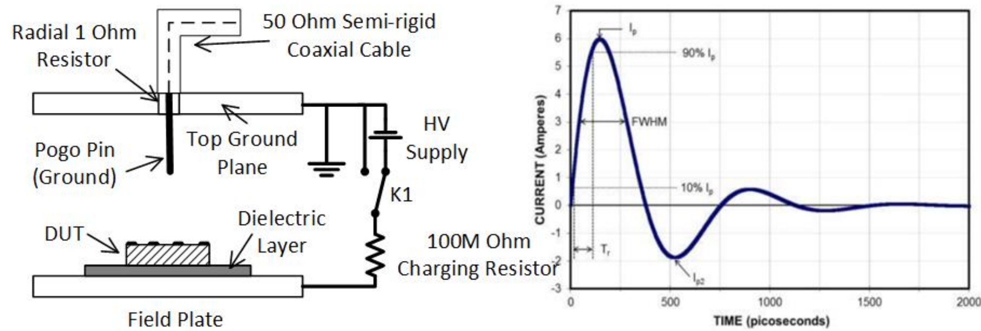


Figure 61 Simplified CDM tester schematic and its ESD discharging waveform.

5.3 Charged Device Model (CDM) ESD Protection

For ICs in advanced technology nodes, CDM ESD protection design is becoming a challenging design task. CDM ESD model describes an ESD discharging phenomenon where electrostatic charges accumulated inside an IC part will discharge when a pin is grounded mentioned in chapter 5.1, resulting in ultra-fast and vert short ESD pulses that will damage ICs [49], [50]. Fig. 61 shows a conceptual schematic for a CDM ESD tester per the ESDA/JEDEC standard and the typical CDM ESD transient source waveform whose first dominating high wave will damage an IC in most cases. In principle, a packaged IC is placed on an insulated plate and is first charged through field induction by a high voltage source, typically to a level of 125/250/500/1000V, then, discharged into the ground through a discharging pogo pin. While designing ESD protection structures to meet the relatively slower human body model (HBM) ESD protection requirements has become well-known, CDM ESD protection design is an emerging challenge to the IC industry because technology scaling-down has made advanced ICs fabricated in nanoscale processes more vulnerable to gate-oxide damages caused by the ultra-fast and very strong CDM ESD surges [51]. Since HBM and CDM ESD zapping tests are destructive and

cannot provide any I-V behaviors for ESD protection structures, the transmission line pulsing (TLP) testing method was developed to emulate the ESD zapping procedures, because TLP testing can reveal instantaneous ESD discharging I-V details, which are very useful for ESD protection design characterization and optimization. TLP testing is used for HBM ESD model that features typical transient waveforms with a rise time (t_r) of 10ns and duration (t_d) of 150ns. The ultra-fast CDM ESD testing is characterized by a very-fast TLP (VFTLP) system that uses ultra-fast pulse waveforms, typically, t_r of 100/200/400ps and t_d of 1-10ns. In practical CDM ESD protection designs, the key task is not only to accurately design the ESD-critical parameters, including the triggering voltage (V_{t1}) and breakdown voltage (V_{t2}), to meet the ESD design window requirement to ensure full-chip ESD protection success [52], but also to evaluate the overshoot voltage (V_{OV}) and response time (t_{re}), which are critical for ultra-fast CDM ESD protection. But there are some limits from VFTLP measurements which requires other tools like TCAD to help analysis and estimate the CDM performance before real package [53].

5.3.1 Combined VFTLP and TCAD analysis method

a) Quasi-static VFTLP testing

To demonstrate the VFTLP & TCAD combined method, SCR and 2 diode triggered DTSCR (2DTSCR) are used as an example whose VFTLP measured I-V curve is showed in Figure 62. As we known, the I-V curve for an ESD device under test (DUT) from TLP and VFTLP testing are considered quasi-static because the voltage and current values across the ESD device are obtained by averaging the center portion (normally an averaging window of 25% to 75%) of the measure transient waveforms in time domain. This average

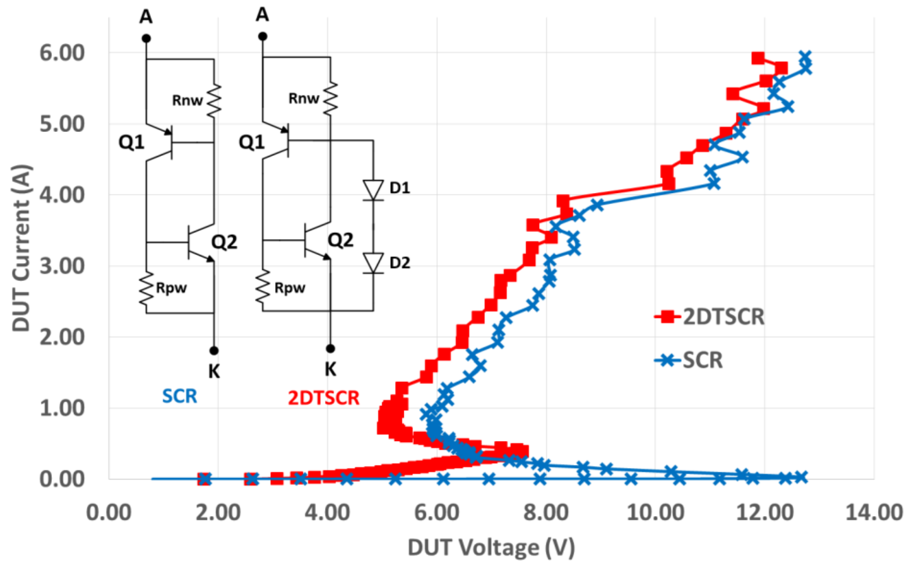


Figure 62 Measured I-V curves for 2DTSCR and SCR structures by VFTLP.

method is good for TLP testing because the pulse duration is long enough (t_d~150ns) and the waveform is typically flat. However, this averaging method for VFTLP testing of SCR and DTCSR structures has substantial errors for several reasons [54], [55]: the VFTLP pulses are very short (t_d ~1ns per CDM standard, and actual t_d~1.25ns for the Barth Model 4012 used in this work) and big overshoots occur at the beginning of the measured voltage

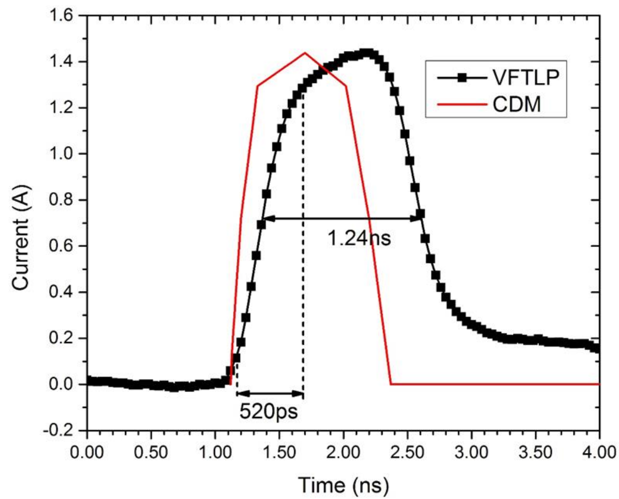


Figure 63 Comparison of ESD source pulse waveforms from CDM standard and VFTLP tester.

waveforms for DTCSRs, causing complex internal transistor structures turn-on speed. The traditional averaging method simply ignores the significant overshooting effect that leads to quasi-static I-V curve accuracy problem.

Fig. 63 compares the standard CDM testing pulse source waveform with that from a widely used Barth 4012 VFTLP tester, which is used to stress an ESD DUT structure during measurement targeting at CDM ESD protection level of 125V. The difference is obvious. While the VFTLP tester was made according to the CDM Standard, i.e., to produce a CDM pulse of t_r~200ps and t_d~1ns, the actual pulse waveform delivered to an ESD DUT structure has t_r~520ps and t_d~1.24ns, which clearly deviates from the JEDEC Standard (t_d=1ns ± 0.5ns and t_r < 400ps) . Since it is very fast and short, such a difference in the pulse source waveform results in significant measurement errors in VFTLP testing. Unfortunately, there is no better testing technique to obtain the required instantaneous I-V curves of ESD structures for CDM ESD model than the existing VFTLP technique before real packaged CDM testing.

b) TCAD transient CDM ESD simulation

Synopsys TCAD tool is used to study the transient behaviors of CDM protection capability with help of build-in physical models. 28nm process parameters and ESD waveform. For CDM ESD simulation, the key is to understand the overshoot voltage and response time of the protection structures in order to evaluate and optimize the design. TCAD transient simulation gives a detailed information with pulse injection including the breakdown and thermal effects. As an example showed in Figure 64, the transient I-V curves for a 2DTSCR and three diodes string (3DS) under CDM pulse. These two

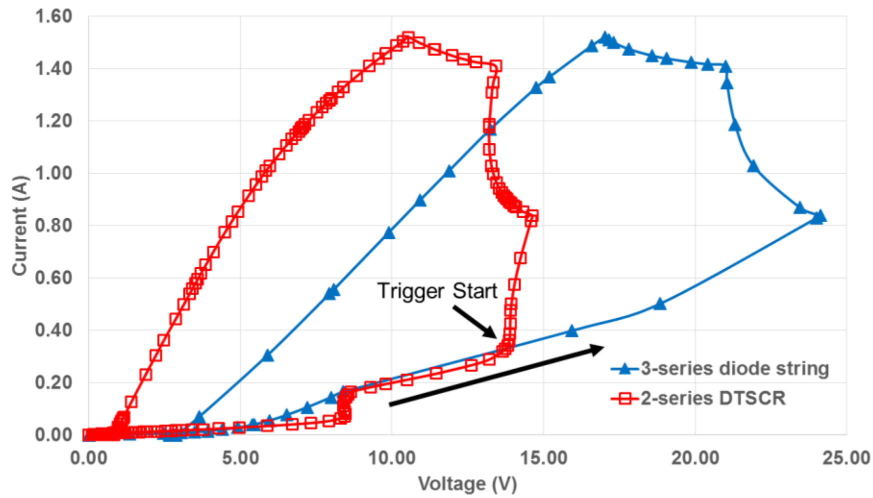


Figure 64 TCAD simulated transient I-V curve for 2DTSCR and 3DS under CDM stressing.

structures have same estimated triggering voltage due to the effective triggering diode in 2DTSCR is 3 (with one intrinsic). The DSTCR triggering mechanism is observed where the portion of I-V curve from the 3DS matches that for the STSCR before the internal SCR triggered.

By comparing TCAD simulated transient I-V curves for the DTSCR and corresponding 3-diodes string in Figure 64, an obvious resistance decrease can be seen when the current increases sharply, which means the SCR starts to conduct current. At the time, the SCR is not fully triggered yet and overshooting typically occurs. The TCAD results in Figure 65 shows the current density of the 2-diode DTSCR cross-section under CDM ESD stressing. The left part of the cross-section shows the triggering diodes, which conduct most current before SCR triggering. As the current increases, the embedded BJTs on the right side will be turned on, leading to SCR conduction (trigger start). After SCR being fully triggered, a conduction path with low ESD discharging R_{ON} will be formed inside the SCR to discharge the large ESD pulse, i.e., the DTSCR starts to work for ESD

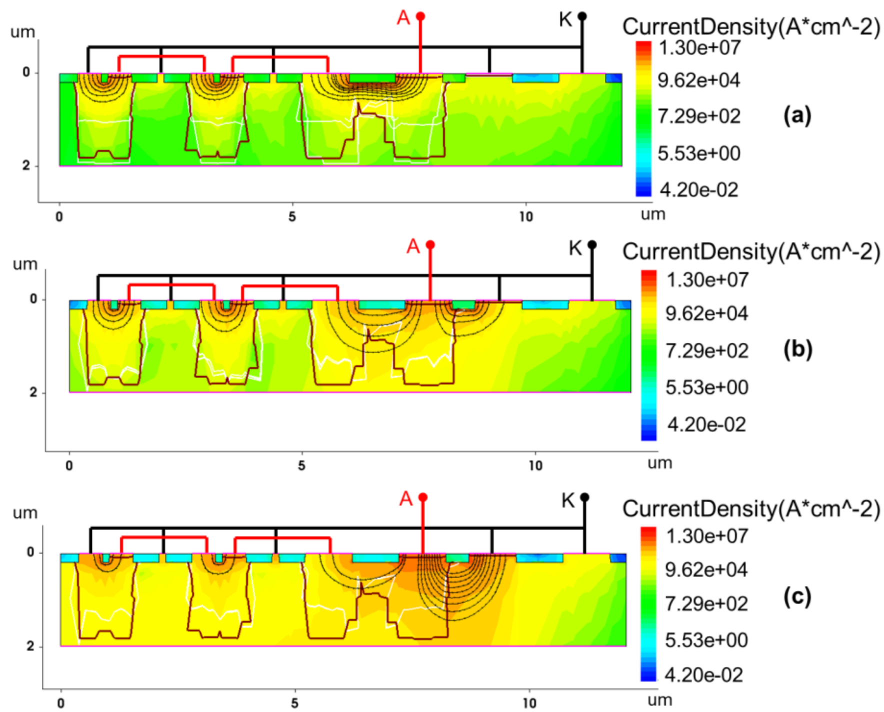


Figure 65 Cross-section views of ESD discharging current distribution across DTSCR by TCAD simulation shows the two-step DTSCR triggering mechanism: (a) current starts in the diodes (left) before triggering the SCR (right), (b) more current steers into SCR at its triggering threshold, and (c) most current conducts through the SCR after DTSCR fully turns-on.

protection. Obviously, there is a turn-on delay from the turn-on of the trigger-assisting diodes (trigger starts) to the full triggering of the SCR. This DTSCR triggering mechanism is depicted by the simulated conduction current distribution for the DTSCR structure by TCAD as shown in Figure 65. It is clear that, under the ESD stressing, the trigger-assisting diodes are turned on first and, as the diode current increases, it then triggers the SCR.

After the DTSCR is fully triggered, most ESD discharging current is steered into the SCR and little current conducts in the diodes. Therefore, we defined that the DTSCR is fully triggered as the moment when the current in the trigger-assisting diodes equals to that in the SCR. The "turn-on time" is from the pulse coming to that moment, and the "delay time" is from SCR trigger start to fully trigger.

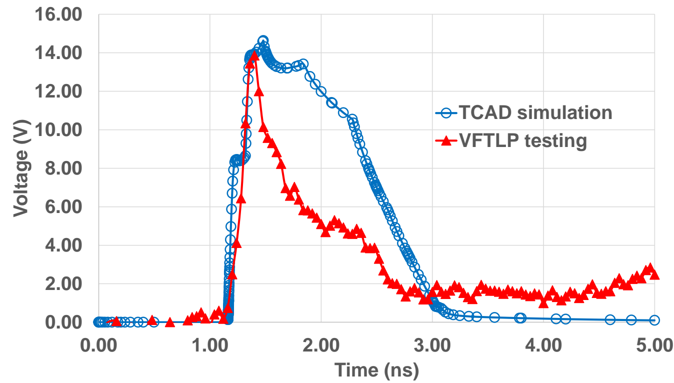


Figure 66 Comparison of voltage transient waveform of DTSCR by testing and TCAD simulation under VFTLP.

As mentioned above, the VFTLP pulse is a bit slower than standard CDM pulse. To estimate the real performance under real CDM testing, two steps with the help of TCAD simulation are needed.

First, a comparison of TCAD simulation with VFTLP pulse versus VFTLP testing results is performed to verify the accuracy of TCAD simulation. Figure 66 shows the voltage waveform of TCAD simulation and testing measurements under same VFTLP pulse. The voltage duration is a little bit over estimate, but the peak voltage is almost same which gives enough information for failure analysis. The peak voltage observed for both cases are about 14.6V, though the voltage peak seems to be sharper in VFTLP testing than that in TCAD simulation, possibly because of the difference of substrate size. Since CDM ESD damage is mainly associated with the overshoot high voltage, the failure risks shall be similar.

Second, a comparison between real CDM pulse and VFTLP by TCAD simulation is performed to estimate the ESD performance under CDM testing. Figure 67 compares the TCAD simulated voltage waveforms of DTSCR for both CDM and VFTLP zapping. Overshooting occurs in both CDM and VFTLP testing, though CDM case features a

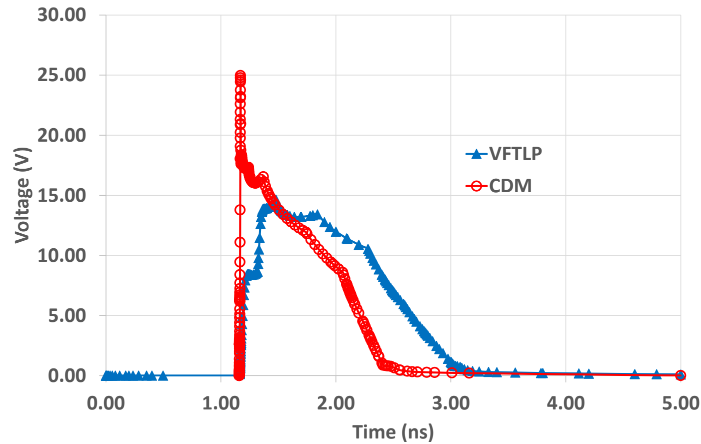


Figure 67 Comparison of TCAD simulated transient volatge waveform for 2DTSCR under VFTLP and CDM pulse.

sharper and higher overshoot. The peak voltage for CDM testing is about 25.0V compared to about 14.6V for VFTLP testing. The waveform duration for VFTLP testing is about 25% longer than that for CDM zapping. The significantly severer overshooting in CDM zapping may be attributed to the fact that the rise time and duration of the CDM source pulse (200ps & 1ns) is shorter than that of VFTLP testing (460ps & 1.25ns) for the same 125V ESD protection target, causing a substantial sharper and higher overshoot in CDM zapping simulation. Based on the VFTLP testing results, about 80% margin is needed for real CDM protection which because of the ultra-fast pulse from real CDM zapping.

This method gives a great reference for CDM protection design before packaged system level CDM testing but VFTLP stand-alone testing with the help of TCAD simulation. The overshoot voltage can be well estimated by comparison of transient waveforms.

5.3.2 Overshoot issue in diode based ESD protections

Particularly, the ultra-thin gates can be sassily damaged by the super-fast CDM transient spikes, which is an emerging challenge in ESD design for advanced ICs and

becomes a sizzling research topic in the field [56]–[58]. Since the overshoot voltage issue is the biggest challenge for CDM ESD protection, understanding of the reason of it is critical to compensate and minimize it during CDM protection design. During the research of CDM protection methodology, I've notice that the diode based ESD structures have more obvious overshoot. As a result, the study of a set of diodes, diode string and diode triggered SCR structures is performed to reveal the cause of overshoot phenomenon.

a) Voltage overshoots in diode structures

It was believed that, at high frequency, a PN junction diode may exhibit inductive characteristics possibly due to conductivity modulation [59]. The transient voltage overshoots observed in some diode ESD protection structures under CDM stressing may be associated to the high-frequency inductive effect due to the very fast rise time and very

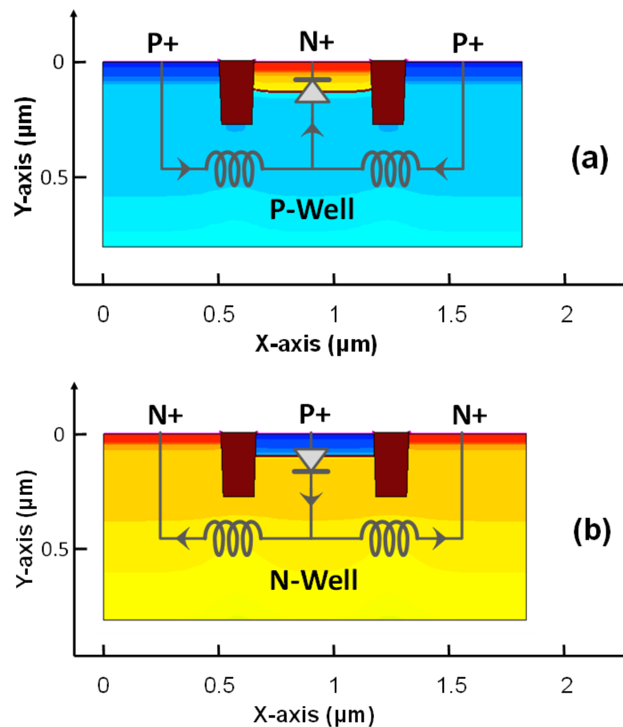


Figure 68 Cross-sections for N+PW (a) and P+NW (b) STI diodes with doping concentration by TCAD simulation.

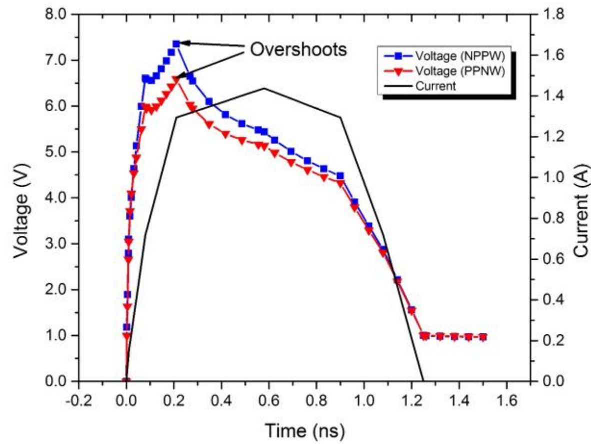


Figure 69 The transient voltage waveforms by TCAD simulation for STI diodes under 125V CDM stressing.

short period of a CDM pulse. To understand this CDM overshooting behaviors, a set of diodes ESD protection devices, including N+/P-well (NPPW) and P+/N-well (PPNW) with both STI and gated isolation were designed and fabricated in a foundry 28nm CMOS technology for investigation. A combined method with TCAD simulation and VFTLP testing is used to understand the phenomena. Figure 68 shows the simulated cross-sectional structures for STI N+PW and P+NW diodes. The equivalent circuit includes the inductive effect modeled by the embedded inductors which may delay the trigger under fast CDM pulse. The transient voltage waveform simulated two devices is presented in Figure 69. From the time-domain CDM simulation results, it is readily observed that voltage overshoots of up to 7V for both ESD didoes occurred well before the CDM pulse peaks at about 200ps. This phenomenon means that it takes some time to fully trigger the ESD diodes under the CDM stressing, likely due to the embedded inductive delay effect, and the pre-triggering high impedance causes fast voltage accumulation across the intrinsic PN junction, therefore, leads to a voltage overshoot. Since the transient voltage overshoot

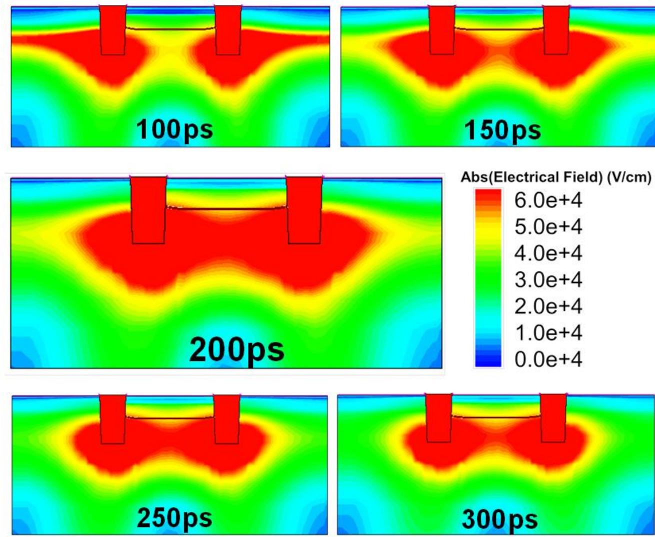


Figure 70 Simulated transient electrical field distribution for N+PW STI diode.

occurs well before the ESD protection structure is turned on, the large voltage spike may cause a dielectric damage to the CMOS gate.

To understand the mechanism of overshoot voltage, I explored the transient electrical behaviors by TCAD simulation. Figure 70 depicts the simulated transient electrical field density inside N+PW diode at different times before and after the peak overshoot around 200ps. Dynamic observation reveals that the high electrical field area gradually moves and reaches it at 200ps when the voltage peaks. This may attribute to the assumption that, due to the embedded inductance along the substrate introduces the delay in triggering the diode under CDM stressing. After the ESD diode is turned on, it creates a low-impedance conduction channel to discharge the CDM pulse quickly, hence, the electrical field density inside the diode will drop accordingly.

Same simulation is applied to gated diode whose cross-section is showed in Figure 71 using N+PW as an example. The transient voltage waveform of gated diode in Figure 72

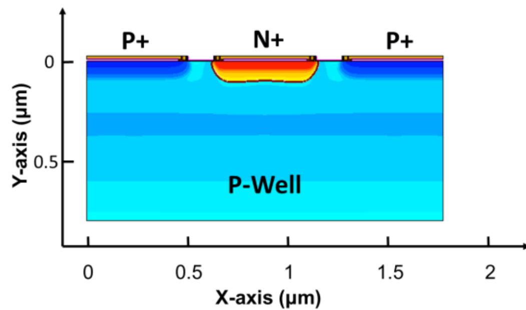


Figure 71 The simulated x-section for a sample N+PW gated diode.

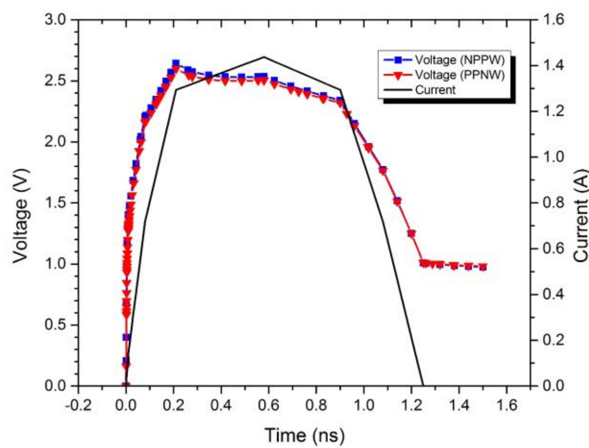


Figure 72 TCAD simulated voltage waveform for gated diodes under 125V CDM pulse.

shows a much less voltage overshoot which confirms my consideration with less path inductance. This can be explained that the actual internal conduction path for STI diode is much longer than that in a gated diode due to the trench depth. Therefore, the possible internal inductive effect is much stronger for STI diodes. This can also be confirmed by the simulated electrical field density depicted in Figure 73 where the dense electrical field reaches the intrinsic PN junction at around 100ps, much earlier than the 200ps observed in STI diode counterpart.

The VFTLP measurement is also applied to these four structures but based on the analysis before, the pulse is slower than real CDM pulse which is less stressful showed in

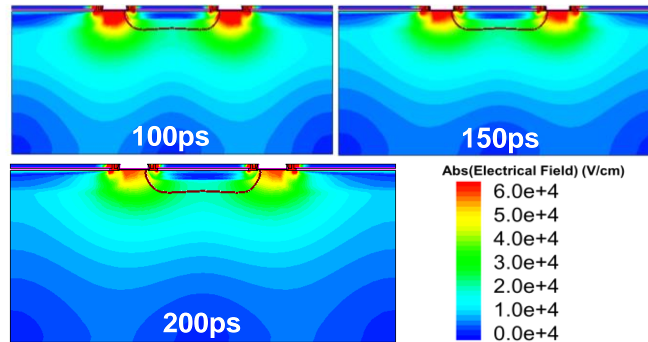


Figure 73 Simulated transient electrical field distribution for a sample N+PW gated diode.

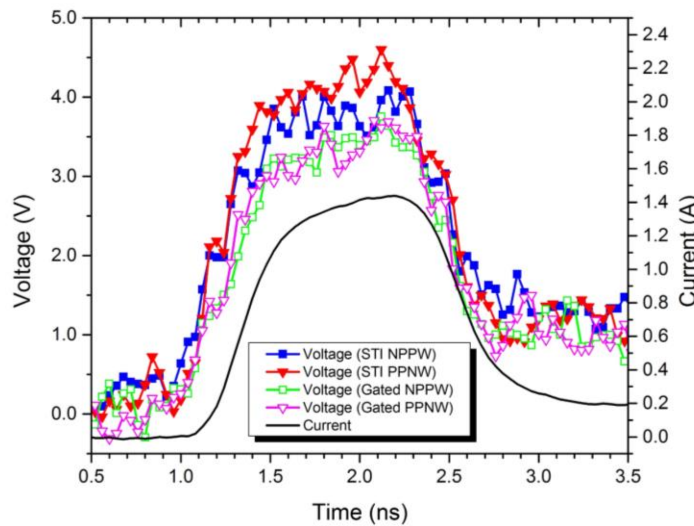


Figure 74 VFTLP-measured transient voltage and current waveforms for sample ESD diodes.

Figure 74. The overshoot effects are much less than simulated CDM circumstances. However, the voltage level of gated diodes is still smaller than STI diode because the turn-on resistance of gated diode is also smaller. As such, the actual VFTLP tester failed to catch the possible transient overshoots in measurement, causing a serious transient estimation error.

b) Diode based protection structures for CDM protection

To further understand the transient voltage overshoot mechanism possibly due to the built-in inductive modulation effect in diodes, we further investigated a few diode

derivatives, i.e., diode string and DTSCR ESD protection structures where the embedded diodes determine the ESD triggering. The parasitic inductance along the internal ESD path to intrinsic PN junction of diode-based structures will be associated with the possible transient voltage overshoots.

I. Diode string structures

A diode string ESD protection structure is used to boost the ESD triggering voltage (V_{t1}) to accommodate the core circuits. While a diode-string is advantageous in terms of simplicity and lower parasitic capacitance, it increases the total series resistance that worsens the overheating effect during ESD stresses. Importantly, it is expected that the

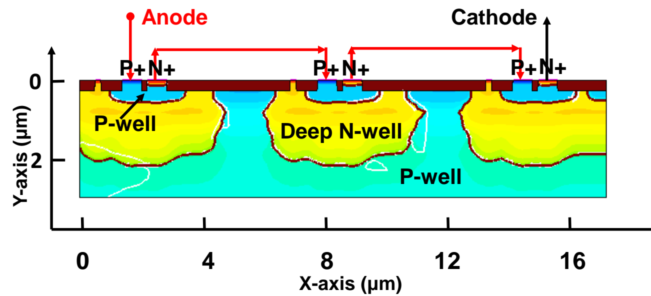


Figure 75 Cross-section of a sample 3-diode string structure with doping concentration.

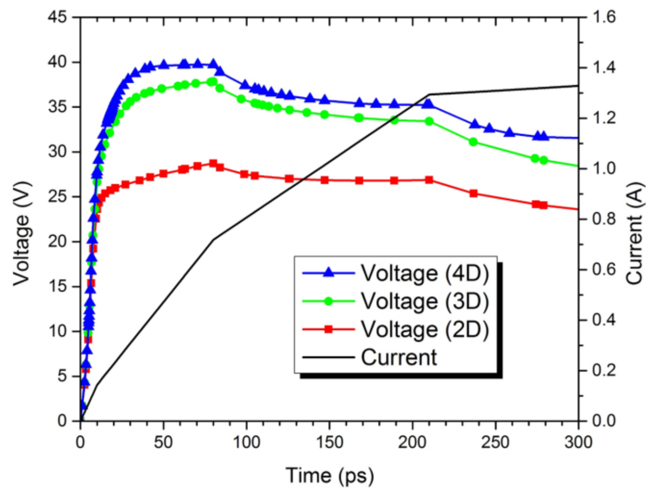


Figure 76 TCAD-simulated voltage waveform for diode string structures with different diode number.

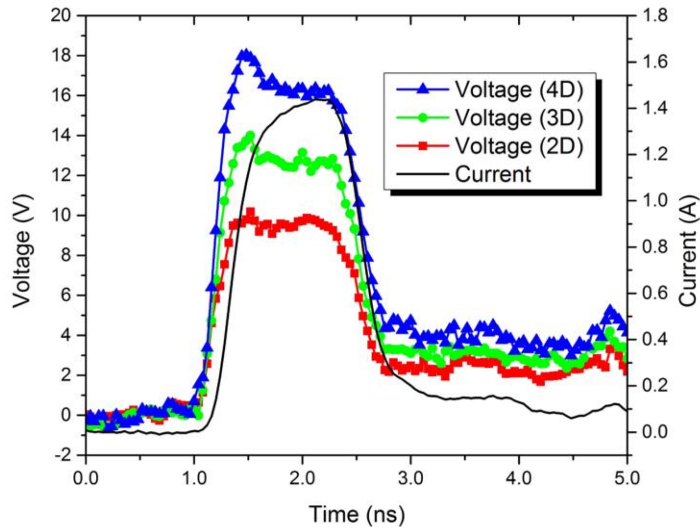


Figure 77 VFTLP-measured transient voltage and current waveforms for sample ESD diode string.

parasitic built-in inductance effect in a diode string ESD protection structure will be worse due to a longer conduction path. Various diode string ESD protection structures with 2, 3 and 4 diodes were designed and fabricated for a comparison study. Figure 75 shows a sample 3 diodes diode string cross-section in TCAD simulation. Figure 76 shows the TCAD-simulated transient voltage responses for the 2/3/4-diode diode string ESD protection structures under 125V CDM stressing. The zoom-in figure within up to the initial 300ps allows to focus on the transient voltage details during the ESD triggering phase. It is observed that the level of transient voltage overshoots varies according to the number of diodes in a diode string and can reach to 40V for a 4-diode diode string per TCAD simulation. The diode-string structures designed all failed 125V CDM ESD stressing due to much higher total series discharging resistance that led to overheating. Figure 77 presents the measured transient voltage waveforms for these diode string structures from VFTLP testing. Clear voltage overshoots were observed for these diode string ESD protection structures, e.g., ~18V for a 4-diode diode string, which is though

lower than that from TCAD simulation. The discrepancy between simulation and testing may be largely attributed to the difference of stress pulse speed. In general, both simulation and VFTLP testing clearly indicate that the transient voltage overshoots are directly related to the number of didoes in a diode string structure, which is due to the varying parasitic inductance effect along the varying-length internal ESD discharging paths.

II. DTSCR structures

An SCR ESD structure is efficient in ESD discharging due to its unique snapback ESD discharging I-V behavior and very low discharging resistance. However, an intrinsic SCR structure typically has high V_{t1}, making SCR ESD protection structure not suitable for low-voltage ICs. A DTSCR utilizes embedded diode(s) to reduce the ESD triggering voltage, hence, attracts lots of interests in advanced IC designs. The embedded didoes may be of various fashions in a SCR ESD protection structure to meet the design requirements. Apparently, the internal parasitic inductance associated the embedded diode(s) may cause

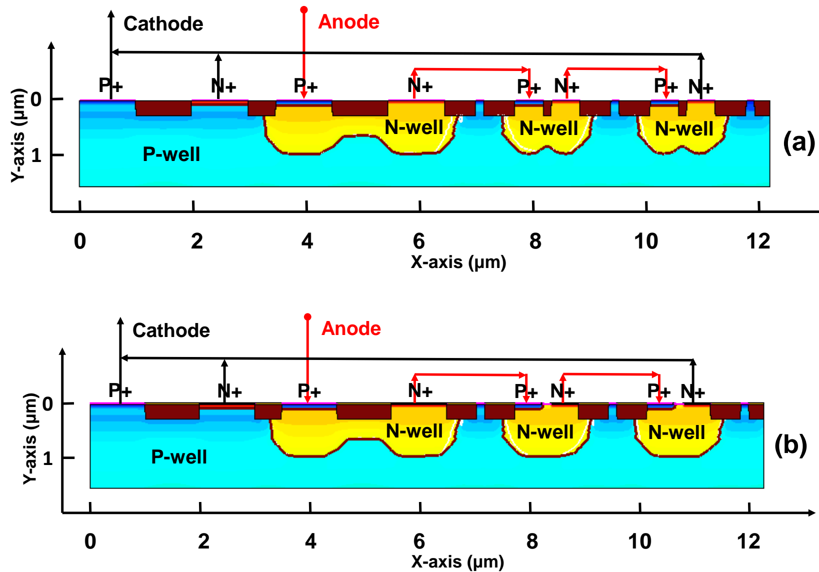


Figure 78 TCAD-simulated cross-sections for sample 2-diode STI DTSCR (a) and 2-diode gated-diode DTSCR (b) ESD protection structures.

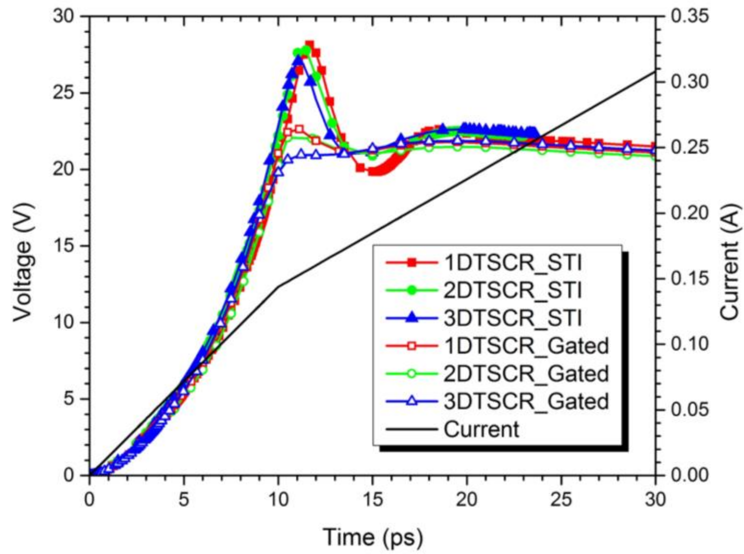


Figure 79 TCAD-simulated voltage waveforms for sample DTSCR ESD protection structures under 125V CDM ESD stressing.

delay in ESD triggering and transient voltage overshoot in a SCR structure. Two kinds of triggering diode including STI and gated diodes are used with different diode numbers are demonstrated with TCAD simulation and VFTLP measurement.

Figure 78 illustrates the cross-section and external connections for a sample DTSCR using two PPNW diodes (2DTSCR) of STI and gated isolation, respectively. During ESD zapping, the embedded didoes must be triggered first, which leads to 2DTSCR turn-on eventually. The TCAD-simulated transient waveforms under 125V CDM zapping is showed in Figure 79, zoomed-in for the initial triggering phase up to 30ps for clarity. Further, the transient voltage overshoots for the STI-diode-triggered DTSCR structures are much higher than that for the gated-diode-triggered DTSCR structures, due to the longer conduction path in the STI-type DTSCR structures than that in gated-type DTSCR structures. Figure 80 presents the VFTLP-measured transient voltage waveforms for the same DTSCR ESD protection structures. Unlike in the diode string ESD protection

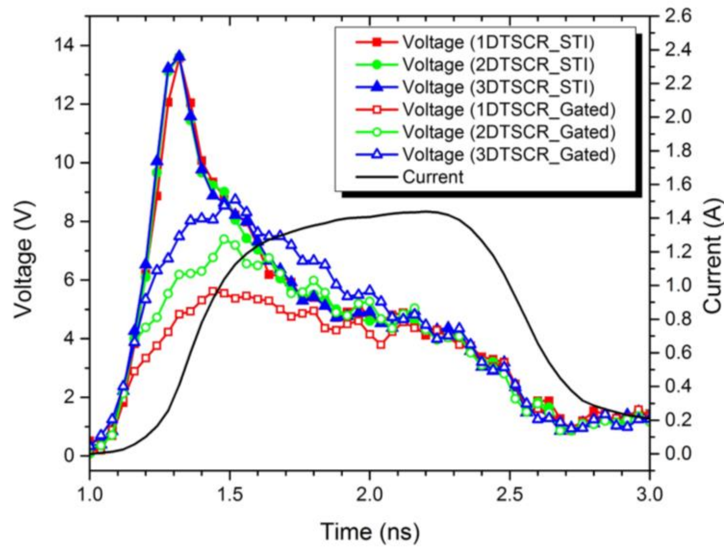


Figure 80 VFTLP-measured voltage waveforms for sample DTSCR ESD protection structures.

structures, the DTSCR structures show clear transient voltage overshoots in VFTLP measurement. This may be because the combination of SCR and didoes in the DTSCR structures slows down the ESD triggering procedure to the point that the VFTLP tester can effectively catch these voltage overshoots in measurements. On the other hand, the substantial difference in the voltage overshoots observed for the STI-type DTSCR structures and the gated-type DTSCR structures further supports the model that the internal parasitic inductance plays a key role in delaying the ESD triggering procedure that, in turn, causes transient voltage overshoots in CDM ESD zapping. The discrepancy between simulation and measurement are attributed to two factors: the incapability of an actual VFTLP tester in catching the ultra-fast CDM pulse details and the insufficient TCAD calibration.

Chapter 6 ESD Parasitic Effects

All ESD protection design for ICs emerges as a major challenge for high frequency and high-speed ICs due to its unavoidable parasitic effects including leakage and capacitance [45], [60]. While different protection structures have varying parameters to meet different design window, they introduce same parasitic effect, C_{ESD}, which will affect the core IC performance more or less. Such influences on capacitance is significant for high data rate and high frequency ICs because the circuit functions, e.g., frequency, data rate, signal integrity, input/output impedance matching, power delivery and bandwidth, etc., are very sensitive to capacitance. Accurate extraction of parasitic capacitance is very important during the design so that it is possible to compensate it before silicon tape out. But different structures have different capacitance levels and different application requires different ways to reduce the parasitic effects.

This chapter explains the parasitic capacitance induced from different ESD protection structures in 28nm CMOS. An enhanced accurate extraction method is introduced to reach a better understanding of parasitic effects [61]. ESD-IC co-design is required to enhance both the ESD capability and I/O ports' performance. Here, the mm-wave switches for 5G application are used to demonstrate the co-design concepts [62].

6.1 ESD Parasitic Capacitance

Normally, to extract the parasitic capacitance of stand-alone ESD devices, a testchip is needed to evaluate the ESD performance and capacitance in the meantime. The capacitance to be analyzed can be divided into three parts:

1) The capacitance between metal routing induced by testing pattern layout called C_{metal}.

2) The capacitance between the pad and substrate called C_{pad}.

3) The capacitance comes from the ESD device itself, like junction capacitance, called C_{ESD}.

For the C_{metal} and C_{pad}, we can estimate the capacitance from the EDA tools with post-layout extraction. But the C_{ESD}, we cannot get it directly from basic simulation but can use de-embedded methodology to measure the capacitance using network analyzer.

To de-embedded the capacitance from pad and metal connection, two test patterns, A and B, showed in Figure 81 are used to extract the stand-alone ESD device capacitance. The ESD structure is removed in pattern B to get the Y parameter of metal and pad. The simplified schematic and Y parameters are showed in Figure 82 from which the C_{ESD} Y parameter can be picked out, de-embedded, by the equation below. The Y parameter of two patterns are measured using E8363B network analyzer.

$$Y_{ESD} = Y_{patternA} - Y_{patternB}$$

$$C_{ESD} = \frac{IM(Y_{ESD})}{2\pi f}$$

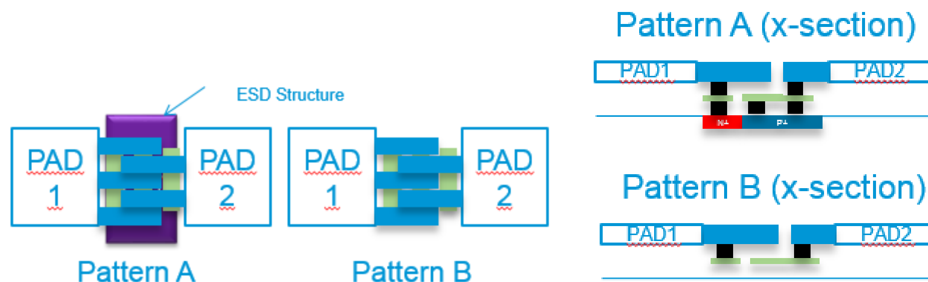


Figure 81 De-embedded method patterns for stand-alone C_{ESD} extraction.

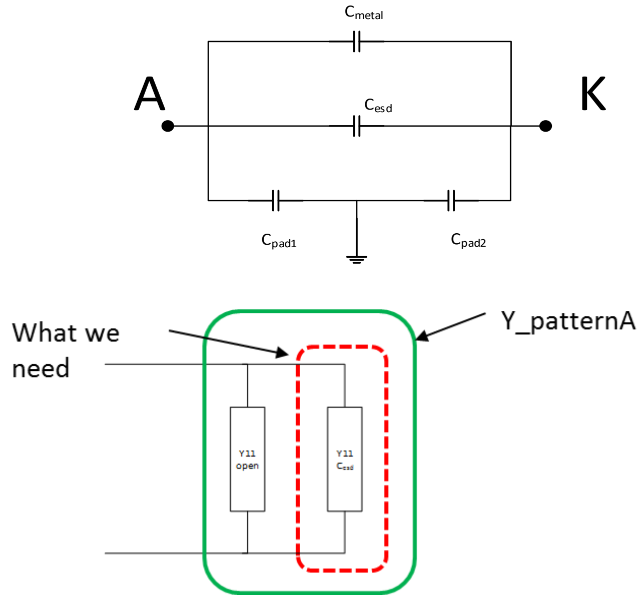


Figure 82 Equivalent circuit of de-embedded method with Y parameter calculation.

As a demonstration, several commonly used ESD structures illustrates the capacitance value with different width and type.

a) STI diode

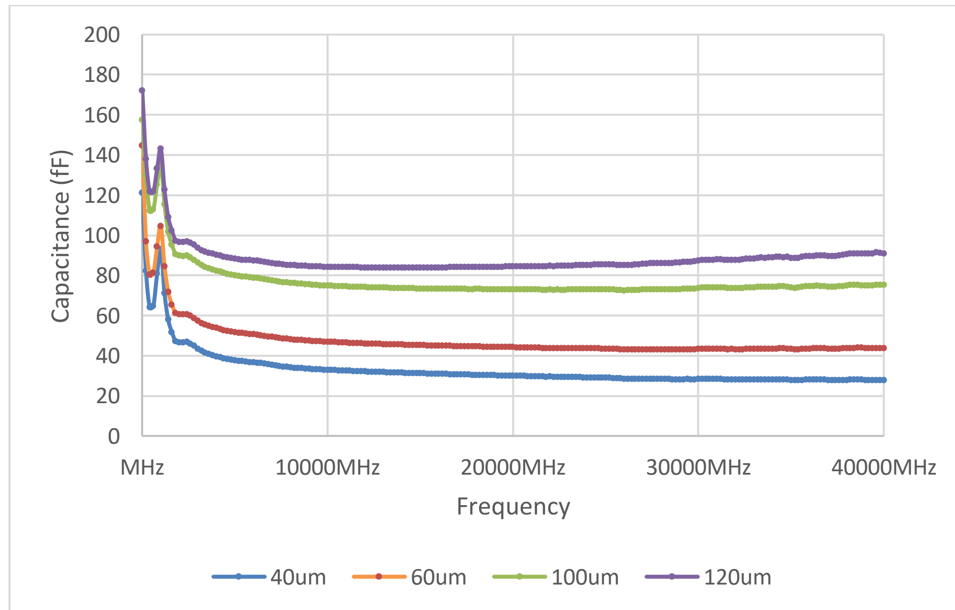


Figure 83 Extracted capacitance of N+PW STI diode from 10MHz to 40GHz.

Cap. (fF) @10GHz	40 um (1KV)	60 um (2KV)	100 um (3KV)	120 um (4KV)
N+PW STI Diode	32.81	46.78	74.76	84.15
P+NW STI Diode	46.77	65.84	105.71	124.85

Table 6-1 De-embedded capacitance of STI diode at 10GHz.

The structure of STI diode are introduced in Chapter 5 with critical paramters. Another important parameter that designers cares a lot is the parasitic capacitance when its on the off-state. Figure 83 shows the de-embedded capacitance of N+PW STI diode structures from 10MHz to 40GHz with different width. It is clear that the capacitance increases with the width (size) increasing because both of the substrate capacitance and junction capacitance are positive related. The capacitance level of STI diode is tens of femtofarad listed in Table 6-1. It can be found that the capacitance of P+NW diode is larger than N+PW which may because of the doping level different between two kinds of junctions.

b) Gated diode

Gated diode with its advantage of turn-on resistance and CDM performance gives more opportunity in ESD protection. However, due to the dummy gate to avoid STI

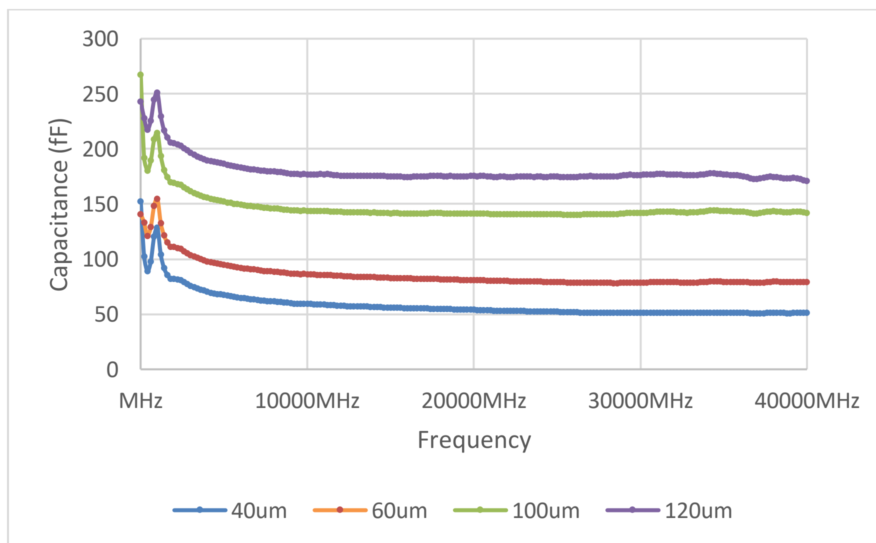


Figure 84 Extracted capacitance of N+PW gated diode from 10MHz to 40GHz.

Cap. (fF) @10GHz	40 um (1KV)	60 um (2KV)	100 um (3KV)	120 um (4KV)
N+PW gated Diode	59.29	86.22	143.71	176.76
P+NW gated Diode	70.28	102.64	154.35	184.56

Table 6-2 De-embedded capacitance of gated diode at 10GHz.

fabrication, the parasitic capacitance is much larger than normal STI diodes. Figure 84 shows the de-embedded capacitance of N+PW gated diode structures from 10MHz to 40GHz with different width. It is clear that the capacitance increases with the width (size) increasing because both of the substrate capacitance and junction capacitance are positive related. The capacitance level of gated diode is about 100fF listed in Table 6-2. It can be found that same as STI diodes the capacitance of P+NW diode is larger than N+PW which may because of the doping level different between two kinds of junctions.

c) Diode string

Diode string with series diodes actually reduce the capacitance by series the capacitors from each diode. However, the extracted capacitance values of 2/3/4 diodes string decreases not obviously from Figure 85. But the capacitance level is less than 20fF with

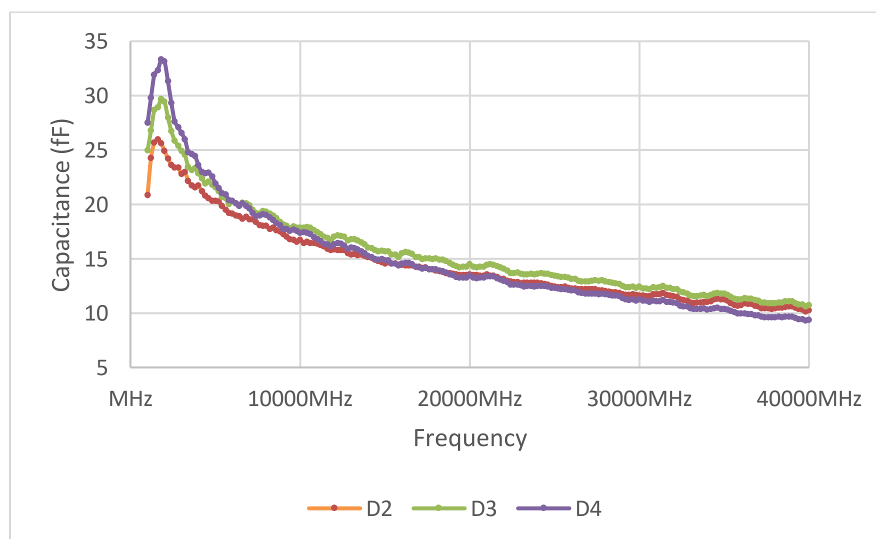


Figure 85 Extracted capacitance of N+PW diode string from 10MHz to 40GHz.

Cap. (fF) @10GHz	1D	2D	3D	4D
N+PW gated Diode	≈ 20.00	16.70	17.85	17.37

Table 6-3 De-embedded capacitance of diode string at 10GHz.

tunable triggering voltage. The capacitance of diode string with different diode number from 1 to 4 are listed in Table 6-3. It shows that the capacitance is almost constant even with diode number keep increasing which may because the junction capacitance is not dominated in this situation.

d) GGNMOS

The parasitic capacitance of GGNMOS is much higher than other structures due to its unavoidable grounded gate with the large size for ESD protection. Figure 86 shows extracted capacitance level is about hundreds of femtofarads which is unafordable for high speed I/O application. It's also found that the capacitance is mainly relied on the size of MOSFET that indicated in the name of those curves. The detailed value of capacitance are listed in Table 6-4 with different width including 240um, 240um with thicker gate, two

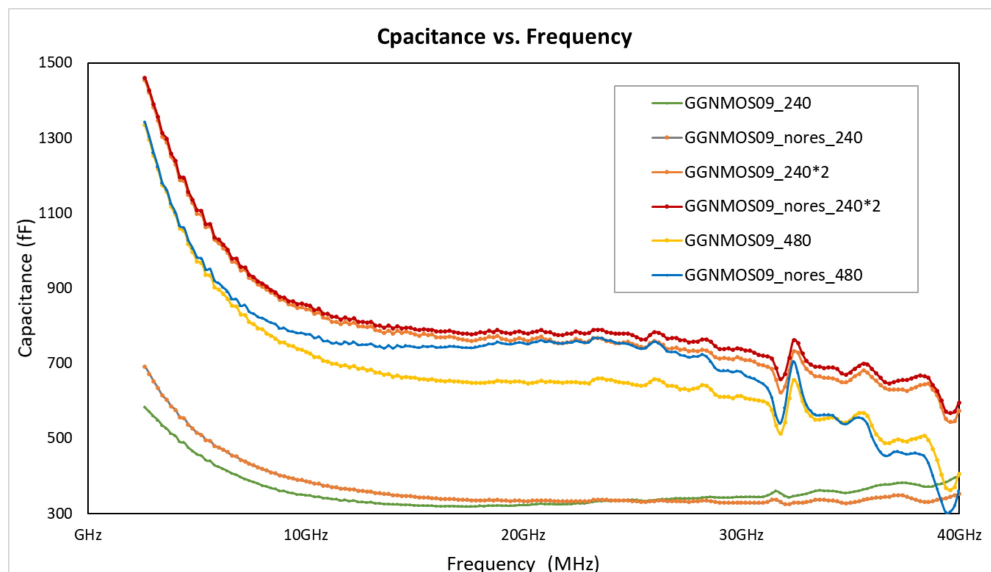


Figure 86 Extracted capacitance of GGNMOS from 10MHz to 40GHz.

Cap. (fF) @10GHz	_240	_240OD	_240*2	_480
GGNMOS18	290	267	603	679

Table 6-4 De-embedded capacitance of GGNMOS at 10GHz.

parralell 240um and 480um. It shows that thicker gate give more distance between two nodes of gate capacitor and twice the width gives about two times capacitance. However, attribute to the snapback I-V curve and great current handling capability, the GGNMOS are still wildly used for DC I/O protection like power supply or control I/O pads.

e) SCR

To balance the capacitance an power handling capability, SCR structures give a low resistance current path with small capacitance level smaller than 100fF showed in Figure 87 with different length. The capacitance at high frequency end is only about 30fF which is very attractive to mm-wave RFICs. The accurate values of each width are listed in Table 6-5 indicates smaller capacitance than STI diodes with same protection level. The only problem comes from its high trigger voltage which is hard to use in mordern low power technology.

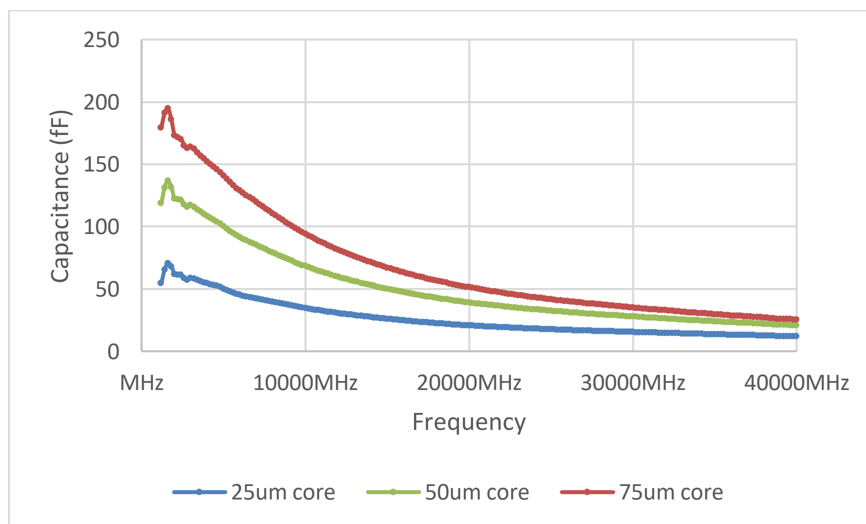


Figure 87 Extracted capacitance of SCR from 10MHz to 40GHz.

Cap. (fF) @10GHz	25um (>1KV)	50um (>2KV)	75um (>3KV)
SCR	34.71	68.44	94.41

Table 6-5 De-embedded capacitance of SCR at 10GHz.

f) DTSCR

If adding the triggering diode to reduce the V_{t1}, the parasitic effect will definitly change due to the size and complex connection. Figure 88 shows the measured capacitance from 1 to 3 diodes triggered SCR whose capacitance is almost same which is about 50fF. The width of SCR is 50um whose stand-alone capacitance is 68.44fF at 10GHz which means the capacitance is actually smaller due to the partially series connection of diode string whose capacitance is only about 20fF. Then the total capacitance decreases and the high triggering voltage issue has been solve in the sametime. This makes DTSCR structure paid more attention by mordern ESD design. Table 6-6 lists the capacitance value of STI

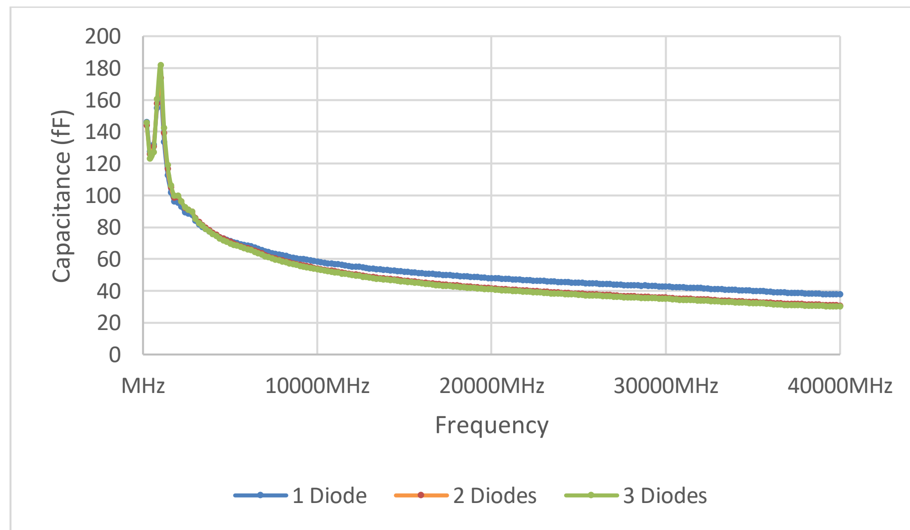


Figure 88 Extracted capacitance of STI DTSCR from 10MHz to 40GHz.

Cap. (fF) @10GHz	1 Diode	2 Diodes	3 Diodes
STI DTSCR	58.11	53.94	53.38
Gated DTSCR	79.45	70.81	70.43

Table 6-6 De-embedded capacitance of DTSCR at 10GHz.

diode triggered SCR and gated diode triggered ones the only difference of which is that the STI is replaced by dummy gate to reduce the length of current flow path. The gate structure gives more capacitance but since the current mainly go through SCR after it's triggered, the current path of tiggering diodes is not critical as in diode only protection.

6.2 Enhanced Parasitic Extraction

As introduced in Chapter 6.1, the de-embedded method for ESD capacitance extraction use the device pattern with pads and interconnections (Pattern A) and open pattern without ESD structure (Pattern B). The calculation is straight forward with Y parameters of two patterns. However, this method assumes that both the connection and ESD structures have only capacitance. While this simplified de-embedding ESD testing method works well for ESD devices with small C_{ESD}, it introduces errors for large C_{ESD} because it ignores the parasitic inductance effect associated with metal interconnects (often long metal wires for GS test patterns). To have a more accurate de-embedded methodology, another short circuit pattern is needed to extract the inductance and resistance form interconnecting metals and pads. But this requires at least three same testchip pattern area which is limited to the budget. An enhanced extraction method with the help of HFSS can improve the efficiency and accuracy of capacitance.

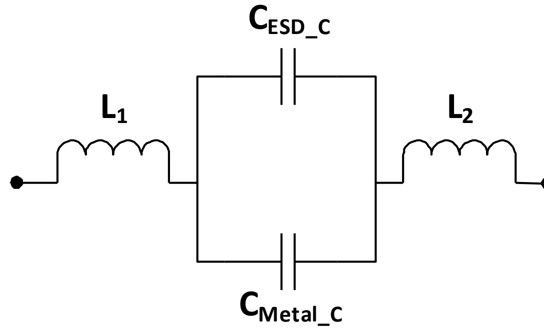


Figure 89 Equivalent circuit models for new de-embedding C_{ESD} extraction methods.

6.2.1 New C_{ESD} Extraction Method

As depicted in Figure 89 for its equivalent circuit, two series inductors associated with the out-of-ESD metal interconnects are added to the existing C_{ESD} extraction circuit model. A calibration procedure using the two series inductors, L_1 and L_2, serves to calibrate the C_{ESD} extracted directly from testing using an "open" dummy only, following the formula below,

$$C_{ESD_C} = \frac{C_{DUT_T}}{1+\omega^2 LC_{DUT_T}} - \frac{C_{Metal_T}}{1+\omega^2 LC_{Metal_T}}$$

$$< \frac{C_{DUT_T} - C_{Metal_T}}{1+\omega^2 LC_{Metal_T}} < C_{ESD_T}$$

where ω is the angular frequency of testing signal, L is the inductance of the metal interconnects (L_1+L_2) and $\omega^2 LC_{Metal_T}>0$. The capacitance labeled as "_T" is the tested capacitance of whole DUT and metal connections. C_{ESD_C} is the corrected capacitance considering the inductance with enhanced method. The new model suggests that C_{ESD_T} is lower than C_{ESD_C} at high frequency due to L_1 and L_2, which may be attributed to the fact that, without the L, the extracted C_{ESD} decreases slightly because of a slower response due

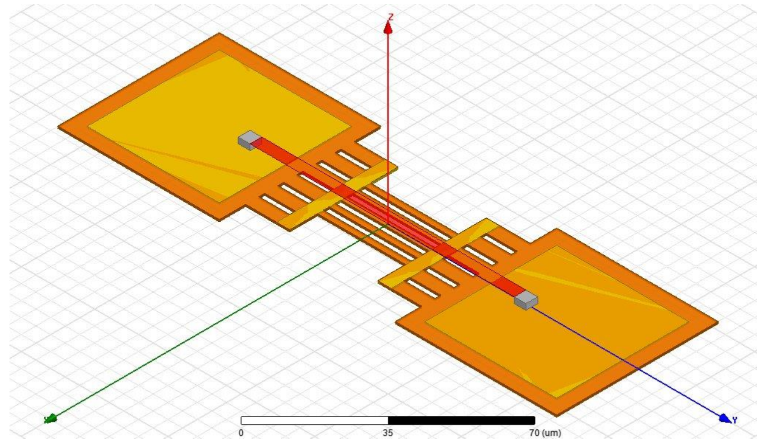


Figure 90 HFSS 3D model of the "short" metal wire.

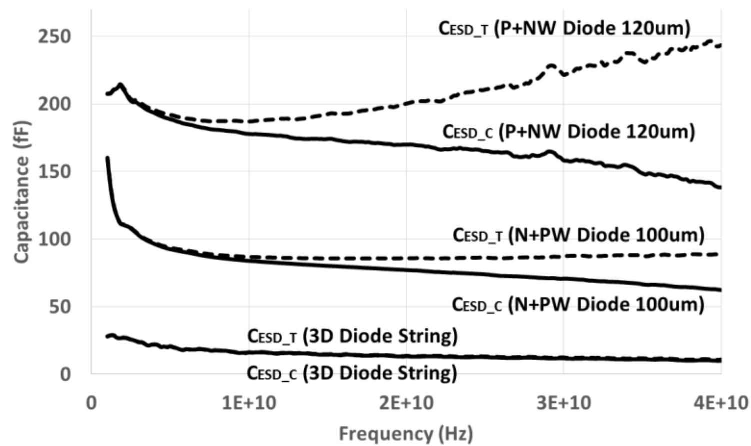


Figure 91 Comparison of C_{ESD_T} with C_{ESD_C} for ESD devices shows extraction errors for C_{ESD_T}.

to PN depletion capacitance [63]. In the new C_{ESD} extraction method, a "short" dummy is omitted to save layout area and the series L associated with the "short" metal can be readily obtained by HFSS field simulation, as depicted in Figure 90. Figure 91 compares the C_{ESD_T} with the C_{ESD_C} for a set of ESD devices with varying sizes (different C_{ESD}), which clearly shows that error margins exist for ESD devices with larger C_{ESD}.

6.2.2 Measurement with High Speed Circuits

a) Ring Oscillator

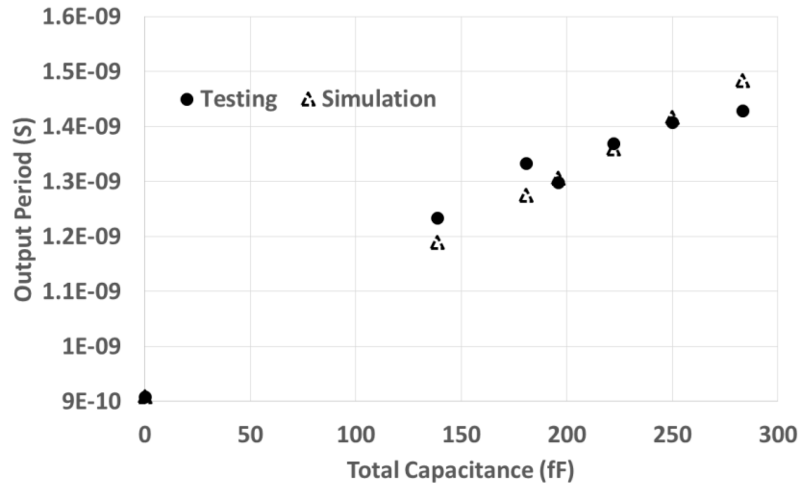


Figure 92 Comparison of tested and simulated output period for a 37-stage ring oscillator using different width ESD devices.

Figure 92 depicts the output period, hence, the frequency, of the ESD-protected 37-stage ring oscillator circuit obtained directly from measurement (solid) and from simulation (hollow) using the C_{ESD_C} values extracted using our new calibrated C_{ESD} extraction technique for the ESD structures used in this work. It readily shows that the new calibrated C_{ESD} extraction method is accurate in predicting circuit performance with ESD influences.

b) I/O Buffer

The dummy input and output buffer circuits are used to evaluate impacts of ESD on return loss performance. The ESD diode protection structures have varying sizes (W=60/100/120μm) for ESD protection target of 2/3/4KV. The C_{ESD} values for different ESD structures were extracted using both the simplified extraction method (C_{ESD_T}) and the new calibrated extraction technique (C_{ESD_C}). The industry standard return loss mask was used to evaluate the data rate performance if the I/O buffers where the return loss (S_{11}) of the I/O circuits was measured by use of the return loss mask matching, hence converted

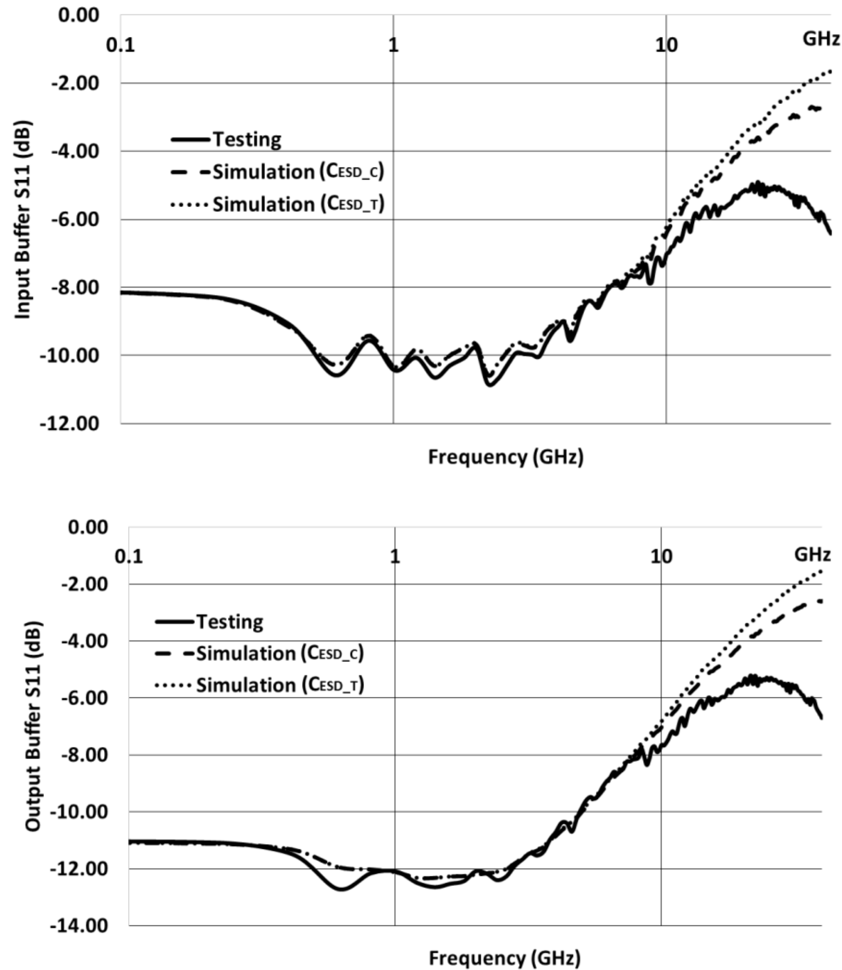


Figure 93 Comparison of S_{11} between measurement and simulation for the input buffer (Upper) and output buffer (Lower) with a 120µm gated diode ESD device.

to the maximum baud rates. Figure 93 compares the measured return loss (S_{11}) with simulation for sample input and output buffers using gated diode ESD devices for 4KV ESD protection, with the baud rate values summarized in Table 6-7. The simulation has two slits, each using the obtained C_{ESD} by the simplified extraction method (C_{ESD_T}) and the new calibrated extraction technique (C_{ESD_C}) for the same I/O circuits. The comparison clearly shows that the new calibrated C_{ESD} extraction technique is more accurate than the simplified C_{ESD} extraction method in the high frequency region for ESD structures with

I/O Buffers		Input			Output		
ESD Levels		2KV	3KV	4KV	2KV	3KV	4KV
SIM	C_{ESD_T}	9.8	7.6	5.9	12.5	10.9	7.7
	C_{ESD_C}	9.9	7.8	6.0	12.6	11.2	7.9
Testing		11.5	9.0	6.1	13.6	10.9	8.3
Error (%)		13.9	13.3	1.6	7.4	2.8	4.8

Table 6-7 A summary for return loss baud rate masks of I/O buffers.

relatively higher C_{ESD} (i.e., for higher ESD protection targets required for advanced ICs). It should be noted that, the simulated S_{11} may over-estimate the data rate degradation in the high frequency section because it only considers the capacitance effects.

As a conclusion, the enhanced method of ESD parasitic extraction, considering the metal-induced series inductance, shows a more accurate estimation of performance degradation of high speed ICs. And it reduces the testchip size, the cost on the other side, with the help of efficient HFSS EM simulation.

6.3 Mm-wave Switch and ESD Protection Parasitic Impacts

Millimeter-wave bands, especially the bands around 28/38GHz, are proposed for 5th generation of mobile networks to ensure ultra-high data rate applications mentioned in Chapter 2.3. Accordingly, the Federal Communications Commission (FFC) already approved these frequency bands for 5G research in United States. As I've introduced the complicity of front-end module for 5G RFICs, more and more RF antenna switched will be used in each mobile phone to achieve low power, low cost and high data rate. At that frequency, the parasitic effect of ESD protection, which is a must for all ICs especially the antenna side exposed to outside world, will be a disaster for RF performance.

To date, many efforts have been devoted to designing mm-wave RF switches in CMOS with various features, such as using floated-body [14], active balun [64] and traveling-

wave techniques [65]. However, the robust on-chip ESD protection is not considered in those designs. ESD induced parasitic effects are inevitable that can seriously effect switch performance due to impedance mismatching and additional path to ground, particularly for series-shunt switches operating at high frequencies.

Here, series-shunt SPDT RF switches of 28GHz and series-only SPDT switches at 38GHz with different kinds of ESD protection are designed to analysis the impacts of ESD parasitic effects for mm-wave frequency application.

6.3.1 Mm-wave Switch Design with ESD Protection

Figure 94 depicts the topology of the 28GHz series-shunt (a) and 38GHz series-only (b) switches designed in this work. MOSFET stacks and large gate resistors are used to achieve higher power handling capability up to 23dBm. Feed forward capacitor (FCC) is

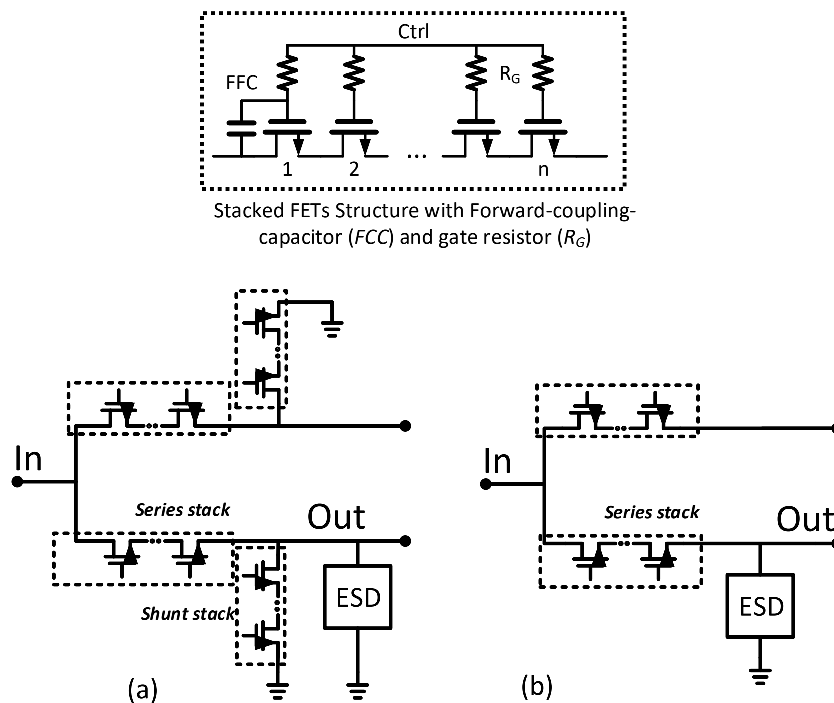


Figure 94 Series-shunt (a) and series-only (b) switch topologies feature stacked MOSFETs with output ESD protection.

Designs	Branch	FET Stack Number	Finger Width (μm)	Finger Number	Total Width (μm)	Gate Resistance (Ω)	FFC (fF)
28GHz SPDT	Series	10	2.508	140	351.12	1K	20
	Shunt	10	2.508	60	150.48	1.4K	20
38GHz SPDT	Series	10	2.508	140	351.12	800	20

Table 6-8 Switch design specifications of 28GHz series-shunt SPDT and 38GHz series-only SPDT.

adopted to balance voltage distribution across MOSFETs in the stacks [5]. These MOSFETs operate in ON and OFF states, and the switch impedance can be readily modeled. R_{lin} is the ON-impedance of a MOSFET in linear region and C_{ov} is the OFF-capacitance of the MOSFET across drain to source. Hence, the OFF/ON impedance (Z_{OFF} and Z_{ON}) ratio is modeled by $1/2\pi f_{signal} C_{ov} R_{lin}$. The Z_{OFF} to Z_{ON} ratio is directly related to the difference between the insertion loss (IL) and the isolation (Iso) of a switch, as given in Equation (1), which should be designed to be as large as possible. Unfortunately, at mm-wave frequency, this ratio drops rapidly, for example, down to 55:1 for a 38GHz switch in 45nm SOI CMOS. For a 50Ω-matching system, the difference between IL and Iso is given as

$$Iso - IL\ (dB) = 20 \log \frac{|Z_{OFF} + 50|}{|Z_{ON} + 50|}$$

Hence, an ON-resistance of 5Ω gives an Iso-to-IL difference of only 14dB, which decreases further if Z_{ON} is smaller. Therefore, for a preferred lower insertion loss, the switch isolation will become very poor, which limits the performance of a high frequency switch. Table 6-8 summaries the design split data for the series and shunt MOSFET stacks optimized for IL and isolation per simulation.

For the 28GHz series-shunt switch, the parasitic drain-to-source diodes in the shunt FETs provide ESD self-protection. The shunt branch was removed for the series-only

38GHz switch to achieve better insertion loss performance, which hence does not have any ESD self-protection. Dedicated ESD protection is designed for the two switches in this work: an ESD protection from I/O to ground is used for the series-only 38GHz switch, and additional ESD protection was added to the series-shunt 28GHz switch for higher ESD robustness on top of its self ESD protection shown in Figure 94 (a) and (b). Two types of dedicated ESD protection structures were used in this design: a 5-diode diode-string (5DS) and a silicon-controlled rectifier (SCR) introduced in Chapter 5, for switch output. The ESD triggering voltage (V_{t1}) was designed to be higher than the maximum signal level of 3.2V (23dBm). The 5DS ESD protection structure utilizes P-plus/N-body silicide blocked (SBLK) diode that has a perimeter of 320um for each diode in the string and its V_{t1} is about 5V. The SCR ESD protection device is a multi-finger SBLK structure featuring a total finger width of 200um and $V_{t1} \approx 9V$ based on breakdown voltage of reversed N-well/P-sub junction.

6.3.2 RF Performance: Insertion Loss and Isolation

RF characterization of insertion loss and isolation was conducted at die level using GSG microwave probes by Agilent E8363B vector network analyzer. Figure 95 and Figure 96 show the measured IL and Iso across a frequency range from 26GHz to 30GHz for 28GHz band and 36GHz to 40GHz for 38GHz band switches, respectively. The negative influences of ESD protection on IL and Iso can be readily observed for the switches. In general, the ESD-induced parasitic effect, including resistance and capacitance from output to ground, increases the insertion loss in ON state, but improves the isolation in OFF state.

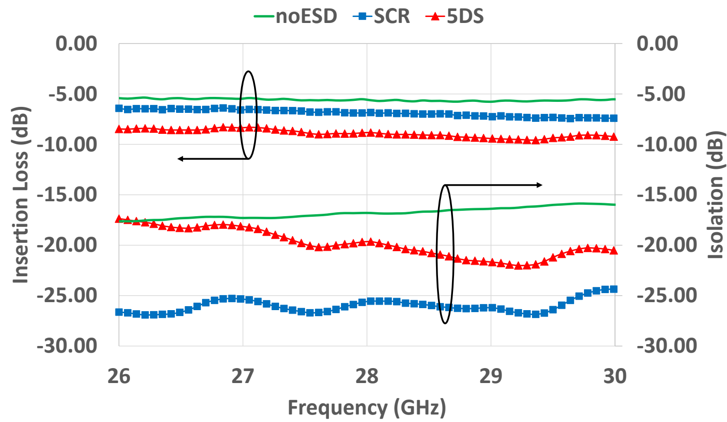


Figure 95 Measured insertion loss and isolation for 28GHz switches with different ESD protection: no dedicated ESD protection (noESD), and 5DS and SCR ESD protection.

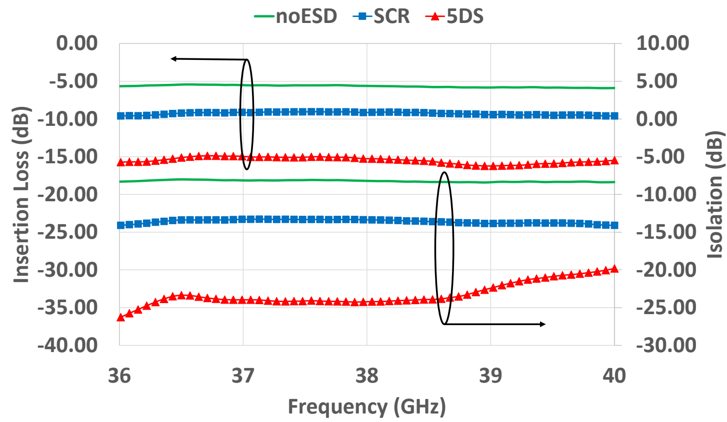


Figure 96 Measured insertion loss and isolation for 38GHz switches with different ESD protection: no dedicated ESD protection (noESD), and 5DS and SCR ESD protection.

In ON state, the diode-string ESD protection affects the switch insertion loss more than that of using SCR ESD protection, because the 5DS ESD structure has smaller off-resistance that introduces more leakage current. On the other hand, both 5DS and SCR ESD structures have parasitic capacitance, which affects IL more severely in higher frequency. Therefore, the measured IL degrades more significantly for the 38GHz switch than for the 28GHz switch due to ESD protection.

Band	P_{1dB}	Specifications	No ESD	5DS	SCR
28GHz	27dBm	Insertion Loss (dB)	5.62	8.83	6.84
		Isolation (dB)	16.8	19.6	25.5
		Iso – IL (dB)	11.18	10.77	18.66
		V_{t1} (V)	3.6	3.59	6.65
		I_{t2} (A)	0.27	2.76	1.63
38GHz	28dBm	Insertion Loss (dB)	5.63	15.26	9.12
		Isolation (dB)	8.22	24.2	13.4
		Iso – IL (dB)	2.59	8.94	4.28
		V_{t1} (V)	-	4.59	8.97
		I_{t2} (A)	~ 0	2.76	1.68

Table 6-9 Measured mm-wave switch RF & ESD performance with different ESD protection.

In OFF state, the extra ESD-induced parasitic shunting path reduces signal through-leakage from input to output, hence improves the isolation for switches with dedicated ESD protection, i.e., using 5DS and SCR ESD structures. From the isolation comparison, measurement shows better isolation of the 28GHz switch using SCR ESD protection, while isolation is better for the 38GHz switch using 5DS ESD protection. This may be attributed to the fact that the SCR ESD structure is insensitive to frequency. However, the 5DS ESD structure is more sensitive to frequency due to larger capacitance, therefore, showing better isolation in the 38GHz band, while much worse isolation than the case using SCR ESD protection in the 28GHz band where resistance dominates. Table 6-9 summarizes the measurement results for the switches using different ESD protection structures.

6.3.3 ESD Performance: TLP measurements

Chip level ESD protection behaviors were characterized using transmission line pulsing (TLP) method (Barth Model 4002 TLP Tester) for human-body-model (HBM) ESD protection. The TLP stressing waveforms used feature a rise time of 10ns and a duration of 100ns. Figure 97 shows the measured instantaneous ESD discharging I-V curves along with the leakage current for the 28GHz and 38GHz switches' self-protection.

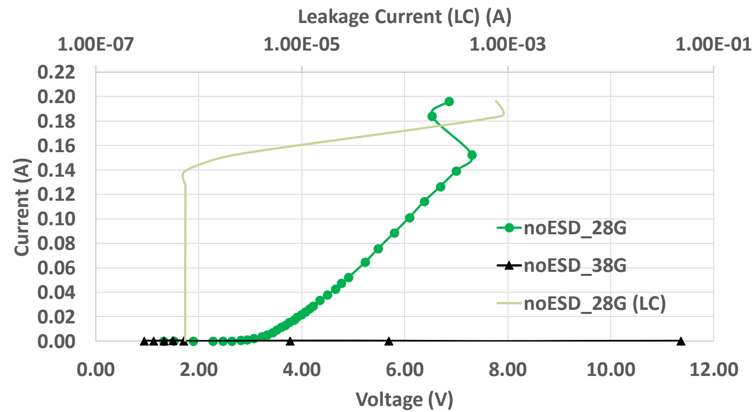


Figure 97 TLP tested I-V curve of 28/38GHz switch self-protection.

The secondary breakdown current (I_{t2}) represents HBM ESD failure threshold where leakage changes abruptly due to the physical failure of devices. It clearly shows that the 28GHz switch without using a dedicated ESD structure has ESD self-protection capability due to the stacked FET shunt branch for ESD discharging. However, the 38GHz switch without using a dedicated ESD structure does not have any ESD self-protection.

Additional ESD protection structures with 5DS and SCR are used to improve the ESD performance of switch design. The designed ESD protection level for the stand-alone ESD structures are I_{t2}~3A for the 5DS ESD device and I_{t2} ~1.8A for the SCR ESD device, respectively.

Figure 98 (a) and (b) show that much higher ESD protection was achieved for the 28GHz and 38GHz switches using dedicated 5DS and SCR ESD protection structures. The designed ESD triggering voltage for the stand-alone 5DS and SCR ESD protection structures are around V_{t1} ~5V and V_{t1} ~9V, respectively. Due to ESD self-protection in the 28GHz switch, the measured V_{t1} for the switch is reduced from its stand-alone device. This is not the case for the 38GHz switch featuring no ESD self-protection. However, I_{t2} does

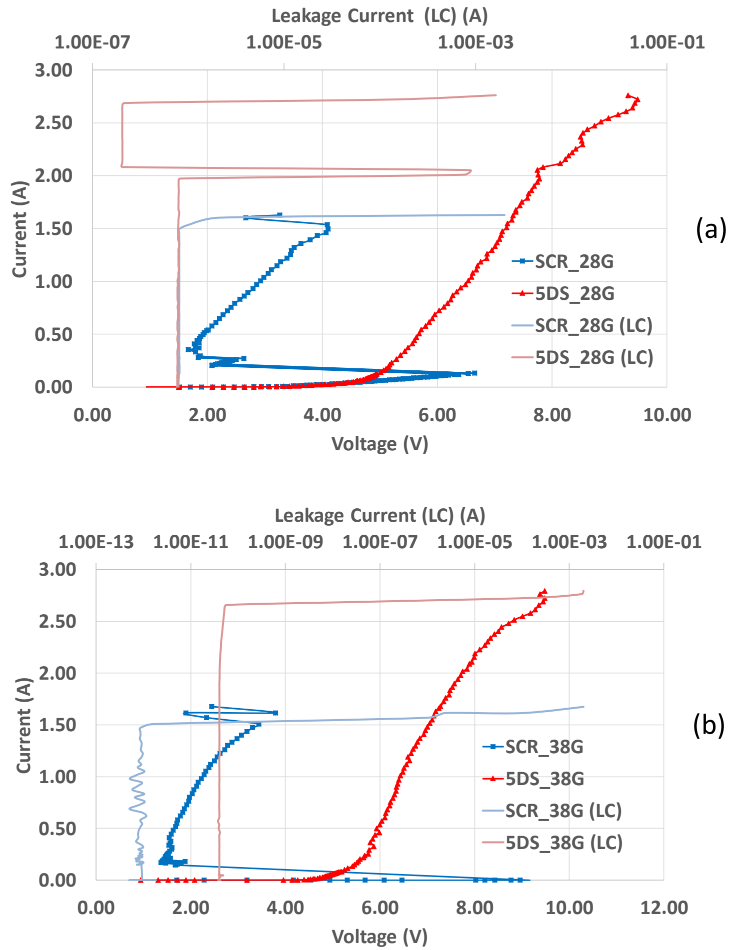


Figure 98 TLP tested I-V curve of 28GHz (a) and 38GHz (b) switch with different ESD protection.

not show a difference between 28GHz and 38GHz switches because most current go through the ESD structure after it turns on. The measured switch ESD protection results with RF performance are summarized in Table 6-9.

The impacts from ESD protection to 28GHz/38GHz mm-wave RF switches for 5G application are serious which will degrade insertion loss a lot. RF and TLP ESD measurements were conducted, which shows that ESD-induced parasitic effects have strong influences on the switches, which becomes stronger in higher frequency. Though diode string and SCR whose parasitic effects are attractive to high frequency application

can give great protection capability, the strong influences still unpredictable which requires more co-design techniques together during the switch design. The observation is critical for designing ESD-protected mm-wave switches for 5G ICs.

Chapter 7 Conclusion

This dissertation reports the novel radio frequency switch design with consideration of interference suppression and electrostatic discharge protection for 5th generation of mobile network. The advanced 45nm SOI CMOS process and 28nm bulk CMOS process are used to implement this critical RF front end block. The challenge and requirements from ultra-fast 5G application drive the RFFE research to reach a new height with higher power efficiency, higher frequency and of course better RF performance.

The interference from massive switches for massive MIMO application is unavoidable based on our research using a 4G/LTE switch array design. Results shows that even with layout optimization and high isolated process, the interference from other switch branches is still unacceptable for sensitive receiver side. As a result, a novel crosstalk suppression technique using through BEOL metal wall which compatible with modern fabrication process can remove more than 98% flying crosstalk. This concept is proven by both simulation and chip-level measurements with post-fabrication.

As a block closed to outside world, antenna switch suffers more threats from electrostatic discharge of both HBM and CDM. Comprehensive understanding of each ESD protection structure including the critical parameters and parasitic effects are indispensable during IC design. The ESD performance especially CDM protection capability, which is more challenge for modern thin oxide, are well studied with the new TLP/VFTLP and TCAD combined method. It gives more detailed ESD performance before and after chip tapeout which is very useful to evaluate protection ability. The parasitic effect, C_{ESD}, that severe degrades core IC performance especially for high speed, high

frequency ICs who is very sensitive to capacitance, is carefully extracted. The enhanced extraction methodology with the help of HFSS simulation improves the accuracy and saves more budget by reducing one third of the testchip area. This gives more room for IC designer to compensate the parasitic effect which will be very serious when reaching mm-wave 5G RFICs. The impacts analysis using 28/38GHz mm-wave RF switches for 5G application demonstrates the terrible performance degradation when adding essential ESD protection like diode string or SCR. The consideration for ESD-induced capacitance compensation is non-negligible during the mm-wave RF design which may be a significant research frequency level for 5G application.

As a future improvement, more front-end blocks including low noise amplifier, power amplifier, filter, etc., should be considered. The build-in through BEOL metal wall fabricated by foundry standard process need to be studied to make further proof of suppression capability. Other narrow band topologies like impedance changing method working as a switch for mm-wave band design can be considered which also give more room for parasitic compensation and ESD-RFIC co-design.

References

[1] T. Lee, *The design of CMOS radio-frequency integrated circuits*. 2003.

[2] W. M. C. Sansen, *Analog design essentials*. Springer, 2006.

[3] J. G. J. G. Andrews *et al.*, "What will 5G be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[4] NGMN Alliance, "NGMN 5G White Paper," 2015.

[5] M. Iwamura, "NGMN View on 5G Architecture," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, 2015, pp. 1–5.

[6] S. Akhtar, "2G-4G Networks: Evolution of Technologies, Standards, and Deployment," 2009.

[7] T. S. Rappaport *et al.*, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.

[8] S. Shinjo, K. Nakatani, K. Tsutsumi, and H. Nakamizo, "Integrating the Front End: A Highly Integrated RF Front End for High-SHF Wide-Band Massive MIMO in 5G," *IEEE Microw. Mag.*, vol. 18, no. 5, pp. 31–40, Jul. 2017.

[9] Y. Huo, X. Dong, and W. Xu, "5G Cellular User Equipment: From Theory to Practical Hardware Design," *IEEE Access*, vol. 5, pp. 13992–14010, 2017.

[10] U. L. Rohde, A. K. Poddar, I. Eisele, and E. Rubiola, "Next generation 5G radio communication NW," in *2017 Joint Conference of the European Frequency and Time Forum and IEEE International Frequency Control Symposium (EFTF/IFC)*, 2017, pp. 113–116.

[11] M. Agiwal, A. Roy, and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," *IEEE Commun. Surv. Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[12] Federal Communications Commission, "Use of Spectrum Bands Above 24 GHz for Mobile Radio Services (FCC-CIRC1711-02)," 2017.

[13] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges," *Wirel. Networks*, vol. 21, no. 8, pp. 2657–2676, Nov. 2015.

[14] Q. Li, Y. P. Zhang, K. S. Yeo, and W. M. Lim, "16.6- and 28-GHz Fully Integrated CMOS RF Switches With Improved Body Floating," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 2, pp. 339–345, 2008.

[15] M. Horstmann *et al.*, "Advanced SOI CMOS transistor technologies for high-performance microprocessor applications," in *2009 IEEE Custom Integrated Circuits Conference*, 2009, pp. 149–152.

[16] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wirel. Commun.*, vol. 21, no. 3, pp. 118–127, Jun. 2014.

[17] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO Architectures for Millimeter Wave Communications: Phase

Shifters or Switches?," *IEEE Access*, vol. 4, pp. 247–267, 2016.

[18] M. H. Abdul Hadi, B. H. Ahmad, P. W. Wong, and N. A. Shairi, "An overview of isolation improvement techniques in RF switch," *ARPN J. Eng. Appl. Sci.*, vol. 9, no. 3, pp. 342–348, 2014.

[19] C. Tinella, J. M. Fournier, D. Belot, and V. Knopik, "A high-performance CMOS-SOI antenna switch for the 2.5-5-GHz band," *IEEE J. Solid-State Circuits*, vol. 38, no. 7, pp. 1279–1283, 2003.

[20] X. J. Li and Y. P. Zhang, "Flipping the CMOS switch," *IEEE Microw. Mag.*, vol. 11, no. 1, pp. 86–96, 2010.

[21] M. Ahn, C. H. Lee, B. S. Kim, and J. Laskar, "A high-power CMOS switch using a novel adaptive voltage swing distribution method in multistack FETs," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 4, pp. 849–858, 2008.

[22] M. Parlak and J. F. Buckwalter, "A 2.5-dB Insertion Loss, DC-60 GHz CMOS SPDT Switch in 45-nm SOI," *2011 IEEE Compd. Semicond. Integr. Circuit Symp.*, no. c, pp. 1–4, 2011.

[23] M. B. Shifrin, P. J. Katzin, and Y. Ayasli, "Monolithic FET structures for high-power control component applications," *IEEE Trans. Microw. Theory Tech.*, vol. 37, no. 12, pp. 2134–2141, 1989.

[24] X. S. Wang and C. P. Yue, "A Dual-Band SP6T T/R Switch in SOI CMOS With 37-dBm P-0.1dB for GSM/W-CDMA Handsets," *IEEE Trans. Microw. Theory Tech.*, vol. 62, no. 4, pp. 861–870, 2014.

[25] A. S. Cardoso, P. Saha, P. S. Chakraborty, D. M. Fleischhauer, and J. D. Cressler, "Low-loss, wideband SPDT switches and switched-line phase shifter in 180-nm RF CMOS on SOI technology," *2014 IEEE Radio Wirel. Symp.*, pp. 199–201, 2014.

[26] S. Tanaka *et al.*, "A 3 V MMIC chip set for 1.9 GHz mobile communication systems," in *Proceedings ISSCC '95 - International Solid-State Circuits Conference*, 1995, p. 144–145,.

[27] X. S. Wang *et al.*, "Concurrent Design Analysis of High-Linearity SP10T Switch With 8.5 kV ESD Protection," *IEEE J. Solid-State Circuits*, vol. 49, no. 9, pp. 1927–1941, Sep. 2014.

[28] M. Ahn, H. W. Kim, C. H. Lee, and J. Laskar, "A 1.8-GHz 33-dBm P 0.1-dB CMOS T/R switch using stacked FETs with feed-forward capacitors in a floated well structure," *IEEE Trans. Microw. Theory Tech.*, vol. 57, no. 11, pp. 2661–2670, 2009.

[29] C. Wang *et al.*, "A Study of Interferences Inside an RF Switch Array in 45nm SOI CMOS," in *SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S), 2017 IEEE*, 2017, pp. 1–3.

[30] J.-P. Raskin, A. Viviani, D. Flandre, and J.-P. Colinge, "Substrate crosstalk reduction using SOI technology," *IEEE Trans. Electron Devices*, vol. 44, no. 12, pp. 2252–2261, 1997.

[31] Han-Su Kim, Kyuchul Chong, Ya-Hong Xie, and K. A. Jenkins, "The importance of distributed grounding in combination with porous Si trenches for the reduction of RF crosstalk through p/sup -/ Si substrate," *IEEE Electron Device Lett.*, vol. 24,

no. 10, pp. 640–642, Oct. 2003.
[32] J. S. Hamel, S. Stefanou, M. Bain, B. M. Armstrong, and H. S. Gamble, “Substrate Crosstalk Suppression Capability of Silicon-on-Insulator Substrates with Buried Ground Planes (GPSOI),” *IEEE Microw. Guid. Wave Lett.*, vol. 10, no. 4, pp. 134–135, 2000.
[33] T. Blalack, Y. Leclercq, and C. P. Yue, “On-chip RF isolation techniques,” in *Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 205–211.
[34] M. J. Loinaz, D. K. Su, and B. A. Wooley, “Experimental results and modeling techniques for switching noise in mixed-signal integrated circuits,” in *1992 Symposium on VLSI Circuits Digest of Technical Papers*, pp. 40–41.
[35] K. Joardar, “A simple approach to modeling cross-talk in integrated circuits,” *IEEE J. Solid-State Circuits*, vol. 29, no. 10, pp. 1212–1219, 1994.
[36] L. Connell *et al.*, “A CMOS broadband tuner IC,” in *2002 IEEE International Solid-State Circuits Conference. Digest of Technical Papers (Cat. No.02CH37315)*, vol. 1, pp. 400–476.
[37] C. P. Yue and S. S. Wong, “On-chip spiral inductors with patterned ground shields for Si-based RF ICs,” *IEEE J. Solid-State Circuits*, vol. 33, no. 5, pp. 743–752, May 1998.
[38] K. Ben Ali, C. Roda Neve, A. Gharsallah, and J.-P. Raskin, “Ultrawide Frequency Range Crosstalk Into Standard and Trap-Rich High Resistivity Silicon Substrates,” *IEEE Trans. Electron Devices*, vol. 58, no. 12, pp. 4258–4264, Dec. 2011.
[39] C. Roda Neve and J.-P. Raskin, “RF Harmonic Distortion of CPW Lines on HR-Si and Trap-Rich HR-Si Substrates,” *IEEE Trans. Electron Devices*, vol. 59, no. 4, pp. 924–932, Apr. 2012.
[40] K. Ben Ali, C. R. Neve, A. Gharsallah, and J.-P. Raskin, “RF Performance of SOI CMOS Technology on Commercial 200-mm Enhanced Signal Integrity High Resistivity SOI Substrate,” *IEEE Trans. Electron Devices*, vol. 61, no. 3, pp. 722–728, Mar. 2014.
[41] N. P. Pham, K. T. Ng, M. Bartek, P. M. Sarro, B. Rejaei, and J. N. Burghartz, “A micromachining post-process module for RF silicon technology,” in *International Electron Devices Meeting 2000. Technical Digest. IEDM (Cat. No.00CH37138)*, pp. 481–484.
[42] S. Stefanou *et al.*, “Ultralow Silicon Substrate Noise Crosstalk Using Metal Faraday Cages in an SOI Technology,” *IEEE Trans. Electron Devices*, vol. 51, no. 3, pp. 486–491, 2004.
[43] J. H. Wu, J. Scholvin, and J. A. delAlamo, “A Through-Wafer Interconnect in Silicon for RFICs,” *IEEE Trans. Electron Devices*, vol. 51, no. 11, pp. 1765–1771, Nov. 2004.
[44] J. H. Wu and J. A. del Alamo, “Design and modeling of Faraday cages for substrate noise isolation,” *Solid. State. Electron.*, vol. 85, pp. 6–11, Jul. 2013.
[45] A. Z. H. Wang *et al.*, “A Review on RF ESD Protection Design,” *IEEE Trans. Electron Devices*, vol. 52, no. 7, pp. 1304–1311, Jul. 2005.
[46] J. Liu *et al.*, “Design and Analysis of Low-Voltage Low-Parasitic ESD Protection

for RF ICs in CMOS," *IEEE J. Solid-State Circuits*, vol. 46, no. 5, pp. 1100–1110, May 2011.

[47] X. Wang *et al.*, "Post-Si Programmable ESD Protection Circuit Design: Mechanisms and Analysis," *IEEE J. Solid-State Circuits*, vol. 48, no. 5, pp. 1237–1249, May 2013.

[48] M.-D. Ker and C.-Y. Lin, "ESD protection consideration in nanoscale CMOS technology," in *2011 11th IEEE International Conference on Nanotechnology*, 2011, pp. 720–723.

[49] A. Wang, *On-chip ESD protection for integrated circuits: an IC design perspective*. Kluwer, 2006.

[50] M.-D. K. M.-D. Ker, J.-J. P. J.-J. Peng, and H.-C. J. H.-C. Jiang, "ESD test methods on integrated circuits: an overview," *ICECS 2001. 8th IEEE Int. Conf. Electron. Circuits Syst. (Cat. No.01EX483)*, vol. 2, pp. 1–4, 2001.

[51] Y. Zhou and J.-J. Hajjar, "CDM ESD failure modes and VFTLP testing for protection evaluation," in *2008 9th International Conference on Solid-State and Integrated-Circuit Technology*, 2008, pp. 333–336.

[52] F. Lu *et al.*, "A Systematic Study of ESD Protection Co-Design with High-Speed and High-Frequency ICs in 28 nm CMOS," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 63, no. 10, pp. 1746–1757, 2016.

[53] C. Wang *et al.*, "A comparison study of DTSCR by TCAD and VFTLP for CDM ESD protection," in *2017 IEEE 24th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 2017, pp. 1–4.

[54] R. J. Gauthier, "Investigation of voltage overshoots in diode triggered silicon controlled rectifiers (DTSCRs) under very fast transmission line pulsing (VFTLP)," *EOS/ESD Symp.*, pp. 1–10, 2009.

[55] G. Wybo, S. Verleye, B. Van Camp, and O. Marichal, "Characterizing the transient device behavior of SCRs by means of VFTLP waveform analysis," *Electr. Overstress/Electrostatic Disch. Symp. Proc.*, pp. 378–387, 2007.

[56] W. Y. Chen, E. Rosenbaum, and M. D. Ker, "Diode-triggered silicon-controlled rectifier with reduced voltage overshoot for CDM ESD protection," *IEEE Trans. Device Mater. Reliab.*, vol. 12, no. 1, pp. 10–14, 2012.

[57] F. Farbiz, A. Appaswamy, A. A. Salman, and G. Boselli, "Overshoot-induced failures in forward-biased diodes: A new challenge to high-speed ESD design," *IEEE Int. Reliab. Phys. Symp. Proc.*, pp. 1–8, 2013.

[58] T. Smedes and N. Guitard, "Harmful voltage overshoots due to turn-on behaviour of ESD protections during fast transients," in *2007 29th Electrical Overstress/Electrostatic Discharge Symposium (EOS/ESD)*, 2007, p. 6A.4-1-6A.4-9.

[59] I. Ladany, "An analysis of inertial inductance in a junction diode," *IRE Trans. Electron Devices*, vol. 7, no. 4, pp. 303–310, Oct. 1960.

[60] Ke Gong, Haigang Feng, Rouying Zhan, and A. Z. H. Wang, "A study of parasitic effects of ESD protection on RF ICs," *IEEE Trans. Microw. Theory Tech.*, vol. 50, no. 1, pp. 393–402, 2002.

[61] Chenkun Wang *et al.*, "A study of accurate extraction of ESD parasitic

capacitance," in *2016 13th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, 2016, pp. 437–440.

[62] C. Wang *et al.*, "A Study of Impacts of ESD Protection on 28 / 38GHz RF Switches in 45nm SOI CMOS for 5G Mobile Applications," in *Proc. IEEE Radio and Wireless Symposium*, 2018, pp. 157–160.

[63] M. L. Lucia, J. L. Hernandez-Rojas, C. Leon, and I. Mártil, "Capacitance measurements of p-n junctions: depletion layer and diffusion capacitance contributions," *Eur. J. Phys.*, vol. 14, no. 2, pp. 86–89, Mar. 1993.

[64] C. Huynh and C. Nguyen, "New ultra-high-isolation RF switch architecture and its use for a 10-38-GHz 0.18-??m BiCMOS ultra-wideband switch," *IEEE Trans. Microw. Theory Tech.*, vol. 59, no. 2, pp. 345–353, 2011.

[65] M. Yeh *et al.*, "A Millimeter-Wave Wideband SPDT Switch with Traveling-Wave Concept Using 0.13-Microm CMOS Process," *2005 IEEE MTT-S Int. Microw. Symp.*, vol. 1, no. C, pp. 53–56, 2005.