High-Performance and Energy Efficient Multi-Band I/O Interface for 3D Stacked Memory

Ahmed Alzahmi
Southern Methodist University, aalzahmi@smu.edu

Follow this and additional works at: https://scholar.smu.edu/engineering_electrical_etds

Recommended Citation
https://scholar.smu.edu/engineering_electrical_etds/14

This Dissertation is brought to you for free and open access by the Electrical Engineering at SMU Scholar. It has been accepted for inclusion in Electrical Engineering Theses and Dissertations by an authorized administrator of SMU Scholar. For more information, please visit http://digitalrepository.smu.edu.
HIGH-PERFORMANCE AND ENERGY EFFICIENT MULTI-BAND I/O INTERFACE
FOR 3D STACKED MEMORY

Approved by:

_______________________________________
Prof. Duncan L. MacFarlane
Professor of Electrical Engineering, Advisor

_______________________________________
Prof. Ronald A. Rohrer
Professor of Electrical Engineering

_______________________________________
Prof. Jennifer Dworak
Associate Professor of Computer Science and Engineering

_______________________________________
Prof. Joseph Camp
Associate Professor of Electrical Engineering

_______________________________________
Prof. Mohammed Khodayar
Assistant Professor of Electrical Engineering
HIGH-PERFORMANCE AND ENERGY EFFICIENT MULTI-BAND I/O INTERFACE
FOR 3D STACKED MEMORY

A Dissertation Presented to the Graduate Faculty of
Lyle School of Engineering
Southern Methodist University
in
Partial Fulfillment of the Requirements
for the degree of
Doctor of Philosophy
with a
Major in Electrical Engineering
by
Ahmed Alzahmi

August 7, 2018
ACKNOWLEDGMENTS

I would like to begin by thanking my former advisor, Dr. Byun, Gyung-Su, and my current advisor, Dr. MacFarlane, Duncan L. for his constant guidance and support. I would also like to thank my committee for the support and valuable comments during my Ph.D. proposal and the entire program. I would like to thank my parents and siblings who supported me all throughout my life. Finally, I give appreciations to my wife Ikram who supported and comforted me through all the degree, Thank you!


High-Performance and Energy Efficient Multi-Band I/O Interface for 3D Stacked Memory

Advisor: Professor Duncan L. MacFarlane

Doctor of Philosophy conferred August 7, 2018

Dissertation completed June 1, 2018

This dissertation describes the development of an energy efficient 3D multi-band I/O interface to meet the demand for high computation and improve battery life for future mobile memory interface exploiting 3D integration. The multi-band I/O (MBI) interface utilizes 3D integration, amplitude shift keying modulation/demodulation scheme for two radio frequency (RF) band transceivers, and CMOS driver with resistive feedback for the baseband (BB) transceiver. It enables transceiving three bands simultaneously that results in significant enhancements in energy efficiency and aggregate bandwidth. The 3D MBI system was implemented in 130nm CMOS process technology. It obtains a high aggregate data rate of 14.4Gb/s and an energy efficiency of 2.6pJ/b compared with prior works.
# TABLE OF CONTENTS

1. LIST OF FIGURES .................................................................................................................. x

2. LIST OF TABLES .................................................................................................................. xiv

1. CHAPTER 1: Introduction .................................................................................................. 1
   1.1. Introduction .................................................................................................................... 1
   1.2. Research Objectives ....................................................................................................... 4
   1.3. Dissertation Organization ............................................................................................... 6

2. CHAPTER 2: 3D Power Delivery Network ....................................................................... 8
   2.1. Introduction .................................................................................................................... 8
   2.2. Proposed Performance Aware 3D PDN Utilizes Common Centroid Regulator Placement ........................................................................................................................................... 11
   2.3. On-chip 3D PDN models ............................................................................................... 15
   2.4. The evolutionary process of 3D PDN .......................................................................... 18
   2.5. Experimental Results .................................................................................................... 20
   2.6. Conclusion ................................................................................................................... 27

3. CHAPTER 3: 3D Baseband ................................................................................................ 29
   3.1. Introduction ................................................................................................................... 29
   3.2. Proposed 3D BB for mobile memory interface ................................................................ 30
   3.3. Experimental Results ................................................................................................... 37
   3.4. Conclusion ................................................................................................................... 39
4.    CHAPTER 4: 3D RF bands ............................................................................................................ 41

4.1.  Introduction .......................................................................................................................... 41

4.2.  Proposed 3D RF Memory interface .................................................................................... 42

4.3.  Simulation Results ............................................................................................................... 60

4.4.  Conclusion .......................................................................................................................... 61

5.    CHAPTER 5: 3D MBI System ................................................................................................ 62

5.1.  Introduction .......................................................................................................................... 62

5.2.  Proposed 3D MBI utilizing TSV and LDO ......................................................................... 64

5.3.  Low noise amplifier design for 30 GHz ............................................................................. 67

5.4.  Band Selective Transformer ............................................................................................... 71

5.5.  Experimental Results .......................................................................................................... 72

5.6.  Conclusion .......................................................................................................................... 74

5.7.  Future Works ...................................................................................................................... 75

6.    CHAPTER 6: Ultra low power 2.4 GHz RF transmitter ....................................................... 77

6.1.  Introduction .......................................................................................................................... 77

6.2.  Proposed 2.4 GHz RF Transmitter ..................................................................................... 79

6.3.  Simulated Results ................................................................................................................. 82

7.    BIBLIOGRAPHY .................................................................................................................... 85
1. LIST OF FIGURES

Figure 1-1 Memory interface for smart mobile phone devices......................................................... 1
Figure 1-2 The projected number of cores, performance, and I/O data rate per pin for future application [95]......................................................................................................................... 2
Figure 1-3 Memory interface structure trends .................................................................................. 3
Figure 2-1 Diagram of a conventional (a) 2D and (b) 3D PDN ...................................................... 8
Figure 2-2 The proposed performance-aware 3-D PDN using a common centroid LDO placement (a) 3-D integration (b) TSV Array (c) TX (d) RX (e) LDO ..................................................... 11
Figure 2-3 The 3D PDN based on 2D mesh power routing and 3D supply TSVs (a) 3D TSV model (b) TSV parameters (c) 3D TSV HFSS Model (d) Wire Model (e) PDN Mesh layout .... 15
Figure 2-4 Time corresponding specs of the TSV fitness function, construction of the GA chromosome string for (b) TSV and (c) proposed 3D PDN sub-block optimization ............... 19
Figure 2-5 Simulated OTA in LDO (a) before and (b) after the evolution...................................... 22
Figure 2-6 Evolution of LDO regulated output signal in 1000 iterations ....................................... 23
Figure 2-7 Evolution of 3D TSV output signal compared to the corresponding 2D signal in 1000 iterations .......................................................................................................................... 23
Figure 2-8 The Experimental setup for C1, C2, C3, C4, C5, C6 for on-chip 3D PDN ............... 24
Figure 2-9 Transient analysis of PDN DC output voltage Vs. input supply noise voltage (Vin tx) of each case ........................................................................................................................................ 26
Figure 2-10 (a) IC prototyping for the proposed 3D-PDN (C6) (b) Microphotograph of each tier die, and (c) measured eye diagram .......................................................................................... 27
Figure 3-1 Diagram of (a) a conventional 2D and (b) 3D memory interface ........................................... 29
Figure 3-2 The proposed 3-D baseband architecture ................................................................. 30
Figure 3-3 Proposed schematic of 3-D baseband transceiver (a) TX (b) RX (c) input buffer (d) output driver (e) output buffer (f) LDO ............................................................... 31
Figure 3-4 5.8 Gb/s baseband signal after the input buffer stage. ................................................. 32
Figure 3-5 Simulation results for the intrinsic bandwidth of output driver with a different value of \( \frac{W_p}{W_n} \) ratio and supply voltage. ....................................................................................... 33
Figure 3-6 Simulation results for power consumption and rise time at different feedback resistance for 5.8 Gb/s data rate. ......................................................................................... 34
Figure 3-7 Simulated baseband signal after (a) TSV channel (b) inverter with a resistor feedback stage. ...................................................................................................................... 35
Figure 3-8 Simulated output signal at receiver end of baseband transceiver. ......................... 36
Figure 3-9 Simulation results of 3D TSV channel (a) signal loss (b) HFSS Model ......................... 36
Figure 3-10 Layout for BB transmitter (a) and receiver (b) ......................................................... 37
Figure 3-11 Simulated waveforms (a) input baseband signal of 5.8 Gb/s (b) output data rate (c) eye diagram of output signal .......................................................................................................... 38
Figure 3-12 Simulated power consumption for 10 Gb/s at different supply voltages in 130nm CMOS technology. .................................................................................................. 39
Figure 4-1 Diagram of a conventional (a) 2D and (b) 3D I/O transceiver. ......................... 41
Figure 4-2 3D RF I/O transmitter .................................................................................................. 43
Figure 4-3 Crossed coupled oscillator configuration ................................................................. 44
Figure 4-4 Spiral inductor (a) Layout (b) simulated series resistance \( R_s \) (i.e., real(Z(1,1)), inductance value (L), and quality factor (Q) of 10 GHz band ........................................... 45
Figure 4-5 Spiral inductor (a) Layout (b) simulated series resistance \( R_s \) (i.e., real(Z(1,1)),
inductane value (L), and quality factor (Q) of 30 GHz band

Figure 4-6 Simulated VCO output voltage and phase noise @10GHz (a) Voltage Swing (b) Phase noise

Figure 4-7 Simulated VCO output voltage and phase noise @ 30GHz (a) Output voltage (b) Phase noise

Figure 4-8 ASK modulator used in both RF bands (10/30 GHz) transmitter

Figure 4-9 Simulated waveforms of (a) 4.2Gb/s input data rate and (b) ASK output for 10GHz RFTX & (c) 4.2Gb/s input data rate and (d) ASK output signal for 30 GHz RFTX

Figure 4-10 Layouts of RF transmitter excluding spiral inductors (b) RFTX for 10 GHz (b) RFTX for 30 GHz

Figure 4-11 3D I/O RF receiver architecture

Figure 4-12 Self Oscillating-differential mixer

Figure 4-13 Simulated results from self-oscillating mixer (a) 10 GHz (b) 30 GHz

Figure 4-14 Differential amplifier schematic in both RF-band (10/30 GHz) receiver

Figure 4-15 Simulated output signal of the differential amplifiers for RF bands (a) 10 GHz (b) 30 GHz

Figure 4-16 Circuit configuration for (a) converter buffer and (b) output driver

Figure 4-17 Simulated output signal for both RF bands receiver (a) 10 GHz (b) 30 GHz

Figure 4-18 Layout for both RF bands receiver (a) 10 GHz (b) 30 GHz

Figure 4-19 Simulated schematic of RF I/O interface for both RF bands (10GHz & 30GHz)

Figure 4-20 Simulated S-parameters (a) S11, (b) S21, and (c) S22 for 10 GHz RF band

Figure 4-21 Simulated S-parameters (a) S11, (b) S21, and (c) S22 for 30 GHz RF band

Figure 4-22 Simulated input & output waveforms at 4.2Gb/s for (a) 10 & (b) 30 GHz I/O interface
2. LIST OF TABLES

Table 2.1 The Evolutionary Specifications of The Proposed Algorithm ............................................. 21
Table 2.2 Recommended TSV Design Parameters for 3D PDN ......................................................... 24
Table 2.3 Simulated Results of Different Cases of 3D PDN Architecture ........................................ 26
Table 3.1 Simulation results of normalized power and rise time at different feedback resistance. ........................................................................................................................................... 34
Table 5.1 Performance comparison of proposed 3D-MBI with prior works......................................... 75
Table 6.1 Performance comparison of proposed 2.4 GHz RF transmitter with Prior Works.............. 83
This is dedicated to my parents, siblings, friends, and lovely wife.
1. CHAPTER 1: Introduction

1.1. Introduction

Advances in mobile devices such as smartphones demand to elaborate and improve computation capabilities and video processing to achieve greater aggregate bandwidth and higher energy efficiency [1]. For instance, Samsung Galaxy Note 7 and OnePlus 3 are equipped with 4 GB and 6 GB DRAM, respectively [41].

Since in future very large system integration (VLSI), the processor shifts to multi-core architecture to increase the performance of a control processing unit (CPU) [42]. However, the mobile device systems consist also of Dynamic Random-Access Memory (DRAM) and I/O interface (channel) which is the link between CPU and DRAM as shown in Figure 1.1. Thus, we need to improve DRAM and I/O interface regarding the number of I/O pins and data rate.

![Diagram of Memory Interface for Smart Mobile Phone Devices](image)

Figure 1-1 Memory interface for smart mobile phone devices

According to International Roadmap Committee (IRC), as shown in Figure 1.2, the number of cores and each core’s frequency are increased by 1.4x and 1.05x per year, respectively. The performance is also projected to increase 1000x in 2024 compared to 2009 [5].
Nevertheless, the I/O pin data rate is only improved by 10x in 2024 compared to 2009 due to the channel loss and crosstalk, and the number of I/O pins increase 2x or 4x at the same period [5], [43]. Thus, the I/O pin bandwidth, i.e., data rate per pin, between cores and memories demands to both scale at the same rate to achieve the best performance possible. Moreover, energy per bit of I/O transceiver has much slower rate compared with the projected I/O pin bandwidth for future mobile memory advances [5]. Thus, significant improvement in energy efficiency and data rate for future advances in mobile memory is required.

Figure 1-2 The projected number of cores, performance, and I/O data rate per pin for future application [95]

The existing and future trends of memory interface architecture are shown in Figure 1.3. To increase the bandwidth of memory especially in high-performance computing and video/graphics applications, the IO pin data rate should increase [49]. However, it is very challenging to increase the data rate beyond 10 Gb/s/pin. For instance, low power double data rate (LPDDR)3 and LPDDR4 accomplished 1.6 Gb/s and 4.2 Gb/s, respectively. The LPDDR3 and LPDDR4 utilize conventional two-dimensional (2D) point to point (P2P) memory interface
as shown in Figure 1.3(a) [44]. Second, 2D multi-drop (MD) memory interface which implements multiple DRAMs could be installed for higher bandwidth in emerging application such as cloud computing and data mining. However, the MD memory interface increases power consumption and leads to routing congestion on PCBs as shown in Figure 1.3(b). 2D MD channels also suffer from high capacitive load. For example, the transmitter (TX) in multi-drop structure is large regarding transistor sizing compared to P2P to support driving considerable load. Thus, the MD I/O interface consumes significant power consumption compared to P2P because of the huge distance that the data travel between CPU and memory which deteriorate the bandwidth. Recent P2P and MD I/O interfaces achieve 5.5 Gb/s/pin at 7.8 mW and 5.8 Gb/s/pin at 14.198 mW, respectively [45], [46]. To obtain high energy efficiency, the I/O transceivers need to operate at
high frequency and low power [2]. However, inter-symbol interference (ISI) and high-frequency distortion decrease the maximum operating frequency of 2D long transmission lines (T-Line) (e.g., 5-30cm) used in P2P and MD I/O [11].

Increasing the number of IO pins could be a solution to increase the performance by utilizing three-dimensional (3D) integration such as wide IO or high bandwidth memory (HBM) as shown in Figure 1.3(c) and (d), respectively. For example, the number of IO pins are 32, 512, and 1024 for conventional DRAM, wide IO, and HBM, respectively [47], [48]. 3D integration enables multiple stacking of chips vertically using very short channel (e.g., 50 – 200 μm), through-silicon-via (TSV) [12], [13], [14]. 3D integrated circuit (IC) design also minimizes signal path lengths, significantly improves bandwidth/power efficiency, and decreases latency, noise, and chip area. Thus, the advances in this thesis are designed by utilizing 3D integration to maximize the bandwidth and energy efficiency of I/O interface to meet the demand for high performance and extend battery life.

1.2. Research Objectives

The goal of this research is developing an energy efficient 3D multi-band I/O interface to support the demand for high computation and improve battery life for future mobile memory interface exploiting 3D integration. The multi-band I/O interface uses 3D integration, amplitude shift keying modulation/demodulation scheme for radio frequency (RF) band transceiver, and CMOS driver with resistive feedback for baseband (BB) transceiver. Thus, the research has been developed in multiple tasks to evaluate the progress and ensure having the intended results for each task. In extra work, an ultra-low power 2.4 GHz RF transmitter for biomedical implants has been developed.
1.2.1. Task 1: Develop a reliable 3D I/O architecture utilizing a power delivery network and TSVs

3D power delivery network (PDN) using a novel common-centroid regulator placement is presented to build a reliable 3D I/O architecture. To analyze the main benefits of the proposed on-chip 3D PDN, the regulator, most crucial block, is presented in the PDN. Through-silicon via and I/O links are also integrated into the 3D PDN to analyze the power noise accurately. The proposed 3-D PDN can significantly improve the energy efficiency, supply noises, and power/signal quality of 3D I/O.

1.2.2. Task 2: Design an energy efficient 3D Baseband Transceiver for I/O interface utilizing CMOS driver with resistive feedback Architecture

Energy and performance-aware 3D baseband transceiver utilizing CMOS driver with resistive feedback architecture is presented. 3D technology utilizing through-silicon via (TSV) is exploited in the proposed design to attain high performance and low power.

1.2.3. Task 3: Develop high-performance and low-power 3D RF Transceiver for I/O interface using amplitude shift keying modulation/demodulation scheme

A high-performance 3D RF transceiver with improved through-silicon via (TSV) geometry and impedance matching for future 3D stacked memory has been introduced. RF band transmitter (TX) comprise LC-voltage control oscillator (VCO) and ASK modulation. RF receiver (RX) comprises self-oscillating mixer, a differential amplifier, and class-AB amplifier.

1.2.4. Task 4: Implement Energy Efficient 3D Multi-Band I/O Interface for Heterogenous Stacked Memory

Combines high-speed and low-power baseband, task 2, and two RF bands, task3, transceivers utilizing 3D integration and LDO, task 1, to maximize data rate and minimize
power consumption. In this design, the three transmitted signals (BB + 2 RFs) have been transmitted and recovered concurrently utilizing selective band transformer. The transmitted signal strength (voltage swing) of 30 GHz suffers from high degradation due to parasitics; thus, low noise amplifier (LNA) is implemented in the receiver side to compensate. To minimize supply noise that has a significant impact on power/signal integrity, the LDO is exploited along with 3D TSV technology.

1.2.5. Extra work: Ultra-low-power 2.4 GHz RF transmitter for medical implants

A high performance and low power 2.4 GHz radio frequency (RF) transmitter for biomedical application is presented. The design utilizes LC voltage control oscillator (VCO), ASK modulator, and a Class-D inverse power amplifier (PA). The ASK-modulates the carrier signal generated by a low phase noise LC-VCO, then the modulated signal is amplified by Class-D inverse PA.

1.3. Dissertation Organization

An energy-efficient 3D I/O transceiver architectures and circuits are developed to support the high demand for intensive computing and video processing for mobile device memory. The thesis is organized as follows. Chapter 2 describes a reliable method of utilizing 3D integration, power delivery network, and I/O interface to support the high-speed data rate and save battery life by utilizing power delivery network to ensure signal/power integrity (i.e., task 1). Design an energy efficient 3D baseband transceiver for I/O interface using CMOS driver with resistive feedback architecture (i.e., task 2) is presented in chapter 3. A high-performance and low-power 3D RF Transceiver for I/O interface employing amplitude shift keying modulation/demodulation scheme (i.e., task 3) is discussed in Chapter 4. A multi-band I/O interface using 2 RF bands, 10
and 30 GHz, and baseband is designed in chapter 5. It enables transmitting/receiving the three signals (BB + 2RFs) by implementing band selective transformer, TSV, and low noise amplifier for the 30 GHz RF band. The conclusion and future work will also be discussed in chapter 5. In chapter 6, the extra work, an ultra-low power 2.4 GHz RF transmitter for medical implants is described.
2. CHAPTER 2: 3D Power Delivery Network

2.1. Introduction

The content of this chapter has been published [65]. In a mixed-signal system on chip (SoC) design, providing a robust power delivery network (PDN) is a critical factor to enhance the overall performance [50]. A well-designed PDN is essential to reduce supply noises, tolerate large variations of load current, and accomplish better power/signal integrity of heterogeneous 3D stacked devices to support failure-resistant chip operations.

The power delivery of a conventional 2D SoC design, as shown in Figure 2.1(a), is a critical challenge in advanced technologies because of higher operating frequency/power density and lower supply voltage. As the operation frequency of heterogeneous devices increases with lower supply voltage, the design of optimized PDN structure becomes more critical. In this chapter, an evolutionary process is utilized to find the best design of a PDN structure.

Figure 2-1 Diagram of a conventional (a) 2D and (b) 3D PDN
Latency and integrity can inherently be improved due to vertically short 3D TSVs (or 2.5D µbumps) channel. However, as the number of 3D TSV I/O channel increases, e.g., 1024 TSV I/Os in an HBM [51], the TSV 3D I/O transceiver should be co-integrated into the PDN to evaluate the PDN performance accurately. Figure 2.1(b) shows an on-chip 3D-PDN which is composed of a phase-locked loop (PLL), TSV I/O array, and clock/signal distribution networks. In this complex configuration of 3D PDN, the critical design parameters such as TSV resistance, capacitance, and inductance can be the most critical design variables for optimization. For example, TSV connection on 3D PDN has small inductance value comparing to traditional bond wire connection. Thus, when the frequency of I/O interface increases reaching gigahertz frequencies such as 2 Gb/s in high bandwidth memory (HBM) [52], the fraction of inductance value (e.g., 200pH) causes disruptive frequency-dependent impedance and entire system failure. In future, the number of 3D I/O array can be increased significantly. Therefore, prudent PDN topology planning can have significant effects on the heterogeneous 3D IC device operation. For instance, severe crosstalk and supply voltage fluctuations occur due to poor PDN topology and dense 3D TSV array.

Integrating PDN and low-dropout voltage regulator (LDO) provide significant improvement to eliminate power supply noises and complete on-chip 3D PDN. Integrating PDN and low-dropout voltage regulator (LDO) minimize power supply noises and complete on-chip 3D PDN. Also, implementing multiple LDOs close to noise sources on a die in a distributive manner reduces power supply noise [53], [54],[55]. However, the complex interaction between the LDOs, TSVs, and I/O interface causes unstable power delivery network results in severe degradation of circuit performance or design failure.
In this chapter, the PDN stability and chip functionality are accomplished by proper LDO placement among multiple tiers. Rigorous analysis of different LDO’s placement and active/passive device sizing is implemented to find the optimum 3D PDN performance. Recently, many works have been done to explore 3D PDN design. Work in [56] shows the impact of TSV sizing and allocation on 3D PDN, and [57] analyzes different TSV densities and aspect ratios. A novel topology for 3D PDN is taken into consideration for TSV placement and package parasitics in [58]. Floor-planning and optimization of PDN are discussed in [59]. Electrothermal modeling and 3D PDN optimization under electrothermal constraints are presented in [60], [73]. Works in [61],[62], [70] discuss power supply noises and decoupling insertion, while [63] investigates placement and planning of power and ground TSVs in 3D PDN. Work in [64] designed and analyzed PDN impedance including TSV effect for HBM interposer. An HBM and microbump (µbump) test methods are explained in [48] for improved testability of the chip. To the best of our knowledge, this is the first work that ensures 3D PDN stability taking into consideration the complex interaction between LDOs, TSV, and I/O interface. It also finds the optimum active/passive devices sizing and placement for LDO and TSV to achieve the best performance for 3D PDN and I/O interface.

Since the academic 3D IC implementation using TSVs or 2.5D through-silicon interposers (TSIs) is challenging, a possible packaging technique such as 3D µbumps [69] with similar dimensions has been utilized to optimize the TSV geometry for the proposed 3D PDN. The proposed 3D PDN architecture could be applied to any 2.5D silicon interposer and future 3D TSV-based ICs.

The rest of the chapter is organized as follows. Section 2.2 outlines the proposed power delivery circuit in 3D IC design. Section 2.3 investigates TSV and wire models. The
evolutionary process is presented in Section 2.4. Section 2.5 presents the optimization method for TSV and sub-blocks. It also discusses the impact of TSV/LDO placement on the 3D PDN performance to find the optimum architecture. Finally, Section 2.6 concludes this work.

2.2. Proposed Performance Aware 3D PDN Utilizes Common Centroid Regulator Placement

To reduce severe crosstalk and supply voltage fluctuations from dense 3D TSV, I/O array, and 2D complex power routing, performance-aware 3D PDN utilizing common centroid regulator placement has been proposed. It comprises an LDO placed in the middle tier (Tier 2), TSVs, and straightforward transmitter and receiver to form I/O interface. Figure 2(a) shows the 3D integration for the proposed architecture. A multi-drop link is formed by CPU connected to multi-DRAM chips through TSV channels exploiting 3D integration technology. Since the distance between CPU and each DRAM is very short, the signal integrity greatly improved compared to traditional 2D DRAM I/O interface [66]. TSV channel enables the output data to be

![Diagram of 3D Integration](image)

Figure 2-2 The proposed performance-aware 3-D PDN using a common centroid LDO placement (a) 3-D integration (b) TSV Array (c) TX (d) RX (e)LDO
transmitted vertically from CPU to DRAM where the data is going to be recovered. Every TSV signal is bounded by TSV grounds to shield the noise-sensitive devices or interconnect which improve 3D signal integrity. 3D TSV pitch of adjacent TSVs can be increased two or three times wider than 3D TSV diameter to minimize crosstalk and timing distortion. Consequently, the proposed architecture uses five TSV channels as shown in Figure 2.2(b); the labeled TSVs represent ground, data, and power channels.

2.2.1. 3D I/O Interface

Figure 2.2(c) shows the proposed 3D transmitter (TX) consists of an input buffer and an output driver. On the transmitter side, the generated data ($D_{input}$) is fed into the input buffer followed by an output driver. The output driver utilizes CMOS logic to drive the TSV channel and enable data transmission. To support high-frequency operation and minimize delay, voltage offset, mismatch, and noise effect, an input buffer using a self-biased differential amplifier is exploited [18].

Regarding the receiver side (RX) as shown in Figure 2.2(d), it consists of a digitally on-die termination (ODT) [22], pre-driver, and output driver. Because TX and RX are communicating by the common mode voltage, ODT is utilized to amplify incoming data from TSV channel. It also controls the common mode voltage and improves signal integrity by eliminating impedance mismatch. The pre-driver utilizes inverters to enhance the recovered data. The output driver uses a push-pull configuration with two resistors in series with transistors to avoid impedance mismatch and reduce sensitivity to process, voltage, and temperature variations (PVT) [23].

2.2.2. Low Dropout Voltage Regulator for 3D PDN

LDO is exploited in many high-performance power supply circuits to provide stable and
reliable voltage [67]. As shown in Figure 2.2(e), the utilized LDO in the proposed architecture consists of an error amplifier, a voltage reference, a pass transistor, and a feedback network which forms the regulator loop. The error amplifier uses a symmetrical operational transconductance amplifier (OTA) in the regulator design. A common source PMOS device is chosen as a pass element. It exhibits the best overall design yielding a good compromise of drop-out voltage, quiescent current flow, output current, and speed [68]. The pass device must be physically large to obtain low drop-out voltage and high output current characteristics. Depending on the reference voltage \( (V_{REF}) \) and feedback resistors \( (R_1 \) and \( R_2) \), the output voltage \( (V_{out}) \) can be calculated as in (1). When the LDO supplied by suitable input voltage \( (V_{IN}) \), the \( V_{OUT} \) altered based on node voltages of the pass transistor. When sum of \( V_{OUT} \) and the threshold voltage of the pass transistor \( (V_{TH}) \) is less than gate voltage of pass transistor, the on-chip decoupling capacitor \( (C_L) \) is charged due to pass transistors is ON. In contract, \( C_L \) is discharged and \( V_{OUT} \) reduced depending on \( V_{REF} \), \( R_1 \) and \( R_2 \) caused by the pass transistor is OFF.

\[
V_{REG} = V_{REF} \times (1 + \frac{R_1}{R_2}) \tag{1}
\]

There are crucial aspects in defining LDO performance which can be concluded into three namely quiescent current, operating voltages, and regulating performance [68]. Drop-out voltage, load regulation, and transient output voltage variation are important specifications of any LDO design. Drop-out voltage \( (V_{DROP-OUT}) \) shown in (2) determines least input/output differential voltage where the circuit fails to regulate depending on switch “on” resistance \( (R_{ON}) \) and load current \( (I_{LOAD}) \). The load regulation is the output impedance \( (R_O) \) of the circuit as seen in (3). \( R_{O-PASS} \) is the output impedance of the pass transistor while \( F \) and \( A_O \) are feedback factor and open loop gain of the system respectively. Last, transient voltage
variations \( (V_{TVV}) \) occurs as sudden load current changes that controlled by closed-loop bandwidth, decoupling capacitor \( (C_L) \) and load current. The worst scenario happens when sudden load current steps from zero to its maximum \( (I_{LOAD\text{-}max}) \) as shown in (4) where \( \Delta t \) is the time needed for LDO to respond, and \( \Delta V_{TVV} \) is output voltage variation. The transient voltage variation must be minimized to fulfil accuracy requirement of the system. In this design, the regulated output voltage \( V_{REG} \) and power supply noises are critical measurements for power/signal integrity [24].

\[
V_{DROP\text{-}OUT} = R_{ON} \times I_{LOAD} \quad (2)
\]

\[
R_O = \frac{\Delta V}{\Delta I} = \frac{R_{O\text{-}PASS}}{1 + A_O \times F} \quad (3)
\]

\[
\Delta V_{TVV} = \frac{\Delta t \times I_{LOAD\text{-}max}}{C_L} \quad (4)
\]

In this LDO architecture, a classical two-stage OTA is utilized [37]. It has an external bias current \( I_{BIAS} \) and single ended output. The first stage of OTA comprises a differential NMOS input pair \( M_1, M_2 \) and two PMOS transistors \( M_3, M_4 \) as an active load. The drain current of \( M_1 \) and \( M_2 \), \( I_{BIAS}/2 \), is mirrored to \( M_5 \) and \( M_6 \), respectively. As shown in (5), the transconductance of OTA \( (G_m) \) is dominated by a gain factor \( B \) and a transconductance of either input pair [37]. Thus, \( B \) can be changed according to the ratio of \( (W/L)_{(5,6)} \) : \( (W/L)_{(3,4)} \) while \( g_{m(1,2)} \) depends on DC current \( (\sqrt{I_{BIAS}/2}) \) which can be modified by \( (W/L)_{(1,2)} \).

Moreover, the voltage gain of OTA \( (A_v) \) is displayed on (6) where \( Z_L \) is the output impedance of the OTA which can be calculated as given in (7). OTA has high output impedance due to \( r_{0(M_8)||r_{0(M_6)}} \) results in easier frequency compensation as \( C_L \) forms only dominate pole. Therefore, modifying \( C_L \) is the way of ensuring stability of the system since it controls the gain bandwidth and hence phase margin.
\[ G_m = 2 \cdot g_{m(1,2)} \cdot B \]  
\[ A_v = G_m \cdot Z_L \]  
\[ Z_L = \frac{R_0 || R_L}{1 + R_0 || R_L \cdot C_L} \]

### 2.3. On-chip 3D PDN models

In the literature, the basic electrical model of TSV is explored in [25], [26], [27], [76], and [77]. TSV coupling has a significant effect on signal integrity and power consumption due to its immense value (tens of femtofarads) [78], [79]. For accurate computation of TSV coupling, TSV-to-TSV or TSV-to-wire distance as well as TSV and wire dimensions have to be considered [80]. Thus, the TSV channel model and power/ground (wire) model are discussed for an accurate optimization method to achieve optimum performance for power and signal integrity.

![Figure 2-3](image_url)

**Figure 2-3** The 3D PDN based on 2D mesh power routing and 3D supply TSVs (a) 3D TSV model (b) TSV parameters (c) 3D TSV HFSS Model (d) Wire Model (e) PDN Mesh layout

### 2.3.1. TSV channel for 3D PDN

To optimize the 3D PDN structure, it is necessary to implement the TSV channel model. The electrical model and design parameters of the signal and power/ground TSV channels are illustrated in Figure 2.3(a) where \( R_{TSV}, L_{TSV}, C_{TSV}, R_{si}, \) and \( C_{si} \) represent TSV resistance, inductance, capacitance, and silicon substrate resistance and capacitance, respectively [28], [29], [81].

15
For analog/mixed signal generally, the noise-sensitive devices or interconnects need to be shielded by ground. Thus, one signal TSV such as data $i$ is surrounded by two ground TSVs inside the silicon substrate. The ground TSVs can provide a shielding for the signal TSVs and make the 3D routing signal integrity less noise-sensitive.

The model includes first coupling effect between two adjacent TSVs that can degrade the 3D channel signal integrity. Each TSV is surrounded by an insulation layer formed for isolating the TSV from the conductive silicon substrate and the corresponding capacitance ($C_{TSV}$), shown in (8), where $\varepsilon_0$, $\varepsilon_{r,ox}$, $l_{TSV}$, $r_{TSV}$, and $t_{ox}$ represent the permittivity of free space, relative permittivity of silicon oxide dielectric, height of TSV, radius of TSV, and insulator thickness, respectively. The TSV resistance is also analytically formulated from the geometric parameters according to (9). The first part reflects DC TSV resistance ($R_{dc}$), and the second term, $R_{ac}$, corresponds to the surface resistance ($R_{ac}$). As the frequency increases, current flows more on the surface of TSVs and causes $R_{ac}$ to be a more dominant term of (9). Thus, $R_{ac}$ should be considered to measure the frequency effect. This parameter is also called skin resistance. The $\rho_{TSV}$, $\mu_{TSV}$, $\sigma_{TSV}$, and $f$ represent TSV electrical resistivity, magnetic permeability, metal conductivity, and operating frequency. The TSV inductance is modeled, as in equation (10) where $\mu_0$ permeability of free space in addition to previously mentioned parameters, with a calculation of the two-wire transmission line [71] as shown in Figure 2.3(b). The analytic equations are derived from the physical configuration including the design parameters. A detailed evaluation is discussed and analyzed in [28-30].

$$C_{TSV} = \varepsilon_0 \varepsilon_{r,ox} \frac{2\pi l_{TSV}}{4 \ln \left( \frac{r_{TSV} + t_{ox}}{r_{TSV}} \right)}$$  (8)
\[ R_{TSV} = R_{dc} + R_{ac} = \left( \frac{\rho_{TSV} l_{TSV}}{\pi (r_{TSV})^2} \right) + \left( \frac{\tau_{TSV} \sqrt{\pi l_{TSV} \sigma_{TSV}}}{r_{TSV} \sigma_{TSV}} \right) \]  

(9)

\[ L_{TSV} = \frac{\mu_0 l_{TSV}}{2\pi} \times \left( \ln\left( \frac{2l_{TSV}}{\pi (r_{TSV})^2} - \frac{3}{4} \right) \right) \]  

(10)

The silicon resistance, \( R_{si} \), and capacitance, \( C_{si} \), can be calculated from (11) and (12), respectively. The \( \varepsilon_{si} \), \( d \), and \( \sigma \) represent silicon permittivity, TSV pitch, and silicon conductivity, respectively. The scalable equivalent electrical lumped model is simulated and verified in HFSS as shown in Figure 2.3(c).

\[ C_{si} = \frac{\pi \varepsilon_{si} l_{TSV}}{\ln\left( \frac{d}{\frac{1}{2} r_{TSV}} \sqrt{\left( \frac{d}{\frac{1}{2} r_{TSV}} \right)^2 - 1} \right)} \]  

(11)

\[ R_{si} = \frac{\varepsilon_{si}}{C_{si} \sigma} \]  

(12)

2.3.2. Power/ground wire model for 3D PDN

Latency, noise, and power consumption optimization in 3D design require TSV, wire model, RC estimation, and routing improvement. To estimate wire parameters shown in Figure 2.3(d), the metal layers assigned to inter-block signals must be defined. In conventional 2D designs, top metals are typically wide and thick. Thus, they are rarely implemented for signal routing. Due to this fact, using a wide portion of these metal layers for power delivery does not impact the overall signal interconnect length. As M4-M6 layers are the commonly used top metal type, we have considered the same metal stacking. The appropriate RC values for power/ground wire are evaluated in \( R_w \) and \( C_w \) in equations (13), (14), and (15) considering copper for wire material, respectively. \( w, t, h, \) and \( d \) are wire width, wire thickness, wire-ground space, and wire-wire space, respectively. For the layout, a conventional power delivery design method is implemented with middle metal layers to connect M1 layer in VDD/VSS rails to the top metal in
On each tier, power delivery network is usually structured as a regular mesh with metal tracks running in perpendicular to each other. Figure 2.3(e) shows the PDN mesh layout structure, its corresponding wire model [71], and a sample layout.

\[ R_w = \rho \frac{l_{wire}}{A_{wire}} \]  
\[ C_w = C_{1w} + C_{2w} \]  
\[ C_{1w} = \varepsilon \left( \left( \frac{w}{h} \right) + 2.97 \left( \frac{l}{h} \right)^{0.232} \right) \]  
\[ C_{2w} = \varepsilon \left( 0.23 \left( \frac{w}{d} \right) + 1.22 \left( \frac{l}{d} \right)^{1.384} \right) \left( \frac{d}{h} \right)^{-0.0398} \]

2.4. The evolutionary process of 3D PDN

To achieve optimum power efficiency, supply noises, and power/signal quality of a heterogeneous 3D PDN, the proposed architecture utilizes an evolutionary method. The optimization technique considers all design parameters, i.e., LDO placement, TSV geometry, and LDO circuit, simultaneously. When considering co-optimization of TSV structure and circuit sub-blocks, different performance parameters like area overhead and coupling effect in TSVs or speed and power consumption in the LDO conflict with one another. Thus, the optimization method requests trade-offs namely Pareto-fronts to find optimum design points among performance functions that compromise a computational efficiency [74], [75]. In this novel co-optimization, the critical TSV geometry and circuit design parameters are individually defined as input variables in MATLAB. The MATLAB-HSPICE toolbox interface is created to simulate the proposed 3D-PDN based transceiver using Genetic Algorithm (GA) as an evolutionary solver. After each simulation, MATLAB uses HSPICE output to evaluate the desired performance parameters and feed the new design to GA [74].

There are previous works on 3D PDN and TSV optimization such as in [61], [63]
respectively. However, to the best of our knowledge and based on the references, this work finds the best LDO placement for on-chip 3D PDN taking into consideration TSV coupling, LDO, and I/O interface by implementing the evolutionary process.

2.4.1. Evolutionary process of TSV

Figure 2.4(a) shows the design variable and specification of the TSV. Given an n-dimensional decision variable vector \( x = \{x_1, \ldots, x_n\} \) in the solution space \( X \), i.e., \( X \) represents the whole design parameter space, the TSV optimization problem can be formulated to find a vector \( x^* \) that minimizes a set of \( K \) objective functions \( z(x^*) = \{z_1(x^*), \ldots, z_k(x^*)\} \). The solution space \( X \) is restricted to a set of equality constraints denoted by \( g_j(x^*) = b_j \) for \( j = 1, \ldots, m \), and bounds on the decision variables. The equation (16) describes the GA’s fitness function including all TSV output signal specifications to obtain the optimum design parameters. To fulfill the desired value of the fitness, TSV needs lower settling time, error band, overshoot percentage, and peak time. It also requires a higher absolute peak value; \( K_j \) is the relevant coefficient for each design variable.

![Diagram of TSV input and output parameters](image)

Figure 2-4 Time corresponding specs of the TSV fitness function, construction of the GA chromosome string for (b) TSV and (c) proposed 3D PDN sub-block optimization
Fitness Function = \( \frac{k_0 \times \text{Settling time} + k_1 \times \text{errorband} + k_2 \times \text{Overshoot} + k_3 \times \text{peak time}}{k_4 \times \text{Peak absolute value}} \) \( (16) \)

Then, GA’s fitness function of (16) tries to extract all relevant TSV parameters by sizing through a reduced set of independent variables, namely genes. These variables shape the chromosome string. Figure 2.4(b) shows the corresponding string where \( d_{\text{bump}}, h_{\text{bump}}, \) and \( h_{\text{imd}} \) represent \( \mu \)bump diameter, \( \mu \)bump height, and IMD height respectively.

2.4.2. Evolutionary process of sub-blocks

To attain the optimum performance of the proposed architecture regarding power consumption, voltage stability, and supply fluctuations, it is essential to find the best design points of the sub-blocks. The optimum design can be achieved for LDO-error amplifier after implementing the optimal TSV geometry obtained from the first evolutionary process. To improve the efficiency and reduce search space, a hierarchical multi-objective optimization is implemented. Figure 2.4(c) shows an example of a sub-block chromosome of the evolutionary process; where, \( WL_{p11}, WL_{p12}, \ldots \) show transistors width (\( W \)) to length (\( L \)) ratio inside each sub-block and \( p \) represents the transistor pair number. The chromosome size would be different depending on each sub-block structure. The best (minimum) fitness value can be obtained using (17).

Fitness Function = \( \frac{k_0 \times \text{pwr} + k_1 \times \text{Vfluctuation}}{k_2 \times \text{VREG}} \) \( (17) \)

where \( \text{pwr} \) and \( \text{Vfluctuation} \) represent the circuit power consumption and output voltage fluctuation, respectively. In fitness function, \( k_0, k_1, \) and \( k_2 \) are relevant coefficients for power consumption, voltage fluctuation, and regulated output voltage. \( k_0, k_1 \) are chosen to be minimized while the \( k_2 \) should be maximized. Then, GA’s fitness function in (17) tries to extract all relevant PDN sub-blocks design parameters to achieve the best performance.

2.5. Experimental Results

TSV geometry, LDO placement/sizing, and TSV/LDO interaction are the factors that
significantly improve power efficiency, supply noises and obtain power/signal integrity. The evolutionary process is implemented on the factors mentioned above. The circuit simulation is conducted by using Cadence Spectre and HSPICE-MATLAB interface in MATLAB R2016a environment. The post-layout extracted simulation was performed considering parasitic capacitance and resistance. Table 2.1 displays a summary of the optimization parameters [81].

We first designed each architecture and found initial design value and range for each variable. Since the initial values are located near the optimal point for the first generation, it can increase the capability of finding the optimal global design points and minimize the computational burden and convergence time. The creation of the first generation is also a trivial process to prevent the optimization of being trapped in local minima instead of the global minima. The convergence time increases considerably without the evaluation of the initial operating point. The entire design took 4 hours using high-performance computing (HPC) server with 128GB RAM 16-core Intel(R) Xeon(R) CPU X5560 @ 2.8GHz 107 processor.

<table>
<thead>
<tr>
<th>Table 2.1 The Evolutionary Specifications of The Proposed Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Variables (sub-blocks/TSV)</strong></td>
</tr>
<tr>
<td><strong>Population Size</strong></td>
</tr>
<tr>
<td><strong>Number of Generation</strong></td>
</tr>
<tr>
<td><strong>Crossover</strong></td>
</tr>
<tr>
<td><strong>Mutation probability</strong></td>
</tr>
<tr>
<td><strong>Stopping Condition</strong></td>
</tr>
<tr>
<td><strong>Total Performance Enhancement (%)</strong></td>
</tr>
<tr>
<td><strong>Total Convergence Time (hr.)</strong></td>
</tr>
</tbody>
</table>

2.5.1. Evolutionary process of LDO design

Before the co-optimization of the entire 3D PDN, the evolutionary process is implemented on LDO to improve performance parameters such as regulated output voltage and
stability for LDO. First, the evolutionary process is implemented on the error amplifier. Figure 2.5 shows the improvement of its phase margin, bandwidth, and DC gain. After the evolutionary process, DC gain, bandwidth, and phase margin are improved from 49dB, 2GHz, and 54° as displayed in Figure 2.5(a) to 53dB, 5GHz, and 66° respectively as seen in Figure 2.5(b).

Second, the evolutionary process is executed on the overall LDO performance. Figure 2.6 shows the regulated output voltage waveform after each optimization iteration. It also displays the output voltage waveform with (/w) and without (w/o) optimization (opt). As expected, the optimized output is significantly improved regarding the regulated voltage and stability.

Figure 2-5 Simulated OTA in LDO (a) before and (b) after the evolution
2.5.2. Evolutionary process of TSV geometry

The evolutionary process is implemented on TSV that can relatively improve the key performance of the heterogeneous 3D stacked devices. Figure 2.7 shows the final output voltage waveform after each optimization iteration. It also displays the output voltage waveform with and without optimization. As predicted, the optimized output voltage is significantly improved regarding latency, power, and overshooting. The black line shows the output voltage signal in the state-of-the-art 2D interconnect.
The recommended TSV channel design parameters, summarized in Table 2.2, are utilized to achieve the best 3D PDN performance.

Table 2.2 Recommended TSV Design Parameters for 3D PDN

<table>
<thead>
<tr>
<th>TSV Length (Height)</th>
<th>35µm</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSV Pitch</td>
<td>45µm</td>
</tr>
<tr>
<td>TSV Diameter</td>
<td>25µm</td>
</tr>
</tbody>
</table>

2.5.3. Impact of LDO placement on 3D PDN performance

Integrating LDO in the on-chip 3D PDN provides significant improvements in power supply noises and signal/power integrity. Thus, six performance aware cases of LDO placement in 3D PDN are simulated as shown in Figure 2.8. In case of C1, (0-0-1), LDO is placed on tier 3. LDO is placed on tier 1 in case of C2, (1-0-0). In case of C3, (1-1-0), LDOs are on tier 1 and 2. In case of C4, (0-1-1), LDOs are on tier 2 and tier 3. LDOs are placed on tier 1 and 3 in case of
C5, (1-0-1). LDO is located on tier 2 in case of C6, (0-1-0). After obtaining the optimized sub-blocks and TSVs, the co-optimization is implemented to find the optimum architecture considering the complex interaction between TSVs, LDO and I/O interface for all cases. Since the power supply noise, i.e., maximum dc voltage droop, is worst at the topmost tier, the experimental results are performed on tier 3. Transient analysis has been carried out to measure the PDN output voltage. Figure 2.9 demonstrates stable waveforms of the output voltage in cases of C3-to-C6. Moreover, to find the best LDO placement of 3D PDN, a performance comparison in power consumption, PDN output voltage, delay, and output voltage fluctuation is summarized in Table 2.3. The peak-to-peak noise introduced to the input supply voltage \( V_{in \ tx} \) is \( \sim 170 \text{ mV} \). The results significantly are improved after implementing the optimization algorithm. As expected, when the number of LDO decreases, the PDN power consumption reduces. In stand-by and power-nap mode of the heterogeneous 3D device application, LDO can be always on, thus the number of utilized LDOs should be minimum. For instance, C5 (1-0-1) has the best performance in terms of voltage fluctuation, but it consumes power of 1.46X higher than C6 (0-1-0) after optimization. Consequently, the 3D PDN utilizing the common centroid placement of LDO (C6) has stable output voltage and minimum power consumption among C3-to-C6. It also has 10.2 mV output voltage fluctuations after optimization. Thus, C6 has the optimum overall performance.
Figure 2-9 Transient analysis of PDN DC output voltage Vs. input supply noise voltage ($V_{in\,tx}$) of each case.

Table 2.3 Simulated Results of Different Cases of 3D PDN Architecture

<table>
<thead>
<tr>
<th>PDN Case #</th>
<th>LDOs Placement T (1-2-3)</th>
<th>Total Power (mW)</th>
<th>PDN DC output (V)</th>
<th>PDN Output Voltage Fluctuation (mV)</th>
<th>Delay(ps)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>w/o opt / w opt</td>
<td>w/o opt / w opt</td>
<td>w/o opt / w opt</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>0-0-1</td>
<td>6.8 / 4.25</td>
<td>0.9 / 0.92</td>
<td>112 / 40</td>
<td>82 / 49</td>
</tr>
<tr>
<td>C2</td>
<td>1-0-0</td>
<td>6.3 / 4.18</td>
<td>0.9 / 0.95</td>
<td>118 / 42</td>
<td>85 / 51</td>
</tr>
<tr>
<td>C3</td>
<td>1-1-0</td>
<td>8.4 / 5.95</td>
<td>0.9 / 0.91</td>
<td>70 / 9</td>
<td>97 / 88</td>
</tr>
<tr>
<td>C4</td>
<td>0-1-1</td>
<td>9.2 / 6.26</td>
<td>0.9 / 0.9</td>
<td>65 / 8</td>
<td>93 / 86</td>
</tr>
<tr>
<td>C5</td>
<td>1-0-1</td>
<td>8.8 / 6.16</td>
<td>0.9 / 0.9</td>
<td>30 / 2</td>
<td>95 / 87</td>
</tr>
<tr>
<td>C6</td>
<td>0-1-0</td>
<td>6.6 / 4.22</td>
<td>0.9 / 0.96</td>
<td>75 / 10.2</td>
<td>78 / 46</td>
</tr>
</tbody>
</table>

For verification, the performance-aware 3D PDN with common centroid placement of LDO has been fabricated in a 65nm CMOS technology. Because 3D TSVs and TSIs prototyping are challenging in academia, we implemented the proposed 3D PDN by using the feasible 3D μbumps which could provide similar dimensions of the latest TSVs. The optimized μbumps geometry can be easily adapted to any 2.5D silicon interposer-and future 3D TSV-based ICs.
The 3D structure block diagram is shown in Figure 2.10(a). Figure 2.10(b) shows die photographs of the fabricated chips including transceivers, LDO, and 3D pads. The data rate and power consumption accomplished 1.9 Gb/s and 4.22 mW, respectively, with bit error rate (BER) \(< 10^{-12}\). Figure 2.10(c) shows the measured eye diagram of the top-most RX output of the proposed architecture to verify that the I/O interface is functioning properly as the power/signal integrity and stability of supply are achieved.

2.6. Conclusion

This work proposed a performance-aware 3D PDN utilizing a common-centroid regulator placement to improve power/signal quality, aggregate I/O data bandwidth, and energy efficiency. For accurate optimization, I/O interface was integrated with LDO and TSVs. An evolutionary algorithm was conducted to achieve minimum energy consumption and power supply noises.
despite the complex interaction among TSVs, LDO, and I/O interface. Consequently, the fabricated chip accomplished 4.22mW total power and 1.9 Gb/s data rate of the integrated I/O interface.
3. CHAPTER 3: 3D Baseband

3.1. Introduction

Advances in mobile devices such as smartphones demand to improve computation capabilities and video processing to achieve more bandwidth and higher energy efficiency. To obtain high energy efficiency, the I/O transceivers need to operate at high frequency and low power. Recent works for the mobile interface exploits traditional baseband (BB) transceiver which consumes power linearly as data rate increased [4]. Moreover, A pulse amplitude modulation (PAM) discussed in [11] is providing high aggregate data since it transceives data concurrently by utilizing multi-level-modulated signal. For instance, 4 PAM transceives twice the data rate for the same sampling rate by encoding every 2-bits into 4-signal levels. However, 4 PAM causes more design complexity and silicon area compared to traditional baseband.

Figure 3-1 Diagram of (a) a conventional 2D and (b) 3D memory interface

Inter-symbol interference (ISI) and high-frequency distortion decrease the maximum operating frequency of two-dimensional (2D) long transmission line (T-Line) (e.g., 10-30cm) as shown in Figure 3.1(a) in conventional 2D I/O interface. Thus, three-dimensional (3D) integration minimizes signal path lengths that significantly improves bandwidth/power efficiency
and decreases latency, noise, and chip area. Moreover, in 3D transceiver systems, shown in Figure 3.1(b), such as high-bandwidth memory (HBM), the pin-count of I/O through TSV channels has been considerably increased. For instance, 3D I/O pins can be expanded from 32 pins to 1024 [48]. Despite their high pin-counts, HBM enables only 2 Gb/s per pin. Since the conventional 2D baseband I/Os interface has power hungry components such as output driver and pre-driver degrades energy efficiency. Thus, a baseband transceiver utilizing an energy efficient and high-speed driver is proposed.

The rest of this chapter is organized as follows. Section II outlines the proposed 3D BB architecture design. The experimental results on the proposed architecture are presented in Section III. Finally, Section IV concludes this chapter.

3.2. Proposed 3D BB for mobile memory interface

3.2.1. 3D BB Mobile Architecture

To enhance data rate and energy efficiency, 3D BB architecture is designed. Figure 3.2 shows a 3D integration where the CPU mounted face down on a package board while DRAM is placed face up on the CPU. Point to point link is created by CPU and one DRAM connected by TSV channels. Since the distance between CPU and each DRAM is very short (~0.2mm), the
signal integrity improved significantly comparing to traditional 2D DRAM I/O interface [11].

CPU, controller in the memory interface architecture is connected to DRAM through TSV channel. TSV channel enables the output data to be transmitted vertically from CPU to DRAM where the data is going to be recovered.

In high-speed application, current mode logic usually exploited instead of CMOS logic [15]. However, CML dissipates more static power comparing to CMOS logic. In this architecture, the CMOS logic has been implemented in all blocks of the data path to increase energy efficiency. To the best of our knowledge, this is the first work that design and implement 3D BB architecture utilizing energy efficient and high-speed output driver. The results show the proposed architecture forming point to point link can operate at 5.8Gb/s/pin through TSV channel.

3.2.2. 3D BB Transceiver

Figure 3.3(a) shows proposed 3D BB transmitter that consists of a high-speed self-biased differential amplifier used as an input buffer, pre-driver, and inverter based output driver.

![3D BB Transceiver Diagram](image)

Figure 3-3 Proposed schematic of 3-D baseband transceiver (a) TX (b) RX (c) input buffer (d) output driver (e) output buffer (f) LDO

The TX is also contained LDO, shown in Figure 3.3(f), to minimize supply noises and improve signal integrity. Since the input signal to the chip has imperfections like fall/rise time results in timing errors and an incorrect output signal, input buffer using a self-biased differential amplifier is implemented. The input buffer, shown in Figure 3.3(c), is exploited to support high-frequency
operation with lower delay and minimize voltage offset, mismatch, and noise [18]. Figure 3.4 shows the 5.8 Gb/s output baseband signal from the input buffer.

![baseband signal](image)

Figure 3-4 5.8 Gb/s baseband signal after the input buffer stage.

The generated data ($D_{input}$) is fed into the input buffer followed by a baseband transmitter (TX) which utilizes output driver, an enhanced logic speed with resistive feedback drives the TSV channel as shown in Figure 3.2(d)[16]. The proposed output driver utilizes multiple stages of CMOS logic with resistive feedback. The output driver is implemented in an interleaving manner to enhance the data rate. The inverter with resistive feedback turns basic inverter into high-gain amplifier that will respond to the small signal and large signal as well. The resistor of feedback loop enables the output and input to settle at half the supply input voltage. Thus, the inverter based driver is implemented for TX due to its high energy efficiency in comparison to CML logic. In CMOS logic, ($W_p/W_n$) ratio usually has 2-3 depending on process technology to balance unequal hole and electron mobilities in digital inverters. Yet, this ratio is not valid anymore since the inverter here is exploited as a linear gain amplifier. Therefore, the ratio should be designed to maximize the overall performance. The intrinsic bandwidth, $f_{int}$, of the output driver is calculated as in (1) as found on [19]:

$$f_{int}$$
\[ f_{int} = \frac{g_{m_n} + g_{m_p}}{2\pi C_{ox}(W_n + W_p)L} \]  

(1)

The size of PMOS and NMOS have to be minimized to have the highest intrinsic bandwidth of the driver. The maximum gain bandwidth of CMOS technology is maximized with the NMOS and PMOS devices are equal in geometric size [20], [21]. In this architecture, output driver utilized at equal width for both devices to have the highest intrinsic bandwidth which confirmed by simulation results shown in Figure 3.5.

![Simulation results for the intrinsic bandwidth of output driver with a different value of \((W_p/W_n)\) ratio and supply voltage.](image)

Figure 3-5 Simulation results for the intrinsic bandwidth of output driver with a different value of \((W_p/W_n)\) ratio and supply voltage.

Because the output driver does not have a substantial capacitive load, the pre-driver stage can have less fan-out to maximize data rate operation by reducing the delay for each stage to be less than the unit interval (UI). Consequently, the fan-out has to be reduced or use CML logic instead results in higher power consumption. However, the feedback resistor of the inverter reduces the time delay and data jitter, so that the data path can handle high-speed data rate at the slight degradation of energy efficiency. Thus, the resistor plays a significant role in decreasing the power consumption as its value increases. Therefore, the feedback resistor has to be designed to meet high-speed bandwidth and energy efficiency which confirmed by simulation results as
displayed in Figure 3.6.

![Graph showing voltage over time for different feedback resistances.]

Figure 3-6 Simulation results for power consumption and rise time at different feedback resistance for 5.8 Gb/s data rate.

<table>
<thead>
<tr>
<th>Feedback Resistance</th>
<th>Normalized Power</th>
<th>Rise Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_f = \infty \Omega )</td>
<td>1</td>
<td>48.6 ps</td>
</tr>
<tr>
<td>( R_f = 10 , k\Omega )</td>
<td>1.97</td>
<td>39.8 ps</td>
</tr>
<tr>
<td>( R_f = 6 , k\Omega )</td>
<td>3.06</td>
<td>25.6 ps</td>
</tr>
</tbody>
</table>

Table 3.1 Simulation results of normalized power and rise time at different feedback resistance.

Regarding the receiver side (BBRX), shown in Figure 3.3(b), it consists of multiple stage inverter with feedback resistance and an output buffer which consists of pre-driver stage and pull-up/pull-down output driver. The transmitted signal goes through TSV channel. Moreover, the inverter with resistive feedback is exploited in the receiver side to enhance further the data rate. Figure 3.7(a) and (b) show the baseband signal after the TSV and inverter with a resistor feedback stage, respectively. The baseband RX utilizes push-pull output driver, shown in Figure 3.3(e), with two resistors in a series with a transistor to avoid impedance mismatch and reduce sensitivity to process, voltage, and temperature variations (PVT) [23]. The output driver schematic is sub-divide into a pull-up and a pull-down branch implemented as a PMOS and NMOS respectively connected in series with poly resistance where the impedance matched the measurement’s device impedance which is usually 50\( \Omega \). The poly resistance should be more
dominant comparing to the transistor impedance since the transistor impedance is not linear and susceptible to process variation. However, the transistor size must be significant to obtain negligible impedance which causes a high load capacitance for the pre-driver stage. Therefore, less power consumption costs a slight accuracy degradation of the pull-up/pull-down branch resistance. The output signal waveforms at the receiver end are shown in Figure 3.8.

Figure 3-7 Simulated baseband signal after (a) TSV channel (b) inverter with a resistor feedback stage.
3.2.3. 3D TSV Model

3D EM solver tool (HFSS) and electrical model of TSV is shown in Figure 3.9(c) [26]-[28], discussed thoroughly on chapter 2 section 3. Each TSV channel consists of TSV and two µbumps. Every signal TSV should be bounded by several grounded TSVs to shield the noise-sensitive devices or interconnect which improve 3D routing signal integrity. 3D TSV pitch of adjacent TSVs can be increased 2 or 3 times wider than 3D TSV diameter to minimize crosstalk and timing distortion. The TSV is designed with 65µm, 65µm, and 150µm for height, diameter, and TSV pitch. Consequently, the TSV parasitics are 280fF, 2mΩ, and 14pH for $C_{TSV}$, $R_{TSV}$, and $L_{TSV}$, respectively. Figure 3.9(a) and (b) show signal loss ($S_{21}$) and the HFSS model for the 3D TSV channel.

Figure 3-8 Simulated output signal at receiver end of baseband transceiver.

Figure 3-9 Simulation results of 3D TSV channel (a) signal loss (b) HFSS Model
3.3. Experimental Results

The proposed 3D baseband using inverter based with resistive feedback driver architecture simulated on 130 nm CMOS technology, and the layout for TX and RX are shown in Figure 3.10. The area of TX and RX are 123\(\mu\)m x 13\(\mu\)m and 94 \(\mu\)m x 32 \(\mu\)m, respectively. Figures 3.11 (a), (b), and (c) show the waveforms of input data, received data, and eye diagram at the receiver, respectively.

(a)

(b)

Figure 3-10 Layout for BB transmitter (a) and receiver (b)

(a)
Figure 3.11 Simulated waveforms (a) input baseband signal of 5.8 Gb/s (b) output data rate (c) eye diagram of output signal

Figure 3.12 shows power consumption according to the supply voltage. The proposed architecture power consumption minimized quadratically as supply voltage decrease due to $P = CV^2f$. The proposed BB memory I/O interface accomplished 5.8 Gb/s/pin with a power consumption of 7.841 mw. Thus, this architecture provides a promising solution to increase power efficiency and data rate for the future mobile memory interface.
Figure 3-12 Simulated power consumption for 10 Gb/s at different supply voltages in 130nm CMOS technology.

Table 1 summarized and compared the proposed 3D BB I/O interface performance with prior work. The proposed 3D BB memory interface achieved energy efficiency and data rate of 4.62 mW and 5.8 Gb/s/pin respectively (E/bit Efficiency=0.79 pJ/b/pin).

Table 3.2 Performance comparison of the proposed baseband I/O interface with published works

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CMOS Tech.</td>
<td>100 nm</td>
<td>30 nm</td>
<td>180 nm</td>
<td>130 nm</td>
<td>22 nm</td>
<td>65nm</td>
<td>130 nm</td>
</tr>
<tr>
<td>Dimension</td>
<td>2D</td>
<td>2D</td>
<td>2D</td>
<td>2D</td>
<td>2D</td>
<td>2D</td>
<td>3D</td>
</tr>
<tr>
<td>DC Supply</td>
<td>1.8 V</td>
<td>1.2 V</td>
<td>1.8 V</td>
<td>1.2 V</td>
<td>N/A</td>
<td>-</td>
<td>1.1</td>
</tr>
<tr>
<td>Data Rate</td>
<td>4Gb/s</td>
<td>1.6 Gb/s</td>
<td>5Gb/s</td>
<td>6Gb/s</td>
<td>10Gb/s</td>
<td>16Gb/s</td>
<td>5.8 Gb/s</td>
</tr>
<tr>
<td>Total Power</td>
<td>28.44 mW</td>
<td>N/A</td>
<td>87 mW</td>
<td>95 mW</td>
<td>12 mW</td>
<td>14.7 mW</td>
<td>4.62 mW</td>
</tr>
<tr>
<td>E/bit Efficiency</td>
<td>7.11</td>
<td>3.7</td>
<td>17.4</td>
<td>15.83</td>
<td>1.2</td>
<td>0.919</td>
<td>0.79</td>
</tr>
</tbody>
</table>

3.4. Conclusion

The proposed Baseband transceiver is utilizing inverter based driver with resistive feedback in 130 nm CMOS technology at 1.1 V supply. Using 3D integration and CMOS logic enable 5.8 Gb/pin/s and 4.62 mW for I/O data bandwidth and energy consumption, respectively.
The proposed baseband transceiver is promising I/O interface for future 3D stacked memory due to high energy efficiency and data rate throughput.
4. CHAPTER 4: 3D RF bands

4.1. Introduction

The content of this chapter is an extended version of the published paper [96]. High bandwidth and energy efficient memory interface is required to meet the demand of high computation and support video/graphics processing. To obtain high energy efficiency, the I/O transceivers need to operate at high frequency and low power. However, inter-symbol interference (ISI) and high-frequency distortion decrease the maximum operating frequency of 2D long transmission line as shown in Figure 4.1(a). Three-dimensional (3D) IC designs, shown in Figure 4.1(b), demonstrate vital advantages over 2D utilizing very short TSV [11] (e.g., 30-70 µm). This technology enables multiple stacking of chips vertically using TSV. 3D integration minimizes signal path lengths, significantly improves bandwidth/power efficiency, and decreases latency, noise, and chip area. Moreover, in 3D transceiver systems, the pin count of 3D high-bandwidth memory (HBM) I/O through TSV channels has considerably increased. For instance, 3D I/O pins can be expanded from 32 or 64 pins to 1024 or 2048 pins.

Figure 4-1 Diagram of a conventional (a) 2D and (b) 3D I/O transceiver.
To improve energy efficiency and performance, the proposed RF I/O interface utilizes ASK modulation/demodulation scheme. Moreover, we propose matching capacitance along with the TSV to improve further signal integrity. Prior works have been conducted to reduce signal reflection by TSV geometry optimization [12]. Decreasing signal reflection by adding capacitance between TSV-interconnect junction and ground is implemented [82]. However, TSV coupling and silicon substrate effects along with a mobile memory I/O utilizing RF band interconnect have not been studied. This chapter considered both effects and improved RF I/O interface regarding power consumption and data rate. It also presents two RF band, i.e., 10 GHz and 30 GHz, a transceiver that can be utilized for 3D I/O interface.

The rest of this chapter is organized as follows. Section II outlines the proposed 3D RF architecture design. Section III discusses TSV and impedance matching of the proposed 3D RF I/O interface. The experimental results of the proposed architecture are presented in Section IV. Finally, Section V concludes this chapter.

4.2. Proposed 3D RF Memory interface

4.2.1. 3D RF Mobile Architecture

The proposed architecture of 3D RF is similar to the 3D BB. As shown in Figure 4.1, the 3D integration enables very short channel where the CPU, tier1, is mounted face up on a package board while DRAM, tier 2, is placed face toward the CPU. Thus, Point to point link is created between CPU and DRAM by using TSV channels. Since the distance between CPU and each DRAM is very short (~0.2mm), the signal integrity improved significantly comparing to traditional 2D DRAM I/O interface [11]. CPU, controller in the memory interface architecture, is connected to DRAM through TSV channel. Therefore, TSV channel enables the output data to be transmitted vertically from CPU to DRAM where the data is going to be recovered.
4.2.2. 3D RF Transmitter for 10/30 GHz

RF band transceiver comprises an RF transmitter (RFTX) and RF receiver (RFRX). RFTX consists of input buffer, voltage control oscillator (VCO), and ASK modulator. The input buffer is similar to what has been used in the baseband that improves the signal and eliminates mismatch, voltage offset, and timing error. The phase margin, bandwidth, and the common mode of output voltage are essential, e.g., the bandwidth of the input buffer should be much larger than the data rate to ensure proper functionality.

![Diagram of 3D RF Transmitter](image)

**Figure 4-2 3D RF I/O transmitter**

Second, a VCO which is typically used to generate the carrier frequency. It has also been implemented in many electronic systems, and its application ranges from clock generation in microprocessors to carrier synthesis in mobile phones. LC oscillator and ring oscillator are commonly used in electronic systems. However, the LC oscillator is used where there is a stringent requirement for phase noise while the ring oscillator is favored when it comes to power consumption compared to LC oscillator. Thus, the LC oscillator is selected in the proposed architecture to minimize the phase noise. High performance and low power oscillator is always a critical challenge [17]. Thus, the topology we used to implement for VCO is crossed coupled oscillator as shown in Figure 4.3 which enables low power consumption, maximum operating frequency and voltage swing [34]. This topology is exploited in both RFTX for 10 and 30 GHz bands.
Figure 4-3 Crossed coupled oscillator configuration.

To establish oscillation, the circuit demands \( R_p \ast g_m \geq 1 \) where \( R_p \) is the equivalent parallel resistance of the load (LC tank). Thus, \( R_p \ast g_m \geq 4 \) is selected to have strong oscillation where \( R_p = Quality\ factor(Q)^2 \ast R_s \) where \( R_s \) is the lossy component of the LC tank. Thus, careful design of LC tank and bias current are crucial to attain strong oscillation and performance targets such as oscillation frequency, phase noise, voltage swing, and power consumption. M1 and M2 (input pair) are chosen to be NMOS, so the oscillator can achieve maximum carrier frequency and voltage swing. M3 and M4, FET capacitance, are used as a varactor (variable capacitance) to make the oscillator tunable. Moreover, \( R_p = Q \ast W_s \ast L \) and voltage swing = \( I_{bias} \ast R_p \) where \( W_s, L, I_{bias} \) are oscillation frequency, inductor of the LC tank, and biasing current, respectively. Thus, using spiral inductor with high quality factor will significantly improve phase noise and voltage swing. The quality factor is limited by the chip area, e.g., increasing the diameter of the spiral inductor will improve further the quality factor. To enhance further the voltage swing, the biasing current can be increased. Thus, Silicon area, power consumption, and voltage swing requirements are the design constrains for LC-VCO.

Figure 4.4 and 4.5 show simulated results for the quality factor (Q), the series resistance of the inductor, and inductor value of spiral inductors that are designed for 10 GHz and 30 GHz.
transmitter, respectively. The spiral inductor with a diameter = 230µm, width = 10µm, and s = 5µm, the distance between two coils, and n = 2 where n is the number of turns is implemented for 10 GHz band. Also, another spiral inductor with 110µm, 10µm, and 5µm for diameters, width, the distance between coil, and number of turns, respectively, is employed for 30 GHz band.

Figure 4-4 Spiral inductor (a) Layout (b) simulated series resistance \( R_s \) (i.e., real(Z(1,1)), inductance value \( L \), and quality factor \( Q \) of 10 GHz band
Figure 4-5 Spiral inductor (a) Layout (b) simulated series resistance ($R_s$) (i.e., real(Z(1,1))), inductance value (L), and quality factor (Q) of 30 GHz band

Note that the oscillation frequency designed for each VCO taking into consideration the highest quality factor at the desired oscillation frequency for each inductor (10 GHz for an inductor with diameter 230µm and 30 GHz for an inductor with a diameter of 110µm) to obtain
maximum voltage swing and phase noise for each RF band. Figure 4.6 and 4.7 show output voltage swing and phase noise of the LC VCO for 10 GHz and 30 GHz respectively. The LC-VCO obtains 2.1V and 1.2V voltage swing for 10 GHz and 30 GHz, respectively. The attained the phase noise for LC-VCO are 107.4 and 107.3 dBc/Hz @ 1 MHZ offset from the carrier frequency of 10 GHz and 30 GHz band, respectively.

Figure 4-6 Simulated VCO output voltage and phase noise @ 10GHz (a) Voltage Swing (b) Phase noise
The last component of RFTX is ASK modulation. The ASK modulation and demodulation is relatively simple and cheap and has less power consumption compared to other sophisticated modulation schemes like FSK [35]. Thus, ASK modulator is utilized in both RF bands (10 GHz and 30 GHz) transmitter. The modulator is used to shift the amplitude of carrier wave proportional to the amplitude of the modulating signal. In ASK modulating signal, the digital signal value either 1 or 0 which indicate the presence or absence of the carrier signal. Figure 4.8 shows the ASK circuits that employed in both RF bands (10 GHz and 30 GHz) transmitter. ASK modulator circuit passes the signal generated by the VCO when the transistor
switch is ON and blocked the VCO signal when the transistor switch is OFF. M1 and M2 where the signal from VCO is arrived to while M3 and M4 operate as switches based on digital data 0, block VCO signal, or 1, pass the VCO signal. The PMOS transistors M5 is utilized to reduce the noise and amplitude of the non-modulated signal.

Figure 4-8 ASK modulator used in both RF bands (10/30 GHz) transmitter

Figure 4.9 shows the simulated input and output waveforms of ASK modulator for 10 GHz and 30 GHz, respectively. The input data (baseband signal) to ASK modulator at 4.2 Gb/s as shown in Figure 4.9(a) and (c) while (b) and (d) show ASK modulator output signal for 10 GHz and 30 GHz, respectively. Figure 4.10 shows the layout of both RF bands transmitter. The area of RFTXs excluding spiral inductors for 10 and 30 GHz bands are 123µm x 13µm and 94 µm x 32 µm, respectively. Both RF bands transmitter are simulated in 130nm CMOS process technology. Note that the RFTX 10 GHz and 30 GHz designed at 12 GHz and 32 GHz instead respectively to handle the frequency degradation after considering parasitic capacitance in layout and obtain the desired transmitted RF bands (10 GHz and 30 GHz).
Figure 4-9 Simulated waveforms of (a) 4.2Gb/s input data rate and (b) ASK output for 10GHz RFTX & (c) 4.2Gb/s input data rate and (d) ASK output signal for 30 GHz RFTX
4.2.3. 3D RF receiver for 10/30 GHz

Figure 4.11 shows RF band receiver (RFRX) that comprises mutual-differential self-oscillating mixer, three-stage differential amplifier, buffer converter, inverter chain, and output driver. This section will discuss each component of the RFRX.

Figure 4-11 3D I/O RF receiver architecture
An ideal mixer has two input, RF and local oscillator (LO) frequencies, and one output frequency to produce sum and difference of RF and LO frequencies. There is two type of mixer circuits can be designed either passive or active mixer. Active mixer outperforms in conversion loss and noise figure, especially when using double sideband comparing to the passive mixer. Thus, differential mutual mixer, shown in Figure 4.12, is utilized to amplify incoming ASK modulated RF carrier and down-convert the RF to the baseband signal and remove undesirable RF and LO signals [36]. The modulated signal coming from the RFTX fed to four inputs M1 to M4 where M1 and M2 with source degeneration used to provide voltage to current conversion. The high pass path through capacitor C1 and C2 to the source of M3 and M4 is LO signal

![Figure 4-12 Self Oscillating-differential mixer](image)

The incoming signals, RF and LO to the mutual mixer, have same carrier frequency (self-oscillating mixer) which results in the double carrier for the additive part and baseband signal for the subtractive part. Thus, the down-converted signal from the mixer is initially a baseband signal that is transmitted from RFTX. The sum part will be reduced and filtered out by subsequent stages. The down-converted baseband signal (transmitted data) amplified rail-to-rail to be recovered via subsequent circuit block in the receiver chain. Figure 4.13 displays simulated
output from the self-oscillator mixer for both RF bands. However, the high-frequency peaks are smoothed by subsequent stages. Three stages of the differential amplifier are shown in Figure 4.14 to amplify the difference between the input voltage signals and reject the common mode voltage that suffers from noise.

![Simulated results from self-oscillating mixer](image)

Figure 4-13 Simulated results from self-oscillating mixer (a) 10 GHz (b) 30 GHz

The output of differential amplifier is given in (1) where $V_{in+}$ and $V_{in-}$ are the input voltages, $A_d$ and $A_c$ are the differential and common mode voltages of the differential amplifier, respectively.

To attain high gain and cancels common mode noise, the common mode rejection ratio defined in (2) should be as high as possible. Since the differential amplifier has a high gain that can be calculated in (3) where $g_m$ is the transconductance, it is utilized in this design. Thus, three stages of the differential amplifiers are exploited to amplify the incoming signal and reject high-frequency peaks as shown in Figure 4.15. Note that the output of the mixer is differential signals and have two different common-mode due to down-conversion implementation that causes a problem for subsequent differential stages.

$$V_{out} = (V_{in+} - V_{in-}) + A_c \frac{(V_{in+} - V_{in-})}{2}$$  (1)

$$CMRR = \frac{A_d}{A_c}$$  (2)
\[ A_d = -g_m R_D \] 

Figure 4-14 Differential amplifier schematic in both RF-band (10/30 GHz) receiver

Figure 4-15 Simulated output signal of the differential amplifiers for RF bands (a) 10 GHz (b) 30 GHz

Thus, buffer converter with RC- feedback as shown in Figure 4.16(a) is implemented to control overall gain and settle the common mode voltage at the half supply voltage. The RC feedback circuit controls transient response for best recoverable data rate. As shown in Figure 4.17, the differential output signal for each band has the same common mode. Then, inverter chains is used to enable rail to rail output signal. Last, an output driver, shown in Figure 4.16 (b), is employed with the equivalent resistance of 50 Ω to match the off-chip components.
Figure 4-16 Circuit configuration for (a) converter buffer and (b) output driver

Figure 4-17 Simulated output signal for both RF bands receiver (a) 10 GHz (b) 30 GHz

The area of RFRXs for 10 and 30 GHz bands are 216µm x 120µm and 214µm x 120µm, respectively, shown in Figure 4.18.
Figure 4-18 Layout for both RF bands receiver (a) 10 GHz (b) 30 GHz

4.2.4. TSV and Impedance Matching

Figure 4.19(a) and (b) show the simulated schematic of RF I/O interface for 10/30 GHz band. To improve speed/power efficiency and signal integrity, TSV and tunable capacitances are employed. This work improves 3D RF I/O interface regarding power consumption and data rate by utilizing impedance matching and TSV including coupling and the silicon substrate. Thus, matching capacitances \( C_{tune1}, C_{tune2} \) for 10 GHz band and \( C_{tune3}, C_{tune4} \) for 30 GHz band are tuned to improve impedance matching. Figure 4.19(c) shows the implemented TSV model and matching network where \( R_{TSV}, L_{TSV}, C_{TSV}, \) and \( R_{si}, C_{si} \) represent TSV resistance, inductance, capacitance, silicon substrate resistance and capacitance, respectively.
The noise-sensitive devices typically are shielded by ground in analog/mixed signal to improve signal integrity. Thus, each signal TSV is bounded by two ground TSVs. TSV parameters are shown in Figure 4.19(d) where \( l_{TSV} \), \( r_{TSV} \), \( t_{ox} \), and \( d_{TSV} \) are TSV height, TSV radius, insulator thickness, and the pitch between two TSVs, respectively. Each TSV channel includes TSV and two \( \mu \) bumps. Figure 4.19(e) shows TSV equivalent electrical lumped model that simulated and verified in HFSS. The matching capacitances \( C_{tune1}, C_{tune2}, C_{tune3}, \) and \( C_{tune4} \) are implemented along with TSV, transformer, and mixer models to ensure matching and improve \( S \)-parameters \( (S_{11}, S_{22}, \) and \( S_{21} \)) for 10/30 GHz RF band. The \( S_{11}, S_{22}, \) and \( S_{21} \) represent return loss or reflection coefficient at port 1, return loss at port 2, and the power transferred from port 1 to port 2, respectively. The results of \( S_{11}, S_{22}, \) and \( S_{21} \), shown in Figure 4.20, have attained -
17.9dB, -6.9dB, and -3.2dB, respectively, showing better filtering with TSV and matching capacitance ($C_{tune1}$, $C_{tune2}$) at the desired 10 GHz band. Similarly, Figure 4.21 shows that $S_{11}$, $S_{22}$, and $S_{21}$ have achieved -17.9dB, -10dB, and -2.9dB, respectively, by proper tuning of matching capacitence ($C_{tune3}$, $C_{tune4}$) at 30 GHz band.
Figure 4-20 Simulated S-parameters (a) $S_{11}$, (b) $S_{21}$, and (c) $S_{22}$ for 10 GHz RF band.
Figure 4-21 Simulated S-parameters (a) $S_{11}$, (b) $S_{21}$, and (c) $S_{22}$ for 30 GHz RF band

4.3. Simulation Results

The proposed architecture has been simulated in 130nm CMOS technology considering resistance and capacitance parasitics. Figure 4.22 shows input/output waveform for each RF band. The simulated eye diagrams of 4.2Gb/s for both RF bands output are shown in Figure 4.23.

Figure 4-22 Simulated input & output waveforms at 4.2Gb/s for (a) 10 & (b) 30 GHz I/O interface
4.4. Conclusion

This chapter presents dual RF bands transceiver that would be a promising solution for future 3D stacked mobile memory interface due to higher energy efficiency and data rate throughput. The both RF transceiver are designed and simulated in 130 nm CMOS technology. They also achieve aggregate I/O data rate of 8.4 Gb/s (i.e., 4.2 Gb/s from each RF band) with power consumption of 14.18 mW and 15.94 mW for 10 GHz and 30 GHz, respectively. The bit error rate of the eye diagram on both bands has $< 10^{-12}$.

Figure 4-23 Simulated eye diagram at 4.2Gb/s for (a) 10 & (b) 30 GHz receiver
5. CHAPTER 5: 3D MBI System

5.1. Introduction

Future mobile devices demand the I/O interface to offer high computational capabilities and energy efficiency [1]. Reducing supply voltage and increasing operating frequency are promising mechanisms but impose strict challenging to maintain an adequate overall performance [2].

Recent works for the mobile interface in [2],[3],[8] exploits the traditional baseband (BB) transceiver which consumes power linearly as data rate increased [1]. Another work [4],[6],[7] has utilized serial link that attains high bandwidth, reduces power dissipation, and needs fewer bus lines. However, serial link suffers from long initialization time (~ 1000 clock cycles) which is incompetent with mobile DRAM I/O that requires a fast switching among active, standby, self-refresh, and power-down operations modes [1]. Moreover, a pulse amplitude modulation (PAM) discussed in [9],[10] is providing high aggregate data since it transceives data concurrently by utilizing a multi-level modulated signal. For example, 4 PAM transceives twice the data rate for the same sampling rate by encoding every 2-bits into 4-signal levels. Another work, dual-band Interconnect (DBI) [22] using baseband (BB) and radio frequency band (RF) to enable simultaneously bidirectional memory I/O through a shared off-chip transmission line (T-Line). It obtains aggregate data of 8.4 Gb/s and consumes 92mW. To obtain high energy efficiency, these transceivers demand to operate at high frequency and low power. However, inter-symbol interference (ISI) and high-frequency distortion limit the maximum operating frequency of 2D long channel (e.g., 10-30cm)[11].
Three-Dimensional (3D) IC designs demonstrate key advantages over 2D using very short through-silicon via (TSV) (i.e 50 – 200 µm) [12],[13],[14]. This technology enables multiple stacking of chips vertically using TSV. 3D integration minimizes signal path lengths and significantly improves delay, noise, and chip area. Moreover, in 3D transceiver systems, the memory bus width through TSV channels has significantly increased. For instance, I/O pins can be increased from 32 or 64 bits to 1024 or 2048 pins.

Figure 5-1 Proposed 3D MBI I/O for mobile memory interface (a) 3D integration (b) 3D MBI architecture

The 3D multi-band (Baseband and Dual RF bands) I/O interface is presented. This chapter combines all the concepts that have been discussed in previous chapters to obtain the objective of the thesis. The primary goal of the thesis is transceiving the three signals simultaneously to achieve high energy efficiency and higher aggregate data rate utilizing 3D TSV technology. To maintain signal/power integrity for the proposed architecture, a low dropout voltage regulator (LDO) and a specific arrangement for TSV array have been utilized to obtain high accuracy of the supply voltage; the voltage variation should be less than 10% of the supply voltage (Vdd). Moreover, low noise amplifier has been added in the chain of RFRX for 30 GHz to improve and filter the signal strength of 30 GHz band further and for recovering the transmitted data successfully.
This chapter is organized as follows. Section II presents the architecture of the 3D multi-band I/O interface. Section III discusses the low noise amplifier for 30 GHz receiver to enable receiving and recovering data simultaneously for both RF receivers. Section IV shows a selective band transformer to enable transmitting BB + 2 RF concurrently. Experimental results, layout, and die photos presented in Section V. Then, we ended this work with a conclusion.

5.2. Proposed 3D MBI utilizing TSV and LDO

To enhance data rate and improve power/signal integrity, multi-band transceiver, TSV, and LDO are implemented in this chapter. It comprises an LDO and multi-band transmitter that are placed on the bottom (Tier 1) and multiband receiver which is placed on the top chip (Tier 2) to form I/O interface. It also has TSVs to enable the communication between the two chips. Figure 5.1(a) shows the 3D integration for the proposed architecture in this chapter. A point to point link is formed by CPU (bottom chip) connected to DRAM chip (top chip) through TSV channels exploiting 3D integration technology. Since the distance between CPU and each DRAM is very short (e.g., 35µm), the signal integrity enhanced relatively compared to traditional 2D DRAM I/O interface [12]. TSV channel enables the output data to be transmitted vertically from CPU (Tier 1) to DRAM (Tier 2) where the data is going to be recovered. Every TSV signal is bounded by TSV grounds to shield the noise-sensitive devices or interconnect which improve 3D signal integrity. 3D TSV pitch of adjacent TSVs can be increased two or three times wider than 3D TSV diameter to minimize crosstalk and timing distortion. Consequently, the proposed architecture uses 10 TSV channels to keep every data TSV is bounded by ground/power TSV to enhance the signal integrity of data channels; the labeled TSVs represent ground, data, and power channels. To boost further signal/power integrity, LDO is utilized to minimize supply noises (i.e., DC voltage fluctuation) to maintain the accuracy of DC supply. Note that the analog circuits are more susceptible to noise compared to digital
circuits, so voltage fluctuation requirement is more stringent in analog circuits [65].

5.2.1. 3D MBI I/O Interface

The design of the baseband and 10/30 GHz RF band are discussed in chapter 3 and 4, respectively. Figure 5.1(b) shows proposed 3D multi-bands (Baseband and Dual RF bands) I/O interface architecture. In the transmitter side (bottom chip), shown in Figure 5.2, comprise baseband (BBTX) and both RF transmitters (RFTX for 10 & 30 GHz). Moreover, the LDO is also utilized in the bottom chip to minimize the most effect of supply noises which typically comes from the bottom chip (Tier 1). The BBTX consists of the input buffer and an output driver as discussed in chapter 3. Moreover, the RFTX 10/30 GHz comprises input buffer, LC-VCO, and ASK modulator to modulate the baseband signal with low power consumption, RFTX is discussed in details in chapter 4. To enable transmitting the three signals (BB+2 RFs) simultaneously, selective band transformer is implemented [22]. It is carefully designed, so the smaller transformer and larger transformer in diameter enable 30 and 10 GHz, respectively.

![Figure 5-2 Schematic of 3D MBI I/O transmitter (Tier 1)](image)

Regarding the receiver side (top chip) as shown in Figure 5.3, Tier 2, the RFRX 10/30
GHz comprises a differential mixer, multiple stages of the differential amplifier, buffer converter, inverter chain, and an output driver. RFRX is described in chapter 4. Since both RF bands are transmitted simultaneously, LNA should be utilized as the first component of receiver chain to provide filtering for the desired frequency (e.g., either 10 GHz or 30 GHz band). However, for the feasibility of 3D integration for academic purpose, the top chip has to be much smaller than a bottom chip to mount top on the bottom and wire-bonding. Thus, narrowband LNA has been only utilized for 30 GHz. The LNA component is used in addition to the self-oscillating mixer, differential amplifier, buffer converter, inverter chain, and output driver to compensate the loss of signal strength of 30 GHz band due to the parasitics. The LNA also enhance filtering to recover the transmitted data rate. Moreover, The BBRX receiver contains an inverter based driver with resistive feedback, inverter chain, and an output driver. BBRX is introduced in chapter 3. Thus, utilizing band selective transformer and low noise amplifier for 30 GHz enable the receiver for BB and Dual RF bands to successfully recovers the original input data streams that sent from the transmitter. The 3D multi-band (BB+2RF interconnect) I/O interface, shown in Figure 5.4, is compatible with CMOS technology utilizing on-chip transformers, 3D integration, and 30-GHz LNA results in minimum layout and high energy efficiency and data rate compared to prior work. Thus, multi-band I/O interface promising for the future advanced memory interface.
5.3. **Low noise amplifier design for 30 GHz**

Low noise amplifier is used to minimize noise distortion from the transmitter part and provide gain. As shown in Figure 5.5 the signal strength of the transmitted signal is 1.2V and
550mV for RFTX at 10 GHz and 30 GHz, respectively, considering resistance and capacitance parasitics. Consequently, the signal strength of 30 GHz RF band is deficient and distorted compared to 10 GHz band which prevents recovering the transmitted data correctly at the receiver end. Thus, common source low noise amplifier has been utilized to amplify the RFTX and improve filtering for 30 GHz band. Figure 5.6 shows common source cascade LNA used in the proposed 3D-MBI because it minimizes the Miller effect at the gate results in better noise performance. Thus, the input impedance can be calculated from (1) assuming the impedance of the previous stage is 50 ohm. The impedance matching network is implemented to enable the LNA to bypass the 30 GHz frequency at resonant. Because we have LC resonance considering \( L_1 \) and parasitic capacitance of M1 (\( C_{gs} \)) in series at 30 GHz frequency, the gain improved by the quality factor of LC resonance as seen in (2). The cascaded transistor M2 increases the output impedance and improves transconductance (gm) to pass the signal from M1 to the transformer. Moreover, transformer with L and variable capacitance (c) is implemented to ensure the transformer resonate at 30 GHz frequency, so the signal pass through to the input of the subsequent circuit. The transformer also enables converting the input signal from being single to differential which relatively reduces layout of the circuit.

![Graph](image)

(a)
Figure 5-5 The voltage swing (signal strength) of the RFTX (a) 10 GHz band (b) 30 GHz band

\[
Z_{in} = \frac{(g_m L_s)}{C_{gs}} + J(w(L_s + L_1) - \frac{1}{W C_{gs}}) 
\]

(1)

\[
\frac{V_{gs}}{V_{in}} = \frac{\sqrt{(L_s + L_1) C_{gs}}}{(g_m L_s) C_{gs}} \frac{1}{C_{gs}} 
\]

(2)

The simulated signal gain \((S_{21})\), return loss \((S_{11})\), and noise figure \((NF)\) are attained 6dB, -38dB, and 3.5dB, respectively as shown in Figure 5.7(a), (b), and (c). Thus, the 30 GHz band voltage swing has been improved and attained 876mV, as shown in Figure 5.7(d) which enable the transmitted data rate to be recovered at the front end.

Figure 5-6 Narrowband Cascade CS-LNA for 30 GHz

69
Figure 5-7 Simulated CS-LNA for 30GHz (a) $S_{21}=6$dB (b) $S_{11}=-38$dB (c) and signal swing after LNA
5.4. **Band Selective Transformer**

To enable transmitting three bands (Baseband and 2 RFs) simultaneously, the band-selective transformer is utilized as shown in Figure 5.8. The metals in olive green and purple are thick aluminum (MA) and thick copper (E) layers, respectively, in 130nm CMOS technology. The MA and E metals are used for transformer layout to obtain high frequency and quality factor. Thus, the differential input for larger and smaller transformer are fed by the modulated signal of RFTX for 10 GHz and 30 GHz, respectively. Then, both RF signal coupled to lower metal labeled in purple E (thick copper) layer in 130nm CMOS technology to enable transmitting the three bands. Thus, the band-selective transformer enables the both RF bands to couple to the lower metal (E), so the output signal that combines both RF bands goes through TSV channel. Note that the RF bands (10 & 30 GHz) are chosen to have enough separation to enable recovering the data transmitted by both RF bands and minimize the area for the passive components. Consequently, the on-chip transformer with a turn ratio of 2:1 has been carefully designed considering the dual-band RF frequencies. Proper sizing for the on-chip transformer for both RF bands and using 3D TSV reduce the reflected signals and enable simulations transmission for three signals (2 RFs band + Baseband) comparing to off-chip/on-Chip transmission line. As shown in Figure 5.8, the first primary coil has outer diameter 230µm and the second primary coil has 110µm for 10 GHz and 30 GHz respectively to enable transmission simultaneously. The baseband signal is going to the center tap of the primary coil and transferred to the channel in common mode then goes through TSV channel. Then, the baseband signal extracted and goes to the baseband receiver input to be recovered at the end. On the other hand, each RF bands fed to the differential ports of the targeted primary coil then coupled to single ended channel through secondary coil. Then, the output of the secondary coil goes through TSV
channel then to the input of each band receiver.

Figure 5-8 Layout of the band-selective transformer.

5.5. Experimental Results

A prototype of 3D MBI transceiver chip is fabricated in the IBM 130nm CMOS process to present simultaneous 3D multi-band data communication. Figure 5.9 shows Die Photo of 3D MBI transceiver which occupies an active area of 0.20mm². The bottom chip has to be larger than top chip for feasible integration and testing. The bottom chip that consists of RFTX 10/30GHz, BBTX, LDO, and band selective transformer as shown in Figure 5.9(a). Figure 5.9 (b) shows that the top chip which comprises RFRX 10/30 GHz, LNA for 30 GHz band, and BBRX.
To complete 3D integration, the printed circuit board (PCB) has been designed for the bottom. Thus, the top chip will be flipped and placed on the bottom chip where the two chips are communicating through TSVs. There are two layers in the PCB design that are used where the top layer used for high-frequency traces. Note that the 3D pads for top and bottom chip are aligned on top of each other. The FR4 material is used for fabricating PCB, and the traces length/width are carefully designed to reduce coupling between the traces, improve signal integrity, and ensure 50 ohms matching impedance. Figure 5.10 shows the PCB layout design.

3D baseband transceiver using inverter based with resistive feedback driver and Dual RF band architecture has been fabricated on 130 nm CMOS technology. Figure 5.11 show waveforms of eye diagram at each receiver end. The results indicate good performance compared to 2D even though the chip is fabricated with 130nm CMOS processor. The architecture can recover 4.2 Gb/s from each RF bands and 5.8 Gb/s from BB consuming 36.4 mW results in high energy efficiency (2.6pJ/bit) and comparison with prior work summarized in Table I.
Figure 5-10 PCB layout design for 3D MBI

![PCB layout design for 3D MBI](image)

Figure 5-11 Measured Eye diagram of 3D (a) Baseband (b) 10GHz RF band (c) 30GHz RF band

![Measured Eye diagram of 3D](image)

### 5.6. Conclusion

The proposed 3D MBI is designed and fabricated in 130 nm CMOS technology and achieves aggregate I/O data rate of 14.4 Gb/s and consumes 36.2 mW. Baseband attains 5.8 Gb/s. On the other hand, the dual RF band achieves 8.4 Gb/s for 10 GHz and 30 GHz, respectively. Utilizing band selective transformer, LNA at the receiver of 30 GHz, and 3D TSV enable simultaneously transceiving three signals (BB+2RFs). The proposed architecture outperforms all works in Table 1 except [40] and [38] because their research has been implemented in a much newer process support less power supply consumption. This work would be adequate for future
3D stacked mobile memory interface due to higher energy efficiency and data rate throughput.

Table 5.1 Performance comparison of proposed 3D-MBI with prior works.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>130nm</td>
<td>40nm</td>
<td>65nm</td>
<td>45nm</td>
<td>65nm</td>
<td>65nm</td>
<td>130nm</td>
</tr>
<tr>
<td>Bands</td>
<td>BB</td>
<td>BB</td>
<td>BB+RF</td>
<td>BB+RF</td>
<td>BB+RF</td>
<td>BB(PAM)+RF</td>
<td>BB+2RF</td>
</tr>
<tr>
<td>Supply</td>
<td>1.2V</td>
<td>1.1V</td>
<td>1.0V</td>
<td>N/A</td>
<td>1.2V</td>
<td>1.2V</td>
<td>1.2V</td>
</tr>
<tr>
<td>Dimension</td>
<td>2D</td>
<td>2D</td>
<td>2D</td>
<td>2D</td>
<td>2D</td>
<td>2D</td>
<td>3D</td>
</tr>
<tr>
<td>T-Line/3D-Channel length</td>
<td>5cm</td>
<td>7cm</td>
<td>10cm</td>
<td>10cm</td>
<td>5cm</td>
<td>5cm</td>
<td>65µm</td>
</tr>
<tr>
<td>Data Rate</td>
<td>6.0Gb/s</td>
<td>4.3Gb/s</td>
<td>8.4Gb/s</td>
<td>15Gb/s</td>
<td>8Gb/s</td>
<td>13.4Gb/s</td>
<td>14.4 Gb/s</td>
</tr>
<tr>
<td>Active Area</td>
<td>0.3mm²</td>
<td>0.9mm²</td>
<td>0.14mm²</td>
<td>NA</td>
<td>0.12mm²</td>
<td>0.13mm²</td>
<td>0.20mm²</td>
</tr>
<tr>
<td>Total power</td>
<td>95mW</td>
<td>14.4mW</td>
<td>21mW</td>
<td>7.47mW</td>
<td>32.2mW</td>
<td>31mW</td>
<td>36.4 mW</td>
</tr>
<tr>
<td>Bit Error Rate</td>
<td>10⁻¹²</td>
<td>NA</td>
<td>&lt;10⁻¹²</td>
<td>NA</td>
<td>&lt;10⁻¹²</td>
<td>&lt;10⁻¹²</td>
<td>&lt;10⁻¹²</td>
</tr>
</tbody>
</table>

5.7. Future Works

Further advances in the system on chip (SoC) are demanded to support higher operation speed and energy efficiency. Thus, the MBI systems will be further improved by utilizing more bands. Figure 5.12 shows an extended version of the MBI system (EMBI) where we can transmit/receive four bands simultaneously to maximize data rate and energy efficiency. The EMBI utilize 3 RF bands (20, 40, and 60 GHz) along with baseband and has been designed and layout in 65nm CMOS process.

Figure 5-12 The EMBI system (a) TX (b) RX.
To enable concurrent transceiving for all bands, LNA for 40 GHz and 60 GHz have been implemented. The EMBI system is expected to obtain 28Gb/s and 43.6 mW for data rate and power consumption, respectively. The EMBI can accomplish much higher speed and energy efficiency compared to prior works.
CHAPTER 6: Ultra low power 2.4 GHz RF transmitter

6.1. Introduction

The content of this chapter has been published [97]. Advances in CMOS technologies enable the application of implantable medical devices (IMDs) [83]-[85]. Due to the ability of these devices to be placed on a human’s body and operate without a wired connection, many works have been conducted in academia and industry toward IMD [86]. Recent systems have implemented a radio frequency (RF) transmitter to enable wireless transfer of the neural information from and to outside of the body. Moreover, the energy has been provided using an inductive link without battery [83], [84]. Such a system supports more mobility freedom and less risk of infection for the human body [87], [88]. However, utilizing wireless solution places stringent prerequisite to relatively reduce power consumption so that the device can have more battery life and be energy efficient. Thus, new structures that are entirely integrated on-chip and have low power consumption are required to enable IMD implementation and further enhancement.

Figure 6.1 shows a typical wireless neural recording system comprises signal pre-conditioning, signal processing unit, and implementable RF transceiver. In the signal pre-conditioning unit, shown in Figure 6.1(a), spike detector captures the neural and biomedical data [84]. Next, analog to digital converter (ADC) samples and sends the data to a processing unit where the neural data is extracted and formed into a data packet for wireless transmission. Next, the data can be sent wirelessly outside the body. Then, in the external unit, shown in Figure 6.1(b), the wireless signal is extracted by an envelope detector to down-convert the modulated
signal to the baseband signal that initially transmitted. Thus, the baseband signal is processed by a computer application.

Figure 6-1 Diagram of typical IMD architecture (a) Implementable device (b) External unit (c) Proposed RFTX architecture

The power consumption of the telemetry transmitter that usually consumes more than 50% of the total power consumption is most critical in the biomedical system [83]. It also can be minimized quadratically as the supply voltage decreases due to $P=CV^2f$ where $P$, $C$, $V$, and $f$ are power consumption, capacitance, voltage supply, and operating frequency, respectively. Thus, reducing supply voltage and utilizing energy efficient components of the transmitter will significantly improve energy consumption. Moreover, the high transmitted data rate is essential to support medical images.

Prior works have been conducted to minimize supply voltage to reduce the power consumption by utilizing on/off keying (OOK) modulator, VCO and PA [90], [91]. Class-E PA has been used for the biomedical implant, but it typically suffers from efficiency due to the finite quality factor of passive components [90]. Recent work achieves 10 Mb/s data rate which is
insufficient for medical imaging [91]. To overcome the limitations above, we propose an energy
efficient and high-performance RF transmitter (RFTX) implementing LC-VCO, ASK modulator
and Class-D$^{-1}$PA as shown in Figure 6.1(c).

The rest of the chapter is organized as follows. Section II outlines the proposed 2.4 GHz
transmitter design. Simulation results are discussed in section III. Finally, section IV concludes
this work.

6.2. Proposed 2.4 GHz RF Transmitter

Since 2.4 GHz is dedicated to industrial, scientific and medical (ISM) band for
implantable medical Microsystems, the proposed RFTX utilizes 2.4 GHz carrier signal [3]. The
proposed transmitter comprises input buffer, LC-VCO, ASK modulator, and Class-D$^{-1}$PA,
shown in Figure 6.2, to achieve a high data rate, output power, and energy efficiency. It is also
entirely integrated on-chip. Because the input signal to the chip has deficiencies like fall/rise
time resulting in timing errors and an improper output signal, the input buffer is implemented. It
utilizes a self-biased differential amplifier to support high-frequency operation with lower delay
and minimize voltage offset, mismatch, and noise [18]. In electronic systems, VCO is
implemented, and its application ranges from clock generation in microprocessors to carrier
synthesis in mobile phones. LC-VCO operates at high frequency and achieves better phase noise
performance comparing to the ring oscillator because the LC resonant can attain high-quality
factor (Q) [89], [92]. Thus, the cross-coupled oscillator topology is applied to enable low power
consumption and maximum operating frequency and voltage swing.

Frequency shift keying (FSK) provides less data rate in comparison to quadrature phase
shift keying (QPSK). Thus, FSK has been preferably implemented in a wireless transceiver due
to the high power consumption and design complexity of the QPSK [93]. However, an FSK still is power hungry like any coherent scheme because it requires phase and frequency detector. Thus, non-coherent ASK modulator has been utilized in the transmitter as shown in Figure 6.2(a). Because the ASK modulation and demodulation circuits are relatively simple, cheap, and have less power consumption, it is a promising modulation/demodulation scheme for IMD. The modulator is used to shift the amplitude of carrier wave proportional to the amplitude of the modulating signal. In ASK modulation, the digital signal value is either 1 or 0 which indicates the presence or absence of the carrier signal, respectively. Thus, M1 and M2 are switching on/off 2.4 GHz RF carrier generated by VCO through M3 and M4. To improve the signal-to-noise ratio, M5 is implemented. Then, the modulated signal will be inductively coupled to an on-chip 4:1 transformer. The tunable capacitance $C1$ is implemented to resonate the 2nd coil of the transformer at 2.4 GHz and improve matching with the subsequent circuit.

An efficient PA allows the transmitter to provide high output power while operating at low supply voltage. To have higher efficiency PA, switching amplifier Class-D/E is typically used instead of the linear amplifier (e.g., Classes- A/AB/B). Class-D PA suffers from power

Figure 6-2 Proposed 2.4 GHz RF Transmitter schematic (a) ASK Modulator (b) Class-D$^{-1}$ PA
dissipation of $CV^2f$ because the drain capacitors are not included in the matching network [94]. Moreover, Class-D implements p-type metal oxide semiconductor (pMOS) which requires twice the size of nMOS to obtain similar on-resistance. Consequently, output driver design is more challenging. On the other hand, Class-E PA mitigates the problem of power dissipation of Class-D, but it demands multiple passive components. As a result, more silicon area and higher cost are required, and the PA efficiency is also degraded due to the finite quality factor of the passive components. Thus, Class-E and voltage mode Class-D PA are undesirable schemes for portable devices. Due to ease of integration, high output power, and efficiency, Class- D$^{-1}$PA, shown in Figure 6.2(b), is exploited [94]. Impedance matching transforming and differential to single ended conversion are performed by 2:1 on-chip transformer composed of parallel LC tank as detailed in [94]. Thus, higher Q of the transformer increases the efficiency. The conventional layout of 2:1 transformer has two downsides. First, to attain low dc series resistance, the width of the primary winding is required to be large which is typically limited by CMOS technology (e.g., 10-12µm). Second, the conventional layout suffers from proximity effect, so the current is not uniformly distributed on the surface of the inductor. To overcome these shortcoming, parallel-primary transformer layout is utilized in Class-D$^{-1}$PA. Thus, the dc resistance is reduced by a factor of 3 when the primary turn divided into three parallel winding. Moreover, the proximity effects minimized by having three parallel-primary winding intertwined symmetrically around the two series-connected secondary windings instead of side by side. The coupled ASK modulated signal goes through M6 and M7, and the PA output impedance is enhanced by using cascade devices (e.g., M8 and M9). To resonate the primary and secondary windings at 2.4 GHz and obtain matching, tunable capacitances $C2$ and $C3$ are employed.
6.3. Simulated Results

The proposed transmitter is designed in a 0.13µm CMOS technology. The post-layout simulation is performed considering parasitic capacitance and resistance. The proposed RFTX is simulated with 0.6V supply and consumes 1.98mW at 130Mb/s data rate. Figure 6.3 shows the simulated input and output waveforms of the RFTX. Figure 6.3(a) shows the VCO output waveform with a voltage swing of 1.2V. As shown in Figure 6.3(b), the input signal is randomly generated from input buffer at 130Mb/s to be modulated by the ASK modulator. Figure 6.3(c) and (d) illustrate that the proposed transmitter obtains ASK modulated and RFTX outputs with a peak-to-peak voltage of 500mV and 175mV, respectively. Thus, the output power is -11.14dBm which is equivalent to 3.88% efficiency (e.g., TX Efficiency=TX output power/power).
Figure 6-3 Simulated 2.4 GHz RFTX (a) LC-VCO swing (b) Input signal (c) ASK modulated signal (d) Class-D−1PA/RFTX output waveform

Figure 6-4 Chip layout for proposed 2.4 GHz RFTX

Table 6.1 Performance comparison of proposed 2.4 GHz RF transmitter with Prior Works

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>0.5 µm</td>
<td>0.13µm</td>
<td>0.13µm</td>
<td>0.18µm</td>
<td>0.13µm</td>
</tr>
<tr>
<td>Frequency (GHz)</td>
<td>0.433</td>
<td>0.403</td>
<td>2.4</td>
<td>2.36</td>
<td>2.4</td>
</tr>
<tr>
<td>Supply Voltage (V)</td>
<td>3.3</td>
<td>0.7</td>
<td>0.5</td>
<td>0.9</td>
<td>0.6</td>
</tr>
<tr>
<td>Power (mW)</td>
<td>6.75</td>
<td>0.78</td>
<td>1.65</td>
<td>1.712</td>
<td>1.98</td>
</tr>
<tr>
<td>TX output power (dBm)</td>
<td>N/A</td>
<td>N/A</td>
<td>-14</td>
<td>-10.63</td>
<td>-11.14</td>
</tr>
<tr>
<td>Data Rate (Mb/s)</td>
<td>0.33</td>
<td>40</td>
<td>100</td>
<td>10</td>
<td>130</td>
</tr>
<tr>
<td>Energy Per Bit (pJ/b)</td>
<td>20450</td>
<td>19.5</td>
<td>16.5</td>
<td>171.2</td>
<td>15.23</td>
</tr>
<tr>
<td>TX output power/power (%)</td>
<td>N/A</td>
<td>N/A</td>
<td>2.41</td>
<td>5.05</td>
<td>3.88</td>
</tr>
</tbody>
</table>
Figure 6.4 shows the layout of the RFTX where occupies 980µm x 980µm. The performance comparison with prior works is summarized in Table I. It indicates that the proposed RFTX has the best data rate and energy per bit of 130Mb/s and 15.23pJ/bit, respectively.
7. BIBLIOGRAPHY


[46] H. W. Lim et al., "A 5.8-Gb/s Adaptive Integrating Duobinary DFE Receiver for Multi-


[54] Z. Zeng, X. Ye, Z. Feng, and P. Li, “Tradeoff analysis and optimization of power delivery


[70] H. He and J. J.-Q. Lu, “Modeling and analysis of PDN impedance and switching noise in


