Analog-Digital Co-existence in 3D-IC

by

Gilad Yahalom

B.Sc., Technion - Israel Institute of Technology (2008)
S.M., Massachusetts Institute of Technology (2012)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2016

© Massachusetts Institute of Technology 2016. All rights reserved.

Author . . . .
Department of Electrical Engineering and Computer Science
January 29, 2016

Certified by . . . .
Anantha P. Chandrakasan
Vannevar Bush Professor of Electrical Engineering
Thesis Supervisor

Accepted by . . . .
Leslie A. Kolodziejski
Chair, Department Committee on Graduate Theses

Analog-Digital Co-existence in 3D-IC

by

Gilad Yahalom

Submitted to the Department of Electrical Engineering and Computer Science
on January 29, 2016, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering

Abstract

Ubiquitous mobile communication creates an increasing demand for high data rates, complex modulation schemes and low power design. The cost and performance benefits of conventional lithographic scaling are diminishing as process cost increases exponentially. 3D integration has the potential to keep driving performance forward while keeping cost down. The possibility to integrate separate dies with low-parasitic, dense interconnect and shorter routing provides area and power benefits. However, new challenges must be addressed in order to enable design in this new dimension and provide system level improvements. This thesis explores the impact, challenges and advantages of using 3D integration for combining digital and analog circuits for RF applications.

The use of a vertical solenoid inductor in a Voltage Controlled Oscillator (VCO) is proposed. The inductor design utilizes the through-silicon-vias of the 3D stack as part of its geometry. The solenoid inductor exhibits a 28% larger inductance and a 6 dB higher quality factor compared to a conventional planar inductor occupying the same area. The VCO circuit phase noise is improved by 6 dB and exhibits an improved immunity to coupling from adjacent digital clock lines routed on the bottom tier of the 3D stack.

An efficient hardware implementation is presented for an LTE uplink channel. The proposed design processes input data for cellular transmission. The core of the computation includes a variable-length, high-order, mixed-radix FFT and IFFT blocks. The use of energy efficient circuits and algorithms enables achieving an energy efficiency of up to 95 pJ/Sample and additional power savings of up to 24% for different operation modes.

Both designs are combined along with digital-to-analog conversion to create a partial cellular transmitter in 3D-IC. Highly flexible and configurable design allows for various partitioning of the system. The 3D design has a digital link energy efficiency of up to 0.37 pJ/bit, compared to the 33.3 pJ/bit consumed in a multiple die partitioning and 0.83 pJ/bit for a 2.5D interposer emulated design. The use of the solenoid VCO along with digital-analog partitioning between the die tiers enables high immunity to noise and reduction of spurs at the VCO output.

Thesis Supervisor: Anantha P. Chandrakasan
Title: Vannevar Bush Professor of Electrical Engineering

Acknowledgments

During my time at MIT I had quite a bit of unique and wonderful experiences. Studying under world-renowned professors, hearing lectures from great thinkers, visionaries and leaders, witnessing some quirky and impressive hacks. But the best feature of MIT is its amazing, spectacular collection of individuals. People from all over the world, with a great passion for knowledge, learning and collaboration. People who I could learn so much from and share so much with. I am so grateful that I had the privilege to meet these people and to be part of this vast energetic vibrant MIT community.

First and foremost, I would like to thank and express my deepest appreciation and gratitude to my research advisor Professor Anantha Chandrakasan. His continual guidance and support were invaluable in my journey. Anantha, despite his extremely busy schedule and myriad of responsibilities, always found the time to discuss my research and is always sure to provide a valuable piece of advice. Whenever I needed, Anantha was there with an interesting direction to explore and encouragement on the best way to frame and present my work. I have learned a lot from Anantha's mentorship and guidance. The wonderful, collaborative atmosphere among the students in Anantha's lab is a true testament to his leadership and character. Thank you Anantha for having faith in me, supporting my research and enabling me to explore new and uncharted territories.

I would like to thank my committee members, Prof. Hae-Seung (Harry) Lee and Prof. Ruonan Han for their advice, feedback and support. I would also like to thank Prof. Dina Katabi and Prof. Luca Daniel for fruitful discussions and advice which helped shape the direction of my thesis work.

This work could not have been done without the tremendous help, cooperation and generosity of MediaTek Inc. and especially Dr. Alice Wang and Dr. Uming Ko. I would like to thank them profusely for helping define, mold and finance this research and their dedication and commitment. I would also like to thank Hugh Meir, Chang Huang and Maria Lawinson of MediaTek's Austin TX offices for help in the design and fabrication of the 3D solenoid inductor test chip. Further thanks go to Stacy Ho, CC Hsiao, Susheel Bhalabhadra, Zoran Zvoran and Lubna Ikram at the offices in Woburn MA. Their continuous help and

support in providing feedback and reference designs made it possible to create such a complex and thorough test chip of the 3D LTE transmitter chain.

A special thanks goes out to all the students, postdocs and other members of the extended Anantha group lab. Thanks to Arun Paidimarri, Phillip Nadeau and Nachiket Desai for countless useful discussions and debugging sessions. Many thanks to Dr. Nathan Ickes for help in design and debug of PCBs and test setups. My deep gratitude goes out to Mehul Tikekar for helping me in promoting and creating the new Anantha group Wiki and repository, and for being my on-call Unix and Git guru. Many thanks also go out to Avishek, Bonnie, Chiraag, Chu, Dina, Frank, Georgios, Ishwarya, Michael, Mohamed, Omid, Preet, Priyanka, Rahul, Sirma, Sungjae, Utsav and Yildiz for being such a great and supportive group, making it a delight to come every day to the lab. A special warm thank you also goes out to Margaret Flaherty, our group's administrator, which skillfully coordinates and manages all the workings of the lab and makes sure everything is operating smoothly and efficiently.

Many other great friends and colleagues also deserve recognition. Radhika, Zhipeng, Yan and SungWon, as well as other students in Prof. Weinstein, Sodini and Lee groups. Sunghyun Park was a close companion in the design of the first 3D test chip and helped on many of the initial trouble-shooting of the layout and fabrication at MediaTek.

This work was carried out both at MediaTek as well as the Microsystems Technology Laboratory at MIT, and I would like to thank Michael Mcllarath for assisting with CAD tool support and setup, Mike Hobbs for assistance with IT and computing resources and Deborah Hodges-Pabon for making MTL such a warm and inviting home to all students, faculty and staff. On the administrative side I am most grateful to Janet Fischer, Alicia Duarte and the rest of the EECS HQ members for being so helpful in assisting in all administrative and pedagogical matters.

Aside from the vast amount of people, which only few of them were mentioned above, and helped further my academic research, others here at MIT helped balance out my day and maintain a semblance of normality. Thank you Sarit and Yair, Noa and Jonathan, Inbal and Roy for being fantastic friends and companions. Thanks for endless hours of card games and countless Friday night dinners. Thanks for hikes, road trips, movie nights and

the occasional award ceremony viewing parties.

Although physically far away, my parents' support helped fuel this long and arduous journey. I know that without the education and values instilled from an early age I would never have embarked on such a journey to begin with. Thank you Mom and Dad for always being there for me.

Lastly, but definitely not least, my biggest thanks, devotion and appreciation go out to my one true friend and soul-mate. My courageous, patient, loving and thoughtful wife Emanuel. All my accomplishments are made possible by her. She makes everything clear and focused and helps put everything in to the right perspective. I would be lost in a vast sea of incomprehensibility without you. Thank you for being by my side always. I love you with all my heart.

Contents

List of Figures **13**

List of Tables **19**

List of Acronyms **21**

1 Introduction **27**

1.1 Motivation 27

1.2 Three Dimensional Integration 28

1.3 Previous work on 3D-IC 33

1.4 Thesis Contributions and Outline 35

2 Passive Structures **41**

2.1 Introduction 41

2.2 Analog-Digital Integration 43

2.2.1 Noise Contributors 43

2.2.2 Isolation Techniques 44

2.2.3 Isolation in 3D-IC 46

2.3 Test Chip Design 47

2.3.1 3D Stacking 47

2.3.2 Inductor Design 48

2.3.3 VCO Design 51

2.3.4 Clock Lines 53

2.3.5 Coupling Analysis 56

2.3.6 Coupling Transfer Function Calculation 59
2.4 Measurements 61
2.5 Conclusion 69

3 Data-Dependent Signal Processing 71
3.1 Introduction 71
3.2 LTE 72
3.3 FFT Overview 78
3.3.1 FFT Algorithm 79
3.3.2 General Radix FFT 81
3.3.3 Bit Reversal 87
3.3.4 Hardware Implementation 88
3.4 Digital Baseband Implementation 97
3.4.1 Hardware Efficient Processing 100
3.4.2 QAM Mapping 102
3.4.3 DFT 104
3.4.3.1 Butterfly Structure 105
3.4.3.2 Delay Lines 107
3.4.3.3 Memory Sharing 109
3.4.3.4 CORDIC Rotation 112
3.4.3.5 Angle Generation 116
3.4.3.6 Index Counter 117
3.4.4 Resource Mapping 119
3.4.4.1 Interpolation 120
3.4.4.2 Transform Decomposition 124
3.4.5 IDFT 126
3.4.6 Cyclic Prefix Addition 127
3.4.7 Upsampling 128
3.4.8 Snapshot Memory 137
3.5 Measurements 138

3.5.1 Data Path 139
3.5.2 Power 144
3.5.3 Comparison 149
3.6 Conclusion 150

4 Analog-Digital 3D Integration 153
4.1 Introduction 153
4.2 Transmitter Chain Design 154
4.2.1 Chip Layout and Partitioning 155
4.2.2 Digital Baseband 164
4.2.3 Digital Buffering and Level Shifting 165
4.2.4 Clock Distribution 167
4.2.5 Digital Routing 171
4.2.6 Analog Routing 173
4.2.7 Digital to Analog Conversion 176
4.2.8 LO Generation 178
4.2.9 Upconversion 180
4.2.10 Bias 182
4.2.11 External I/O 183
4.3 Measurements 186
4.3.1 Digital Baseband 187
4.3.2 DAC 187
4.3.3 VCO 191
4.3.4 Mixer 192
4.3.5 System Partitioning 194
4.4 Conclusions 198

5 Conclusion and Future Directions 201
5.1 Summary of Contributions 202
5.1.1 Passive Devices in 3D-IC 202
5.1.2 Data-Dependent Signal Processing 203

5.1.3 Analog-Digital 3D Integration 204
5.2 Future Directions 206
5.2.1 Full Transceiver Design 206
5.2.2 Power Generation in 3D-IC 208
5.2.3 Heterogeneous Integration 209

A QAM Mapping **211**

B Configuration **215**
B.1 Scan Chain Architecture 215
B.2 3D Scan Chain 216
B.3 Configuration Options 219

Bibliography **231**

List of Figures

1-1 Peak data rates for various wireless communication standards 28

1-2 (a) Gate count and wafer cost per technology node leading to trend change in (b) cost per gate 29

1-3 Illustration of different stacking approaches (not to scale) (a) Interposer stacking (2.5D), (b) 3D stack using F2F and (c) 3D stack using B2F assembly 32

2-1 Illustration of Back-to-Face 3D die stack (not to scale) 47

2-2 3D-IC stack illustration 48

2-3 Monolithic integrated inductor equivalent circuit model 50

2-4 Simulated inductor parameters extracted from field solver simulation results 51

2-5 CMOS VCO core schematic used with both the planar and solenoid inductor structures 52

2-6 Frequency tuning of VCO achieved via (a) continuous MOS varactor and (b) switched capacitor bank unit cell 53

2-7 Differential capacitive divider source follower output buffer schematic . . . 54

2-8 Clock input buffer (a) differential amplifier and (b) pseudo differential buffer chain schematics 55

2-9 Clock divide by 16 circuit schematic 55

2-10 Controllable clock lines location relative to inductors' dimension and schematic of line switch control 56

2-11 Simulated mutual inductance between clock lines and inductor structures as a function of normalized clock line offset location 57

2-12 Simplified coupling model between clock lines and VCO LC tank 58

2-13 Die micrograph (a) top tier (b) bottom tier 61
2-14 PCB used for testing and measurements 62
2-15 Measured tuning range for planar and solenoid VCOs 63
2-16 Measured phase noise of planar and solenoid VCOs 64
2-17 VCO output power spectral density with spurious tones caused by coupling from digital clock lines 65
2-18 Clock signal spur power as a function of relative frequency to VCO free-running frequency 65
2-19 Clock signal spur power as a function of clock line location 66
2-20 Clock signal spur power as a function of coupling strength 67
2-21 Clock signal spur power vs. aggregate coupling strength 68
2-22 Phase noise measurement with noise coupling from a 1 MHz clock signal . 69
2-23 Output power spectral density of VCO with close-in clock line noise. Planar inductor VCO exhibits injection locking to the clock signal 70

3-1 LTE FDD frame structure 73
3-2 Time domain slot structure with (a) normal and (b) extended cyclic prefix addition to SC-FDMA symbols 74
3-3 LTE uplink (a) resource grid and (b) channel configuration illustration . . . 75
3-4 LTE PHY layer 77
3-5 (a) Localized and (b) Interleaved SC-FDMA subcarrier mapping 79
3-6 Radix-2 decimation in frequency FFT 82
3-7 Radix-2 butterfly 83
3-8 Radix-3 butterfly structure 84
3-9 Radix-5 butterfly structure 85
3-10 Radix-4 butterfly structure 86
3-11 FFT memory architecture block diagram 90
3-12 FFT pipeline architecture block diagram 90
3-13 R2MDC block diagram for 16 point FFT 91
3-14 R2SDF block diagram for 16 point FFT 91

3-15 Radix 2 type I butterfly 92
3-16 R4SDF block diagram for 16 point FFT 92
3-17 R4MDC block diagram for 16 point FFT 93
3-18 R4SDC block diagram for 16 point FFT 93
3-19 Radix 2 (a) type II butterfly and (b) 90° rotation block diagrams 94
3-20 $R2^2$SDF block diagram for 16 point FFT 94
3-21 Radix 2 (a) type III butterfly and (b) 45° rotation block diagrams 95
3-22 $R2^3$SDF block diagram for 64 point FFT 95
3-23 Baseband block simplified architecture 97
3-24 Double buffer handshake process (a) simplified schematic and (b) timing diagram 99
3-25 Constellation mapping block diagram 104
3-26 Variable size UL LTE DFT pipeline topology 106
3-27 Control signal timing for (a) radix-2, (b) radix-3 and (c) radix-5 butterflies with a delay of T_{delay} clock cycles 106
3-28 SRAM based delay line schematic 108
3-29 Latch based memory schematic 109
3-30 CORDIC block schematic 115
3-31 CORDIC schematic of (a) quadrant rotator and (b) micro-rotator block . . . 116
3-32 Angle generation for CORDIC following (a) general radix butterfly and (b) simplified architecture for radix 2^3 butterfly block 117
3-33 Mixed radix reverse digit counter for DFT output indexing 118
3-34 Simplified multiplier blocks for (a) 2 bit and (b) 3 bit multipliers 119
3-35 Block diagram of filter re-sampling by a factor of N/M 122
3-36 Reconstructed samples from resampling filter with length (a) L = 12801 and (b) L = 128001 123
3-37 EVM of reconstructed samples as a function of resampling filter length. 10 sets of 1200 random 64-QAM points using a 2048 point IDFT 123
3-38 Transform decomposition procedure for an FFT calculation with a subset of the inputs 125

3-39 Fraction of operations required in TD-FFT as a function of non-zero sample ratio 126
3-40 IDFT block architecture 127
3-41 Zero order hold impulse response in the (a) time and (b) frequency domain . 129
3-42 Power spectrum of 10 MHz channel LTE signal in (a) baseband) and (b) at RF around a 2 GHz carrier, with ZOH D/A with no oversampling 130
3-43 Power spectrum of 10 MHz channel LTE signal in (a) baseband) and (b) at RF around a 2 GHz carrier, with ZOH D/A with x8 oversampling and interpolation 132
3-44 Integer rate conversion upsampling system 133
3-45 Low pass interpolation FIR filter (a) time and (b) frequency response 133
3-46 N tap FIR segment 134
3-47 Expander filter Noble identity 135
3-48 Polyphase decomposition for an interpolating filter using the Noble identity 136
3-49 Polyphase implementation of the x8 interpolation filter 137
3-50 Micrograph of (a) chip and (b) zoom-in of implemented LTE digital baseband chip 138
3-51 Measured 480 64-QAM modulated input symbols 139
3-52 FFT output symbol index 140
3-53 FFT output after zero padding (a) w/o and (b) w/ the use of transform decomposition 141
3-54 IFFT output symbol index 142
3-55 IFFT output power spectral density 142
3-56 Interpolator output power spectral density 143
3-57 Internal trigger signals (a) w/o and (b) w/ the use of transform decomposition 144
3-58 FFT block power consumption (a) all FFT sizes and (b) separated into homogeneous radix butterfly blocks 146
3-59 Relative total power when using transform decomposition as a function of FFT size (i.e. number of non-zero samples at input) 147
3-60 Energy per sample for each LTE bandwidth operating mode 148

4-1 Partial transmitter chain block diagram 154
4-2 Schematic illustration of 3D-IC packaging process (dimensions not to scale) 156
4-3 Block diagram of single cube tier in 3D stack 157
4-4 Single die, SoC-like operation emulation (only top tier of cube shown) . . . 158
4-5 Two dies, partition at (a) digital boundary and (b) analog boundary (only top tier of cubes shown) 160
4-6 2.5D operation emulation with partitioning at (a) digital boundary and (b) analog boundary 161
4-7 3D-IC operation with partitioning at (a) digital boundary and (b) analog boundary 162
4-8 Schematic of (a) tri-state buffer and (b) scalable controlled buffer 166
4-9 High-VDD to low-VDD level shifter schematic using thick-oxide inverter cell 166
4-10 Low-VDD to high-VDD level shifter schematic 167
4-11 Clock input differential amplifier schematic 168
4-12 Clock divider schematic 168
4-13 Clock distribution for (a) single and (b) two cube operation 170
4-14 Clock generation and distribution block diagram for two tiers in 3D-IC . . . 171
4-15 Digital crossbar (a) Tx column and (b) Rx column 172
4-16 Digital signal multiplexing block diagram 173
4-17 Analog crossbar schematic 174
4-18 Voltage-to-Current conversion and signal multiplexing schematic 175
4-19 Analog signal path multiplexing block diagram 176
4-20 General schematic of binary-unary current steering DAC segmentation topology 177
4-21 DAC data and clock synchronization schematic 177
4-22 First-order, adjustable low-pass filter schematic 178
4-23 Quadrature VCO core schematic 179
4-24 VCO differential buffer driver schematic 180
4-25 Simplified RF mixing DAC schematic 181
4-26 Partitioning of mixing DAC operation in 3D-IC 182

4-27 High-speed I/O (a) serializer and (b) deserializer schematics 185
4-28 Micrograph of test chip (same layout used for both top and bottom tiers) . . 187
4-29 DAC voltage output for 11 bit codeword input 188
4-30 Calculated DAC (a) DNL and (b) INL 189
4-31 DAC output spectrum for 5 MHz BW LTE signal 190
4-32 VCO frequency tuning range 191
4-33 Measured VCO phase noise for Planar and Solenoid inductor structures . . 192
4-34 Example of time domain output of solenoid quadrature VCO 193
4-35 Phase difference between VCO quadrature outputs 193
4-36 External mixer output spectrum of 5 MHz BW LTE signal 195
4-37 Digital link power consumption comparison 196
4-38 VCO output spectrum using different partitioning schemes 198

B-1 Scan chain unit cell for (a) writing configuration bits and (b) reading chip data 215
B-2 Scan chain topology for 3D test chip 217
B-3 Tie-high cell schematic 218

List of Tables

1.1 Summary of Select Previous Work on 3D-IC 39

2.1 Summary and comparison of results 64

3.1 Transmitter available resource blocks for different channel bandwidths . . . 74

3.2 Radix-2 FFT output bit reversal 87

3.3 Mixed radix digit reversal example 89

3.4 Pipeline FFT architecture resource comparison 96

3.5 Valid DFT sizes for LTE UL symbol generation 105

3.6 FFT butterfly memory requirements 111

3.7 FFT memory bank size and grouping 113

3.8 Design Specification Summary 139

3.9 Power and energy consumption summary 149

3.10 Comparison to other LTE OFDMA signal generation processors 151

4.1 System partitioning summary 163

4.2 DAC specification summary 188

4.3 Summary of VCO performance 194

A.1 BPSK modulation mapping 211

A.2 QPSK modulation mapping 212

A.3 16-QAM modulation mapping 212

A.4 64-QAM modulation mapping 213

B.1 Scan chain block order 218

B.2 General configuration bits 219
B.3 Baseband module configuration bits 221
B.4 DAC module configuration bits 226
B.5 VCO module configuration bits 227
B.6 Mixer module configuration bits 227
B.7 Bias module configuration bits 228
B.8 Transmitter test chip I/O pads 229

List of Acronyms

3D-IC Three Dimensional Integrated Circuits

3GPP 3rd Generation Partnership Project

AC Alternating Current

ACLR Adjacent Channel Leakage Ratio

ADC Analog to Digital Converter

AMPS Advanced Mobile Phone System

ASIC Application Specific IC

B2B Back-to-Back

B2F Back-to-Face

BEOL Back End Of Line

BPSK Binary Phase Shift Keying

BS Base Station

BW Bandwidth

C4 Controlled Collapse Chip Connection

CAD Computer Aided Design

CMOS Complimentary Metal-Oxide-Semiconductor

CORDIC COordinate Rotation Digital Computer

CP Cyclic Prefix

CPU Central Processing Unit

CSD Canonical Signed Digit

DAC Digital to Analog Converter

DC Direct Current

DDR Double Data Rate

DFT Discrete Fourier Transform

DL Downlink

DMIPS Dhrystone Million Instructions Per Second

DRAM Dynamic RAM

DNL Differential Non-Linearity

DSP Digital Signal Processing

DVFS Dynamic Voltage-Frequency Scaling

EDGE Enhanced Data Rates for GSM Evolution

EM Electro-magnetic

ESD Electrostatic Discharge

E-UTRA Evolved Universal Terrestrial Radio Access

EVM Error Vector Magnitude

F2F Face-to-Face

FDD Frequency Division Duplexing

FDMA Frequency Division Multiple Access

FEC Forward Error Correction

FEOL Front End Of Line

FF Flip Flop

FFT Fast Fourier Transform

FIR Finite Impulse Response

FOM Figure of Merit

FPGA Field Programmable Gate Array

GPIO General Purpose I/O

GPS Global Positioning System

GSM Global System for Mobile Communications

HSPA High Speed Packet Access

IC Integrated Circuit

ICI Inter-Carrier Interference

IDFT Inverse Discrete Fourier Transform

IF Intermediate Frequency

IFDMA Interleaved FDMA

IFFT Inverse Fast Fourier Transform

IIR Infinite Impulse Response

INL Integral Non-Linearity

I/O Input/Output

IoT Internet of Things

ISI Inter-Symbol Interference

ISM Industrial, Scientific and Medical

KVL Kirchhoff's Voltage Law

LDO Low Dropout Regulator

LFDMA Localized FDMA

LNA Low Noise Amplifier

LO Local Oscillator

LPF Low Pass Filter

LSB Least Significant Bit

LTE Long Term Evolution

LTE-A LTE Advanced

MCM Multi-Chip Module

MEMS Micro-Electro-Mechanical System

MIMO Multiple Input Multiple Output

MOM Metal-Oxide-Metal

MOS Metal-Oxide-Semiconductor

MSB Most Significant Bit

MU-MIMO Multi-User MIMO

NMOS N-type MOS

NoC Network on Chip

OFDM Orthogonal Frequency Division Multiplex

OFDMA Orthogonal Frequency Division Multiple Access

OpAmp Operational Amplifier

PA Power Amplifier

PAPR Peak to Average Power Ratio

PC Personal Computer

PCB Printed Circuit Board

PE Processing Element

PHY Physical Layer

PLL Phase Lock Loop

PMOS P-type MOS

PUSCH Physical Uplink Shared Channel

QAM Quadrature Amplitude Modulation

QFN Quad Flat No-leads

QPSK Quadrature Phase Shift Keying

R2MDC Radix-2 Multi-path Delay Commutator

R2SDF Radix-2 Single-path Delay Feedback

R2^{2}SDF Radix-2^2 Single-path Delay Feedback

R2^{3}SDF Radix-2^3 Single-path Delay Feedback

R4MDC Radix-4 Multi-path Delay Commutator

R4SDC Radix-4 Single-path Delay Commutator

R4SDF Radix-4 Single-path Delay Feedback

RAM Random Access Memory

RB Resource Block

RF Radio Frequency

SC Subcarrier

SDF Serial Delay Feedback

SC-FDMA Single Carrier Frequency Division Multiple Access

SerDes Serializer/Deserializer

SiP System in Package

SoC System on Chip

SRAM Static Random Access Memory

TCAM Ternary Content-Addressable Memory

TD Transform Decomposition

TDD Time Division Duplexing

TSV Through Silicon Via

UE User Equipment

UL Uplink

UMTS Universal Mobile Telecommunication System

VCO Voltage Controlled Oscillator

ZOH Zero Order Hold

Chapter 1

Introduction

1.1 Motivation

The development of integrated circuits has allowed for the proliferation of powerful personal computing and communication solutions. The prediction (or observation) set forth by Gordon Moore that the transistor density will double every 18 months [1] proved to be both quite accurate and a major force in driving this revolution. Cellular and wireless communication today is ubiquitous and albeit its relative short history it's hard to imagine our lives today without being continuously connected to one another.

Along with the advancement in lithographic scaling of integrated circuits, the demand for increased data rate and lower latency increased tremendously as well. From humble beginnings of the analog cellular network at 1984 in the form of the Advanced Mobile Phone System (AMPS), where peak data rates were a mere 14.4 Kb/s [2], the evolution of cellular and wireless communication took on exponential form, much like the growth in semiconductor technology. Fig. 1-1 plots the peak Downlink (DL) data rates for various cellular and wireless communication protocols throughout history. We see that today, current standards such as LTE Advanced (LTE-A) and WiFI 802.11ad call for data rates in the range of a few Gb/s.

However, evidence from recent years suggest that the trends of semiconductor scaling can no longer keep up with such a demanding scaling rate. Most importantly, it may not be economically beneficial to do so even if the technological obstacles can be overcome. Fig.

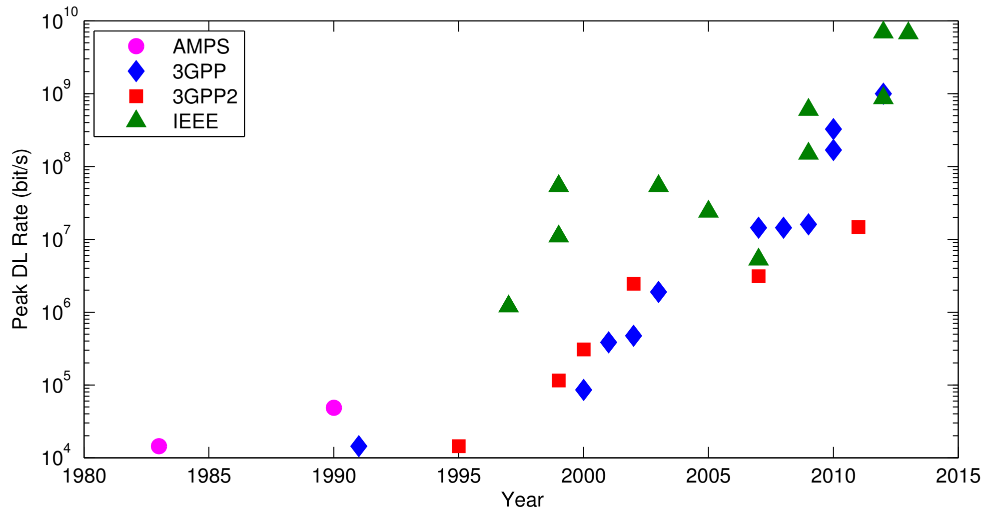
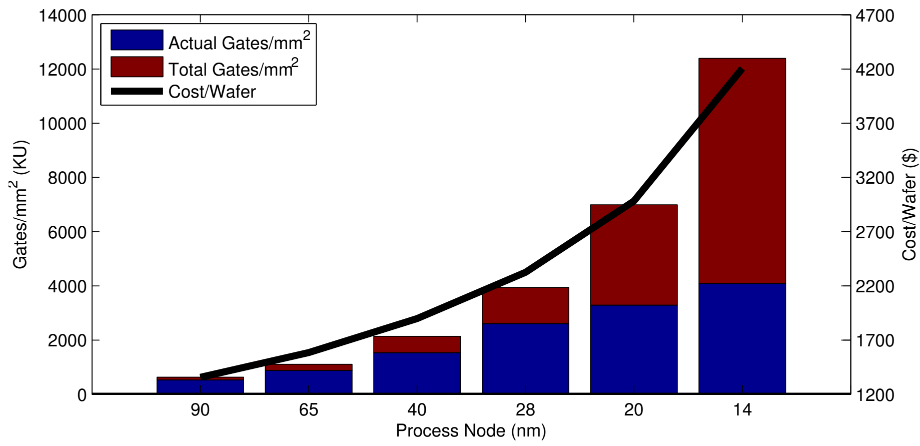


Figure 1-1: Peak data rates for various wireless communication standards. Reproduced using data from [3]

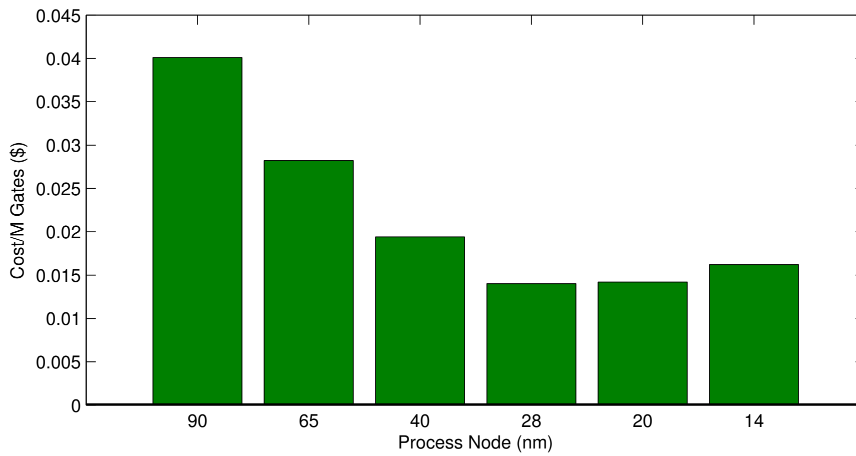
1-2a plots the cost per wafer for different technology nodes along with the amount of usable transistors per wafer due to reduced yield and effective usable area. While the amount of fabricated gates rises as the feature dimension becomes smaller, issues such as defect density, leakage, doping uniformity, line edge roughness and other physical parameters which are sensitive to minute variation reduce the amount of actual usable gates in the fabrication process. The cost per wafer continues to rise sharply, but the amount of usable devices does not scale as fast. As a result, as can be seen in Fig. 1-2b, the price per transistor is not scaling as it used to in the past. And according to some estimations might even be reversing the trend, making it impractical to continue the strive for further scaling of the process node below 22nm [4].

1.2 Three Dimensional Integration

An alternative to further process scaling is functional diversification. By adding higher degrees of functionality and achieving more diverse integrated systems we can overcome this bottleneck and continue pushing the performance envelope further out. One of the key areas which can enable this approach is 3D packaging. Three Dimensional Integrated Circuits (3D-IC) are an emerging solution to achieve higher functionality diversification in



(a)



(b)

Figure 1-2: (a) Gate count and wafer cost per technology node leading to trend change in (b) cost per gate. Reproduced from data by [4]

integrated circuits [5].

3D-IC is a relatively new technique, and as such it does not refer to any one specific type of processing, but rather to a broader set of packaging and post-processing techniques used [6, 7]. In general, it is customary to distinguish between package level technologies and technologies which include the use of Through Silicon Vias (TSVs). In package level technologies, which combine several dies in a single package (also known as System in Package (SiP) or Multi-Chip Module (MCM)), the connections are made at the package level interconnect, such as wirebonds and flip-chip bumps (also known as Controlled Collapse Chip Connection (C4)), resulting in connections on the order of a few hundreds of micrometers. Use of TSV technology allows connecting of several dies to a common passive silicon interposer substrate (usually referred to as "2.5D integration") or vertically stacking several dies. This interconnect level may be below ten micrometers in size resulting in highly dense, low-parasitic interconnect. As the technology matures, we will expect to see the TSV diameter and pitch shrink to the size of a mere few micrometers [8].

Even within the vertical "True 3D" integration process family one may find a myriad of different approaches and variants. In one approach, the dies are fabricated separately, the wafer of one tier is thinned before inserting the TSVs, and only then are the tiers stacked. In another approach, the tiers are grown one on top of each other, creating a new active silicon layer on top of a formed wafer. This allows for denser interconnect and avoids the processing and handling of the thinned wafer. The former approach may be viewed as a parallel manufacturing process resulting in a polylithic structure, while the latter is a sequential process, resulting in a monolithic structure.

Further differences are found at the direction of the die stacking. Whether two adjacent dies will be stacked Face-to-Face (F2F), through micro-bumps, where the top metal layer of each die is connected to the other, and the signals are routed outside of the die stack using TSVs. Or alternatively Back-to-Back (B2B) connection, through the back side silicon substrate of each die or alternatively connected Back-to-Face (B2F). Each connection style entails different constraints over connection size, pitch, density, impact on available active die area as well as the scalability of the connection [9].

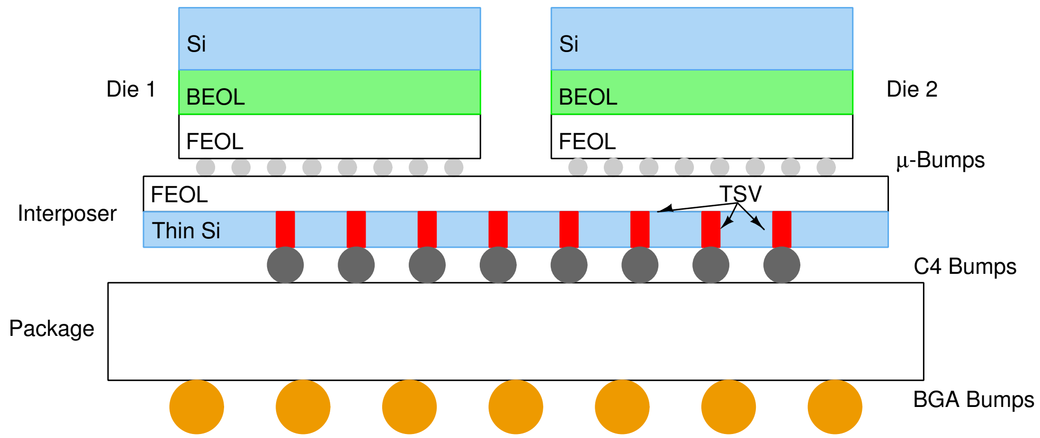
Different approaches also include the stacking process, whether it is wafer-to-wafer,

die-to-wafer or die-to-die. Again, each choice results in a different trade-off in the manufacturing process which affects yield, testability and flexibility. The TSVs themselves may be added at different points in the fabrication process, dubbed via-first, via-middle and via-last. This pertains to the step in the fabrication process when the TSVs are inserted. Via-first adds the TSVs before the Front End Of Line (FEOL) step and the addition of active devices, via-middle adds the TSVs between the FEOL and Back End Of Line (BEOL) steps. This is the most common approach since it usually gives the best trade-off in terms of cost, yield and manufacturing complexity. The via-last option adds the TSV structures after the BEOL step has completed and all metal routing layers were formed on top of the active die area. Figure 1-3 illustrates schematically a few of the aforementioned techniques and approaches to 3D-IC integration. The figure only illustrates a few representative scenarios for illustration purposes, and only depicts two tier die stacking. The above mentioned 3D stacking techniques are also utilized to stack a larger number of die tiers (as much as 32 in some cases [10]).

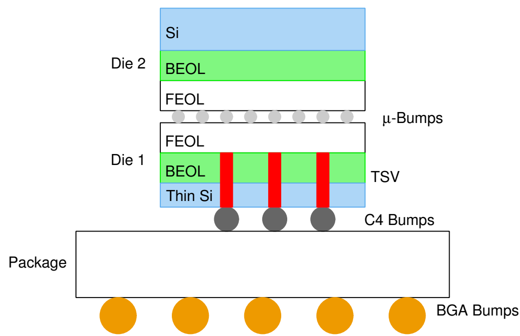
Apart from the up-front reduction in foot-print area of the 3D stack compared to conventional planar 2D dies, 3D-IC offers a great advantage in reduced interconnect length which may dominate the chip's power consumption [11, 12], limit its clock frequency [13] and incur area overhead [14].

Three dimensional integration further opens up the possibility for closely integrating units specialized for different functionalities. This leads to heterogeneous integration, either in the sense of different functionality, different process nodes or completely different materials. Such examples include integrating logic with memory [15–17], optics [18–20], Micro-Electro-Mechanical System (MEMS) [21–23], power [24, 25], as well as III-V materials with Si [26, 27] and so forth. The close vertical integration of active silicon layers gives the designer a new degree of freedom in the system and block level partitioning [28, 29] which enables new architectures for both traditional logic blocks [30], and more complex systems such as Network on Chip (NoC) [31, 32] and Field Programmable Gate Arrays (FPGAs) [33, 34].

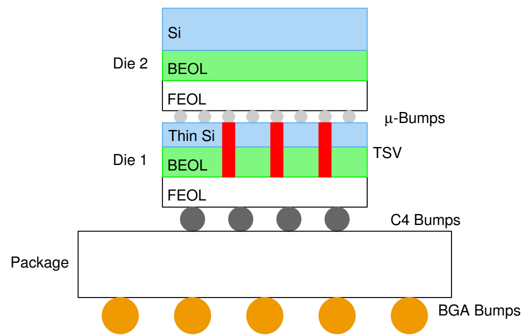
Along with these opportunities and advantages 3D integration introduces new challenges [35]. Some of these challenges stem from the new processing techniques required to



(a)



(b)



(c)

Figure 1-3: Illustration of different stacking approaches (not to scale) (a) Interposer stacking (2.5D), (b) 3D stack using F2F and (c) 3D stack using B2F assembly

correctly create the die stacks with acceptable yield [36, 37]. The stacking of dies can create thermal hot-spots which might degrade performance. The TSVs and back-grinding of the silicon wafer can cause mechanical stresses and local variation of device parameters [38–44]. Coupling of TSVs amongst each other [45–47] and to substrate [48–50] as well as capacitive and inductive coupling between the two dies [51,52] may give rise to power and signal integrity issues [53, 54]. Furthermore, reliability and yield might increase overall cost [55–57]. Adding to these issues is the increased complexity of testing [58], as well as relatively new and immature Computer Aided Design (CAD) tools and software for design and validation [59–62].

1.3 Previous work on 3D-IC

There is abundant research on improvement in timing and power for digital systems due to shortened interconnect [28, 63–65], as well as achieving higher system level performance. Such improvements were shown through architectural re-partitioning of digital blocks [66–68]. Integration of different functional blocks in a 3D stack also shows great potential as seen in examples of integrating digital logic and memory. Below is a partial review of notable recent work which has a greater focus on the circuit and system level potential and challenges of 3D-IC design.

As early as 2004 there has been research carried out by Intel corp. to explore the potential benefits of 3D die stacking for improved system performance in microprocessor design. These initial studies, exploring re-partitioning of deeply pipelined IA32 microprocessors have shown the potential to gain improvements on the order of 15% while reducing power by 15% due to the potential shortening of critical paths and reduced clock routing [66]. Further research focused on integration of logic and memory, combining a multi-core die with an SRAM memory chip to achieve a very high speed interconnect of 1.62 terabits/s and 3x reduction of off-die memory [69].

Professor Eby Friedman's group at Rochester university has been exploring three dimensional integrated circuits for over a decade and have made significant contributions to the field. Among the circuit level exploration, they have demonstrated the use of distributed transmission

lines in a 3D stack to be used as the inductive element in a DC-DC buck converter, enabling reduction of required capacitor size [70]. This technique allows integration of the power control and generation modules on die without the need for external energy storage components. Also explored were architectures for 3D NoC, analyzing the different trade-offs involved and demonstrating a potential of up to 33% performance improvement over 2D NoCs [31] and designs with 12.8 Gb/s bandwidth and down to 0.9 pJ/bit I/O link power efficiency [71].

Two additional topics explored, which have a tight connection to the proposed research are clock distribution topologies in 3D-IC and power distribution and integrity issues. Their research analyzes the impacts of clock tree topology on metrics such as clock frequency, skew, delay and power dissipation [29]. Similarly, they have shown the impact of different power routing topologies on the overall power delivery response of the design [72]. Others have shown potential for power supply integration and routing in 3D demonstrating such benefits as 45% reduction in DC noise and 65% reduction in overhead power consumed by power supply [73], 13.3% reduction in dynamic noise [74] and 12% improvement in power efficiency when using interposer thick metal lines [25].

Davis *et. al* demonstrated the potential benefits of re-partitioning conventional 2D circuits in a 3D topology. They showed that by re-designing the circuit architecture across several die tiers, splitting functionality and memory between them they are able to show power and performance benefits. Such improvements include a 22% reduction in cycle-time and 18% reduction in energy/FFT for an 8192 point FFT processor [75] and a 23% power reduction for a Ternary Content-Addressable Memory (TCAM) implementation [67]. Other similar circuit re-partitioning research has shown 16% reduction in delay time for a 3D floating point adder [30], as well as 2x speed improvement and 2.8x reduced interconnect in 3D tree-based FPGA design [34] compared to their 2D counterparts. These results are mainly driven as explained previously by the possibility to significantly reduce the circuit interconnect length in the 3D stack, which can be reduced by up to 51% [76].

Implementing the concepts previously discussed, Dae Hyun Kim *et. al* [15] demonstrated the benefits of integrating a multi-core processor and memory in a single 3D stack. This massively parallel integration allowed achieving a bandwidth of up to 64 GB/s while consuming 4W of power. Similarly, David Fick *et. al* [77] demonstrated a similar concept of parallel

cores stacked with memory to achieve 3930 DMIPS/W.

Several companies, including IBM [78], Samsung [79] and Micron [80], have also designed commercial products utilizing 2.5D and 3D technology for memory applications including DRAM and flash. As demonstrated previously they claim to achieve high data bandwidth and high memory density while maintaining low power consumption.

An example of integrating digital and analog functionality has been demonstrated by Xilinx in their recent FPGA products. In this work they have used a silicon interposer to place four 28nm FPGA dies side by side, as well as additional Analog to Digital Converter (ADC) and Digital to Analog Converter (DAC) circuits fabricated at a 65nm process node [81]. This integration enabled them to achieve high data rates of up to 400 Gb/s with an interface power consumption of 0.3 pJ/bit and demonstrate competitive performance of the analog blocks in the presence of the digital FPGA dies. It should be emphasized that this work consists of 2.5D, silicon interposer integration and not true 3D vertical die stacking.

Several other publications dealt, as this work does, with potential noise caused in 3D-IC by TSV-TSV or TSV-device coupling and ways to mitigate it. Several techniques, which we will discuss in greater length in Chapter 2, provide some degree of isolation. Use of ground planes between tiers results in a crosstalk reduction of 8 dB [82]. Guard ring structures provide up to 17 dB coupling reduction at 100 MHz and up to 10 dB reduction at 1 GHz [83].

1.4 Thesis Contributions and Outline

This thesis will attempt to take forward the existing body of research by exploring the potential benefits and key challenges of 3D integration of digital and analog circuits. Whereas much research has been done on logic and memory integration in 3D there is still relatively many unknown aspects of mixed signal integration. The research will demonstrate possible new passive structures enabled by 3D integration, their attributes as well as how they can be incorporated and modeled in circuit design. Issues of circuit, block and system level partitioning will be studied to identify key trade-offs and create better design methodologies. Complementary issues of efficient signal processing will be presented which will enable

taking advantage of the high bandwidth and low power benefits promised by 3D integration.

These goals are achieved by building the work on top of previously done research in the literature as well as creating incrementally more complex designs. Findings and lessons learnt from earlier work and test chips will be re-used and implemented in larger systems to both utilize, re-validate and build upon them. This approach has enabled to create a relatively complex and broad system to help better explore the main goals of the research to determine key potential benefits and challenges of system level integration of Radio Frequency (RF) circuits in 3D-IC.

The main contributions of this thesis are in the following areas

- **Passive structures in 3D-IC**. The analysis of inductive coupling between passive structures in 3D-IC is presented in Chapter 2. Voltage Controlled Oscillator (VCO) circuits are ubiquitous in many cellular and RF communication systems and are a prime example of a frequency selective, sensitive analog block, which is at a great risk to suffer from spurious noise caused by digital circuits. Noise coupling between noisy digital blocks and sensitive analog circuits is a key design challenge in many highly integrated systems. In this chapter we explore key contributors to noise and various techniques to model and mitigate noise propagation and coupling. A circuit modeling approach is presented and studied in oder to aid in the design of new structures utilizing the 3D vertical domain for noise mitigation. A proposed inductor structure helps reduce the noise coupling from adjacent clock lines while still maintaining and even improving circuit performance. The small scale test chip helps illustrate both the potential hazard of neglecting to model and consider such effects as electromagnetic coupling in 3D design as well as the potential benefits of taking advantage of the new vertical dimension. The proposed approach exhibits more than an order of magnitude reduction in noise coupling compared to a conventional planar inductor, while also providing twice the quality factor and phase noise performance.

- **Energy-efficient data-dependent signal processing**. Modern communication protocols employ extensive, complex modulation and coding schemes in order to achieve high spectral efficiency and optimal utilization of the available spectrum. Such elaborate

computation also leads in turn to specific signals which share common statistics and characteristics unique to them. An approach designed to take advantage of such properties in the signal data in order to carry out energy efficient computation in hardware is presented in Chapter 3. The work emphasizes the use of the signal generation process to identify key points where power saving can be achieved. This is done at the algorithm, block and circuit level. Savings are achieved by utilizing specialized algorithms which have computational benefit for the specific conditions and data properties. Recognizing the limits of the modes of operation enables to re-use and share hardware blocks more efficiently. Implementing specialized, computation circuits dedicated for the data being processed creates a robust, low-power computation processor for the desired signals. This approach was implemented in the design of a Long Term Evolution (LTE) baseband signal processor to create Single Carrier Frequency Division Multiple Access (SC-FDMA) signals. The combination of these techniques, along with conventional energy-efficient strategies such as clock gating and voltage and frequency scaling enabled to create a very power efficient design. The proposed processor exhibits more than 4x reduction in energy per sample compared to other similar processors reported. Although not directly related to 3D integration, this is a necessary step in the overall system level optimization which is the ultimate goal of the work. In order to demonstrate system level benefits we must make sure that we are designing all system blocks, from the passive structures, through the digital processing and the system level architecture to be as efficient as possible.

- **System level analog-digital integration in 3D-IC**. The potential benefits of 3D integration for form factor reduction and power saving are an appealing feature for future low-power systems if the new arising challenges can be dealt with. The design of part of a direct conversion to RF transmitter for LTE cellular communication in 3D-IC is presented in Chapter 4. The design demonstrates the superior performance of the system and improved isolation by utilizing the 3D integrated process. By partitioning the circuit blocks as to separate the frequency domain operations of the different segments and utilizing the inherent built-in substrate separation of the

3D process, we are able to achieve the required performance without the use of area-consuming isolation techniques while still maintaining high bandwidth and a small form factor and area cost. In order to benefit from overall system level power savings, all parts of the chain must be optimized for power efficiency. In such designs, the dense, low-power interconnect gives an overall system level power benefit. The proposed 3D design exhibits a 0.37 pJ/bit digital link energy efficiency compared to 33.3 pJ/bit for a multiple-chip partitioning and 0.83 pJ/bit for a 2.5D interposer emulated design.

Table 1.1 summarizes several of the key previous research discussed in Section 1.3 which touches along the same themes as this thesis. The main contributions in 3D integration of this work are also listed for comparison. We are able to present both circuit as well as system level improvements by using a holistic, bottom-up approach in our work.

Each chapter in the thesis begins with an introduction and overview of the background information regarding the discussed topic. An extensive account of the analysis and design procedure of each implementation is given, followed by key measurement results. A brief summary of the main findings is given at the conclusion of every chapter. A final summary of all conclusions and several topics which show promise for taking this work even further are presented in Chapter 5.

Table 1.1: Summary of Select Previous Work on 3D-IC

Work	Description	Key Results
Ishida *et. al* [25]	2.5D interposer inductor design for a buck converter	12% improvement in power efficiency by utilizing interposer thick metal
Kim *et. al* [82]	Ground planes between active layers in 3D for noise reduction	8 dB crosstalk reduction
Cho *et. al* [83]	Guard ring structure to mitigate TSV-TSV and TSV-device coupling	Up to 17 dB coupling reduction at 100 MHz and 10 dB reduction at 1 GHz
Erdman *et. al* [81]	2.5D integration of FPGA, ADC and DAC	400 Gb/s BW, 0.3 pJ/bit interface, 1.6 GS/s DAC, 500 MS/s ADC
Dutoit *et. al* [71]	Stacked NoC and memory	12.8 GB/s BW, 0.9 pJ/bit I/O link power efficiency
This work	Vertical solenoid inductor for RF VCO	2x higher quality factor, -6 dB phase noise, -14 dB coupling from adjacent digital clock lines at 1.9 GHz compared to planar inductor
	3D digital-analog transmitter chain integration	0.37 pJ/bit digital power link, reduction in noise coupling compared to one-tier implementation

Chapter 2

Passive Structures

2.1 Introduction

3D-IC have the potential to meet the demand for higher system performance and data rates while avoiding the increase in cost of scaled CMOS technologies. 3D integration opens up the possibility for closely integrating units specialized for different functionalities, such as logic [77], memory [15], optics, power and so forth. The close vertical integration of active silicon layers gives the designer a new degree of freedom in the system and block level partitioning which enables new architectures [28, 29]. The short interconnect between device layers enables high throughput of data and lower power usage due the smaller interconnect parasitics.

The three dimensional environment, including the added Through Silicon Vias (TSVs), enable fabrication of new types of passive devices. Previous work has extensively researched physical and thermal properties of TSVs [84]. The electrical properties of TSVs were also thoroughly analyzed offering many empirical formulas as well as circuit level models at varying degrees of complexity which capture the self parasitic properties of the structures at various frequencies [43, 85–87]. Extended models also allow capturing of the coupling effects between TSVs at different frequency ranges [88–90]. Furthermore, passive structures utilizing TSVs have been explored and analyzed, among them the use of TSVs in creating vertical solenoid inductors [91–96]. These previous studies explore the modeling of such structures as well as demonstrate the potential for higher quality factor inductors and

reduced inductive coupling compared to their planar counterparts.

In this work we explore 3D integration of logic devices with RF circuits. Such coexistence might be hindered due to inductive and capacitive coupling between the tiers. Noise coupling is not unique to such 3D structures and is present also in 2D System on Chip (SoC) design, mainly through substrate noise [97]. The common solution for such issues is to have a high degree of spatial separation between the circuits along with large guard rings [83], resulting in area consumption and increased interconnect length. 3D integration allows creating a high degree of separation and isolation of substrate between domains on the one hand, but on the other hand it brings the two domains into very close vertical proximity of a few dozen microns due to the vertical stacking which may otherwise degrade circuit performance [98]. Large passive structures as on-chip inductors are a potential source of such coupling and performance degradation [99]. A technique is presented to utilize a vertical solenoid inductor in 3D-IC in order to improve the quality factor of the structure and minimize coupling between tiers.

Full 3D Electro-magnetic (EM) solvers, as well as fast field solves and analytic formulas will be used for the analysis of structures and building of intuitive understanding. The results will be used to formulate design guidelines as well as suggest techniques for mitigating coupling and interference between circuit blocks. The new 3D structures will also be compared and contrasted to more conventional passive structures such as planar integrated inductors in silicon.

The outline of the chapter is as follows. An overview of noise contribution and propagation mechanisms as well as a review of commonly used isolation techniques is given in section 2.2. The details of the technology, circuit implementation and analysis are presented in section 2.3. Section 2.4 presents the measurement setup and results, as well as a comparison to the theoretical derivations and simulations. We present our final conclusions in section 2.5.

2.2 Analog-Digital Integration

Modern high performance communication systems are required to perform many functions while maintaining high signal integrity with stringent dynamic range requirements, often in excess of 100 dB. The signal processing blocks consist of such elements as Phase Lock Loop (PLL), VCO, Mixer, Low Noise Amplifier (LNA), Digital baseband, DAC, ADC and power conversion among others.

The signal processing functions can be physically separated into different Integrated Circuits (ICs), packaged individually, as is commonly done in many applications today such as hand held mobile devices. In this manner it is possible to obtain the required isolation between the blocks, however we incur a penalty in terms of the required system area, placing separate packages on a Printed Circuit Board (PCB) as well as possible maximum bandwidth due to the need to traverse through package and board interconnect with high parasitic inductance and capacitance.

2.2.1 Noise Contributors

Integrating several blocks onto one die, forming a more complete SoC, allows reduction of the area and bandwidth limitations. However, this solution introduces a myriad of new design issues to consider including thermal distribution, power integrity and signal integrity. Focusing on signal and power integrity issues, these mainly stem from the degraded isolation between the blocks. The most dominant aggressors in such a system will be the Power Amplifier (PA) and the digital clock. While the PA is relatively frequency selective, since it is usually a highly tuned circuit, the clock has a wide spectral range due to its square wave shape and sharp transition edges. The digital I/O signals may also act as a noise source similar to the digital clock, however the clock will likely be more dominant in the general case due to its wider routing and high power. Furthermore, the nature of the clock signal, desired to swing rail-to-rail leads to the fact that even for a clock operating at a few hundreds of MHz there are harmonic components with non-negligible power levels at RF. The noise created by these harmonics is propagated to the rest of the die through EM coupling (via line-to-line capacitance or radiating inductive elements) and substrate

coupling via the shared bulk. Following are some typical examples of issues that might arise from such coupling and noise injection [100, 101].

Large signal tones which lie inside the RF band of interest can couple into parts of the front-end. This will cause for example elevation of the noise floor for a LNA, or even de-sensitization leading to saturation if the coupling is strong and the LNA linearity is not high. It will also appear as spurs at the output of a VCO and degrade the linearity of the mixer or PA. The noise generated by the digital system clock is coupled into the die substrate and propagates to couple into other system elements.

The coupled noise will create a DC offset at the output of devices such as an ADC or mixer. In direct-conversion systems this is a more severe problem since the baseband data contains a DC component which can be affected by such an offset.

In some cases, when the noise frequency is close enough to the frequency of the local oscillator and powerful enough, it may cause frequency pulling, causing the oscillator to lock to a different frequency than desired due to the injection of the interferer [102].

Additionally, coupling through the substrate might yield oscillatory loops and unwanted feedback between separate system parts. This is especially true in direct-conversion systems since they have a high gain at a single frequency band. Heterodyne systems alleviate this problem somewhat due to the fact that the system gain is partitioned over several frequencies.

2.2.2 Isolation Techniques

In order to mitigate these effects several design approaches are used. Presented below are several common techniques which aim to minimize the interference by separation in frequency, time or space. These include layout, circuit and system architecture modifications.

In order to improve isolation it is common to ensure a substantial amount of physical planar separation between noisy digital blocks and sensitive analog blocks on the die. This however incurs an area penalty and also causes the interconnect between the blocks to be much longer, thus degrading the bandwidth of the system and increasing the dissipated power.

Another isolation technique often found in such mixed-signal chips is the use of guard rings around sensitive or noisy blocks in order to improve the isolation. In this regard it is important to distinguish between epitaxial (epi) and non-epitaxial (non-epi) wafers. In epi wafers, there is a shallow ($\sim 10\mu$m) epitaxial layer which is lightly doped presenting a high resistivity of approximately 10-20$\Omega \cdot$cm on top of a highly doped, low resistivity (10-20m$\Omega \cdot$cm) thick bulk. The non-epi wafer offers a bulk which is entirely lightly doped with high resistivity. The epi wafer offers better avoidance from latch-up issues by preventing parasitic voltage drops over resistive substrate, where as the non-epi wafer reduces inductor losses due to substrate eddy currents and presents better substrate isolation by avoiding the equi-potential bulk node. Therefore, in an epi wafer, guard rings will offer little to no protection and isolation between noise sources and circuits [103]. In a non-epi wafer, a guard ring around the noisy or sensitive circuit offers some degree of isolation which is dependent on the size of the guard ring, the distance between the blocks, the size of the blocks themselves and further parameters such as the number of ground connections to the guard ring and their individual inductance path to ground [104]. Overall typical values are in the range of 20-25 dB of isolation [105]. The addition of the large guard rings however, requires a substantial area cost, in some cases up to about 10% of the total die area [104].

Further physical isolation can be gained in some processes which allow for the use of deep N-well (or triple well) [106]. Separating the NMOS device's bulk connection from the general substrate typically achieves an isolation of about 40-50 dB. The use of such a process does require though additional mask layers in the BEOL process for fabrication.

Since these isolation techniques are not always sufficient in achieving the desired sensitivity and dynamic range, other approaches should be used. One such approach is to ensure that the digital clock frequency is offset to a frequency which has minimal harmonic content at the operating frequency and harmonics of the analog/RF blocks [102]. This implies a dependence between the operation of the domains (RF, Intermediate Frequency (IF) and baseband) and will also limit the possible operating frequencies of the system. This is likely to be a drawback in modern systems where we see a greater tendency towards multi-band, multi-standard systems which are expected to handle various frequency scenarios.

Other techniques include reducing the harmonic content of the clock by introducing intentional jitter [102] (and as a consequence sacrificing digital performance). A variant of the former includes adjusting the phase of the clock during operation to avoid pulling of the oscillator [107]. And yet another technique would be minimizing the impact of even harmonics by using differential topologies with high common mode rejection ratios.

The aforementioned techniques and isolation methods are used in conjunction in order to achieve the required sensitivity levels and performance metrics of the system. However, much of the information regarding such adverse effects is available only at relatively late stages of the design, requiring macro models for blocks and physical layout for parasitic extraction and will also require additional EM modeling in order to capture the effects and ensure that the system requirements are met [108].

2.2.3 Isolation in 3D-IC

One of the potential advantages of 3D-IC is the separation of substrate between die tiers and thus alleviating some of the coupling and noise sources discussed previously. The physical separation however is not enough in order to mitigate coupling since signal and power lines may still connect the different tiers and capacitive and inductive coupling may still exist. The fact that the die tiers are in very close vertical proximity of a few dozen micrometers makes coupling effects potentially more dominant and also difficult to model and predict.

Several techniques have been proposed to mitigate such coupling effects. These include mostly shielding in various topologies in order to confine the electric field lines from propagating and capacitively coupling to other metal lines and the substrate. One approach is to create a complete “Faraday Cage” around sensitive TSVs [109]. Other approaches try to reduce the area overhead by implementing a more limited approach such as using shielding metal ground planes [82], or a modified TSV formation process to create a miniature coaxial style structure [110]. Other approaches utilize known approaches from conventional planar ICs and utilize guard rings to protect sensitive circuits [83].

These techniques however do not assist in the mitigation of inductive, magnetic coupling. Since the magnetic field is not terminated on conductive metal shields these are not suitable

to mitigate such effects. In this work we will explore alternative options to reduce the effects of inductive coupling between die tiers among other noise coupling mechanisms.

2.3 Test Chip Design

2.3.1 3D Stacking

The 3D stack used in this work is depicted in Fig. 2-1. Each die in the 3D stack was designed and fabricated separately in a conventional planar 28 nm bulk CMOS process with 7 metal layers and a redistribution layer. The stack consists of two tiers. The top silicon wafer is back grinded and the silicon substrate is thinned to approximately 60 μm. TSVs are drilled from the backside, filled with copper and connect to the polysilicon layer of the top die. On the backside, the TSV connect via micro-bumps to the top metal layer of the bottom tier die. This process may be extended further to include more tiers stacked vertically to extend the system.

The TSV diameter used in this process was 10 μm with a TSV-to-TSV pitch of 40 μm. The TSVs enable passing of signals between the die tiers and allows utilizing both stack tiers as active silicon layers for circuit fabrication. The overall die stack results in close proximity of the die circuits and metal layers to each other on the order of a few dozen micrometers.

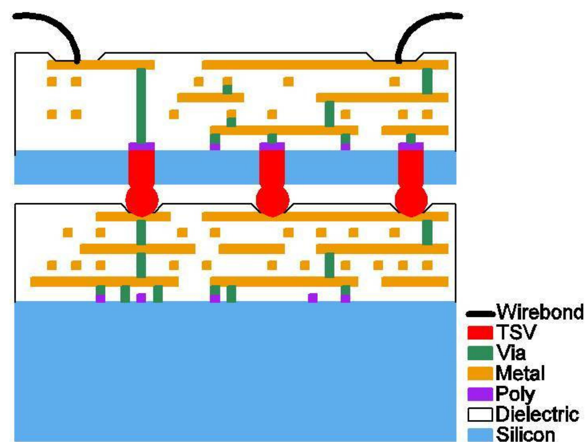


Figure 2-1: Illustration of Back-to-Face 3D die stack (not to scale)

2.3.2 Inductor Design

Fig. 2-2 shows an illustration of two stacked dies. The bottom tier die contains signal lines which emulate a part of a high-speed digital clock tree which would be an integral part of any large digital block and would likely be the most dominant aggressor in such a scenario to sensitive analog circuits. Directly above the clock lines, two different integrated inductor structures were fabricated - the proposed solenoid for noise mitigation and a conventional planar structure as a reference design.

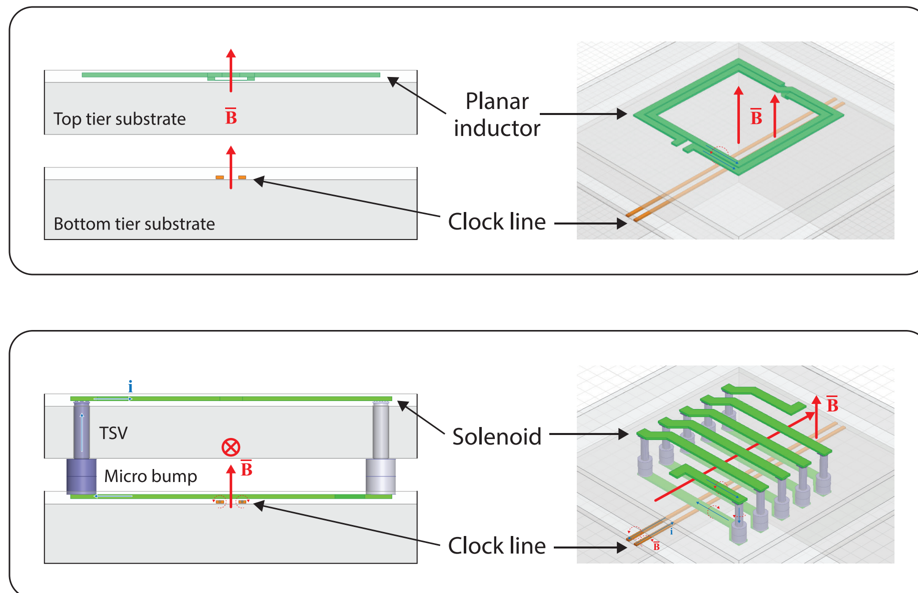


Figure 2-2: Illustration of 3D-IC stack with clock lines, planar and solenoid inductor structures, along with current directions and magnetic field lines (figure not to scale)

A typical shielded clock line, consisting of a signal line and a return path will generate a magnetic field around the lines. Due to the proximity of the lines, the strongest magnetic field will appear between the lines and will be directed in the vertical Z direction, while the magnetic field outside that area will be much weaker due to the cancellation caused by the opposite currents of the signal and return path. When routed underneath a structure such as a planar inductor which also has a strong vertical magnetic field component, we expect a high degree of coupling between them causing spurious tones on one to be coupled to the other [111]. This is usually an undesired effect and is often avoided by simply clearing the entire area underneath the inductor [112] at the cost of unused area, similar to the 2D

spatial separation. The solenoid structure however has a magnetic field mainly aligned in the XY plane along its axis, perpendicular to the clock line magnetic field, thus minimizing their coupling. There still exists coupling due to the outside fields, and mainly due to edge effects where the fields do not cancel each other. This will result in greater coupling when the clock lines are parallel to the solenoid turns in contrast to an orthogonal pattern. Analysis of the topologies using field solvers shows that due to the perpendicular magnetic field, the inductive coupling to orthogonal lines routed beneath the solenoid would be at least twice as weak than that of a conventional planar inductor.

The planar inductor utilizes the top metal redistribution layer and has a patterned ground shield in the polysilicon layer. The solenoid inductor uses the TSVs themselves as part of the inductor structure and the redistribution layers on both the top and bottom die tiers. In both cases the area directly underneath the inductor structure on the bottom die is free for use for other active devices with the exception of the top metal layer in the case of the solenoid inductor. Both inductors were designed to occupy an area of 200 μm x 200 μm.

The use of the vertical domain further enables us to gain a substantially larger loop area for the solenoid and thus obtain a higher inductance per the given occupied planar area. Furthermore, the use of the TSVs as part of the inductor structure allows reduction of the serial resistance of the inductor and obtain a higher quality factor.

The monolithic inductors are electrically modeled as shown in Fig. 2-3 [113]. The two-port admittance matrix is written as

$$Y = \begin{pmatrix} Y_{sub} + Y_{ind} & -Y_{ind} \\ -Y_{ind} & Y_{sub} + Y_{ind} \end{pmatrix}. \tag{2.1}$$

The component values were derived from a two port 3D EM field solver simulation via the following equations [114]:

$$Y_{ind} = -Y_{12} = \frac{1}{R_s + j\omega L_s} \tag{2.2}$$

$$Y_{sub} = Y_{11} + Y_{12} = \frac{j\omega C_{ox}(1 + j\omega R_{sub} C_{sub})}{1 + j\omega R_{sub}(C_{ox} + C_{sub})} \tag{2.3}$$

$$L_s = \frac{\Im\left\{{}^{1}/_{Y_{ind}}\right\}}{\omega} \tag{2.4}$$

$$R_s = \Re\left\{{}^{1}/_{Y_{ind}}\right\} \tag{2.5}$$

$$Q = \frac{\Im\{Y_{11}\}}{\Re\{Y_{11}\}}. \tag{2.6}$$

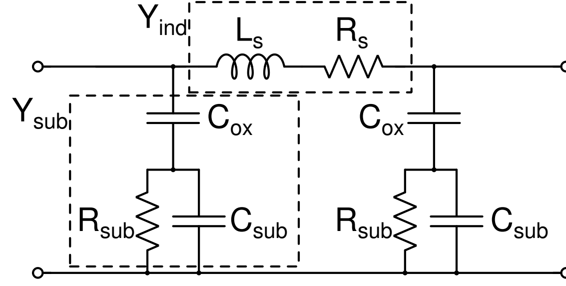


Figure 2-3: Monolithic integrated inductor equivalent circuit model

The oxide capacitance C_{ox} models the capacitance from the inductor windings through the dielectric layers to the substrate in the planar inductor case, and also the TSV capacitance to substrate for the solenoid inductor. C_{sub} and R_{sub} represent the substrate lossy capacitance. The series resistor R_s captures both the series metal resistance of the inductor windings as well as the losses due to the substrate induced eddy currents. It should be noted that the use of the patterned ground shield [115] of the planar inductor helps to reduce the substrate capacitance but does not reduce the eddy current induced losses since the ground shield terminates the electric field but does not block the magnetic field from entering the substrate.

Fig. 2-4 plots the simulated inductance and quality factor of the structures as obtained from a field solver and derived using Eq. (2.1)-(2.6). The solenoid structure indeed exhibits a superior quality factor by almost 2x around 2 GHz consistent with previous findings [93, 116].

The main drawback of using such a solenoid inductor compared to a planar one is the fact that the structure is now dependent on both die tiers. This means that the top metal layer of the bottom tier die is appropriated to the inductor structure and may not be used by the designer of the bottom die. Furthermore, this topology is more susceptible to

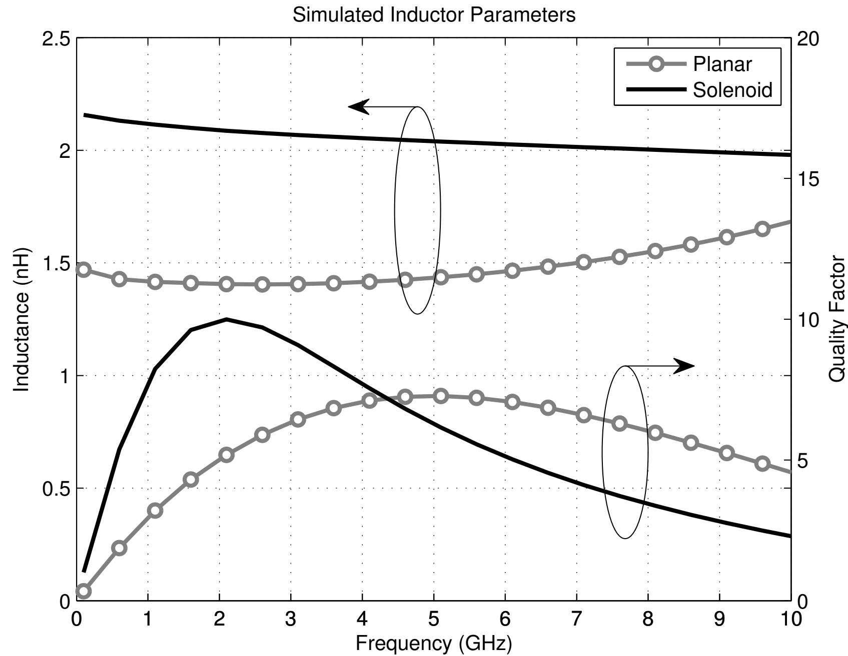


Figure 2-4: Simulated inductor parameters extracted from field solver simulation results

assembly variation than a planar inductor since the inductor parameters are also affected by the precision of die alignment and stacking separation. Finally, it will not be possible to test the structure, and any circuit dependent on it as separate die tiers before the entire stack has been assembled, potentially negatively affecting yield.

2.3.3 VCO Design

In order to demonstrate the aforementioned benefits and act as a proof-of-concept for digital and analog integration, a VCO was designed in 28 nm CMOS to use each of the proposed inductor structures. Fig. 2-5 shows the circuit diagram of the VCO, replicated once with a conventional planar inductor and once with the proposed solenoid structure.

The VCO core consists of a cross-coupled pair of NMOS and PMOS devices [117]. The VCO operates from a 1 V supply rail and draws 5 mA of tail current. The VCO core is connected to the resonant LC tank in order to set the free running oscillation frequency. The tank is tuned via a varactor and switched capacitor bank.

A biased MOS varactor was used in order to fine tune the VCO frequency. The varactor

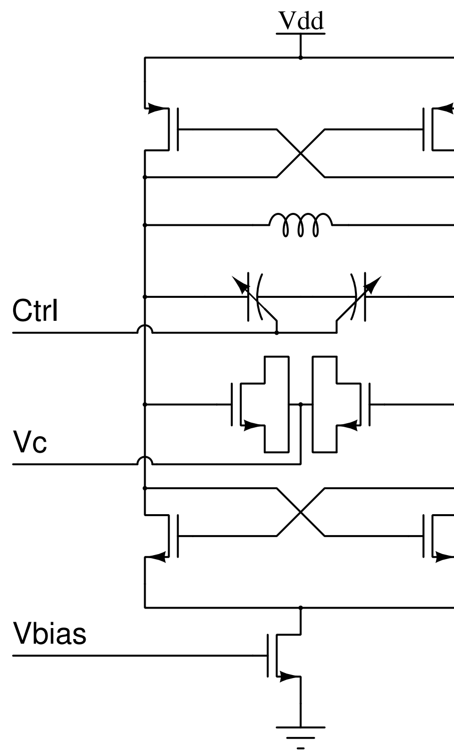


Figure 2-5: CMOS VCO core schematic used with both the planar and solenoid inductor structures

circuit schematic is shown in Fig. 2-6a. Decoupling capacitors are used to set the bias point at mid-rail of the supply voltage in order to maximize the frequency range of the tuning varactors. The varactor was designed to vary around 470 fF over a 1 V tuning range.

In order to be able to compare the two inductor structures we wish to operate the VCOs around the same frequencies despite the difference in self inductance of the structures. Therefore, an additional 3-bit binary weighted MOM capacitor bank is used in conjunction with the continuously tunable MOS varactor to extend the available VCO center frequency tuning range. The differential switched capacitor bank unit cell is depicted in Fig. 2-6b. Each unit cell was designed to provide 400 fF of capacitance in order to include some overlap between the varactor tuning coverage and the switched capacitor bank.

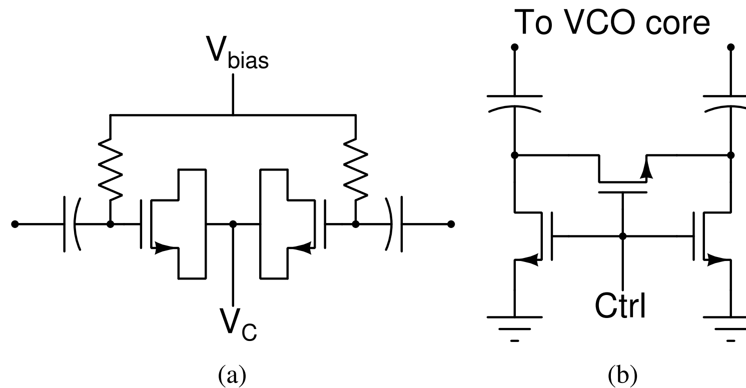


Figure 2-6: Frequency tuning of VCO achieved via (a) continuous MOS varactor and (b) switched capacitor bank unit cell

The output of the VCO was connected to a capacitive divider leading to a source-follower output as shown in Fig. 2-7. The capacitive divider enabled reducing the signal level to avoid distortion at the output as well as re-bias it using the resistive divider for input to the output amplifier. The output however is amplified individually for each branch and is not truly differential at the output,

2.3.4 Clock Lines

In order to excite the on-die clock grid, an external high-speed clock was supplied to the chip via external pads. Two variants of the clock input buffer were designed to ensure

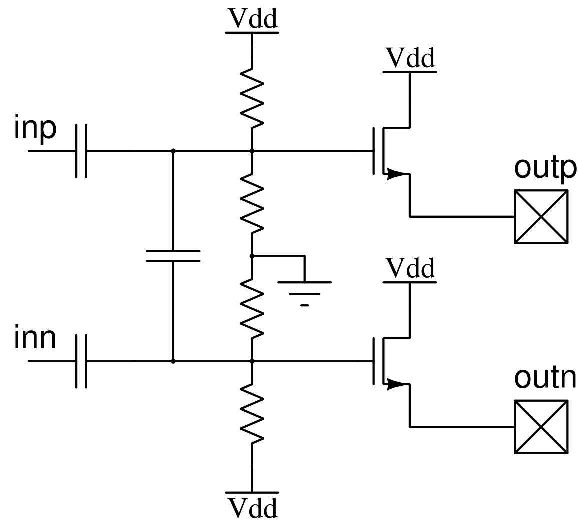
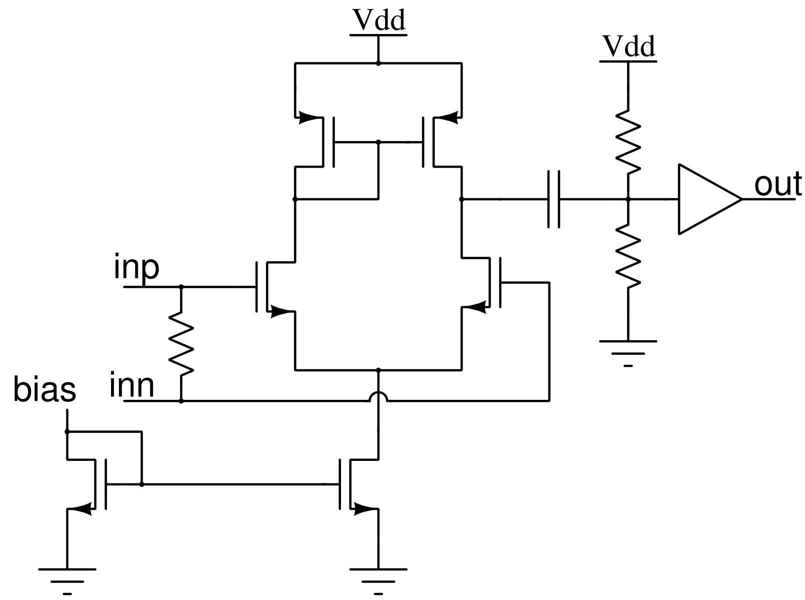


Figure 2-7: Differential capacitive divider source follower output buffer schematic

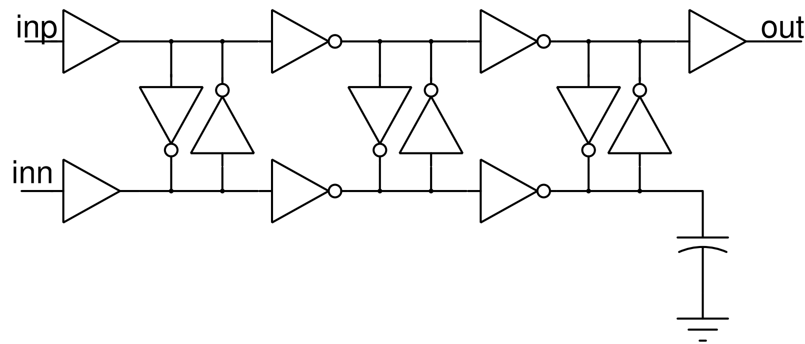
proper operation. A differential, active load Operational Amplifier (OpAmp) was implemented with 50 Ω differential input termination as shown in Fig. 2-8a. The single ended output was DC coupled to a standard cell buffer and from there distributed to the clock network. The second input buffer is shown in Fig. 2-8b and consists of a pseudo-differential inverter chain, with back-to-back balancing inverters. The chain utilizes standard digital inverter and buffer cells. One output branch is tied to a dummy capacitive load, while the other is used to feed the clock network, The choice between the two input buffers was done by a control configuration bit loaded to the test chip. No significant difference between the operation of the two was detected for the frequency range of interest used in theses tests.

In order to verify the clock behavior and the input buffer performance, the clock was also fed through a divide-by-16 chain as shown in Fig. 2-9. The output of the clock divider was routed off chip for external measurement and to confirm proper clock signal generation, timing and duty cycle performance.

Several controllable clock lines were implemented beneath each inductor structure at various offset positions. Each metal line is 1 μm wide and runs along the entire length of the inductor structure, i.e. 200 μm. A schematic representing the clock lines and their respective control signals is illustrated in Fig. 2-10. This control scheme will allows individual control of the amount and location of the coupling clock lines, thus adjust for



(a)



(b)

Figure 2-8: Clock input buffer (a) differential amplifier and (b) pseudo differential buffer chain schematics

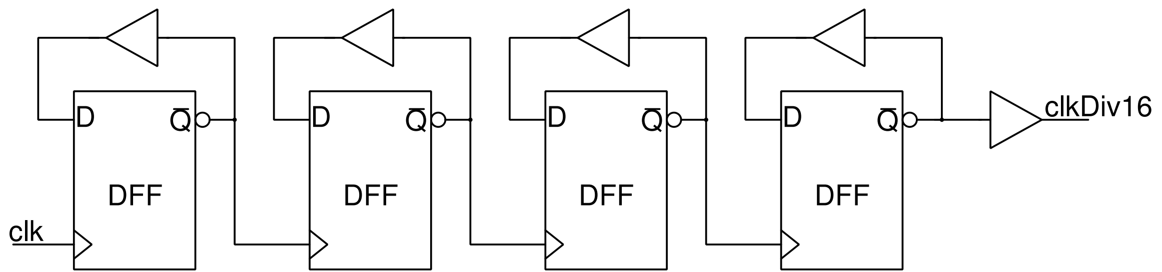


Figure 2-9: Clock divide by 16 circuit schematic

the desired coupling strength.

All clock lines are fed via a balanced distributed clock tree so as to ensure in-phase operation. This method allows for easier study of the effects of the various clock line combinations. Since the mutual coupling effect is linear, it is possible to use superposition to analyze the combined effects of several clock lines operating simultaneously on the VCO performance.

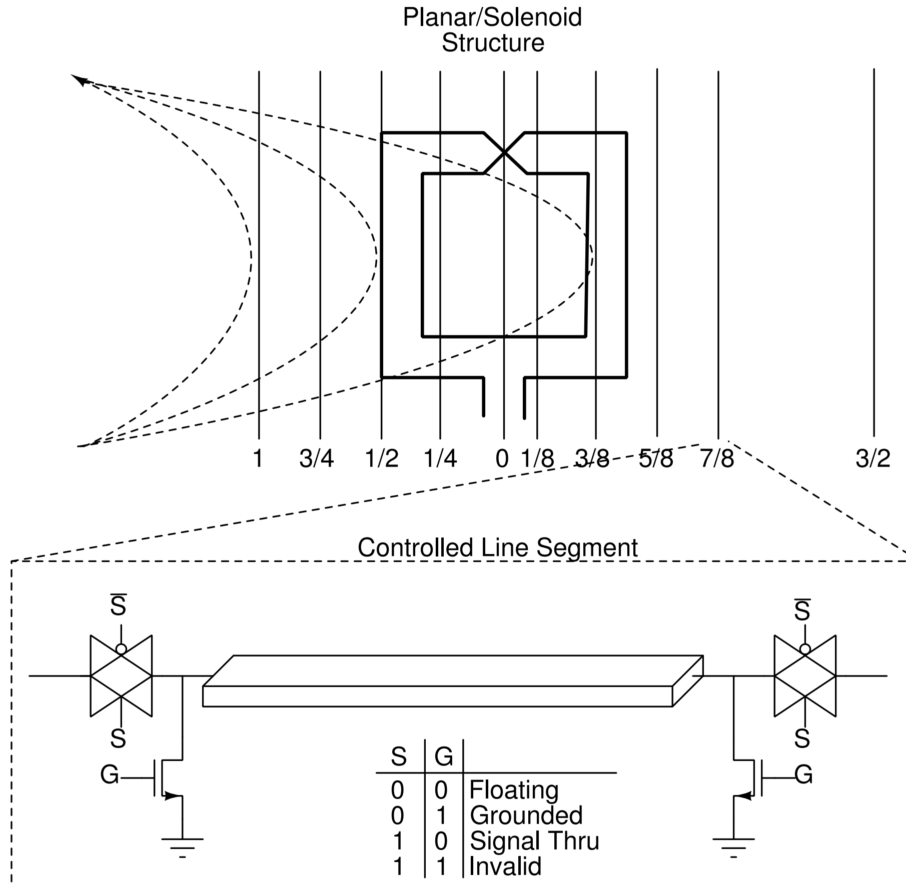


Figure 2-10: Controllable clock lines location relative to inductors' dimension and schematic of line switch control

2.3.5 Coupling Analysis

In order to analyze the impact of coupling between the digital clock lines and the VCO performance we must first estimate the coupling magnitude of different inductor structures

to other metal line routings at various offset locations. Since we are only interested in inductive coupling between metal geometries, there is no need to carry out lengthy full 3D EM simulations. Other, faster approaches can be used to obtain the coupling coefficients such as FastHenry [118]. Fig. 2-11 plots the simulated mutual inductance values for various clock line locations relative to the inductor structures for both the planar and solenoid inductors. It can be seen that the solenoid inductor exhibits a consistently lower coupling to the clock lines throughout up to an order of magnitude smaller coupling.

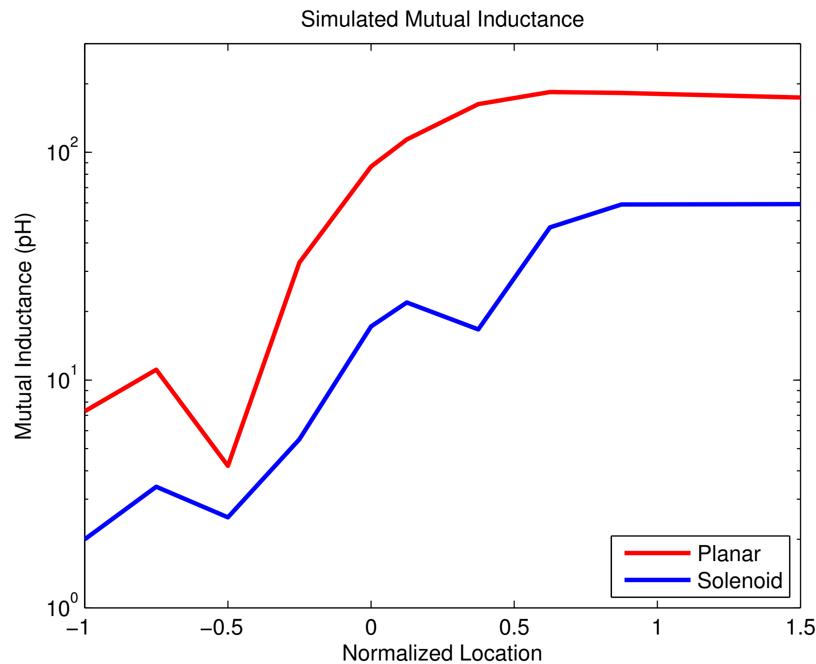


Figure 2-11: Simulated mutual inductance between clock lines and inductor structures as a function of normalized clock line offset location

To gain a better understanding of the expected impact of coupling on the VCO output, we construct a simplified model of the scenario as depicted in Fig. 2-12. In this simplified scheme, the clock lines are modeled as simple series RLC circuits driven by an AC source. The self resonant frequency of the clock lines is naturally much higher than the typical driving frequencies which we are considering, causing an attenuation of the resulting currents. The VCO is simplified down to its resonant LC tank along with the equivalent parallel resistance modeling the finite closed-loop quality factor of the tank. It is important to

emphasize that this resistance includes the effect of the negative resistance created by the active VCO core devices to achieve an overall higher quality factor than the inherent one of the LC tank itself. The coupling between the clock line and the VCO tank is modeled by the mutual inductance coefficient k and mutual inductance value M, where

$$M = k\sqrt{LL_C}. \tag{2.7}$$

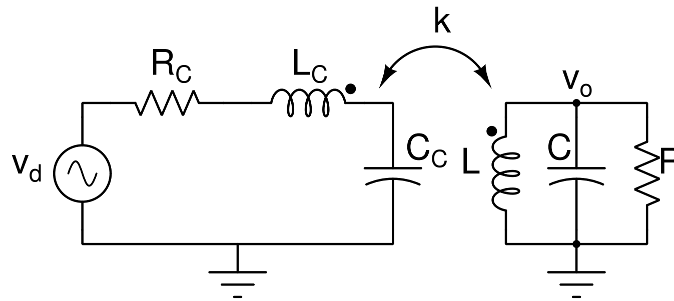


Figure 2-12: Simplified coupling model between clock lines and VCO LC tank

We now calculate the impact on the tank output voltage (and therefore the power output spectrum) caused by the driven coupled clock line as approximately

$$P_{spur} \propto \left| \frac{M\omega^2}{1 - \left(\frac{\omega}{\omega_0}\right)^2} \right|^2 \tag{2.8}$$

when considering the impact of harmonic tones close to the free running frequency of the VCO (see section 2.3.6), where ω_0 is the resonant frequency of the tuned VCO. Since the tank voltage is linearly proportional to the mutual inductance and the power spectrum is proportional to the square of the tank voltage we expect a quadratic dependency between the mutual coupling and the power of the output spur. Furthermore, when considering the frequency of the driven clock signal, we observe a resonant amplification behavior around the VCO's natural frequency.

When considering a more realistic scenario, with a large digital block with many metal routings it would be unreasonable to model every line in order to calculate the coupling

effect as described previously in this section. However, it is important to remember that the inductive coupling effect is relatively local and is only dominant in the 3D stack due to the very close vertical proximity between die tiers. Therefore we limit our modeling to circuits within 2-3 times the inductor's dimensions. In this area, the most dominant aggressor is most likely the clock lines which oscillate at a constant frequency and have significant power due to driving a greater capacitive load. Modeling the main clock routing lines is likely sufficient in order to obtain a good approximation for the coupling between the digital circuit and the main inductor structure.

2.3.6 Coupling Transfer Function Calculation

From the circuit model described in Fig. 2-12 we write the sinusoidal steady state circuit KVL equations in matrix form to obtain

$$\begin{pmatrix} v_d \\ v_o \end{pmatrix} = \begin{pmatrix} \frac{1}{sC_c} + sL_c + R_c & sM \\ sM & sL \end{pmatrix} \begin{pmatrix} i_c \\ i_L \end{pmatrix} \tag{2.9}$$

Where i_c and i_L are the currents in the clock line circuit, namely the current through the clock line capacitance and the current through the VCO inductor respectively. Writing the values of these currents using the tank output voltage and the clock line capacitance voltage we obtain

$$\begin{pmatrix} i_c \\ i_L \end{pmatrix} = \begin{pmatrix} sC_c & 0 \\ 0 & -sC - \frac{1}{R} \end{pmatrix} \begin{pmatrix} v_c \\ v_o \end{pmatrix} \tag{2.10}$$

Combining (2.9) and (2.10) yields

$$\begin{pmatrix} v_d \\ v_o \end{pmatrix} = \begin{pmatrix} 1 + sR_cC_c + s^2L_cC_c & -s\frac{M}{R} - s^2MC \\ s^2MC_c & -s\frac{L}{R} - s^2LC \end{pmatrix} \begin{pmatrix} v_c \\ v_o \end{pmatrix} \tag{2.11}$$

Rearranging (2.11) and substituting the following relations

$$\omega_0 = \frac{1}{\sqrt{LC}} \tag{2.12}$$

$$Q = R\sqrt{\frac{C}{L}} \tag{2.13}$$

$$\omega_c = \frac{1}{\sqrt{L_c C_c}} \tag{2.14}$$

$$Q_c = \frac{1}{R_c}\sqrt{\frac{L_c}{C_c}} \tag{2.15}$$

results in

$$\begin{pmatrix} v_d \\ 0 \end{pmatrix} = \begin{pmatrix} 1 + \frac{s}{\omega_c Q_c} + \frac{s^2}{\omega_c^2} & s\frac{M}{R} + s^2 MC \\ s^2 MC_C & 1 + \frac{s}{\omega_0 Q} + \frac{s^2}{\omega_0^2} \end{pmatrix} \begin{pmatrix} v_c \\ -v_o \end{pmatrix} \tag{2.16}$$

Multiplying (2.16) by the inverse matrix, we derive the transfer function from the driving clock signal to the VCO tank output.

$$\frac{v_o}{v_d} = \frac{-\omega^2 MC_c}{H(j\omega)H_c(j\omega) - k^2\left(\frac{\omega}{\omega_0}\right)^2\left(\frac{\omega}{\omega_c}\right)^2\left[1 + \frac{j\omega_0}{\omega Q}\right]} \tag{2.17}$$

Where $H(j\omega)$ and $H_c(j\omega)$ are the characteristic functions of the parallel LC tank and serial clock circuit respectively, defined as:

$$H(j\omega) = 1 - \left(\frac{\omega}{\omega_0}\right)^2 + \frac{j\omega}{\omega_0 Q} \tag{2.18}$$

$$H_c(j\omega) = 1 - \left(\frac{\omega}{\omega_c}\right)^2 + \frac{j\omega}{\omega_c Q_c} \tag{2.19}$$

This result is simplified further under the assumptions that the coupling is relatively weak, i.e. $k \ll 1$. We also assume that the self resonance frequency of the clock line is much higher than that of the VCO frequency and that we are mainly concerned in the behavior of the output at frequencies which are close to this frequency, implying $\omega \approx \omega_0 \ll \omega_c$. Finally, recalling that Q represents the closed-loop quality factor of the VCO, and therefore $Q \gg 1$, we approximate the characteristic functions as $H(j\omega) \approx 1 - \left(\frac{\omega}{\omega_0}\right)^2$ and $H_c(j\omega) \approx 1$ as well as neglecting the coupling term in the denominator so as to approximate (2.17) as

$$\frac{v_o}{v_d} \approx \frac{-\omega^2 MC_c}{1 - \left(\frac{\omega}{\omega_0}\right)^2} \tag{2.20}$$

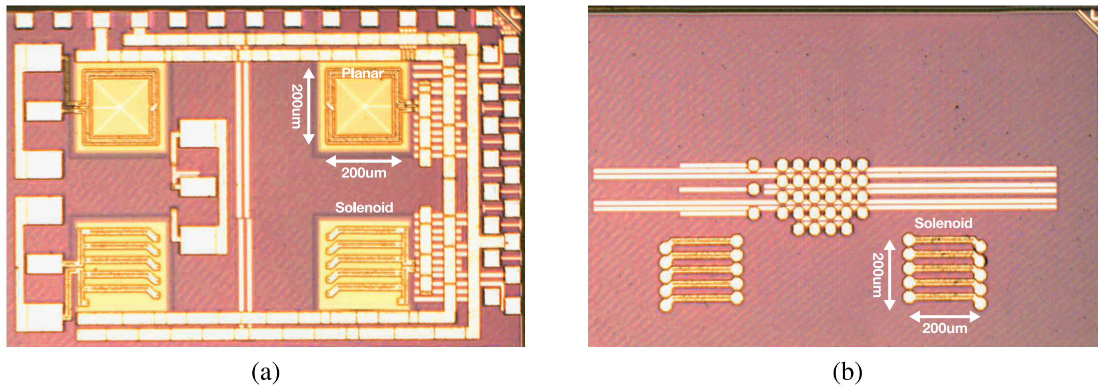
The coupled spur's power will be proportional to the square of the signal's voltage and therefore we conclude that

$$P_{spur} \approx K \left| \frac{M\omega^2}{1 - \left(\frac{\omega}{\omega_0}\right)^2} \right|^2 \tag{2.21}$$

where K is a proportionality factor combining both the physical clock line capacitance, the clock signal voltage level and conversion of the voltage quantities to power.

2.4 Measurements

The proposed solenoid and planar inductor structures, as well as the test VCO and controllable clock lines were fabricated in a two-tier 28nm CMOS stack. The TSV pitch used was 40 μm with a diameter of 10 μm. The stack height between the tiers' top metals is roughly 60 μm. A die photo of the top and bottom tiers is shown in Fig. 2-13.



(a) (b)

Figure 2-13: Die micrograph (a) top tier (b) bottom tier

The 3D die stack was packaged via wire-bonds in a Quad Flat No-leads (QFN) package and assembled on a PCB for testing. Supply and bias voltages are supplied to the test chip via Low Dropout Regulator (LDO) components. The clock signal is supplied through an RF generator to the chip and the VCO output is measured through the use of a real time oscilloscope and a spectrum analyzer. The various configuration options for enabling the clock lines and controlling the VCO capacitor bank are controlled via an FPGA communicating with a PC. Fig. 2-14 shows a picture of the test PCB and the various components.

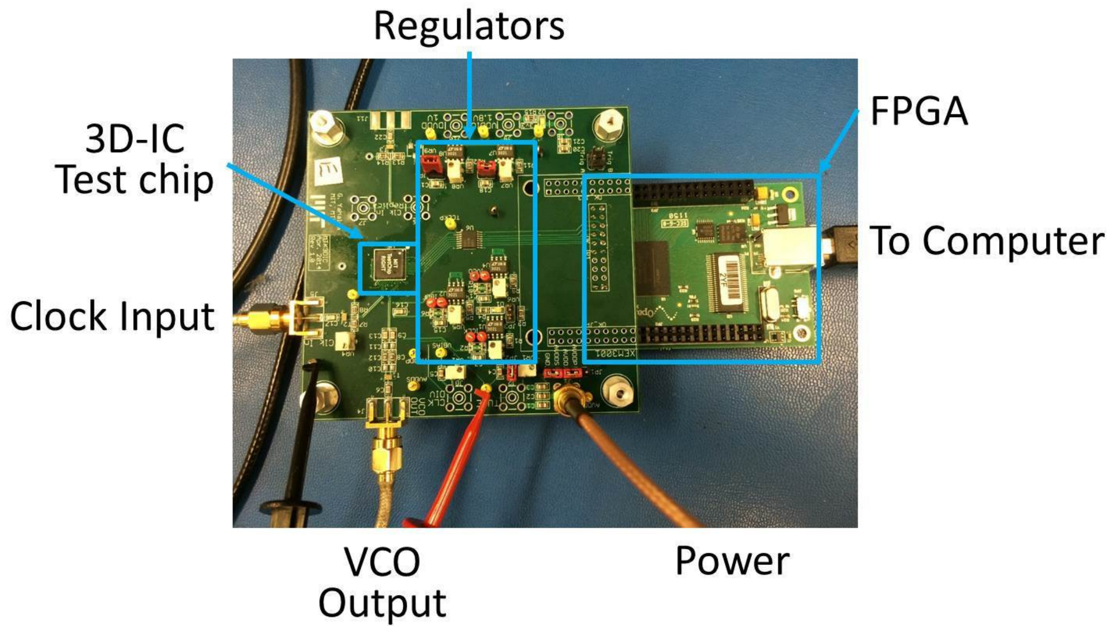


Figure 2-14: PCB used for testing and measurements

The VCO frequency output and tuning range are shown in the measurement results of Fig. 2-15. As can be seen, the tuning range of the VCO with the planar inductor spans from 1.88 GHz to 2.37 GHz (TR = 23%) and the range for the solenoid structure is from 1.67 GHz to 2.09 GHz, as would be expected from their different inductance value and substrate capacitance. The overlap in the frequency range allows for an easier comparison of the structure performance.

Fig. 2-16 shows measurement results for the phase noise. The phase noise of the VCO is plotted for a carrier frequency of 1.9 GHz for both VCO inductor structures. The improved quality factor of the solenoid structure is manifested in approximately 6 dB improvement of the phase noise [119], improving from -116 dBc/Hz at 1 MHz offset for the planar inductor to -122 dBc/Hz for the solenoid. The VCO core consumes 5 mA from a 1 V supply in both cases resulting in a Figure of Merit (FOM) of 175 and 181 dBc/Hz for the planar and solenoid structures respectively. These results are comparable to other similar CMOS VCOs, and the same design techniques and concepts described may be readily adopted and used with other such designs to achieve an even higher FOM [120]

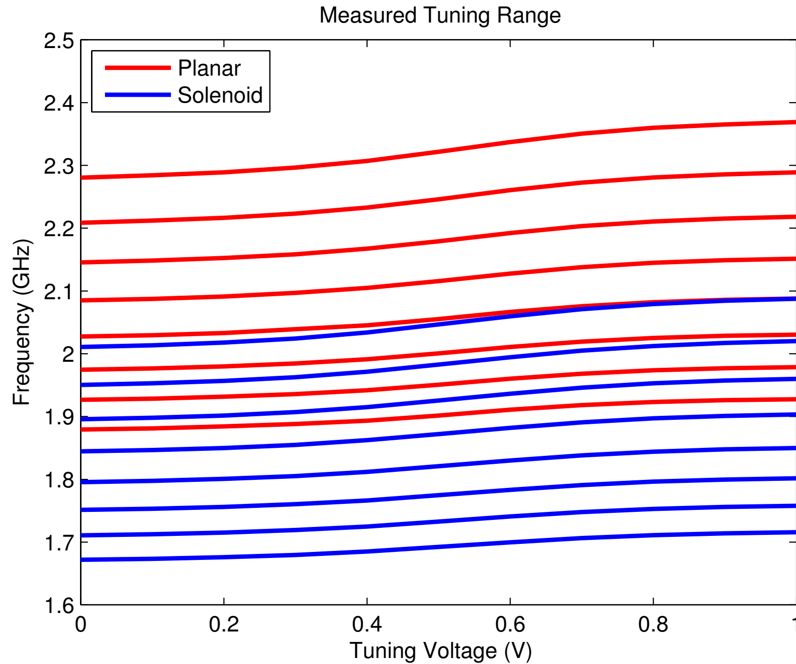


Figure 2-15: Measured tuning range for planar and solenoid VCOs

while maintaining the benefits of the improved noise immunity. A summary of the key metrics and measurement results of the VCO are presented in Table 2.1.

Although it was not possible to measure the mutual inductance between the clock lines and the inductors directly, we are able to obtain an indication for it through the power of the output spurs in the VCO spectrum as implied by (2.8). A digital rail-to-rail clock signal around 633 MHz was used in order to stimulate a 3rd harmonic signal close to the VCO free running frequency of 1.9 GHz. Clear spurious tones are observed in the output spectrum of the VCO. An example of these clock spurs along with the VCO main output frequency is depicted in Fig. 2-17.

Varying the frequency of the clock signal we expect to observe a resonant-like behavior in the output power of the spur as indicated by (2.8). The measurement results of a frequency sweep close to the VCO free running frequency is shown in Fig. 2-18, along with the theoretical curve. As can be seen, there is a very good match between the measured results and the theoretical curve line.

Fig. 2-19 plots the measured output spur power of the clock signal on the VCO

Table 2.1: Summary and comparison of results

	Planar	**Solenoid**
Technology	3D stacked 28nm LP CMOS	
Frequency Range (GHz)	1.88 ∼ 2.37 (23%)	1.67 ∼ 2.09 (22%)
Power (mW)	5	5
Phase Noise @1MHz (dBc/Hz)	-116	-122
FoM[1]	175	181

[1] $\mathrm{FoM} = 20 \cdot \log\left(\frac{\omega_0}{\Delta\omega}\right) - 10 \cdot \log\left(\frac{P}{1\,\mathrm{mW}}\right) - PN\left(\Delta\omega\right)$

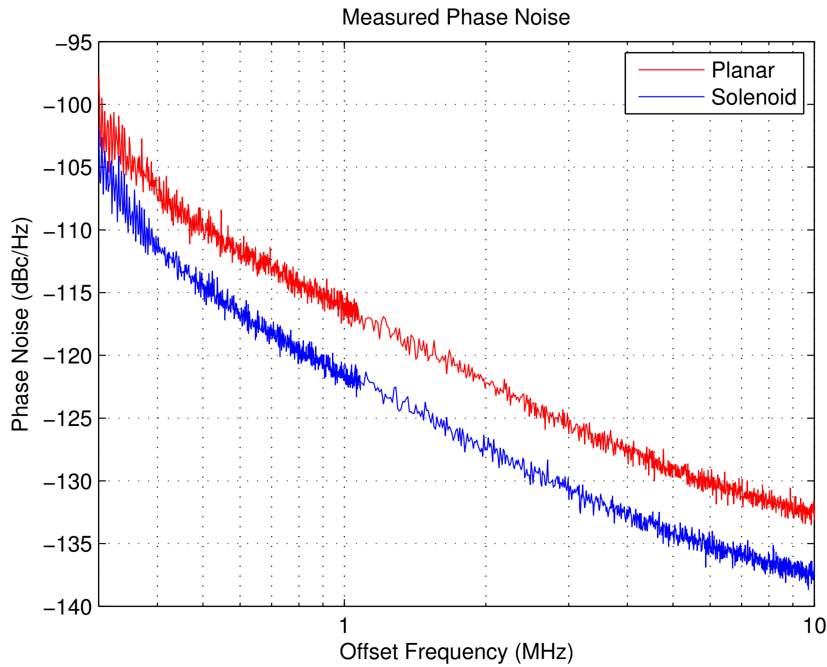


Figure 2-16: Measured phase noise of planar and solenoid VCOs

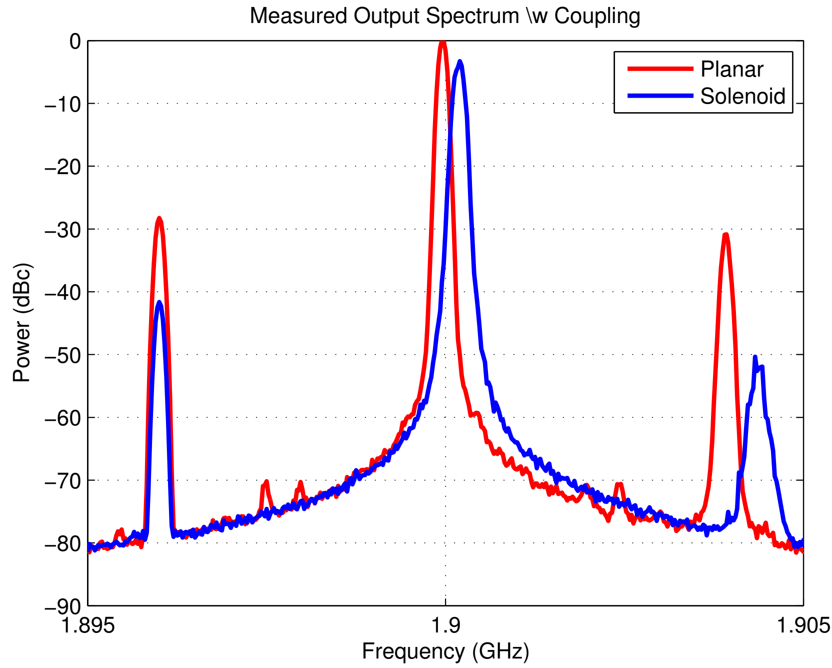


Figure 2-17: VCO output power spectral density with spurious tones caused by coupling from digital clock lines

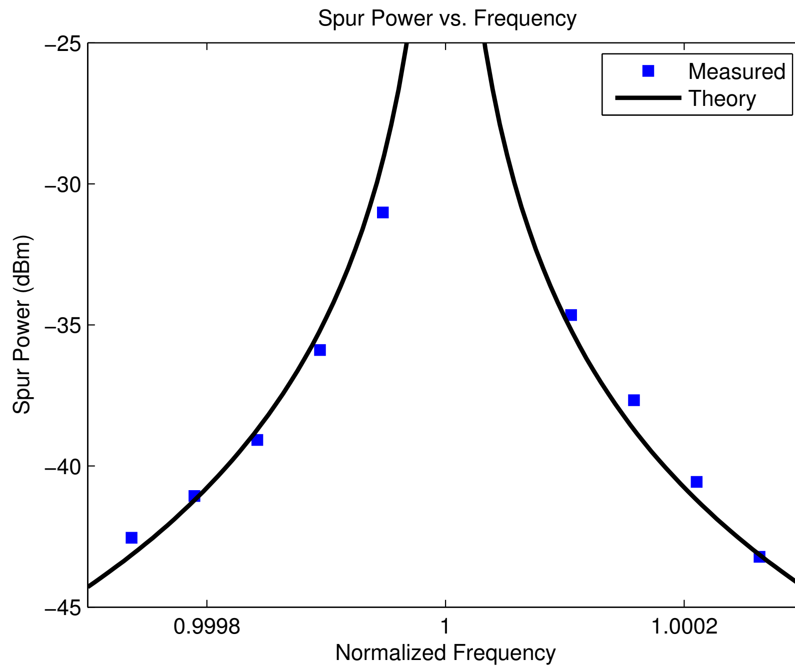


Figure 2-18: Clock signal spur power as a function of relative frequency to VCO free-running frequency

spectrum for both inductors while activating different clock lines on the bottom die tier. It can be observed that as we expect, the solenoid structure exhibits a consistently lower coupling to the clock lines than the planar inductor. This relation matches the general shape as the one observed in simulation shown in Fig. 2-11, though not precisely. The reason for this is that the spurs observed include the total coupling between the structures and therefore is only a proxy of the simulated values. It is still a reasonable assumption to consider the actual mutual coupling to have a linear relation to the simulated values, i.e.

$$M_{total} = \alpha M_{sim} + \beta. \tag{2.22}$$

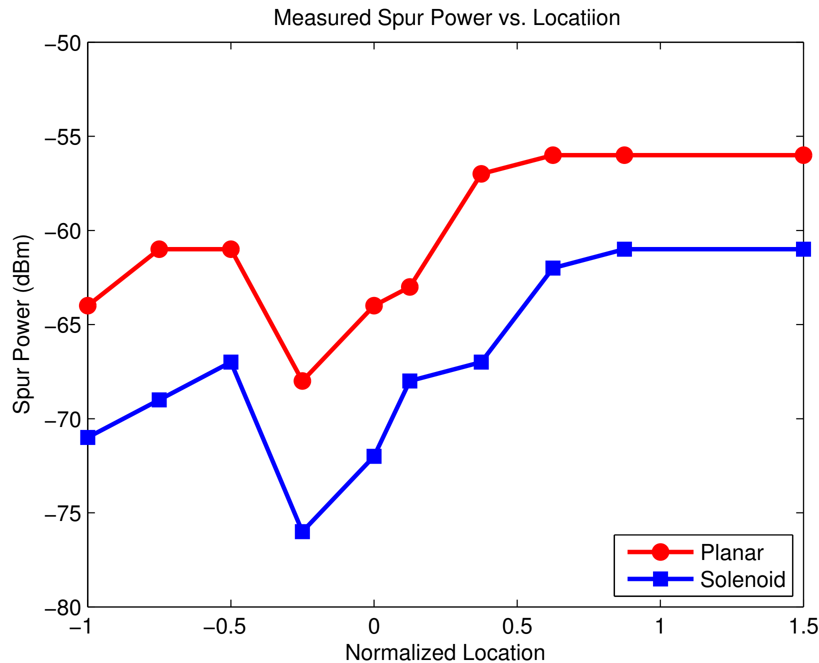


Figure 2-19: Clock signal spur power as a function of clock line location

From (2.8) we expect to observe a quadratic dependency between the mutual inductance and the output spur power. This relationship will remain true both in respect to the total mutual coupling as well as the simulated one through the relationship described in (2.22). Plotting the measured output spur power when using each clock line location as a function of the simulated mutual inductance for that location will result in a second order polynomial

behavior. This measurement is plotted in Fig. 2-20 along with the fitted second order polynomial which minimizes the least square errors from the measured data points. The R^2 correlation coefficient of the fit is also displayed and is equal to 0.97 and 0.96 for the planar inductor and solenoid inductor measurements respectively. These results indicate a good fit between the measured data points and the quadratic dependency assumption.

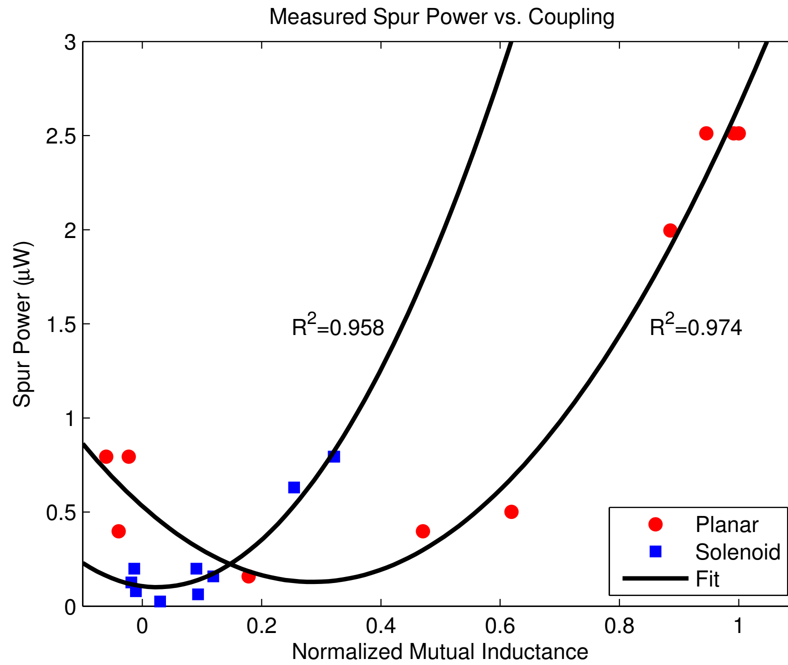


Figure 2-20: Clock signal spur power as a function of coupling strength

As previously discussed, all clock lines are driven in-phase, therefore we drive several lines simultaneously to obtain further data points with different coupling values. Assuming the previously derived relations hold, the aggregate effect will consist of a superposition of the individual contribution of each clock line, which would be a simple addition in the case of in-phase signals. Fig. 2-21 plots the measurement points obtained from driving several clock lines in parallel. Each point represents an additional line being activated adding to the overall mutual inductance. A second order polynomial is again fitted to the data points for both inductors. The resulting R^2 for both plots is approximately 0.96 demonstrating a good fit to the data. The coefficients of the curve are slightly different than those observed in Fig. 2-20 due to numerical inaccuracies as well as imperfect phase matching between

the different clock lines.

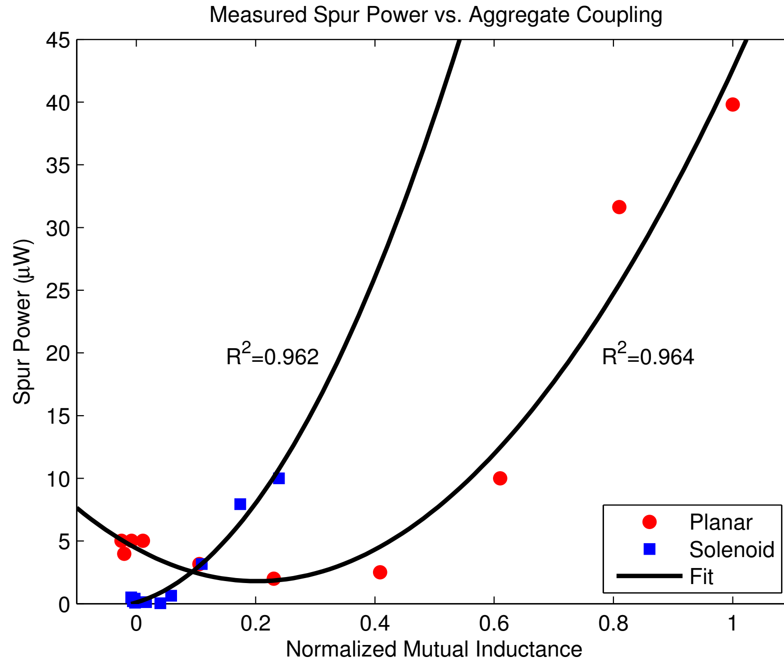


Figure 2-21: Clock signal spur power vs. aggregate coupling strength

The coupling between the clock lines and VCO tank is not limited to frequencies near the VCO operating frequency. Allowing for a slow speed 1 MHz clock to operate in the vicinity of the VCO may also couple and mix with the VCO output as illustrated in the phase noise measurement of Fig 2-22. Clear spurs can be observed in the phase noise spectrum of the planar inductor, while these are completely suppressed by the improved noise immunity of the solenoid.

As previously indicated, as the interfering clock line signal frequency approaches that of the VCO there is an increase in the coupling strength and spur power. If the clock signal power is strong enough, and the frequency close enough we will eventually observe injection locking and the VCO will begin to oscillate at the same frequency as the injected clock signal, i.e. it will no longer appear as an additional spur at the spectrum output of the VCO [121]. Fig. 2-23 depicts such a scenario where the planar inductor VCO has undergone injection locking and is pulled by the clock signal. The solenoid VCO's output spectrum, though highly corrupt, demonstrates a higher immunity to the injection locking

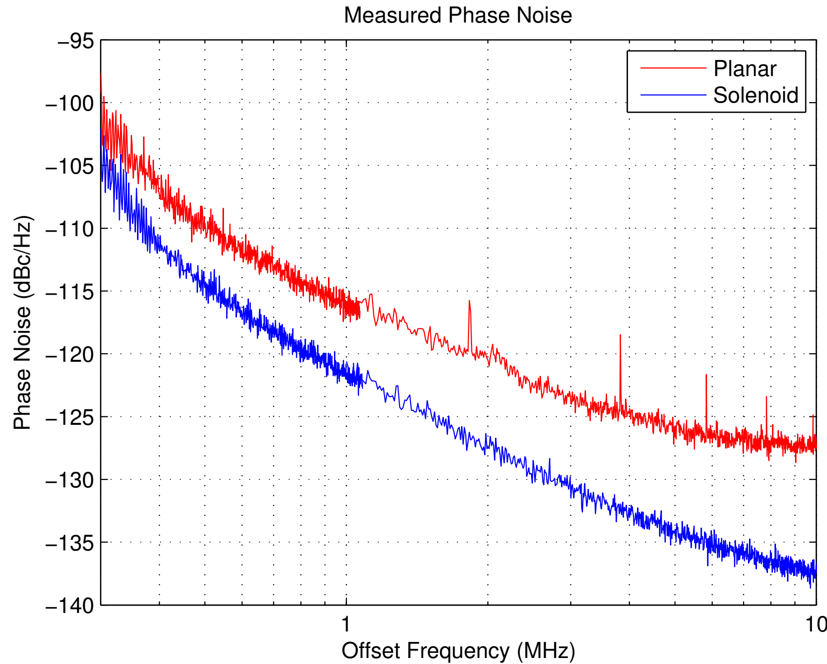


Figure 2-22: Phase noise measurement with noise coupling from a 1 MHz clock signal

and still exhibits oscillation around the tank's resonance frequency along with a spur at the clock frequency. The solenoid VCO is in a state of quasi-lock and therefore exhibits sidebands mostly above the injection frequency as predicted by [121].

2.5 Conclusion

In this chapter we reviewed some of the potential benefits and key challenges which arise when closely integrating sensitive analog blocks over noisy digital clock lines in a 3D stack. We began with a review of noise sources and impact in conventional 2D design. We reviewed existing techniques which aim to deal with such noise in integrated systems which include noise aggressors and sensitive blocks. We then concentrated on the implications of noise and isolation in a 3D stack and specifically reviewed how different structures couple differently due to the difference in the magnetic field line directions. Electrical models were presented to estimate the impact of coupling on the performance of a VCO. A vertical solenoid structure utilizing the interconnecting TSVs of the 3D stack itself was proposed

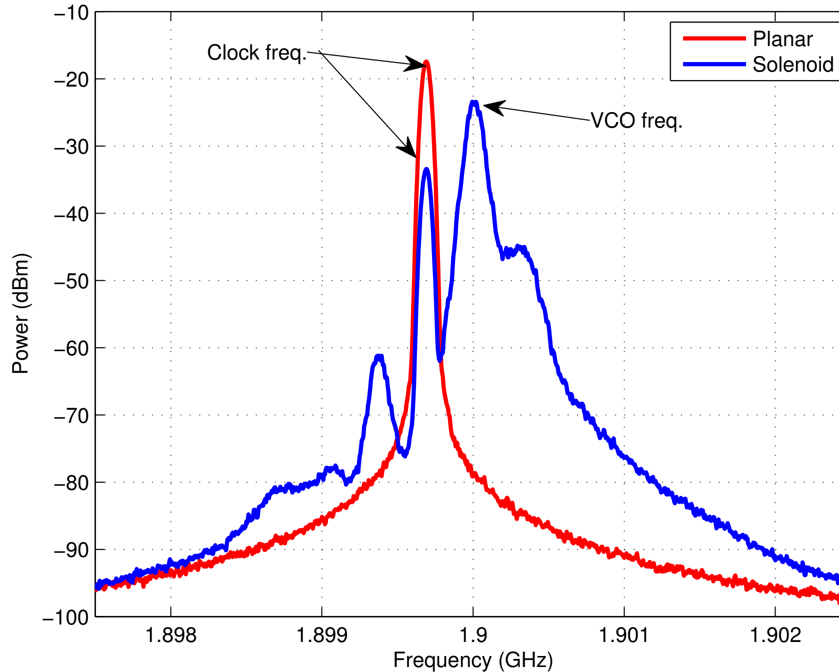


Figure 2-23: Output power spectral density of VCO with close-in clock line noise. Planar inductor VCO exhibits injection locking to the clock signal

to mitigate the effect of such coupling.

The presented solenoid vertical inductor for use in logic-RF 3D integration exhibits a 2x higher quality factor, 6 dB improvement in phase noise and an order of magnitude reduction in noise coupling compared to its conventional planar counterpart occupying the same area. The derived theoretical analysis demonstrates the impact of coupling on the VCO output and performance as well as the impact of clock signal frequency. The use of such structures allows more complex integration of future systems to maximize other benefits of 3D design in power and bandwidth while still maintaining acceptable circuit isolation and performance.

The ability to model, simulate and accordingly design structures to mitigate undesired coupling effects will enable further 3D integration of more complex systems achieving better functionality and energy efficiency. Addressing the new challenges arising from 3D integration will enable us to also benefit from the added opportunities it brings with it.

Chapter 3

Data-Dependent Signal Processing

3.1 Introduction

Next generation wireless mobile devices will be required to overcome challenges for achieving as much as a 1000x higher demand for network capacity [122]. Ubiquitous personal mobile communication have become a major driving force in the development of new communication standards. The demand for higher data rates, with high throughput and lower latency and higher mobility has led to the rapid evolution of the communication standards in use today. Such new communication standards, such as LTE attempt to address these growing needs with extensive solutions from the network layer down to the physical radio-communication layer.

In order to comply with the growing complexity of communication protocols, new hardware architectures must be used to keep up with the required computation complexity and speed as well as preform these in a low power budget to support low-power mobile devices. Efficient processing, based on the inherent properties of the communication protocol's signal statistics allows better utilization of the available hardware and achieves improved system efficiency. These issues are not specific to 3D integration, but in order to achieve our ultimate goal of creating an efficient RF system in 3D-IC we must optimize all aspects of the system. The digital portion is a dominant part of such systems and special care must be taken in its design in order to allow achieving overall system level benefits in the final design.

In this work we present an efficient hardware implementation of part of a digital baseband for an LTE User Equipment (UE) Uplink (UL) channel. We explore how to utilize specific data processing requirements in the LTE standard in order to design the processing hardware computation chain to support high data rates and throughput while minimizing the required power and area. We further take advantage of specific key properties of typical signals in LTE in order to further customize the design and make it more power efficient. An emphasis is given in the design to the implementation of the Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (IFFT) blocks due to their dominant role in the signal processing chain, as well as their general importance in many communication schemes and computation in general. This same design may be re-used with slight modifications to also act as the digital processor for the LTE DL since they contain many similar blocks.

The outline of this chapter is as follows. A brief overview of the LTE communication standard [123] is presented in section 3.2. Following is an overview of the FFT algorithm and its hardware implementations in section 3.3. The details of the implemented LTE transmitter baseband are given in section 3.4. Section 3.5 presents the measurement setup and results, as well as a comparison to other state-of-the-art equivalent implementations. We present our final conclusions in Section 3.6.

3.2 LTE

LTE is a standard for wireless communication of high-speed data for mobile devices. It was defined and developed by the 3rd Generation Partnership Project (3GPP) and is specified in its Release 8 document series 36 (frozen December 2008). The main motivation behind the development of LTE was the increasing demand for reliable, high-speed mobile communication, while reducing the network cost and complexity to improve upon the existing technologies of GSM/EDGE and UMTS/HSPA.

LTE includes many improvements and unique features across all layers of the network, we will focus on the features that comprise the lowest level - layer 1, the Physical Layer (PHY) [124]. The PHY of the Evolved Universal Terrestrial Radio Access (E-UTRA) is based on Orthogonal Frequency Division Multiple Access (OFDMA) communication for

the DL (from Base Station (BS) to mobile) and SC-FDMA for the UL (from mobile to BS). It supports both Frequency Division Duplexing (FDD) and Time Division Duplexing (TDD), modulation of up to 64-QAM and a scalable channel bandwidth of up to 20 MHz. Furthermore, the standard supports spatial multiplexing in the DL of up to 4×4 Multiple Input Multiple Output (MIMO) and also supports Multi-User MIMO (MU-MIMO) in the UL [125]. These capabilities allow for the desired high data rates of up to 300 Mb/s and 75 Mb/s in the DL and UL respectively, while maintaining high spectral efficiency.

LTE defines two frame structures, one used in FDD networks and the other in TDD networks. In our work we have focused on the implementation of a FDD system so we will focus our discussion on its implementation. Details on the TDD frame structure can be obtained from [126]. As seen in Fig. 3-1 the LTE radio frame's duration is 10 ms, and comprises of 10, 1 ms subframes, which each in turn consists of two 0.5 ms slots numbered from 0 to 19. The radio frame structure uses a basic sample unit size equal to $T_s = 1/(15000 \times 2048) \approx 32.55\,\text{ns}$.

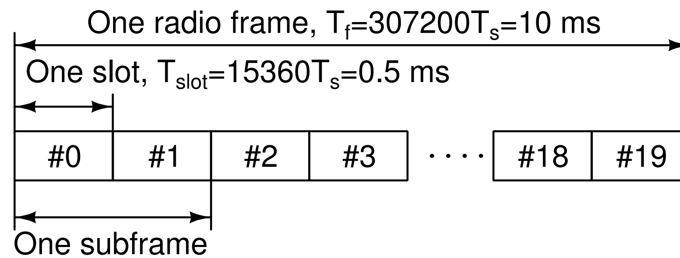


Figure 3-1: LTE FDD frame structure

Each transmitted slot comprises of several SC-FDMA symbols. A Cyclic Prefix (CP) is added to each symbol by replicating the tail portion of the symbol and attaching it to the beginning of the symbol. LTE supports two modes of CP addition, regular and extended. A schematic illustration of the slot structure for the two modes is shown in Fig. 3-2. In the regular mode there are a total of 7 symbols in each slot with a different CP extension for the first symbol and the other 6 symbols in the slot. When using extended CP mode the slot is built up by 6 symbols each having the same size extended CP addition at the head of each symbol.

The SC-FDMA symbols comprise of resource elements which are part of a time-frequency

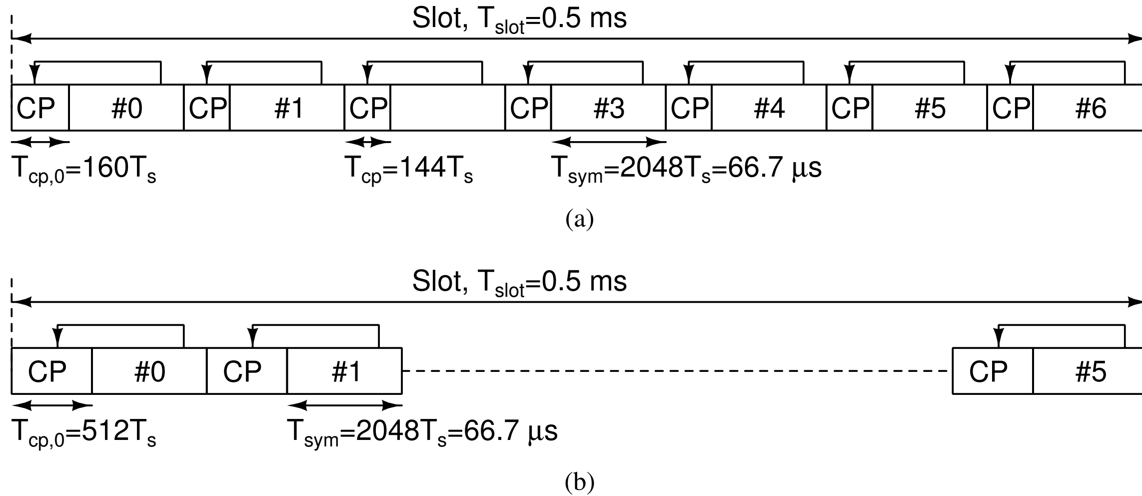


Figure 3-2: Time domain slot structure with (a) normal and (b) extended cyclic prefix addition to SC-FDMA symbols

Table 3.1: Transmitter available resource blocks for different channel bandwidths

Channel Bandwidth (MHz)	1.4	3	5	10	15	20
Transmitter Bandwidth Configuration (N_{RB})	6	15	25	50	75	100

resource grid. Fig. 3-3a illustrates the resource grid which comprises the available spectrum for transmission. A resource element represents a Subcarrier (SC) with 15 kHz bandwidth. 12 adjacent SCs in frequency along the period of one slot comprise a Resource Block (RB). The number of available resource blocks depends on the channel bandwidth and is summarized in table 3.1. Each UE is assigned to use a subset of the available resource blocks and those will be considered active. Fig. 3-3b illustrates the spectrum allocation in LTE.

In this work we have focused on realizing one of the key channels used for data transmission in LTE - the Physical Uplink Shared Channel (PUSCH). The baseband signal processed by the PUSCH is defined by the following steps:

- scrambling
- modulation of scrambled bits to generate complex valued symbols

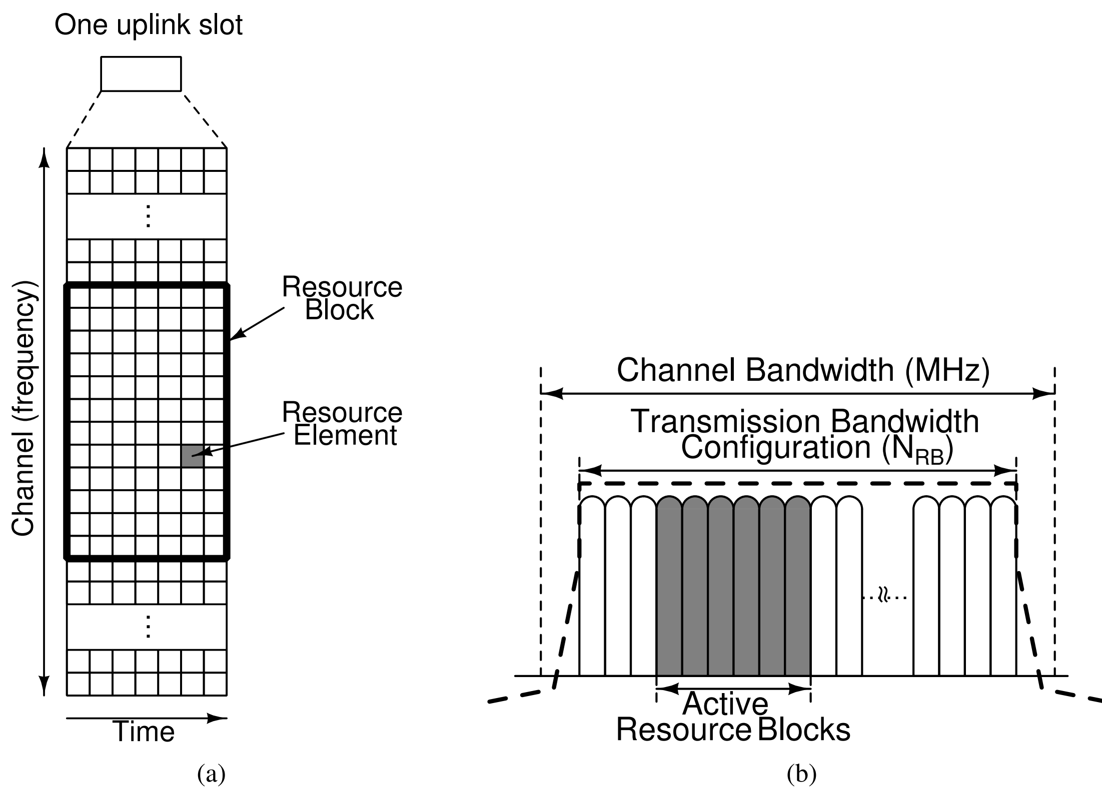


Figure 3-3: LTE uplink (a) resource grid and (b) channel configuration illustration

- mapping of complex valued symbols onto one or more transmission layers
- transform precoding to generate complex valued symbols (essentially a Discrete Fourier Transform (DFT) performed on the complex valued symbols)
- precoding of the complex valued symbols
- mapping of precoded complex valued symbols to resource elements
- generation of complex valued time-domain SC-FDMA signal for each antenna port (by adding a cyclic prefix and performing an Inverse Discrete Fourier Transform (IDFT) on the complex valued symbols)

In this work we will focus on signal generation for one transmission layer and one port, though the key ideas may be extended to achieve the full scale functionality if desired. Therefore, the steps of mapping the modulated symbols to transmission layers, precoding for spatial multiplexing and mapping to antenna ports are all trivial steps since we are dealing with only one layer. Fig. 3-4 summarizes the main steps in the signal generation process for the PUSCH. We will further focus our analysis on the modulation steps involved in the signal generation scheme and leave the details of the data coding steps outside the scope of this work. The interested reader may refer to [127] for further implementation details of the coding process.

Orthogonal Frequency Division Multiplex (OFDM) signals in general suffer from large Peak to Average Power Ratio (PAPR). Since the OFDM symbol comprises of a superposition of N sinusoids on different subcarriers. On average the emitted power is linearly proportional to N. However, the different signals may add up constructively such that the total *amplitude* is proportional to N and thus the power proportional to N^2. In the worst case we observe cases where the PAPR of the system increases with the number of subcarriers. This in turn requires a highly linear PA at the output in order to be able to transmit all required symbols properly without distortion. Linear PAs usually suffer from low efficiency and high power consumption [2]. If the PA is not sufficiently linear then the resulting distortion of the output will cause loss pf orthogonality between the subcarriers as well as out-of-band emissions and spectral regrowth.

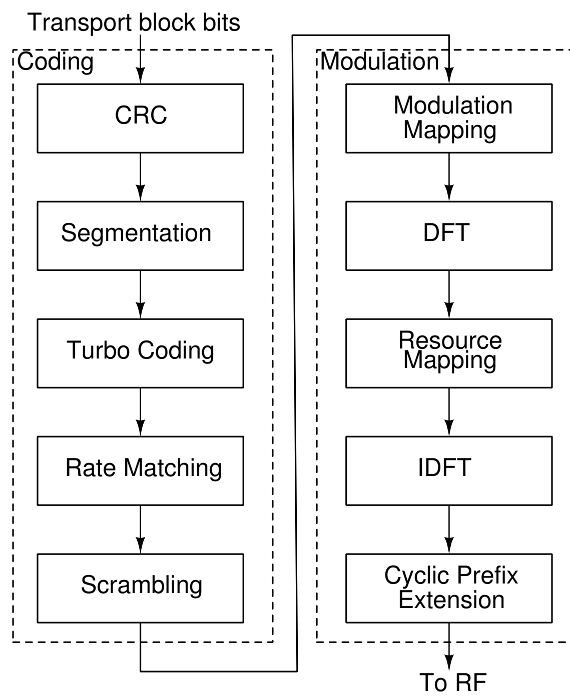


Figure 3-4: LTE PHY layer

The BS is less power constrained than the UE mobile units, therefore the added cost of implementing highly linear transmitters in the BS is acceptable and OFDM is used in the DL of LTE. However, to relax the requirements on the UE transmitter, a variant of this scheme is used - Single Carrier Frequency Division Multiple Access (SC-FDMA). SC-FDMA modulation adds an additional step of spreading the data across the spectrum before generating the orthogonal modulated subcarriers. This is achieved by performing a DFT processing step on the data before mapping to the subcarriers. The multiple access part is realized by assigning zero values to the subcarriers not used by the user and thus lowering the interference to other users. The reason to use such spreading of the data is that this results in an overall reduction of the PAPR [128]. This allows having a simpler, more power efficient PA in the mobile device which is inherently more power constrained. Furthermore, the spreading of the data is similar to Forward Error Correction (FEC) coding. If a subcarrier is affected by a deep fade in the channel response and is not received, the data is not completely lost, since it can be recovered by the added redundancy and coding used to spread the original data over many subcarriers which most of them arrive intact,

When allocating the used SCs over the entire bandwidth there are two main approaches. One is to allocate a continuous set of SCs to the user, and set all other values to zero as shown in Fig. 3-5a, this is known as Localized FDMA (LFDMA). A second approach is to distribute the used SCs over the entire bandwidth and interleave them with zero values as shown in Fig. 3-5b, this technique is known as Interleaved FDMA (IFDMA). IFDMA has a generally better PAPR than LFDMA, which in turn is still better than OFDMA. However, LFDMA offers better performance and overall higher system throughput by allowing dynamic, channel dependent scheduling of SCs to different users depending on the channel fading characteristics [129]. Therefore, localized SC-FDMA is the modulation scheme used for the UL in LTE.

3.3 FFT Overview

As described in the previous sections, the DFT plays a dominant role in the implementation of OFDM modulation as part of the LTE communication standard. The miniaturization

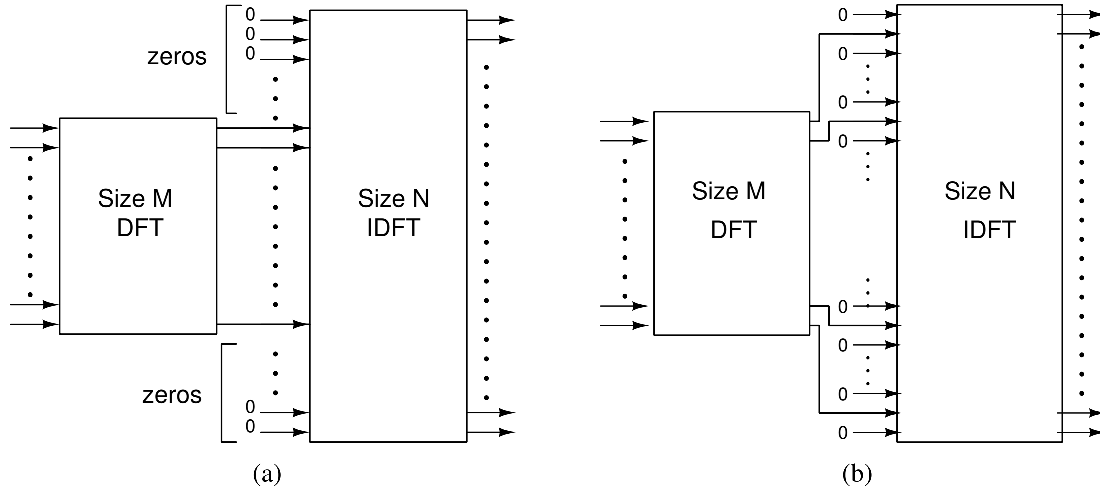


Figure 3-5: (a) Localized and (b) Interleaved SC-FDMA subcarrier mapping

of semiconductor devices allowing for small, efficient complex Digital Signal Processing (DSP), along with efficient algorithms derived for the computation of the DFT is what allowed these communication standard to go from theory to practice. In this section we give a brief overview of the DFT calculation in general and the FFT algorithm specifically for its efficient computation. We also consider several prevailing techniques for implementing the FFT computation efficiently in hardware and the various trade-offs between these approaches.

3.3.1 FFT Algorithm

Recalling the DFT analysis equation for a discrete signal of length N,

$$X[k] = \sum_{n=0}^{N-1} x[n] W_N^{nk} \tag{3.1}$$

Where

$$W_N = e^{-j\frac{2\pi}{N}}. \tag{3.2}$$

We observe that a direct calculation would require N multiplications and additions for each sample of the output. Since there are N output values (the same number as the input samples), the overall direct computation complexity appears to be $O(N^2)$.

The FFT is a well known algorithm derived in [130] for the efficient computation of

the DFT with a complexity of only $O(N \log N)$. One way to understand the operation of the algorithm is to consider a decomposition known as decimation in frequency. In this decomposition we separately calculate the even and odd frequency components of the DFT.

Decomposing (3.1) into its even and odd parts we obtain for the even components, where $k = 2k', \quad 0 \leq k' \leq \frac{N}{2} - 1$

$$\begin{aligned} X[2k'] &= \sum_{n=0}^{N/2-1} x[n] W_N^{2nk'} + \sum_{n=N/2}^{N-1} x[n] W_N^{2nk'} \\ &= \sum_{n=0}^{N/2-1} x[n] W_{N/2}^{nk'} + \sum_{n=0}^{N/2-1} x\left[n + \frac{N}{2}\right] W_{N/2}^{nk'} \\ &= \sum_{n=0}^{N/2-1} \left(x[n] + x\left[n + \frac{N}{2}\right] \right) W_{N/2}^{nk'} \\ &= \text{DFT}_{\frac{N}{2}} \left\{ x[n] + x\left[n + \frac{N}{2}\right] \right\} \end{aligned} \tag{3.3}$$

Similarly, for the odd components where $k = 2k' + 1$

$$\begin{aligned} X[2k'+1] &= \sum_{n=0}^{N/2-1} x[n] W_N^{n(2k'+1)} + \sum_{n=N/2}^{N-1} x[n] W_N^{n(2k'+1)} \\ &= \sum_{n=0}^{N/2-1} x[n] W_N^n W_{N/2}^{nk'} + \sum_{n=0}^{N/2-1} x\left[n + \frac{N}{2}\right] W_N^n W_N^{N/2} W_{N/2}^{nk'} \\ &= \sum_{n=0}^{N/2-1} \left(x[n] - x\left[n + \frac{N}{2}\right] \right) W_N^n W_{N/2}^{nk'} \\ &= \text{DFT}_{\frac{N}{2}} \left\{ \left(x[n] - x\left[n + \frac{N}{2}\right] \right) W_N^n \right\} \end{aligned} \tag{3.4}$$

We observe that the even components are simply the DFT of the new, half-length, series which adds the signal values offset by half of the total signal length. The odd components are similarly a DFT of half the size of the difference between the offset components, with a multiplication by a "twiddle" factor W_N^n. This result is what allows the efficient radix-2 calculation of the FFT algorithm which only has $N \log_2 N$ operations instead of N^2. This is due to the fact that we may repeat the above process, and calculate any N equals power-of-2 DFT by calculating the $N/2$ DFT on a slightly modified series down to the trivial case of

$N = 2$ where

$$X[0] = x[0] + x[1] \tag{3.5}$$

$$X[1] = x[0] - x[1] \tag{3.6}$$

An illustration of such a decimation in frequency for the case of 8 point DFT is shown in Fig. 3-6. The inherent "butterfly" structure is repeatedly used in many implementations of the FFT algorithm. Such a radix-2 butterfly is shown in Fig. 3-7.

3.3.2 General Radix FFT

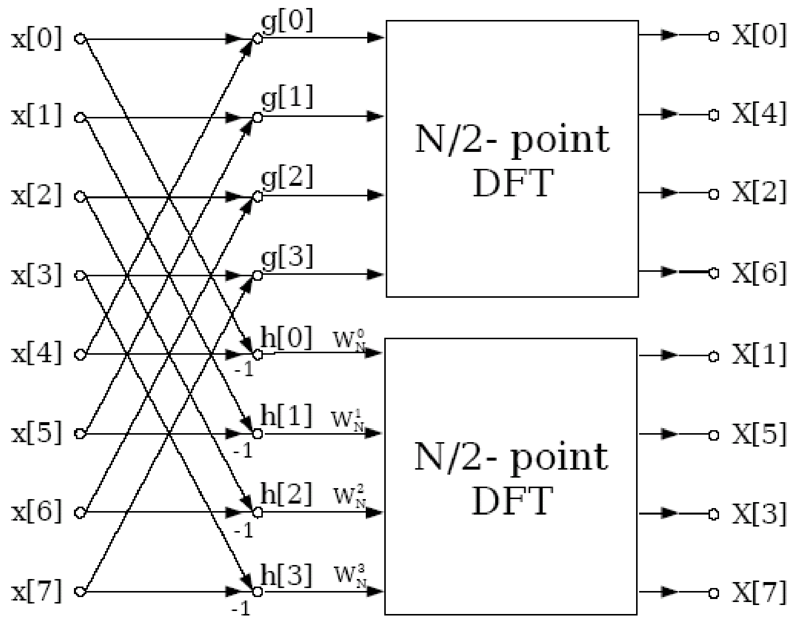
The radix-2 decomposition described previously can be extended to other radices as well in order to utilize the efficient calculation of the DFT for series of various sizes, which are not necessarily an integer power of 2. Assuming that a radix R is a factor of the number of points N, we decompose the DFT equation as follows

$$n = \left\langle \frac{N}{R} n_1 + n_2 \right\rangle_N \qquad 0 \le n_1 \le R-1, \quad 0 \le n_2 \le \frac{N}{R} - 1 \tag{3.7}$$

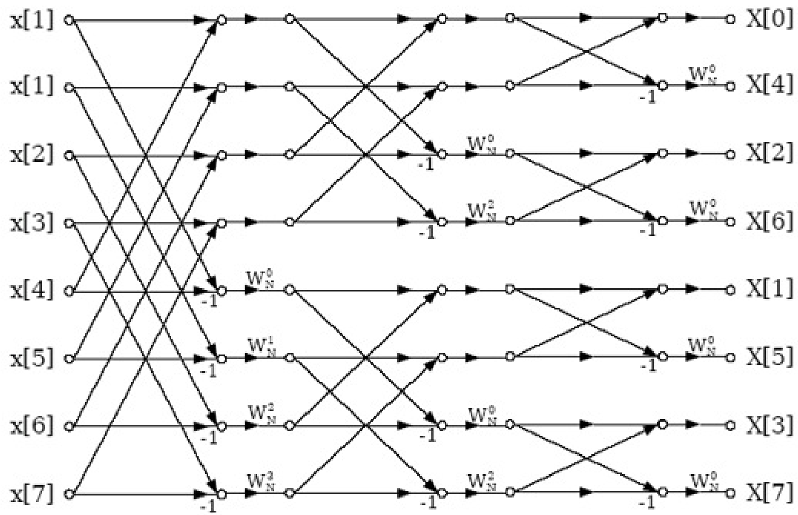
$$k = \langle k_1 + R k_2 \rangle_N \qquad 0 \le k_1 \le R-1, \quad 0 \le k_2 \le \frac{N}{R} - 1 \tag{3.8}$$

Plugging these into (3.1) we obtain

$$\begin{aligned} X[k_1 + Rk_2] &= \sum_{n_2=0}^{N/R-1} \sum_{n_1=0}^{R-1} x\left[n_2 + \frac{N}{R} n_1\right] W_N^{\left(n_2 + \frac{N}{R} n_1\right)(k_1 + Rk_2)} \\ &= \sum_{n_2=0}^{N/R-1} \left[\left[\left[\sum_{n_1=0}^{R-1} x\left[n_2 + \frac{N}{R} n_1\right] W_R^{n_1 k_1} \right] W_N^{n_2 k_1} \right] W_{N/R}^{n_2 k_2} \right. \\ &= \mathrm{DFT}_{\frac{N}{R}} \left\{ \underbrace{\left[\sum_{n_1=0}^{R-1} x\left[n_2 + \frac{N}{R} n_1\right] W_R^{n_1 k_1} \right]}_{\text{sum of rotated offset elements}} \underbrace{W_N^{n_2 k_1}}_{\text{twiddle factor}} \right\} \end{aligned} \tag{3.9}$$



(a)



(b)

Figure 3-6: Radix-2 decimation in frequency FFT

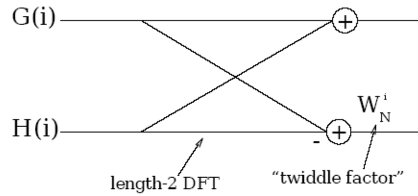


Figure 3-7: Radix-2 butterfly

Using the following shorthand

$$x_i = x\left[n + i\frac{N}{R}\right] \tag{3.10}$$

We rewrite (3.9) with i and r in the range $[0, R-1]$ as

$$X[Rk+r] = \mathrm{DFT}_{N/R}\left\{W_N^{nr}\sum_i x_i W_R^{ir}\right\} \tag{3.11}$$

Replacing $R = 2$ we will obtain (3.3) and (3.4) (for $r = 0$ and $r = 1$ respectively). A few other specific cases to be noted are for radix 3 which results in

$$\begin{aligned} X[3k] &= \mathrm{DFT}_{N/3}\{x_0 + x_1 + x_2\} \\ X[3k+1] &= \mathrm{DFT}_{N/3}\left\{\left(x_0 + x_1 W_3^1 + x_2 W_3^2\right) W_N^n\right\} \\ X[3k+2] &= \mathrm{DFT}_{N/3}\left\{\left(x_0 + x_1 W_3^2 + x_2 W_3^1\right) W_N^{2n}\right\} \end{aligned} \tag{3.12}$$

Realizing that

$$\Re\left\{W_3^1\right\} = \Re\left\{W_3^2\right\} \quad = -\frac{1}{2} \tag{3.13}$$

$$\Im\left\{W_3^1\right\} = -\Im\left\{W_3^2\right\} = -\frac{\sqrt{3}}{2} \tag{3.14}$$

will allow some simplification of the flow graph. The butterfly structure which satisfies these equations is shown in Fig. 3-8.

It is possible even to eliminate the complex multiplication in the radix-3 butterfly (as

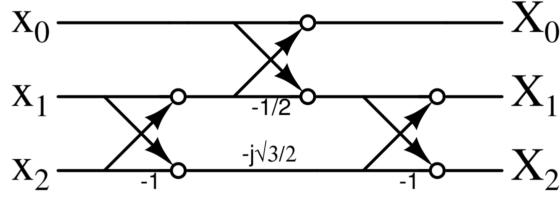


Figure 3-8: Radix-3 butterfly structure

well as theoretically in other radices) if we are willing to work with a different complex number system which includes the cube root of -1, and not the square root [131]. This however will require transitions back and forth from different bases depending on the FFT size so is likely to be more useful to cases where the FFT comprises of radix-3 units only.

For the case of radix 5 we evaluate (3.11) as

$$\begin{aligned}
X[5k] &= \mathrm{DFT}_{N/5}\left\{x_0+x_1+x_2+x_3+x_4\right\} \\
X[5k+1] &= \mathrm{DFT}_{N/5}\left\{\left(x_0+x_1W_5^1+x_2W_5^2+x_3W_5^3+x_4W_5^4\right)W_N^n\right\} \\
X[5k+2] &= \mathrm{DFT}_{N/5}\left\{\left(x_0+x_1W_5^2+x_2W_5^4+x_3W_5^1+x_4W_5^3\right)W_N^{2n}\right\} \\
X[5k+3] &= \mathrm{DFT}_{N/5}\left\{\left(x_0+x_1W_5^3+x_2W_5^1+x_3W_5^4+x_4W_5^2\right)W_N^{3n}\right\} \\
X[5k+4] &= \mathrm{DFT}_{N/5}\left\{\left(x_0+x_1W_5^4+x_2W_5^3+x_3W_5^2+x_4W_5^1\right)W_N^{4n}\right\}
\end{aligned} \tag{3.15}$$

Using the following relationships

$$\Re\left\{W_5^1\right\} = \Re\left\{W_5^4\right\} \quad = \cos\left(\frac{2\pi}{5}\right) \tag{3.16}$$

$$\Re\left\{W_5^2\right\} = \Re\left\{W_5^3\right\} \quad = \cos\left(\frac{4\pi}{5}\right) \tag{3.17}$$

$$\Re\left\{W_5^1\right\} + \Re\left\{W_5^2\right\} \quad = -\frac{1}{2} \tag{3.18}$$

$$\Im\left\{W_5^1\right\} = -\Im\left\{W_5^4\right\} = -\sin\left(\frac{2\pi}{5}\right) \tag{3.19}$$

$$\Im\left\{W_5^2\right\} = -\Im\left\{W_5^3\right\} = -\sin\left(\frac{4\pi}{5}\right) \tag{3.20}$$

we simplify the radix 5 butterfly structure to that shown in Fig. 3-9 [132]. Where we

denoted the coefficients

$$k_0 = \cos\left(\frac{4\pi}{5}\right) \tag{3.21}$$

$$k_{10} = -\sin\left(\frac{4\pi}{5}\right) \tag{3.22}$$

$$k_{11} = \sin\left(\frac{4\pi}{5}\right) - \sin\left(\frac{2\pi}{5}\right) \tag{3.23}$$

$$k_{12} = \sin\left(\frac{4\pi}{5}\right) + \sin\left(\frac{2\pi}{5}\right) \tag{3.24}$$

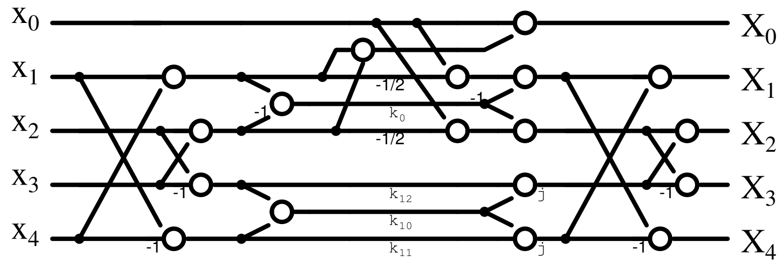


Figure 3-9: Radix-5 butterfly structure

In general, the radix need not be a prime number. For the case of $R = 4$ for example we will obtain

$$\begin{aligned}
X[4k] &= \mathrm{DFT}_{N/4}\{x_0 + x_1 + x_2 + x_3\} \\
X[4k+1] &= \mathrm{DFT}_{N/4}\left\{\left(x_0 + x_1W_4^1 + x_2W_4^2 + x_3W_4^3\right)W_N^n\right\} \\
X[4k+2] &= \mathrm{DFT}_{N/4}\left\{\left(x_0 + x_1W_4^2 + x_2 + x_3W_4^2\right)W_N^{2n}\right\} \\
X[4k+3] &= \mathrm{DFT}_{N/4}\left\{\left(x_0 + x_1W_4^3 + x_2W_4^2 + x_3W_4^1\right)W_N^{3n}\right\}
\end{aligned} \tag{3.25}$$

Using some simple substitutions as before, utilizing the relations

$$W_4^1 = -W_4^3 = -j \tag{3.26}$$

$$W_4^2 = -1 \tag{3.27}$$

We plot the radix-4 butterfly structure as shown in Fig. 3-10.

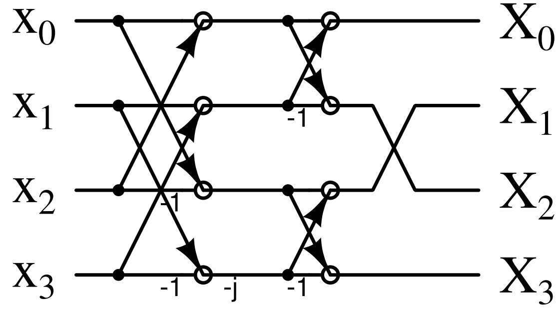


Figure 3-10: Radix-4 butterfly structure

An alternative way to write (3.25) would be in a nested fashion as abbreviated here

$$\begin{aligned} X[4k+r] &= \mathrm{DFT}_{N/4}\left\{\left(x_0 + x_1 W_4^r + x_2 W_4^{2r} + x_3 W_4^{3r}\right) W_N^{nr}\right\} \\ &= \mathrm{DFT}_{N/4}\left\{\left[\left(x_0 + x_2 W_2^r\right) + W_4^r\left(x_1 + x_3 W_2^r\right)\right] W_N^{nr}\right\} \\ &= \mathrm{DFT}_{N/4}\left\{\mathrm{BF}_{\mathrm{II}}^r\left(\mathrm{BF}_{\mathrm{I}}^r\left(x[n]\right)\right) W_N^{nr}\right\} \end{aligned} \tag{3.28}$$

Where

$$\mathrm{BF}_{\mathrm{I}}^r\left(x[n]\right) = x[n] + (-1)^r x\left[n + \frac{N}{2}\right] \tag{3.29}$$

$$\mathrm{BF}_{\mathrm{II}}^r\left(x[n]\right) = x[n] + (-j)^r x\left[n + \frac{N}{4}\right] \tag{3.30}$$

Similarly, for $N = 8$ we derive

$$X[8k+r] = \mathrm{DFT}_{N/8}\left\{\mathrm{BF}_{\mathrm{III}}^r\left(\mathrm{BF}_{\mathrm{II}}^r\left(\mathrm{BF}_{\mathrm{I}}^r\left(x[n]\right)\right)\right) W_N^{nr}\right\} \tag{3.31}$$

With the addition of

$$\mathrm{BF}_{\mathrm{III}}^r\left(x[n]\right) = x[n] + \left(\frac{1-j}{\sqrt{2}}\right)^r x\left[n + \frac{N}{8}\right] \tag{3.32}$$

This form, of nested butterfly structures is what enables the use of fewer twiddle multipliers in the FFT chain when utilizing power-of-two radices.

Table 3.2: Radix-2 FFT output bit reversal

Decimal	Binary	Bit Reversal	Reverse Decimal
0	000	000	0
1	001	100	4
2	010	010	2
3	011	110	6
4	100	001	1
5	101	101	5
6	110	011	3
7	111	111	7

3.3.3 Bit Reversal

As seen previously, if the input data is in order, the output of the FFT algorithm utilizing decimation in frequency is bit reversed. In a radix-2 algorithm this surmounts to simply reversing the bit order of the output index in order to obtain the correct output index. A simple illustration of this for the case of N=8 is shown in table 3.2.

For the case of a mixed radix FFT computation, the same basic principle applies. The output data is received in reverse digit order, where the digits are each represented in their corresponding basis along the FFT pipeline. If for example a 12 point FFT is calculated using two radix-2 and one radix-3 butterflies sequentially, we write the index using these bases. The first two digits will be in base 2 and the last will be in base 3. The reversed output will consist of numbers where the first digit is in base 3 and the following two are in base 2. In general a number is represented by N digits with a designated order like so

$$D \triangleq d_{N-1}d_{N-2}\ldots d_1d_0 \tag{3.33}$$

Each digit location i corresponds to a specific base r_i. Therefore we calculate the value

represented by this notation as

$$D = \sum_{i=0}^{N-1} d_i \prod_{k=0}^{i-1} r_k \tag{3.34}$$

Where the product for the Least Significant Bit (LSB) ($i = 0$) is defined as 1. For the most common case where all digit locations have the same base (as is usually the case in decimal or binary notation), i.e.

$$r_k \equiv r \quad \forall k \tag{3.35}$$

we simplify (3.34) to the more familiar form

$$D = \sum_{i=0}^{N-1} r^i d_i \tag{3.36}$$

Table 3.3 illustrates the digit reversal concept for the example discussed previously of a 12 point FFT output constructed with two radix-2 butterflies and one radix-3 butterfly. Note the change in the multiplicand of the digit based not only on its location but also on the preceding digit bases as implied by (3.34).

3.3.4 Hardware Implementation

The ability to create efficient, low power, low area FFT processors in CPUs, FPGAs and ASICs enabled the proliferation of advanced communication standards as WiFi and LTE. The implementation in hardware of the FFT algorithm heavily utilizes the decomposition described in section 3.3.1 in order to carry out the DFT calculation efficiently. Most common hardware implementations rely on two basic building blocks. One is a memory unit used to store the input and output data as well as the currently processed data and also acts as a time delay element. The other key building block is a core Processing Element (PE), usually in the form of a butterfly structure as described earlier. Different combinations of these building blocks enable creating several FFT processing architectures with varying degrees of area, power consumption, latency, throughput, block utilization and control complexity.

The most area efficient implementation can utilize only one butterfly (such as a radix-2 butterfly) as a PE and a large memory to store all data. The data is read from memory,

Table 3.3: Mixed radix digit reversal example

Decimal	Mixed Radix	Digit Reversal	Reverse Decimal
0	$0_2 0_2 0_3$	$0_3 0_2 0_2$	0
1	$0_2 0_2 1_3$	$1_3 0_2 0_2$	4
2	$0_2 0_2 2_3$	$2_3 0_2 0_2$	8
3	$0_2 1_2 0_3$	$0_3 1_2 0_2$	2
4	$0_2 1_2 1_3$	$1_3 1_2 0_2$	6
5	$0_2 1_2 2_3$	$2_3 1_2 0_2$	10
6	$1_2 0_2 0_3$	$0_3 0_2 1_2$	1
7	$1_2 0_2 1_3$	$1_3 0_2 1_2$	5
8	$1_2 0_2 2_3$	$2_3 0_2 1_2$	9
9	$1_2 1_2 0_3$	$0_3 1_2 1_2$	3
10	$1_2 1_2 1_3$	$1_3 1_2 1_2$	7
11	$1_2 1_2 2_3$	$2_3 1_2 1_2$	11

processed step by step by the butterfly and the intermediate calculation result is stored in memory as shown in Fig. 3-11. After completing $N \log N$ such steps the calculation is complete and is read out of memory before new data is loaded. This approach is very minimal in the required area consumption given that a highly dense memory unit may be used since we are using only one PE at any given time and spreading the calculation over time, thus paying with latency and throughput for this design. This approach has variations with memory units which have either multiple read/write ports [133], several single port memory banks to increase the calculation speed [134, 135] or additional high-speed cache as an auxiliary memory [136]. This methodology may also be extended to allow mixed radix calculation by adding different options to choose from in the PE, though the hardware utilization will be low since at any given step only one butterfly is used at a time [137–139].

An opposite approach to the FFT architecture would be to have a completely "flat", "rolled-out" style implementation. In this scenario the entire FFT flow graph structure is implemented as shown for example in Fig. 3-6 for an 8-point calculation. In this

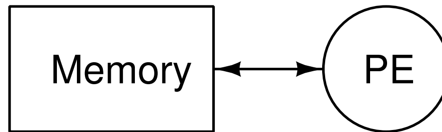


Figure 3-11: FFT memory architecture block diagram

implementation there is no need for any memory storage neither for the data nor for the twiddle factors. Furthermore, all multipliers may be hard-coded in the design and efficiently implemented. The calculation can be pipelined but it offers very high throughput and low latency compared to the memory architecture with only one PE. The main drawback of this approach is its large area consumption due to the $N \log N$ PEs required. Furthermore, The growth in wiring required as the number of points increases eventually causes large power consumption due to routing complexity and long data paths. These drawbacks make this approach suitable only when a relatively small FFT size is required.

Between these two extremes there is another set of popular FFT architectures which implement a pipeline solution. In this approach we break down the processing into several steps which are processed in parallel thus striking a balance between the two approaches obtaining both good timing and reasonable area and power consumption [140]. Fig. 3-12 illustrates the basic block diagram of the pipeline approach. Since we implement $\log N$ stages, then the entire calculation will require on the order of N cycles to compute the entire FFT.

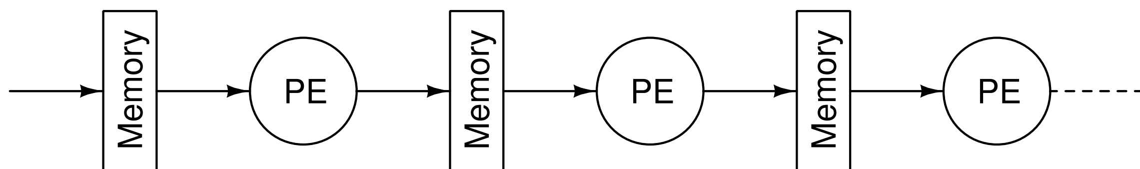


Figure 3-12: FFT pipeline architecture block diagram

Within the pipeline group of FFT implementations, several popular topologies exist. They are characterized by the radix used in the butterfly elements and the way data is propagated through the chain. Different topologies vary by the amount of memory required

to store data along the pipeline, the number of multipliers and adders required, their throughput and control complexity.

The Radix-2 Multi-path Delay Commutator (R2MDC) is one of the most classical approaches to the pipelined FFT. In this method, the input data is split into two parallel streams with the correct "distance" between samples and fed along the pipeline, entering the butterflies with the correct delays. Both butterflies and multipliers have a 50% utilization. $\log N - 2$ multipliers are required, $\log N$ butterflies and $3/2N - 2$ delay elements (corresponding to memory size) are needed. Fig. 3-13 illustrates the R2MDC pipeline for a 16 point FFT.

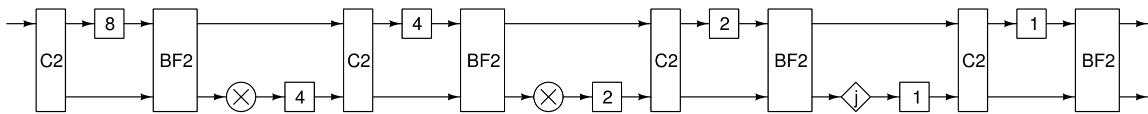


Figure 3-13: R2MDC block diagram for 16 point FFT

A different approach to R2MDC is Radix-2 Single-path Delay Feedback (R2SDF) [141], where the memory registers are used more efficiently as feedback elements to the butterflies, acting both as their input and output. A single data stream is fed into the pipeline. It has the same amount of multipliers and butterflies as R2MDC with the same 50% utilization, but requires only $N - 1$ memory registers, which is the minimal required memory for such a calculation. Fig. 3-14 illustrates the R2SDF topology for a 16 point FFT along with the counter based control signals.

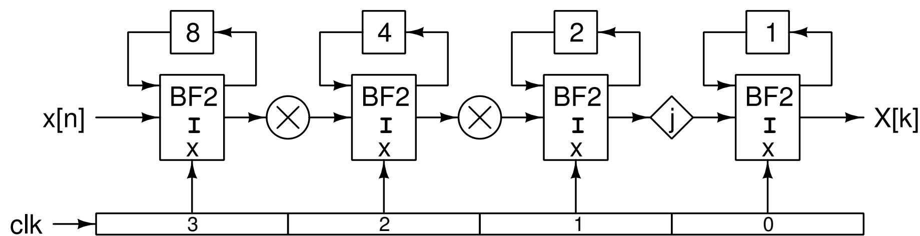


Figure 3-14: R2SDF block diagram for 16 point FFT

Each butterfly in the R2SDF topology is implemented as shown in Fig. 3-15. The control signal which dictates whether the butterfly will pass along the data for storage or

preform the desired butterfly calculation is an N bit counter, where each bit corresponds to a butterfly unit control signal. This results in the 50% utilization discussed previously.

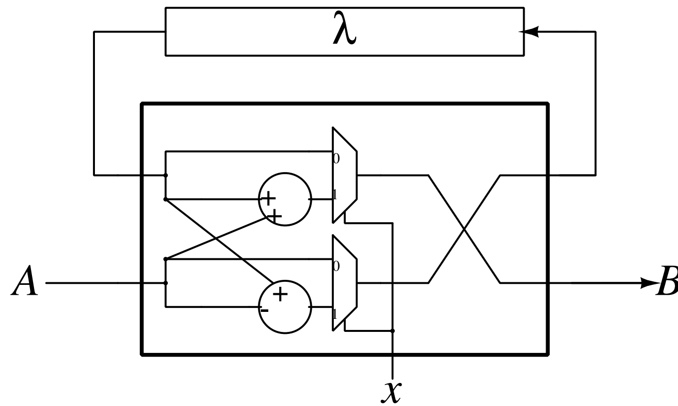


Figure 3-15: Radix 2 type I butterfly

To improve on the 50% utilization of the butterflies and multipliers we choose to use a higher radix for the design if possible. For example, using radix-4 instead of radix-2 will allows increasing the utilization of the multipliers to 75%. A R4SDF structure using radix-4 butterflies [142] is illustrated in Fig. 3-16. The revised butterfly uses a structure as shown in Fig. 3-10. The utilization of the butterflies however has dropped to only 25%, while increasing the complexity of the butterfly structure and its control scheme. As with R2SDF this topology requires only $N-1$ memory registers, and we are able to reduce the number of multipliers to $\log_4 N - 1 = {}^1\!/_2 \log_2 N - 1$ and the number of butterflies to $\log_4 N = {}^1\!/_2 \log_2 N$.

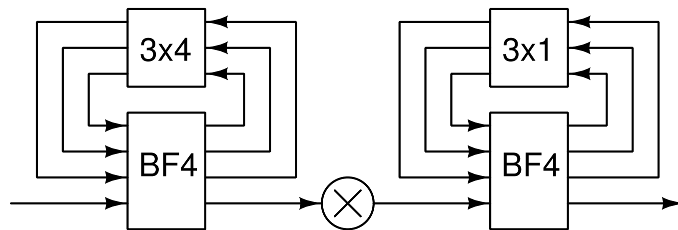


Figure 3-16: R4SDF block diagram for 16 point FFT

R4MDC [143] is the radix-4 version of R2MDC described previously. It has a parallel output of 4 bits leading to high throughput, however it suffers from a low utilization of 25%

of all components, which can be compensated only in specific scenarios where several FFTs need to be computed in parallel [144]. This topology requires $3/2 \log N$ multipliers, $1/2 \log N$ butterflies and relatively large memory requirement of $5/2 N - 4$ registers. Fig. 3-17 shows a block diagram of the R4MDC topology for a 16 point FFT.

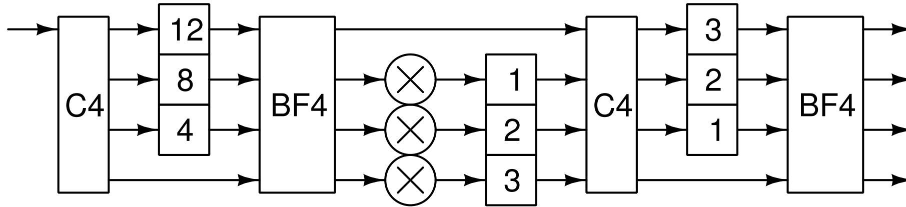


Figure 3-17: R4MDC block diagram for 16 point FFT

Radix-4 Single-path Delay Commutator (R4SDC) [145] uses a modified version of the radix-4 algorithm in order to mimic the R4MDC topology with a single data path. The loss in throughput is compensated by a much better multiplier utilization of 75% and a reduction in the required memory to $2N - 2$ registers compared with the R4MDC topology. Fig. 3-18 shows a highly simplified block diagram of the pipeline.

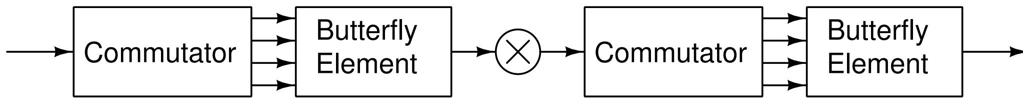


Figure 3-18: R4SDC block diagram for 16 point FFT

As we explored in section 3.3.2, we rewrite the radix-4 algorithm as a nested version of two radix-2 butterflies with a 90° rotation. This relationship was derived in (3.28) and we interpret (3.29) as the "regular" radix-2 butterfly as shown in Fig. 3-15 and using another butterfly structure as shown in Fig. 3-19 to satisfy (3.30). Using such a divide-and-conquer approach for cascading the pipeline allows construction of a new variant - $R2^2SDF$ [146]. This approach has a similar multiplicative complexity as the R4SDF topology, but with the same architecture and control complexity as the R2SDF approach. Since we are using the cascade of two butterfly structures we only require a non-trivial multiplication every other element resulting in only $1/2 \log N - 1$ multipliers as in R4SDF. Fig. 3-20 shows the block

diagram for a 16 point FFT using R2^{2}SDF along with the slightly modified counter based control signals.

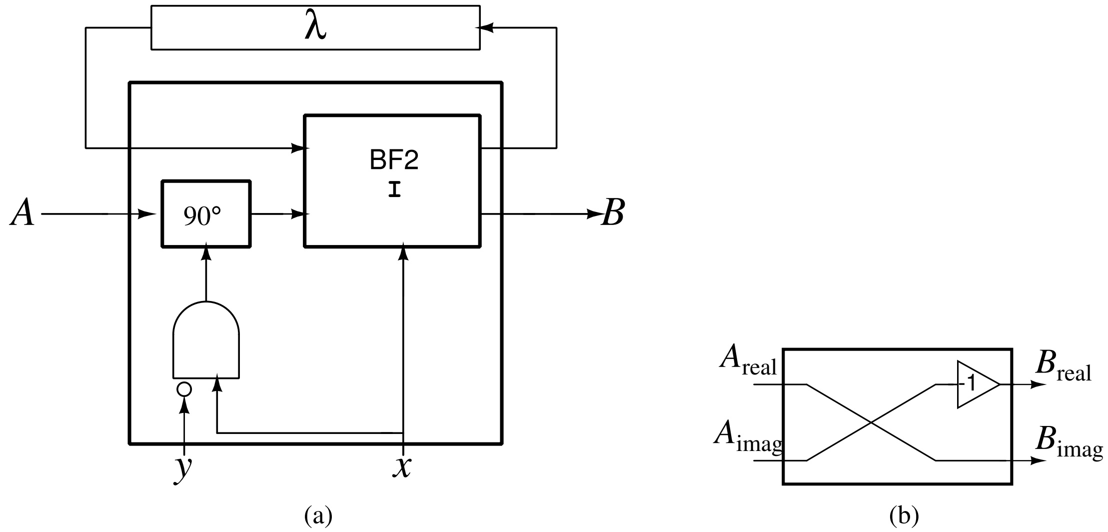


Figure 3-19: Radix 2 (a) type II butterfly and (b) 90° rotation block diagrams

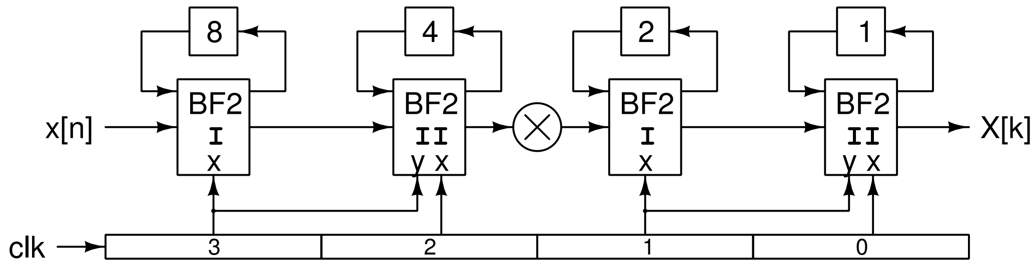


Figure 3-20: R2^{2}SDF block diagram for 16 point FFT

We further expand this concept as derived in (3.31) and add a third butterfly type which precedes a type II butterfly with a 45° rotator block as shown in Fig. 3-21. Cascading the 3 types of radix-2 butterflies will enable us to create a R2^{3}SDF topology. This will enable, as in the R2^{2}SDF and R2SDF approaches, to maintain the very simple, counter based control scheme and low memory requirement along with a reduction of the number of required non-trivial multipliers to $\log_8 N - 1 = 1/3 \log N - 1$. Fig. 3-22 shows a block diagram for a 64 point FFT using R2^{3}SDF along with the correspondingly modified counter based control signals. If the total FFT size is not divisible by 8 we use less butterflies as

needed converting them to other types as necessary.

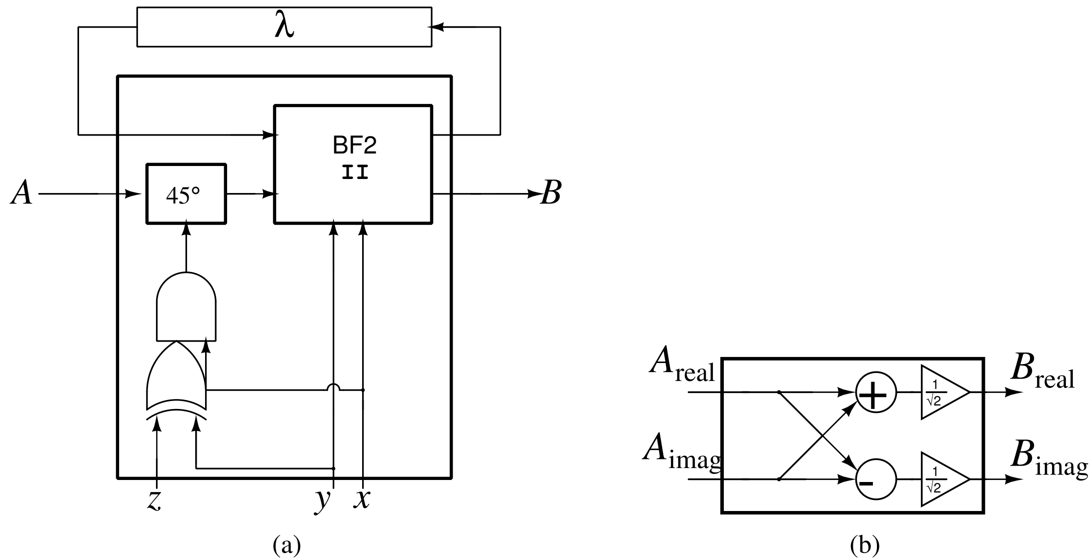


Figure 3-21: Radix 2 (a) type III butterfly and (b) 45° rotation block diagrams

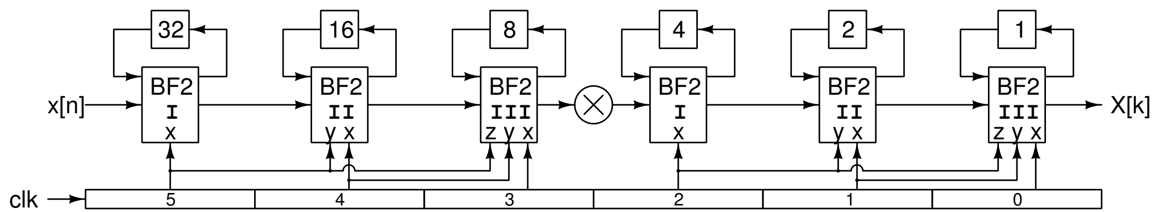


Figure 3-22: R2^{3}SDF block diagram for 64 point FFT

Another advantage of using such cascaded, high-radix pipeline approaches is that it is very easy to bypass parts of the pipeline as necessary in order to accommodate for different FFT sizes. It is also very simple to convert a butterfly from a higher order to a lower order. For example, to convert a type III butterfly to a type II butterfly we simply need to set the top two control signals to be equal ($z \equiv y$). By doing so we effectively disable the 45° rotator block and we are left with a regular type II butterfly. To convert from type II to type I, we simply set the appropriate control signal to a fixed high value ($y =' 1'$), in this manner again we disable the 90° rotator block and retain a type I butterfly. In all of these cases, the basic counter control signal remains the same and the FFT pipeline will perform as

Table 3.4: Pipeline FFT architecture resource comparison

Topology	# Multipliers	# Adders	Memory Size	Control
R2MDC	$\log N - 2$	$2\log N$	$\frac{3}{2}N - 2$	Simple
R2SDF	$\log N - 2$	$2\log N$	$N - 1$	Simple
R4SDF	$\frac{1}{2}\log N - 1$	$4\log N$	$N - 1$	Medium
R4MDC	$\frac{3}{2}\log N - 3$	$4\log N$	$\frac{5}{2}N - 4$	Simple
R4SDC	$\frac{1}{2}\log N - 1$	$\frac{3}{2}\log N$	$2N - 2$	Complex
$R2^2SDF$	$\frac{1}{2}\log N - 1$	$2\log N$	$N - 1$	Simple
$R2^3SDF$	$\frac{1}{3}\log N - 1$	$\frac{16}{3}\log N$	$N - 1$	Simple

expected for the revised calculation size. This makes this approach very robust and flexible with minimal overhead for design and control.

We are not limited to using a single butterfly type along the pipeline, as suggested by the $R2^2SDF$ and $R2^3SDF$ approaches, we create mixed radix FFTs by including various radix butterfly components [147, 148]. These will require slightly more complex control. The control will still be counter based, but since we are not using base 2, we cannot simply use the bit values of the counter binary output. The mixed radix approach allows creation of efficient FFT pipeline implementations for specific cases while maintaining high flexibility in determining the total FFT size as needed.

Table 3.4 summarizes the key attributes for the various FFT pipeline architectures. We can see that the high radix cascaded approaches, such as $R2^3SDF$ offer a promising direction using simple control and basic PEs along with a very low requirement for multipliers and memory. We will opt to use this approach for the implementation of the IDFT block in the SC-FDMA signal generation since it is required to support power of 2 calculations[1]. For the DFT block we will use a mixed radix version of the SDF pipeline architecture.

These architectures are obviously not the only ones in existence, but are a representation of the most popular and widely used architectures. Other variants of these approaches have been implemented previously, along with combinations of the described approaches

[1]We actually also need to support a 1536 point FFT which we will do by adding an optional radix-3 butterfly.

in order to gain the benefits from a particular approach while offsetting the drawbacks under certain scenarios [149].

3.4 Digital Baseband Implementation

The function of the baseband digital processing component is to create correctly modulated LTE SC-FDMA signals for transmission. The baseband operates in a pipeline structure continuously processing raw input data and forming modulated OFDM output symbols. A simplified flow graph is shown in Fig. 3-23 highlighting the key blocks of the system. A scan chain is used to load various control and configuration options to set the various parameters of the system and to demonstrate its flexibility for a wide range of signal scenarios. The system closely follows the signal generation chain for LTE as described in section 3.2 The raw data is stored in the main memory which can hold up to an entire radio frame of data of a 20 MHz channel. The data is then mapped to the appropriate Quadrature Amplitude Modulation (QAM) constellation and spread across the available user spectrum via DFT. Following, the data is mapped to the appropriate resource block locations and followed by orthogonal frequency duplexing via an IDFT. Finally, a cyclic prefix is added to the data and the entire signal is upsampled by a factor of 8 and interpolated between samples to improve the spectral output. The following subsections describe in detail the various blocks and their implementation details.

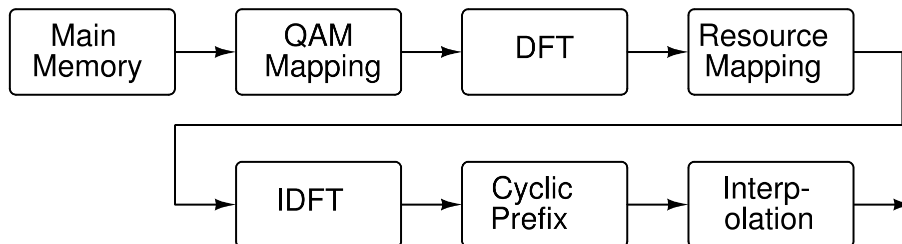


Figure 3-23: Baseband block simplified architecture

The entire signal processing chain operates in a pipeline fashion, timed such that a new value is sent to the DAC and RF blocks at each clock cycle from the output of the oversampling filter. This clock rate is adjusted to be eight times higher (due to the

x8 oversampling) than the symbol period required for the specific LTE bandwidth used. Therefore, e.g. for a 20 MHz channel, consisting of 2048 point IDFT of 15 kHz subcarriers (of which up to 1200 are occupied), a clock rate of 245.76 MHz is used. The data path itself, before the oversampling filter uses a clock which is derived from the main clock by a division by 8 (30.72 MHz in our previous example). If operating at a different channel size, and hence using less SCs, the clock is scaled accordingly. That is, for operation in a 5 MHz channel for instance, the datapath clock will be scaled by 4 to 7.68 MHz and the oversampling clock will be 61.44 MHz.

Since the clock rate is adjusted to match the output rate of the data points from the IDFT block we will encounter rate matching issues along the interface of the various blocks. Since, for example, both the DFT and IDFT blocks operate in a pipeline fashion, but consist of different size symbols, one would finish processing while the other is still working. We overcome this by careful matching of different clock rates for each block, but this will require very accurate skew and timing control, and the issue will also become more complicated since the block symbol size, and therefore the processing time, will vary greatly according to the specific operating scenario and RB allocation.

Alternatively, we employ a handshake technique using double buffers along the datapath to avoid accurate clock rate derivation and matching. The blocks' data rate increases as we proceed further down the pipeline, therefore each stage can indicate to the previous block whether it is done processing the current symbol and is ready to accept new data. A double buffer, consisting of two memory banks is used to store and time the data between two different rate blocks. Once the faster first block is finished writing the new output data it holds operation, until it receives indication from the next block that it has also finished and is ready to receive new data. The two blocks then switch memory bank usage in the double buffer, swapping the allocation of the memory for each block and repeat the process. The basic handshake process circuit schematic and timing diagram is illustrated in Fig. 3-24. This flow allows for the output to be always active at full rate with no gaps in the data output, and also allows for simple clock domain control since the data does not need to cross clock domain boundaries and further enables power saving by placing inactive blocks in sleep mode for lower power consumption when not in use.

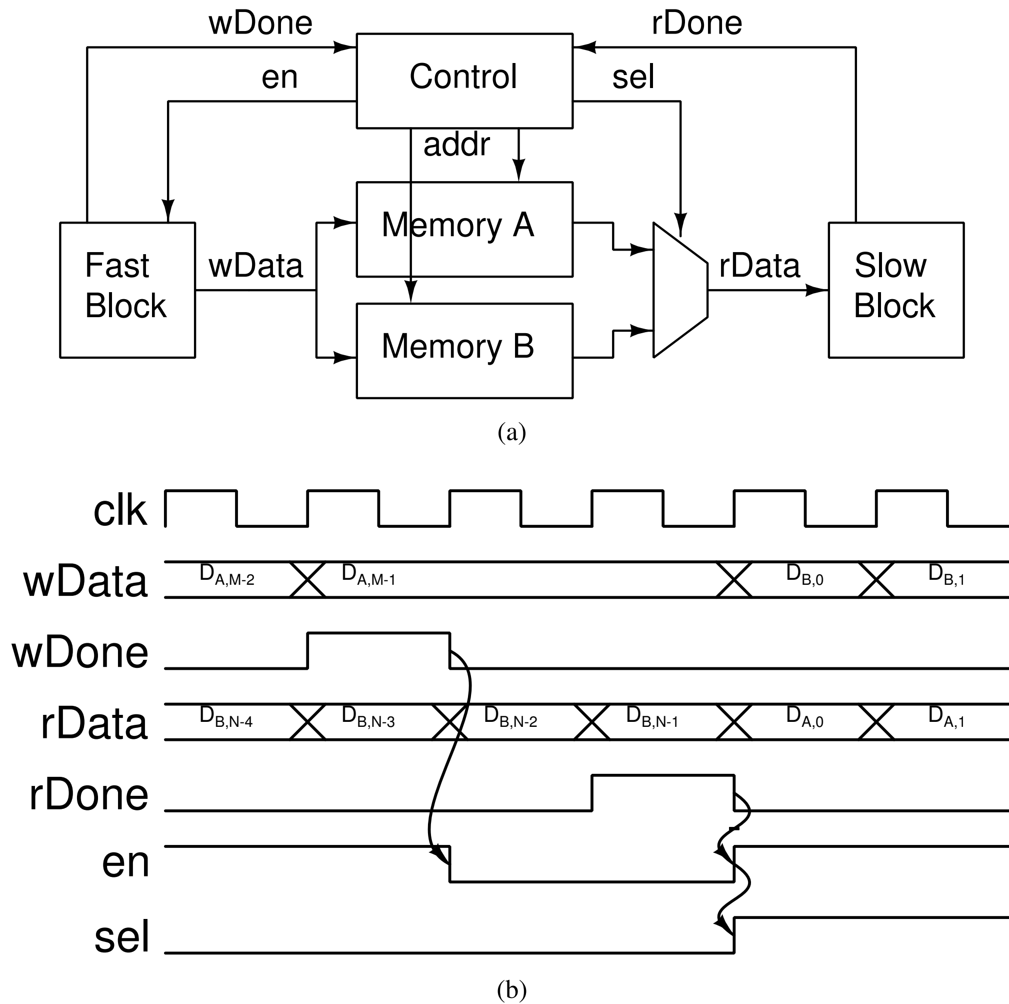


Figure 3-24: Double buffer handshake process (a) simplified schematic and (b) timing diagram

The digital block contains a total of 120 configuration bits loaded into the scan chain write blocks, with an additional 32 bits being read from the read blocks coming from the snapshot memory (see appendix B). Table B.3 summarizes the various scan chain configuration bits, their functionality in the signal processing block and their default values. The scan chain is loaded onto the chip using dedicated pads on die and controlled by an external FPGA communicating with a computer to set the various modes of operation.

3.4.1 Hardware Efficient Processing

Throughout the design of the digital block components we wish to utilize as much as possible efficient hardware implementation of the required signal processing, while also leveraging several of the key attributes specific to LTE signals. The datapath itself is designed in a pipeline fashion, reducing the complexity of necessary arithmetic operations required between adjacent sequential blocks. This allows reduction of the clock rate to the minimum required and allows for greater voltage scaling of the supply to reduce switching power losses [150–152]. Dynamic Voltage-Frequency Scaling (DVFS) utilizes such scaling in a dynamic fashion in order to account for temperature and process variations [153]. Furthermore, due to the handshake style operation described earlier, and since the entire chain is highly flexible allowing change of the symbol size to be processed, there are many times where certain blocks are not needed and are powered down. Clock gating is used extensively to reduce power losses and toggling to the unused blocks while still maintaining data retention in the memory and sequential elements to be used when operation in the blocks resume.

Noting the fact that the DFT operation performed on the LTE signals is restricted to a finite radix base of 2, 3 and 5 we chose to implement the DFT processing core as a pipeline stage consisting of these blocks with the ability to power down unused elements for different size calculations, thus optimizing the datapath and saving power when it is not required [139]. Similarly, the IDFT processor uses mainly radix-2 butterflies and a single radix-3 butterfly structure to support the 128-2048 along with 1536 point calculations.

A prominent aspect of the data being processed in LTE SC-FDMA signals is the fact

that there are many null entires in the data. These appear due to the multiple-access nature of the signal generation and the resource mapping utilized in the protocol. Therefore, we encounter many instances along the data path calculation where zero values are present in the calculation. Special indicator flags are introduced in the various processing blocks along the datapath in order to recognize whether one or more of the inputs is zero in order to enable bypassing of the block's operation and reduce overall power consumption [154]. This technique is utilized in both adders and multipliers along the path as well as in the CORDIC rotators (see section 3.4.3.4) and in the implementation of the resource mapping block prior to the IDFT processing (see section 3.4.4).

In general, a multiplication of two complex numbers would require two real additions and 4 real multiplications, since

$$(a+jb)(c+jd) = (ac-bd)+j(ad+bc) \tag{3.37}$$

. This calculation may be rewritten as

$$(a+jb)(c+jd) = [a(c+d)-d(a+b)]+j[a(c+d)+c(b-a)] \tag{3.38}$$

Note the recurring component in the real and imaginary parts which allows for the calculation of the multiplication using five additions and only three real multiplications.

In several instances along the data path it is required to perform multiplications of the data by a fixed value. In general, fixed value multiplication can be implemented more efficiently than a general purpose multiplier. The fixed value is represented in its binary form, and a series of bit shifts and adds are used to create the desired multiplication. An improvement upon this approach is gained by using the Canonical Signed Digit (CSD) representation [155] of the fixed value instead of its binary representation. CSD representation uses a ternary set instead of a binary one for number representation, therefore each bit location can accept one of three values instead of two - plus (+), minus(-) or zero (0). Which indicates whether the weighted binary value of the bit in that position should be added (+), subtracted (-) or not used (0).

For example, the integer value 15 is represented as the 4 bit binary value *1111*, which

will require 4 shifts and additions in a fixed multiplier.

$$a \times 15 = a \times 4'b1111 = (a \ll 3) + (a \ll 2) + (a \ll 1) + a \tag{3.39}$$

The same number can be represented as 15 = 16 - 1 = *+000-* in CSD notation. This in turn implies that we realize the same multiplier with one one subtractor instead of three adders.

$$a \times 15 = a \times (16 - 1) = (a \ll 4) - a \tag{3.40}$$

In general signed digit representation of numbers is not unique. However CSD notation implies several properties which ensure a unique representation for values. These also imply that no two consecutive digits are non-zero. On average, CSD notation includes 33% fewer non-zero bits than regular two's complement notation saving power and area in the constant multiplier blocks.

The CSD multiplication is used in particular in the radix-3 and radix-5 butterflies as well as in the output of the CORDIC rotator (see section 3.4.3.4) for the constant multiplications. The use of CSD enables reduction of the required multipliers in these instances to 4 or 5 shift and add operations instead of a full 11 bit multiplier, thus saving more than 50% of required operations.

3.4.2 QAM Mapping

In order to conform with the required modulation constellations available in LTE communication, the input bit stream was first mapped to one of several possible constellations. The mapping was performed to either BPSK (essentially no mapping), QPSK, 16-QAM or 64-QAM. Tables A.1-A.4 in appendix A detail the appropriate mapping from bits to symbols. The symbol in-phase and quadrature parts (i.e. real and imaginary parts) are normalized such that the average power over all constellation symbols is equal to 1. This value is calculated as:

$$P = \sqrt{\frac{1}{N}\sum_i I_i^2 + Q_i^2} \stackrel{!}{=} 1 \tag{3.41}$$

Where N is the number of constellation points and I_i and Q_i represent the constellation points' real and imaginary values. Applying odd positive values for each constellation point component we rewrite (3.41) as

$$P = \sqrt{\frac{4}{N} \sum_{\substack{x=1 \\ x \text{ odd}}}^{\sqrt{N}-1} \sum_{\substack{y=1 \\ y \text{ odd}}}^{\sqrt{N}-1} x^2 + y^2} \tag{3.42}$$

Combining the two sums and utilizing the formulas for the sum of square odd integers

$$\sum_{k=1}^{n} (2k-1)^2 = \frac{1}{3} n(2n-1)(2n+1) \tag{3.43}$$

We simplify the total power normalization factor to be

$$P_N = \sqrt{\frac{2}{3}(N-1)} \tag{3.44}$$

Plugging in the various constellation sizes we obtain the normalization values square root of 2, 10 and 42 for QPSK, 16-QAM and 64-QAM constellations respectively.

From table A.4 we can observe the pattern which maps each 6 bit codeword to a symbol. When considering only the symbol value numerator (without the normalization factor calculated previously), we observe that the Most Significant Bit (MSB) (bit 5) corresponds to the sign of the in-phase component, where a '0' value corresponds to a positive value and a value of '1' corresponds to a negative value. Similarly, bit 4 dictates the sign of the quadrature component of the symbol. Following, bits 3 and 1 combined determine the in-phase value and bits 2 and 0 determine that of the quadrature component. This observation allows us to create a block which performs constellation mapping for all modulation types with minimum hardware redundancy and minimizing memory and lookup tables.

The constellation mapping block diagram is shown in Fig. 3-25. The data in the main memory is stored in 24 bit wide words. Each line is read and fed to the constellation mapping block. According to the selected modulation scheme and QAM order the bit-select block translates any required bit values to a fixed 6 bit size code word. The way this is done

for different mappings is as follows. The codewords are mapped as

$$\text{BPSK:} \quad b_0 \mapsto b_0 b_0 0011 \tag{3.45}$$

$$\text{QPSK:} \quad b_1 b_0 \mapsto b_1 b_0 0011 \tag{3.46}$$

$$\text{16-QAM:} \quad b_3 b_2 b_1 b_0 \mapsto b_3 b_2 00 \bar{b}_1 \bar{b}_0 \tag{3.47}$$

Afterwards, the fixed length 6 bit pattern is mapped as described previously and scaled according to the constellation size following (3.44). The output of the mapping block is muxed between the mapped symbol and the sign-extended bare codeword directly from memory using an enable signal. The output symbol length comprises of a real and imaginary part, each 16 bits long, for a total of 32 bits.

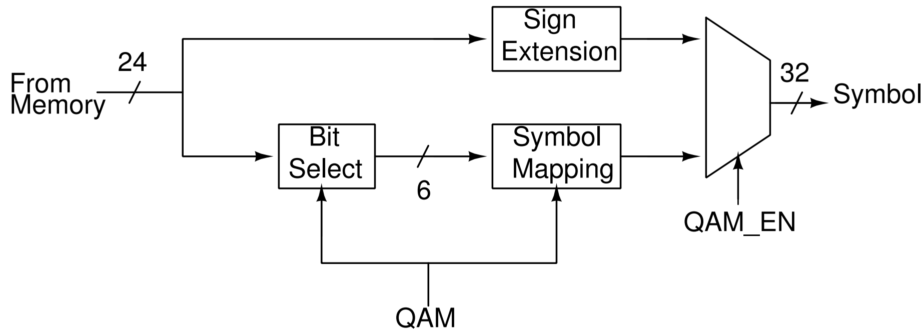


Figure 3-25: Constellation mapping block diagram

3.4.3 DFT

The DFT block is used to perform the spectrum spreading required in the LTE uplink. As mentioned previously, the block should support a variable size Fourier transform limited to the set defined by

$$\text{DFT Size} = \left\{ x | x = 12 \times 2^{\alpha} 3^{\beta} 5^{\gamma}, 12 \leq x \leq 1200,\ \alpha, \beta, \gamma \in \mathbb{N} \right\} \tag{3.48}$$

This corresponds to the fact the DFT size will be equal to the number of SCs in the uplink channel. Recalling that there are 12 SCs per RB and between 1 and 100 RBs.

Table 3.5: Valid DFT sizes for LTE UL symbol generation

RB	α	β	γ	Size	RB	α	β	γ	Size
1	0	0	0	12	30	1	1	1	360
2	1	0	0	24	32	5	0	0	384
3	0	1	0	36	36	2	2	0	432
4	2	0	0	48	40	3	0	1	480
5	0	0	1	60	45	0	2	1	540
6	1	1	0	72	48	4	1	0	576
8	3	0	0	96	50	1	0	2	600
9	0	2	0	108	54	1	3	0	648
10	1	0	1	120	60	2	1	1	720
12	2	1	0	144	64	6	0	0	768
15	0	1	1	180	72	3	2	0	864
16	4	0	0	192	75	0	1	2	900
18	1	2	0	216	80	4	0	1	960
20	2	0	1	240	81	0	4	0	972
24	3	1	0	288	90	1	2	1	1080
25	0	0	2	300	96	5	1	0	1152
27	0	3	0	324	100	2	0	2	1200

Observing this decomposition of the required DFT size we observe that we will only need to accommodate for radix 2, 3 and 5 in the DFT calculation. Table 3.5 lists all valid DFT sizes according to (3.48) along with the corresponding radix order.

3.4.3.1 Butterfly Structure

From table 3.5 we observe that using a pipelined DFT topology will require at most eight radix-2 butterflies, five radix-3 butterflies and two radix-5 butterflies. This also takes into account the fixed multiplication by 12 ($= 2^2 \times 3$) which is inherent in the fact that each RB consists of 12 SCs. The different radix butterflies were implemented as described in sections 3.3.2 and 3.3.4. Furthermore, when using a pipelined structure we utilize higher order butterfly radix-2 as detailed in the same section. The overall DFT pipeline structure is

illustrated in Fig. 3-26. All butterflies which are not needed for a specific DFT calculation size are bypassed and clock-gated.

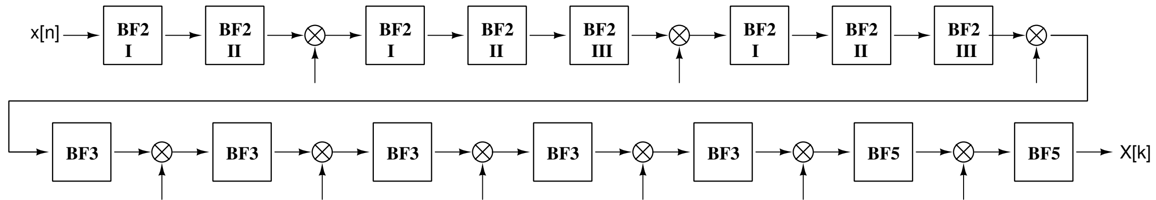


Figure 3-26: Variable size UL LTE DFT pipeline topology

The control signals for each butterfly are generated by simple counters as explained in section 3.3.4. The control signal is high for a duration equal to the butterfly's delay, and has a duty cycle ratio equal to the butterfly radix, i.e. the control signal will have a 50% (1 in 2) duty cycle for the radix-2 butterfly, 33% (1 in 3) for the radix-3 and 20% (1 in 5) for the radix-5 butterfly. For the radix 2^2 and 2^3 butterfly blocks, the control signals are also pipelined in order to maintain proper timing with the data path signals. An illustration of the control signals for different butterflies is given in Fig. 3-27.

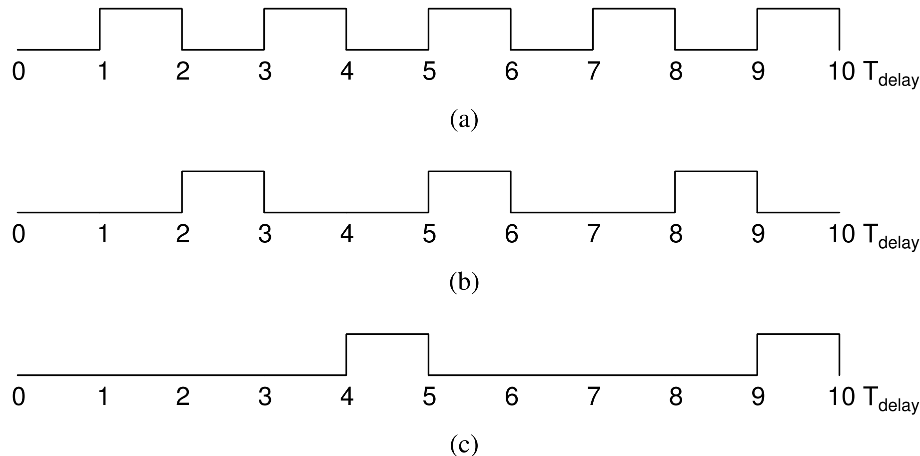


Figure 3-27: Control signal timing for (a) radix-2, (b) radix-3 and (c) radix-5 butterflies with a delay of T_{delay} clock cycles

3.4.3.2 Delay Lines

The butterfly delay lines may be implemented in several ways. One approach would be to use edge triggered flip flops acting as a shift register. Each clock cycle all samples will be shifted by one position along the shift register. This approach however simple, is very inefficient, although only one new sample is added to the queue and only one sample is read out to be processed, all delay line elements are active in every cycle consuming switching power. Another approach would be to use a form of Random Access Memory (RAM). In this approach a read and write counter increment every cycle in order to store and retrieve the appropriate symbols each cycle without changing the state of all other storage elements. The counter cycle size will determine the length of the actual delay obtained.

One way to implement the RAM delay line is with the use of Static Random Access Memory (SRAM) cells. The advantage of using SRAM cells as memory storage elements is that they are fairly dense with low area requirements and optimized for the given technology process. The disadvantage of using them is their fixed area overhead (making them more attractive for larger memory sizes than small delays), their power consumption and limited ability to operate at scaled down supply voltages. Usually SRAM cells require a minimum supply voltage of around 0.7 V. In order to accommodate the requirements to use them as a delay line, we require both read and write operations to occur simultaneously at each clock cycle. This can be achieved either by using a dual port SRAM or by using two single port SRAMs. Single port SRAM generally has a smaller footprint and lower power consumption and since in our scenario we ensure that we will not need to read and write to the same memory at once we utilize this split approach.

Fig. 3-28 depicts the schematic of a delay line unit implemented using two single port SRAM blocks. The incoming data is fed to both memory blocks, a counter (equal in length to the desired delay) is incremented each cycle. The LSB of the counter is used to select which memory block to write the data to, and accordingly read from the other. The rest of the counter bits are used to form the write address (which will change only every other cycle due to the truncation of the LSB). The read address is derived from the write address by adding the fixed value of 3 and again removing the LSB. The addition of 3 ensures

that we first read a value before it is overwritten. The counter LSB value also acts as a write-enable signal for one memory block and its inverse to the other. The output of the two memories are muxed and the output is selected again according to the counter LSB which toggles every clock cycle between one memory block and the other.

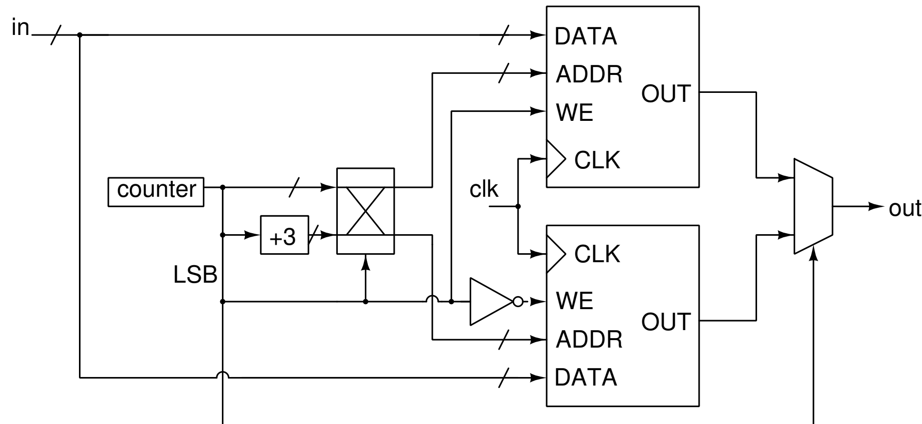


Figure 3-28: SRAM based delay line schematic

As described previously, SRAM based delay lines allow for a relatively compact and area efficient implementation of the delay operation for longer delays requiring large memory. However, for shorter delay lines other approaches are more power efficient. One other such approach is the use of a register file to act as the basis for the RAM operation. One may use storage elements such as Flip Flops (FFs) and use decoders to enable a specific element according to a given address. A more power efficient solution would be the use of transparent latches instead of FFs. Basing the delay line on latch memory units will not be as area efficient as SRAM based memory, but will allow for greater utilization of DVFS due to their ability to operate at much lower supply voltages [156].

A schematic of the latch based delay line cell is depicted in Fig. 3-29. A one-shot counter toggles every cycle ensuring that only one latch output is read each time by using a combination of AND and OR gates with the latch and counter output. A delayed version of the counter is used as the write-enable signal to the latch units, such that each cycle the data is written to the latch which was read the previous clock cycle.

The point at which one approach becomes more efficient, in the sense of their area-power

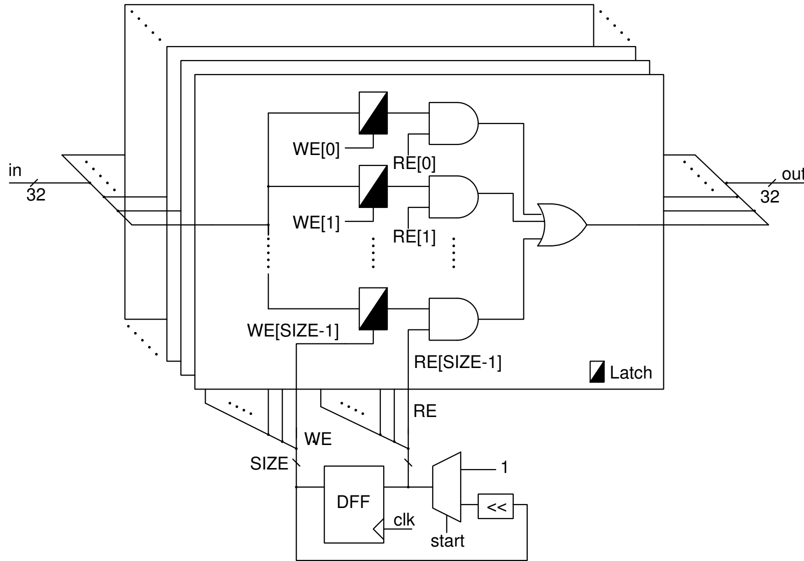


Figure 3-29: Latch based memory schematic

product, varies depending on the specific technology. In our process, similarly to other reported results [157], it appears that the SRAM based delay line is favorable for delay buffer lengths of 512 and above. In our DFT block implementation using memory sharing as will be explained in section 3.4.3.3, the largest delay line used is of length 300. Therefore we chose to implement all delay lines as latch-based memory cells to allow lowering of the supply voltage as needed and enable further power savings.

3.4.3.3 Memory Sharing

Each butterfly unit in the FFT pipeline requires a memory bank in order to preform the required shift delay of the Serial Delay Feedback (SDF) algorithm. The total memory requirement for each size N FFT is N-1 symbols. Table 3.6 lists the required memory size for each butterfly in the pipeline for each of the valid 34 possible FFT sizes. It should be noted that the radix-3 butterflies use twice the memory size written due to its structure, and the radix-5 butterflies use four times the size written. Indeed each row of the table sums to one less of the according FFT size (after accounting for the radix-3 doubling and radix-5 quadrupling as mentioned). However, if we design the pipeline to have memory allocated to each butterfly so as to meet its maximum memory requirement across all possible scenarios

(as shown on the last row of the table), the overall memory allocated would consist of 26 memory banks ($= 8 \times 1 + 5 \times 2 + 2 \times 4$) with a total size of 4271 symbols. This is more than 3.5 times larger than the theoretical minimum required size of 1199 (for the largest FFT size of 1200), and almost two times more separate memory banks required from the maximum of 14 (for the case of size 900 and 1200 points).

On the other extreme, we imagine a solution which indeed includes only the bare minimum of required memory banks and size, i.e. 14 banks with a total size of 1199. We will then require a very large crossbar matrix to map the possible 26 butterfly inputs/outputs to the appropriate memory banks which needs to be reconfigured for every FFT size as well as every clock cycle to manage the size allocation. This solution will dramatically increase the control complexity of the algorithm and will require a substantial amount of routing and buffering due to the large number of possible combinations.

Between these two extreme options we consider several alternatives which will reduce the amount of memory required on the one hand, but will not incur an overly complicated control scheme in order to manage it on the other. First we note that not all butterflies are active and require memory simultaneously. We identify groups of butterflies which their operation is mutually exclusive in some or all cases. It is possible to construct a system where the memory is shared within such mutually exclusive groups. In order to simplify the memory addressing control scheme we will allow neighboring radix-2 butterflies to share their memory, such that each butterfly can utilize the memory of up to two preceding butterflies.

The memory used in the FFT block is all latch based due to each banks' relatively small size (not exceeding 300 symbols as we will see shortly). Using this control scheme we design the memory banks as follows. The first radix-2 butterfly (R_8^2) has no preceding neighbors so it must support the maximum memory requirement, which is 384 symbols (for N=768). R_7^2 has a maximum requirement of 576 symbols (for N=1152), however in that case R_8^2 is not in use and its memory may be used, therefore we only need the memory to have 192 symbols. This value is also sufficient for cases where both R_7^2 and R_8^2 are used simultaneously. Similarly, the next butterfly R_6^2 will have a memory size of only 288 symbols, using an additional 192 from the previous memory bank when needed to reach a

Table 3.6: FFT butterfly memory requirements

Size	R_1^5	R_2^5	R_1^3	R_2^3	R_3^3	R_4^3	R_5^3	R_1^2	R_2^2	R_3^2	R_4^2	R_5^2	R_6^2	R_7^2	R_8^2
12			1					3	6						
24			1					3	6	12					
36			1	3				9	18						
48			1					3	6	12	24				
60	1		5					15	30						
72			1	3				9	18	36					
96			1					3	6	12	24	48			
108			1	3	9			27	54						
120	1		5					15	30	60					
144			1	3				9	18	36	72				
180	1		5	15				45	90						
192			1					3	6	12	24	48	96		
216			1	3	9			27	54	108					
240	1		5					15	30	60	120				
288			1	3				9	18	36	72	144			
300	1	5	25					75	150						
324			1	3	9	27		81	162						
360	1		5	15				45	90	180					
384			1					3	6	12	24	48	96	192	
432			1	3	9			27	54	108	216				
480	1		5					15	30	60	120	240			
540	1		5	15	45			135	270						
576			1	3				9	18	36	72	144	288		
600	1	5	25					75	150	300					
648			1	3	9	27		81	162	324					
720	1		5	15				45	90	180	360				
768			1					3	6	12	24	48	96	192	384
864			1	3	9			27	54	108	216	432			
900	1	5	25	75				225	450						
960	1		5					15	30	60	120	240	480		
972			1	3	9	27	81	243	486						
1080	1		5	15	45			135	270	540					
1152			1	3				9	18	36	72	144	288	576	
1200	1	5	25					75	150	300	600				
Max	**1**	**5**	**25**	**75**	**45**	**27**	**81**	**243**	**486**	**540**	**600**	**432**	**480**	**576**	**384**

capacity of 480 symbols. Continuing with this method we allocate further memory banks to R_5^2, R_4^2, R_3^2, R_2^2 and R_1^2 of 240, 216, 300, 270 and 243 symbols respectively.

The radix-3 butterflies reuse the already existing memory banks which were designed earlier since they usually do not operate together with the earlier radix-2 stages. Apart from the last radix-3 butterfly (R_1^3) which is always on, and therefore requires its own separate memory bank of 2×25 symbols, the others may share. Therefore, R_2^3 utilizes the same memory bank as R_8^2 and R_3^3, R_4^3 and R_5^3 can share with R_7^2, R_6^2 and R_5^2 respectively. Since the radix-3 butterflies require two memory blocks for each butterfly, the corresponding memories will be split into two halves, each having their own read/write access ports. A small correction must be made for R_2^3 due to the special case of an FFT of size 1152. In this case R_7^2 is using both its memory as well as the one allocated for R_8^2 therefore R_2^3 can not use it and requires a small extension memory of 3×2 symbols to accommodate for this specific case.

For the last two radix-5 butterflies, dedicated latch based memory blocks were used since the memory requirements for them are minimal - 5×4 and 1×4. This was also done to avoid further complication of the control scheme and memory addressing.

Table 3.7 summarizes the details of the memory banks implemented in the FFT design along with their size and which butterfly blocks use them. Each symbol represents a 16 bit word length complex number, therefore each word line is 32 bits long. The total memory size used is 2213 symbols. This is still about 84% higher than the theoretical minimum of 1199, but represents a nearly 50% reduction compared to the straightforward over-design using the maximum requirement for each butterfly, while maintaining reasonably simple control, memory addressing and routing.

3.4.3.4 CORDIC Rotation

A key part of the FFT calculation involves complex multiplication by the various twiddle factors along the pipeline stages. If implemented using a regular complex multiplier, the twiddle factors would need to be stored in a memory bank or lookup table in order to be used. Since we would also like to support variable size FFT calculations we would need to store the appropriate twiddle factors for all scenarios which would significantly increase

Table 3.7: FFT memory bank size and grouping

Memory index	Size (Symbols)	Used by
1	2×192	R_8^2, R_7^2, R_2^3
2	2×96	R_7^2, R_6^2, R_3^3
3	2×144	R_6^2, R_5^2, R_4^2, R_4^3
4	2×120	R_5^2, R_4^2, R_3^2, R_5^3
5	216	R_4^2, R_3^2
6	300	R_3^2, R_2^2
7	270	R_2^2
8	243	R_1^2
9	2×3	R_2^3
10	2×25	R_1^3
11	4×5	R_2^5
12	4×1	R_1^5

the memory requirements. The address control to retrieve them would add additional complexity and overhead to the design.

The main motivation behind reducing the number of multiplications is in order to reduce the pipeline depth and reduce the overall power of the circuit since a multiply is generally slower and consumes more power than an addition operation. This drives various attempts to reduce the number of multipliers in the FFT implementation, such as radix-2^2 and 2^3 butterfly usage as well as other multiplier-less implementations using various shift-add-subtract techniques [158] in order to reduce or remove completely the number of required complex multipliers.

While the twiddle factor complex value changes in a non-linear fashion throughout the FFT calculation, the angle (i.e. the complex exponent power) varies linearly with a constant step. This is observed for example by examining the twiddle factor powers which appear on the flow graph in Fig. 3-6. The use of the angle value will allow using a simple counter or adder instead of needing to store the twiddle factors in memory. However, we would still like to avoid calculating the value of the twiddle factor from the angle value which would

require a trigonometric function or an approximation for it [159] and would still require the complex multiplication. Noting that the multiplication by the twiddle factor used is in fact a rotation of the signal in the complex plain, we choose to use a COordinate Rotation Digital Computer (CORDIC) [160] block in order to rotate the data signal instead of using multipliers.

CORDIC rotation is based on the observation that a rotation by θ degrees of a complex vector is written as

$$u = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} v \tag{3.49}$$

Using the trigonometric identities

$$\sin\theta = \frac{\tan\theta}{\sqrt{1+\tan^2\theta}} \tag{3.50}$$

$$\cos\theta = \frac{1}{\sqrt{1+\tan^2\theta}} \tag{3.51}$$

(3.49) may be written as

$$u = \frac{1}{\sqrt{1+\tan^2\theta}} \begin{pmatrix} 1 & -\tan\theta \\ \tan\theta & 1 \end{pmatrix} v \tag{3.52}$$

The CORDIC algorithm uses a series of such rotations in a pipeline form where the angle of rotation θ for each stage is chosen such that

$$\tan\theta_i = 2^{-i} \tag{3.53}$$

This, in turn, will simplify (3.52) to

$$u_i = K_i \begin{pmatrix} 1 & -\sigma_i 2^{-i} \\ \sigma_i 2^{-i} & 1 \end{pmatrix} v_i \tag{3.54}$$

Where σ_i could be either +1 or -1 (for positive or negative rotation), and K_i is the stage's

scaling factor equal to

$$K_i = \frac{1}{\sqrt{1+2^{-2i}}} \tag{3.55}$$

This choice of angle θ lends itself to simple implementation in hardware, since division by powers of two is simply attained by bit shifting the signal and does not require complex calculation. Furthermore, the scaling factor is accumulated across all stages as the product of the individual scaling factors and applied only at the end to reduce the complexity. Therefore each stage consists of only bit shifts and two adders. Note that all stages rotate either clock-wise or counter-clockwise, there is no "no-rotation" state in this topology (though such implementations have been suggested [161] with additional cost to control complexity), therefore the scaling factor is constant and does not depend on the angle of rotation and is implemented using CSD multiplication of the final result.

Fig. 3-30 shows the schematic of the implemented CORDIC block in the system. It consists of 6 CORDIC rotation stages with an additional quadrant rotation block to handle the simple cases of 90° rotations. The following rotation blocks actually begin with $i = 1$, which corresponds to a rotation angle of $\theta = 26.6°$, and not from $i = 0$ ($\theta = 45°$) since the sum of all stages already covers this angle range so it is redundant. The overall scaling factor required is

$$K = \prod_{i=1}^{i=6} K_i = \prod_{i=1}^{i=6} \frac{1}{\sqrt{1+2^{-2i}}} \approx 0.86 \tag{3.56}$$

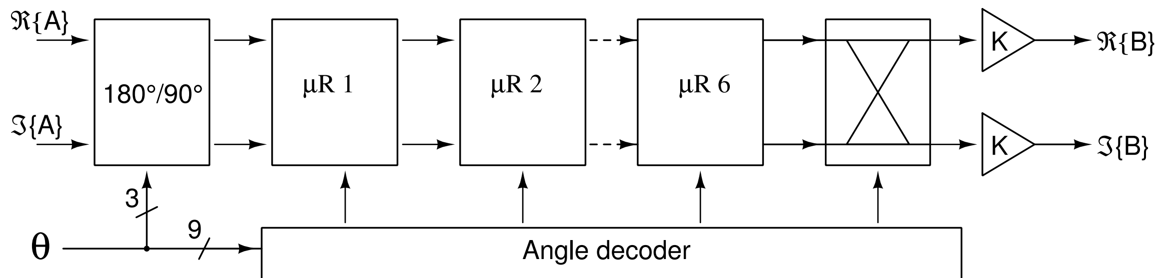


Figure 3-30: CORDIC block schematic

The control signal to decide in which direction each stage should be used for a given

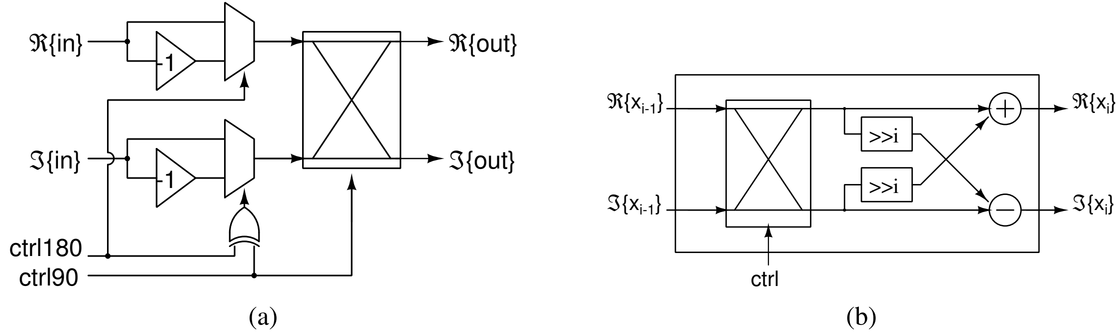


Figure 3-31: CORDIC schematic of (a) quadrant rotator and (b) micro-rotator block

desired rotation angle is calculated by mapping the possible angle range to the CORDIC rotation angle space and thus removing the need of additional memory storage for the CORDIC rotation values [162]. In our design, the angle codeword consists of 10 bits mapping one quadrant of 90 degrees. The CORDIC block consists of 6 bits (one for each stage) and the total degree coverage is

$$\sum_{i=1}^{6} \theta_i = \sum_{i=1}^{6} \arctan 2^{-i} \approx 108 \tag{3.57}$$

Each rotation block implements the reverse rotation simply by swapping the real and imaginary components at the input and output of the block. In order to save on time and hardware, only an input swapping block is used at each stage and the output swap is performed by the following stage. If two consecutive stages wish to rotate in the same direction then no swap is required in between them. Therefore the rotation vector calculated previously is XOR'd with a shifted version of itself to determine the final swap control signal for each rotator stage.

3.4.3.5 Angle Generation

The specific angle which should be used at each CORDIC rotator depends on the location of the rotator in the pipeline chain and on the current clock cycle within the DFT calculation. Fig. 3-32a illustrates the basic block diagram architecture for generating a general angle for a given rotator. The angle is incremented at each cycle by a fixed value (denoted *inc* in the

diagram). The increment value itself is increased by a fixed amount (denoted *scale*), which is determined by the location of the rotator in the pipeline, and depends on the accrued delay up to it. The increment value is reset by a counter set to count up to the relevant butterfly radix. The entire process is repeated each time a global counter reaches the required delay.

For the special case of 3 butterflies configured as a radix-2^3, this architecture is simplified as shown in Fig. 3-32b. Here, we use a 3 bit counter as the radix counter, but utilize the counter output itself. Reversing its bits, and multiplying by the required *scale* we obtain the desired increment for the angle. Since it is only a 3 bit counter we use the fixed 3 bit multiplier as described in section 3.4.3.6 and shown in Fig. 3-34b. The entire process is again reset by a global counter when the proper butterfly delay value is reached.

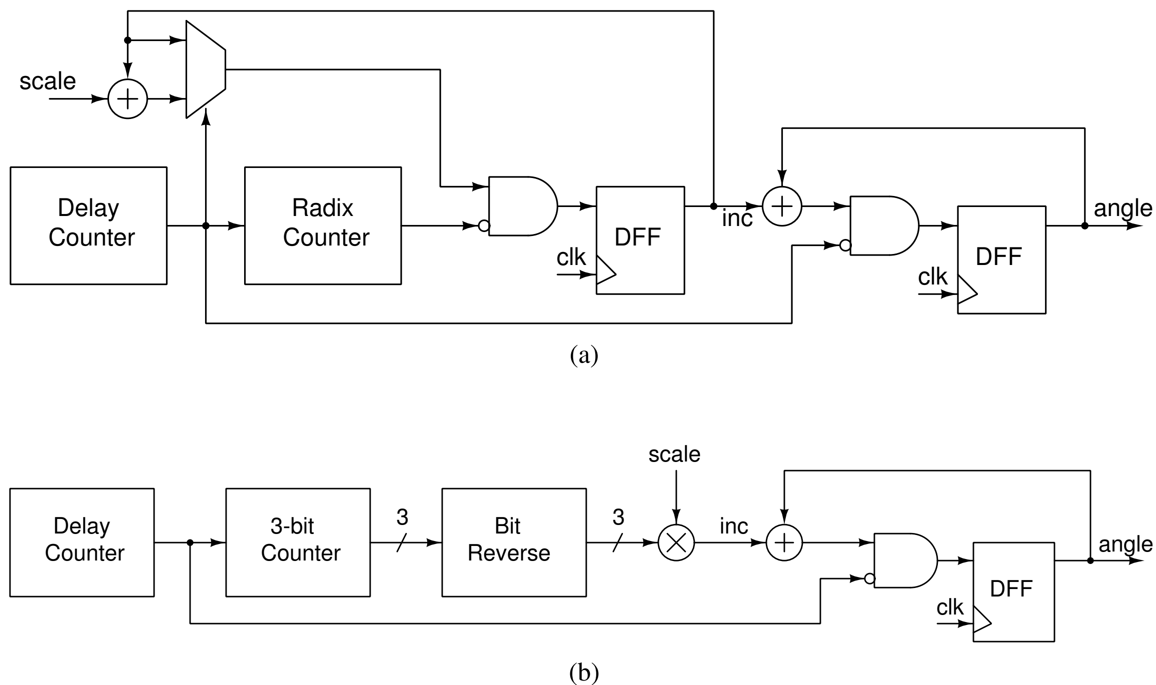


Figure 3-32: Angle generation for CORDIC following (a) general radix butterfly and (b) simplified architecture for radix 2^3 butterfly block

3.4.3.6 Index Counter

As described in section 3.3.3, the FFT pipeline output emerges out of order in a reversed digit format which depends on the actual butterflies used and their appropriate radix. In

order to sort the data before entering the following pipeline blocks an index counter must be used to correctly attribute each data point with its matching index. Since the FFT calculation is not radix 2 only, this is slightly more involved than a simple bit reversal of a counter. The index calculation is performed similarly to that described in the aforementioned section and specifically follows (3.34).

A schematic of the basic mixed-radix counter circuit used to calculate the appropriate index as the data points become available is shown in Fig. 3-33. Enable signals for each block (derived from the corresponding butterfly block enable signal) control which counters are active and if data needs to be bypassed as necessary. The bypass mechanism is not shown in the schematic to avoid clutter in the diagram. Each counter is triggered to increment one step when the previous counter has completed a count to its full range. Each counter output is scaled according to the proper factor based on the following active units and their radix and all results are eventually summed. The simple bit reversal of the radix-2 butterfly counter is adjusted to accommodate for cases where not all of the units are used. The butterflies' enable signal is negated, bit reversed and decoded from thermometer code to binary to indicate the required left shift of the counter before its own bit reversal. This allows for correct bit reversal of the full 8 bit counter, accounting for the fact that several of the top bits are sometimes unused.

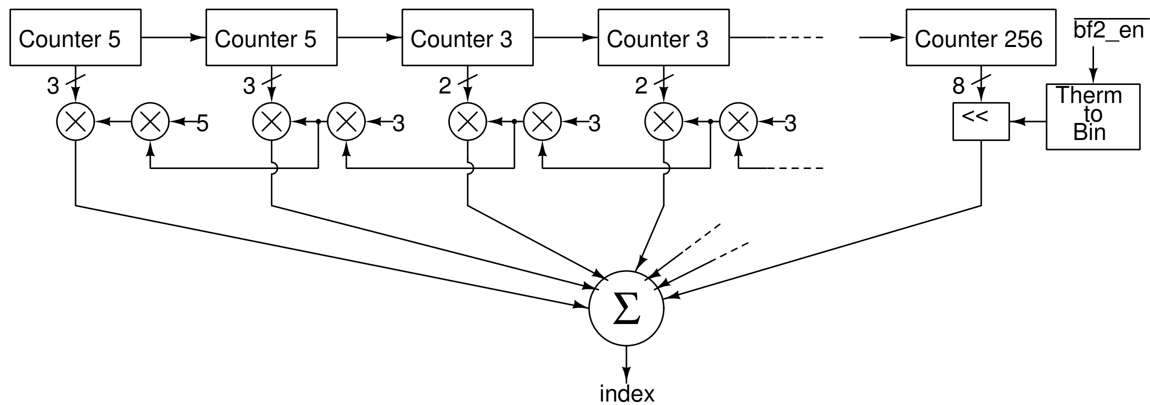


Figure 3-33: Mixed radix reverse digit counter for DFT output indexing

It would seem that this topology uses as many as 14 multipliers, however we note that 7 of these are constant multipliers by 3 or 5, which are implemented using a shift left by

one or two respectively and using only one adder to add the original value to the shifted result. We also note that all other multipliers have their input from the 3/5 counters which are only 2/3 bits wide which allows use of a much more simplified multiplier blocks using only one or two adders, respectively. The schematic for these 2 and 3 bit multipliers are shown in Fig. 3-34a and 3-34b, respectively.

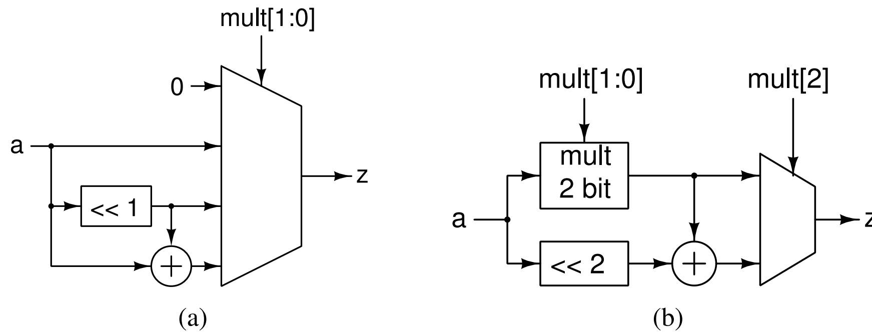


Figure 3-34: Simplified multiplier blocks for (a) 2 bit and (b) 3 bit multipliers

3.4.4 Resource Mapping

After the data is spread across the bandwidth via the use of the DFT block, we wish to map the different subcarriers to the available spectrum and assign the orthogonal frequency duplexing which will be preformed in the following IDFT block. In the uplink channel in LTE communication, each UE is allocated a continuous set of RBs in the frequency domain. In order to minimize interference to other transmitters in the same cell, all other subcarrier values are set to zero before the IDFT calculation is performed. Since the IDFT size is inherently different than the DFT (due to the added zero padding) we will also need to accommodate for the difference in data rates in the pipeline architecture.

Following the DFT block, a double buffer is used to synchronize the data flow. Two SRAM banks, each with a size of 1200 symbols are used to buffer and re-time the data from the DFT output to the IDFT input. The buffer control would allow either passing the data stored in one bank, while continuing to save the DFT output data to the second bank, or passing zero values as required for the null subcarriers. Working with the same clock implies that once the symbol data output from the DFT is complete, that block is no longer

enabled and holds until all data and null values are read and passed on to the IDFT stage. This approach allows using one global clock for the entire chain and avoids exact clock rate matching between the various pipeline stages which could cause data integrity issues when crossing clock domain boundaries.

One may question whether it is possible to exploit in some manner the fact that many of the data inputs to the IDFT block are constant zero values. Next we will explore such opportunities which arise from the unique characteristics of the typical LTE signal statistics. As discussed earlier, the DFT is performed on a signal with length equal to the amount of SCs allocated to the specific user which could be any one of 34 values ranging from 12 (for 1 RB) to 1200 (for 100 RBs) as detailed in table 3.5. The IDFT on the other hand is performed at only a few lengths which correspond to the LTE bandwidth allocations between 1.4 MHz and 20 MHz. Therefore the amount of non-zero data values in the IDFT calculation may be as low as 0.6% of the total data input in the extreme case of allocating 1 RB to a user in a 20 MHz channel.

3.4.4.1 Interpolation

First we derive the time domain interpretation for the mathematical process of applying an IDFT operation on a zero padded signal which underwent the reverse DFT operation. If there was no zero padding between the two stages we would have obtained an exact reconstruction of the original time domain data, i.e.

$$\mathscr{F}_M^{-1}\left\{\mathscr{F}_M\left\{x[n]\right\}\right\} = x[n] \tag{3.58}$$

In general, taking the Fourier transform of a zero padded signal is equivalent to band-limited interpolation of the frequency domain content [163]. Therefore, in this case, we would expect the output result to be some form of interpolation of the original time domain data. Verifying this, we denote the zero padded sequence of the original data $x[n]$ after the DFT block as

$$X[k] = \mathscr{F}_M\left\{x[n]\right\} \tag{3.59}$$

$$Y[k] = \begin{cases} X[k] & 0 \le k < M \\ 0 & M \le k < N \end{cases} \tag{3.60}$$

Note that we have set all the zero values to occur at the end of the original data sequence. This is not necessarily the case (and will actually never happen due to the guard bands used at the edges of the channel bandwidth), but translating the final result to any circularly shifted version of the signal is a trivial rotation of the result since

$$\mathscr{F}_N\{x[n+L]\} = \mathscr{F}_N\{x[n]\} W_N^{-kL} \tag{3.61}$$

Using the zero padded series defined in (3.60) as the N-length input to the IDFT block we write the output as

$$\begin{aligned} y[n] &= \frac{1}{N}\sum_{k=0}^{N-1} Y[k] W_N^{-nk} = \frac{1}{N}\sum_{k=0}^{M-1} X[k] W_N^{-nk} \\ &= \frac{1}{N}\sum_{k=0}^{M-1}\sum_{m=0}^{M-1} x[m] W_M^{mk} W_N^{-nk} \\ &= \frac{1}{N}\sum_{m=0}^{M-1} x[m] \sum_{k=0}^{M-1} W_M^{-k\left(\frac{M}{N}n-m\right)} \\ &= \frac{1}{N}\sum_{m=0}^{M-1} x[m] W_M^{\frac{M-1}{2}\left(m-\frac{M}{N}n\right)} \frac{\sin\left[\pi\left(m-\frac{M}{N}n\right)\right]}{\sin\left[\frac{\pi}{M}\left(m-\frac{M}{N}n\right)\right]} \end{aligned} \tag{3.62}$$

Using this direct calculation, we wish to compute the total output using the result found in (3.62). The computation complexity is on the order of $O(MN)$ but also requires complex exponentiation and trigonometric function calculation for each point which might result in fairly complex and power-expensive hardware implementation. Alternatively, we view (3.62) as a sort of band-limited interpolation filter, where we compute the result of a convolution of the original signal with a filter and then resample it at non-integer sample intervals. That is

$$y[n] \approx \frac{M}{N}(x*h)\left[\frac{M}{N}n\right] \tag{3.63}$$

where

$$h[n] = \frac{\sin(\pi n)}{M \sin\left(\frac{\pi}{M} n\right)} W_M^{\frac{M-1}{2} n} \tag{3.64}$$

Furthermore, this can be viewed as a standard multi-rate resampling of a signal, such as depicted in Fig. 3-35. The filter is an ideal reconstruction filter in the ideal case (i.e. a brick-wall frequency response). Therefore, we see that the final signal output, after performing an M-size Fourier transform, zero padding, shifting by L points and taking the inverse transform of size N results in a scaled, rotated band-limited interpolation of the rotated original data series, concisely expressed as

$$y[n] \approx \frac{M}{N} W_N^{-nL} x\left[\frac{M}{N} n\right] \tag{3.65}$$

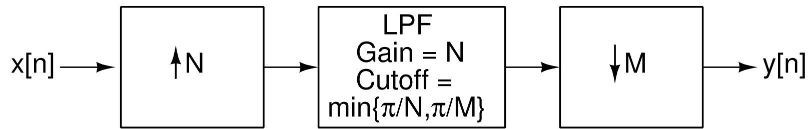
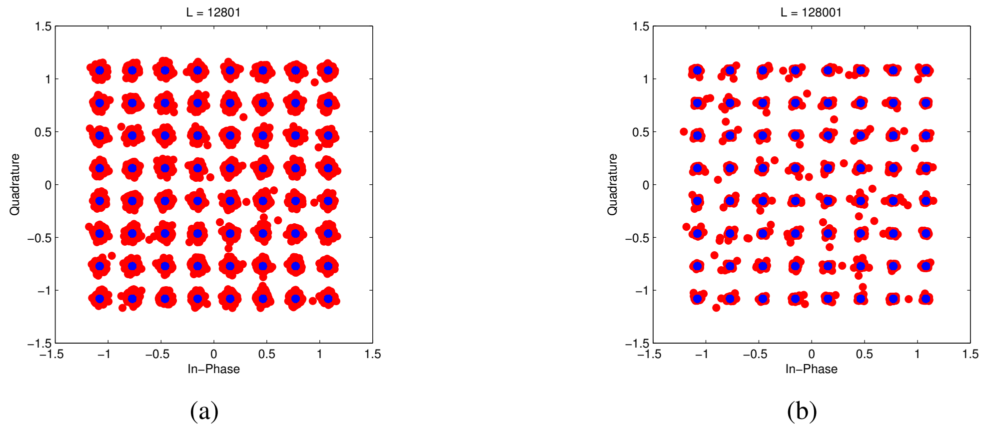


Figure 3-35: Block diagram of filter re-sampling by a factor of N/M

This observation implies that one can obtain the desired output data values by interpolating the input data. A typical approach to obtain a non-integer re-sampling of a signal would involve upconverting the signal, low-pass filtering the result in order to interpolate between the samples, and then decimating to obtain the desired new sampling rate. The advantage of using this scheme compared to the original calculation is the potential of using a fixed low pass interpolation filter, which in turn can be realized by having constant real multiplications of the signal instead of the more complex direct calculation.

The reconstructed constellation points are shown in Fig. 3-36 after calculating the output vector using the proposed multi-rate resampling approach with various fixed Low Pass Filter (LPF) designs. It can be seen that we indeed approximate the ideal signal though with some degree of degradation due to the approximations used and the fact that we wish to use a finite filter length which is not equivalent to the ideal, infinite LPF. Fig. 3-37 plots the reconstructed symbol Error Vector Magnitude (EVM) achieved for various filter lengths.



(a)

(b)

Figure 3-36: Reconstructed samples from resampling filter with length (a) L = 12801 and (b) L = 128001

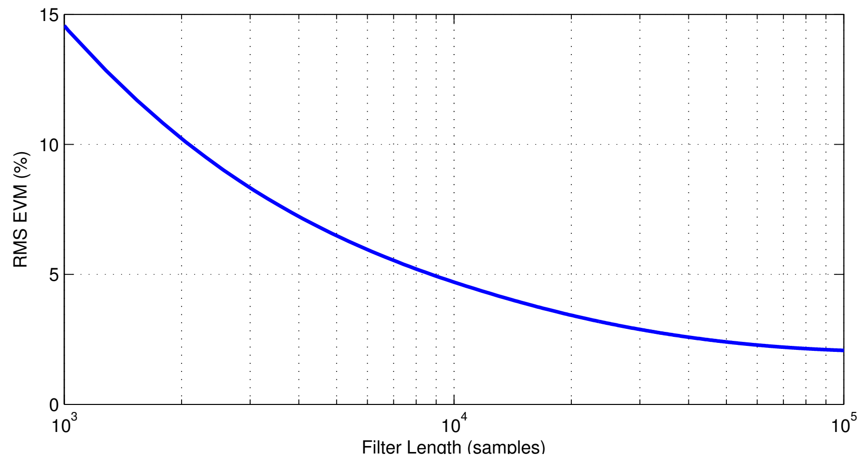


Figure 3-37: EVM of reconstructed samples as a function of resampling filter length. 10 sets of 1200 random 64-QAM points using a 2048 point IDFT

This approach, although feasible as seen, does not seem to ultimately have an advantage over the original approach of calculating the forward and inverse Fourier transforms. This is due to the fact that in order to obtain accurate interpolation values, we will require a filter with as many taps as on the order of the signal length. A typical filter calculation requires $O(N^2)$ operations, which would be much higher than that required when using the FFT algorithm for the two blocks, which would only require $O(N \log N)$ operations.

3.4.4.2 Transform Decomposition

Another approach we might consider in our case is the use of Transform Decomposition (TD) [164] for the calculation of the inverse Fourier transform with only a subset of the input points. We consider the case where only L inputs are non-zero and assume there exists a value $P \geq L$ which divides the DFT size N and define $R = N/P$. Recalling our result of (3.11), we observe that in this case $x_i \neq 0$ only for the case of $i = 0$ (where x_i is defined in (3.10)), which would reduce the result to

$$X[Rk+r] = \mathscr{F}_P\{x[n]W_N^{nr}\} \qquad \begin{array}{l} n = 0,1,\ldots,P-1 \\ k = 0,1,\ldots,P-1 \\ r = 0,1,\ldots,R-1 \end{array} \tag{3.66}$$

Therefore, in this case we only need to compute the P-length DFT R times, instead of the N-length DFT. This in turn results in a reduced complexity of $O(R \times P \log P) = O(N \log P)$ instead of $O(N \log N)$. Fig. 3-38 illustrates the TD calculation. In the case of LTE SC-FDMA signal processing, we would expect to observe a benefit from using this technique when the amount of non-zero elements in the sequence is below 50% which would allow using an IDFT block size which is lower than the originally required length.

The main benefit of such a manipulation is the reduction in the amount of required calculations. It can be shown [164] that whereas for a standard complex N point FFT the total number of required operations is on the order of

$$\mathrm{OP}_{\mathrm{FFT}} = 6N \log N - 6N \tag{3.67}$$

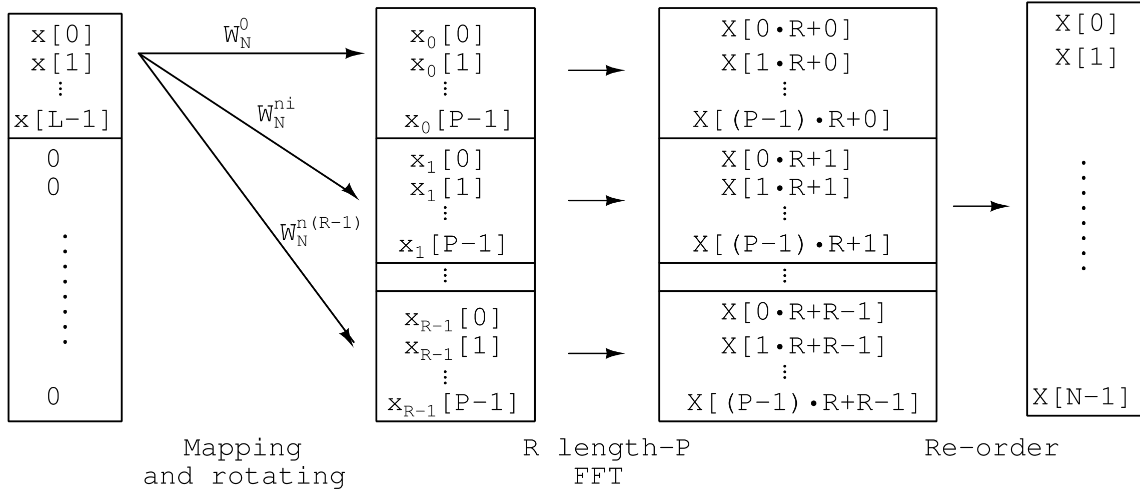


Figure 3-38: Transform decomposition procedure for an FFT calculation with a subset of the inputs

With the use of TD, choosing P to be the nearest power of 2 larger than L (the number of non-zero samples in the data), we obtain a total operation count on the order of

$$\mathrm{OP}_{\mathrm{TD}} = 4N\log P + 6\frac{LN}{P} + 8\frac{N}{P} - 6(N+L) \tag{3.68}$$

A plot of the ratio between the two expressions in (3.68) and (3.67) as a function of the fraction of non-zero samples is shown in Fig. 3-39. We note that a reduction in the number of operations is observed when the number of samples is less than half the total FFT size, since naturally this is the first point where we can actually use a value of P which is a power of 2 smaller than the original value of N and divide our data set into two groups. Further reduction in the number of operations is observed as the number of non-zero samples decreases, achieving as much as a 60% reduction in the number of operations for the extreme cases where for example only 1 RB is used in the LTE frame of a 20 MHz channel.

In order to prepare the data for calculation, an additional CORDIC rotator is placed before the IDFT block which rotates the incoming data according to (3.66). The output of the IDFT block is rotated once more in order to account for the offset of the data within the channel. If the use of TD is not desired, the first rotator is simply bypassed, passing along

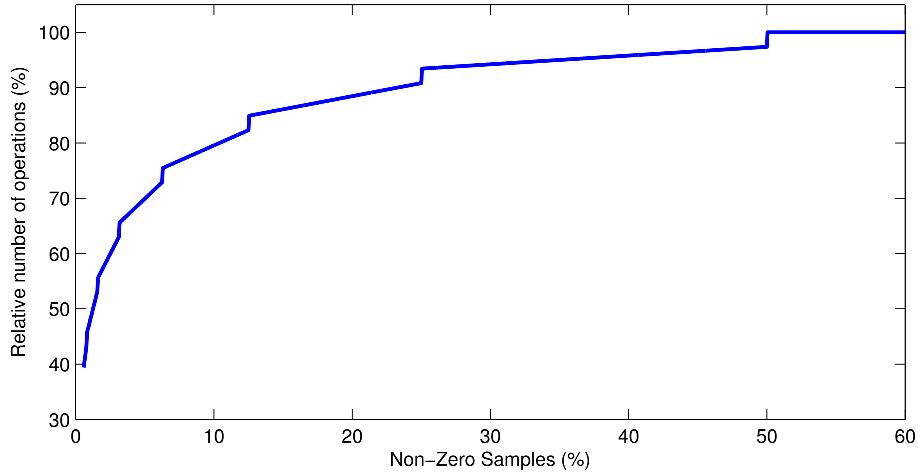


Figure 3-39: Fraction of operations required in TD-FFT as a function of non-zero sample ratio

the unmodified original data to the IDFT processor. As for the twiddle factors, the angle increase of the multiplication is linear and therefore is implemented with a simple counter with no need for additional memory storage.

3.4.5 IDFT

The IDFT block required as part of the LTE symbol generation scheme requires a variable size inverse Fourier transform of either 128, 256, 512, 1024, 1536 or 2048 points. These are all powers of two apart from the 1536 point case. Although, strictly speaking, it is possible to use the 2048 point transform for all cases in the LTE transmission by simply zero padding all other inputs, this will not allow benefiting from reduced hardware usage and reduced clock rate when lower bandwidth is needed for transmission. Therefore we would like to be able to change the IDFT size as needed.

The IDFT block was implemented via a $R2^3SDF$ pipeline. An additional radix-3 block was added in order to comply with the 1536 point case. The overall block architecture of the IDFT chain is shown in Fig. 3-40. The radix-3 and radix-2 butterflies of all types are identical to the ones used in the FFT chain described in section 3.4.3 and shown in Fig. 3-15, 3-19 and 3-21. Due to the use of high order butterflies, we are able to use only three twiddle factor multipliers along the chain instead of 10 which would be needed in

the conventional R2SDF. The twiddle multipliers were again implemented as CORDIC rotators as in the DFT implementation. Since we are mainly using radix-2 elements, the memory allocation, control scheme and CORDIC angle generation are much simpler in this block compared to the DFT block.

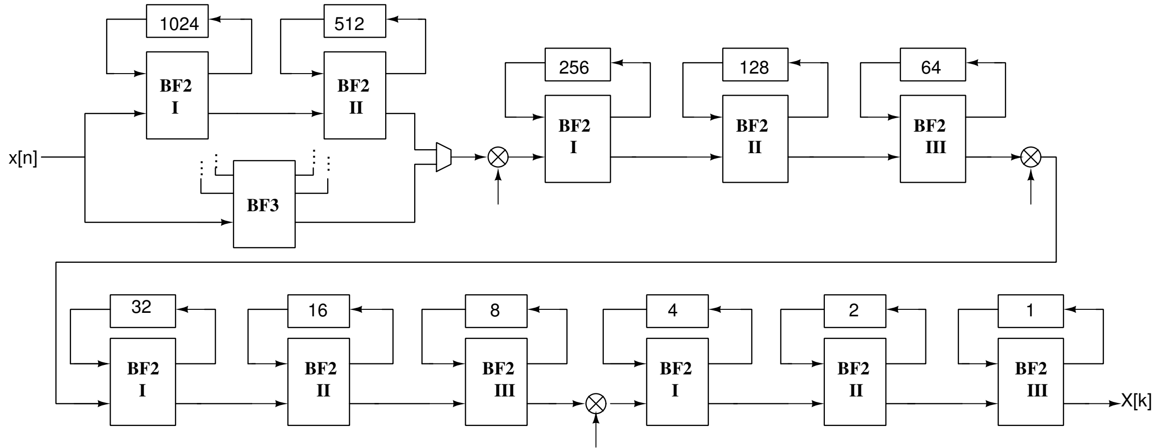


Figure 3-40: IDFT block architecture

The delay lines used for the shorter delays of 256 and below were implemented as a latch based delay line. The two largest delay blocks of 512 and 1024 samples were implemented as SRAM based delay lines as described in section 3.4.3.2. The top two SRAM delay lines were also used as the double, 512 sample, delay line required for the radix-3 butterfly when performing a 1536 point IDFT calculation. In that case, the top two radix-2 butterflies were powered down, and their memory delay line blocks were used by the radix-3 butterfly in parallel for the required dual delay path.

3.4.6 Cyclic Prefix Addition

The addition of the cyclic prefix to the signal after frequency duplexing is achieved by use of an additional double buffer consisting of two SRAM memory banks, each of size 2048 to accommodate for the largest output bandwidth of 20 MHz. A simple cyclic counter both reads the output data in-order after accounting for the IDFT bit reversal and also transmits a duplicate of the symbol end according to the LTE protocol requirements. When using an

extended cyclic prefix, the last $N_{\text{CP,ext}}$ symbols are repeated, where

$$N_{\text{CP,ext}} = 512 \times \frac{N_{\text{IDFT}}}{2048} \tag{3.69}$$

Which account for 25% of the symbol length. For regular cyclic prefix addition, we differentiate between the first OFDM symbol in a slot and the remaining 6 symbols, such that the amount of values repeated is

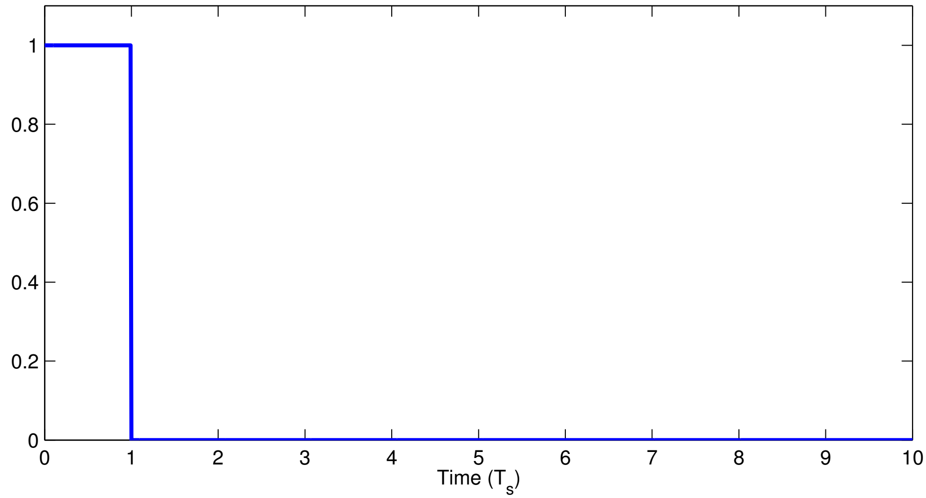
$$\begin{aligned} N_{\text{CP,0}} &= 160 \times \frac{N_{\text{IDFT}}}{2048} \\ N_{\text{CP,i}} &= 144 \times \frac{N_{\text{IDFT}}}{2048} \qquad i = 1, 2, \ldots, 6 \end{aligned} \tag{3.70}$$

3.4.7 Upsampling

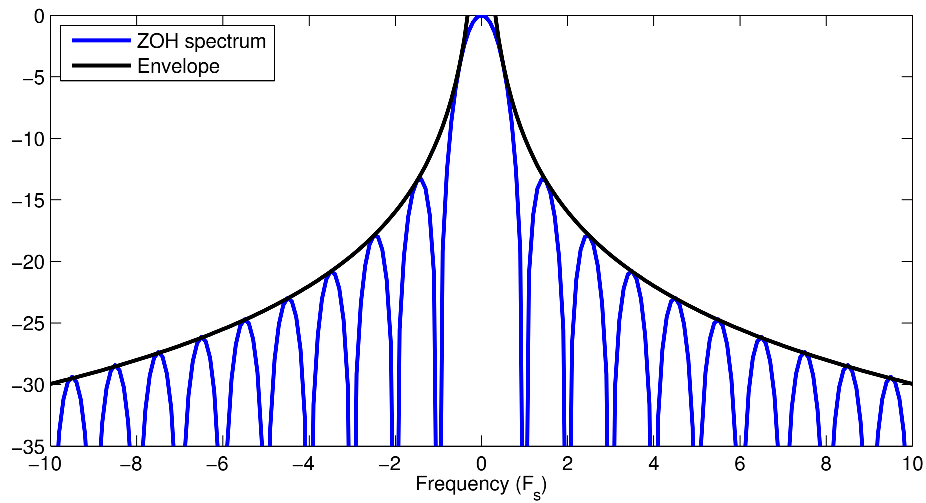
Following the addition of the cyclic prefix to the OFDM symbol the data is ready to be converted from digital samples to analog and converted from the baseband to the RF frequency (f_c) for transmission over the channel. The process of conversion from digital to analog is equivalent to replacing each symbol with an appropriate continuous time impulse. In the simple case of a Zero Order Hold (ZOH) DAC, this impulse response is simply a rectangular pulse in the time domain, which has a *sinc* shape in the frequency domain as shown in Fig. 3-41. The signal's spectrum in the continuous time domain consists of replication of the discrete spectrum at intervals equal to the sampling frequency (F_s). The convolution with the D-to-A impulse response is equivalent to multiplication in the frequency domain. Therefore, we would expect to see after upconversion to RF, the output spectrum of the baseband centered around the carrier, replicated at sampling intervals with a decay following a *sinc* envelope, which falls off as

$$\text{ZOH}_{\text{envelope}} = \frac{F_s}{|f - f_c|\pi} \tag{3.71}$$

To illustrate this effect, a 600 SC LTE signal occupying a 10 MHz channel around a 2 GHz center frequency is simulated in Matlab. Fig. 3-42a depicts the signal spectrum at baseband occupying 10 MHz out of the entire 15.36 MHz spectrum covered by the OFDM



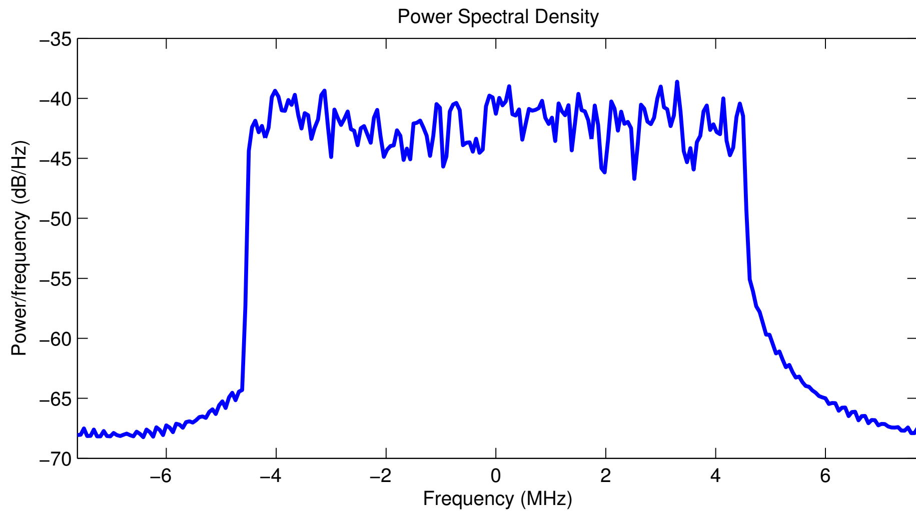
(a)



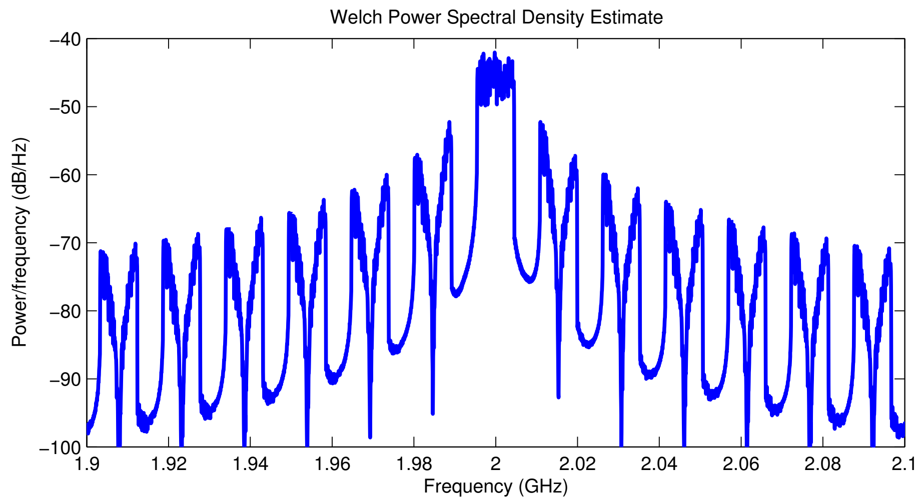
(b)

Figure 3-41: Zero order hold impulse response in the (a) time and (b) frequency domain

symbol. Fig. 3-42b shows the signal after ZOH conversion to analog and upmixing to an RF frequency of 2 GHz. The replication of the spectrum is visible with intervals of 15.36 MHz, which are equal to the sampling frequency (1024 SCs of 15 kHz each), along with the envelope decay associated with the ZOH. This implies that additional aggressive channel filtering is required before the transmitter output in order to maintain the required Adjacent Channel Leakage Ratio (ACLR) mask and avoid interference to other channels.



(a)



(b)

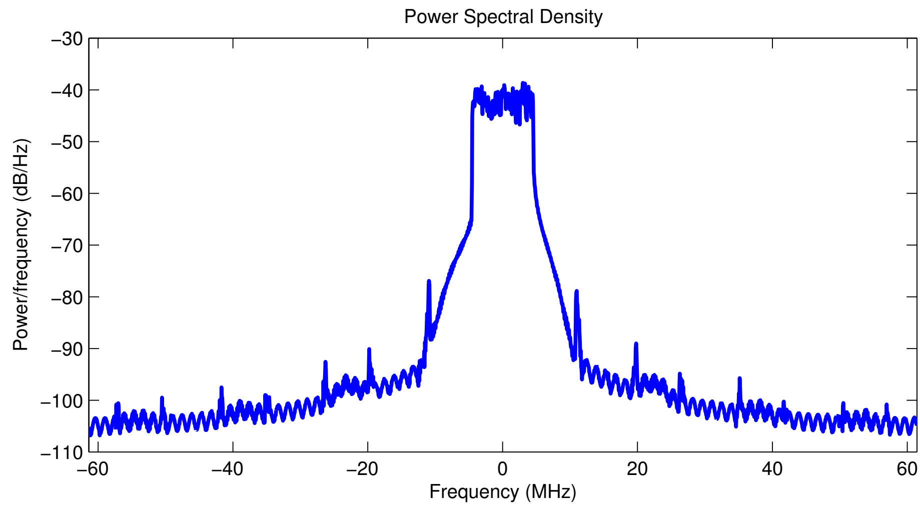
Figure 3-42: Power spectrum of 10 MHz channel LTE signal in (a) baseband) and (b) at RF around a 2 GHz carrier, with ZOH D/A with no oversampling

To avoid the need for high-Q filters we use upsampling allowing us to shape the signal's

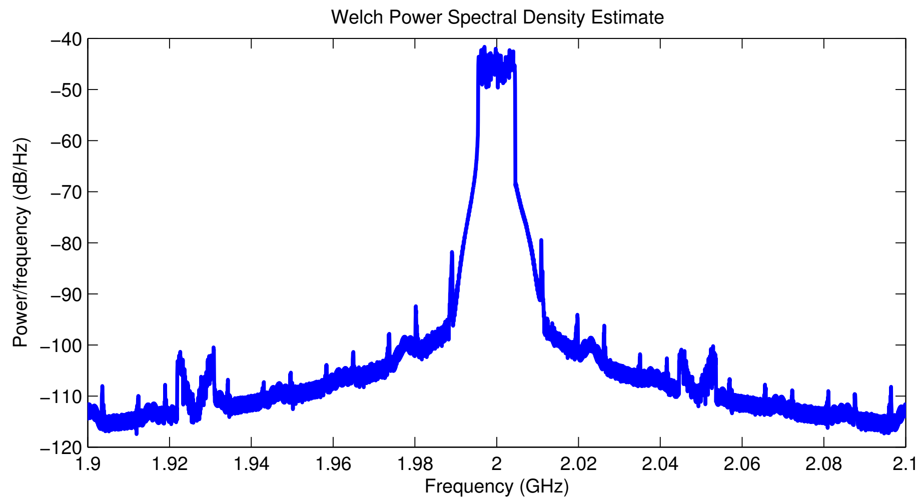
output spectrum over a wider frequency range. By upconverting the signal's sampling rate while keeping the signal spectrum the same, we shape the spectrum and thus reduce the requirements of the output filter. For example, upsampling the previous example signal by a factor of 8 via interpolation, we create a new signal with 8 times more data points which we then again convert to analog and upconvert to RF, but at a sampling frequency which is 8 times higher than before, i.e. 122.88 MHz. The output spectrum of the upsampled signal in baseband and at RF are shown in Fig. 3-43. There will still be replication of the baseband signal at the output spectrum, but these will now occur at 8 times the distance from the carrier than previously and will also be attenuated by the same ZOH *sinc* roll-off. This will make the output filter much simpler since it will not need to be as narrow band and will require a less steep roll-off in order to meet the required output spectral mask.

In order to achieve the desired upsampling, the basic topology shown in Fig. 3-44 was used. This topology consists of an expander block, inserting L-1 zero values (in our case L=8) between every two samples of the original signal and then passing the newly expanded signal through a LPF in order to interpolate between the samples.

An ideal "brick-wall" LPF cannot be realized in a physical causal system since its limited bandwidth implies it is infinite in the time domain. However other causal filters can be designed which are close in their response to the ideal filter. In implementing the filter in our system we chose to implement a Finite Impulse Response (FIR) filter, This is mostly due to the fact FIR filters, compared to Infinite Impulse Response (IIR) filters are inherently stable, and more importantly are easily designed to have a linear phase and therefore will not cause distortion of the signal due to frequency dependent group delay [163]. Matlab was used to obtain the desired interpolation filter coefficients for the upsamling filter. A compromise between filter length and the sharpness of the frequency response resulted in the choice of a 49 tap filter, which corresponds to a 6 symbol wide x8 oversampling filter. Fig. 3-45a plots the filter impulse response in the time domain, note that the impulse response is zero at the integer symbol location of $8n$ so the filter will not cause Inter-Symbol Interference (ISI). The filter's frequency response is depicted in Fig. 3-45b, it can be observed that we are indeed mimicking the ideal LPF interpolation filter, having a gain of 8 equal to the upsample rate and a 3 dB cutoff at $\pm\pi/8$.



(a)



(b)

Figure 3-43: Power spectrum of 10 MHz channel LTE signal in (a) baseband) and (b) at RF around a 2 GHz carrier, with ZOH D/A with x8 oversampling and interpolation

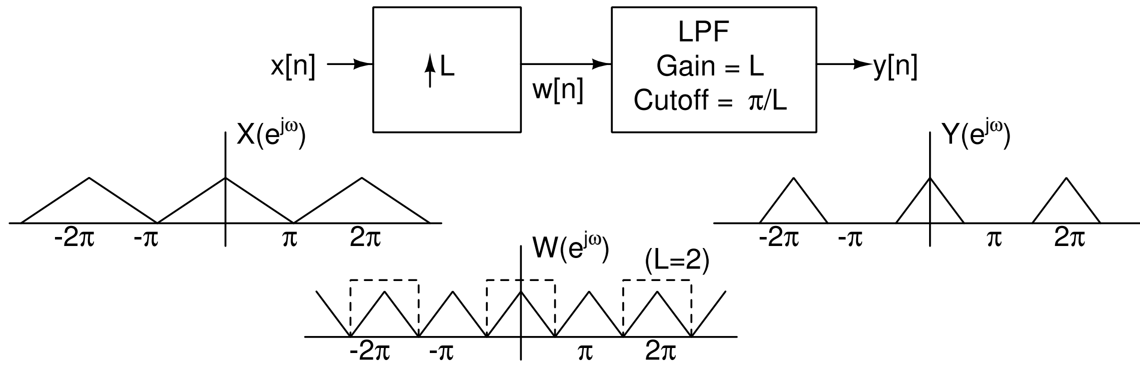
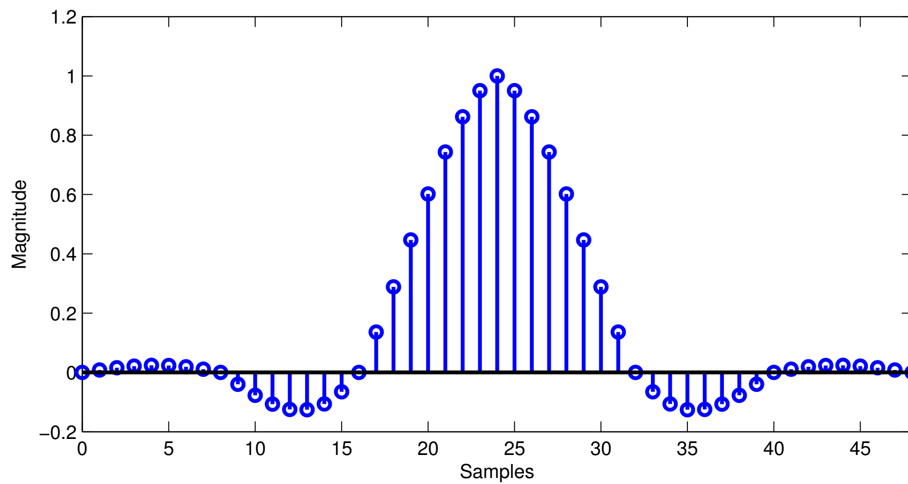
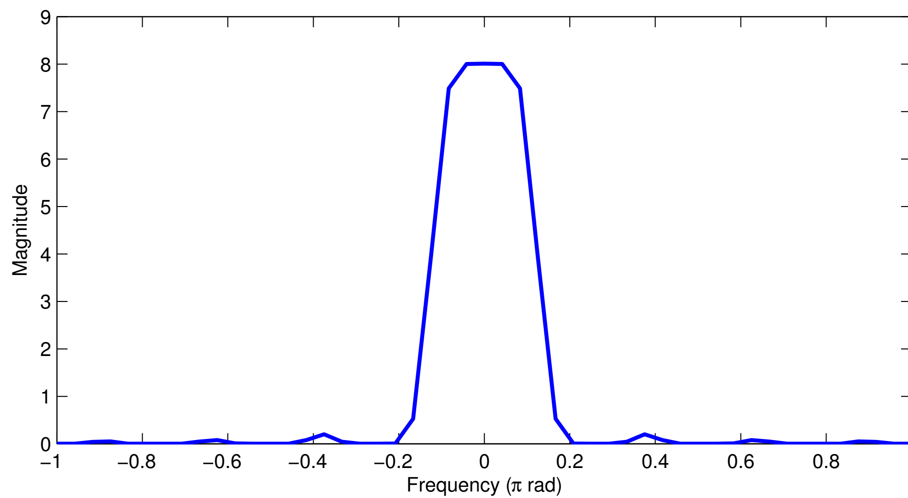


Figure 3-44: Integer rate conversion upsampling system



(a)



(b)

Figure 3-45: Low pass interpolation FIR filter (a) time and (b) frequency response

A typical hardware implementation of an FIR filter is shown in Fig. 3-46. The filter coefficients $h[n]$ are realized by constant multipliers using CSD multiplication. The delay units z^{-1} are easily realized by sequential elements such as flip flops. However, recalling the configuration of the upsampling system as shown in Fig. 3-44 we see that the filter will need to operate at the rate of the output, performing N multiplications and additions every cycle at the faster x8 clock rate of the oversampled output. Furthermore, the filter will inherently process many samples which are known to be zero which is redundant. In order to avoid these issues, we will utilize two techniques to improve the processing efficiency - The Noble identity and signal polyphase decomposition.

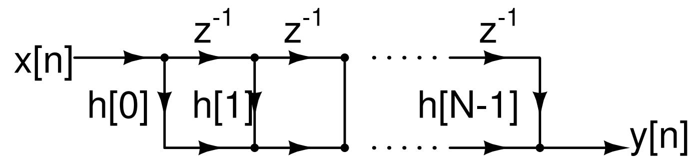


Figure 3-46: N tap FIR segment

The Noble identity states that the two systems depicted in Fig. 3-47a are identical mathematically. This can be easily proved, since in general, an expansion of a signal by a factor of L will cause its frequency spectrum to contract by the same factor, such that

$$Y_L\left(e^{j\omega}\right) = Y\left(e^{j\omega L}\right) \tag{3.72}$$

Therefore, the system output of Fig. 3-47 may be written as

$$Y\left(e^{j\omega}\right) = X\left(e^{j\omega L}\right) H\left(e^{j\omega L}\right) \tag{3.73}$$

Which is equivalent to the output of the system depicted in Fig. 3-47b. Note that the identity does not claim that it is at all possible to transition from the system depicted in Fig. 3-47b to that of 3-47a in all cases. It might not necessarily be possible to express the required filter as an L root of the original filter. The identity simply states that if the filter can be written in this form, then the two systems are equivalent.

To take advantage of the previous derivation, we turn to another aspect of the signal

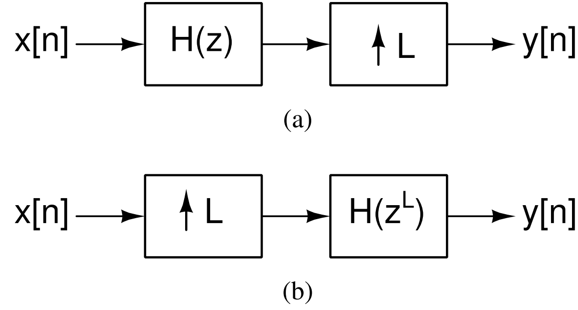


Figure 3-47: Expander filter Noble identity

processing which is polyphase decomposition. This simply implies that we may rewrite a signal, or filter, as the sum of lower order functions such that

$$H(z) = \sum_{k} E_k(z^L) z^{-k} \tag{3.74}$$

Given a causal, length-N FIR filter of the form

$$H(z) = \sum_{n=0}^{N-1} h[n] z^{-n} \tag{3.75}$$

We choose a value L which divides N (if necessary, the original filter is zero padded so that the new length is divisible by L) such that

$$n = mL + k \qquad \begin{array}{l} k = 0, 1, \ldots, L-1 \\ m = 0, 1, \ldots, \frac{N}{L} - 1 \end{array} \tag{3.76}$$

Substituting (3.76) into (3.75) we obtain

$$H(z) = \sum_{k=0}^{L-1} \sum_{m=0}^{N/L-1} h[mL+k] z^{-(mL+k)} \tag{3.77}$$

Which is exactly in the form of (3.74) if we denote

$$E_k(z^L) = \sum_{m=0}^{N/L-1} h[mL+k] z^{-mL} \tag{3.78}$$

Using the above decomposition for the interpolation filter used in Fig. 3-44 will result in the system depicted in Fig. 3-48a. We now use the Noble identity derived earlier to transform the system to that shown in Fig. 3-48b. In this case, we are guaranteed that this transition is possible due to the way the sub filters were constructed and we recognize the new sub-filters as

$$E_k(z) = \sum_{m=0}^{N/L-1} h[mL+k]z^{-m} \tag{3.79}$$

This is simply a set of L filters, each comprising of samples of the original filter taken at an interval of L samples. For example, for an upsample by 8 as in our case, the first sub filter will consist of the original filter samples $h[0]$, $h[8]$, $h[16]$, $\ldots$, the second filter will consist of the points $h[1]$, $h[9]$, $h[17]$, $\ldots$ and so on until the last sub-filter which comprises of the points $h[7]$, $h[15]$, $h[23]$, $\ldots$.

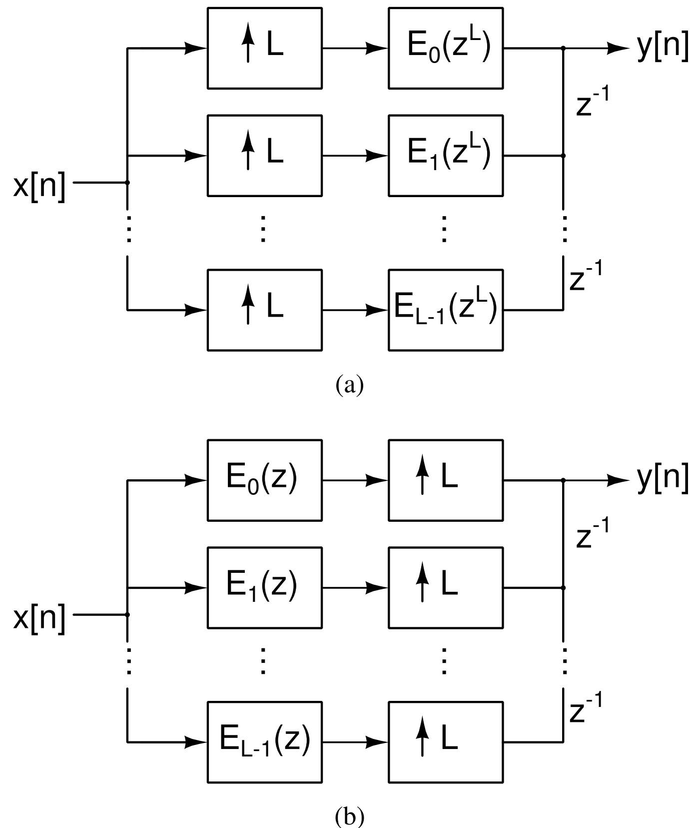


Figure 3-48: Polyphase decomposition for an interpolating filter using the Noble identity

Each of the aforementioned sub-filters is realized as explained earlier and depicted in Fig. 3-46. However now we are processing the data before the expander, therefore we are not conducting any redundant calculations (there are no inherent zero values during the filtering process). In addition, the computation is now performed at the clock rate of the original data before the increase in the sample rate, which means we carry out N multiplications and additions at the slower clock rate frequency, reducing the area and power requirements of the hardware implementation of the filter.

The technique described above was used to implement the 49 tap interpolation filter used to gain the x8 upsampling of the output signal from the baseband block before entering the DAC. The function following the sub-filters, i.e. the expansion, delay and addition was simply implemented as a MUX which selected the following sample output from one of the sub-filter bank outputs according to a rotating counter operating at eight times the baseband clock frequency as illustrated in Fig. 3-49. Thus, the only component required to operate at the higher data rate is a simple 8-to-1 multiplexer.

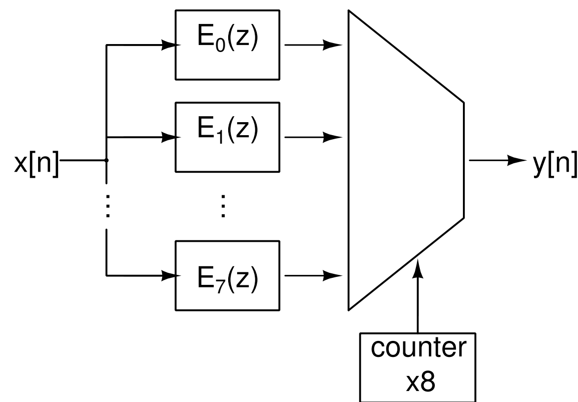


Figure 3-49: Polyphase implementation of the x8 interpolation filter

3.4.8 Snapshot Memory

To gain visibility into the inner workings of the various sub-blocks of the digital signal processing chain, an inner "snapshot" memory was added. The memory is an SRAM unit of 2048 words, which are each 32 bits long. This allows capturing of the longest sequence of

one OFDM symbol with a complex representation of 16 bit words. The snapshot memory is configured to capture either the DFT or IDFT block's input or output for a specific desired symbol. The data is replicated and stored in the snapshot memory during the regular digital block signal processing flow as the data arrives and is stored there for later extraction. This allows reading the data using an external, low speed clock through slower I/O pads instead of the full rate digital block clock speed.

3.5 Measurements

The digital LTE baseband design was implemented in a 28 nm LP CMOS process. The total die area is $2.8 \times 2.9\ \text{mm}^2$, with a core active area of $1.17 \times 0.68\ \text{mm}^2$. The FFT and IFFT blocks occupy within the core an area of $0.31\ \text{mm}^2$. An annotated photo of the die is shown in Fig. 3-50. The die was packaged in a QFN 88 lead package. The packaged die was mounted onto a test PCB which enabled communication with an FPGA board and a PC which were used to load data and configuration options to the chip as well as read the snapshot memory and high speed outputs. The key features of the digital block are summarized in table 3.8.

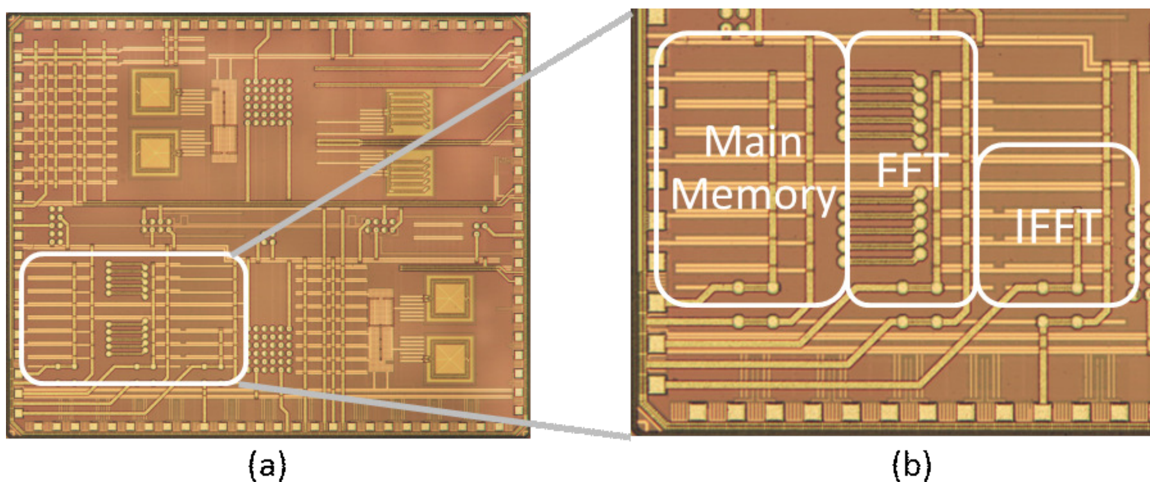


Figure 3-50: Micrograph of (a) chip and (b) zoom-in of implemented LTE digital baseband chip

Table 3.8: Design Specification Summary

	FFT	**IFFT**	**Interpolator**
Technology	28 nm LP CMOS		
Package	QFN 88		
Modulation	BPSK, QPSK, 16-QAM, 64-QAM		
Word width	2×16 bits	2×16 bits	2×11 bits
Memory size	2213 words	2047 words	48 words
Clock frequency (MHz)	$1.92 \sim 30.72$	$1.92 \sim 30.72$	$15.36 \sim 245.76$
FFT size	$12 \sim 1200$ Radix-2/3/5	$128 \sim 2048$, 1536 Radix-2/3	-
Gate count	511K	170K	22K

3.5.1 Data Path

To illustrate the operation of the baseband processor, we will closely examine the data output at various points along the data path as read from the test chip via the on-board snapshot memory and high-speed I/O pads. We will consider the example of transmitting 40 RBs (corresponding to an occupied bandwidth of 7.2 MHz) of 64-QAM modulated random data in a 10 MHz bandwidth. The input data symbols, as captured from the SRAM memory output after QAM mapping, are shown in Fig. 3-51. The data is then processed via a 480 point FFT which correspond to 40 RBs, each containing 12 SCs of 15 kHz bandwidth.

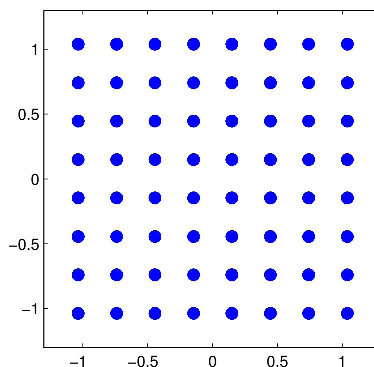


Figure 3-51: Measured 480 64-QAM modulated input symbols

The symbols come out of the FFT out of order and need to be sorted as detailed in section 3.3.3. The 480 point FFT comprises of 5 radix-2 butterflies, 2 radix-3 butterflies and one radix-5 butterfly in the pipeline. Therefore the index of each symbol will be a digit reversed pattern from 0 to 479 in that basis in reverse. The output of the FFT index counter is shown in fig. 3-52, as can be seen we observe the periodicity in base 5, within base 3 periods which is in itself in a base 2 repetition pattern representing the index counter output.

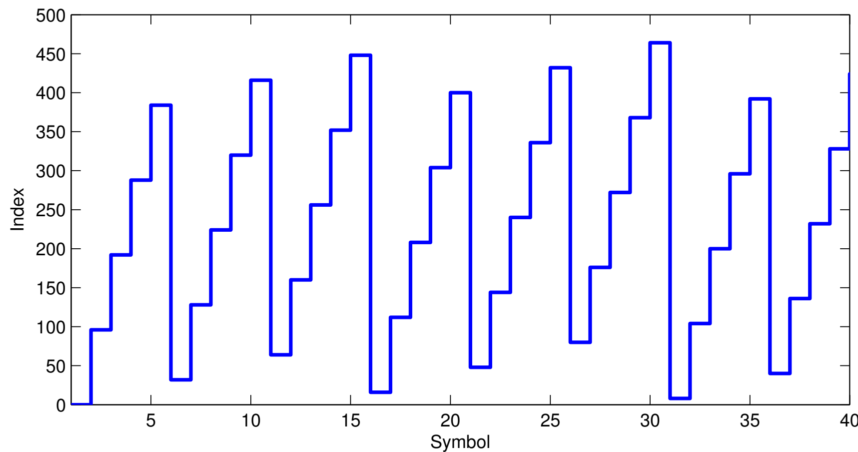
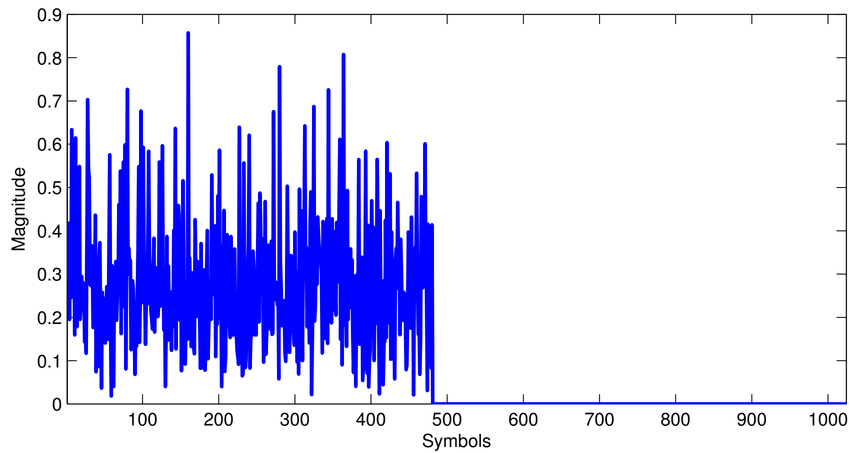


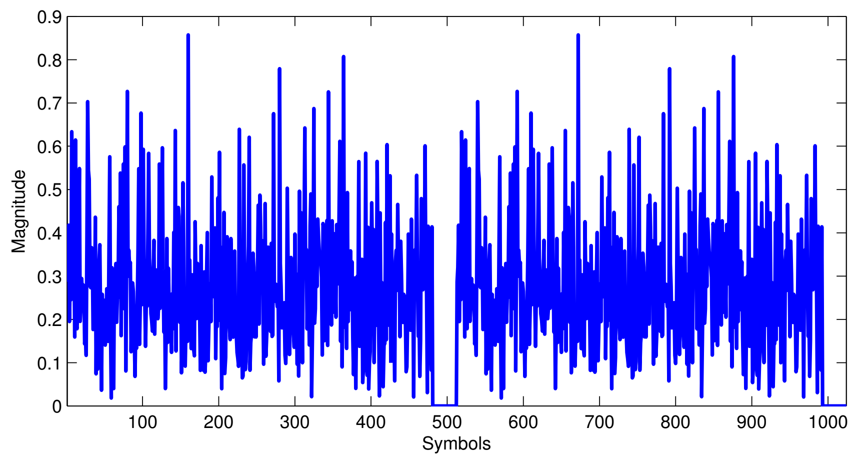
Figure 3-52: FFT output symbol index

The output of the FFT is zero padded before entering the IFFT block. The way the zero padding is carried out depends if we are enabling the use of TD for the calculation of the 1024 point IFFT. The conventional resource mapping method would entail appending an additional 544 zero values to the FFT output as shown in Fig. 3-53a. The location of the zero values is not important since a circular shift may be applied via a constant rotation factor applied to the IFFT output on top of the half subcarrier shift required in order to avoid having a subcarrier around DC. However, utilizing the fact that the number of non-zero samples in our input signal is less than half of the total IFFT size, we employ TD to reduce the complexity and power requirements of the calculation by preforming two 512 point IFFT calculations instead of a single 1024 point calculation. In order to do so, we need to pad the original signal by only 32 zero values and then re-transmit the signal again and apply a fixed rotation at the IFFT block input. This zero padding scheme is

demonstrated in the captured output data magnitude shown in Fig. 3-53b.



(a)



(b)

Figure 3-53: FFT output after zero padding (a) w/o and (b) w/ the use of transform decomposition

Upon exit from the IFFT block we must again re-order the output symbols. This time the index bit reversal is much simpler since the IFFT size factors comprise of radix-2 only, therefore the index is simply a binary bit reversal of a regular 10 bit counter output. The captured output index value for the first few dozen samples is shown in Fig. 3-54. A cyclic prefix is now added to the data by re-iterating the last samples of the OFDM symbol and appending them to the beginning of the symbol. This slightly degrades the overall output spectrum but helps to mitigate effects of delay path spreading leading to ISI and Inter-Carrier Interference (ICI). The baseband output power spectral density of the

measured IFFT output symbols is plotted in Fig. 3-55 for the case of adding a cyclic prefix and without. As can be seen, the total occupied bandwidth of the signal is 7.2 MHz as desired. It is also important to note that this is a two-sided spectrum and it is not symmetric around DC, since the originating baseband signal is complex and not real-valued.

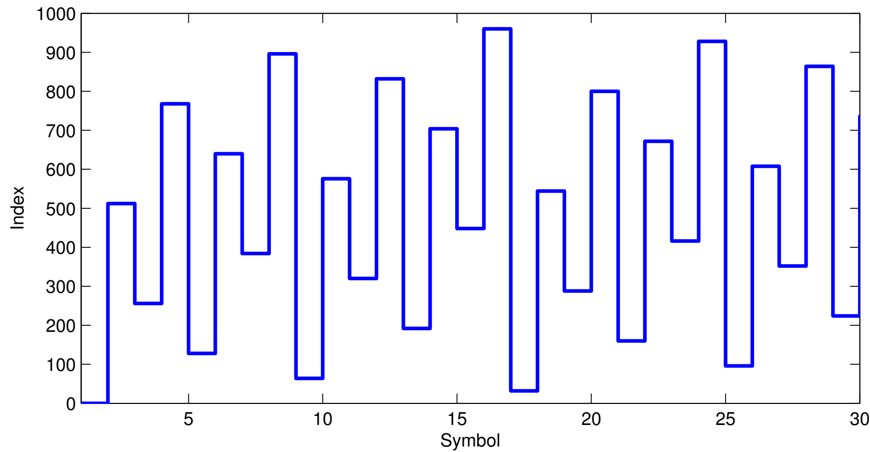


Figure 3-54: IFFT output symbol index

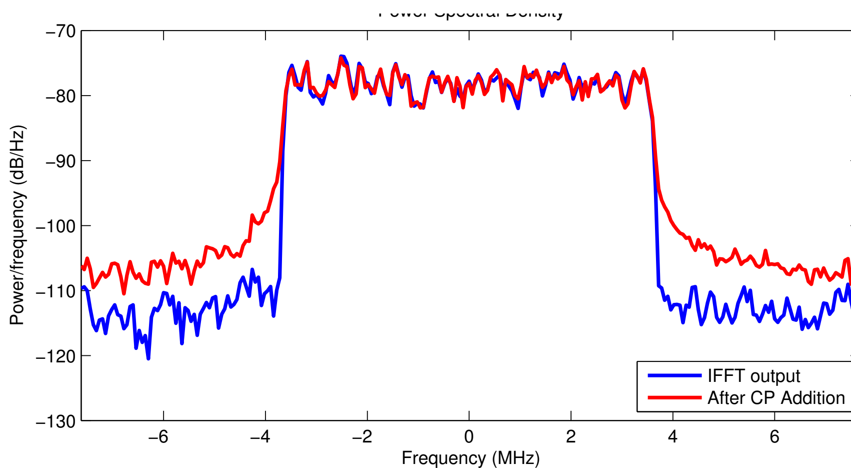


Figure 3-55: IFFT output power spectral density

The final step in the signal processing flow is the upsampling by 8 and interpolation filter. Each polyphase filter operates at the same digital baseband clock frequency, which is 15.36 MHz in this test scenario. The filter output is interleaved between the 8 filter banks at 122.88 MHz. The output spectrum of the output is plotted in Fig. 3-56. The data occupied

bandwidth remains 7.2 MHz, however it is shaped over a 122.88 MHz range.

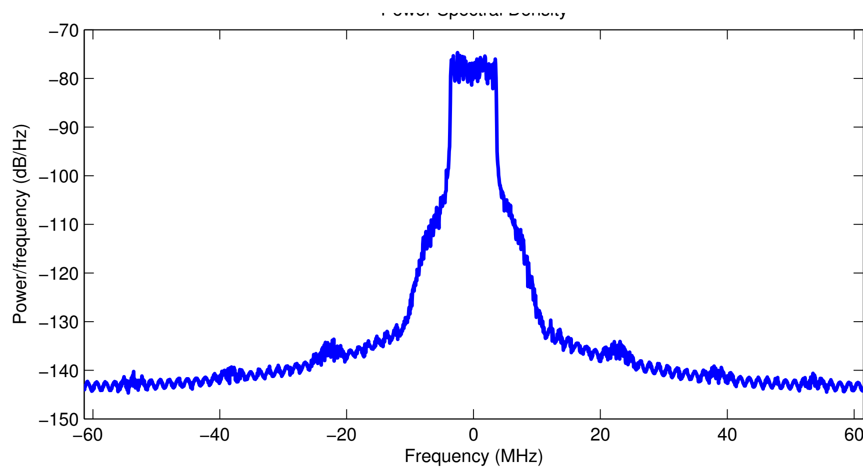


Figure 3-56: Interpolator output power spectral density

Output triggers captured from the chip during operation illustrate the signal processing flow. The following signals were captured during the calculation of 4 RBs in a 1.4 MHz LTE bandwidth. This setting was chosen to be different than the example above simply to reduce the time between processing events for easier illustration of the control signals. The various control signals can be seen in Fig. 3-57. The *FFT Start* signal indicates the beginning of an FFT calculation, and it occurs every OFDM symbol time period (66.67 μs when there is no cyclic prefix). The *FFT Done* signal indicates that the FFT block has completed its calculation (of 48 points in this example). The *FFT Enable* signal indicates when the FFT block is active and processing symbols, it can be seen that the enable signal goes low and shuts off the FFT block whenever it is done calculating and while waiting for the IFFT and CP addition blocks to conclude their operation. Similarly the signals *IFFT Start, Done* and *Enable* all correspond to the same logic for the IFFT block. The *Zero Signal* indicates when high, that zero values are fed to the input of the IFFT. We observe the behavior of the signal in the presence of using TD for the signal processing, resulting in a shorter zero padding signal high time occurring twice in this case since we are performing two IFFT calculations instead of one.

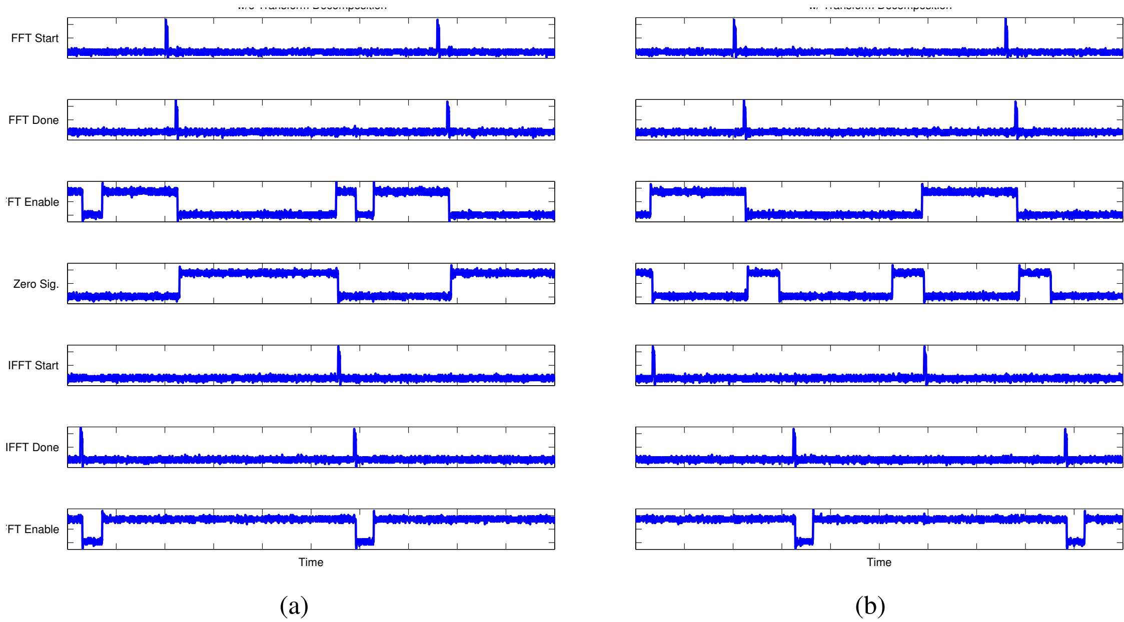


Figure 3-57: Internal trigger signals (a) w/o and (b) w/ the use of transform decomposition

3.5.2 Power

The scaling of the OFDM symbol sizes and throughput rate as a function of the LTE bandwidth being used allows reducing the power supply voltage from its nominal value of 1 V as the frequency scales. Thus, we still achieve full functionality of the processor with a supply voltage of 0.61, 0.68, 0.75, 0.89, 0.96 and 1 V for a bandwidth of 1.4, 3, 5, 10, 15 and 20 MHz respectively. The dynamic switching power of the circuit is modeled as

$$P_{dyn} = \alpha C V^2 f_{clk} \tag{3.80}$$

With C being the effective gate capacitance, V the supply voltage, f the clock frequency and α is the activity factor for the circuit operation. We see that for lower bandwidth (corresponding to lower clock frequencies) we achieve reduction in power consumption from both the scaling of the frequency as well as from lowering of the supply voltage.

The total energy consumed in the calculation of an OFDM symbol will scale as well, since the energy consumption can be approximated from the power via the execution time

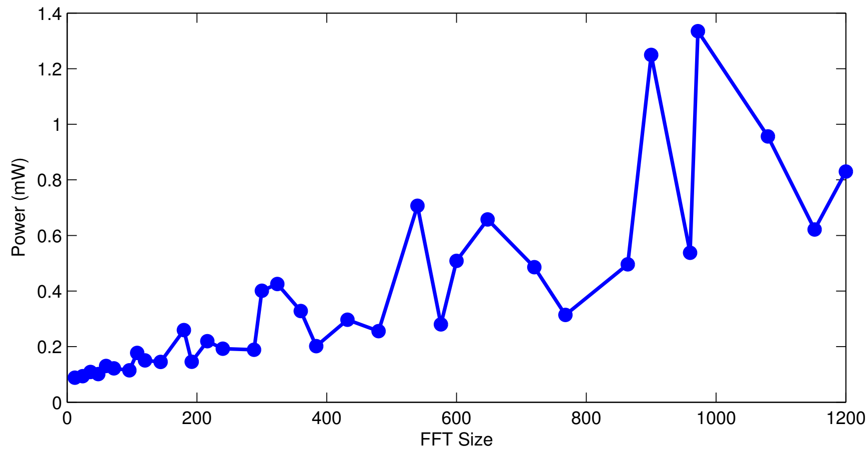
of the calculation in N clock cycles as

$$E_{dyn} = P_{dyn} T_{clk} N_{cycles} = \alpha C V^2 N_{cycles} \tag{3.81}$$

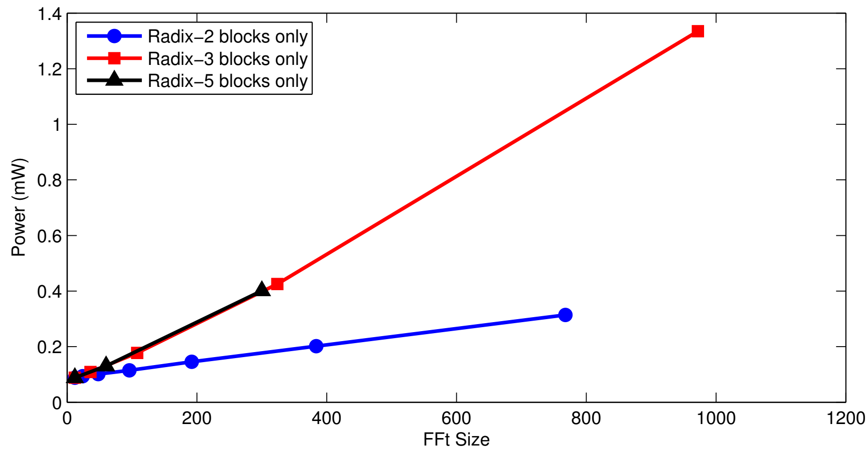
In general, simply reducing the clock frequency of a circuit will not result in the reduction of the energy consumption as seen since the execution time will be longer. However, in our LTE application the number of clock cycles scales along with the clock frequency to maintain a fixed execution time therefore we still benefit from both frequency scaling as well as supply voltage reduction for energy savings. These calculations neglected the important contribution of leakage power which will eventually limit the above benefits and the possibility for energy reduction.

The power consumption of the FFT block was measured across all FFT calculation sizes ranging from 12 points (1 RB) up to 1200 (100 RBs) at the maximum bandwidth and clock speed of 30.72 MHz to represent the worst case scenario. The results of the power measurement are plotted in Fig. 3-58a. The power consumption seems erratic at first glance and not rising monotonically as one would perhaps expect as the FFT size increases. However, it is important to recall that the FFT block comprises of a chain of different radix FFT butterflies and each point corresponds to a different factorization and butterfly usage. Since the radix-3 and radix-5 butterflies are considerably more complex and consume more power, their use is more “expensive” in an FFT calculation. This point is illustrated if we separate the power measurements to include sets of only homogeneous butterfly types. In Fig. 3-58b we see such a breakdown, where we examine the FFT power for 1, 2, 4, 8, 16, 32 and 64 RBs to observe the effect of adding further radix-2 butterflies to the calculation. Similarly we inspect the power for 1, 3, 9, 27 and 81 RBs and 1, 5 and 25 RBs to uncover the underlying effect of the addition of radix-3 and radix-5 butterflies respectively. This measurement indeed validates our assumption and demonstrates the expected linear increase in power with larger FFT size. We also observe the fact that the power consumption of one radix-5 butterfly is greater than that of one radix-3 butterfly, which in turn exceeds that of a simple, standard radix-2 butterfly.

The overall power measurement of the signal processing core including both the FFT



(a)



(b)

Figure 3-58: FFT block power consumption (a) all FFT sizes and (b) separated into homogeneous radix butterfly blocks

spreading and IFFT mapping was conducted using the same voltage scaling as described previously. The power consumed at the largest bandwidth and IFFT size was measured for each possible FFT size ranging from 12 to 1200 without the use of TD. A second measurement was then carried out while utilizing TD to reduce the required size of the IFFT calculation for lower size FFTs. The ratio between the power consumption with and without the use of TD is plotted in Fig. 3-59. The different power of 2 IFFT sizes are marked on the plot with dashed vertical lines. It can be clearly seen that the use of TD contributes to power saving in a step-wise fashion due to the discrete jumps from one IFFT size to the other. Up to 24% power savings at the extreme case of using only one RB in a 20 MHz bandwidth is achieved. The power savings reduce as the number of non-zero samples increases until they are larger than half the IFFT size and there is no power saving compared to the regular mode of operation.

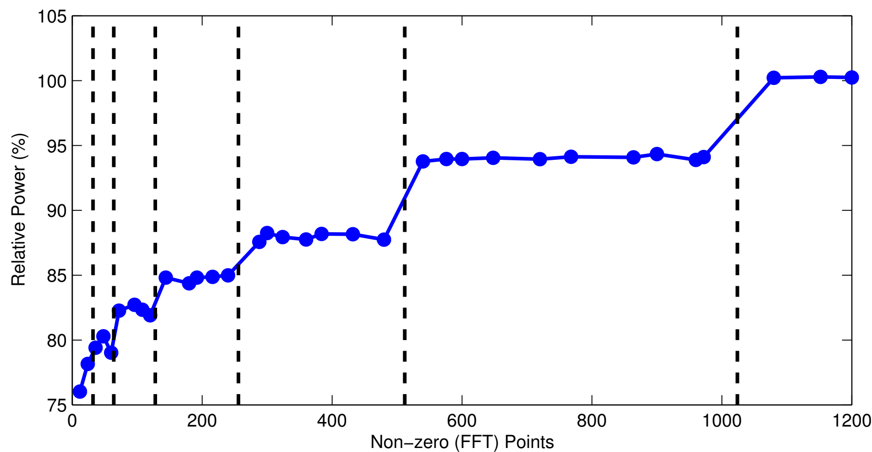


Figure 3-59: Relative total power when using transform decomposition as a function of FFT size (i.e. number of non-zero samples at input)

The overall power was also measured for all combinations of number of RBs and IFFT sizes (i.e. LTE bandwidth). The maximum power consumption measured was 0.08, 0.16, 0.40, 1.17, 2.06 and 2.93 mW for each LTE bandwidth respectively. The total energy consumption was calculated by multiplying the power consumption by the execution time, which is identical to all cases at 66.67 μs, resulting in energy consumption ranging from 5.2 to 195.3 nJ. Furthermore, the energy per sample was calculated in order to have a better

comparison between channels and to other designs in the literature by dividing the energy by the number of samples. This effectively eliminates the impact of the frequency scaling and only reflects the benefit of voltage scaling. The energy per sample calculated was between 40.3 to 95.4 pJ and is plotted in Fig. 3-60.

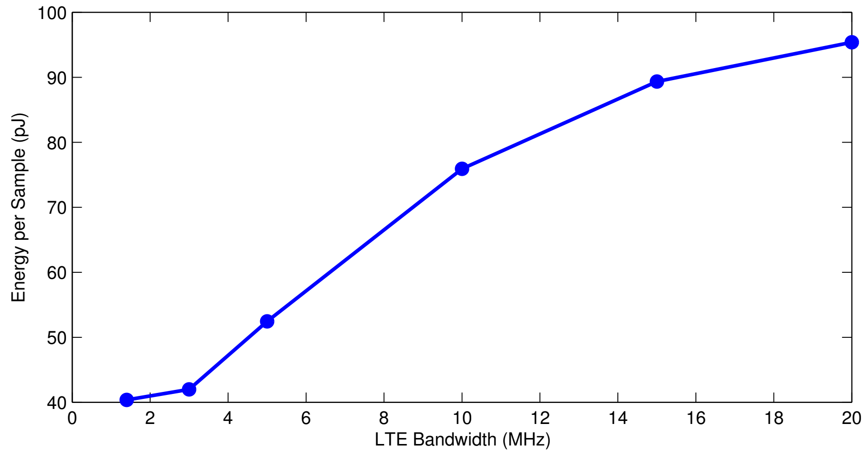


Figure 3-60: Energy per sample for each LTE bandwidth operating mode

The power consumption of the interpolation FIR filter was measured to be between 0.02 and 0.92 mW for the various LTE modes. This illustrates also the benefit of using such a filter compared to using a larger size IFFT calculation. For example, for a 3 MHz channel, using 15 RBs, requiring a 256 point IFFT we shape the spectrum out to 30.72 MHz by using a base clock rate of 3.84 MHz, the 256 point IFFT and the interpolation filter, consuming $0.16+0.05=0.21$ mW. Alternatively, we use the 2048 point IFFT with a base clock rate of 30.72 MHz and without using the interpolation filter, but then we would consume 2.18 mW of power (less than the maximum of 2.93 mW since we use TD on the maximum 15 RBs of the 3 MHz channel), which is an order of magnitude higher than the first option using the interpolation filter.

The power and energy measurement and calculation results are summarized in table 3.9. These results are consistent with the expected trends denoted in (3.80) and (3.81).

Table 3.9: Power and energy consumption summary

LTE BW (MHz)	1.4	3	5	10	15	20
IFFT size	128	256	512	1024	1536	2048
Max FFT size	72	180	300	600	900	1200
Clock freq. (MHz)	1.92	3.84	7.68	15.36	23.04	30.72
Supply voltage (V)	0.61	0.68	0.75	0.89	0.96	1
Filter Power (mW)	0.02	0.05	0.13	0.36	0.63	0.92
FFT/IFFT Power (mW)	0.08	0.16	0.40	1.17	2.06	2.93
Energy/FFT (nJ)	5.2	10.7	26.9	77.7	137.2	195.3
Energy/Sample (pJ)	40.3	42.0	52.4	75.9	89.3	95.4

3.5.3 Comparison

A comparison of the proposed design to other similar work targeted at LTE signal generation is presented in table 3.10. The key features and results of each work is detailed. In an attempt to perform a fair comparison across designs, the area value is normalized to the 28 nm process node used in this work via

$$\text{Normalized Area} = \frac{\text{Area}}{(\text{Tech}/28\text{ nm})^2} \tag{3.82}$$

The energy consumption for each design was also normalized in order to remove the impact due to difference caused by the process technology node, the FFT size and the width of the datapath [136] via

$$\text{Normalized Energy} = \frac{\text{Energy/FFT}}{\text{FFT Size} \times \left(\frac{2}{3}\frac{\text{Wordlength}}{16} + \frac{1}{3}\left(\frac{\text{Wordlength}}{16}\right)^2\right) \times \left(\frac{\text{Tech}}{28\text{ nm}}\right)} \tag{3.83}$$

The property values for the proposed design are mentioned for the total of both the FFT and IFFT blocks together. Where relevant, the breakdown for the individual blocks is denoted in parentheses for the IFFT and FFT blocks respectively.

It is important to note that most other FFT processor designs found in the literature

focus on radix-2 implementations, which simplifies the design considerably and do not satisfy as-is the complete requirements for the SC-FDMA signal generation portion of LTE UL. The most relevant design for comparison is the one found in [139] which implements the same functionality as the proposed design for LTE applications performing the full SC-FDMA signal generation (although not implementing the 1536 point, 15 MHz mode). The proposed design offers a reduction in gate count as well as 2x reduction in area and 4.3x improvement in energy efficiency compared to the previous design.

The proposed design presents a comparable or lower area compared also to radix-2 only pipeline designs (especially if we only take into account the IFFT block area), the design in [165] uses a cached FFT architecture and therefore naturally has a smaller area compared to pipelined designs in exchange for a lower throughput. The energy efficiency is also better than most other designs, except for the one reported in [157] which has a normalized energy per sample of 12-32 pJ. However it should be emphasized once more that the proposed design carries out both FFT and IFFT calculations concurrently and handles a much wider range of non-trivial FFT sizes.

3.6 Conclusion

In this chapter we presented an overview of some of the requirements needed to support future high data rate communication schemes. An overview of the fourth generation cellular communication protocol LTE detailed the need for advanced signal processing in order to meet the ever increasing demand of high bandwidth efficiency and utilization. The use of OFDMA inspired the examination of efficient techniques to preform DFT calculations in hardware and a review of the various algorithms and existing hardware topologies.

An LTE digital baseband processor was designed to assist in the energy efficient creation of LTE compatible SC-FDMA signals. The processor enables mapping of digital bits to modulated complex BPSK, QPSK, 16-QAM and 64-QAM symbols. It supports transform precoding of the symbols via a variable mixed-radix pipeline FFT of size 12-1200. The processor enables RB mapping to support LTE OFDMA and SC-FDMA signal generation by performing a variable length radix-2/3 IFFT of size 128-2048 or 1536 points. A cyclic

Table 3.10: Comparison to other LTE OFDMA signal generation processors

Work	**Proposed**	**[139]**	**[157]**	**[165]**
Architecture	SDF	DEM	SDF	Cached FFT
FFT size	128 $\sim$ 2048, 1536 + 12 $\sim$ 1200 DFT	128 $\sim$ 2048 + 12 $\sim$ 1296 DFT	128 $\sim$ 2048/1536	128 $\sim$ 2048/1536
Technology	28 nm	0.18 μm	65 nm	0.18 μm
Word width (bits)	2×16	2×16	2×12	2×16
Output sorting	Yes	Yes	No	No
Voltage (V)	0.61 $\sim$ 1	1	0.45	1.8
Gate Count	681K (170K + 511K)	798K (316K + 482K)	1100K	98K
Area (mm^2)	0.31 (0.09 + 0.22)	25	1.375	1.932
Normalized Area (mm^2)		0.60	0.26	0.05
Clock (MHz)	1.92 $\sim$ 30.72	122.88	20	35
Throughput (MS/s)	1.92 $\sim$ 30.72	122.88	80	8.72
Power (mW)	0.08 $\sim$ 2.93	320	4.05	11.29
Energy/FFT (nJ)	5.1 $\sim$ 195.3	5333.3	2.5 $\sim$ 103.7	64 $\sim$ 2641.9
Normalized Energy/Sample (pJ)	40 $\sim$ 95	405	12 $\sim$ 32	78 $\sim$ 201

prefix is added to the symbols to mitigate delay spread and reduce ISI and ICI. The output samples are upsampled and interpolated by a factor of 8 to improve the output power spectrum of the signal.

The design utilized efficient hardware techniques such as pipelining, clock gating, voltage and frequency scaling, CSD constant multipliers and efficient complex multiplication techniques to improve the system efficiency. Furthermore, use of CORDIC rotators instead of complex multipliers, utilization of high order radix butterflies, latch-based delay lines and memory sharing among butterflies in the DFT blocks contributed to the overall system efficiency, simplicity, reduction of memory and area savings. Finally, Taking advantage of the specific data statistics of LTE signals and their generation scheme we implemented the use of the TD algorithm to reduce power consumption whenever possible.

The proposed design was fabricated in 28 nm CMOS and the core active area is 1.17 $\times$ 0.675 mm^2, with the FFT and IFFT blocks occupying a total area of 0.31 mm^2. The DFT calculations consumes between 0.08 to 2.93 mW from a 0.61 to 1 V supply at an operating frequency between 1.92 to 30.72 MHz for the respective LTE bandwidth from 1.4 to 20 MHz. The energy per sample ranges from 40 to 95 pJ for the worst case scenario. The use of TD enables reducing the consumed power by up to 24% over the various cases of RB utilization. The interpolating filter consumes between 0.02 to 0.92 mW and enables extended shaping of the signal output spectrum to reduce the effects of the ZOH sampling. The design presents a 4.3x improvement in energy efficiency and 2x reduction in area over state-of-the-art full SC-FDMA UL LTE signal generation designs.

These techniques along with other published work on low power, high throughput systems will enable future complex system design. Enabling such advancements helps to reduce the cost of hardware and make it widely available for more people around the world to connect and share information with ease. Combined with further improvements in transistor fabrication, packaging and heterogeneous integration, new system level benefits are achieved to deliver ever increasing performance and functionality. This brings us another step closer in the design of an energy efficient RF system integrated in a 3D package. Optimizing the digital block and integrating it along with analog circuits required for RF applications will enable creation of improved, energy-efficient 3D-ICs.

Chapter 4

Analog-Digital 3D Integration

4.1 Introduction

Three Dimensional Integrated Circuits (3D-IC) have the potential to allow further advancement in system functionality and power efficiency beyond those offered by conventional lithography scaling. The ability to stack several dies vertically and connect them using short, low-parasitic Through Silicon Vias (TSVs) enables higher level of functional integration with shorter interconnect. This potentially enables lower power consumption as well as smaller foot print and also heterogeneous integration of different functionalities and process technologies in a single stack.

3D-IC have shown the potential to improve system efficiency in areas such as memory [79], processors [15,77], optics [18–20] and more. These benefits however come at a cost of new design challenges and issues including thermal, mechanical, power and signal integrity as well as electromagnetic noise coupling as explored in chapter 2. RF communication is an application which may greatly benefit from such 3D integration due to the increasing demands for data rate at low power levels as described in chapter 3. These applications might suffer however from many issues due to the close integration of high-speed, noisy digital circuits along with sensitive, analog and RF circuit blocks.

In this work we explore the system level benefits, as well as challenges encountered in the design of a 3D-IC for RF applications. Various partitioning techniques are explored for the integration of digital, analog and RF circuits in a two tier, 3D die stack using

TSVs. The design implements a part of the LTE UE UL channel and includes baseband signal processing functionality as described in detail in chapter 3 as well as digital to analog conversion, upsampling and conversion to RF. The design is also compared to other partitioning solutions, emulating single 2D die implementations and 2.5D interposer solutions.

The outline of the chapter is as follows. The details of the technology, circuit implementation and analysis are presented in section 4.2. Section 4.3 presents the measurement setup and results. We present our final conclusions in section 4.4.

4.2 Transmitter Chain Design

In order to explore the impact of digital and analog circuit co-existence in 3D-IC, the various effects of vertical close proximity and the overall potential benefits of such an integration, we have decided to focus on the design of part of a cellular transmitter chain. The block diagram of the components implemented in this design are shown in Fig. 4-1. The implementation includes part of the digital baseband, which is responsible for modulation and constellation mapping, digital to analog conversion, Local Oscillator (LO) generation and upconversion mixing to RF. The preceding digital baseband stages of coding the digital data as well as the RF power amplification, RF band filtering and antenna ports were not integrated in this design.

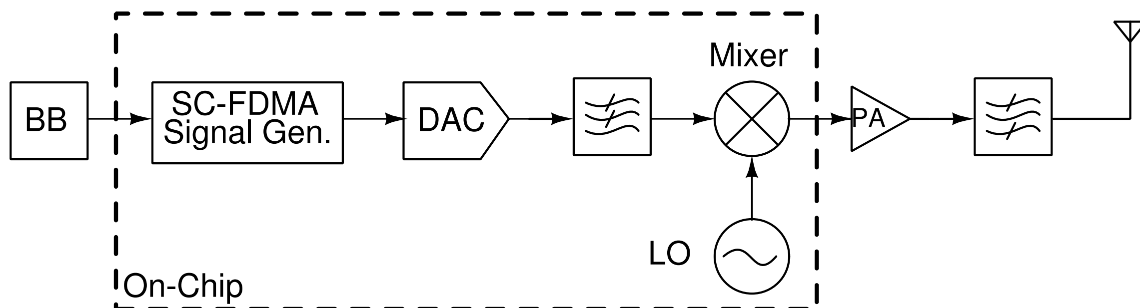


Figure 4-1: Partial transmitter chain block diagram

The core components of the system which are implemented are meant to act as a proof of concept and a representative example of some of the core circuits and sub-systems which

are prevalent in many communication systems. Our specific implementation focuses on LTE cellular communication, but uses it only as a framework and reference point for the design and specifications, and a leading example to high data rate modern communication protocols. However, the main topics of investigation, namely efficient signal processing discussed earlier and analog-digital integration in 3D-IC are general issues which may be analyzed in this context.

The SC-FDMA generation block transforms raw data bits into complex, modulated SC-FDMA signals for transmission. These in turn are converted to an analog signal via the DAC module, which is filtered and then upconverted to RF via the mixer and LO output. The final modulated RF output is driven off chip for measurement and analysis.

4.2.1 Chip Layout and Partitioning

In order to explore the various aspects of efficient signal processing as well as digital and analog co-existence in a 3D-IC RF communication system, the proposed partial transmitter chain was designed for a 3D-IC stack. The design focuses on implementing the key blocks described earlier - some digital baseband signal processing, conversion to analog, filtering, LO signal generation and upconversion to RF. We further wish to explore two main aspects of the analog-digital integration for 3D RF applications. First, we wish to be able to compare the 3D-IC topology and approach to other possible system partitioning approaches such as multi-chip, SoC and 2.5D (Silicon interposer). Secondly, we wish to examine within these different packaging options the trade-offs between different partitioning possibilities between the digital and analog domains.

The main area of research will focus on aspects of new design challenges arising from the use of 3D integration such as noise coupling effects on power and signal integrity and how to overcome them. The second main aspect will focus on the potential for new design structures and topologies which will enable us to gain new benefits in 3D design over other approaches in areas such as scaling and power, not only in the circuit and block level, but also on the system scale level.

The 3D stacking process used in our design is a via-last ,back-to-face, die-to-die post

processing step. There is only one die layout which is created and fabricated in a regular 2D, 28 nm, bulk CMOS process. Half of the wafers are back-grinded, and the backside Silicon layer is thinned to around 60 μm. The thinned wafers are back-drilled to place the TSVs, while the other non-thinned wafer has corresponding microbumps added. The wafers are later diced, and the individual die pairs are rotated by 180° relative to each other, and placed in a 3D stack to create the 3D-IC where each TSV in the thinned, rotated top die tier is bonded to a microbump on the bottom tier die. The die cubes are later placed in a package and wire-bonded via the exposed top-tier die pads to the package pads to be used on a PCB for testing. This process is crudely illustrated in Fig. 4-2.

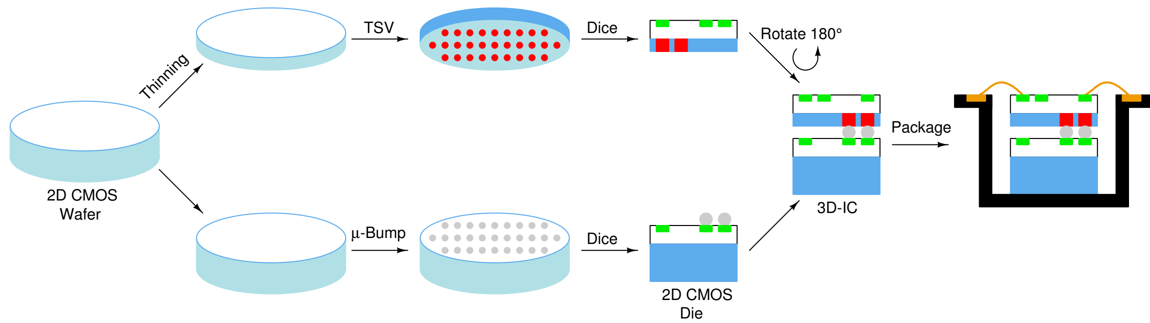


Figure 4-2: Schematic illustration of 3D-IC packaging process (dimensions not to scale)

The fact that we only utilize one set of wafer masks for the 3D-IC fabrication, and utilize a 180° rotation to create the die stack is a constraint not normally found in commercial 3D-IC fabrication, but is necessary in these preliminary research experiments for cost savings. This does incur some "waste" in the final design since not all of the silicon cube area is fully used, however we have made an effort in the design methodology described here to utilize as much of the area as possible. One approach to the die layout would be to designate strict areas which will act as "top" and "bottom" tiers. This is a simple approach to implement, however it will result in 50% of unused die area, since in the final cube layout we will have both the desired top-over-bottom portion, but also an equivalent useless bottom-over-top area. In contrast, in our design we have used a more modular approach, instantiating different key circuit blocks across the die and allowing flexible cross-routing to enable a much higher degree of die area utilization and a more expansive set of possible

testing scenarios.

A simplified block diagram of the die layout is shown in Fig. 4-3. The TSV and microbump locations are not illustrated on this figure to retain clarity, they are mostly concentrated in the crossbar module regions. The die also includes a pad ring consisting of 93 wirebonded pads, though these are not illustrated as well in order to avoid excess clutter in the image. Lastly, the clock and configuration scan chain blocks are not illustrated as well and will be detailed separately later in this chapter. All of the main key blocks shown in the transmitter schematic of Fig. 4-1 can be seen in the chip layout, namely a digital baseband block for generation and processing of SC-FDMA signals for LTE UL communication, two DAC instances, three quadrature VCOs and upconversion mixers for translating the baseband signals to RF. Along with these blocks, some simple crossbar blocks represent the ability to manipulate the chain configuration in order to achieve different routing and partitioning options as well as emulate other packaging topologies. In the following examples we will review a partial subset of the possible chip routing configurations to achieve these various options. The routings will be highlighted on top of the basic die layout in order to illustrate how the different configurations are achieved from this basic single die layout.

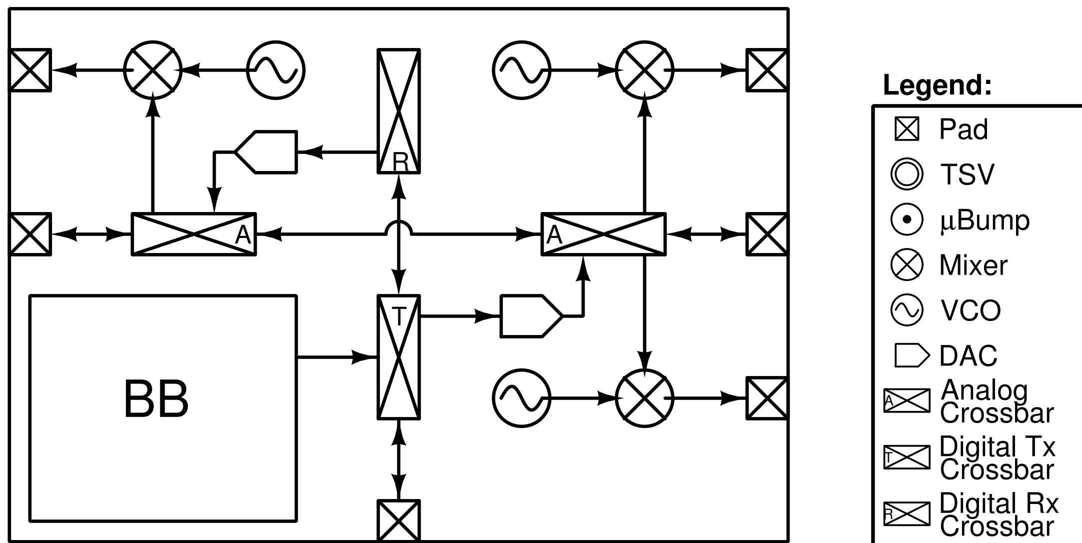


Figure 4-3: Block diagram of single cube tier in 3D stack

One of the popular approaches, gaining many applications in recent years for the close

integration of complex systems is SoC. In this approach we simply wish to integrate several key system components which were traditionally separated into separate dies onto one single, 2D die. This enables close integration and use of high density silicon level routing, however all system blocks share the same die and substrate and are potently susceptible to noise through substrate coupling. These are somewhat alleviated by maintaining large spatial distance between the digital and sensitive analog blocks, as well as large guard ring structures to mitigate the noise, however that incurs a cost in die footprint as well as long routing which could cause high power consumption for high data throughput. Another disadvantage of complete SoC integration is the fact that we are inherently constrained to use only one material and process node for the entire system, so it cannot be optimized for different blocks and functionality. The single chip SoC-style implementation scenario is obtained in our die via the routing shown in Fig. 4-4. All required components of the transmitter chain are active on one die tier. Namely the top die tier is active while the bottom die tier in the stack is completely powered off and left unused. Different instances of the DAC, VCO and mixers may be used for this configuration, resulting in different wiring lengths and separation between the various blocks.

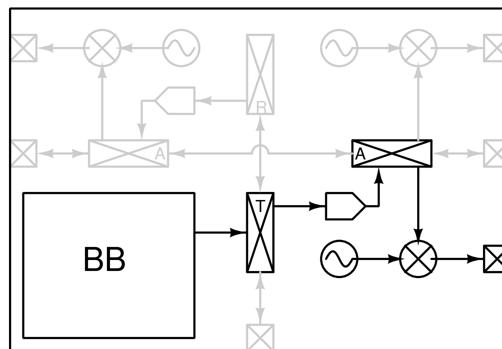


Figure 4-4: Single die, SoC-like operation emulation (only top tier of cube shown)

Predating the SoC solution is the older, conventional multi-chip solutions, where the different building blocks are each fabricated on a separately packaged die, and the system is integrated on the board level. This allows for a very large degree of separation between the different blocks and minimize the coupling and interference between them. However, in this approach the data is transmitted and needs to drive board level traces and associated

package parasitics which requires larger, power hungry drivers. We are able though to manufacture each component in its ideal, optimal processing and packaging technology. Within the multi-chip approach we also distinguish where the data partitioning occurs, whether it is after the digital baseband processing, requiring to drive multiple high-speed bit lines from chip to chip but with high levels of noise immunity, or alternatively, to transmit the baseband signals after conversion to an analog signal. This saves on pin count, since we do not need to drive a large bus, and also saves on bandwidth, but the signals are now more sensitive to noise and interference while being transmitted from the baseband chip to the RF conversion chip. A sample routing topology for the multi-chip scenario, for both partitioning options, is shown in Fig. 4-5. In this scenario we again utilize only the top tier die of each cube, while leaving the bottom tier powered down since we are still emulating regular 2D chips. Several other internal routings are possible in each die for slightly different routing options between the blocks.

Utilizing the advent of TSVs, we may create the system as several dies on top of a silicon interposer, a packaging technique also referred to as 2.5D. Interposer packaging allows having the benefit of heterogeneous integration of separate, different process dies in close proximity while using chip and silicon level interconnect to pass data between the dies. The silicon interposer layer has no active devices of its own, and merely acts as a "bridge" between the different dies, capable of supporting dense, low-parasitic interconnect. The routing however still needs to traverse along a 2D planar route between dies, and there is no area or footprint reduction compared to the SoC solution, and might be even larger due to the interconnect overhead. In our design there is no real passive interposer layer, however we may view the bottom tier die as a passive layer, only used for routing of signals without any active circuits and devices. In this manner we create configuration scenarios emulating the 2.5D interposer set-ups with the existing dies. The data will be passed through TSVs to the bottom tier, and then routed to another crossbar location where it is brought back up again to the top tier for further processing. We can again support different partitioning options by selecting the point where the data is transferred via the bottom "interposer" after either the digital or analog conversion as shown in Fig. 4-6. It should be noted that this mode of operation is not completely identical to a true interposer solution since

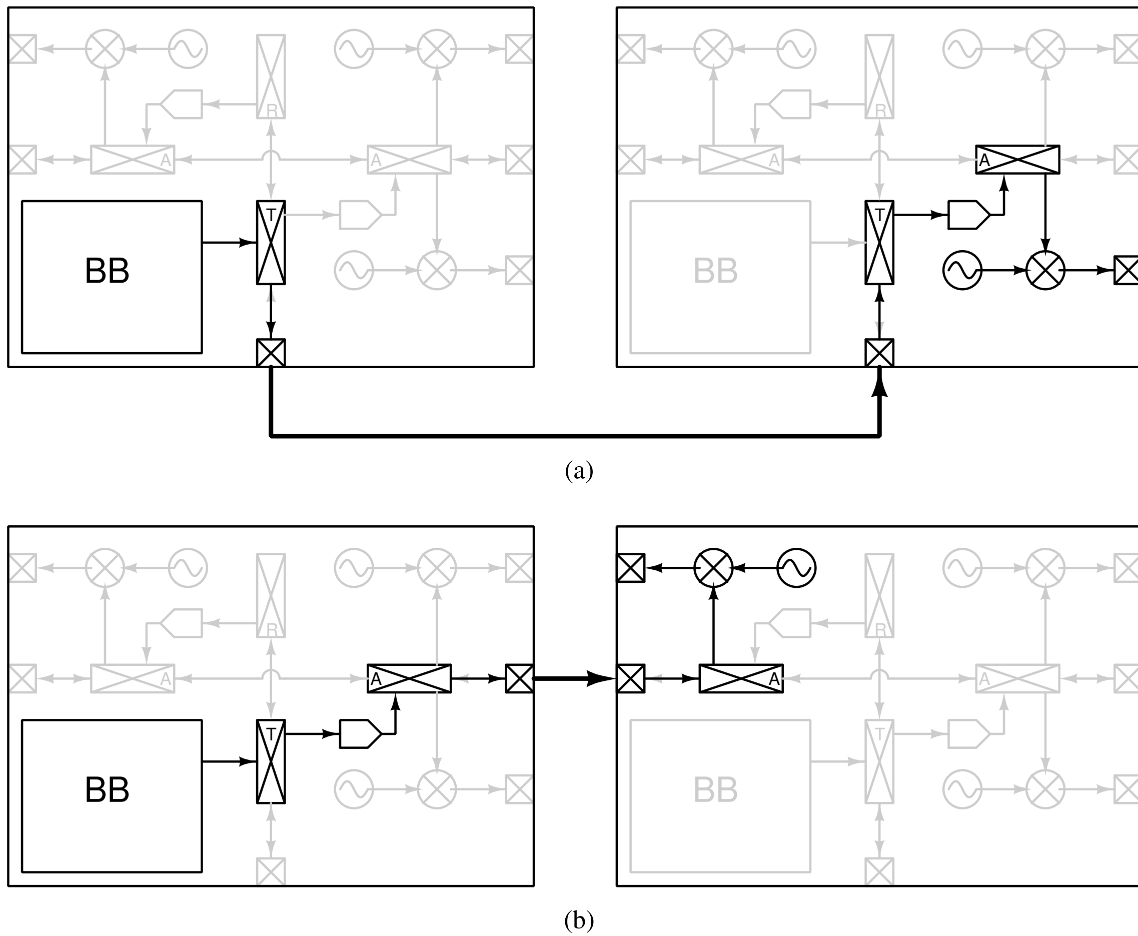


Figure 4-5: Two dies, partition at (a) digital boundary and (b) analog boundary (only top tier of cubes shown)

ultimately the "start" and "end" circuit blocks are still on the same die and share the same substrate and are not truly separated as they would be in a 2.5D solution. However, the multiple block instances at our disposal do allow for some degree of spatial separation and a reduction of these effects to some degree and allow for a reasonable comparison of link power consumption in this case.

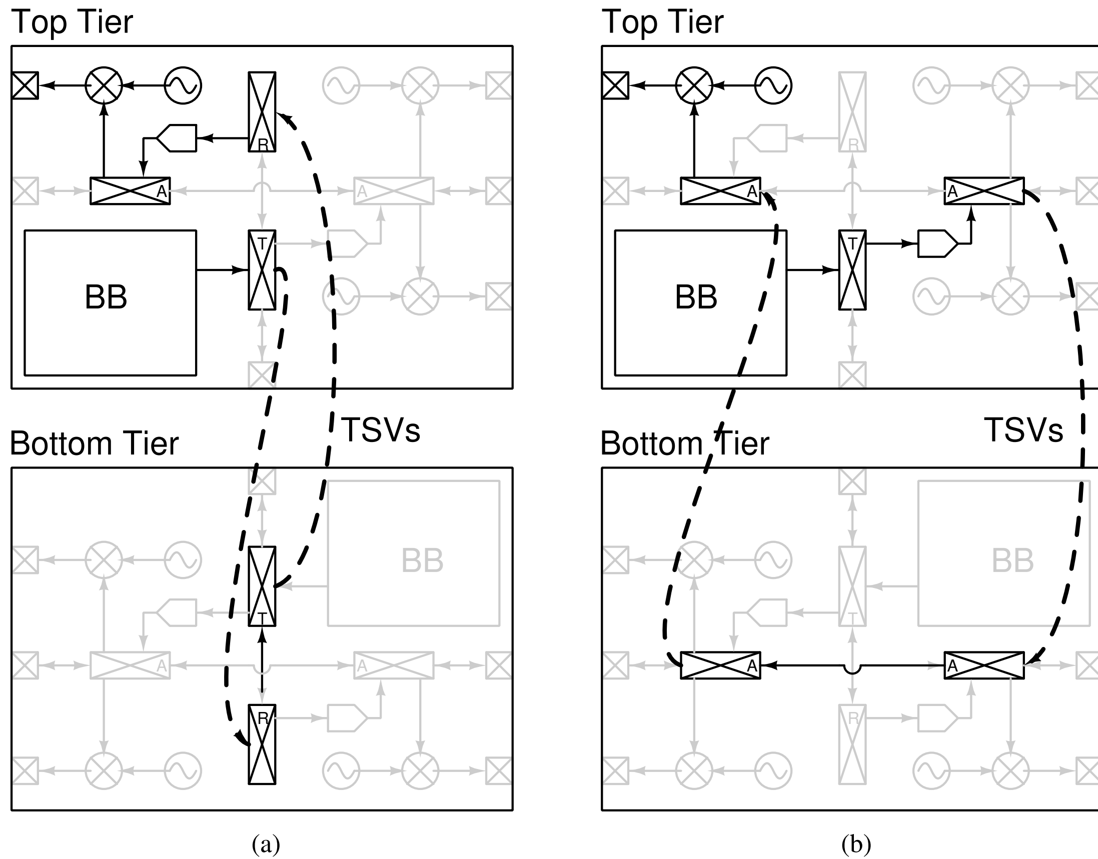


Figure 4-6: 2.5D operation emulation with partitioning at (a) digital boundary and (b) analog boundary

Finally, we are able to configure the die to act as originally designed, as a 3D-IC system. In a 3D stack we can potentially integrate heterogeneous die processes (although not done so in this work due to the specific constraints detailed earlier regarding die stacking). We can obtain very good separation of substrates between the dies, as well as very short, low-parasitic, dense vertical interconnect between the die tiers. The question of digital/analog partitioning remains an important one in this topology as well and the possible routing configuration, including an active bottom tier die, is illustrated in Fig. 4-7. The digital data

is generated and processed on the bottom tier die in both partitioning scenarios, and we can either choose to convert the digital data to analog baseband signals or pass them as-is to the top tier die via TSV for further processing and output,

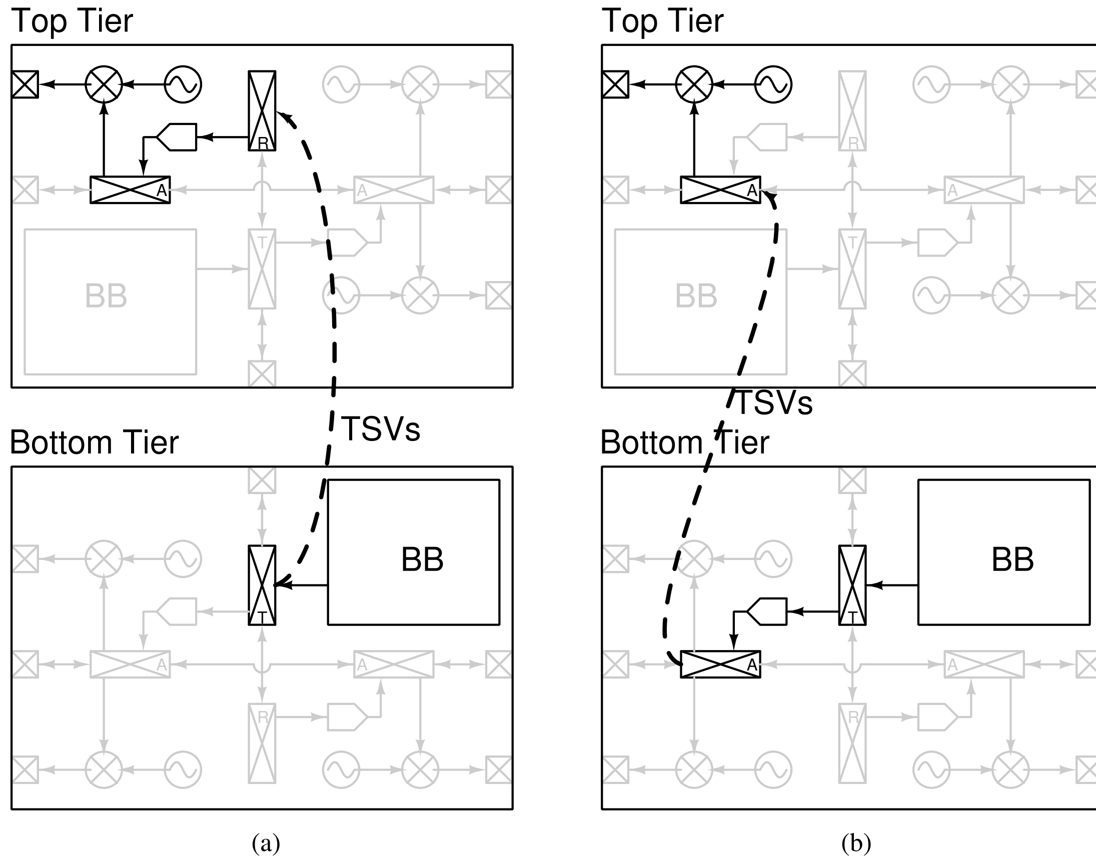


Figure 4-7: 3D-IC operation with partitioning at (a) digital boundary and (b) analog boundary

The different partition options described above are summarized in table 4.1. The table lists each topology along with its partition boundary, which circuit blocks are active on each tier of each chip as well as the type of TSVs used for communication if at all. Routing digital data through the TSV is denoted as *Hi-speed*, whereas routing the analog DAC output across the TSV is marked as *Baseband* to emphasize the bandwidth of the signals passing between the stack tiers. The partition index will be used later on for easy reference to the different options studied.

The ability to support this myriad of operation modes with the exact same basic die layout and building blocks will enable us to make a more accurate, fair comparison between

Table 4.1: System partitioning summary

#	Topology	Partition Boundary	Chip A Top	Chip A Bottom	Chip B Top	TSV	Illustration
1	SoC	None	BB, DAC, VCO, Mixer	Disabled	Disabled	None	Fig. 4-4
2	Separate Dies	Digital	BB, Digital pads	Disabled	Digital pads, DAC, VCO, Mixer	None	Fig. 4-5a
3	Separate Dies	Analog	BB, DAC, Analog pads	Disabled	Analog pads, VCO, Mixer	None	Fig. 4-5b
4	2.5D	Digital	BB, DAC, VCO, Mixer	Digital routing	Disabled	Hi-speed	Fig. 4-6a
5	2.5D	Analog	BB, DAC, VCO, Mixer	Analog routing	Disabled	Baseband	Fig. 4-6b
6	3D	Digital	DAC, VCO, Mixer	BB	Disabled	Hi-speed	Fig. 4-7a
7	3D	Analog	VCO, Mixer	BB, DAC	Disabled	Baseband	Fig. 4-7b

the pros and cons of each topology. The ability to support different partitioning techniques will further allow exploring the impact of each choice in every scenario and have a good basis for comparison. Further yet, the 3D integration architecture will allow investigation of the key issues of interest in this research involving the impact of placing digital and sensitive analog circuits in close vertical proximity. Observing the impact of digital clock noise on the output spectrum and linearity of analog circuits such as the DAC, VCO and mixer and reviewing different techniques for mitigating these effects. We thus obtain a broader, fuller picture of the potential overall system level benefits of pursuing 3D integration for RF applications and in which scenarios does it seem to offer the greatest benefit.

The high degree of flexibility and reconfiguration does though come at a cost. The added circuitry, routing options and multiple block instantiations requires both much more area than would be necessary in any one specific design and also burdens somewhat on the signal routing. The design tries to minimize these affects and account for them. Overall, the ability to compare the same design at the same technology process for different partitioning schemes and topologies outweighs these drawbacks.

4.2.2 Digital Baseband

The digital baseband module generates the desired modulated signal vectors which encode the communication data. The baseband implements part of the LTE communication protocol for generating UL SC-FDMA signals from raw data bits stored in on-board SRAM memory cells. The data is read from memory, modulated by a QAM constellation mapping to complex symbols. The symbols are then precoded using an FFT operation and allocated to appropriate frequency bins. Following this stage the symbols are mapped to the frequency bins via an IFFT operation and a cyclic prefix is added to each symbol. Finally, the output is upsampled by a factor of 8 and interpolated between data samples. The output of the digital baseband block is 22 bit codewords consisting of real and complex 11 bit words per symbol. The sample rate depends on the channel bandwidth used and ranges from 15.36 MS/s for a 1.4 MHz channel bandwidth, and up to 245.76 MS/s for a 20 MHz channel bandwidth allocation. A detailed description of the digital baseband module and its implementation

details can be found in section 3.4.

4.2.3 Digital Buffering and Level Shifting

Digital data is buffered and driven throughout the chip in various scenarios and across a wide range of conditions. Data is required to traverse long metal routings when emulating integrated SoC operation or be transferred from one die tier to another in the 3D stack through TSV and microbump connections. Different routing options require different drive strength depending on the effective driven capacitance of the load lines. This in turn yields different size and power requirements from the digital buffers along the data path. In order to allow flexibility for testing, as well as enable power saving demonstration, a scalable tri-state buffer was designed for use as the main buffer throughout the chip. The tri-state gated operation also allows bi-directional communication and selection of the directionality of the data flow.

A schematic of the tri-state buffer design used in this work is shown in Fig. 4-8a. When the enable signal is low, the output is disconnected via the high output impedance of the MOS devices. Furthermore, the input signal is also set to a low logic value via an AND gate in order to reduce switching effects at the output coupled from the top and bottom transistors. A set of four, binary weighted scaled tri-state buffers are connected in parallel to form a scalable controlled buffer as seen in Fig. 4-8b. This buffer can be scaled in strength by 4 bits of control, and is also placed back-to-back with an identical buffer to create a bi-directional link, where only one of the buffers is active at any given point and the other is left open. The bi-directional configuration is later used in the digital crossbar connections to set the various routing options (see section 4.2.5). The various routing options and buffer strength settings are controlled via configuration bits loaded by the global chip scan chain (see appendix B).

In general there are two main voltage domains in the chip. An I/O voltage supply of 1.8 V, and a core voltage supply of 1 V. The I/O supply is also used for some of the DAC configuration controls as well as the bias control values. When crossing operating voltage domains some care must be taken to ensure proper operation and ensuring device reliability.

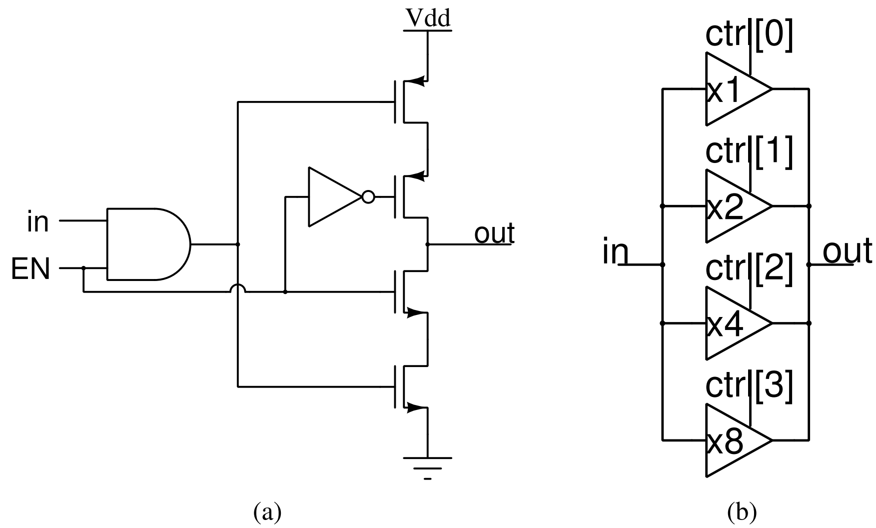


Figure 4-8: Schematic of (a) tri-state buffer and (b) scalable controlled buffer

When crossing from a high voltage domain to a lower one we simply use a standard cell buffer (or inverter) which comprises of appropriate thick oxide devices which can endure the large voltage swing, but connecting it to the lower power supply. This will ensure that the device output will swing between the lower domain voltage rails, but the input devices can withstand the incurred stress of the high voltage input values. An illustration of the node voltages and stress across the transistors for this scenario is shown in Fig. 4-9.

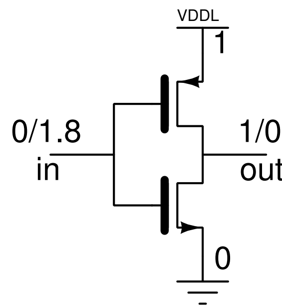


Figure 4-9: High-VDD to low-VDD level shifter schematic using thick-oxide inverter cell

On the other hand, when crossing from a low side voltage domain (such as the core voltage) to a higher supply domain, a more careful approach must be taken. We cannot use the simple thick-oxide inverter technique we used previously, since although this will ensure there is no excess stress across the devices, when the input is high we will not

adequately switch off the upper PMOS device. With a high voltage at its source, and the lower supply at its gate, the transistor will only be partly cut off and still conduct while the bottom NMOS transistor will be open resulting in a large shoot-through current flowing from the high supply to ground through the inverter transistors.

Alternatively, we use a level shifting topology as shown in Fig. 4-10 in order to achieve proper domain crossing from low to high voltage. The input value is fed to inverters connected to the lower supply rail, with their output feeding the gate and drain of a thick oxide NMOS device. Connecting to the NMOS drain assists in faster voltage transitions of the converter [166]. A cross-coupled thick-oxide PMOS device pair configured as a latch with positive feedback assists in making the transitions occur faster. A final thick-oxide inverter at the output is used for buffering and making the voltage transitions sharper at the proper output domain high voltage. Across all points in its operation cycle, this topology ensures that core devices do not endure a voltage drop greater than the core supply voltage between any of their terminals and avoids shoot-through current.

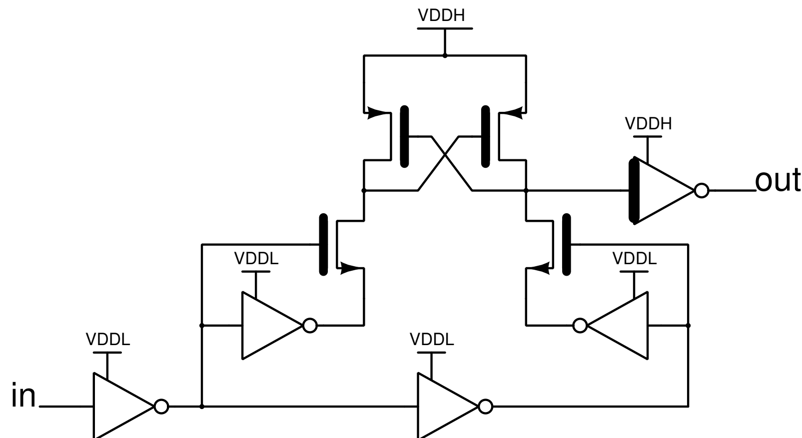


Figure 4-10: Low-VDD to high-VDD level shifter schematic

4.2.4 Clock Distribution

The clock for the chip operation is not generated on-die, but is fed from an external source. A differential clock signal is used to support high speed clock rates of up to 983 MHz. The clock signal is received on die via a differential common source amplifier with an active

load and converted to a single ended output. The amplifier is followed by a PMOS common source amplifier and finally fed to an inverter to adjust it to the core supply voltage rails. A schematic of the clock input amplifier is shown in Fig. 4-11. In order to ensure a 50% duty cycle, the output of the clock buffer is divided by 2 in order to create a cleaner clock output. The clock is then divided again by a factor of two to support dual chip operation as will be explained shortly. A slightly delayed version of the clock is also created in order to enable safe latching of the digital block output by the DAC input. The clock division block schematic is shown in Fig. 4-12.

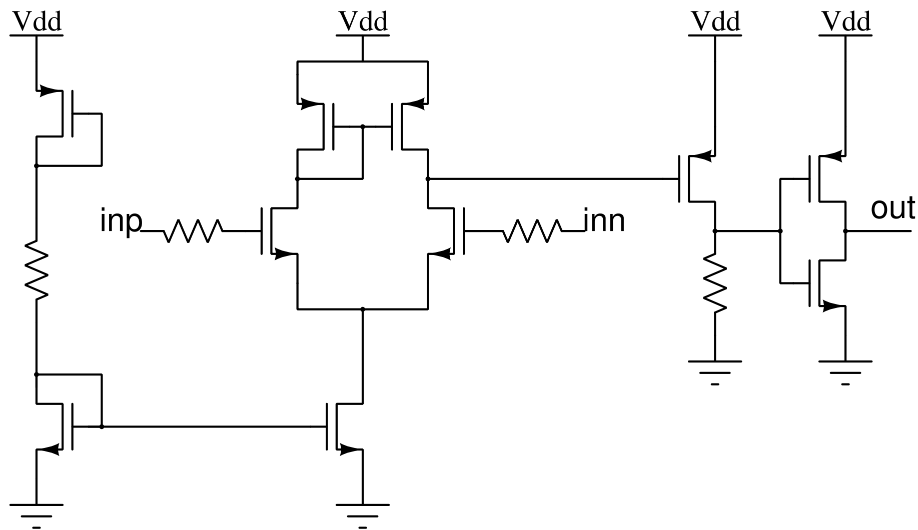


Figure 4-11: Clock input differential amplifier schematic

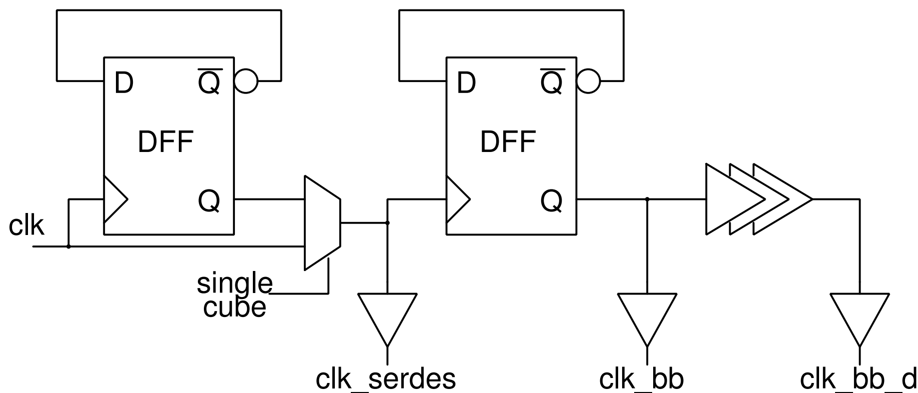
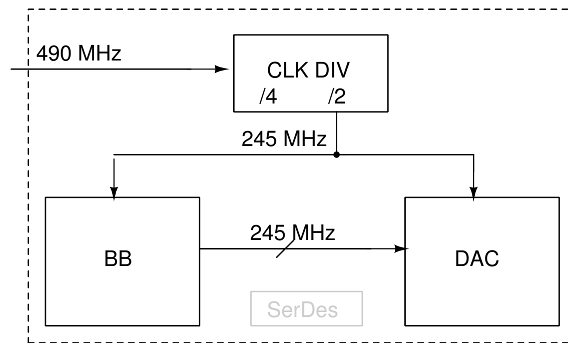


Figure 4-12: Clock divider schematic

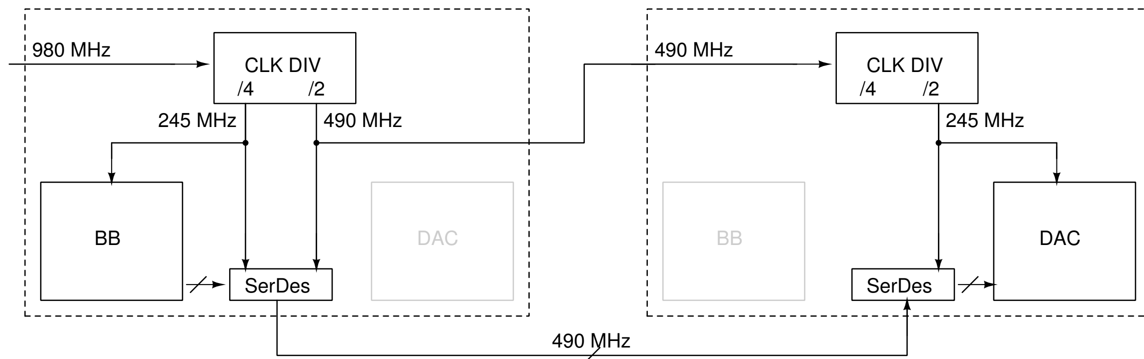
As described in section 4.2.2, the nominal clock frequency of the baseband output is 245.76 MHz (internally divided by 8). This represents the oversampled symbol rate of the digital block as well as the DAC sampling frequency. Therefore, an input frequency of 491.52 MHz is required prior to the division by two in order to support such single die operation as shown in Fig. 4-13a. When using chip-to-chip communication, the serial output multiplexes the real and imaginary parts of the symbol data in order to save on I/O pad count. This requires serial operation at twice the symbol rate, i.e. 491.52 MHz, which in turn requires the input clock to operate at a frequency of 983.04 MHz. In this scenario, the input clock will be divided by two and supplied to the Serializer/Deserializer (SerDes) block, following an additional division by two the clock will be supplied to the digital baseband. The clock divided by two which was used to serialize the data will also be used as an output clock to be sent along with the data and fed as the input clock to the second chip, where it will in turn be divided by two and used to deserialize the data and be used as the input clock for the DAC. The two chip digital communication operation mode is illustrated in the block diagram of Fig. 4-13b.

The clock frequencies discussed previously are the maximum used in this system for the case of a 20 MHz channel. If a smaller channel bandwidth is to be used, these frequencies are scaled accordingly. Thus, e.g. if the channel bandwidth is 5 MHz, then the baseband base symbol frequency will be 7.68 MHz, the x8 oversampled data rate will be 61.44 MHz. This in turn will require a 122.88 MHz frequency clock when operating in single chip mode and 245.76 MHz when operating in dual chip mode.

When operating in a 3D configuration mode, the clock must be distributed properly to all necessary die tiers. Since in our die stack we are re-using the same die layout rotated by 180°, we must be able to support such clock distribution with the same hardware blocks on both die tiers. The block schematic of the 3D clock distribution topology is shown in Fig. 4-14. Note that both die tiers basically contain the same circuits in a rotated form. The signals indicating which tier is the top tier and whether to use the vertical clock connection are supplied by the global chip scan chain (see appendix B). These signals are used to select the clock input origin (either from external pad or from microbump) and whether to transmit the clock between die tiers. Thus, in a 3D configuration, the clock is always



(a)



(b)

Figure 4-13: Clock distribution for (a) single and (b) two cube operation

generated at the top tier die, as a divided version of the external clock supplied through the pads, and transmitted via TSV to the bottom tier die.

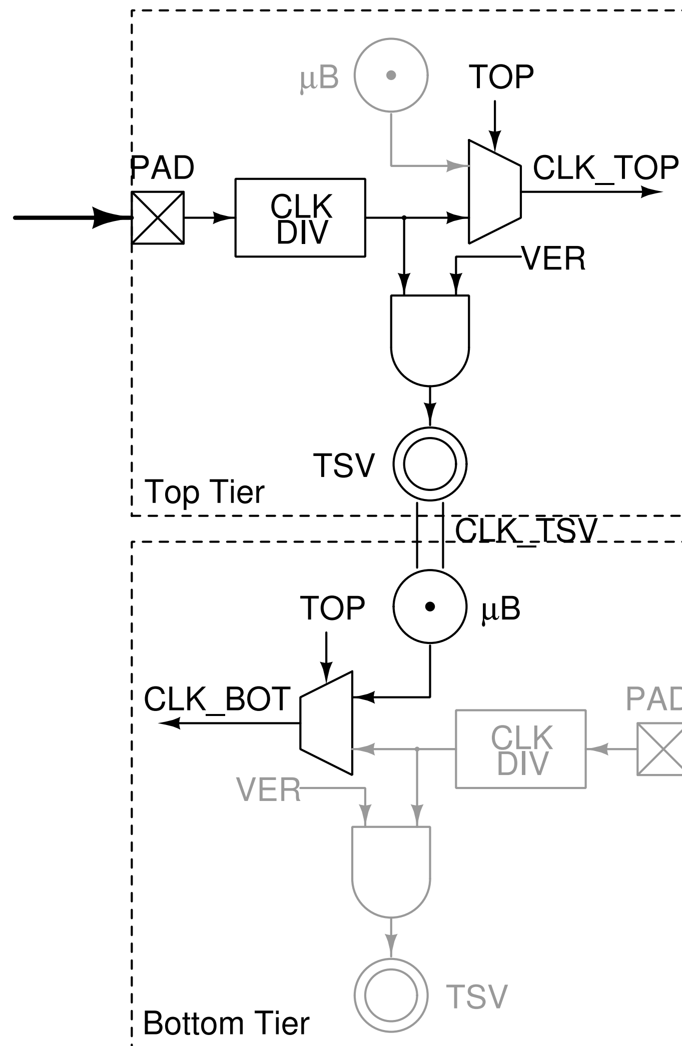


Figure 4-14: Clock generation and distribution block diagram for two tiers in 3D-IC

4.2.5 Digital Routing

The output of the digital baseband processor is composed of 11-bit complex symbols (total of 22 bits) which are to be ultimately routed to the DAC block in order to convert the digital symbol into its analog representation. The digital data is routed across the chip via the use of digital crossbar modules at several locations. These modules serve to route the data

according to the desired mode of operation enabling the versatility of testing desired for the chip.

The digital crossbars allow routing of the baseband data either to, or from external pads to allow multi-chip operation, routing to pass-through lines to other parts of the die as well as routing to and from micro-bumps and TSVs to support 2.5D and 3D operation modes. The data is routed through the use of controllable tri-state buffers as described in section 4.2.3.

Due to the 3D stack structure of the test chip, utilizing a 180° rotation of the die and stacking onto itself, we identify two types of digital crossbars - A “transmitting” crossbar, which only sends data to microbumps and TSVs as shown in Fig. 4-15a. And “receiving” crossbars, which only receive data from microbumps and TSVs as shown in Fig. 4-15b. There is no need for crossbars with bi-directional data transmission with microbumps and TSVs.

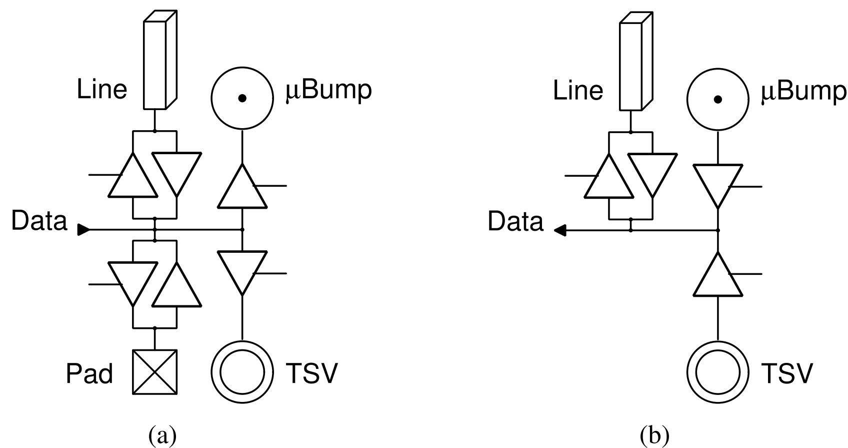


Figure 4-15: Digital crossbar (a) Tx column and (b) Rx column

Fig. 4-16 illustrates the overall routing and muxing of the digital data across the chip. The output of the digital baseband is routed to the “Tx” crossbar, where it can either be sent directly to the nearest DAC, routed across the chip, routed to external pads, sent via TSV to the bottom tier die (in case the baseband operating is on the top tier die) or send the data through the microbump (when the baseband is operating on the bottom tier die).

The second, "Rx" crossbar, is located on the other side of the die and receives the data sent across from the first crossbar, or receive it via microbump (when acting as bottom tier) or TSV (when acting as top tier). When multi-chip operation is used, the first die sends the digital data through the crossbar to the external pads, while the second chip receives the data through identical external pads and is able to route it as mentioned previously to different DAC instantiations across the die. The various routing options are controlled via configuration bits loaded by the global chip scan chain (see appendix B).

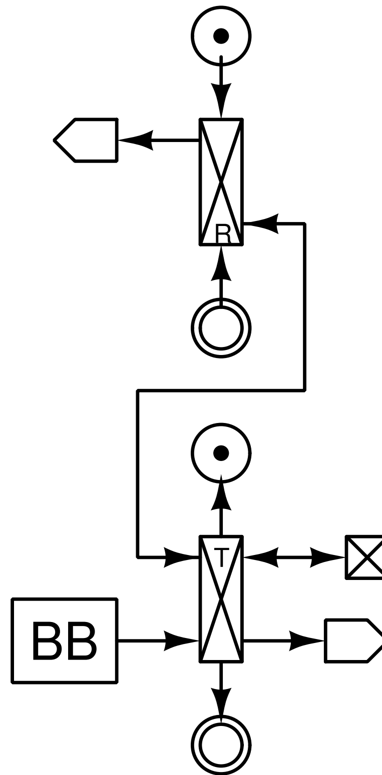


Figure 4-16: Digital signal multiplexing block diagram

4.2.6 Analog Routing

The analog crossbar modules serve a function similar to that of the digital crossbars described in section 4.2.5, but for the case of analog signal routing instead of digital signals. The modules are used to route the output of the DAC either off-chip through pads or to the mixer in order to upconvert the analog baseband signal to RF.

A schematic of the analog crossbars is shown in Fig. 4-17. Complementary NMOS and PMOS transistors are used as passgates to control routing of the analog signals either to pads, to through-pass metal routing, or vertically between die tiers via microbumps and TSVs.

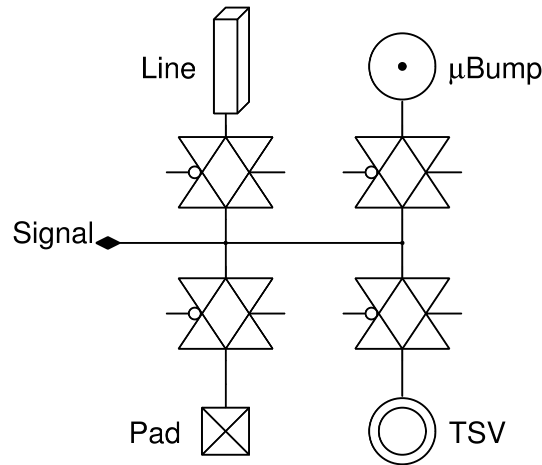


Figure 4-17: Analog crossbar schematic

The output of the DAC goes through an active LPF (see section 4.2.7) and has a common mode of about 0.6 V with an amplitude of roughly $1\,V_{p-p}$. This output is then fed into a secondary routing block which acts as a switch and as a voltage-to-current conversion stage before entering the mixer. A schematic of the switching block following the DAC is shown in Fig. 4-18 for one side of the differential output. This block enables connection between the analog crossbar, the DAC output and the mixer's IF input (see section 4.2.9) in order to allow various connection types. Enabling control signal *AG* or *BG* ties the internal or external IF signals respectively to the gate of an NMOS device to convert the voltage signal to current. The *AO* and *BO* control signals tie the internal and external IF signals directly to the output of the block and the input stage of the mixer. Alternatively, the *AB* signal shorts the external and internal paths in order to route the internal DAC output to the crossbar and from there it is routed across or off the chip as desired. Not all control signal combinations are permissible and need to be configured according to the desired routing topology.

The four control signals (*AG, AO, BG* and *BO*) are controlled via the chip scan chain

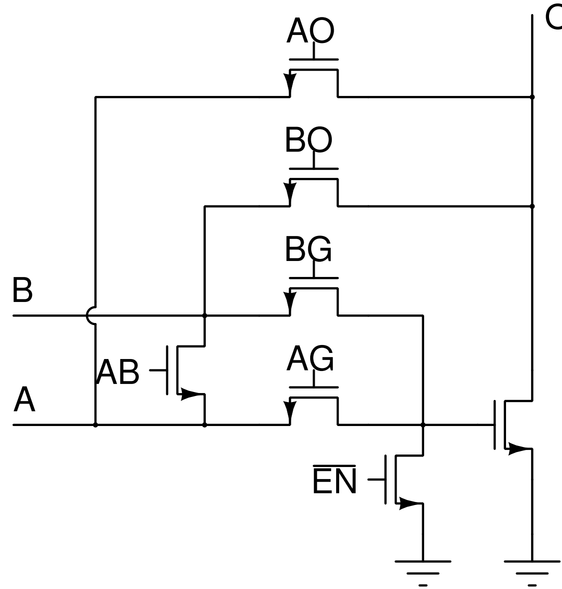


Figure 4-18: Voltage-to-Current conversion and signal multiplexing schematic

configuration (see appendix B), as detailed in table B.6. The control signal *AB*, used to bypass the current conversion block and route between the DAC output and the analog crossbar is not explicitly externally set, but instead derived from the other control signals via the relation $AB = \overline{AG + AO + BG + BO}$, i.e. it is active if and only if all other control signals are disabled. These scan chain configuration modules are allocated along with the various upconversion mixer instantiations across the chip.

A block diagram of the overall analog routing matrix is shown in Fig. 4-19. The use of the analog crossbars along with the switching blocks allows operating in all desired modes of operation (multi-chip, single die, 2.5D and 3D) with different partitioning scenarios between analog and digital data transfer. This allows passing along the analog DAC baseband output to another chip or to several mixers on the same die or between die tiers in the 2.5D interposer or 3D operation modes. The various routing options are controlled via configuration bits loaded by the global chip scan chain (see appendix B).

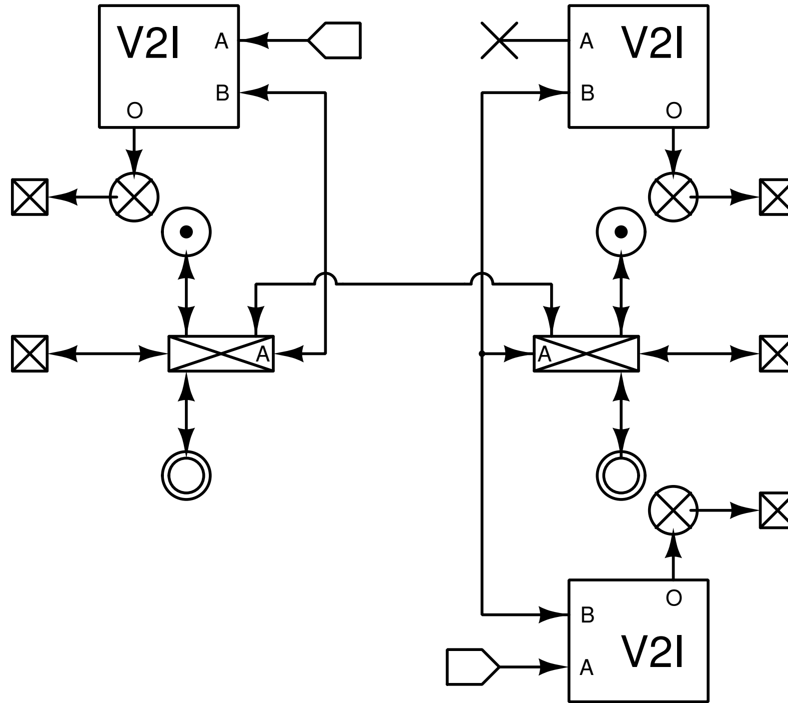


Figure 4-19: Analog signal path multiplexing block diagram

4.2.7 Digital to Analog Conversion

Conversion of the digital baseband digital codewords to an analog signal is performed by two, 11 bit, 3-8 segmented current steering DACs [167], one for the real part and one for the imaginary of the digital codeword symbol. The DAC module was designed by MediaTek and reused with their permission in this work. The DAC architecture is similar to the one reported in [168]. The maximum sampling rate of the DAC is 416 MS/s, which exceeds the maximum required in our design of 245.76 MS/s.

The basic principle of a current steering DAC is the complementary control of two branches of a current source by a digital signal, directing the current into one branch or the other. In a mixed M-L segmentation, an N bit DAC is implemented by using L binary weighted current sources for the LSB bits of the digital vector. An additional $2^M - 1$ unit current sources, each scaled by 2^L compared to the base current source are used to implement the remaining M bits, such that $N = M + L$. This topology is illustrated in Fig. 4-20. This technique maintains a reasonable scale ratio between the smallest and

largest current sources which helps with practical matching of the unit cells. As mentioned, the DAC design used in this work utilizes a 3-8 segmentation to implement the 11 bit resolution.

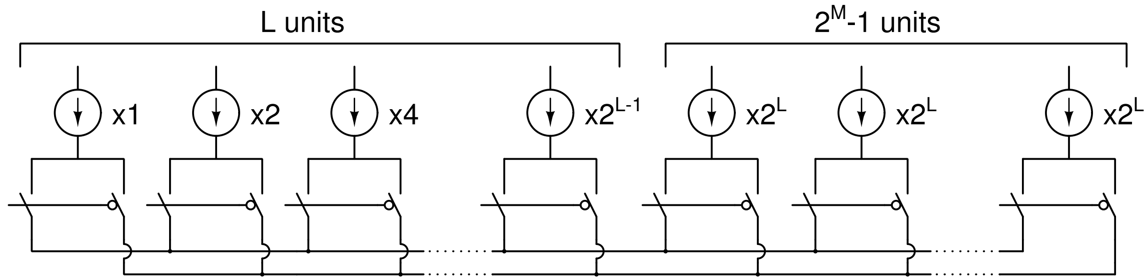


Figure 4-20: General schematic of binary-unary current steering DAC segmentation topology

The digital data arriving from the baseband module is latched and synchronized with the DAC clock by a series of flip-flops as shown in Fig. 4-21. This ensures that all bits of the digital codeword are re-aligned in order to toggle the DAC current sources simultaneously and avoid glitches at the DAC output.

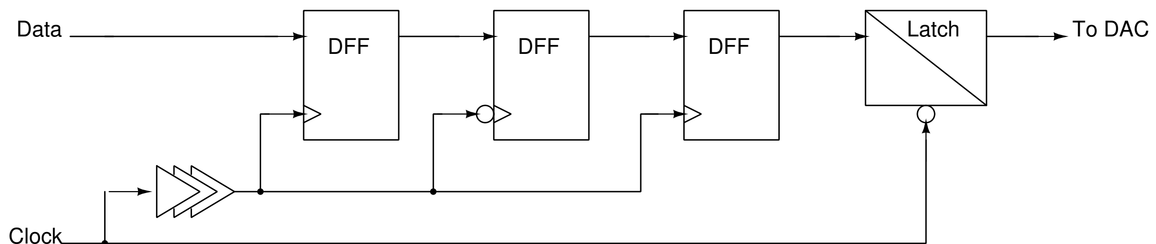


Figure 4-21: DAC data and clock synchronization schematic

The output of the DAC passes through a differential, first order, active RC filter as illustrated in Fig. 4-22. The filter consists of a 4 $k\Omega$ resistor and a switched capacitor bank which can be varied from 0.63 to 3.74 pF. This allows for the filter to have an adjustable bandwidth between 10 to about 60 MHz, to accommodate the required maximal bandwidth of LTE channels of 20 MHz. The DAC analog output has a voltage swing of 1.3 V_{p-p}, centered around a common mode voltage of 0.65 V

The DAC operates from the 1.8 V analog supply, and draws a current of 4 mA. It uses

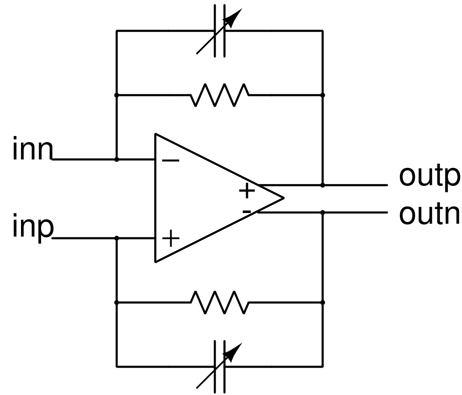


Figure 4-22: First-order, adjustable low-pass filter schematic

two LDOs to generate the required 1.5 V and 1 V needed for its operation. The LDOs are implemented via a regulated PMOS device acting as a power gate. This ensures a quiet, steady supply for the DAC operation, minimizing interference from the digital block supply.

The various control and configuration options for the DAC operation are controlled via the on-chip scan chain (see appendix B). Table B.4 summarizes the available configuration control signals and their functionality. These are repeated twice on each die tier, once for every DAC instantiation.

4.2.8 LO Generation

In order to transform the baseband modulated data to the higher RF band for transmission over the network, we require an accurate tunable LO. The generated LO signal will be used for mixing and upconverting the data. As such, it is required to be tunable and stable. Since the data modulated in our LTE system design is complex, quadrature upconversion is required, therefore we wish to have quadrature instances of the LO with 90° offset between the carriers.

The LO signals are generated via a quadrature LC-VCO. The VCO core design is similar to the one described previously in section 2.3.3, but extended to support quadrature signal generation as shown in Fig. 4-23. The two basic core cross-coupled VCO structures are connected amongst themselves to ensure the quadrature signal generation [169].

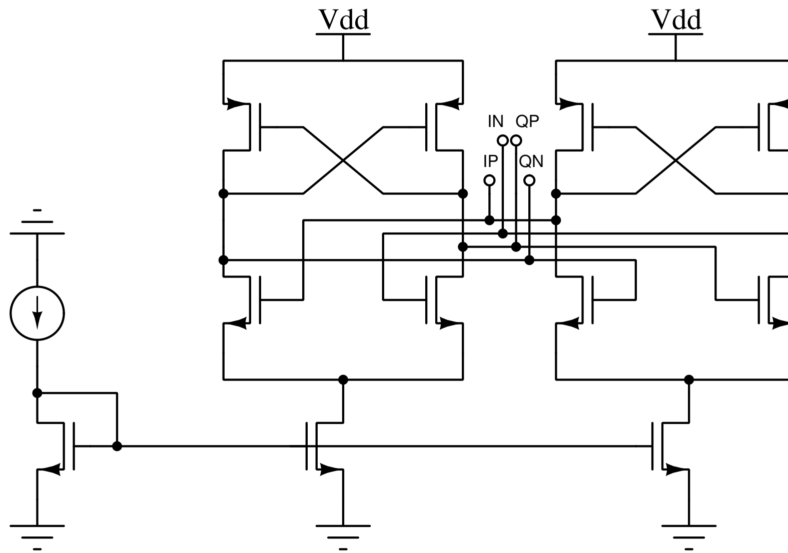


Figure 4-23: Quadrature VCO core schematic

The total fixed tank capacitance was reduced by half in order to shift the VCO center frequency to a slightly higher frequency range in order to include the 2.4 GHz Industrial, Scientific and Medical (ISM) band in the available tuning range. As before, the VCO core is tuned via both a 3 bit digital binary weighted switched capacitor bank to allow for coarse tuning, and a continuous fine-grain tuning via a varactor. These allow tuning the VCO output carrier frequency to the desired frequency for operation. Ideally the VCO should be placed as part of an overall PLL structure to ensure stability and locking of the carrier frequency to a set reference. However in this design only the VCO itself was implemented and the tuning voltage and loop feedback were controlled from off-chip. For a more complete, stand-alone system we would wish to integrate this functionality on chip as well.

Three instances of the quadrature LC-VCO were implemented across the chip. These different instantiations allow examining the effects of coupling between the digital block and various signal routing topologies to different spatial locations of the VCOs. Furthermore, as described in detail in section 2.3.2, both the planar inductor structure as well as the vertical solenoid inductor structures were implemented to examine the trade-offs and benefits of improved noise immunity of the solenoid structure to digital noise coupling as part of

the complete transmitter system. The vertical solenoid inductor VCO was instantiated such that it is located on the top tier die directly above the digital baseband module in the bottom tier die, potentially suffering from the most noise coupling from the digital clock signal.

The VCO output are two pairs of differential sinusoidal signal at the carrier frequency. In order to route the signals to the mixers, while also providing buffering, gain and drive strength (as well as enable driving the signals off chip for measurement purposes), a differential buffer was used as shown in Fig. 4-24. The buffer inputs are AC coupled to the VCO output via a blocking capacitor. The signal's common mode is set by a self-biased inverter, which sets it around the device's tripping point. The signals are then buffered via standard CMOS inverters. The 180° phase shift between the differential branches is maintained by using back-to-back inverter pairs. This balancing, along with AC decoupling between repeated such buffer stages helps to mitigate issues of variation between the buffers setting the output common mode value.

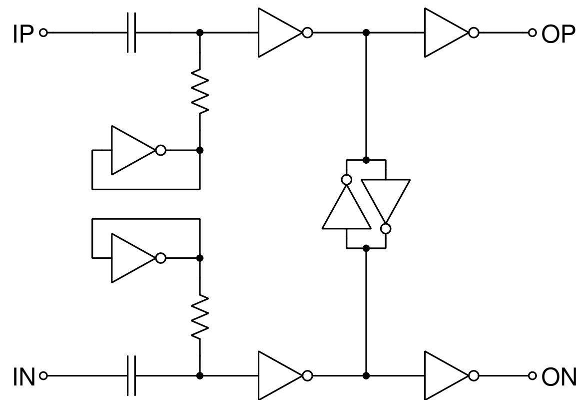


Figure 4-24: VCO differential buffer driver schematic

The few digital control signals required to configure the VCO operation are loaded via the scan chain (see appendix B). Table B.5 summarizes the configuration bits used in the VCO module instances across the chip.

4.2.9 Upconversion

The upconversion of the modulated data from baseband to RF is achieved via a double balanced Gilbert-cell mixer. The differential IF signal is mixed along with differential

quadrature versions of the LO carrier and then added together. The use of the double balanced mixer enables higher immunity to LO noise [169]. This upconversion is repeated for both real and imaginary parts of the IF baseband, each time with a different LO quadrature component, and combined at the output to generate the final RF signal output as shown in Fig. 4-25.

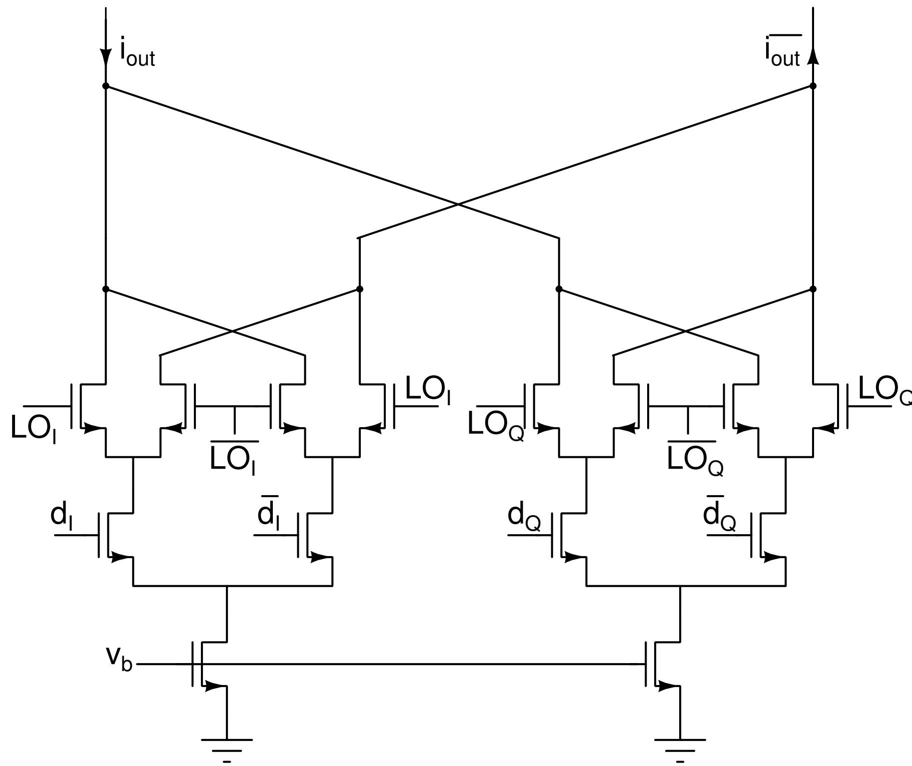


Figure 4-25: Simplified RF mixing DAC schematic

The operation of the upconversion mixer is closely tied to that of the DAC, forming what is known as a mixing-DAC (or RF-DAC) [170, 171]. Apart from design decisions whether to preform the mixing locally or globally [172], in the 3D environment available in our design we consider further partitioning possibilities. Such a specific partitioning possibility is illustrated in Fig. 4-26 where the digital codeword is first converted to analog via the current steering DAC, then summed and carried through to another die tier via TSVs for further mixing and upconversion. A different partitioning would be to transmit each digital codeword bit, or either each corresponding analog current and summing them on the next tier. The first approach has the benefit of requiring less TSVs to carry the data

due to the analog aggregation inherent in it as well as potentially reduced bandwidth of the overall signal compared to its individual comprising components.

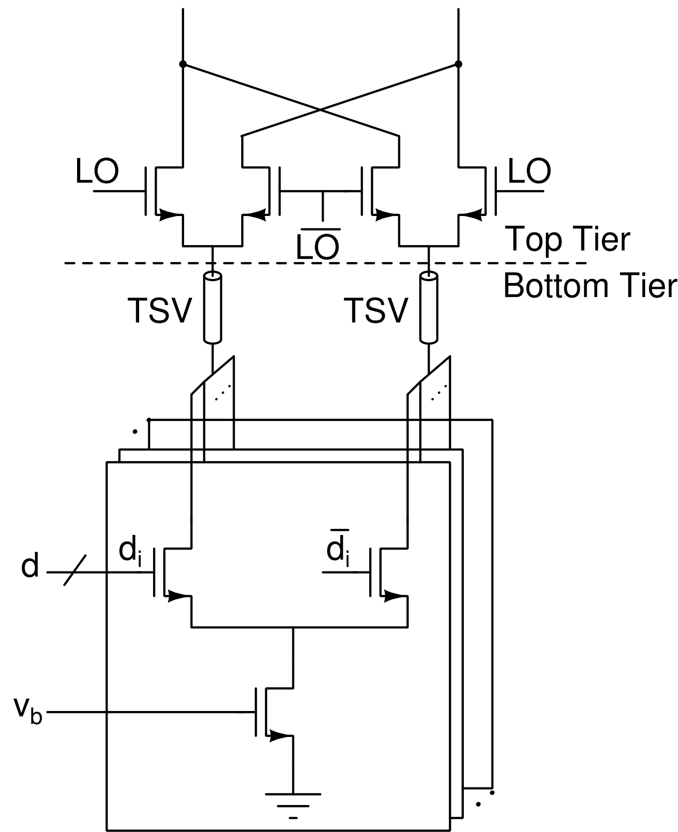


Figure 4-26: Partitioning of mixing DAC operation in 3D-IC

These various partitioning options were implemented in the transmitter design as previously described in section 4.2.6 by using the voltage-to-current passgate structures to control routing of the DAC IF output before entering the mixers. This allowed for better comparison of the different possibilities as described above.

4.2.10 Bias

The various bias conditions required for the different modules are mostly supplied by a central bias module. This module acts as a global current mirror to reproduce the required bias currents needed. The use of bias currents in contrast to bias voltages is preferred since they can be routed across the chip to the different modules while being less affected by IR

voltage drops across the metal routing. When using a bias voltage as a reference, there is some small amount of current leakage on the reference line. Coupled by the potentially large resistance of long metal wirings this leads to voltage drops and change in bias points in voltage bias routings. Current bias on the other hand is less susceptible to such effects of line resistance [173] and therefore more suitable to such applications where the bias needs to be routed from a central location to the other units.

The bias module receives an external reference voltage of 1.2 V which it uses as the basis for the PMOS side current mirror topology to generate the various bias currents. Each bias current acts as an independent current source. Many of the bias currents can be scaled by using digital configuration bits stored in the scan chain (see appendix B). The current scaling is done by aggregating binary or unary weighted current mirror cells for each bias current. A summary of the available configuration options for the bias module is shown in table B.7.

4.2.11 External I/O

The external interface to the test chip is conducted via several different I/O pad structures which differ by the signal types, voltages and data rate. Only the top tier die pads are exposed and are accessible. The pads have an equivalent capacitance to ground of roughly 100 fF. The pads are wire bonded to an 88 pin QFN package, which in turn is placed on the test PCB either via a socket or by direct soldering. The wirebonds present an added parasitic inductance of 2-3 nH. The QFN package's short leads and central ground plane allows for reduced pin parasitic inductance and better electric and thermal performance [174].

One side of the die pads is dedicated to low-speed digital data input, mostly dominated by the scan chain control and data signals (see appendix B). These signals are digital signals operating at a few MHz frequency or below at a supply voltage of 1.8 V. These signals are mainly used for the scan chain clock, control and data signals, used to load the necessary configuration bits to the various blocks throughout the chip as well as read out data stored in the baseband snapshot memory. Internally generated baseband signals such as internal control signals and flags are also accessible through these General Purpose I/O (GPIO)

pads.

The differential, upconverted RF output of the mixers is connected to the pads directly along with Electrostatic Discharge (ESD) protection. This configuration is also used for the analog signals of the DAC output, both real and imaginary parts. The DAC output pads are also used as an input if the on-die DAC is disabled. This is used when operating in dual-chip mode and transmitting the analog baseband signals between chips. Alternatively, the pads are used as an input when manually generating the baseband signals as an input to the mixer.

The in-phase and quadrature differential VCO output is tied to an open-drain NMOS buffer. The open drain output can be disabled, leaving the pad at high impedance and reverting its function to operate as an input node instead of an output. In this manner we override the internal LO with an externally generated version.

The digital symbol data which needs to be either transmitted or received in dual-chip mode operation (or for testing purposes) must allow for high speed data transfer rates of up to 5.4 Gb/s. This is due to the fact that the maximum baseband symbol rate is 245.76 MS/s after a factor of 8 oversampling. Each symbol consists of 22 bits, 11 bit words for the real and complex parts. To support high speed communication we wish to use differential signaling which will increase the required pad count. In order to keep the number of pads at a reasonable amount for packaging, and balance the required effective data rate we chose to implement interleaving in the digital data output.

11 differential pad pairs are designated for the digital data input and output. The real and complex words are interleaved during transmission and reception and Double Data Rate (DDR) is used. Therefore the maximum data rate for each differential pair is 491.52 MS/s. The SerDes block used in the design is illustrated in Fig. 4-27. The architecture acts as a sort of DDR system, sampling on both the rising and falling edges of the clock signal and later aligning the output at double the clock frequency. Deserializing of the data is performed in the reverse order sampling again on both clock edges and re-aligning the data to the rising clock edge. Both serializing and deserializing blocks can be tri-stated and are placed in parallel to allow communication in either direction.

Each digital differential output pair uses a single ended-to-differential output amplifier

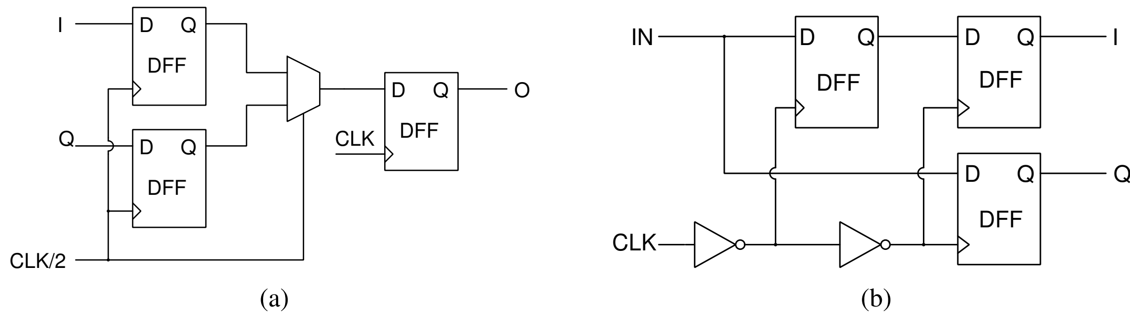


Figure 4-27: High-speed I/O (a) serializer and (b) deserializer schematics

to convert the single ended data bits to a differential output signal around the supply voltage. On the receiver side, a differential amplifier translates the signal back to full rail single ended form. The polarity of the data may also be inverted in order to accommodate scenarios where the interconnect between the sender and receiver chips are reversed.

The high speed clock has designated differential input and output pads to support high speed operation. The clock input pads are general, analog pads, which connect to the clock input differential amplifier as described in section 4.2.4. The clock output is driven by a differential amplifier identical to the one used for the high speed digital data transmission described above.

Bias voltages and power are supplied, and measured, by standard analog pads. These are used to supply the tuning voltage to the VCOs (shared by all VCO instances) as well as to supply the 1.2 V reference voltage to the bias circuitry. The digital pads' ESD protection circuitry is connected to the digital I/O supply voltage of 1.8 V, and the analog pads are connected to an analog I/O supply voltage of 1.8 V as well. The analog supply is also used as the main power source for the DAC and bias circuits to generate the various internal analog supplies and reference currents. Each VCO instance has a separate designated analog supply with a nominal value of 1 V, which can be disabled to shut down a specific instance as well as measure the consumed power of each VCO. The various digital buffers and scan chain are connected to the digital core voltage supply of 1 V. The core baseband circuits on the top tier die are separated into 3 different supplies - one for the FFT block, one for the IFFT block and one for the rest of the baseband digital circuits. All three supplies are nominally tied to the same core voltage of 1 V, however their separation allows for

easier power consumption analysis. These three domains are shorted and supplied together in the bottom tier die and are connected via a dedicated supply pin.

Table B.8 summarizes the various I/O pads which are used in the test chip. In addition to the previously described pads, several ground pads are used across the die. All of the ground pads are downbonded to the QFN package's central thermal pad in order to reduce the wirebond length and obtain a short, low-impedance path to ground.

4.3 Measurements

The 3D LTE transmitter design was implemented in a 28 nm LP CMOS process and then the die was attached to a 180° rotated version of itself via TSVs in a B2F form through micro-bumps on the top metal redistribution layer. The die stack was wirebonded and packaged in an 88 lead QFN package and mounted on a PCB for testing. The total die area is 2.8×2.9 mm^2, the TSV diameter is 10 μm with a minimum pitch of 40 μm. The total die stack height is roughly 200 μm. An annotated micrograph of the die (acting as both top and bottom tiers) is shown in Fig. 4-28. Communication with the test chip was performed via an FPGA which loaded data and configuration to the chip as well as read data from snapshot memory. The output of the DACs, VCOs and mixers was connected to a real time oscilloscope and spectrum analyzer for probing. A second test chip was also mounted on the same PCB to allow for testing of chip-to-chip communication and data transfer in order to mimic separate die configuration operation. The different power supplies were either generated by on board LDOs or supplied externally from a source meter which allowed for power consumption measurements.

The following sections detail the measurement results for the key individual circuit blocks in the transmitter chain, as well as a review of the overall system performance. A comparison follows of different configuration scenarios enabling emulation of SoC, separate dies, 2.5D and 3D operation.

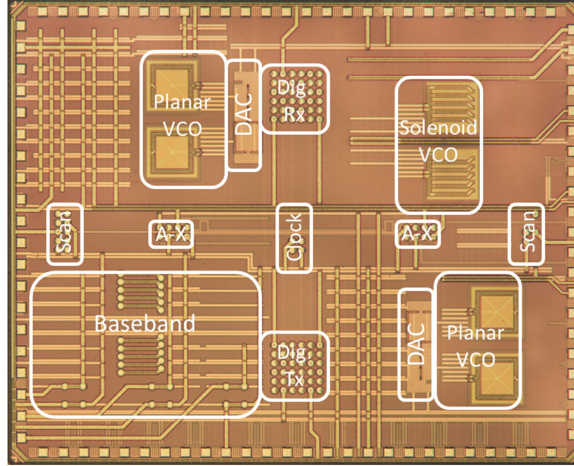


Figure 4-28: Micrograph of test chip (same layout used for both top and bottom tiers)

4.3.1 Digital Baseband

The digital baseband operation was thoroughly tested and the measurements results are reported in detail in section 3.5. The DFT calculations consume between 0.08 to 2.93 mW from a 0.61 to 1 V supply at an operating frequency between 1.92 to 30.72 MHz for the respective LTE bandwidth from 1.4 to 20 MHz. The energy per sample ranges from 40 to 95 pJ for the worst case scenario. The use of TD enables reducing the consumed power by up to 24% over the various cases of RB utilization. The interpolating filter consumes between 0.02 to 0.92 mW and enables extended shaping of the signal output spectrum to reduce the effects of the ZOH sampling.

4.3.2 DAC

The DAC output voltage was measured for each codeword of the 2048, 11 bit possibilities and plotted in Fig. 4-29. The Differential Non-Linearity (DNL) and Integral Non-Linearity (INL) measurements were calculated as follows

$$\mathrm{DNL}(i) = \frac{V_{\mathrm{out}}(i) - V_{\mathrm{out}}(i-1)}{\mathrm{LSB}} - 1 \tag{4.1}$$

$$\mathrm{INL}(i) = \frac{V_{\mathrm{out}}(i) - V_{\mathrm{out}}(0)}{\mathrm{LSB}} - i \tag{4.2}$$

Table 4.2: DAC specification summary

Bits	11
Sampling rate (MS/s)	245.76
Segmentation (Unary-Binary)	3-8
DNL (LSB)	± 0.8
INL (LSB)	± 2
Power (mW)	1.35

$$\text{where LSB} = \frac{V_{\text{out}}(2047) - V_{\text{out}}(0)}{2047}$$

and plotted in Fig. 4-30a and 4-30b respectively. The DAC performs as expected with a DNL measure of ± 0.8 LSB and an INL of ± 2 LSB.

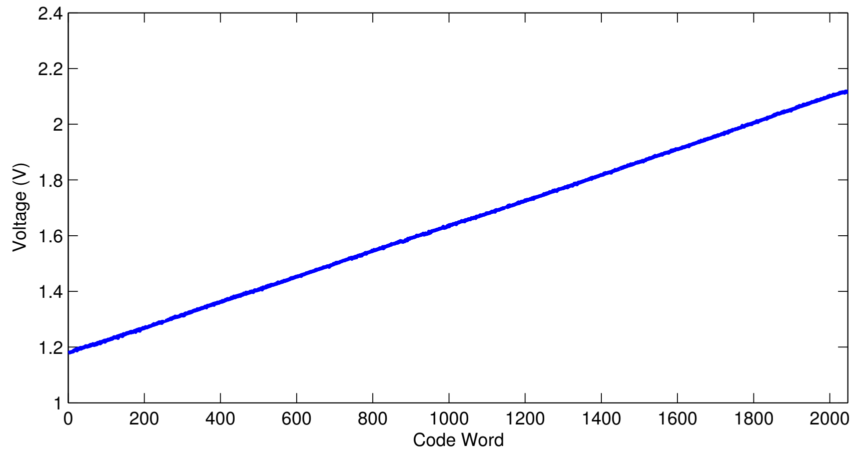
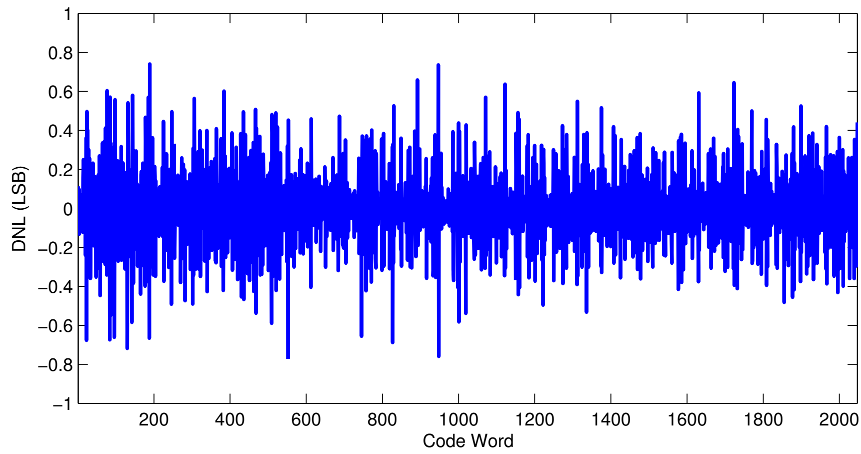


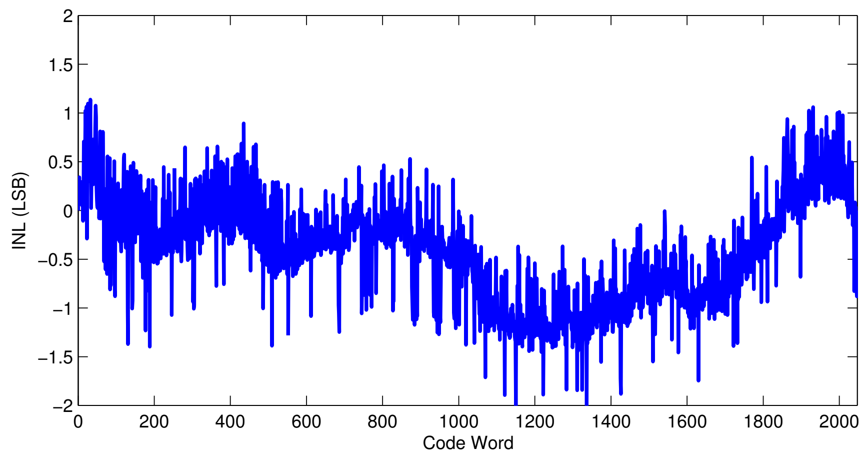
Figure 4-29: DAC voltage output for 11 bit codeword input

Each DAC of the two used for the in-phase and quadrature signals consumes 0.75 mA from a 1.8 V supply. Each LPF consumes an additional 1.25 mA from the same supply, since it was designed to drive the DAC output off chip. Therefore, the overall power consumption of the quadrature DAC block is 7.2 mW. A summary of the DAC specifications is listed in table 4.2.

The output of the two quadrature DACs was recorded for a test signal of one slot (7



(a)



(b)

Figure 4-30: Calculated DAC (a) DNL and (b) INL

OFDM symbols) of a 5 MHz bandwidth LTE signal. The signal was captured once without interpolation, consisting of a total of 3,840 complex symbols at 7.68 MS/s, and once with oversampling and interpolation by a factor of 8 resulting in 30,720 complex samples at 61.44 MS/s. The frequency scaling ensures the same slot duration of 0.5 ms. The power spectral density of the captured DAC output was calculated and is shown in Fig. 4-31. As can be seen, the signal in both cases indeed occupies the main active bandwidth of 25 RBs, or equivalently 4.5 MHz out of the 5 MHz channel. However, we observe that without interpolation, there are fairly strong replicas of the main signal at a sampling frequency of 7.68 MHz, shaped by the *sinc* windowing caused by the ZOH operation of the DAC sampling. The interpolated signal on the other hand is able to shape the spectrum over a much wider frequency range as expected.

It is also important to note that use of the interpolation filter is much more efficient than using a more aggressive LPF. As discussed in section 3.5 the interpolation filter consumes up to 0.92 mW (and only 0.13 mW for the 5 MHz BW case) which is much lower than the power consumption of the LPF and has a much greater suppression of the images than the roll off of the 1st order RC filter.

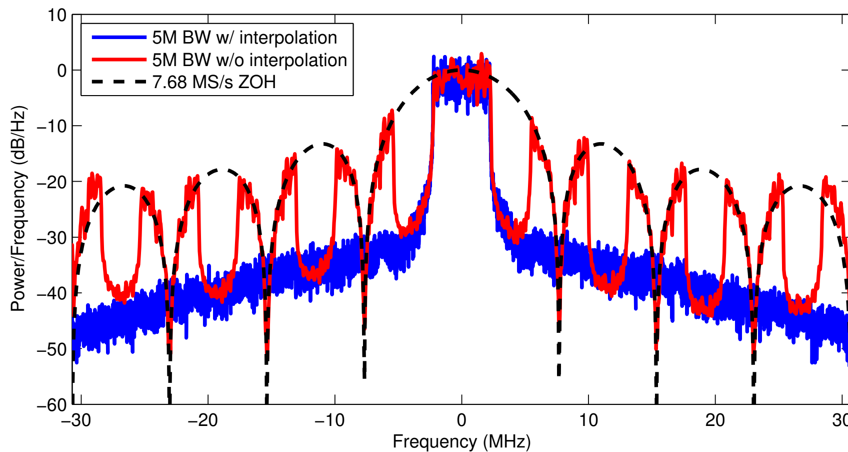


Figure 4-31: DAC output spectrum for 5 MHz BW LTE signal

4.3.3 VCO

Using the 3 bit binary weighted capacitor bank and tunable varactor, the VCO center frequency was varied. Both the planar and solenoid VCO tuning range was measured and is plotted in Fig. 4-32. The planar VCO covers a frequency range between 2.33 and 2.97 GHz, while the solenoid structure VCO is tunable from 2.06 to 2.62 GHz, giving them both a 24% tuning range and indicating that the solenoid inductor has a 28% larger inductance value than the planar inductor occupying the same die area. These results are consistent with the ones measured from the passive structure test chip described in section 2.4, apart from the higher center frequency in this case due to the reduction in the overall tank capacitance in order to cover the 2.4 GHz ISM band as well as LTE bands 7 and 30.

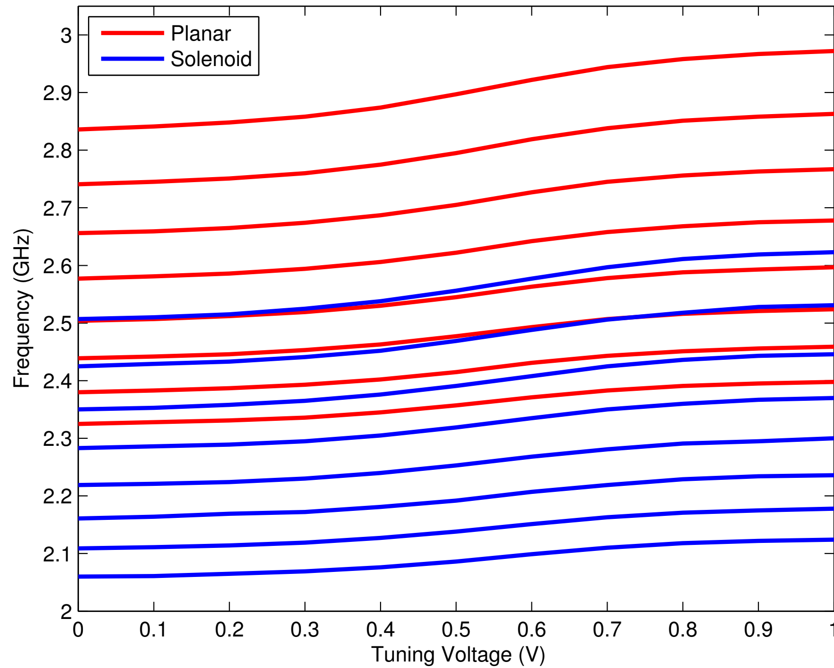


Figure 4-32: VCO frequency tuning range

The measured phase noise of the VCOs is plotted in Fig. 4-33 for a center frequency of 2.45 GHz for both planar and solenoid inductor cases. The phase noise is -113 and -119 dBc/Hz at 1 MHz offset from the carrier for the planar and solenoid inductor VCOs respectively. Each VCO core consumes 5 mA from a 1 V supply for an overall 10 mW power consumption. These results are comparable to the ones obtained in the passive

structure test chip accounting for the slight change in center frequency. The FOM is therefore calculated to be the same - 175 and 181 dBc/Hz for the planar and solenoid VCOs respectively.

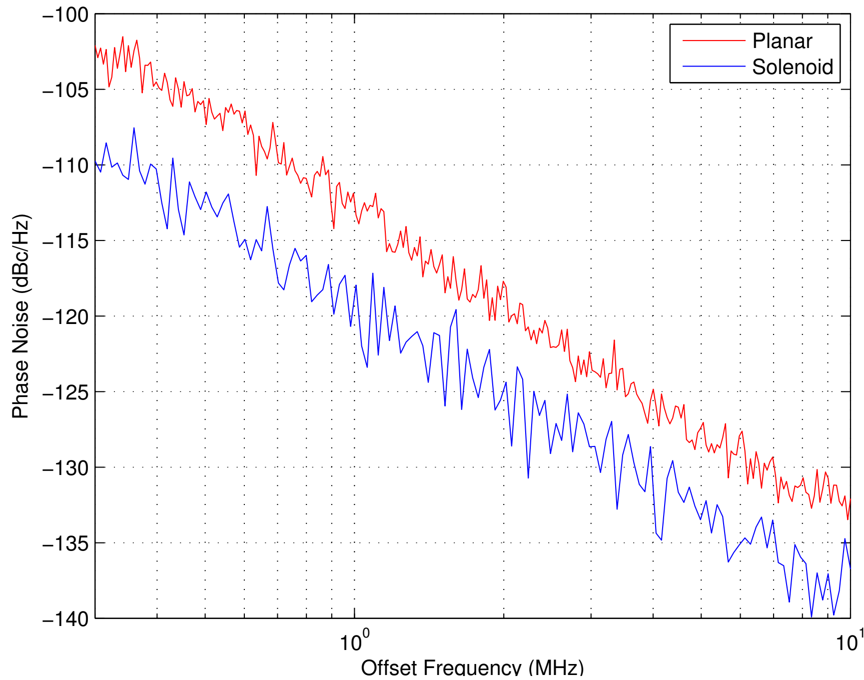


Figure 4-33: Measured VCO phase noise for Planar and Solenoid inductor structures

The output of the two phases of the main solenoid VCO were recorded via a high speed sampling oscilloscope and are plotted in Fig. 4-34. The 90° phase difference between the two outputs is clearly observed. The phase difference between the branches was calculated by estimating the mean crossing of each waveform for multiple time domain measurements and averaging. A plot of the calculated phase difference for the planar and solenoid VCOs over their respective tuning range is shown in Fig. 4-35. A good balance of ±2° between the two outputs was measured. A summary of the performance for the two VCO structures is presented in table 4.3.

4.3.4 Mixer

Measurements from the output of the mixer resulted in a very low signal output which could not be measured reliably. Re-examining the design it appears that both the switching

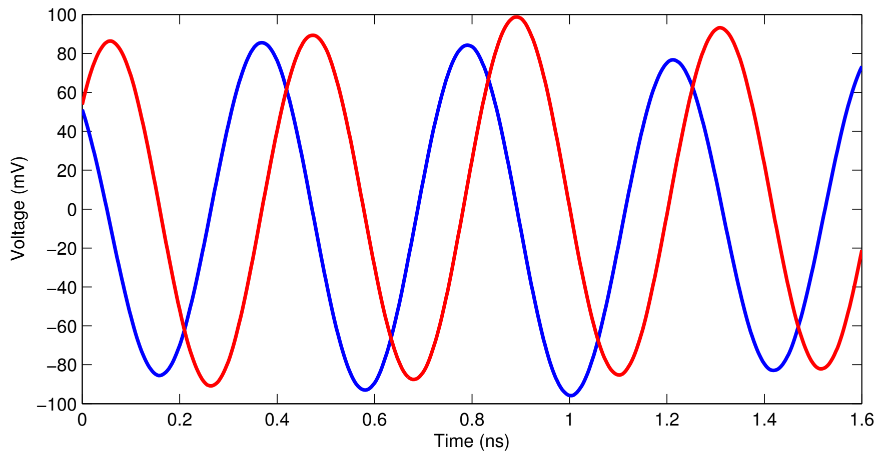


Figure 4-34: Example of time domain output of solenoid quadrature VCO

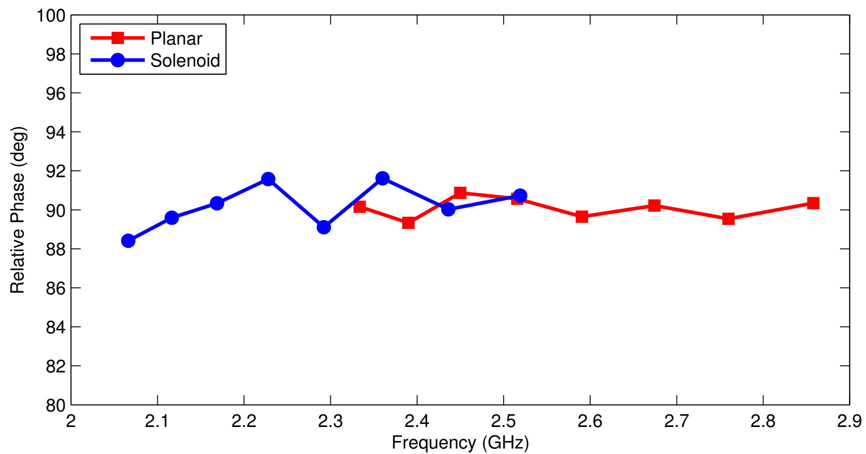


Figure 4-35: Phase difference between VCO quadrature outputs

Table 4.3: Summary of VCO performance

	Planar	**Solenoid**
Frequency Range (GHz)	$2.33 \sim 2.97$	$2.06 \sim 2.62$
Tuning Range	24%	24%
Quadrature Phase (deg)	90 ± 1	90 ± 2
Power (mW)	5	5
Phase Noise @ 1 MHz (dBc/Hz)	-113	-119
FoM[1]	175	181

[1] $FoM = 20 \cdot \log\left(\frac{\omega_0}{\Delta\omega}\right) - 10 \cdot \log\left(\frac{P}{1\,\text{mW}}\right) - PN\,(\Delta\omega)$

devices and output buffers were not sized sufficiently for measurements with 50 Ohm termination equipment. This is mainly due to the very limited time available during the design phase which did not allow for sufficient testing and validation of the design across a wide testing scenario. Although unfortunate, this does not in itself impact the rest of the design described and we can still analyze the various blocks and even most of the effects of the 3D design on the system performance.

In order to illustrate the system functionality an external mixer was used to upconvert the baseband DAC outputs along with the VCO quadrature signal. This enables us to observe the desired mixing behavior which was intended to be performed on-chip. The output power spectrum of a 5 MHz bandwidth LTE signal oversampled by a factor of 8 and upconverted to 2.45 GHz is plotted in Fig. 4-36. The output spectrum appears as expected, occupying 4.5 MHz (equivalent to 25 RBs) with a sharp roll-off characteristic of OFDM modulated signals. The x8 interpolation helps shape the spectrum further out beyond the original sample rate of 7.68 MS/s.

4.3.5 System Partitioning

The power consumption of the digital power supply was measured for different scenarios of operation modeled to mimic two separate dies communicating, single die SoC, 2.5D

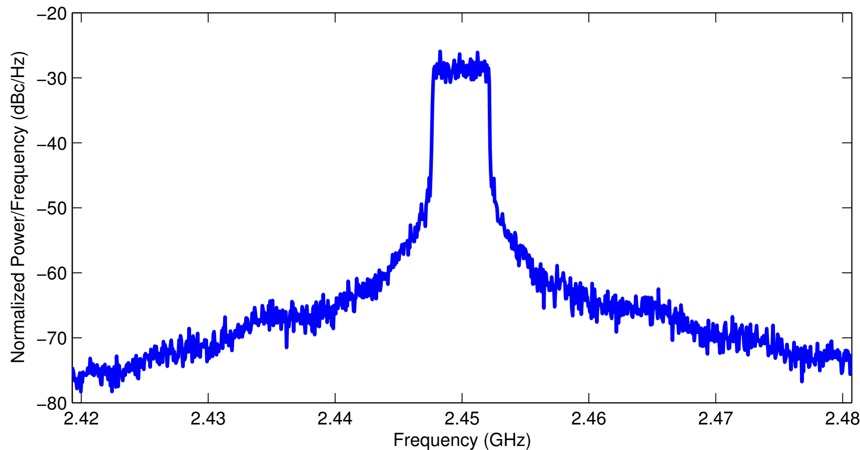


Figure 4-36: External mixer output spectrum of 5 MHz BW LTE signal

interposer scenario using the bottom tier die as a "dummy" interposer substrate for signal relay and a full, true 3D solution as detailed in table 4.1. For the following tests, the most extreme case was considered, transmitting at the largest LTE bandwidth of 20 MHz, at an x8 oversampling rate using 22 bits per complex sample yielding an aggregated required data rate of 5.4 Gb/s.

It is no surprise that when operating with two separate dies, one for the digital baseband and the other for conversion to RF (partition #2), the power consumption required to transmit the data from one die to another through the lossy PCB is very high. The overall power required from both transmitting and receiving ends was measured to be 180 mW. If the conversion to analog is done on the baseband chip (partition #3) then we only need to drive the analog signals. This will save considerable power and will require a much more modest increase in the DAC consumption by implementing a strong driver as detailed in section 4.2.7, increasing the DAC power consumption from 2.7 to 7.2 mW. This scheme however is much more susceptible to noise and coupling to the signal while it is being transmitted through the board, since the analog waveform will have a much lower noise tolerance than the digital bits.

The emulation of the 2.5D solution was achieved by using the TSVs to transmit the data from the top tier die, through the bottom tier and then back up again to the top tier (partition #4). Thus, using the bottom tier die as a passive silicon layer for interconnect

only. This is not a perfect representation of the situation since ultimately both ends of the signal route are on the same die tier and share a substrate and the distances are smaller than typically encountered in 2.5D interposer designs. The total digital power measured when transmitting the digital bits via this topology was 4.5 mW.

Enabling all circuit blocks to operate on the same die tier acted as an emulation of an integrated SoC solution (partition #1). The power consumption depended slightly on the relative location of the active circuit blocks which led to different routing length, however on average the total power was roughly 2.5 mW for a typical block configuration.

Finally, using both die tiers to comprise the entire chain, and also minimizing the routing length by activating blocks vertically adjacent (partition #6), the power consumption was minimized. The overall digital power measured was 2 mW for the maximum data rate of 5.4 Gb/s at a clock rate of 245.76 MHz. These power measurement results are summarized and plotted in Fig. 4-37. As can be seen there is a large reduction in power going from a separate die solution to a 2.5D interposer solution, and an additional reduction when implementing the 3D topology partitioning.

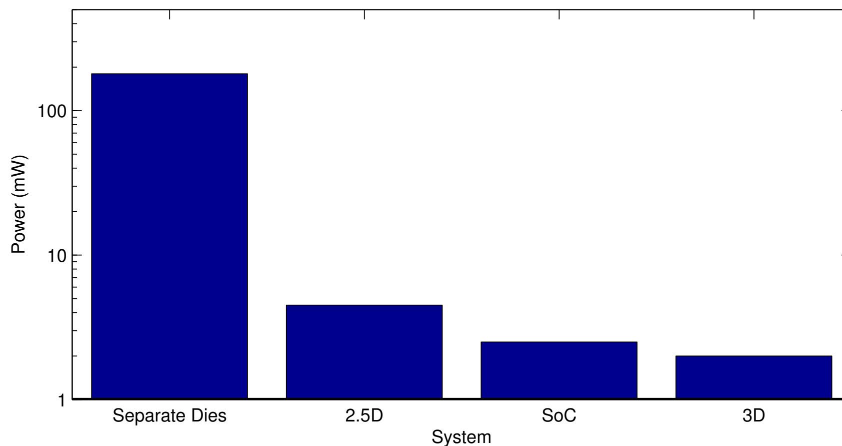


Figure 4-37: Digital link power consumption comparison

It should be noted that the serial link implemented for the separate die solution is far from optimized and is not the best solution achievable. However, even if we consider state-of-the-art digital transceiver implementations, they exhibit an energy efficiency of roughly 20 pJ/bit for typical scenarios as the one under test [175–177]. This energy

efficiency would translate in our case to a total power consumption of about 108 mW, still much higher than the power consumption of the 3D solution. For interposer and 3D links, prior publications indicate a similar energy efficiency to the one measured here, between 0.3 to 1 pJ/bit [71, 178–180]. In these cases the more dominant factor in the actual total power consumption is the total routing length which varies considerably between designs and is typically much shorter in the vertically stacked 3D case.

These differences will be even more pronounced as we move towards more advanced communication schemes with higher data rates. LTE-A for example allows for the aggregation of up to 5 separate 20 MHz channels, effectively increasing the data rate by 5 times. This will both increase the overall power consumption, but will also likely degrade the effective energy per bit due to the higher losses at higher Nyquist frequencies.

Whereas in the separate die scenario it is clear that first converting the digital signal to the analog baseband will result in a significant power reduction, this is not so straight forward in the other cases. Since the power consumption in the SoC, 2.5D and 3D cases is much lower and also depends highly on the actual routing length in the design, other considerations may prove to be more dominant in the choice of system partitioning. Namely, the use of TSVs for digital transmission will incur a significant overhead in terms of die area due to the large amount of digital signals compared to the analog waveforms. Furthermore, issues of noise coupling through the substrate or via capacitive and inductive electromagnetic coupling may interfere with the operation of the analog blocks and dictate a preferred partitioning scheme.

We have previously demonstrated the superior performance of the solenoid inductor structure in the 3D stacked topology to minimize coupling from nearby signal and clock lines (see Section 2.3.5). Adding this attribute to the inherent substrate separation which exists in the two tier stack, we observe much better immunity to noise in the analog blocks. This relative immunity is illustrated in Fig. 4-38. When operating as a single die tier emulating an SoC (partition #1), the digital baseband clock operation creates large noise on the power supplies that also couples through the substrate and appears as visible spurs at the VCO output at -45 dB below the VCO main carrier signal. It is important to note that since we have placed the VCO far enough from the digital block along with proper

guard ring structures, this noise does not couple and mix with the VCO but simply appears as additive noise at its output. This can be observed by noting that the noise spurs are at multiple integers of the baseband digital circuits of 30.72 MHz, and are not up-converted around the carrier. When taking advantage of the 3D topology we obtain a much better separation between the digital tier and its noisy clock and the VCO tier, this separation results in a great reduction of the spurious tones to almost the noise floor level measured at -90 dBc.

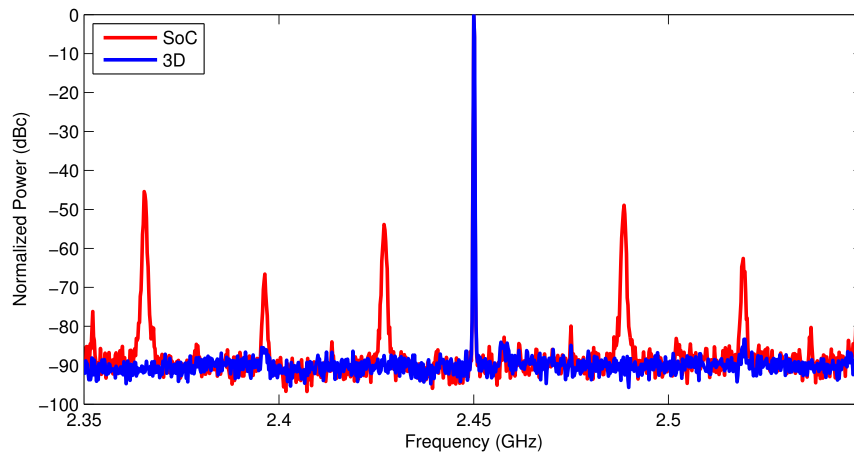


Figure 4-38: VCO output spectrum using different partitioning schemes

4.4 Conclusions

In this chapter we described the key building blocks required for an integrated RF transmitter incorporating some of the digital baseband logic. We have focused specifically on an application of a mobile unit LTE UL channel. We then reviewed how by utilizing the available process and 3D stacking technology available we can configure and emulate various scenarios for system partitioning including single chip SoC, separate dies, interposer based 2.5D integration as well as 3D vertical integration. The potential benefits and challenges of each approach were discussed.

The transmitter chain was fabricated in a 28 nm CMOS process and later vertically stacked and connected via TSVs. Measurements of each individual block were performed.

The baseband used efficient, signal-dependent processing and hardware efficient design to achieve a scalable low power consumption of up to 2.93 mW, providing all required possible LTE SC-FDMA modulated signals up to a maximum sampling rate of 245.76 MS/s. The 11-bit DAC which was re-used courtesy of MediaTek enabled translation of the digital data to an analog output at the maximum data rate. The DAC exhibited a DNL of ± 0.8 LSB and an INL of ± 2 LSB. The quadrature VCO which was a modified implementation of the ones designed for the passive structure test chip performed as expected and measured previously with a tuning range of 24% around 2.65 and 2.34 GHz for the planar and solenoid structures respectively. The phase noise was measured to be -113 and -119 dBc/Hz at 1 MHz offset respectively.

The use of 3D partitioning along with specialized passive structures allowed for better energy efficiency and noise immunity. The digital power link consumed 2 mW from a 1 V supply resulting in an energy efficiency of 0.37 pJ/bit. This is a substantial reduction in power compared to a separate-die operating scheme with 33.3 pJ/bit and compared to a 2.5D interposer design emulation having an energy efficiency of 0.83 pJ/bit. Utilizing the separation of die substrates between the digital and analog domain, along with the vertical solenoid structure employing a magnetic field aligned mostly in parallel with the substrate, we are able to reduce spurious tones in the VCO output by up to 45 dB.

Overall this design illustrates the great potential of future use of 3D integration to create complex, functionally heterogeneous systems to achieve better energy efficiency and small form factor. The use of the vertical dimension for integration enables using specialized process nodes for each tier, optimizing each functional block. It also allows reducing interconnect length and thus reduce the required signal buffering along the data path. Separation of substrate between tiers helps to minimize coupling and cross-talk through the substrate between noisy and sensitive circuits. The vertical integration does introduce new, unwanted effects of capacitive and inductive coupling, however these can be modeled, mitigated and reduced to achieve good overall system performance.

Going forward, 3D integration holds the potential to bring even more functionality and diversification into closely integrated mobile systems with high-end performance. These advancements coincide with conventional lithographic scaling and add another dimension

of added performance and cost reduction. As stacking and packaging technology matures, and demand for energy efficient, high data rate and small form factor systems increase, 3D integration appears to be poised at the intersection to fulfill these needs and beyond.

Chapter 5

Conclusion and Future Directions

The exploding demand for communication and the ever increasing demand for high data rates while operating on mobile, low power systems keeps driving technology to its limits. The conventional source of circuit performance improvement and cost reduction - lithographic scaling is beginning to lose its momentum and the cost per transistor is not scaling as quickly as it did in the past few decades. New approaches however are poised to pick up where Moore's law is somewhat lacking. 3D integration has the potential to offer the added benefit in system power and performance to drive technology to meet these new, ever increasing challenges.

The addition of a new dimension into our design environment calls for the examination of the circuit and system level partitioning in the design. The ability to have a high density, low-parasitic interconnect allows for different choices in system block allocation among tiers in a 3D design and perhaps also in the partitioning of of the circuit blocks themselves across tiers. When considering the potential added benefit of such integration techniques as 3D die stacking, it is most evident that it offers the most benefits in future designs where the data rates and required bandwidth are high, while the benefits of a low power design are greatest such as in mobile devices.

This thesis set out to examine what are the new possibilities, benefits and challenges of developing complex, low-power systems which incorporate digital, analog and RF circuits for cellular applications. The thesis examined this topic beginning from the circuit and component level up to the system level while developing tools, methods and insights along

the way. Following is a brief summary of the main conclusions and contributions this thesis has developed along this journey.

5.1 Summary of Contributions

The thesis emphasizes the design and exploration of future packaging technologies by re-examining and analyzing circuits from the block to the system level through models, simulations and fundamental principles. The broad goal remains throughout to assess the final, system-level benefits of using such circuit and system techniques for high-level applications. Below is a summary of the main contributions of this thesis in the various topics of research.

5.1.1 Passive Devices in 3D-IC

We began by turning our focus to some basic principles and new possibilities in 3D integration. Examining the low level impact of close vertical integration of digital and analog circuits led to the derivation of simplified circuit models as well as analytic analysis of coupling effects. A simple use of fast 3D field solvers was suggested for the fast design and simulation of various passive structures and their impact on the circuit design.

We proposed the use of a vertical solenoid inductor, utilizing the 3D environment itself in order to improve circuit performance and immunity. Our analysis and measurements demonstrated that use of such a solenoid inductor in a VCO structure yielded an improved inductance per given die area, 2x better quality factor and thus also a 6 dB improvement in phase noise performance compared to the classical conventional 2D planar inductor structure widely in use today.

Further analysis demonstrated the improved immunity to coupling of noise from adjacent digital clock lines. Acting as an aggressor, such clock lines represent a common issue in many modern SoCs. The predominant solution in such integrated systems today is to waste area by creating large spatial distances and employ large guard ring structures. The proposed solution we presented on the other hand still enables the use of a small form factor

benefiting from the close vertical integration of 3D stacking while still exhibiting improved immunity to such noise coupling.

These techniques help illustrate the concept of using the new 3D environment to our advantage. The stacking process gives us more than just more layers of interconnect, it opens up the door to new possibilities, designs and structures which were simply not possible before in planar CMOS. Exploring these new possibilities with a combination of fundamental analysis, models and circuit simulations will allow deriving the greatest benefits from this new dimension.

5.1.2 Data-Dependent Signal Processing

The second aspect we turned our focus to was data-dependent energy-efficient signal processing. This topic is not limited to 3D design and is widely used and the topic of much research in general. But with the over-arching objective of investigating system level benefits of integrating analog and digital circuits for RF applications it becomes necessary to optimize all aspects of the system design in order to obtain an overall performance gain and energy savings.

We focused our research on the implementation of the LTE basebandUL signal processor for signal generation for our future 3D-IC transmitter design. This application is a prime example where such 3D integration proves beneficial due to the high computation complexity involved, high data rates required along with the stringent power budget desired due to the mobile platform used. This architecture also acts as a good general test case since at its heart are FFT calculation blocks which have a very wide use in several other communication protocols such as WiFi and GPS as well as countless other applications in image processing, video coding, sensing and more.

We presented a baseband processor which enables bit mapping to complex constellations, processing via a mixed-radix variable length FFT and IFFT blocks, cyclic-prefix addition and oversampling via a polyphase FIR filter. We implemented circuit level efficient hardware techniques such as voltage scaling, frequency scaling and clock gating to save power when possible. We implemented efficient hardware circuits to carry out dedicated computation

in the design such as complex multiplication and vector rotation. Furthermore we utilized our knowledge of the communication protocol and the system to efficiently share resources and minimize redundancy in the design. A-priori knowledge of the signal statistics also allowed optimizing the computational algorithm being used and avoid using a general case solution, but rather one tailored to our system, thus providing further power savings.

The implemented design operated up to a rate of 30.72 MHz for the baseband operation and 245.76 MHz for the x8 oversampling filter output. The FFT and IFFT blocks operated reliably down to a voltage of 0.61 V for the lowest required bandwidth and consumed up to 2.9 mW along with a power consumption of up to 0.9 mW for the interpolating filter. The design achieved an energy efficiency of 95 pJ per sample for the worst case, which corresponds to more than 4 times improvement compared to other such designs.

Use of these approaches is crucial in order to optimize the design and is an integral step in the higher-level system design. Discussing the overall benefits of 3D integration for RF applications would not be possible without considering the impact of such a major component in the system such as the digital block. The use of a true signal generation block also serves as a better test case for a realistic component existing in the final system along with its accompanying noise and area consumption, helping to validate the results and findings of the entire work.

5.1.3 Analog-Digital 3D Integration

The final "dimension" of this thesis was focused on bringing all the previous elements and learnings together to create a high-level system comprising of digital, analog and RF circuits to co-exist together in harmony. This final part represents the attempt to address the underlying question of this research work - when and where, if at all, is it beneficial to use 3D integration for RF applications? Bringing together the various circuit blocks to create a higher order system helped examine this question from several different aspects and preform comparisons.

A specific application was chosen to help answer these questions, part of an LTE UL transmitter chain. As mentioned previously, LTE is a widely used cellular communication

protocol which is also the basis for currently developed future protocols. It represents a good example of complex integrated systems which require highly optimized components to achieve good system performance and energy efficiency. This design was also able to benefit from the first two research topics described and implement them together in the system to build upon them and gain further insight.

The transmitter design included the digital baseband processor block described above as well as digital-to-analog conversion circuits, digital and analog routing, LO generation using solenoid inductor VCO designs as previously described, and upconversion mixers. The entire design was made highly configurable and flexible. This flexibility is what enabled the extensive examination and comparison of the proposed design and system 3D partitioning throughout the 3D stack. The existence of replications of the various circuit blocks along with the comprehensive digital and analog routing and reconfiguration options not only allowed testing of the system performance, but also evaluating the exact same system configured in completely different ways.

Measurement results indicated that partitioning between the digital and analog sections after the digital baseband inherently consumes more energy due to the need to transmit the digital data over some serial link. This data transmission is obviously most expensive in the separate die scenario, where different packaged dies on the board fill the functions of digital baseband and analog/RF blocks. It is shown that a reduction in link power can be achieved by integrating the system into one package. The energy efficiency of transmitting digital data in a 2.5D interposer solution or SoC platform are roughly the same as that of a 3D stacked solution. However, the inherently shorter interconnect length presented by vertically stacking die tiers one on top of each other is what yields a further reduction in power.

Furthermore, the 3D stack offers the ability to have a good separation between the digital and RF operating domains by using different die tiers which do not share a substrate. This property, along with the use of conventional guard ring structures and implementing the previously proposed solenoid vertical inductor structure for better inductive noise coupling immunity yield a far better isolation in the 3D case than that of an SoC solution. Alternatively, to achieve such isolation, an SoC die will require a much larger area and foot-print in order

to separate the blocks and avoid noise and coupling. This area increase will also increase in turn the power consumption due to even longer routing and buffering requirements.

Overall, we have demonstrated a strong case in favor of use of 3D integration for such RF applications. We have demonstrated that issues of noise and coupling can be modeled, mitigated and minimized to achieve acceptable circuit performance. The added benefits of 3D will most likely shine in applications where a very small form factor is desirable and when data rates become very high such that the interconnect and link budget become a dominant factor in the overall system performance. Both of these criteria will likely be met by the future cellular communication protocols of the years to come and the proliferation of wireless and connected devices inspired by the Internet of Things (IoT).

5.2 Future Directions

The work in this thesis is built upon a great deal of research previously conducted. Analysis of circuit integration, noise coupling, system design as well as advances in 3D packaging and TSV modeling allows broadening the horizon of possibilities in this field. However, we have barely scratched the surface of possibilities and potential. There are numerous ways in which we may extend this research work and take it forward. Below are a few topics which are of potential interest for further research and expansion on the topic of 3D integration.

5.2.1 Full Transceiver Design

The main target application described in this thesis work was part of an LTE UL transmitter chain for mobile devices. Due to the limited time and resources, only the main parts of the chain were implemented and the receiver side was not designed. These main components were sufficient in developing the concepts desired and allowed carrying out research on the main topics of analog-digital integration and system partitioning in 3D-IC. It is however desirable to complete the entire system design in order to be able to fully capture all effects and evaluate more comprehensively the impact on the entire system.

A continuation of the work presented could strive to implement a full mobile transceiver.

This work could reuse many of the blocks and concepts described in this thesis work and expand upon them. The digital baseband block can be reused and also duplicated and slightly modified to accommodate for the baseband processing of the DL. The DL side of LTE consists of performing a variable length 128-2048/1536 point DFT calculation depending on the channel size and then selecting from the output only the relevant RBs. This is the inverse process of the one described in this work and calls for a calculation of a DFT with only a subset of the outputs being of interest. Therefore, transform decomposition can be used as well in this case with slight modifications [164]. The mixed radix variable DFT block is not required for the receiver side since the DL in LTE utilizes conventional OFDMA and not SC-FDMA.

The DAC blocks can be re-used, and ADC blocks need to be designed for the receiver. The VCO blocks can be utilized as-is, however for a full system, a full closed-loop PLL must be designed to accurately control the carrier frequency for both transmission and reception. The mixer block should be redesigned with more thorough testing to achieve adequate performance for both the up and down-conversion process.

The addition of a PA and LNA as the front end components will be most desirable and will also allow extending even further the research into 3D integration. The PA is a high-power, frequency-selective block and may benefit from 3D partitioning as well as pose some new interesting challenges to the overall system design. The LNA on the other hand is another good example of a sensitive, low-power, low-noise analog circuit which could also benefit from the advantages of 3D vertical stacking.

Creating such a complete, stand-alone, system will be another great step forward in show-casing the possible future potential of 3D integration for complex, high data-rate systems. The work will allow exploring many new topics and challenges in system and block level partitioning and design. The work done in this thesis will help to make the work on such a future endeavor a much more tractable mission given the broad foundation laid forth.

5.2.2 Power Generation in 3D-IC

Other applications which have potential to benefit from 3D integration are the topics of power management and supply. Enabling integration of the main voltage supply circuits and passives onto a single die stack will enjoy the benefits of foot print reduction, system functionality diversification and potentially reduced passive component values. The work could investigate the potential benefit and challenges of integrating in a 3D stack both digital control and power management (and perhaps energy harvesting) circuits along with high power switches as well as passive energy storing components.

This will continue the line of research which we began in this thesis, exploring digital and analog circuit integration. However, unlike this work which focused on the low-power, RF domain, the new work can target high-power, low frequency systems. Expanding on the research of analog-digital coupling at high frequencies, the new work could investigate in more depth the issues of thermal stress, its affects on system performance as well as methods for mitigation in 3D-IC when implementing high-power circuits.

Furthermore, a future project could examine the use of the solenoid inductor proposed in this work, as well as other possible structures in 3D-IC for use as the core magnetic components of voltage regulators. The work could also potentially compare the 3D stacked die solutions to other proposed 2.5D solutions as well as implementations where the inductors are integrated into the package. Issues such as footprint, losses, frequency of operation, current handling and more can be compared and evaluated.

By targeting other application spaces we will expand our understanding of the true benefits and challenges of 3D integration. The integration of power generation and management onto the die stack is highly desirable since it is a ubiquitous function that will exist in any system at all, no matter the ultimate application space. No system will be truly complete and self-sufficient without the integration of these building blocks, which make them a prime candidate for future 3D integration.

5.2.3 Heterogeneous Integration

In this work, we have mentioned many times the potential benefits of 3D integration to allow for heterogeneous integration. We have mainly focused in this work on functional diversification, bringing together circuits of different functionality such as digital, analog and RF to create a more complex system design. However, a potential big advantage of 3D post-processing is to also incorporate different technologies and dies fabricated in different process nodes in the vertical 3D stack. Combining different process nodes allows customizing and optimizing the process for each functionality. We can choose an advanced, high-end node for the digital part, and a slightly more mature, less variable node for the analog blocks. We may further choose other technologies to suit different goals such as power delivery, optics, memory and so on. Such heterogeneous integration is already currently widely used in applications of memory controllers and memory stacks as well as small, mobile camera optics and sensors in mobile phone cameras.

A future project may attempt to investigate, as the main goal of this research did, the system level benefits of using heterogeneous integration along with exploring the main challenges associated with it. These challenges will most likely include both challenges in terms of material, yield and cost as well as the issues of interfacing the different domains as well as designing a robust system which can tolerate the greater extent of variability.

The proposed work will leverage key findings from this thesis regarding system level integration and issues such as impact of noise and coupling between different die tiers to better assist in the new designs. Furthermore, several of the designed circuits and blocks can be re-used in a design or adapted to a different process node for integration.

Such research will extend and examine further the extent to which 3D heterogeneous integration may indeed offer a true benefit, and in which specific areas to system design. The research will also allow identifying and addressing new challenges that arise, which are unique to such integration. Solving these challenges and understanding in which cases we derive the greatest benefit from 3D integration will help bring it one step closer to a viable solution for an ever increasing number of applications.

Appendix A

QAM Mapping

The following tables list the mapping between data bits to complex valued symbols based on the desired QAM constellation. The values are normalized for unity average power over the entire constellation.

Table A.1: BPSK modulation mapping

b_i	**I**	**Q**
0	$1/\sqrt{2}$	$1/\sqrt{2}$
1	$-1/\sqrt{2}$	$-1/\sqrt{2}$

Table A.2: QPSK modulation mapping

b_i, b_{i+1}	I	Q
00	$1/\sqrt{2}$	$1/\sqrt{2}$
01	$1/\sqrt{2}$	$-1/\sqrt{2}$
10	$-1/\sqrt{2}$	$1/\sqrt{2}$
11	$-1/\sqrt{2}$	$-1/\sqrt{2}$

Table A.3: 16-QAM modulation mapping

$b_i, b_{i+1}, b_{i+2}, b_{i+3}$	I	Q	$b_i, b_{i+1}, b_{i+2}, b_{i+3}$	I	Q
0000	$1/\sqrt{10}$	$1/\sqrt{10}$	1000	$-1/\sqrt{10}$	$1/\sqrt{10}$
0001	$1/\sqrt{10}$	$3/\sqrt{10}$	1001	$-1/\sqrt{10}$	$3/\sqrt{10}$
0010	$3/\sqrt{10}$	$1/\sqrt{10}$	1010	$-3/\sqrt{10}$	$1/\sqrt{10}$
0011	$3/\sqrt{10}$	$3/\sqrt{10}$	1011	$-3/\sqrt{10}$	$3/\sqrt{10}$
0100	$1/\sqrt{10}$	$-1/\sqrt{10}$	1100	$-1/\sqrt{10}$	$-1/\sqrt{10}$
0101	$1/\sqrt{10}$	$-3/\sqrt{10}$	1101	$-1/\sqrt{10}$	$-3/\sqrt{10}$
0110	$3/\sqrt{10}$	$-1/\sqrt{10}$	1110	$-3/\sqrt{10}$	$-1/\sqrt{10}$
0111	$3/\sqrt{10}$	$-3/\sqrt{10}$	1111	$-3/\sqrt{10}$	$-3/\sqrt{10}$

Table A.4: 64-QAM modulation mapping

$b_i,\ldots,b_{i+5}$	I	Q	$b_i,\ldots,b_{i+5}$	I	Q
000000	$3/\sqrt{42}$	$3/\sqrt{42}$	100000	$-3/\sqrt{42}$	$3/\sqrt{42}$
000001	$3/\sqrt{42}$	$1/\sqrt{42}$	100001	$-3/\sqrt{42}$	$1/\sqrt{42}$
000010	$1/\sqrt{42}$	$3/\sqrt{42}$	100010	$-1/\sqrt{42}$	$3/\sqrt{42}$
000011	$1/\sqrt{42}$	$1/\sqrt{42}$	100011	$-1/\sqrt{42}$	$1/\sqrt{42}$
000100	$3/\sqrt{42}$	$5/\sqrt{42}$	100100	$-3/\sqrt{42}$	$5/\sqrt{42}$
000101	$3/\sqrt{42}$	$7/\sqrt{42}$	100101	$-3/\sqrt{42}$	$7/\sqrt{42}$
000110	$1/\sqrt{42}$	$5/\sqrt{42}$	100110	$-1/\sqrt{42}$	$5/\sqrt{42}$
000111	$1/\sqrt{42}$	$7/\sqrt{42}$	100111	$-1/\sqrt{42}$	$7/\sqrt{42}$
001000	$5/\sqrt{42}$	$3/\sqrt{42}$	101000	$-5/\sqrt{42}$	$3/\sqrt{42}$
001001	$5/\sqrt{42}$	$1/\sqrt{42}$	101001	$-5/\sqrt{42}$	$1/\sqrt{42}$
001010	$7/\sqrt{42}$	$3/\sqrt{42}$	101010	$-7/\sqrt{42}$	$3/\sqrt{42}$
001011	$7/\sqrt{42}$	$1/\sqrt{42}$	101011	$-7/\sqrt{42}$	$1/\sqrt{42}$
001100	$5/\sqrt{42}$	$5/\sqrt{42}$	101100	$-5/\sqrt{42}$	$5/\sqrt{42}$
001101	$5/\sqrt{42}$	$7/\sqrt{42}$	101101	$-5/\sqrt{42}$	$7/\sqrt{42}$
001110	$7/\sqrt{42}$	$5/\sqrt{42}$	101110	$-7/\sqrt{42}$	$5/\sqrt{42}$
001111	$7/\sqrt{42}$	$7/\sqrt{42}$	101111	$-7/\sqrt{42}$	$7/\sqrt{42}$
010000	$3/\sqrt{42}$	$-3/\sqrt{42}$	110000	$-3/\sqrt{42}$	$-3/\sqrt{42}$
010001	$3/\sqrt{42}$	$-1/\sqrt{42}$	110001	$-3/\sqrt{42}$	$-1/\sqrt{42}$
010010	$1/\sqrt{42}$	$-3/\sqrt{42}$	110010	$-1/\sqrt{42}$	$-3/\sqrt{42}$
010011	$1/\sqrt{42}$	$-1/\sqrt{42}$	110011	$-1/\sqrt{42}$	$-1/\sqrt{42}$
010100	$3/\sqrt{42}$	$-5/\sqrt{42}$	110100	$-3/\sqrt{42}$	$-5/\sqrt{42}$
010101	$3/\sqrt{42}$	$-7/\sqrt{42}$	110101	$-3/\sqrt{42}$	$-7/\sqrt{42}$
010110	$1/\sqrt{42}$	$-5/\sqrt{42}$	110110	$-1/\sqrt{42}$	$-5/\sqrt{42}$
010111	$1/\sqrt{42}$	$-7/\sqrt{42}$	110111	$-1/\sqrt{42}$	$-7/\sqrt{42}$
011000	$3/\sqrt{42}$	$-3/\sqrt{42}$	111000	$-3/\sqrt{42}$	$-3/\sqrt{42}$
011001	$3/\sqrt{42}$	$-1/\sqrt{42}$	111001	$-3/\sqrt{42}$	$-1/\sqrt{42}$
011010	$1/\sqrt{42}$	$-3/\sqrt{42}$	111010	$-1/\sqrt{42}$	$-3/\sqrt{42}$
011011	$1/\sqrt{42}$	$-1/\sqrt{42}$	111011	$-1/\sqrt{42}$	$-1/\sqrt{42}$
011100	$3/\sqrt{42}$	$-5/\sqrt{42}$	111100	$-3/\sqrt{42}$	$-5/\sqrt{42}$
011101	$3/\sqrt{42}$	$-7/\sqrt{42}$	111101	$-3/\sqrt{42}$	$-7/\sqrt{42}$
011110	$1/\sqrt{42}$	$-5/\sqrt{42}$	111110	$-1/\sqrt{42}$	$-5/\sqrt{42}$
011111	$1/\sqrt{42}$	$-7/\sqrt{42}$	111111	$-1/\sqrt{42}$	$-7/\sqrt{42}$

Appendix B

Configuration

B.1 Scan Chain Architecture

In order to control the various options and setting of the digital signal processing block and various circuits in the 3D transmitter test chip, a set of programmable configuration bits are used. The configuration bits are loaded sequentially into the chip via a shift register scan chain. The basic unit cell topology for the scan chain is shown in Fig. B-1a. In addition the snapshot memory output is stored in a set of shift registers and read out by concatenating these read cells to the scan chain write cells. The schematic of the read cells is shown in Fig. B-1b.

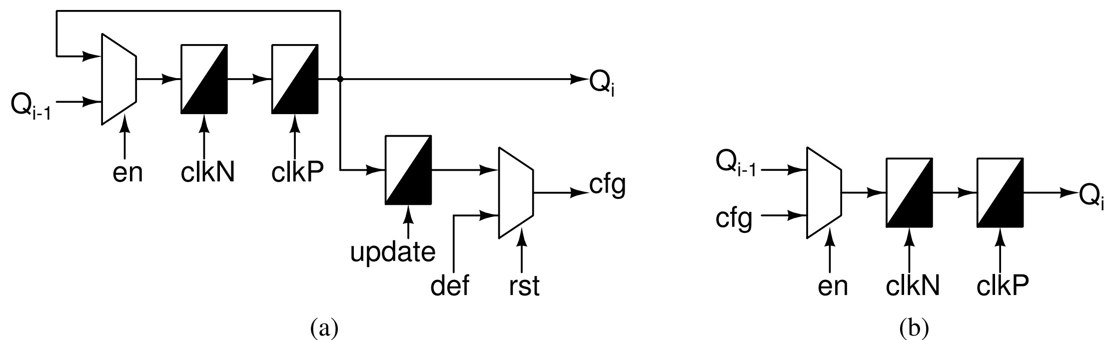


Figure B-1: Scan chain unit cell for (a) writing configuration bits and (b) reading chip data

The scan chain uses two non-overlapping clock signals - *clkP* and *clkN*. The use of two separate non-overlapping signals ensures that there will be no hold timing violations in

the scan chain operation since we are able to increase as necessary the delay between the positive clock active time and that of the negative clock. A separate *update* signal is used to latch the scan chain data and make it visible to the system. This technique is used in order to avoid unwanted and unexpected temporary configuration values being presented to the system while the scan chain is being loaded. Each write cell also has a default value stored which is initially loaded to the bit cell when the system is reset.

B.2 3D Scan Chain

In order to support creating a scan chain which also loads the required configuration to both the top and bottom tier dies, a special layout was used which takes advantage of the specific configuration used for die stacking where the same die layout is rotated 180° and functions as both the top and bottom tier dies in the 3D stack. Fig. B-2 illustrates the conceptual block diagram of the 3D scan chain. The input scan data, clock and control bits coming from the external pad on the top tier die enters a MUX and selected in the top tier die, it then passes to the top tier scan blocks (only two illustrated in the figure). At the output of the last scan block the signals are sent both to a microbump (which is disconnected on the top tier) and a TSV. The TSV in turn is connected to the input microbump and MUX as it is a flipped version of the top tier. In the bottom tier the input from the microbump is selected (since the pad is disconnected on the bottom tier). The signals again pass through all scan blocks on the bottom tier as they did on the top and eventually connect to the microbump/TSV pair. This time the TSV is disconnected, but the microbump routes the output data to an external pad on the top tier which is used to verify and monitor the scan chain integrity.

A critical point to note is the input MUX selection signal. Unlike other MUXes across the chip and selection signals which differentiate the top and bottom tiers by using control bits loaded to the scan chain, here we are not able to use the scan chain itself to load an indicator bit to whether the specific die is acting as the top or bottom tier. Therefore we require another indicator which would help determine this. In order to avoid adding an explicit extra signal for this use, we use the digital I/O supply voltage. This power supply

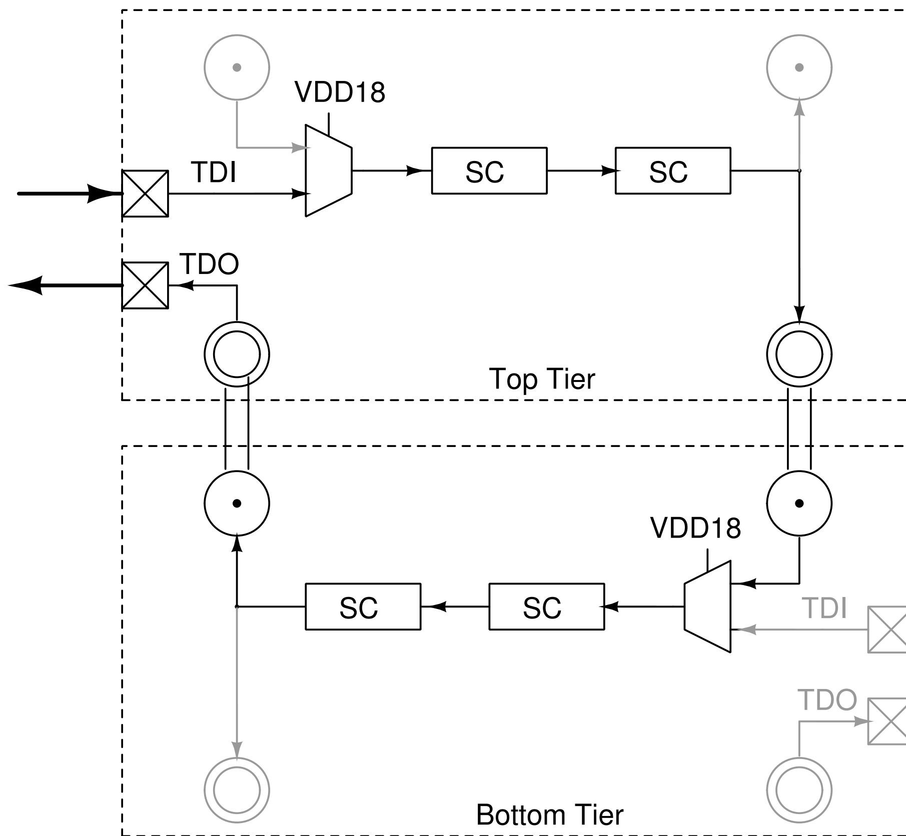


Figure B-2: Scan chain topology for 3D test chip

Table B.1: Scan chain block order

Index	Module	Reference	Index	Module	Reference
1	Baseband	Table B.3	8	Bias instance 2	Table B.7
2	General control	Table B.2	9	Mixer instance 2	Table B.6
3	DAC instance 1	Table B.4	10	DAC instance 2	Table B.4
4	VCO instance 1	Table B.5	11	VCO instance 3	Table B.5
5	Bias instance 1	Table B.7	12	Bias instance 3	Table B.7
6	Mixer instance 1	Table B.6	13	Mixer instance 3	Table B.6
7	VCO instance 2	Table B.5			

is used to power the digital pads and is only used on the top tier, since only the top tier has exposed external pads. This supply voltage is therefore not passed on to the bottom tier die through any TSVs and is only active on the top die. We use a tie-high standard cell as shown in Fig. B-3, which is pulled to a logical high by this supply voltage, and passed through a level shifter (as explained in section 4.2.3) to convert it to the digital core voltage of 1 V in order to act as our desired selection signal for the scan chain input MUX. This will ensure the proper path selection of the scan signals on both the top and bottom tier dies.

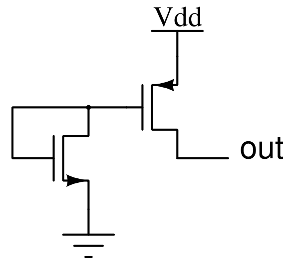


Figure B-3: Tie-high cell schematic

The scan chain blocks on each die tier are connected in the order indicated in table B.1. These blocks are repeated again for the bottom tier die. The various configuration options for each module are detailed in section B.3. The overall scan chain size is 700 bits (350 for each tier).

B.3 Configuration Options

Table B.2: General configuration bits

Name	Description	Default	# bits
Top tier indicator	Signal indicating current die is the top tier die. 0x0 - bottom tier 0x1 - top tier	0x0	1
Clock divider enable	Enable clock divider and input amplifier	disabled	1
Single chip clock	Indicate single or dual chip mode operation. 0x0 - Dual chip, divide clock by 4 0x1 - Single chip, divide clock by 2	0x1	1
Vertical clock path	Enable passing of clock signal to bottom tier die	disabled	1
Buffer control	Tri-state digital buffer drive strength	0x8	4
Tx Xbar vertical	Enable digital Tx crossbar vertical routing	disabled	1
Rx Xbar vertical	Enable digital Rx crossbar vertical routing	disabled	1
Digital Xbar line enable	Enable line routing between Tx and Rx crossbars	disabled	1
Digital Xbar line direction	Indicate line routing direction. 0x0 - From Rx to Tx crossbar 0x1 - From Tx to Rx crossbar	0x1	1
Analog Xbar line enable	Enable line routing between analog crossbars	disabled	1

Continued on next page

Table B.2 – continued from previous page

Name	Description	Default	# bits
Analog Xbar pad enable	Enable pad connection to each analog crossbar	disabled	2
Analog Xbar vertical	Enable vertical routing for each analog crossbar	disabled	2
Hi-speed I/O enable	Enable high-speed digital I/O pads	disabled	1
Hi-speed I/O direction	Indicate I/O pad direction. 0x0 - Receive, use deserializer 0x1 - Transmit, use serializer	0x1	1
Hi-speed I/O inversion	Invert received bits from I/O pads	disabled	1
Hi-speed I/O pull-up	Use internal differential 50 Ω pull-up resistor with I/O pad output amplifier	enabled	1
Hi-speed I/O bias select	I/O transmitter amplifier bias current select. 0x0 - 6.25 μA 0x1 - 12.5 μA 0x2 - 25 μA 0x3 - 50 μA	0x1	2
Hi-speed I/O current mirror	I/O transmitter amplifier current mirror scale ratio for bias current. 0x0 - x1000 0x1 - x2000 0x2 - x3000 0x3 - x4000	0x1	2
VCO output enable	Enable each of the VCO instance's open drain output buffers	disabled	3

Continued on next page

Table B.2 – continued from previous page

Name	Description	Default	# bits
External LO enable	Enable an external LO to be supplied to bypass internal VCOs	disabled	1
Reserved	Unused reserved configuration bits	0x0	7

Table B.3: Baseband module configuration bits

Name	Description	Default	# bits
Memory address	Used for both main memory address for data load and snapshot memory data read	0	16
Memory data	Data input to be loaded in main memory	0	24
Memory write enable	Enable signal for loading data	disabled	1
Memory power down	Power down memory	disabled	1
Memory delay select	Set memory buffer strength	0xa	4
SC-FDMA enable	Enable SC-FDMA processing or bypass all DFT/IDFT and feed memory output directly to FIR filter	enabled	1
QAM enable	Enable QAM mapping of bits to symbols	enabled	1
QAM select	Select QAM mapping to use 0x0 - Binary (1 bit per symbol) 0x1 - QPSK (2 bits per symbol) 0x2 - 16-QAM (4 bits per symbol) 0x3 - 64-QAM (6 bits per symbol)	0x3	2

Continued on next page

Table B.3 – continued from previous page

Name	Description	Default	# bits
Resource blocks	Number of resource blocks to use (N_{rb}). This Will determine FFT size $\rightarrow M = 12 \times N_{rb}$. Must satisfy $N_{rb} = 2^{\alpha} \times 3^{\beta} \times 5^{\gamma}$ Where α, β and γ are non-negative integers and $1 \leq N_{rb} \leq 100$	0x64	7
Auto offset enable	Enable automatic offset of data by phase shifting. Cannot be disabled when using Transform Decomposition	enabled	1
Offset	Amount of offset in subcarriers. If auto offset is disabled, Offset is negative from DC (use M/2 to center). Else, Offset is positive from band edge (use (N-M)/2 to center)	0x1ab	11
IDFT size index	IDFT size (or bandwidth) index. Must be larger than DFT size ($12 \times N_{rb}$). 0x0 - 128 (1.4 MHz) 0x1 - 256 (3 MHz) 0x2 - 512 (5 MHz) 0x3 - 1024 (10 MHz) 0x4 - 2048 (20 MHz) 0x5, 0x6 - INVALID 0x7 - 1536 (15 MHz)	0x4	3

Continued on next page

Table B.3 – continued from previous page

Name	Description	Default	# bits
Transform enable	Enable use of transform decomposition. This will only have an effect when $M < 2^{\log N - 1}$. Must use auto offset when enabled. Do not use with N=1536 point IDFT	enabled	1
Cyclic prefix enable	Enable cyclic prefix addition to data	enabled	1
Extended cyclic prefix	Enable extended cyclic prefix	disabled	1
OFDM symbols	Number of OFDM symbols to cycle through (N_s). There are 7 OFDM symbols in a slot (6 when using extend CP) - 0.5 ms. There are 20 slots in a radio frame - 10 ms. Make sure that total size does not exceed memory capacity $N_s \times M \times bps \leq 1,011,840$ bits	0x8c	8
Filter enable	Enable x8 FIR interpolation filter	enabled	1
Snapshot enable	Enable snapshot memory	enabled	1
Snapshot read	Read from snapshot memory	disabled	1
Snapshot select	Select output to be saved to snapshot memory. 0x0 - DFT input 0x1 - DFT output 0x2 - IDFT input 0x3 - IDFT output	0x1	2

Continued on next page

Table B.3 – continued from previous page

Name	Description	Default	# bits
Snapshot symbol	Select symbol to be saved to snapshot (zero based index). Must be smaller than no. of symbols ($< N_s$)	0	8
Trigger select	Select output for each trigger (repeated for 4 triggers). 0x0 - DFT enable 0x1 - DFT start 0x2 - DFT out begin 0x3 - DFT done 0x4 - IDFT en 0x5 - IDFT start 0x6 - IDFT out begin 0x7 - IDFT done 0x8 - Zero signal 0x9 - Zero padding done 0xa - Cyclic prefix done 0xb - FIR start 0xc - FIR out begin 0xd - Snapshot write enable 0xe - Snapshot done 0xf - Ground	0xf	4

Continued on next page

Table B.3 – continued from previous page

Name	Description	Default	# bits
Output select	High speed 22 bit digital output selection. 0x0 - Output after FIR 0x1 - IDFT input 0x2 - DFT output 0x3 - DFT index and zero pad index 0x4 - IDFT index and cyclic prefix index 0x5, 0x6, 0x7 - Ground	0x0	3
Reverse output	Reverse output bit order for die-to-die mirroring	disabled	1
Bandwidth clock divide	Divide clock according to bandwidth (N). Divide by $2^{11-\log N}$	disabled	1
Filter clock divide	Divide clock by 8 from FIR filter clock	enabled	1
Baseband enable	Global enable for baseband digital block	disabled	1
Output clock select	Bypass clock output from digital block output	no-bypass	1
Snapshot output	Registers for reading snapshot memory		32

Table B.4: DAC module configuration bits

Name	Description	Default	# bits
DAC enable	Enable DAC operation	disabled	1
DAC gain	I/Q gain balance	0x0	2
LPF enable	Enable LPF operation or bypass	enabled	1
LPF double C	Double the capacitor size. Used to toggle between 20/40 MHz nominal bandwidth. 0x0 - No doubling (40 MHz) 0x1 - Double capacitance (20 MHz)	0x0	1
LPF common mode	LPF common mode voltage select. 0x0 - 611 mV 0x1 - 635 mV 0x2 - 660 mV 0x3 - 684 mV 0x4 - 709 mV 0x5 - 733 mV 0x6 - 782 mV 0x7 - 812 mV	0x2	3
LPF C select	LPF capacitor adjustment. This capacitance is doubled when BW = 20 MHz 0.63 - 1.87 pF LSB $\approx$ 9.8 fF	0x26 (1 pF)	7
LPF R select	LPF resistor adjustment. 0x0 - x1 0x1, 0x2 - x0.992 0x3 - x0.987	0x0	4
LDO 1.5 V output	LDO output voltage adjustment. 0x0 - 1.49 V 0x1, 0x3 - 1.58 V 0x2 - 1.55 V	0x2	2
LDO 1.5 V reference	LDO reference adjustment. 0x0 - 849 mV 0x1 - 837 mV 0x2 - 824 mV 0x3 - 812 mV	0x2	2
DC monitor select	Output probe selection. 0x0 - Ground 0x1 - Common mode voltage 0x2 - LDO 1.5 V output 0x3 - LDO 1 V output	0x3	2

Table B.5: VCO module configuration bits

Name	**Description**	**Default**	**# bits**
VCO enable	Enable VCO operation	disabled	1
VCO control	Switched capacitor bank control signal. Varies tank capacitance between approximately 0.96 - 3.3 pF	0x0	3

Table B.6: Mixer module configuration bits

Name	**Description**	**Default**	**# bits**
Mixer AG control	Connect internal DAC IF to amplifier gate	disconnected	1
Mixer AO control	Connect internal DAC IF to mixer input	disconnected	1
Mixer BG control	Connect external IF from crossbar to amplifier gate	disconnected	1
Mixer BO control	Connect external IF from crossbar to mixer input	disconnected	1

Table B.7: Bias module configuration bits

Name	**Description**	**Default**	**# bits**
VCO bias enable	Enable VCO tail current bias	disabled	1
VCO bias	Select signal for VCO bias. Consists of 4 parallel current sources. Each two bits control a current source. 0x0 - 12.5 μA 0x1 - 25 μA 0x2 - 50 μA 0x3 - 100 μA	0xaa (200 μA)	8
DAC bias enable	Enable DAC bias current	disabled	1
LDO bias 1 V	DAC's LDO 1 V bias current. 0x0 - 6.25 μA 0x1 - 12.5 μA 0x2 - 18.75 μA 0x3 - 25 μA	0x1	2
LDO bias 1.5 V	DAC's LDO 1.5 V bas current. 0x0 - 23.4375 μA 0x1 - 25 μA 0x2 - 26.5625 μA 0x3 - 28.125 μA	0x1	2
DAC bias	DAC voltage-to-current bias select. 0x0 - 12.5 μA 0x1 - 14.0625 μA	0x0	1
LPF bias	LPF bias current select. 0x0 - 12.5 μA 0x1 - 25 μA 0x2 - 37.5 μA 0x3 - 50 μA	0x1	2
Reserved	Unused reserved configuration bits	0x0	3

Table B.8: Transmitter test chip I/O pads

Group	Name	Type	Dir.	Voltage	Freq.	#	Comments
Scan	TCKP	Digital	IN	1.8 V	1 MHz	1	Clock phase 0
	TCKN		IN			1	Clock phase 1
	TRST		IN			1	Global reset
	TEN		IN			1	Scan enable
	TUP		IN			1	Scan update
	TDI		IN			1	Scan data in
	START		IN			1	Baseband start
	TDO		OUT			1	Scan data out
Baseband	CLKIN	Digital	IN	1 V	980 MHz	2	Hi-speed clock input
	CLKOUT		OUT	1 V	490 MHz	2	Hi-speed clock output
	CLKDIV		OUT	1.8 V	30 MHz	1	Divided clock output
	DATA[10:0]		I/O	1 V	490 MHz	22	Hi-speed BB data
	TRIG[3:0]		OUT	1.8 V	30 MHz	4	Aux. BB triggers
DAC	DAC0	Analog	I/O	1.8 V	245 MHz	4	Analog BB symbols
	DAC1		I/O		245 MHz	4	
	DCMON		OUT		DC	1	Analog DAC probe
VCO	VCO0	Analog	I/O	1 V	2 GHz	4	LO signal
	VCO1		I/O		2 GHz	4	
	VCO2		I/O		2 GHz	4	
	VTUNE		IN		DC	1	VCO varactor bias
Mixer	MIX0	Analog	OUT	1 V	2 GHz	2	RF output
	MIX1					2	
	MIX2					2	
Bias	REF	Analog	IN	1.2 V	DC	2	Bias reference voltage
Power	DVDD	Power	I/O	1 V	DC	2	Digital supply
	DVDDC			1 V		1	Top BB core
	DVDDF			1 V		1	Top BB FFT
	DVDDI			1 V		1	Top BB IFFT
	DVDDB			1 V		1	Bottom BB
	DVDDIO			1.8 V		1	Digital I/O
	A0VDD			1 V		1	VCO0, MIX0 supply
	A1VDD			1 V		1	VCO1, MIX1 supply
	A2VDD			1 V		1	VCO2, MIX2 supply
	AVDDIO			1.8 V		2	Analog I/O

Bibliography

[1] G. E. Moore, “Cramming more components onto integrated circuits,” *Electronics*, vol. 38, Apr. 1965.

[2] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits, Second Edition*. Cambridge University Press, 2 ed., Dec. 2003.

[3] S. Narendra, L. Fujino, and K. Smith, “Through the Looking Glass?The 2015 Edition: Trends in Solid-State Circuits from ISSCC,” *IEEE Solid-State Circuits Magazine*, vol. 7, no. 1, pp. 14–24, 2015.

[4] J. Handel, “Why Migration to 20nm Bulk CMOS and 16-14nm FinFETs is Not Best Approach for Semiconductor Industry,” Jan. 2014.

[5] W. Arden, M. Brillout, P. Cogez, M. Graef, B. Huizing, and R. Mahnkopf, “More-than-Moore white paper,” *International Technical Roadmap for Semiconductors*, vol. 2, p. 14, 2010.

[6] V. F. Pavlidis, *Three-dimensional integrated circuit design*. Amsterdam ; Boston: Elsevier/Morgan Kaufmann, 2009.

[7] P. E. Garrou, M. Koyanagi, and P. Ramm, *Handbook of 3D Integration : 3D process technology*. Weinheim, Germany : Wiley-VCH, 2014., 2014.

[8] “Interconnect,” tech. rep., International Technology Roadmap for Semiconductors (ITRS), 2013.

[9] Y. Lamy, J. Colonna, G. Simon, P. Leduc, S. Cheramy, and C. Laviron, “Which interconnects for which 3d applications? Status and perspectives,” in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–6, Oct. 2013.

[10] J.-W. Im, W.-P. Jeong, D.-H. Kim, S.-W. Nam, D.-K. Shim, M.-H. Choi, H.-J. Yoon, D.-H. Kim, Y.-S. Kim, H.-W. Park, D.-H. Kwak, S.-W. Park, S.-M. Yoon, W.-G. Hahn, J.-H. Ryu, S.-W. Shim, K.-T. Kang, S.-H. Choi, J.-D. Ihm, Y.-S. Min, I.-M. Kim, D.-S. Lee, J.-H. Cho, O.-S. Kwon, J.-S. Lee, M.-S. Kim, S.-H. Joo, J.-H. Jang, S.-W. Hwang, D.-S. Byeon, H.-J. Yang, K.-T. Park, K.-h. Kyung, and J.-H. Choi, “A 128gb 3b/cell V-NAND flash memory with 1gb/s I/O rate,” in *Solid- State Circuits Conference - (ISSCC), 2015 IEEE International*, pp. 1–3, Feb. 2015.

[11] Z. Or-Bach, "The monolithic 3d advantage: Monolithic 3d is far more than just an alternative to 0.7x scaling," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–7, Oct. 2013.

[12] W. Wahby, A. Dembla, and M. Bakir, "Evaluation of 3dics and fabrication of monolithic interlayer vias," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–6, Oct. 2013.

[13] R. Patti, "Three-Dimensional Integrated Circuits and the Future of System-on-Chip Designs," *Proceedings of the IEEE*, vol. 94, no. 6, pp. 1214–1224, 2006.

[14] A. Rahman and R. Reif, "System-level performance evaluation of three-dimensional integrated circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 8, no. 6, pp. 671–678, 2000.

[15] D. H. Kim, K. Athikulwongse, M. Healy, M. Hossain, M. Jung, I. Khorosh, G. Kumar, Y.-J. Lee, D. Lewis, T.-W. Lin, C. Liu, S. Panth, M. Pathak, M. Ren, G. Shen, T. Song, D. H. Woo, X. Zhao, J. Kim, H. Choi, G. Loh, H.-H. Lee, and S. K. Lim, "3d-MAPS: 3d Massively parallel processor with stacked memory," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, pp. 188 –190, Feb. 2012.

[16] R. Egawa, M. Sato, J. Tada, and H. Kobayashi, "Vertically integrated processor and memory module design for vector supercomputers," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–6, Oct. 2013.

[17] Q. Zhu, B. Akin, H. Sumbul, F. Sadi, J. Hoe, L. Pileggi, and F. Franchetti, "A 3d-stacked logic-in-memory accelerator for application-specific data intensive computing," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–7, Oct. 2013.

[18] X. Zhang, S. Chen, and E. Culurciello, "A second generation 3d integrated feature-extracting image sensor," in *2011 IEEE Sensors*, pp. 1933–1936, Oct. 2011.

[19] V. Re, M. Manghisoni, G. Traversi, L. Gaioni, A. Manazza, and L. Ratti, "Active pixel sensors with enhanced pixel-level analog and digital functionalities in a 2-tier 3d CMOS technology," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–7, Oct. 2013.

[20] D. Temple, D. Malta, E. Vick, M. Lueck, S. Goodwin, M. Muzilla, C. Masterjohn, and M. Skokan, "Advanced 3d mixed-signal processor for infrared focal plane arrays: Fabrication and test," in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–7, Dec. 2014.

[21] T. Bieniek, G. Janczyk, M. Ekwinska, T. Budzynski, G. Gluszko, P. Grabiec, and A. Kociubinski, "Novel methodology for 3d MEMS-IC design and co-simulation on MEMS microphone smart system example," in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–5, Dec. 2014.

[22] H. S. Yang and M. Bakir, "3d integration of CMOS and MEMS using mechanically flexible interconnects (MFI) and through silicon vias (TSV)," in *Electronic Components and Technology Conference (ECTC), 2010 Proceedings 60th*, pp. 822–828, June 2010.

[23] W. Vitale, M. Fernandez-Bolanos, and A. Ionescu, "High-Q 3d embedded inductors using TSV for RF MEMS tunable bandpass filters (4.65-6.8 GHz)," in *Microwave Integrated Circuits Conference (EuMIC), 2012 7th European*, pp. 822–825, 2012.

[24] R. Anthony, S. Kulkarni, N. Wang, and C. Mathuna, "Advanced processing for high efficiency inductors for 2.5d/3d Power Supply in Package," in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–4, Dec. 2014.

[25] K. Ishida, K. Takemura, K. Baba, M. Takamiya, and T. Sakurai, "3d stacked buck converter with 15 um thick spiral inductor on silicon interposer for fine-grain power-supply voltage control in SiP's," in *3D Systems Integration Conference (3DIC), 2010 IEEE International*, pp. 1–4, Nov. 2010.

[26] C. Chen, D.-R. Yost, J. Knecht, D. Chapman, D. Oakley, L. Mahoney, J. Donnelly, A. Soares, V. Suntharalingam, R. Berger, V. Bolkhovsky, W. Hu, B. Wheeler, C. Keast, and D. Shaver, "Wafer-scale 3d integration of InGaAs image sensors with Si readout circuits," in *IEEE International Conference on 3D System Integration, 2009. 3DIC 2009*, pp. 1–4, Sept. 2009.

[27] T. Irisawa, M. Oda, Y. Kamimuta, Y. Moriyama, K. Ikeda, E. Mieda, W. Jevasuwan, T. Maeda, O. Ichikawa, T. Osada, M. Hata, and T. Tezuka, "3d integration of high mobility InGaAs nFETs and Ge pFETs for ultra low power and high performance CMOS," in *2013 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, pp. 1–2, Oct. 2013.

[28] V. Pavlidis and E. Friedman, "Interconnect-Based Design Methodologies for Three-Dimensional Integrated Circuits," *Proceedings of the IEEE*, vol. 97, pp. 123 –140, Jan. 2009.

[29] V. Pavlidis, I. Savidis, and E. Friedman, "Clock distribution networks in 3-D integrated systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 19, pp. 2256 –2266, Dec. 2011.

[30] J. Tada, R. Egawa, and H. Kobayashi, "Design of a 3-D stacked floating-point adder," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–4, Oct. 2013.

[31] V. Pavlidis and E. Friedman, "3-D Topologies for Networks-on-Chip," in *SOC Conference, 2006 IEEE International*, pp. 285–288, Sept. 2006.

[32] Q. Zou, J. Zhan, F. Ge, M. Poremba, and Y. Xie, "Designing vertical bandwidth reconfigurable 3d NoCs for many core systems," in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–7, Dec. 2014.

[33] E. Basha, *Fast Fourier transform on a 3D FPGA*. PhD thesis, Massachusetts Institute of Technology, 2005.

[34] V. Pangracious, H. Mehrez, and Z. Marakchi, "Designing a 3d tree-based FPGA: Optimization of butterfly programmable interconnect topology using 3d technology," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–8, Oct. 2013.

[35] G. Van der Plas, P. Limaye, I. Loi, A. Mercha, H. Oprins, C. Torregiani, S. Thijs, D. Linten, M. Stucchi, G. Katti, D. Velenis, V. Cherman, B. Vandevelde, V. Simons, I. De Wolf, R. Labie, D. Perry, S. Bronckers, N. Minas, M. Cupac, W. Ruythooren, J. Van Olmen, A. Phommahaxay, M. de Potter de ten Broeck, A. Opdebeeck, M. Rakowski, B. De Wachter, M. Dehan, M. Nelis, R. Agarwal, A. Pullini, F. Angiolini, L. Benini, W. Dehaene, Y. Travaly, E. Beyne, and P. Marchal, "Design issues and considerations for low-cost 3-D TSV IC technology," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 293–307, 2011.

[36] J. De Vos, V. Cherman, M. Detalle, T. Wang, A. Salahouelhadj, R. Daily, G. Van der Plas, and E. Beyne, "Comparative study of 3d stacked IC and 3d interposer integration: Processing and assembly challenges," in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–7, Dec. 2014.

[37] A.-C. Hsieh and T. Hwang, "TSV Redundancy: Architecture and Design Issues in 3-D IC," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 20, no. 4, pp. 711–722, 2012.

[38] E. Beyne, "Electrical, thermal and mechanical impact of 3d TSV and 3d stacking technology on advanced CMOS devices - Technology directions," in *3D Systems Integration Conference (3DIC), 2011 IEEE International*, pp. 1–6, Jan. 2012.

[39] Y. Mizushima, Y. Kim, T. Nakamura, S. Kodama, N. Maeda, K. Fujimoto, and T. Ohba, "Impact of Thermomechanical Stresses on Ultra-thin Si Stacked Structure," in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–5, Dec. 2014.

[40] J. Zhang, L. Zhang, Y. Dong, H. Li, C. Tan, G. Xia, and C. Tan, "The dependency of TSV keep-out zone (KOZ) on Si crystal direction and liner material," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–5, Oct. 2013.

[41] K. Matsumoto, S. Ibaraki, K. Sueoka, K. Sakuma, H. Kikuchi, H. Mori, Y. Orii, F. Yamada, K. Fujihara, J. Takamatsu, and K. Kondo, "Thermal design guideline and new cooling solution for a three-dimensional (3d) chip stack," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–8, Oct. 2013.

[42] M. Saeidi, K. Samadi, A. Mittal, and R. Mittal, "Thermal implications of mobile 3d-ICs," in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–7, Dec. 2014.

[43] C. Selvanayagam, J. Lau, X. Zhang, S. Seah, K. Vaidyanathan, and T. Chai, “Nonlinear Thermal Stress/Strain Analyses of Copper Filled TSV (Through Silicon Via) and Their Flip-Chip Microbumps,” *IEEE Transactions on Advanced Packaging*, vol. 32, pp. 720–728, Nov. 2009.

[44] C. Santos, P. Vivet, J.-P. Colonna, P. Coudrain, and R. Reis, “Thermal performance of 3d ICs: Analysis and alternatives,” in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–7, Dec. 2014.

[45] D. Gope, S. Chatterjee, D. de Araujo, S. Chakraborty, J. Pingenot, and R. Camposano, “Device physics aware 3d electromagnetic simulation of Through-Silicon-Vias in system modeling,” in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–5, Oct. 2013.

[46] Z. Xu, A. Beece, D. Zhang, Q. Chen, K.-N. Chen, K. Rose, and J.-Q. Lu, “Crosstalk evaluation, suppression and modeling in 3d through-strata-via (TSV) network,” in *3D Systems Integration Conference (3DIC), 2010 IEEE International*, pp. 1–8, Nov. 2010.

[47] C. Liu, T. Song, J. Cho, J. Kim, J. Kim, and S.-K. Lim, “Full-chip TSV-to-TSV coupling analysis and optimization in 3d IC,” in *2011 48th ACM/EDAC/IEEE Design Automation Conference (DAC)*, pp. 783–788, 2011.

[48] M. Watanabe, M.-A. Fukase, M. Imai, N. Niioka, T. Kobayashi, R. Karel, and A. Kurokawa, “Modeling of substrate contacts in TSV-based 3d ICs,” in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–4, Dec. 2014.

[49] E. Eid, T. Lacrevaz, S. de Rivaz, C. Bermond, B. Flechet, F. Calmon, C. Gontrand, A. Farcy, L. Cadix, and P. Ancey, “Predictive High Frequency effects of substrate coupling in 3d integrated circuits stacking,” in *IEEE International Conference on 3D System Integration, 2009. 3DIC 2009*, pp. 1–6, Sept. 2009.

[50] M. Brocard, C. Bermond, T. Lacrevaz, A. Farcy, P. Le Maitre, P. Scheer, P. Leduc, S. Cheramy, and B. Flechet, “RF characterization of substrate coupling between TSV and MOS transistors in 3d integrated circuits,” in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–8, Oct. 2013.

[51] S. Lee, G. Kim, J. Kim, T. Song, J. Lee, H. Lee, K. Park, and J. Kim, “Modeling and analysis of die-to-die vertical coupling in 3-D IC,” in *Electronics Packaging Technology Conference, 2009. EPTC ’09. 11th*, pp. 707–711, 2009.

[52] K. Koo, M. Kim, J. Kim, J. Kim, and J. Kim, “Vertical Noise Coupling From On-Chip Switching-Mode Power Supply in a Mixed-Signal Stacked 3-D-IC,” *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 3, no. 3, pp. 476–488, 2013.

[53] Y. Kim, J. Cho, K. Kim, H. Kim, J. Kim, S. Sitaraman, V. Sundaram, and R. Tummala, “Analysis and optimization of a power distribution network in 2.5d

IC with glass interposer," in *3D Systems Integration Conference (3DIC), 2014 International*, pp. 1–4, Dec. 2014.

[54] M. Corbalan, A. Keval, T. Toms, D. Lisk, R. Radojcic, and M. Nowak, "Power and signal integrity challenges in 3d systems," in *2013 50th ACM / EDAC / IEEE Design Automation Conference (DAC)*, pp. 1–4, May 2013.

[55] C. Zhang and G. Sun, "Fabrication cost analysis for 2d, 2.5d, and 3d IC designs," in *3D Systems Integration Conference (3DIC), 2011 IEEE International*, pp. 1–4, Jan. 2012.

[56] D. Velenis, E. Marinissen, and E. Beyne, "Cost effectiveness of 3d integration options," in *3D Systems Integration Conference (3DIC), 2010 IEEE International*, pp. 1–6, Nov. 2010.

[57] J. Lau, "TSV manufacturing yield and hidden costs for 3d IC integration," in *Electronic Components and Technology Conference (ECTC), 2010 Proceedings 60th*, pp. 1031–1042, June 2010.

[58] H.-H. Lee and K. Chakrabarty, "Test Challenges for 3d Integrated Circuits," *IEEE Design Test of Computers*, vol. 26, pp. 26–35, Sept. 2009.

[59] G. Neela and J. Draper, "Challenges in 3dic implementation of a design using current CAD tools," in *2012 IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 478–481, Aug. 2012.

[60] C. Chiang and S. Sinha, "The road to 3d EDA tool readiness," in *Design Automation Conference, 2009. ASP-DAC 2009. Asia and South Pacific*, pp. 429–436, Jan. 2009.

[61] R. Puri and D. Kung, "The Dawn of 22nm Era: Design and CAD Challenges," in *23rd International Conference on VLSI Design, 2010. VLSID '10*, pp. 429–433, Jan. 2010.

[62] D. Milojevic, P. Marchal, E. Marinissen, G. Van der Plas, D. Verkest, and E. Beyne, "Design issues in heterogeneous 3d/2.5d integration," in *Design Automation Conference (ASP-DAC), 2013 18th Asia and South Pacific*, pp. 403–410, Jan. 2013.

[63] K. Banerjee, S. Souri, P. Kapur, and K. Saraswat, "3-D ICs: a novel chip design for improving deep-submicrometer interconnect performance and systems-on-chip integration," *Proceedings of the IEEE*, vol. 89, no. 5, pp. 602–633, 2001.

[64] K. Saraswat, "3-D ICs: Motivation, performance analysis, technology and applications," in *Physical and Failure Analysis of Integrated Circuits (IPFA), 2010 17th IEEE International Symposium on the*, pp. 1–6, 2010.

[65] S. Fujita, K. Abe, K. Nomura, S. Yasuda, and T. Tanamoto, "Perspectives and issues in 3d-IC from designers' point of view," in *IEEE International Symposium on Circuits and Systems, 2009. ISCAS 2009*, pp. 73–76, 2009.

[66] B. Black, D. Nelson, C. Webb, and N. Samra, “3d processing technology and its impact on iA32 microprocessors,” in *IEEE International Conference on Computer Design: VLSI in Computers and Processors, 2004. ICCD 2004. Proceedings*, pp. 316 – 318, Oct. 2004.

[67] W. Davis, E. Oh, A. Sule, and P. Franzon, “Application Exploration for 3-D Integrated Circuits: TCAM, FIFO, and FFT Case Studies,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 17, no. 4, pp. 496–506, 2009.

[68] C. Liu, J.-H. Chen, R. Manohar, and S. Tiwari, “Mapping system-on-chip designs from 2-D to 3-D ICs,” in *IEEE International Symposium on Circuits and Systems, 2005. ISCAS 2005*, pp. 2939–2942 Vol. 3, 2005.

[69] J. Bautista, “Tera-scale computing and interconnect challenges,” in *45th ACM/IEEE Design Automation Conference, 2008. DAC 2008*, pp. 665 –667, June 2008.

[70] J. Rosenfeld and E. Friedman, “On-chip DC-DC converters for three-dimensional ICs,” in *Quality of Electronic Design, 2009. ISQED 2009. Quality Electronic Design*, pp. 759–764, Mar. 2009.

[71] D. Dutoit, C. Bernard, S. Cheramy, F. Clermidy, Y. Thonnart, P. Vivet, C. Freund, V. Guerin, S. Guilhot, S. Lecomte, G. Qualizza, J. Pruvost, Y. Dodo, N. Hotelier, and J. Michailos, “A 0.9 pJ/bit, 12.8 GByte/s WideIO memory interface in a 3d-IC NoC-based MPSoC,” in *2013 Symposium on VLSI Circuits (VLSIC)*, pp. C22–C23, June 2013.

[72] I. Savidis, S. Kose, and E. G. Friedman, “Power Noise in TSV-Based 3-D Integrated Circuits,” *IEEE Journal of Solid-State Circuits*, vol. PP, no. 99, pp. 1 –11, 2012.

[73] P. Jain, T.-H. Kim, J. Keane, and C. Kim, “A multi-story power delivery technique for 3d integrated circuits,” in *2008 ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED)*, pp. 57 –62, Aug. 2008.

[74] M. Healy and S.-K. Lim, “Power-supply-network design in 3d integrated systems,” in *2011 12th International Symposium on Quality Electronic Design (ISQED)*, pp. 1–6, 2011.

[75] W. Davis, A. Sule, and P. Franzon, “An 8192-point fast fourier transform 3d-IC case study,” in *51st Midwest Symposium on Circuits and Systems, 2008. MWSCAS 2008*, pp. 438–441, 2008.

[76] S. Das, A. Chandrakasan, and R. Reif, “Design tools for 3-D integrated circuits,” in *Design Automation Conference, 2003. Proceedings of the ASP-DAC 2003. Asia and South Pacific*, pp. 53–56, Jan. 2003.

[77] D. Fick, R. Dreslinski, B. Giridhar, G. Kim, S. Seo, M. Fojtik, S. Satpathy, Y. Lee, D. Kim, N. Liu, M. Wieckowski, G. Chen, T. Mudge, D. Sylvester, and D. Blaauw, “Centip3de: A 3930dmips/W configurable near-threshold 3d stacked system with

64 ARM Cortex-M3 cores," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, pp. 190 –192, Feb. 2012.

[78] T. Kirihata, "Three dimensional dynamic random access memory," in *2014 4th IEEE International Workshop on Low Temperature Bonding for 3D Integration (LTB-3D)*, pp. 2–2, July 2014.

[79] K.-T. Park, J.-M. Han, D. Kim, S. Nam, K. Choi, M.-S. Kim, P. Kwak, D. Lee, Y.-H. Choi, K.-M. Kang, M.-H. Choi, D.-H. Kwak, H.-w. Park, S.-w. Shim, H.-J. Yoon, D. Kim, S.-w. Park, K. Lee, K. Ko, D.-K. Shim, Y.-L. Ahn, J. Park, J. Ryu, D. Kim, K. Yun, J. Kwon, S. Shin, D. Youn, W.-T. Kim, T. Kim, S.-J. Kim, S. Seo, H.-G. Kim, D.-S. Byeon, H.-J. Yang, M. Kim, M.-S. Kim, J. Yeon, J. Jang, H.-S. Kim, W. Lee, D. Song, S. Lee, K.-H. Kyung, and J.-H. Choi, "Three-dimensional 128gb MLC vertical NAND Flash-memory with 24-WL stacked layers and 50mb/s high-speed programming," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International*, pp. 334–335, Feb. 2014.

[80] J. Jeddeloh and B. Keeth, "Hybrid memory cube new DRAM architecture increases density and performance," in *2012 Symposium on VLSI Technology (VLSIT)*, pp. 87–88, June 2012.

[81] C. Erdmann, D. Lowney, A. Lynam, A. Keady, J. McGrath, E. Cullen, D. Breathnach, D. Keane, P. Lynch, M. De La Torre, R. De La Torre, P. Lim, A. Collins, B. Farley, and L. Madden, "A Heterogeneous 3d-IC consisting of two 28nm FPGA die and 32 reconfigurable high-performance data converters," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International*, pp. 120–121, Feb. 2014.

[82] S. Kim, C. Liu, L. Xue, and S. Tiwari, "Crosstalk reduction in mixed-signal 3-D integrated circuits with interdevice layer ground planes," *IEEE Transactions on Electron Devices*, vol. 52, no. 7, pp. 1459–1467, 2005.

[83] J. Cho, E. Song, K. Yoon, J. S. Pak, J. Kim, W. Lee, T. Song, K. Kim, J. Lee, H. Lee, K. Park, S. Yang, M. Suh, K. Byun, and J. Kim, "Modeling and Analysis of Through-Silicon Via (TSV) Noise Coupling and Suppression Using a Guard Ring," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 1, no. 2, pp. 220–233, 2011.

[84] G. Katti, M. Stucchi, K. de Meyer, and W. Dehaene, "Electrical Modeling and Characterization of Through Silicon via for Three-Dimensional ICs," *IEEE Transactions on Electron Devices*, vol. 57, pp. 256–262, Jan. 2010.

[85] J. Kim, J. S. Pak, J. Cho, E. Song, J. Cho, H. Kim, T. Song, J. Lee, H. Lee, K. Park, S. Yang, M.-S. Suh, K.-Y. Byun, and J. Kim, "High-Frequency Scalable Electrical Model and Analysis of a Through Silicon Via (TSV)," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 1, pp. 181–195, Feb. 2011.

[86] L. Cadix, A. Farcy, C. Bermond, C. Fuchs, P. Leduc, M. Rousseau, M. Assous, A. Valentian, J. Roullard, E. Eid, N. Sillon, B. Flechet, and P. Ancey, "Modelling of Through Silicon Via RF performance and impact on signal transmission in 3d integrated circuits," in *IEEE International Conference on 3D System Integration, 2009. 3DIC 2009*, pp. 1–7, Sept. 2009.

[87] Y. Yi and Y. Zhou, "A novel circuit model for multiple Through Silicon Vias (TSVs) in 3d IC," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–4, Oct. 2013.

[88] J. Kim, J. Cho, and J. Kim, "TSV modeling and noise coupling in 3d IC," in *Electronic System-Integration Technology Conference (ESTC), 2010 3rd*, pp. 1–6, 2010.

[89] A. Engin and S. Narasimhan, "Modeling of Crosstalk in Through Silicon Vias," *IEEE Transactions on Electromagnetic Compatibility*, vol. 55, no. 1, pp. 149–158, 2013.

[90] R. Weerasekera, M. Grange, D. Pamunuwa, H. Tenhunen, and L.-R. Zheng, "Compact modelling of Through-Silicon Vias (TSVs) in three-dimensional (3-D) integrated circuits," in *IEEE International Conference on 3D System Integration, 2009. 3DIC 2009*, pp. 1–8, Sept. 2009.

[91] K. Salah, A. EL-Rouby, H. Ragai, and Y. Ismail, "A closed form expression for TSV-based on-chip spiral inductor," in *2012 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2325–2328, 2012.

[92] Y.-C. Tseng, C.-B. Chang, C.-K. Tang, C.-H. Cheng, Y.-C. Lu, K.-Y. Lin, T.-L. Wu, and R.-B. Wu, "Design considerations for radio frequency 3dics," in *2012 IEEE Electrical Design of Advanced Packaging and Systems Symposium (EDAPS)*, pp. 197–200, 2012.

[93] M. Duplessis, O. Tesson, F. Neuilly, J. Tenailleau, and P. Descamps, "Physical implementation of 3d integrated solenoids within silicon substrate for hybrid IC applications," in *Microwave Conference, 2009. EuMC 2009. European*, pp. 1006–1009, 2009.

[94] K. Liu, R. Frye, M. Hlaing, Y. Lee, H. Kim, G. Kim, and B. Ahn, "Investigation of integrated passive device with through-silicon via," in *Electronic Components and Technology Conference (ECTC), 2012 IEEE 62nd*, pp. 1833–1839, 2012.

[95] U. Tida, R. Yang, C. Zhuo, and Y. Shi, "On the Efficacy of Through-Silicon-Via Inductors," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. PP, no. 99, pp. 1–1, 2014.

[96] T. Ebefors, J. Fredlund, D. Perttu, R. van Dijk, L. Cifola, M. Kaunisto, P. Rantakari, and T. Vaha-Heikkila, "The development and evaluation of RF TSV for 3d IPD applications," in *3D Systems Integration Conference (3DIC), 2013 IEEE International*, pp. 1–8, Oct. 2013.

[97] C. Soens, G. Van der Plas, P. Wambacq, S. Donnay, and M. Kuijk, "Performance degradation of LC-tank VCOs by impact of digital switching noise in lightly doped substrates," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 7, pp. 1472–1481, 2005.

[98] S. Lee, K. Koo, and J. Kim, "Analysis of die-to-die vertical crosstalk between Clock-tree and Voltage Controlled Oscillator in 3-D IC," in *2011 IEEE International Symposium on Electromagnetic Compatibility (EMC)*, pp. 28–32, 2011.

[99] N. Nastos and Y. Papananos, "RF operation of MOSFETs under integrated inductors," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 5, pp. 2106–2117, 2006.

[100] E. Charbon, *Substrate noise: analysis and optimization for IC design*. Boston: Kluwer Academic Publishers, 2001.

[101] C. Soens, G. Van der Plas, M. Badaroglu, P. Wambacq, S. Donnay, Y. Rolain, and M. Kuijk, "Modeling of Substrate Noise Generation, Isolation, and Impact for an LC-VCO and a Digital Modem on a Lightly-Doped Substrate," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 9, pp. 2040–2051, 2006.

[102] H. Lakdawala, M. Schaecher, C. Fu, R. Limaye, J. Duster, Y. Tan, A. Balankutty, E. Alpman, C. Lee, K. Nguyen, H. Lee, A. Ravi, S. Suzuki, B. Carlton, H. Kim, M. Verhelst, S. Pellerano, T. Kim, S. Venkatesan, D. Srivastava, P. Vandervoorn, J. Rizk, C. Jan, S. Ramamurthy, R. Yavatkar, and K. Soumyanath, "A 32 nm SoC With Dual Core ATOM Processor and RF WiFi Transceiver," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 1, pp. 91–103, 2013.

[103] D. M. W. Leenaerts, "Reducing Substrate Bounce in CMOS RF-Circuitry," in *Substrate Noise Coupling in Mixed-Signal ASICs* (S. Donnay and G. Gielen, eds.), pp. 271–287, Springer US, Jan. 2003.

[104] P. van Zeijl, J.-W. Eikenbroek, P.-P. Vervoort, S. Setty, J. Tangenherg, G. Shipton, E. Kooistra, I. Keekstra, D. Belot, K. Visser, E. Bosma, and S. Blaakmeer, "A Bluetooth radio in 0.18um CMOS," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 12, pp. 1679–1687, 2002.

[105] N. Checka, D. Wentzloff, A. Chandrakasan, and R. Reif, "The effect of substrate noise on VCO performance," in *2005 IEEE Radio Frequency integrated Circuits (RFIC) Symposium, 2005. Digest of Papers*, pp. 523–526, 2005.

[106] L. Franca-Neto, P. Pardy, M. P. Ly, R. Rangel, S. Suthar, T. Syed, B. Bloechel, S. Lee, C. Burnett, D. Cho, D. Kau, A. Fazio, and K. Soumyanath, "Enabling high-performance mixed-signal system-on-a-chip (SoC) in high performance logic CMOS technology," in *Symposium on VLSI Circuits Digest of Technical Papers, 2002*, pp. 164–167, 2002.

[107] I. Bashir, R. Staszewski, O. Eliezer, B. Banerjee, and P. Balsara, "A Novel Approach for Mitigation of RF Oscillator Pulling in a Polar Transmitter," *IEEE Journal of Solid-State Circuits*, vol. 46, pp. 403 –415, Feb. 2011.

[108] S. Bronckers, K. Scheir, G. Van der Plas, G. Vandersteen, and Y. Rolain, "A Methodology to Predict the Impact of Substrate Noise in Analog/RF Systems," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 28, no. 11, pp. 1613–1626, 2009.

[109] J. H. J. H.-S. Wu, *Through-substrate interconnects for 3-D integration and RF systems*. Thesis, Massachusetts Institute of Technology, 2006. Thesis (Ph. D.)–Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, February 2007.

[110] N. Khan, S. Alam, and S. Hassoun, "Through-Silicon Via (TSV)-induced noise characterization and noise mitigation using coaxial TSVs," in *IEEE International Conference on 3D System Integration, 2009. 3DIC 2009*, pp. 1–7, Sept. 2009.

[111] M. Beattie and L. Pileggi, "On-chip induction modeling: basics and advanced methods," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 10, no. 6, pp. 712–729, 2002.

[112] J. Borremans, P. Wambacq, M. Kuijk, G. Carchon, and S. Decoutere, "A 400 uW 4.7-to-6.4ghz VCO under an above-IC inductor in 45nm CMOS," in *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*, pp. 536–634, 2008.

[113] C. Yue and S. Wong, "Design strategy of on-chip inductors for highly integrated RF systems," in *Design Automation Conference, 1999. Proceedings. 36th*, pp. 982–987, 1999.

[114] K. O, "Estimation methods for quality factors of inductors fabricated in silicon integrated circuit process technologies," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 8, pp. 1249–1252, 1998.

[115] C. Yue and S. Wong, "On-chip spiral inductors with patterned ground shields for Si-based RF IC's," in *, 1997 Symposium on VLSI Circuits, 1997. Digest of Technical Papers*, pp. 85–86, 1997.

[116] Y. Bontzios, M. Dimopoulos, and A. Hatzopoulos, "Prospects of 3d inductors on through silicon vias processes for 3d ICs," in *2011 IEEE/IFIP 19th International Conference on VLSI and System-on-Chip (VLSI-SoC)*, pp. 90–93, Oct. 2011.

[117] A. Hajimiri and T. Lee, "Design issues in CMOS differential LC oscillators," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 5, pp. 717–724, 1999.

[118] M. Kamon, M. Tsuk, and J. White, "FASTHENRY: a multipole-accelerated 3-D inductance extraction program," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, pp. 1750–1758, Sept. 1994.

[119] A. Hajimiri and T. Lee, "A general theory of phase noise in electrical oscillators," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 2, pp. 179–194, 1998.

[120] L. Fanori, A. Liscidini, and P. Andreani, "A 6.7-to-9.2ghz 55nm CMOS hybrid Class-B/Class-C cellular TX VCO," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, pp. 354–356, Feb. 2012.

[121] B. Razavi, "A study of injection locking and pulling in oscillators," *IEEE Journal of Solid-State Circuits*, vol. 39, pp. 1415 – 1424, Sept. 2004.

[122] R. Gilmore, "System design considerations for next generation wireless mobile devices," in *2013 Symposium on VLSI Technology (VLSIT)*, pp. T8–T13, June 2013.

[123] C. Cox, *An introduction to LTE LTE, LTE-advanced, SAE, and 4G mobile communications*. Chichester, West Sussex, U.K: John Wiley & Sons, 2012.

[124] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA): LTE physical layer general description," Tech. Rep. TS 36.201 v8.3.0, Apr. 2009.

[125] J. Mietzner, R. Schober, L. Lampe, W. Gerstacker, and P. Hoeher, "Multiple-antenna techniques for wireless communications - a comprehensive literature survey," *IEEE Communications Surveys Tutorials*, vol. 11, no. 2, pp. 87 –105, 2009.

[126] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation," Tech. Rep. TS 36.211 v10.5.0, July 2012.

[127] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding," Tech. Rep. TS 36.212 v10.6.0, July 2012.

[128] H. Myung, J. Lim, and D. Goodman, "Peak-To-Average Power Ratio of Single Carrier FDMA Signals with Pulse Shaping," in *2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1 –5, Sept. 2006.

[129] H. G. Myung, J. Lim, and D. J. Goodman, "Single carrier FDMA for uplink wireless transmission," *IEEE Vehicular Technology Magazine*, vol. 1, pp. 30 –38, Sept. 2006.

[130] J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," *Mathematics of Computation*, vol. 19, p. 297, Apr. 1965.

[131] E. Dubois and A. Venetsanopoulos, "A new algorithm for the radix-3 FFT," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 3, pp. 222–225, 1978.

[132] J. Lofgren and P. Nilsson, "On hardware implementation of radix 3 and radix 5 FFT kernels for LTE systems," in *NORCHIP, 2011*, pp. 1 –4, Nov. 2011.

[133] A. Ferizi, B. Hoeher, M. Jung, G. Fischer, and A. Koelpin, "Design and implementation of a fixed-point radix-4 FFT optimized for local positioning in wireless sensor networks," in *2012 9th International Multi-Conference on Systems, Signals and Devices (SSD)*, pp. 1 –4, Mar. 2012.

[134] J. Kwong and A. Chandrakasan, "An Energy-Efficient Biomedical Signal Processing Platform," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 7, pp. 1742–1753, 2011.

[135] D. Revanna, O. Anjum, M. Cucchi, R. Airoldi, and J. Nurmi, "A scalable FFT processor architecture for OFDM based communication systems," in *2013 International Conference on Embedded Computer Systems: Architectures, Modeling, and Simulation (SAMOS XIII)*, pp. 19–27, 2013.

[136] B. Baas, "A low-power, high-performance, 1024-point FFT processor," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 380 –387, Mar. 1999.

[137] C.-F. Hsiao, Y. Chen, and C.-Y. Lee, "A Generalized Mixed-Radix Algorithm for Memory-Based FFT Processors," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 57, pp. 26–30, Jan. 2010.

[138] T. Patyk, D. Guevorkian, T. Pitkanen, P. Jaaskelainen, and J. Takala, "Low-power application-specific FFT processor for LTE applications," in *2013 International Conference on Embedded Computer Systems: Architectures, Modeling, and Simulation (SAMOS XIII)*, pp. 28–32, 2013.

[139] J. Chen, J. Hu, S. Lee, and G. Sobelman, "Hardware Efficient Mixed Radix-25/16/9 FFT for LTE Systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, pp. 221–229, Feb. 2015.

[140] S. He and M. Torkelson, "Designing pipeline FFT processor for OFDM (de)modulation," in *1998 URSI International Symposium on Signals, Systems, and Electronics, 1998. ISSSE 98*, pp. 257 –262, Oct. 1998.

[141] H. Groginsky and G. Works, "A Pipeline Fast Fourier Transform," *IEEE Transactions on Computers*, vol. C-19, pp. 1015 – 1019, Nov. 1970.

[142] A. Despain, "Fourier Transform Computers Using CORDIC Iterations," *IEEE Transactions on Computers*, vol. C-23, pp. 993–1001, Oct. 1974.

[143] E. Swartzlander, W. Young, and S. Joseph, "A radix 4 delay commutator for fast Fourier transform processor implementation," *IEEE Journal of Solid-State Circuits*, vol. 19, pp. 702–709, Oct. 1984.

[144] K.-J. Yang, S.-H. Tsai, and G. Chuang, "MDC FFT/IFFT Processor With Variable Length for MIMO-OFDM Systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 4, pp. 720–731, 2013.

[145] G. Bi and E. Jones, “A pipelined FFT processor for word-sequential data,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, pp. 1982–1985, Dec. 1989.

[146] S. He and M. Torkelson, “A new approach to pipeline FFT processor,” in *Parallel Processing Symposium, 1996., Proceedings of IPPS ’96, The 10th International*, pp. 766 –770, Apr. 1996.

[147] E. Wold and A. Despain, “Pipeline and Parallel-Pipeline FFT Processors for VLSI Implementations,” *IEEE Transactions on Computers*, vol. C-33, pp. 414–426, May 1984.

[148] G. Liu and Q. Feng, “ASIC design of low-power reconfigurable FFT processor,” in *7th International Conference on ASIC, 2007. ASICON ’07*, pp. 44 –47, Oct. 2007.

[149] T. Adiono and R. Mareta, “Low latency parallel-pipelined configurable FFT-IFFT 128/256/512/1024/2048 for LTE,” in *2012 4th International Conference on Intelligent and Advanced Systems (ICIAS)*, vol. 2, pp. 768 –773, June 2012.

[150] A. Chandrakasan, S. Sheng, and R. Brodersen, “Low-power CMOS digital design,” *IEEE Journal of Solid-State Circuits*, vol. 27, pp. 473 –484, Apr. 1992.

[151] R. Gonzalez, B. Gordon, and M. Horowitz, “Supply and threshold voltage scaling for low power CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 1210 –1216, Aug. 1997.

[152] A. Wang, A. Chandrakasan, and S. Kosonocky, “Optimal supply and threshold scaling for subthreshold CMOS circuits,” in *IEEE Computer Society Annual Symposium on VLSI, 2002. Proceedings*, pp. 5 –9, 2002.

[153] Y. Chen, Y.-W. Lin, Y.-C. Tsao, and C.-Y. Lee, “A 2.4-Gsample/s DVFS FFT Processor for MIMO OFDM Communication Systems,” *IEEE Journal of Solid-State Circuits*, vol. 43, pp. 1260 –1273, May 2008.

[154] I.-G. Jang, Z.-Y. Piao, Z.-H. Dong, J.-G. Chung, and K.-Y. Lee, “Low-power FFT design for NC-OFDM in cognitive radio systems,” in *2011 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2449 –2452, May 2011.

[155] I. Koren, *Computer arithmetic algorithms*. Natick, MA : A K Peters, c2002., 2002.

[156] D. Jeon, M. Seok, C. Chakrabarti, D. Blaauw, and D. Sylvester, “A Super-Pipelined Energy Efficient Subthreshold 240 MS/s FFT Core in 65 nm CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 47, pp. 23 –34, Jan. 2012.

[157] C.-H. Yang, T.-H. Yu, and D. Markovic, “Power and Area Minimization of Reconfigurable FFT Processors: A 3gpp-LTE Example,” *IEEE Journal of Solid-State Circuits*, vol. 47, pp. 757 –768, Mar. 2012.

[158] N. H. Cuong, N. T. Lam, and N. D. Minh, “Multiplier-less based architecture for variable-length FFT hardware implementation,” in *2012 Fourth International Conference on Communications and Electronics (ICCE)*, pp. 489 –494, Aug. 2012.

[159] T.-D. Chiueh and I.-W. Lai., *Baseband receiver design for wireless MIMO-OFDM communications*. Hoboken, N.J: Wiley, 2nd ed ed., 2012.

[160] J. E. Volder, “The CORDIC Trigonometric Computing Technique,” *IRE Transactions on Electronic Computers*, vol. EC-8, pp. 330 –334, Sept. 1959.

[161] C.-S. Wu and A.-Y. Wu, “Modified vector rotational CORDIC (MVR-CORDIC) algorithm and its application to FFT,” in *The 2000 IEEE International Symposium on Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva*, vol. 4, pp. 529 –532 vol.4, 2000.

[162] M. Garrido and J. Grajal, “Efficient Memoryless Cordic for FFT Computation,” in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. ICASSP 2007*, vol. 2, pp. II–113 –II–116, Apr. 2007.

[163] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing (3rd Edition)*. Prentice Hall, 3 ed., Aug. 2009.

[164] H. Sorensen and C. Burrus, “Efficient computation of the DFT with only a subset of input or output points,” *IEEE Transactions on Signal Processing*, vol. 41, pp. 1184 –1200, Mar. 1993.

[165] S.-Y. Peng, K.-T. Shr, C.-M. Chen, and Y.-H. Huang, “Energy-efficient 128 -2048/1536-point FFT processor with resource block mapping for 3gpp-LTE system,” in *2010 International Conference on Green Circuits and Systems (ICGCS)*, pp. 14 –17, June 2010.

[166] B. Zhang, L. Liang, and X. Wang, “A New Level Shifter with Low Power in Multi-Voltage System,” in *8th International Conference on Solid-State and Integrated Circuit Technology, 2006. ICSICT ’06*, pp. 1857–1859, 2006.

[167] B. Razavi, *Principles of data conversion system design*. New York : IEEE Press, c1995., 1995.

[168] W.-H. Tseng and P.-C. Chiu, “A 960ms/s DAC with 80db SFDR in 20nm CMOS for multi-mode baseband wireless transmitter,” in *2014 Symposium on VLSI Circuits Digest of Technical Papers*, pp. 1–2, June 2014.

[169] B. Razavi, *RF microelectronics*. Upper Saddle River, N.J. : Pearson Education, 2012., 2012.

[170] S. Luschas, R. Schreier, and H.-S. Lee, “Radio frequency digital-to-analog converter,” *IEEE Journal of Solid-State Circuits*, vol. 39, no. 9, pp. 1462–1467, 2004.

[171] G. Engel, D. Fague, and A. Toledano, “RF digital-to-analog converters enable direct synthesis of communications signals,” *IEEE Communications Magazine*, vol. 50, no. 10, pp. 108–116, 2012.

[172] E. Bechthum, G. Radulov, J. Briaire, G. Geelen, and A. van Roermund, “Systematic analysis of the impact of mixing locality on Mixing-DAC linearity for multicarrier GSM,” in *2012 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 241–244, 2012.

[173] B. Razavi, *Design of analog CMOS integrated circuits*. Boston, MA : McGraw-Hill, c2001., 2001.

[174] A. Lu, D. Xie, Z. Shi, and W. Ryu, “Electrical and thermal modelling of QFN packages,” in *Electronics Packaging Technology Conference, 2000. (EPTC 2000). Proceedings of 3rd*, pp. 352–356, 2000.

[175] J. Savoj, K.-H. Hsieh, F.-T. An, J. Gong, J. Im, X. Jiang, A. Jose, V. Kireev, S.-W. Lim, A. Roldan, D. Turker, P. Upadhyaya, D. Wu, and K. Chang, “A Low-Power 0.5-6.6 Gb/s Wireline Transceiver Embedded in Low-Cost 28 nm FPGAs,” *IEEE Journal of Solid-State Circuits*, vol. 48, pp. 2582–2594, Nov. 2013.

[176] Y. Frans, D. Carey, M. Erett, H. Amir-Aslanzadeh, W. Fang, D. Turker, A. Jose, A. Bekele, J. Im, P. Upadhyaya, Z. Wu, K. Hsieh, J. Savoj, and K. Chang, “A 0.5-16.3 Gb/s Fully Adaptive Flexible-Reach Transceiver for FPGA in 20 nm CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 50, pp. 1932–1944, Aug. 2015.

[177] R. Navid, E.-H. Chen, M. Hossain, B. Leibowitz, J. Ren, C.-H. Chou, B. Daly, M. Aleksic, B. Su, S. Li, M. Shirasgaonkar, F. Heaton, J. Zerbe, and J. Eble, “A 40 Gb/s Serial Link Transceiver in 28 nm CMOS Technology,” *IEEE Journal of Solid-State Circuits*, vol. 50, pp. 814–827, Apr. 2015.

[178] C. Erdmann, D. Lowney, A. Lynam, A. Keady, J. McGrath, E. Cullen, D. Breathnach, D. Keane, P. Lynch, M. De La Torre, R. De La Torre, P. Lim, A. Collins, B. Farley, and L. Madden, “A Heterogeneous 3d-IC Consisting of Two 28 nm FPGA Die and 32 Reconfigurable High-Performance Data Converters,” *IEEE Journal of Solid-State Circuits*, vol. 50, pp. 258–269, Jan. 2015.

[179] S.-H. Lee, S.-K. Lee, B. Kim, H.-J. Park, and J.-Y. Sim, “Current-Mode Transceiver for Silicon Interposer Channel,” *IEEE Journal of Solid-State Circuits*, vol. 49, pp. 2044–2053, Sept. 2014.

[180] J. Poulton, W. Dally, X. Chen, J. Eyles, T. Greer, S. Tell, J. Wilson, and C. Gray, “A 0.54 pJ/b 20 Gb/s Ground-Referenced Single-Ended Short-Reach Serial Link in 28 nm CMOS for Advanced Packaging Applications,” *IEEE Journal of Solid-State Circuits*, vol. 48, pp. 3206–3218, Dec. 2013.