

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Time-Domain Ultra-Wideband Synthetic Imager in Silicon

Permalink

<https://escholarship.org/uc/item/19f071g6>

Author

Arbabian, Mohammad Amin

Publication Date

2011

Peer reviewed|Thesis/dissertation

Time-Domain Ultra-Wideband Synthetic Imager in Silicon

by

Mohammad Amin Arbabian

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering—Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ali M. Niknejad, Chair
Professor Jan M. Rabaey
Professor Eli Yablonovitch
Professor Steven Conolly

Fall 2011

Time-Domain Ultra-Wideband Synthetic Imager in Silicon

Copyright 2011
by
Mohammad Amin Arbabian

Abstract

Time-Domain Ultra-Wideband Synthetic Imager in Silicon

by

Mohammad Amin Arbabian

Doctor of Philosophy in Engineering–Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Ali M. Niknejad, Chair

Low-cost and portable medical devices will play a more significant role in wellness, healthcare and medicine. While consumer electronics have become ubiquitous and inexpensive, medical devices, by contrast, are still primarily found only in hospitals. There is a great potential benefit in using techniques developed in the consumer electronic industry and applying them to the healthcare market. To do this, substantial innovation is required to develop new sensors and devices that are fundamentally less invasive and use profoundly different physical phenomena to address medical applications.

This research aims at designing a non-invasive, low-cost, and portable imaging device for cancer screening. Detection in early stages has proven to be essential for reducing the mortality rate in cancer. This requires pursuit of modalities that could be widespread and are safe to be used for more frequent screening. This research uses the available contrast in microwave frequencies to detect abnormalities.

Conceptualization, architectural and system-level design, and finally implementation of the system called TUSI, Time-Domain Ultra-Wideband Synthetic Imager, are addressed. Using an array of closely controlled radiating silicon chips, acting as transceivers in microwave /mm-wave frequencies, this device transmits short “beam-steered” pulses and picks up reflections from tissue abnormalities (e.g. cancerous tissue). By processing the data from multiple transceivers, a larger aperture is synthesized. In essence, this imager probes the “electrical” properties of the tissue. Various challenges related to generating, controlling, transmitting, and detecting these coherent ultra-short pulses are examined and new solutions proposed. A pixel-scalable integrated transceiver consisting of elements from antenna-to-antenna is designed and implemented in a SiGe BiCMOS process.

Contents

List of Figures	iv
List of Tables	viii
1 Introduction	1
1.1 Silicon-Based Medical Imager	1
1.2 Low-Cost Breast Cancer Screening	2
1.3 Skin Cancer Detection	3
1.4 Other Applications	4
1.5 Organization of the Dissertation	4
2 Background and System Design	5
2.1 Radar Imager	5
2.2 Contrast	6
2.3 Resolution Basics	8
2.3.1 Depth Resolution	8
2.3.2 Lateral Resolution	9
2.4 Pulsed-Based Antenna Arrays	10
2.5 Signal-to-Noise Ratio	11
2.6 Resolution Limits	12
2.7 TUSI Imager	15
2.8 Undesirable Coupling of SNR with Resolution	16
2.9 Limits of Integration for Closing the SNR Gap	18
2.9.1 Stability of the Imaging Setup	19
2.9.2 Limitations on Increasing the PRF	19
2.9.3 Accurate Delay Generation	19
2.10 Coherent Radar Phase Information	22
2.11 Spatial Coding to Suppress Spatial Leakage	24
2.12 TUSI in the Large-Scale	24
2.12.1 TUSI for Intelligent Surfaces	25

3	Antenna Design	28
3.1	Integrated Antennas	28
3.1.1	Slot and Dipole Antennas	30
3.2	Antentronics	34
3.2.1	Folded Slot Dipole Antentronic Structure	34
3.2.2	Dual-Loop Antentronic Structure	37
3.3	Wideband Antennas	43
4	TUSI Transmitters	45
4.1	Process Overview	45
4.2	Transmitter Architecture	47
4.2.1	Pulse Modulating the RF Carrier	47
4.2.2	Hybrid Switching	47
4.3	High-Speed Timing Circuitry	50
4.3.1	Mode Selection Circuitry	51
4.3.2	Pulse Generator	54
4.3.3	Programmable Delay	55
4.3.4	Impact of Jitter in Pulse Generation	57
4.3.5	High-Speed Output Buffer	58
4.4	Quadrature Voltage Controlled Oscillator (QVCO)	62
4.5	Switched Power Amplifier	65
4.6	Voltage Switching Transmitter	68
4.6.1	Voltage Switching Antentronic Network	68
4.6.2	Experimental Results	69
4.7	Current Switching Transmitter	76
4.7.1	Power Tuning Capability	76
4.7.2	Current Switching Scheme	77
4.7.3	Experimental Results	77
5	Receiver	84
5.1	94GHz Receiver Overview	84
5.1.1	Distortion	84
5.2	Wideband Amplification	86
5.2.1	Distributed Amplifiers	86
5.2.2	Passive Design	88
5.2.3	Active Element Design	94
5.2.4	Distributed Amplifier with Internal Feedback	97
5.2.5	Tapered-Cascaded Multi-Stage Distributed Amplifier	112
5.2.6	SiGe BiCMOS Distributed Amplifier for the Imager Front-End	120
5.3	Wideband Conversion to Differential	127
5.4	Mixer Design	130

5.4.1	Overview of Challenges	130
5.4.2	Mixer Core Design	132
5.4.3	LO Buffer	133
5.5	Baseband Amplification	134
6	Integrated 94GHz Radar Transceiver	137
6.1	System Overview	137
6.1.1	Receiver Architecture	137
6.1.2	Transmitter Architecture	139
6.1.3	Frequency Generation and Distribution	140
6.1.4	Pulse Position and Width Programming	142
6.2	End-to-End Measurement Results	144
7	Conclusion	152
	Bibliography	154

List of Figures

2.1	TUSI imager concept.	7
2.2	Dielectric-properties data for normal and malignant breast tissue (from [1]).	8
2.3	Depth resolution in pulsed radar.	9
2.4	Frequency-domain array with grating lobes.	10
2.5	Comparison of pulsed and frequency-domain arrays.	11
2.6	Penetration depth and resolution.	12
2.7	Resolution reduction effect from large variations in amplitude.	13
2.8	Effect of pulse integrity on performance	13
2.9	Received pulsed edge with noise.	14
2.10	TUSI array imager on flexible substrate.	15
2.11	Proposed array architecture and system block diagram	17
2.12	Delay-Locked Loop architecture for generating fine time steps.	17
2.13	System simulations of the transmit/receive chain.	18
2.14	Range and position ambiguity from high PRF.	20
2.15	Linear antenna array model.	20
2.16	Effect of delay step quantization.	23
2.17	Spatial sectorization in the TUSI Imager.	24
2.18	Large-scale TUSI array for intelligent surfaces.	26
3.1	Elementary slot and dipole antennas on silicon substrate.	29
3.2	Electric and magnetic sources close to a ground plane.	30
3.3	Various forms of integrated antennas on dielectric substrates.	31
3.4	Directions of propagation of surface-waves for a dipole antenna.	32
3.5	Direction of reflected E-fields in the dipole on grounded substrate.	32
3.6	Antentronic Structure with folded slot dipole antenna.	35
3.7	Non-radiating mode current distribution.	35
3.8	Simulated input impedance of the antenna in transmit mode.	36
3.9	Loop antenna together with radiation pattern in cases of small loop	37
3.10	Resonant loop antenna equivalent to two half-wavelength dipoles.	38
3.11	Loop antenna on ground plane.	38
3.12	Radiation efficiency, broadside gain and input resistance of loop antenna.	39

3.13	Simulated E and H-plane pattern of the loop antenna.	39
3.14	Slot antenna on a dielectric substrate.	40
3.15	Dual-loop antenetric structure.	41
3.16	Dual loop antenna under non-radiating drive conditions.	42
3.17	Some wideband antenna designs.	43
3.18	Tapered loop antenna on silicon substrate.	44
3.19	Single-element tapered loop antenna pattern.	44
4.1	Collector-Emitter breakdown voltage of a $5\mu m$ device.	46
4.2	Conceptual switching options for pulse modulating the carrier.	48
4.3	Pulse generation options.	49
4.4	Input equivalent circuit of the pulse driver load.	50
4.5	System block diagram of the TUSI transmitter.	51
4.6	Schematic of ECL logic gates used in design of High-Speed Timing Circuitry.	52
4.7	Operation modes of TUSI transmitter.	53
4.8	Block diagram of Mode Selection circuitry (programmed for Mode 1).	53
4.9	Simplified schematic of pulse generator.	54
4.10	Detailed block diagram of pulse generator.	54
4.11	Plot of simulated pulse widths at the output of PA driver.	56
4.12	Transient simulation comparing pulse generator output to PA driver output.	56
4.13	Illustration of timing uncertainty in hybrid mode.	57
4.14	Simulated period jitter (rms) of variable delay buffer.	58
4.15	Circuit schematic of the high-speed buffer.	59
4.16	Simplified circuit model for buffer circuit including the trace.	60
4.17	Response of a capacitively terminated transmission line.	61
4.18	Design curves for pulse driver circuit parameters.	63
4.19	Simplified schematic of the Colpitts oscillator.	64
4.20	Complete schematic of QVCO.	64
4.21	Schematic of the output buffer.	65
4.22	Measured VCO tuning range as well as the measured spectrum at 74 GHz.	66
4.23	Two-stage transformer-coupled power amplifier.	67
4.24	Stability factor of the PA under different presented loads to the output.	68
4.25	Chip micrograph of the TUSI VS transmitter.	69
4.26	Transmitter measurement setup.	70
4.27	Ideal unlocked RF pulse response in BW limited regime.	71
4.28	Measured pulse in mode 6 with (a) 50ps/div and (b) 10ps/div.	71
4.29	Time domain measurements of mode 3 (53ps) and mode5 (35ps)	72
4.30	Spectrum measurements of positive pulse in mode 3 (hybrid).	72
4.31	Time domain measurements from Mode 1 (hybrid).	73
4.32	Hybrid mode measurements/	73
4.33	Measurements with a metallic reflector and infinite setting on oscilloscope.	74

4.34	System level block diagram of the current-switched transmitter.	76
4.35	Circuit schematics of the current switching scheme transmitter.	78
4.36	TUSI CS chip micrograph.	79
4.37	Measured PSD of various pulses.	80
4.38	Time-domain measurements for different settings.	81
4.39	Bistatic reflection measurements.	82
5.1	Block diagram of the direct-conversion TUSI receiver chain.	84
5.2	Block diagram of the external down-converter.	85
5.3	Comparing a single stage common-source to a distributed amplifier.	87
5.4	Line bandwidth limitation due to the initial Z_0 of transmission line.	90
5.5	Equivalent loss tangent of CPW transmission lines.	90
5.6	Conceptual layout of shielded elevated line with current flow.	90
5.7	HFSS simulation of elevated line characteristics.	91
5.8	HFSS simulations of loss for various elevations with respect to frequency. . .	92
5.9	Effective dielectric constant of the transmission line.	92
5.10	E-CPW line measurements.	94
5.11	Measurements of a shielded elevated CPW with M1 and poly as filaments. . .	94
5.12	Schematic diagram of cascode device with important parasitic elements. . . .	95
5.13	Simulated DA gain under varying device size and number of devices.	96
5.14	Simulated DA gain vs. number of devices for various number of fingers. . . .	96
5.15	Proposed DA architecture for improved gain-bandwidth product.	97
5.16	MATLAB simulations of reverse gain in hypothetical 3 and 10 stage DAs. . . .	99
5.17	Simulations for forward and reverse gain of 8-stage DA.	99
5.18	DA forward and reverse gain simulations.	100
5.19	Calculated normalized gain for a conceptual feedback DA.	102
5.20	Theoretical NF of DA in terms of frequency for different n	104
5.21	Simulated gain and NF of a 16 stage DA in level 1 MOS models.	107
5.22	Simulated gain and NF of a FBDA with the same number of stages.	108
5.23	Measured s-parameters of the FBDA.	109
5.24	Measured noise and P_{1dB} of the FBDA.	110
5.25	Chip micrograph of the FBDA.	110
5.26	T-section of a synthesized transmission line.	114
5.27	Schematic diagram of a tapered distributed amplifier.	114
5.28	Simulated gain with different tapering strategies.	115
5.29	Schematics of the T-CMSDA.	116
5.30	S-parameter simulation and measurements of the T-CMSDA.	117
5.31	Chip micrograph of the of the T-CMSDA.	117
5.32	Output compression point measurements of the T-CMSDA.	118
5.33	Measured group delay of the amplifier in frequency band of interest.	118
5.34	Single gain stage equivalent circuit.	122

5.35	Common emitter stage with degeneration and series capacitor.	122
5.36	DA cascode gain element.	124
5.37	Input loading from the DA gain element.	125
5.38	Cascaded choke architecture.	125
5.39	Chip micrograph of the SiGe distributed amplifier.	126
5.40	S-parameter measurements of the SiGe DA.	126
5.41	Active balun circuit schematic.	127
5.42	Circuit schematic of the active balun with degeneration and base resistances.	128
5.43	Phase response of the active balun.	129
5.44	Magnitude response of the active balun.	129
5.45	CMRR of the core active balun circuit.	129
5.46	Schematic of the quadrature mixer.	130
5.47	Mixer core with input offset voltage.	132
5.48	Mixer gain	133
5.49	The schematic of the single-ended to differential LO buffer.	134
5.50	LO buffer output power.	135
5.51	Input impedance at the input of the mixer.	135
5.52	The schematic of the baseband buffer.	136
5.53	Baseband buffer voltage gain and group delay simulations.	136
6.1	Block-level schematics of a pulsed-based radar transceiver.	138
6.2	Chip micrograph of the radar transceiver.	138
6.3	Layout of the baseband section.	140
6.4	Two-stage transformer-coupled power amplifier.	141
6.5	PA pulse driver schematics.	141
6.6	PLL architecture implemented in TUSI transceiver.	142
6.7	Conceptual schematics of delay-locked loop (DLL).	143
6.8	Picture of the RF board in the measurement setup.	144
6.9	Measurement setup for testing the radar transceiver.	145
6.10	Frequency domain measurements of the transceiver.	146
6.11	Time-domain measurements with a non-locked PLL.	147
6.12	PLL locking range based on various VCO frequency bands	148
6.13	Measurements of TRX pulse output waveform.	148
6.14	Pulse delay profile in coarse setting.	150
6.15	Measured pulse delay (referenced to initial setting).	151
6.16	Intercepted pulse from the transmitter.	151

List of Tables

4.1	Measurements Summary for TUSI VS	75
4.2	Measurements Summary for TUSI CS	83
5.1	Performance comparison of the FBDA to the prior state of the art.	111
5.2	Comparison table for the tapered DA	119
5.3	Transmission line parameters.	121
6.1	Summary of radar transceiver performance.	149

Acknowledgments

It is difficult for me to believe that my time at Berkeley has come to an end. My Cal experience has been an amazing journey- one that has shaped my life in many different ways. I was never planning on spending my whole life as a Cal grad student but this has been enjoyable and fulfilling and it is definitely strange for me not to be a Cal student anymore! The past few years would not have been what they were without the help and support of many people. Thank you to all!

First, I would like to thank my partner both in school and in life, Sahar. For the past six and a half years that we have been married (and nearly a decade of being together) she has enriched my life in many many different ways. We have grown up together and experienced the toughest transitions and periods side by side. She has always been a constant source of inspiration and encouragement. Words cannot express how grateful I am. My research and my thesis would not have been possible without her support at different levels; emotionally and intellectually. In research, she is the first person to hear my (at times crazy and dumb) ideas and that has been an amazing experience. Thank you for everything!

I could not have asked for a better adviser in supporting my ambitions and goals. Ali has been a great friend, mentor and teacher as well as academic adviser and has provided constant and generous support. I think he has a close to ideal balance in guiding graduate students. Initially, he provided substantial help to get me started. After that, he not only played a big role in technical guidance but also helped me learn how to learn. He has never “over-advised” me if a term like that exists. I have been given freedom to pursue different paths and I thank him for trusting me. I have learned many things from Ali both in science and engineering as well as the academic path in general. I also want to thank him for his excellent teaching of EE142, EE240, and EE242. He is knowledgeable, thoughtful, extremely intelligent and original and at the same time great-hearted, high-minded and considerate. It is not very common to find these characteristics in one person.

I would like to thank Professor Eli Yablonovitch for being on my qualifying exam and dissertation committees and also for his advice and support on many occasions during the past few years. I have always enjoyed our discussions and have learned a great deal from them. His unique and powerful way of approaching problems amazed me during classes and research discussions. Taking Eli’s EE290B and EE236 courses were life-changing experiences. These were among the best courses I have ever taken and have had lasting effects on my education.

I would also like to thank Jan Rabaey and Steven Conolly for being on my PhD qualifying exam and dissertation committees. I have learned a lot from discussing various technologies and future wireless trends with Jan. He has been extremely supportive of my research. I also thank him for developing a collaborative and positive research environment at BWRC. Brainstorm meetings with Steve have been very interesting and intellectually fulfilling. His insight and experience in biomedical imaging and MRI have been a great source of information. I would also like to thank him for his excellent teaching of the medical imaging and

MRI courses at Berkeley.

My warmest gratitude goes to Haideh Khorramabadi, a great teacher and mentor. She has supported us in many different ways and has always been there for help and advice. She is one of the kindest and most respectable people I have come to know here at Berkeley. Thank you Haideh!

An important part of my Berkeley experience was the magnificent courses that are offered across campus. In this regard, in addition to the people mentioned previously, my sincere gratitude goes to professors Khorramabadi, Markovic, Rabaey, Bokor, Bahai, Carmena and Kuroda in the EECS department, professors Ganor and Cummings in the Physics department, professors Wu and Gronsky in the MSE department and professor S. Lee in the Bioengineering department.

My gratitude goes to all the people who helped me in this project. I owe thanks to Bagher (Ali) Afshar for his help in the design of power amplifiers. I would also like to thank the “Junior” graduate students Jun-Chau Chien, Steven Callender, and Shinwon Kang as well as the undergraduate students (Alex Pai, Alan Wu, Yasin Alipour) who helped me in various designs and were all at some point part of the “TUSI team”. It was a privilege working with such smart and talented students.

Words can hardly do justice in showing the extent of my gratitude to many friends that I have been blessed with knowing at Berkeley or before. My warmest thanks go to Ashkan Borna, Ehsan Adabi, Ali Afshar and Arash Parsa for great discussions and being awesome friends at BWRC and beyond! I learned a great deal from Ehsan and Ali when I joined BWRC. My first tapeout would not have gone anywhere without their help. I also enjoyed the random but interesting discussions with Omar Bakr. One day our thoughts might actually converge to something great! I would also like to thank other members of Ali’s group: Mounir Bohsali, Babak Heydari, Cristian Marcu, Debopriyo Chowdhury, Mohan Dunga, Wei-Hung Chen, Stefano Dal Toso, Peter Haldi, Patrick Reynaert, and Jiashu Chen. Many thanks to Matthew Spencer and Siva Thyagarajan for being outstanding TAs for the EE142 course.

I am also thankful to many BWRC students and friends including Simone Gambini, Louis Alarcon, Jesse Richmond, David Chen, Michael Mark, Tsung-Te Liu, Vahid Majidzadeh, Rikky Muller, Mervin John, Lingkai Kong, Chintan Thakkar, Antoine Frappe, and many others. I owe many thanks to BWRC and EECS faculty and staff including Tom Budinger, Elad Alon, Gary Kelson, Tom Boot, Sue Mellers, Brian Richards, Kevin Zimmerman, Brad Krebs, Bira Coelho, Mary Byrnes, Ruth Gjerde, Dana Jantz, and Patrick Hernan. My special thanks goes to my older friends Ashkan, Alireza, Hooman, Kia, Ehsan (Roosta), and Aria (the great!). It is amazing to see how we are all slowly converging to the same geographical location after so long.

I would like to thank Ali Tassoudji, Bo Sun, and Jorge Garcia, my mentors from Qualcomm CRD during the fellowship and the internship. I learned a great deal from my discussions with Ali on different antenna issues. Bo was just a wealth of knowledge and experience in every aspect of IC design. I also enjoyed my other internships at Tagarray Inc.

and would like to thank Farokh Eskafi and Kouros Pahlavan for the exciting opportunities.

I would like to acknowledge and thank professors Sharif-Bakhtiar, Bastani and Jahanbagloo from my undergraduate at Sharif University. They helped shape my understanding of engineering as well as what it would take to continue down the academic path to PhD. I learned much of what I know of electronic and communication systems from them and I am very grateful for their great advice and teaching.

My greatest thanks and gratitude go to my parents Homayoun and Nasrin and also my brothers Iman and Alireza. This work is the result of their endless love and constant support. They must have seen something in me which I am yet to find myself!

Chapter 1

Introduction

1.1 Silicon-Based Medical Imager

Scaling has allowed silicon technology to enter exciting areas of research such as high-speed data communications, radar and imaging [2, 3, 4, 5, 6, 7, 8]. Due to its higher yield and reliability, silicon has outperformed III-V semiconductors in many millimeter-wave applications and is establishing itself as a market standard for high performance reliable systems. Moore's law is pushing silicon towards yet another application: bio-medical devices. Starting from 1970-80's, silicon technology has successfully transformed different areas from measurement instrumentation to personal computers, communication devices and more recently consumer electronics and ubiquitous computation/communication devices. The next big impact is thought to be in the field of bio-medical devices.

Information technology and medical devices will play a more significant role in wellness, healthcare and medicine. While consumer electronics have become ubiquitous and inexpensive, medical devices, by contrast, are still primarily found only in hospitals. There is a great potential benefit in using techniques developed in the consumer electronic industry and applying them to the healthcare market. To do this, substantial innovation is required to develop new sensors and devices that are fundamentally less invasive and use profoundly different physical phenomena to address medical applications. These devices will provide better service to a rapidly aging population at a fraction of the cost of today's technologies.

Some of the efforts are concentrated on designing new instruments and functionalities that have been made possible largely due to the effectiveness and miniaturization of electronics (e.g. brain-machine interfaces and large array neural-sensors). There is also an effort to tackle more conventional medical problems. Imaging is one such problem that can greatly benefit from cost-reduction and/or miniaturization. Today, imaging devices are largely confined to health-centers and some only in large hospitals. The dynamics of patient health-provider (doctors or emergency care units) interaction will radically change if these diagnostic tools are available at the local clinics rather than at long distances. Ultimately,

some of these tools will be available at homes and much closer to the patients.

The path to better utilization of silicon capabilities in medical imaging may well take us towards less conventional sensing and detection schemes. Devices and modalities that have the potential to benefit the most from the silicon advances are to be examined. Here, we focus on using the dielectric properties of tissue for imaging. Although with great potential, electrical properties and polarization characteristics in the microwave/mm-wave spectrum have not yet received much attention for diagnostic and screening applications. There are many specific cases where electrical properties can greatly enhance image quality or provide alternative contrast to conventional techniques. Electromagnetic signatures have been used for medical imaging applications [9] [10] [11]. Breast cancer screening technology is an example that has been pursued by various researchers in the past one or two decades [1] [10] [12]. One of the main limitations with this modality has been obtaining adequate signal to noise levels as well as sufficient cost/size reduction to be appealing compared to conventional modalities. These are areas where there is a great potential improvement to be made in using the techniques developed in consumer electronics and by leveraging Moore's law.

The goal of a dielectric imager is to detect boundaries with different electrical properties. This could be achieved by various signaling schemes, such as frequency chirp methods or pulse based techniques [13] [14] [7]. The focus of this work is on Time-Domain Ultra-wideband Synthetic Imaging (TUSI) where short pulses are transmitted and the reflections are recorded from multiple sites. The final imaging device is portable and battery operated and can be used to address a variety of applications.

We aim to realize an RF to millimeter-wave fully integrated radar transceiver with close to 25 ps modulated pulse in silicon technology [15] [16] [11]. To obtain short pulses, a new hybrid technique is employed in the transmitter that allows for independent control of start/stop time of the pulse in addition to potentially enabling finer tuning of pulse width without compromising the pulse amplitude. In addition, the concept of *antentronics* is introduced, where integrated circuits and antennas are intelligently combined to enhance desired antenna performance.

Extremely wideband signal amplification is another challenge for a pulsed-based radar. A large gain-bandwidth (GBW) product as well as minimal distortion in terms of amplitude ripple or group-delay variations are desirable. Various techniques to achieve wideband amplification in CMOS are described and the design and measurements of three different distributed amplifiers discussed.

In the end, a single-chip pulsed-based coherent radar imager implemented in SiGe BiCMOS process will be described.

1.2 Low-Cost Breast Cancer Screening

Worldwide, breast cancer is the second most common type of cancer after lung cancer. Women in the United States have the highest incidence of breast cancer in the world and

among them this is the most common cancer and the second most common cause of death [17] [18].

Currently, mammography (X-ray imaging of compressed breast) is the dominant modality in early detection schemes. The mammogram essentially is a map of breast density. The resolution of the obtained image could be high, however, the contrast is low. It is difficult, even impossible in many cases, to localize a tumorous structure of few milliliters in size or to distinguish between benign and malignant tumors. According to [17] [18], some of the limitations of mammography are missing up to 15% of breast cancers, difficulty with women with dense breasts, and large false positives (as high as 10-15%). 10% of mammograms result in inconclusive data, only 10% of which turns out to be malignancies. The technique causes discomfort, is not widely available, and due to the ionizing radiation, poses health risks. Ionizing radiation limits the frequency of test to 18-24 months and is only performed for the specific age bracket of +40 years old. These are major obstacles in detecting fast-growing invasive cancers.

Using microwaves to detect cancerous tissue relies on the difference in dielectric properties of malignant and normal tissue. There is evidence that tumors have higher water content than the normal tissue and for certain types of tumor, such as breast carcinoma surrounded by fatty tissue, the difference could be considerable. The obtained data available in the literature [1] [12] [10] [19] show the existence of contrast between cancerous tissue in the breast and the surrounding normal tissue. The contrast is considerably larger for adipose-dominated tissue in which the conductivity is much lower. This allows for a reliable detection scheme. The characteristic difference in the dielectric parameters (permittivity and conductivity) is mainly due to increased water content from increased protein hydration and vascularization of the tumorous tissue.

1.3 Skin Cancer Detection

Another application of this imaging modality is in the detection of skin cancer. Around 160,000 new cases of melanoma are diagnosed worldwide each year with 48,000 melanoma-related deaths [17]. Malignant melanomas account for 75% of all deaths associated with skin cancer. Visual examination is often used to detect melanomas in the initial stage. Diagnosis is made using factors such as size, shape, color, border irregularities, presence of ulcer, tendency to bleed and other factors. Skin biopsy under local anesthesia is used on suspicious sites to determine the characteristics of the mole. However, there are multiple problems with the current methods in practice. Visual examination is generally prone to human errors. The accuracy of detection is somewhere between 55% and 82% [20, 21, 22, 23]. It may result in many suspicious sites especially in older patients and this often leads to waiting for the progress of the moles. This increases the risk of metastases or other complications. Also, the biopsy procedure is often inconvenient and produces many false alarm cases. Increasing the threshold for requiring a biopsy on the other hand increases the

risks of late detection. Other melanomas are amelanotic (colorless or flesh-colored) and this complicates the procedures.

Other skin cancer types are sometimes even more difficult to detect visually (e.g. Merkel cell carcinoma). This and the uncertainty in diagnosis and the issues involved with multiple biopsies leads to a significant requirement for an alternative method. The method needs to be simple, non-invasive, conclusive, easy to interpret, low-cost and able to be widespread. Millimeter-wave imaging is a good candidate since the tumor is close to the skin and hence the losses are low. Also, use of the higher frequency band in TUSI allows for improved depth and lateral resolution and could significantly reduce the number of biopsies. Patients with multiple suspicious sites (not rare in older patients) are able to be diagnosed at early stages.

1.4 Other Applications

The TUSI array is designed to be pixel-scalable. This means that the single transceiver is design to be as agnostic as possible to the size of the array. Some of the other applications for this large and scalable array imager are in security imaging (e.g. airports) and for smart surfaces. The TUSI array can be used in large-scale smart and interactive surface that can detect 3D position of objects and by that facilitate various applications in human-computer interfaces. More details are provided in the next chapter.

1.5 Organization of the Dissertation

The rest of the dissertation is organized as follows. Chapter 2 presents an overview of the time-domain pulsed-based imager. Design concepts as well as the main challenges are described. System level simulations are provided. Chapter 3 discusses the design of various antenna elements for the 90 GHz transceiver. Antentronics is introduced as a technique to synthesize the impulse response of the antenna for desired radiation characteristics. Chapter 4 discusses the design and implementation of transmitter blocks that are capable of generating sub-50ps pulses from the antenna. Two different transmitter topologies together with the measurement results are presented. Chapter 5 reviews the design of the receiver circuits for the detection of the wideband pulse. Various amplifier topologies that provide large gain-bandwidth product are analyzed and compared to the state of the art. In the end, a 1.5 THz gain-bandwidth SiGe amplifier is presented for the final transceiver. In Chapter 6 we present the design, implementation, and measurements of a single chip radar transceiver in a SiGe BiCMOS process. Finally, the dissertation concludes with a summary of the key results and a discussion of future research in Chapter 7.

Chapter 2

Background and System Design

In this section some of the system-level challenges associated with the design of an array imager are briefly discussed.

A silicon-based imaging array for remote measurements of complex permittivity of tissue is introduced. Using a coherent pulsed measurement approach, this time-frequency resolved technique recovers the three dimensional mapping of electrical properties of the subject in the microwave/millimeter-wave frequency spectrum. Some of the major challenges in the design of the system will be examined.

In any imaging modality, a specific property of the tissue is mapped to the final image (e.g. tissue density/atomic number for X-ray or water content for MRI). We target electrical properties of tissue in the RF to mm-wave spectrum for cancer detection. These are the tissue responses to relatively weak electromagnetic fields dominated by electric properties of ensemble of cells. Charge asymmetry and rotational mobilities cause electric dipole moments that respond to the external electric field differently based on the properties of tissue.

The main obstacle in measurement of dielectric properties has been the low penetration depth of microwave fields in human body. Higher frequencies that yield better resolution have even higher losses. The dynamic range for obtaining such images has not been attainable with standard measurement instruments. With the accelerated pace of growth in silicon and electronic devices, reaching the required dynamic range is becoming a possibility.

2.1 Radar Imager

A high-resolution, phase-coherent, miniature-radar is designed for the detection of dielectric boundaries in various tissue structures. Historically, high-resolution radar came about primarily with the ability to control and detect the phase of the radar signal. Coherent sources facilitated this control over phase. Early on, after the realization of phase-coherent radar, concepts like synthetic aperture radar (SAR) came about. SAR uses the motion of radar antenna to synthesize a larger effective aperture [24].

A pulsed radar transmits short pulses and measures the reflections to determine the objects located in the vicinity. Pulses are sent periodically with a period of PRI (pulse repetition interval). In addition, these pulses are often sent in bursts depending on the exact radar application. As an example a burst might consist of 1,000 pulse intervals. In general we could describe the transmitted pulse train by:

$$P_{TX}(t) = \sum_{i=1}^K \sum_{j=1}^N x(t - iT_B - jT_p) \quad (2.1)$$

where

$$x(t) = A(t) \cos(2\pi f_0 t + \phi_{ij}) \quad (2.2)$$

Here T_p is the PRI and T_B is the burst period. The pulses are sent, will get to the target and reflect back to the receiver. The received waveform will consist of a combination of pulses within the PRI depending on the shape and position of reflectors. If the center frequency of the pulse is locked to the PRI, then the phase of the received pulses are preserved and the phase information could be used to improve resolution.

In the array, in addition to the pulse arriving to the point from which it was transmitted, it will also arrive in all of the receivers in the array. From the time position of the pulses arriving to various array receivers, we can identify the dominant reflection points. For the processing, if sufficient SNR exists after averaging, delay-sum or other array processing algorithms could be used to construct the image. In addition, with the system architecture providing control over frequency, pulse width and position, one could also perform time-reversal or diffraction tomography algorithms [25] [26].

Fig. 2.1 shows the general concept of the array imager. Four physical elements provide the resolution for this imager: Large instantaneous bandwidth (short pulse), wide frequency tuning in frequency generation, large synthesized aperture, and phase coherency of the array. In this figure, each square represents a single element of the array which is implemented as a transceiver in silicon technology. This could be thought of as a single “pixel” of the imager. These pixels are then integrated on a common board to form the imager (Fig. 2.1). The single element is designed for scalability. Depending on the specific application, different array sizes or number of elements may be required. The element is designed to be as agnostic as possible to the size of the overall array.

2.2 Contrast

Contrast is the observable and detectable difference in the resulting image of two different tissue types. Higher contrast images are better in distinguishing different objects.

In an imaging system, contrast depends upon the physical mechanism used for imaging. As an example, in X-ray imaging, the tissue density and attenuation are mapped to the image. It does not provide a good soft tissue contrast and is mainly suitable for imaging

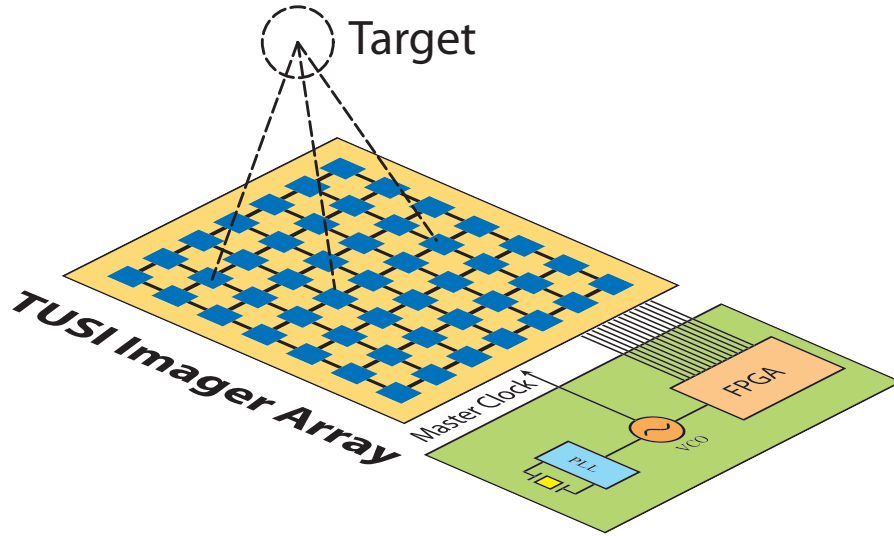


Figure 2.1: TUSI imager concept.

skeletal properties. In MRI, the system measures the spin density and this allows for a much better soft tissue contrast. Given the dominance of water in the body and the spin associated with hydrogen, MRI maps the water content in the body.

Dielectric properties of tissue can also be used for direct reflection imaging. These properties have been studied in the literature [27, 28, 29, 30]. Water is an important part for determining the tissue response at RF to mm-wave frequencies. In biological materials, water is a solvent for salts, protein, nucleic acids, and smaller molecules. Water's response can be modeled by a single-pole Debye relaxation equation with a primary dispersion in the GHz range ($\tau \simeq 10\text{ps}$). At higher frequencies, another dispersion at approximately 670 GHz is observable. Besides water, the majority of biological macromolecules (e.g. proteins) act as polar molecules with permanent or induced dipole moment. The molecular structure, configuration, and size of these molecules determine the size of the dipole moments.

Measurements of tissue properties show several broad dispersions (e.g. α , β and γ dispersions) [27, 28, 29, 30]. The α dispersion is at low frequency and is characterized by very high permittivity values. The β dispersion occurs at intermediate frequencies (radio frequencies) and originates mostly from the capacitive charging of the cellular membranes. At higher frequencies where the water response is dominant, the γ dispersion due to the dipolar polarization of tissue water takes place. The data for various tissue types is available from [27, 28, 29, 30].

For screening and diagnostic imaging applications, response contrast between normal and cancerous tissue is important. Various studies have shown this contrast for different tissues. For example, previous studies [1] [12] [10] have shown normal and malignant breast cancer tissue to have considerable contrast in the microwave frequency range. The char-

characteristic difference in the dielectric parameters (permittivity and conductivity) is mainly due to increased water content from increased protein hydration and vascularization of the tumorous tissue [31]. It can also be affected by other factors such as membrane potential differences, changes to cell connectivity, and sodium concentration. Necrosis in tumorous tissue leads to breakdown of cell membrane and affects low-frequency conductivity.

Fig. 2.2 shows the available dielectric contrast for normal breast tissue and cancerous tissue (from [1]). Based on the conductivity levels reported for normal tissue, the shown contrast is mainly between cancerous tissue and fat-dominated tissue. For connective and gland tissue in the breast, the contrast is smaller than what is reported here. Larger variability of dielectric properties for normal breast tissue has been reported in [32].

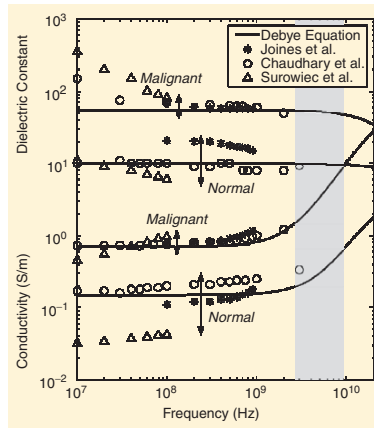


Figure 2.2: Dielectric-properties data for normal and malignant breast tissue (from [1]).

2.3 Resolution Basics

2.3.1 Depth Resolution

Resolution is a measure of the ability to distinguish two small objects that are close to each other. Spatial resolution is often described in terms of the point spread function or PSF. PSF describes the response of an imaging system to a point source or point object. The PSF can be thought of as the system's spatial impulse response [33].

Here, we will analyze resolution in radar terms. Ultimately, the goal is to define the general resolution in the sense of detecting two close by targets. Pulse-based radar sends out short time-domain pulses and detects the echoes/reflections from nearby scattering points. The depth/range resolution is given by [13]:

$$\delta R = \frac{c\tau}{2} \quad (2.3)$$

where c is the speed of light and τ is the pulse width (Fig. 2.3).

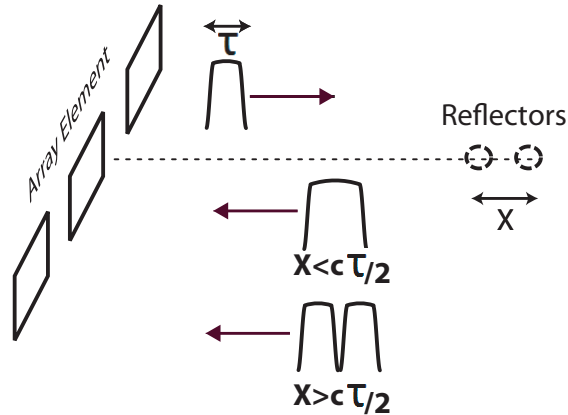


Figure 2.3: Depth resolution in pulsed radar.

It is clear that a large BW is required to achieve mm-range resolution level. The wave velocity being lower in tissue helps reduce this requirement but this is often negated by the increase in dispersion and pulse-spread. For applications such as tumor diagnostic scans, detection of mm-size tumors is desirable. As will be discussed later, in this system a pulse width close to 25 ps is targeted. This will provide resolution close to 1- 2 mm in tissue. Generating, controlling, transmitting, and detecting narrow pulses down to 25 ps poses many circuit and system design challenges that will be described in future chapters.

2.3.2 Lateral Resolution

Lateral or cross resolution is set by antenna parameters and more specifically the aperture of the imager. In order for the waves from two lateral points to be distinguishable, the received wave from the two points has to have some phase-shift across the antenna aperture. If the phase-shift is too small, the two points will be indistinguishable. Assuming this phase shift to be 180 degrees, we can derive a relationship between the lateral resolution, δx , the range, R , the aperture, D , and the wavelength, λ [13]

$$\delta x = \frac{\lambda R}{D} \quad (2.4)$$

One can also derive a similar relationship by taking into account an approximation for the beamwidth, $\theta = \frac{\lambda}{D}$, and multiplying that by the range to get the lateral resolution, $\delta x = R\theta$.

The end result is that for an antenna to have a lateral resolution close to the wavelength, the overall aperture and the distance to the image need to be of the same order. It has to

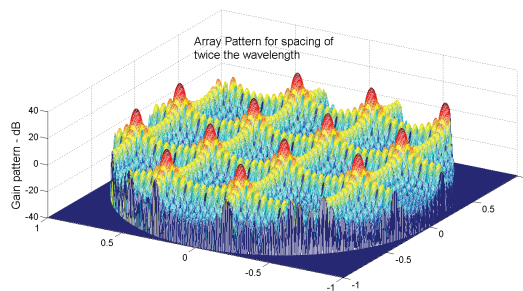


Figure 2.4: Frequency-domain array with grating lobes. Array uses a 2λ element spacing. Grating lobes are visible.

be pointed out that for time-based pulsed radar with a large aperture the pulse-width as well as delay/phase resolution will also play a role and the equations only serve as a first order approximation for lateral resolution.

In our system, the high-frequency band (94 GHz) is designed for the best resolution level. For example, with an aperture of 3 cm and maximum range of 3-5 cm, we need an effective wavelength of close to 1.5 mm in tissue to obtain a preprocessed resolution of 2-3 mm. With these numbers and an effective dielectric constant of 4, a 94 GHz carrier can provide the required cross-resolution. Post-processing can further enhance the resolution to obtain closer to 1 mm voxels.

2.4 Pulsed-Based Antenna Arrays

In conventional antenna arrays, as previously discussed, the broadside beam-width or resolution improves with array aperture [34]. This shows that to improve resolution we need to increase the number of elements or effectively the size of the array. To avoid increasing total system cost or complexity, we can use spatial sub-sampling and increase the spacing between array elements so that the same aperture will use less number of elements. However, in a conventional frequency domain array, this leads to *grating lobes* which are essentially secondary main-lobes that produce ambiguity in the direction of arrival [34]. An example of this is shown in in Fig. 2.4.

Pulsed based arrays use pulses to send and receive data from various array elements [35, 36]. In contrast to sine-based systems, these arrays do not possess side-lobes and have a side-level that is dependent on how well the space is sampled. The main-lobe width of these arrays is dependent on how many samples are taken. Qualitatively, this could be described as follows: In a sine based array, the signals have sidelobes since they are extended in time and can interfere (constructively) at an angle that is *off-center*. If the antenna spacing meets the Nyquist criteria then this interference only shows a sidelobe (as compared to a

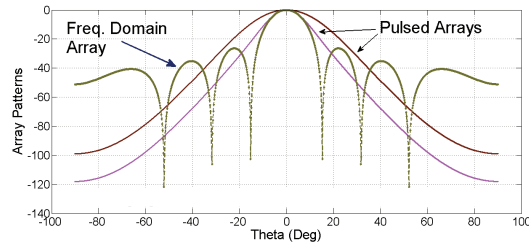


Figure 2.5: Comparison of pulsed and frequency-domain arrays.

grating lobe). In pulse arrays, however, the signals are so short in time so that the result is either addition of all (in the main beam) or just a sum of signals added incoherently. An example of response of a pulsed array with only side-levels is shown in Fig. 2.5.

In pulsed based-systems instead of frequency and wavelength, pulse width and observation window are well defined. Therefore, comparisons of array element spacing need to compare to quantities such as vT in which v is the wave velocity and T is the observation time. Later in the chapter we will describe the issue of false illumination points in a pulsed array.

2.5 Signal-to-Noise Ratio

Another challenge in obtaining effective images through reflections in the microwave and mm-wave regime is the large signal loss in human body. Given that unlike other modalities these losses are the main contrast mechanisms, the system should work with very small SNR levels. Fig. 2.6 shows simulated signal penetration and raw resolution in body fat and fluid based on published models [30]. Penetration depth is plotted for 100 dB of tolerated total loss and resolution is based on signals with 30% relative BW to carrier. This is a first-order approximation for the preprocessed resolution obtainable for each frequency. It is clear that, especially for images involving body fluids, the obtainable resolution becomes smaller than penetration depth for higher frequencies and therefore the low frequency regime has to be simultaneously leveraged. Also, depending on the cross section of the target, the actual loss may be more by 20-30 dB.

For the mentioned 100 dB loss, frequency BW close to 30-40 GHz, and transmit peak power close to 20 dBm, the obtained SNR is close to -30 to -40 dB. There is a need for large number of array elements as well as long integration windows to improve the SNR levels. Synchronized arrays can generate an increase in transmitted signal power by a maximum of N^2 . However, array losses due to lack of perfect synchronization or quantization in delay generation techniques, could result in both power losses as well as pulse widening or “spatial dispersion” (described in section 2.9.3). Beamforming will also require a wideband delay element to shift the modulated pulse signal in time domain. Narrowband approximations to

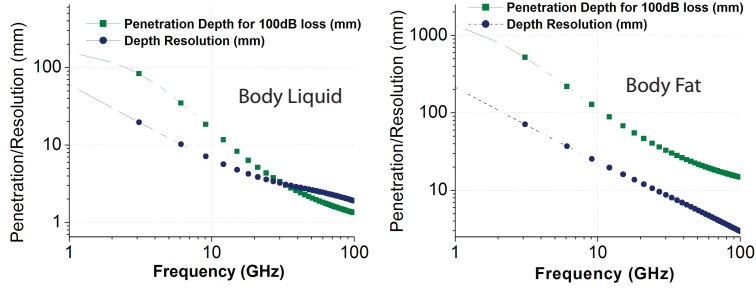


Figure 2.6: Penetration depth and resolution for body fat and body liquid assuming 30% relative BW and 100dB two-way loss.

this delay (fixed phase shifts) will encounter unacceptable errors for larger arrays occupying a wide frequency bandwidth. In this imager, the transmitter uses a delay-locked-loop to move the envelope of the pulse before it is multiplied by the carrier. The carrier phase need to also be shifted by a narrowband phase shifter. After combining (multiplying) the carrier and the envelope, a delayed version of the modulated pulse is obtained and the need for a wideband phase shifter is eliminated.

2.6 Resolution Limits

In the case of reflection measurements for medical applications, the resolution limit is tighter than the classical two-point radar resolution. This is due to the large dynamic range of signal levels at the receiver. Distinguishing multiple targets with largely different amplitudes is more difficult than when the amplitude levels cover a smaller range.

Resolution limits generally address two issues: ability to distinguish close objects (with same or different amplitudes) and the ability to pinpoint the position of an object in the presence of noise. It will be shown that both resolution limits improve with enhancing the signal bandwidth (BW).

As previously discussed, for two equi-amplitude signals, the resolution will be $\frac{c\tau}{2}$ (for pulsewidth of τ) or equivalently $\frac{c}{2BW}$ (where BW is the signal bandwidth for an alternatively modulated radar signal) [13].

If the two signals have different amplitudes, especially in the case where one completely dominates the other, this becomes more restricting. Fig. 2.7 shows the effect of obtaining lower resolution with larger amplitude ratios. For example, if a target is 20 dB lower than a main reflector (deeper in a lossy medium or smaller cross-section), the resolution decreases by approximately a factor of 3. Therefore, we can conclude that resolution in a lossy medium is also a function of depth due to the differences in signal amplitudes.

In order to counter this effect, we need to either push for narrower pulse generation

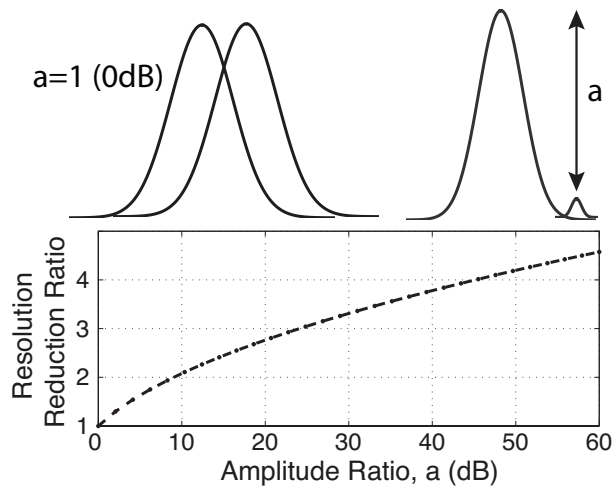


Figure 2.7: Resolution reduction effect from large variations in amplitude. Here, the resolution reduction ratio is shown normalized to the case of $a = 0\text{dB}$.

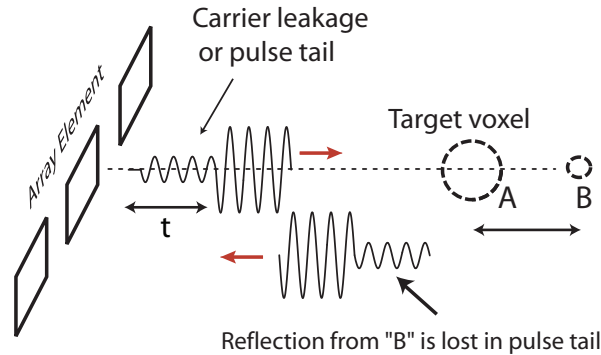


Figure 2.8: Effect of pulse integrity on performance (echo effect shown here).

(more BW) or compensate the effect with assumed prior knowledge of pulse shape and time-gated post-processing. The former is difficult especially that this puts a limit on pulse fidelity and pulse amplitude roll-off in time domain. The latter requires better knowledge of dispersion characteristics. With that knowledge, one could also apply pre-processing to the transmit waveform to improve resolution. In all cases, it is necessary for this system to utilize calibration/compensation algorithms to achieve the resolution limit.

Pulse fidelity is also important for resolution. Fig. 2.8 illustrates the effect of a pulse tail on detection. A large leakage or pulse tail manifests as a false echo that can completely overwhelm a reflection from a smaller object.

Resolution is also restricted from accurately identifying a pulse edge position. We can calculate the Cramer-Rao lower bound (CRLB) for the estimation of the pulse arrival time

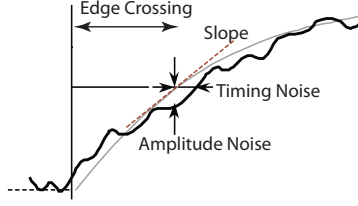


Figure 2.9: Received pulsed edge with noise.

(and from that the position through $R_0 = vt_0/2$ where v is the wave propagation velocity). We will assume a white noise or equivalently colored noise with sampling at nulls of the auto-correlation function. The lower bound on the timing uncertainty can be derived as [37]:

$$\overline{(t_0)^2} = \frac{1}{\beta^2 \frac{E}{N_0/2}} \quad (2.5)$$

where E is the pulse energy, N_0 the noise spectral density and β is an effective bandwidth [37]. Once again, a sharper and narrower pulse results in a better decision on the pulse arrival time.

This result can be explained by Fig. 2.9. An edge from a pulse of width τ and rise time t_r is shown together with additive noise. The quantity to be estimated is the arrival edge of the pulse. The amplitude noise gets “translated” to timing uncertainty through the slope of the pulse edge and hence:

$$\overline{t_0^2} = \frac{\overline{n^2}}{slope^2} = \frac{\overline{n^2}}{A^2/t_r^2} = \frac{t_r^2}{SNR} \quad (2.6)$$

where A is the pulse amplitude and $\overline{n^2}$ is the noise variance (N_0B). Given that signal power (S) is $\frac{E}{\tau}$ and noise BW is B , equation (2.6) can be rewritten as:

$$\overline{t_0^2} = \frac{t_r^2 B \tau}{\frac{E}{N_0}} = \frac{1}{B_{eff}^2 \frac{E}{N_0}} \quad (2.7)$$

where the last part is due to the relationship between τ , t_r and noise BW (B) resulting in an effective bandwidth (B_{eff}) dependent on the specific pulse shape. Thermal noise is not the only limiting factor in obtaining the resolution. Direct timing jitter from the transmitter (σ_{TX}), receiver (σ_{RX}) or from phase noise on reference oscillators (σ_{PN}) could also add to the final variance.

$$\overline{t_0^2} = \frac{1}{\beta^2 \frac{E}{N_0}} + \sigma_{TX}^2 + \sigma_{RX}^2 + \sigma_{PN}^2 \quad (2.8)$$

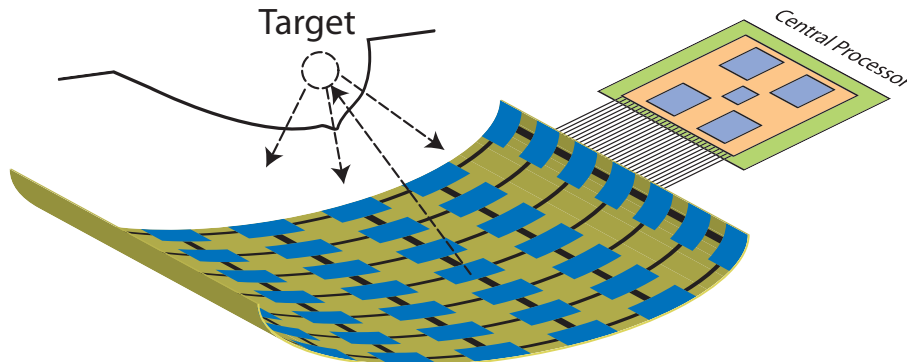


Figure 2.10: TUSI array imager on flexible substrate.

2.7 TUSI Imager

This dissertation describes the design of a Time-Domain Ultra-Wideband Synthetic Imager or TUSI. TUSI implements a synchronized time-domain array with a two-tier integration scheme. In the first tier it integrates single coherent transceiver elements on a chip. This is a bi-static setup (separate TX and RX) to reduce direct leakage effects. Time-gating eliminates “delayed” leakage from TX.

In the second tier, which enables a scalable array, single-element chips are integrated on daughterboards, multiples of which are then mounted on a common motherboard. The motherboard also integrates clock and power distribution circuits as well as the central processor (Fig. 2.1) . Distributed sampling with partial local computation reduces the aggregate data rate provided to the central processor to a few hundred Mbps.

Depending on the specific application, the array may use a flexible substrate. In that case, the processing unit will have a remote connection to the array itself. The flexible-substrate imager will reside on the body separated by a matching layer. The conceptual shape of the imager on the flexible substrate is shown in Fig. 2.10.

Using the individual transceivers in the array, the TUSI system is capable of imaging in both coherent multi-static and coherent MIMO modes. Several techniques are used to improve the resolution beyond the raw pulse-width limited region. A high pulse repetition frequency (PRF) results in better averaging gain in a given window. TUSI uses frequencies spanning from low microwave to mm-waves to provide flexibility in face of diverse loss mechanisms in different tissue types. In addition, at each of the spectral bands, the center frequency is swept on a batch-by-batch basis to cover the frequency spaces (in the output comb) caused by the high PRF.

The block diagram of the system is shown in Fig. 2.11. This is the block diagram of the high-frequency channel that has both of the antennas integrated on the chip. On the receiver the signal goes through a wideband antenna system. The final antenna is designed

as a wideband but resonant tapered loop antenna. The non-idealities of the receiver antenna can be partially compensated for in the processing. The signal then goes through a wideband gain stage. This is the common gain element that is shared between various frequency bands. For the 94 GHz band, the amplifier is followed by a micro-mixer. In the baseband, the quadrature signals go through 10 dB gain and buffering. Signal conditioning, filtering, and digitization happens on the chip. Partial analog/digital integration reduces the interface data-rate to about 1-10Mbps.

In the transmitter the carrier signal from the PLL goes through a narrow-band phase shifter followed by buffer and power amplifier. The carrier is multiplied by the envelope both in the PA as well as the antenna element. As will be described in future chapters, switching happens either in each of the PA or antenna or through a hybrid switching technique. The antenna element in the transmitter, therefore, must incorporate switching functionality. This is the topic of chapter 3.

A 3 GHz reference clock is distributed to each of the transceivers. The reference clock is used to drive central timing circuits, the DLL and acts as the reference clock for the PLL. The DLL sets the PRF and adjusts the position of the pulse. An interpolation scheme is utilized in the DLL architecture (Fig. 2.12). A pulser generates activation signals that drive the PA or the antenna as well as the sampling circuits. In one implementation, an additional “fast clock” is extracted from the PLL chain and is used for further fine-tune synchronization. To reduce the signal dynamic range incident upon the electronic elements, a time-stretched gating circuit is employed. During the averaging period, each “path” will work with a fixed signal level and this reduces the dynamic range.

To enhance the SNR level at the receiver, extensive signal averaging is performed. A sigma-delta modulator is used for data conversion. In the averaging scheme, the quantization noise is shaped through the filter while the white input noise from the receiver sees a “brick-wall” filter (no shaping). For the averaging window of larger than 1 ms, as targeted here, the quantization component of noise will be extensively suppressed. The thermal input noise provides dithering to the sigma-delta converter and this reduces the tones from the limit-cycle oscillations.

2.8 Undesirable Coupling of SNR with Resolution

In many imaging systems the resolution and SNR are tightly coupled through change of signal energy with pixel size. Improving the resolution often requires better SNR levels. This is also true for super-resolution techniques that leverage prior knowledge of the waveform. In this imaging modality, jitter limits the number of averaging cycles for a specific target resolution. Sampling and accumulating with a non-ideal clock signal is equivalent to convolving the signal with the probability distribution function (p.d.f.) of the clock arrival edge (in the limit) [38]. For a gaussian p.d.f. of timing jitter with variance of σ^2 , the signal is effectively low-pass filtered by a gaussian filter of an equivalent bandwidth ($\propto \frac{1}{\sigma}$).

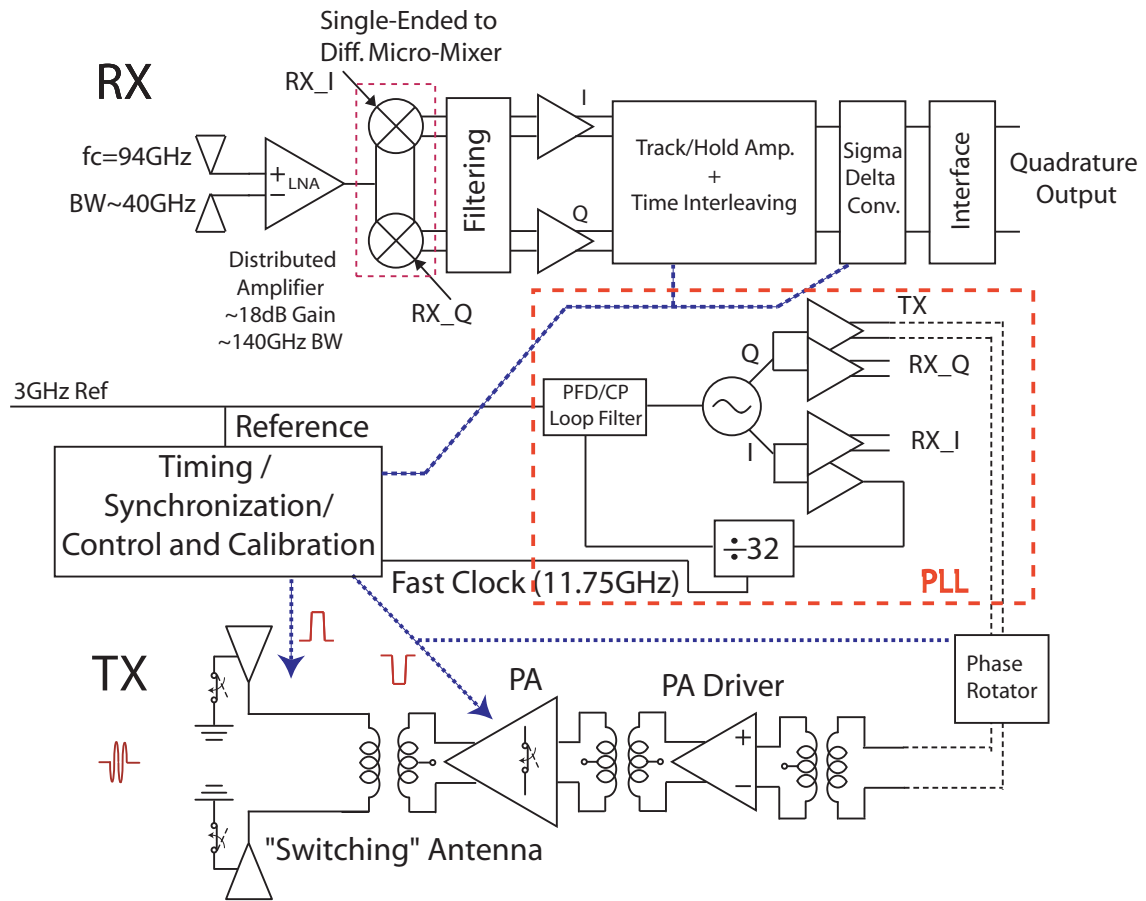


Figure 2.11: Proposed array architecture and the block diagram of the TUSI transceiver for the high frequency band.

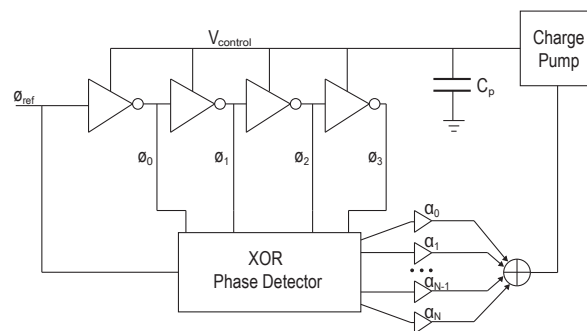


Figure 2.12: Delay-Locked Loop architecture for generating fine time steps.

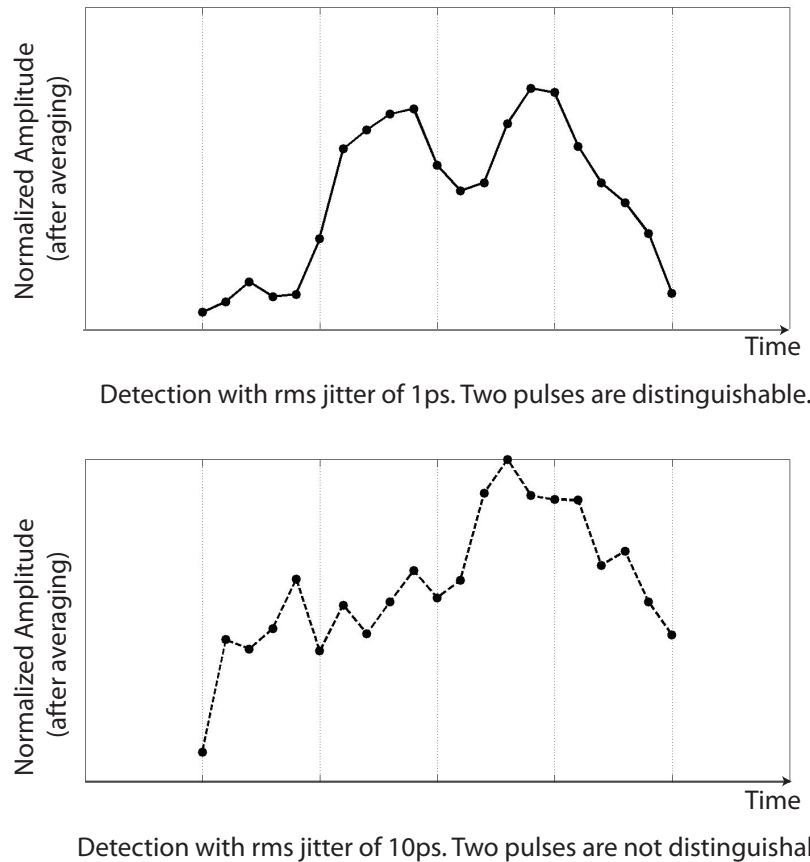


Figure 2.13: System simulations of the transmit/receive chain. LO phase noise, transmit and receive jitter, thermal noise and quantization noise effects are included. Two pulses are 25ps wide and are separated by 60ps. The averaging window is assumed to be 1ms with PRF=1GHz. In one case the system has a total rms jitter of 1ps and in the other 10ps.

Therefore, for long integration times, the image resolution is reduced. This limits the total jitter budget of the transmitter and the receiver to 0.5-5ps depending on range resolution required. Fig. 2.13 shows this effect through transmit/receive system simulations. Averaging has been performed to bring the signal out of noise.

2.9 Limits of Integration for Closing the SNR Gap

The SNR is improved by $10\log(N)$ with having N pulses integrated coherently. The receiver does not use an RF phased-array architecture due to lack of scalability for large apertures. Phase/amplitude errors (from narrow-band approximation as well as from other sources) will limit the resolution in case of RF phase-shifters. Since processing is performed

in the digital domain, extensive averaging is performed on the RX to compensate for the low SNR. However, there are limitations in increasing the averaging window indefinitely.

2.9.1 Stability of the Imaging Setup

Similar to microscopic vibrations that limit optical resolution, macroscopic vibrations lead to spatial filtering of the object and limit the obtainable resolution. Without stabilization techniques, averaging windows are limited to about 10ms-1s, especially for a hand-held imager that is the target of this study.

2.9.2 Limitations on Increasing the PRF

A repetitive pulse transmission approach is being utilized in which short pulses of width τ are transmitted with a repetition frequency of PRF. Here, given a fixed total integration window, we can increase the number of pulses being averaged by increasing the PRF. However, there are multiple limitations for this approach. First of all, the maximum unambiguous range can limit high PRF as it does in classical radar [13]. In a time-based array system, as proposed here, another limitation is due to the cross-range ambiguity from high PRF. Intuitively, this is related to the fact that for large arrays and beam angles, the delay spread between first and last element can be quite large. For a far-field array this difference is $(N - 1)d \sin(\theta)/v$ where N is the number of elements in array, d is the elements spacing and θ the beam angle from broadside. As an example, with an aperture of 10 cm, 45 degrees angle, and $v = 1.5 \times 10^8$, we need 466 ps delays. With the PRF approaching a few GHz this will potentially cause ambiguity in angle of incidence. If, for example, the PRI is set to 466 ps, the signal from element at the near point of the array to the desired direction combines with the second PRF signal of the element at the far end for a sidelobe in $\theta = 0^\circ$ direction.

If the object resides in the near-field of the array, this relationship will be different but in essence one can assume that the delay gets larger with array size. Once this delay difference becomes comparable to the pulse repetition interval (PRI or $1/\text{PRF}$), secondary “illumination” points will appear and cause ambiguity. Fig. 2.14 shows this effect. This is a gray scale coding of the received amplitude versus horizontal position and time. The arrays are focused to $x=X_{\text{max}}$ (Fig. 2.15). In one case (left) the period is chosen to avoid range ambiguity. The second case is for a much higher PRF (same pulse width and distance). Multiple focused points are observed. For illustration purposes a sparse array (large antenna spacing) is chosen to exaggerate secondary focus point amplitudes. In conventional arrays the actual peaks will be smaller.

2.9.3 Accurate Delay Generation

The spatial resolution is a function of the pulse width as well as the delay steps that perform beamforming. The time-step quantization smears multiple reflection points and

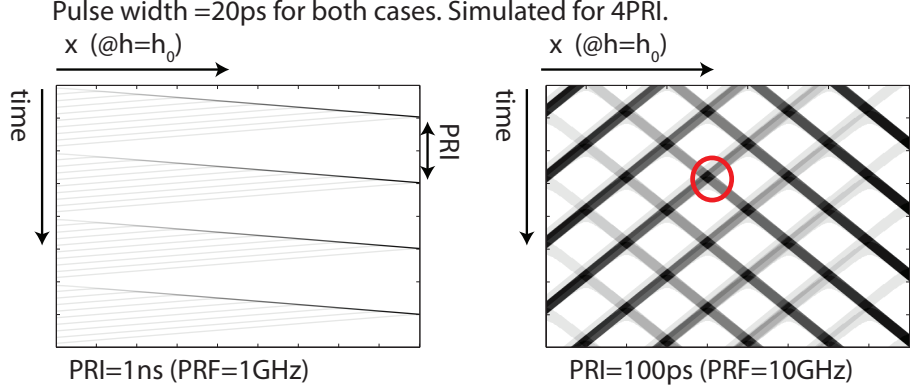


Figure 2.14: Received signal (gray-scale coded, dark=large amplitude) at a fixed vertical distance from a 10-element linear sparse-array (at $h = h_0$). The X-axis is the horizontal position (x) and the y axis is time. The array is focused at $x=X_{max}$ and simulations are done for 4PRI periods. The circle shows one of the ambiguous points.

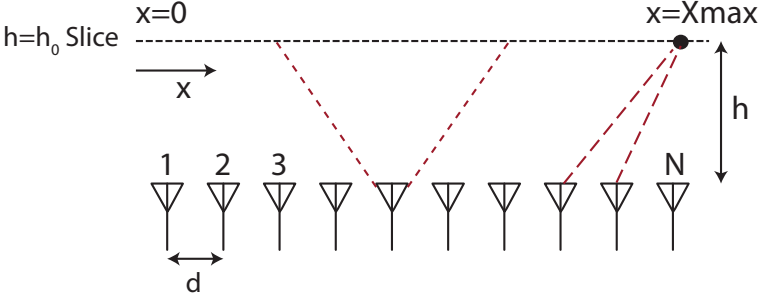


Figure 2.15: Linear antenna array with element spacing d . Here, a slice at vertical distance h_0 from the array is being analyzed.

knocks down or distorts high spatial-frequency components of the image. To calculate the required time-steps, we will first assume a full 180° field-of-view (FOV) for individual antenna elements. With this, we can identify two timing-step requirements. One is related to the relative delay required on two antenna elements when they are focusing on a single point ($\vec{r} = \vec{r}_0$). The second is for one antenna element focusing on two adjacent resolution pixels. It could be seen that these problems are geometrically equivalent as long as the array spacing and the pixel sizes are approximately the same dimension. We will analyze the case of adjacent antennas focusing on a single point.

To find the time step we will take the case where we are focusing at a slice with vertical distance h from the array and at $x = 0$ (Fig. 2.15). For $h \gg d$ the worst case will be the delay difference between the first two elements (directly beneath the object). We can

approximate this as:

$$\Delta\tau_{min} \simeq \frac{d}{2v} \times \frac{d}{h} = \frac{d}{\alpha} \quad (2.9)$$

where d is the antenna spacing. The other extreme case is where $h \ll (N - 1)d$ in which the delay difference approaches d/v in the limit. Obviously (2.9) is the limiting factor in this case. Assuming $PRF = T$ and center frequency of f_0 , the α factor with an antenna spacing of half-wavelength (material wavelength or λ_g) is going to be:

$$\alpha = \frac{h_{max}}{d} = \frac{h_{max}}{\lambda/2} = \frac{vT/2}{v/2f_0} = f_0 \times T \quad (2.10)$$

Here, the maximum range is replaced with a function of PRF. Intuitively, larger T dictates longer distances which leads to a smaller observation angle. This leads to a finer delay step requirement (larger α). In addition to that, a larger center frequency will also lead to smaller antenna spacing and hence a smaller observation angles. As an example, for $PRF = 1GHz$ and $f_0 = 90GHz$ we have $\alpha = 90$, and we need a delay step of:

$$\Delta\tau_{min} \simeq \frac{\lambda/2}{\alpha} = \frac{1}{4f_0\alpha} \quad (2.11)$$

This will translate to a raw delay step requirement of $30fs$ which of course is beyond reach. In the calculated scenario, under worst range/cross-range conditions this delay resolution results in no quantization noise. However, numerous frequency/time domain signal processing techniques can mitigate the error from non-ideal steps.

We can also see this through computing the range inaccuracy incurred if there is a quantization noise in delay step. This is yet another way through which resolution (range accuracy) and SNR (ability to energy-combine sources coherently) are coupled. Trading off resolution for SNR is needed especially in cases where deeper signal penetration in tissue is required.

Fig. 2.16 uses a normalized Gaussian pulse to illustrate this “spatial distortion”. When the array elements focus on a single spatial position, we will have a signal that is ideally N times larger than the signal from one element. This holds true only if there are no timing errors from these array elements. In reality, we do not have ideal delay elements and inevitably there will be timing errors from delay quantization. We have used a simplified model to show the effect of this quantization. A delay error step is defined as the smallest timing error in the array. We have also assumed a linear and uniform model for the delay errors. For example, for an array of 100 elements, the delay error spreads uniformly between one error step and 100 error steps. Other distributions could have also been used depending on the quantization method. We have also normalized this error step to the width of the Gaussian pulse. Fig. 2.16(a) shows the effect of delay quantization. The solid line in this figure shows the ideal case in which 100 pulses are summed from all array elements with no timing errors. The other dashed lines show the summed signal in presence of (increasing)

timing quantization error. As seen in the figure, there are four issues arising due to delay quantization. The pulse peak amplitude drops (SNR loss), the peak amplitude moves in position (this needs calibration), the slope reduces (increasing noise effects and increasing the CRLB on ranging accuracy according to (2.6)) and the pulse width increases (reducing resolution in distinguishing multiple pulses).

Fig. 2.16(b) shows the effect on pulse width and slope. The x-axis is the delay error step. For an array of 100 elements, the largest error on the last element will be 100 times the delay step shown here. It is evident that if the error is within 1% of the pulse width, the errors in the width and slope of the final obtained pulse are within 10% of the ideal case. This result could be generalized to larger or smaller arrays as long as the error bound is scaled accordingly. Our aim is to target a pulse width of 25ps and a delay resolution step close to 1ps (less than 5 %). Since in many practical scenarios the number of elements is smaller than 100 (on each axis), this delay step will in fact be adequate for the desired resolution. If better resolution is required, the array will be sub-sampled at the expense of lower SNR.

2.10 Coherent Radar Phase Information

One of the characteristics that sets coherent radar aside from optical imaging is that we know both the amplitude and the phase of the backscattered signal. The TUSI transceiver is designed to be completely coherent and phase information could be used to enhance resolution. A simple explanation is presented here. Thorough description of phase-based imaging and interferometry is available in the literature [24].

The transmitted signal can be presented by:

$$V_t = A(t) \cos(\omega t + \phi_0) \quad (2.12)$$

The received signal from a single reflector is then represented as:

$$V_r = A(t - \tau_g) \cos[\omega(t - \tau_p) + \phi_0 + \phi_1] \quad (2.13)$$

where τ_g and τ_p are the group and phase delay through the medium and ϕ_1 is the extra phase accumulation (e.g. by target reflection). For the moment we can assume a non-dispersive response through which the phase and group velocities are the same. In this case, $\tau = \frac{2\Delta l}{c}$ where Δl is the distance to target and c is the propagation velocity in the medium (air or complex media). Coherent quadrature detection leads to two baseband signals that carry the phase information. For example, assuming negligible delay through the system, the filtered and detected in-phase component will turn out to be:

$$V_I = A(t - \tau) \cos[\omega \frac{2\Delta l}{c} + \phi_0 + \phi_1] \quad (2.14)$$

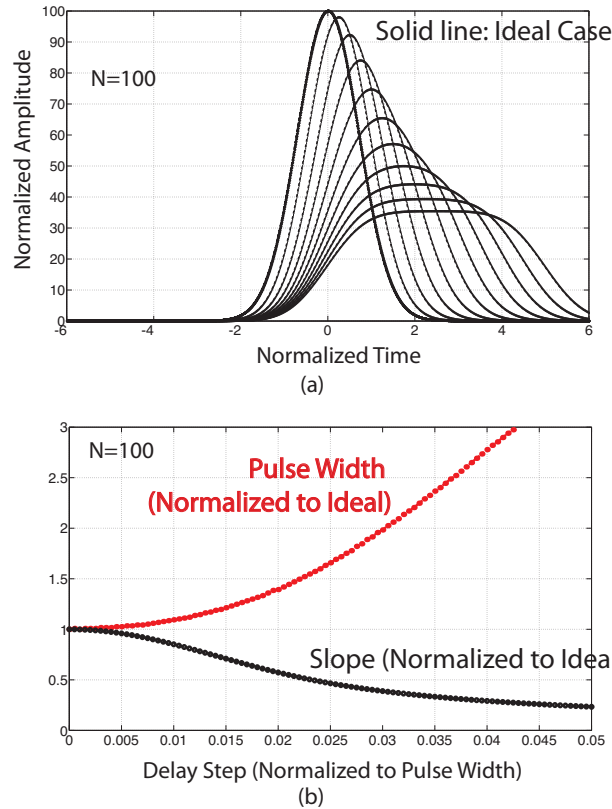


Figure 2.16: Effect of minimum delay resolution on array performance. (a) shows the increase in PW as well as change from ideal position. (b) shows the error in PW and slope (compared to ideal) versus minimum delay step. Number of array elements (N) is chosen to be 100 for simulations.

Neglecting the constant phase factors, we see that the output voltage (either of the quadrature outputs or the equivalent phase) is a periodic function of the distance. Therefore, by accurately measuring the phase (for example by measurement of both quadrature outputs), we can deduce the distance minus a phase ambiguity of or $n\lambda/2$. There are standard ways of eliminating the phase ambiguity. Frequency domain or time-domain techniques could be used. In the time-domain, as previously discussed, the pulse position ($A(t)$) determines the correct cycle of the carrier. In the frequency domain, multiple frequencies are transmitted and used to remove phase ambiguity. For example, frequencies $f_n = f_s + n\Delta f$ are used to illuminate the object. Both phase and amplitude of the reflections are monitored. In the current system, we have close to 10GHz of PLL lock range as well as pulse

positioning capability that could be used for this purpose. Frequency selection accuracy, phase noise and presence of random media can further complicate these algorithms.

2.11 Spatial Coding to Suppress Spatial Leakage

We propose a spatial coding scheme that can accommodate simultaneous focusing of the beam on multiple locations. Different segments of the two dimensional array concentrate their beams on different parts of the tissue and by that reduce the total scan time that is required for the FOV (Fig. 2.17). To alleviate leakage from adjacent cells, a spatial phase code is enforced on TX/RX on a region by region basis. For example, the phase of the carrier could be adjusted between $0^\circ/180^\circ$ on the TX. The RX in the correct region will also integrate with the right phase coding thereby suppressing the unwanted signals and enforcing the desired reflections. All other adjacent regions will suppress signals intended for this region since their code polarity is different. This technique is complicated by the nature of the object under test (phase of reflections).

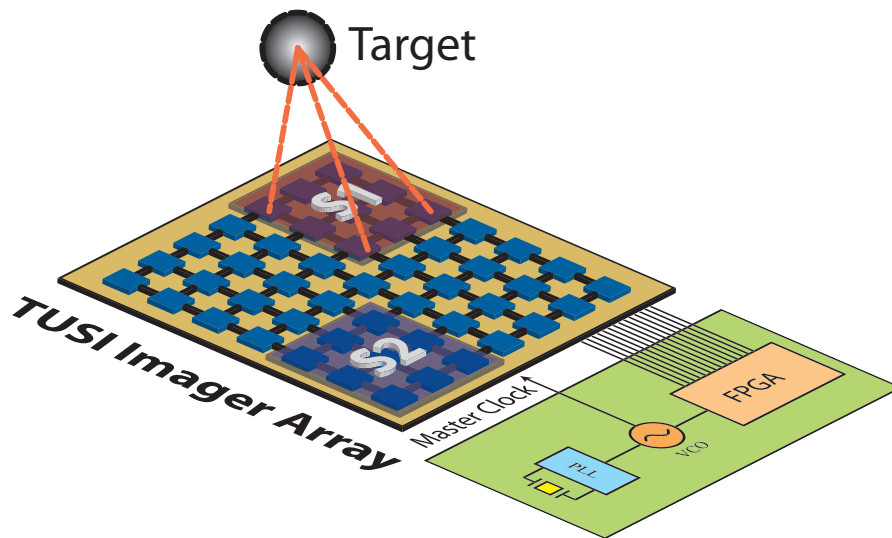


Figure 2.17: Spatial sectorization in the TUSI Imager. Each of the sectors can independently and simultaneously image their own tissue segments.

2.12 TUSI in the Large-Scale

The focus of this work is to design a pixel-scalable imaging array in the microwave to mm-wave part of the spectrum that can address a variety of applications. The first imager is envisioned as a portable and handheld device for localized medical imaging scenarios.

Some of the challenges and issues related to the portable imager were discussed. However, TUSI can also scale to larger areas for medical and non-medical applications.

2.12.1 TUSI for Intelligent Surfaces

The original TUSI system was designed for approximately an aperture of 10 cm by 10 cm. The size and spatial sampling (element spacing) depend on the specific application in terms of penetration depth, resolution and field-of-view (FOV). As previously pointed out, in many cases a sparse array (sub-sampled array) is required. Time-domain and pulse-based approaches can mitigate issues with ambiguous grating lobes in the array.

TUSI can also be used in larger arrays for 3D body imaging as well as positioning of people and objects. The applications of this large-scale imaging array for surfaces are in immersion applications, vital-signs monitoring, body temperature measurements, virtual meetings, gaming, and high-throughput and simultaneous data transfer. For example, hand gesture recognition based on radar imaging can be used for human-computer interfaces. 3D mapping of bodies can assist current 3D imaging techniques in the visible range to obtain finer details and/or reduce complexity.

Scaling TUSI

The scaling of TUSI as a pixel-scalable technology to these large-scale systems is described next. As the diagnostic tool, TUSI is designed for large averaging gains in “static” environments where movements are minimal. But the same system can be used to track object in a real-time scenario by trading off averaging for imaging speed. Often, the SNR levels are much larger in these applications and therefore, less averaging is required in the first place. The designed baseband incorporates flexibility in allowing various averaging ratios (by several orders of magnitude).

For example, a “TUSI tile” could have an aperture of 1 m by 1 m. We will compare this to the 94 GHz band in TUSI. Scaling up the aperture by a factor of 10 on each side (from 10 cm to 1m) results in a factor of 10 increase in range for the same lateral resolution (refer to section 2.3). This will lead to a range of 1-2 m. In many of the large-scale applications, the resolution requirements are less stringent. Also, the SNR is larger leading to the possibility of leveraging post-processing for resolution enhancement. Therefore, we can also scale down frequency to achieve better SNR and/or lower the power consumption. Scaling down the center frequency (and hence scaling up lateral resolution) by a factor of 3 will lead to a center frequency of around 30 GHz. The number of elements will also scale from approximately 100 elements in the 10 cm (94 GHz) array to 3,000 elements in the 1 m (30 GHz) array. Further sub-sampling can be used to obtain between 300-1,000 elements (for an area of 1 m²). This is shown in Fig. 2.18. Due to using sparse arrays, pulsed techniques as well as spatial filtering are used to remove or reduce ambiguous secondary illumination points in space (similar to grating lobes). This was previously explained in this chapter.

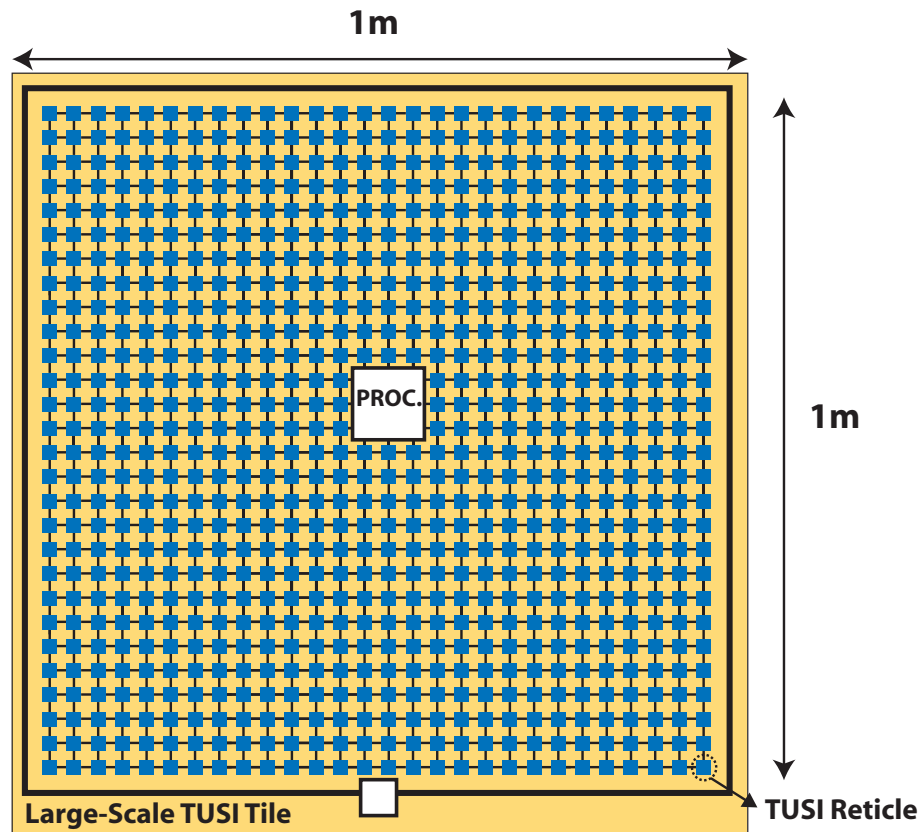


Figure 2.18: Large-scale TUSI array for intelligent surfaces. The array could be 1-2 m on each side or even larger. This figure shows an sparse-array of 1 m by 1 m which contains close to 1,000 TUSI elements. Energy transfer can use the large aperture of the system to recover or deliver energy (for example, as shown here, a large loop antenna can be placed around the imager for this and low-frequency communication purposes).

Depending on the size of the system, assembly/fabrication costs and power consumption are important aspects to consider. Wireless power transfer can be used to and from this system (depending on the availability of power sources). For integration on electronic devices (e.g. laptops), power is mostly available from the host device. In some scenarios, larger arrays can facilitate energy recovery from various external sources (vibrational, ambient or intentional RF, solar and etc) as well as being “wall-powered”. For example, as shown in Fig. 2.18, a large array for ceilings or walls could incorporate a large antenna (or a combination of smaller antennas) to harvest or to provide RF energy. Wireless powering is conceivable once the size of the loop is in the meter range [39]. In addition to radar imaging, other sensing and detection schemes as well as high-throughput communication with multiple and simultaneous beams can also be incorporated in this system. The system could have applications in office environments, digital information kiosks, health-centers

(patient monitoring), and places where seamless monitoring, tracking, and digital connection is important. In outdoor environments this could function as a low-cost radar platform (for example for safety and security applications).

To decrease cost and increase fabrication and assembly yields, smaller array segments can be designed as single modules and then multiples of these modules can cover larger surfaces. As an example, a 30 cm array with 100 elements can be the “unit element” to cover a given surface. A larger synthetic aperture is obtained by placing multiples of this element side by side.

Chapter 3

Antenna Design

In this chapter, design challenges related to on-chip integrated antennas will be discussed. Antenna efficiency and bandwidth will be examined. Several antenna implementations some of which incorporate switching functionality will be addressed. Antentronics is introduced as a merged design of antennas and electronics where the traditional boundaries do not exist. This methodology can be used to obtain the desired impulse response from the antenna structure.

3.1 Integrated Antennas

With the scaling of silicon technology and further increase in chip sizes especially for large scale transceivers, integration of antenna structures have been made possible for mm-wave circuits [2, 40, 41]. On-chip antennas will eliminate the need for complex, lossy interconnects to off-chip antennas and provide the ultimate level of integration. Integrated antennas eliminate the requirement for electrostatic discharge (ESD) protection circuitry on RF pads, a great source of signal loss and bandwidth reduction at mm-wave frequencies. However, the silicon substrate is not friendly to antenna integration due to several loss mechanisms that impact antenna characteristics in multiple ways. The two most important effects are the low resistivity ($10 \Omega\cdot\text{cm}$ compared to $10^7 - 10^9 \Omega\cdot\text{cm}$ for GaAs) and the high permittivity of the silicon substrate ($\epsilon=11.9$). The former introduces electric field losses which reduces the efficiency and radiation resistance of the antenna. Although high permittivity can be beneficial for some antenna designs, here, it leads to absorption of the fields in the substrate as well as undesirable substrate modes (for finite height of substrate) that reduce efficiency substantially [42] [43].

The lowest order substrate mode (TM₀) has a zero cutoff frequency. Substrate thinning will reduce the effect of substrate modes by eliminating higher order modes as well as suppressing TM₀ [42] [43]. Fig. 3.1 shows, to first order, the effect of substrate thinning on the peak broadside gain of elementary dipole and slot antennas. At low thicknesses, the

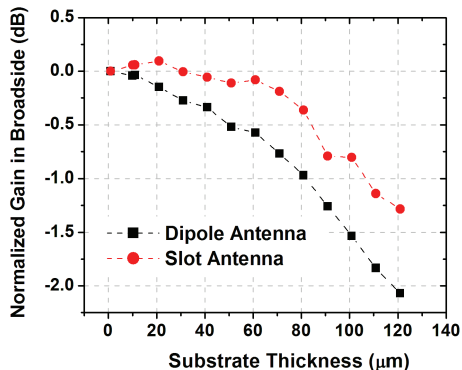


Figure 3.1: First order simulations of normalized broadside gain for elementary dipole and slot antennas for varying substrate thickness at 90 GHz. The slot is simulated with finite ground for closer resemblance to practical case. The antenna is on the top metal layer.

loss due to the high permittivity of substrate is minimized.

Several options have been pursued to improve antenna efficiency. Among them are use of Micro-electromechanical (MEMs) structures [44], substrate dielectric lens [2] [45], superstrate lens [46], and integrated dielectric resonators . The common problem with using these techniques is the extra complexity which negates the benefits of the integrated antennas.

As an example, the silicon dielectric lens provides an infinitely thick substrate. The two surfaces are no longer coplanar and hence the guided wave losses are mostly eliminated. With this method the antenna will radiate most of its energy towards the substrate (higher permittivity). There are, however, challenges in this method including cost, inferior beam patterns, requirement for large lens to reduce off-axis aberrations, and the fabrication of the matching layer as well as the required accuracy for placement of the lens [43] [45]. Also, often the silicon substrate is thinned down before adding the lens to reduce substrate losses. These steps add significant complexity and increase total integration costs. On the other hand, as previously mentioned, if wafer thinning is possible, this step by itself eliminates the problem of surface waves once the TM_0 mode is suppressed sufficiently. It has been shown that substrate thickness smaller than $0.04\lambda_d$ for slot and $0.01\lambda_d$ for dipole is adequate for acceptable low-loss operation [42]. However, substantial thinning (possibly down to 0.04λ or less) may be costly or challenging for mechanical reasons. The other problem with this method is that once thinned, the chip is still sensitive to material on the back side below the chip. The antenna radiates into both upper and lower hemispheres and care must be taken into choosing suitable material to physically support the chip.

3.1.1 Slot and Dipole Antennas

Slot and dipole are complementary antennas. The impedance of a slot antenna is related to its complementary counterpart by Babinet’s principle [34]

$$Z_1 Z_2 = (\eta/2)^2 \tag{3.1}$$

where Z_1 and Z_2 are the slot and dipole impedances, respectively, and η is the characteristic impedance in free space (377Ω). This relationship holds true for slot and dipoles in air and is an approximation if a dielectric substrate is introduced. In that case, an effective η has to be introduced to include the effect of the dielectric substrate.

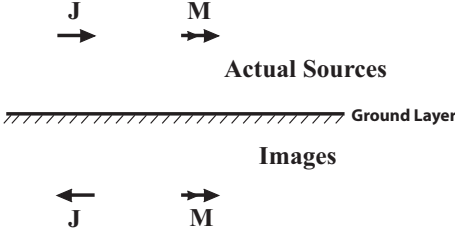


Figure 3.2: Electric and magnetic sources and their images from a ground plane.

Integrated dipoles and slots are briefly compared in terms of various loss mechanisms and other vulnerabilities. The first issue to be addressed is the substrate conductivity. Fig. 3.2 shows electric and magnetic dipoles with their images over an ideal ground plane [34]. For an electric dipole source (dipole antenna) sitting horizontally over the substrate (assuming the substrate is a fully conductive “ground” plane), the image is in the opposite direction to that of the dipole itself and hence an odd-mode excitation exists. The radiation impedance of the dipole is $Z = Z_{11} - Z_{12}$. In the limit of the ground plane being very close to the antenna, $Z_{12} = Z_{11}$ and the antenna is “shorted”. For a small slot antenna, to first order, the excitation is modeled by a fictitious magnetic source (since the electric current travels around the slot) and hence the conductive substrate does not directly lead to an odd-mode excitation. We should note that since the silicon substrate is not a perfect ground plane, this argument regarding the dipole is just an approximation to show the impedance drop effect.

The other issue is related to the high permittivity of the silicon substrate. A substrate dipole or slot antenna can be designed in several ways. Fig. 3.3 shows the different possible options. A dipole or a slot can sit on a substrate that may or may not be grounded on the other side. In addition to that, there are various possible options for the preferred radiation direction (direct to air or through substrate if not grounded). These options will suffer from the existence of substrate modes in different ways [47, 43, 42, 48].

In Fig. 3.3, option (a) will basically be a dielectric waveguide and will support various TE and TM modes. Specifically, this structure will have both TE₀ and TM₀ modes as the

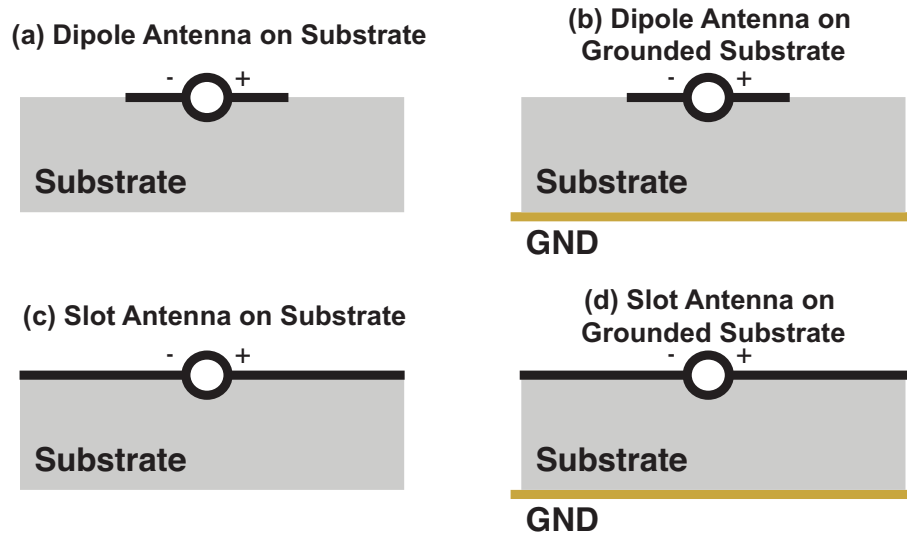


Figure 3.3: Various forms of integrated antennas on dielectric substrates.

substrate thickness goes to zero (there is no cutoff for these two modes). It will therefore have a more abrupt turning on of the substrate modes once the dielectric is introduced. This is observed in the analysis of [42] that the dipole antenna becomes susceptible to surface modes at a smaller substrate thickness compared to the slot antenna (case (c) in Fig 3.3). Fig. 3.1 shows a similar effect. In the slot structure (case (c)), one side of the substrate is covered with a metal plane and therefore TE₀ has a cutoff [47] [49]. TM₀, however, does not have a cutoff and will be present for small substrates.

A dipole on the ground plane (case (b) or a microstrip dipole) will have modes of a substrate with ground on one side and will be similar to case (c) in that regard. The TE₀ mode will have a cutoff. We can therefore conclude that, for elementary dipoles, the addition of the ground plane underneath the substrate structure can potentially lead to better gain and efficiency compared to a dipole on an ungrounded substrate. In the dipole without the ground plane, the turning on of the TE₀ mode is quite abrupt with substrate thickness [42], and if this is to be eliminated with the addition of the ground plane, there is an opportunity to have smaller substrate mode losses. This, of course, will require operation below the TE₀ cutoff of the substrate in the microstrip dipole. For elementary dipoles on grounded substrate, TE₀ will initiate when the thickness is approximately $0.3\lambda_d$ (the dielectric wavelength) and will be one of the dominant sources of loss for thickness around $0.4-0.6\lambda_d$.

Another important point is the direction of propagation for TM and TE modes from slots and dipoles. This is best seen from looking at the E and H field patterns of dipole and slot antennas (which are complementary). Fig. 3.4 shows the propagation directions of TE and TM surface-wave modes in a dipole structure. The slot has complementary E and H

fields and with that the TE and TM directions are interchanged. In a dipole, the TE mode is initiated from the broadside direction and TM will start from the two ends. In the slot, the opposite situation exists.

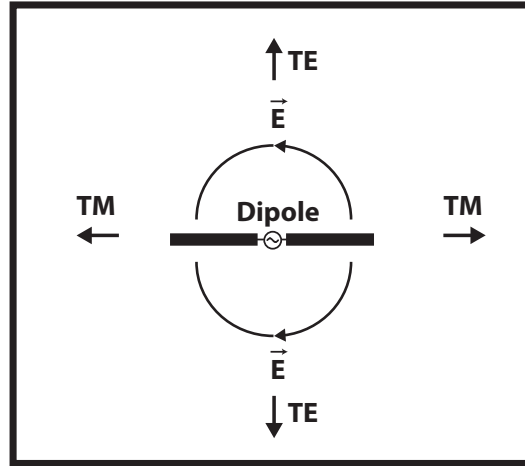


Figure 3.4: Directions of propagation of surface-waves for a dipole antenna.

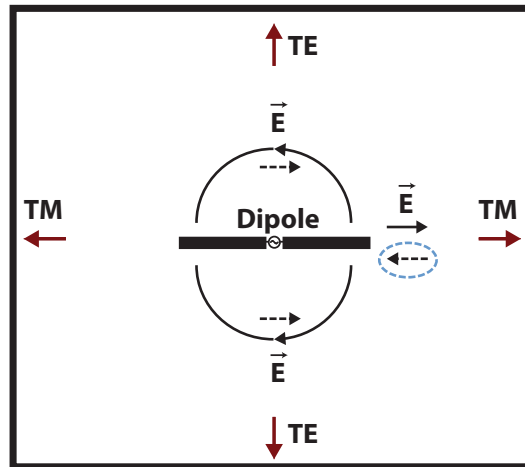


Figure 3.5: Direction of reflected E-fields in the dipole on grounded substrate.

The reflected field polarity and direction also play a part to explain the microstrip dipole efficiency [42]. The TE surface-wave mode propagates from the broadside direction of the dipole. The fields that reflect from the bottom ground plane will undergo a 180° phase shift. Fig. 3.5 shows the microstrip dipole and the associated E-fields in TM and TE mode directions. The reflected E-fields are shown in dashed lines. After reflection, some of the fields escape the substrate. These are radiated with angles that are smaller than the critical

angle [42]. The TE reflected rays are in the opposite direction to direct radiation fields. This leads to cancellation and hence reduction in radiation efficiency. The TM reflected rays (circled in Fig. 3.5), however, will mostly add in phase to the direct radiation field. TM₀ has a zero cutoff in this structure but the TE₀ does not and is avoidable for smaller substrate thicknesses. As the TE₀ is initiated, the efficiency drops rapidly.

The last case (d) is a slot antenna on a grounded substrate. This is basically a slot antenna in a parallel-plate waveguide [47]. In this case the TEM mode is excited and does not have a cutoff frequency. The losses for a thin parallel-plate waveguide grow as the substrate thickness is reduced. Here, a substrate thickness equal to half of the dielectric wavelength will lead to better efficiency due to presenting the slot antenna with a lower impedance (half-wavelength away from a “short”).

It is important to note that the previous arguments mostly ignore the substrate losses in silicon. This process has a resistivity in the range of 10 $\Omega\cdot\text{cm}$ to 20 $\Omega\cdot\text{cm}$. At 100 GHz, this translates to a large loss tangent (between 0.07 and 0.15) and the description of the radiation and loss mechanisms argued above need to take this into account.

Due to the tradeoff between various antenna topologies, several structures are used and compared. In the first antenna structure, we have adopted a slot design to accommodate substrate thinning for future integrated arrays. As previously mentioned, slot antennas have a more graceful substrate-mode initiation for thin substrate as compared to dipoles on (ungrounded) substrates. Our current measurements do not use additional substrate thinning and a thickness of 375 μm provided by the foundry is used. The slot antenna can also be used in a twin slot configuration to suppress the TM₀ mode [50]. This configuration has not been pursued here but provides an option for increased efficiency, especially for design with smaller bandwidths. Dipoles on grounded substrates are also explored and will be discussed in the next sections.

Another source of losses in integrated antennas arises from the low conductivity of thin metal layers available in silicon processes. Slot antennas, being predominantly covered by metal, have lower conductive losses (typically wider paths for current). Dipoles (especially thin wire planar dipoles) are susceptible to these losses. Also, in terms of pattern vulnerability, slot antennas are also less prone to the surrounding metals, complicated feedlines, and other conductive components especially if the large ground plane is placed on the top layer. With dipoles, since the surrounding environment (in the near field) is assumed to be depleted from conductive segments, metal routing or other circuit components can negatively affect the radiation pattern. A folded slot dipole is studied first. The impedance level of a folded slot dipole is better suited for the power amplifier whereas a folded dipoles radiation resistance is too high to be driven efficiently.

3.2 Antentronics

In this section we will describe antentronics as a merged design of antennas and electronics. The general concept is introduced and after that two antentronic structures are proposed for generation of short pulses.

3.2.1 Folded Slot Dipole Antentronic Structure

Previous research has demonstrated the effectiveness of modulating the properties of antenna for the desired characteristic [51, 52, 53, 54, 55]. This modulation of properties is used to achieve the desired antenna pattern or to perform direct data modulation on the antenna. For example, [52] uses parasitically switched reflectors to directly modulate the data on the continuous-wave signal from the antenna. CMOS switches are used to implement the shorting elements between the parasitic reflectors. This provides a high level of integration in the antenna modulated transmitter.

In [54] the authors use embedded diode functionality in the antenna to directly modulate the antenna with data. However, the switching speed is limited by the chosen architecture and the “discrete” nature of the experiment. In [55] authors use an array of patch antennas and switches implemented between the patches to generate the desired pattern. The reconfigurable aperture antenna uses discrete FET transistors to form electrical connection between the patch antennas. If the aperture reconfiguration is performed in a non-real time manner, then switch speed will not be a limiting factor. However, to obtain some of the more interesting dynamic effects in the antenna, it would be necessary to alter antenna’s characteristics at signal or modulation frequency.

Building on previous research, we realized the integrated antennas by embedding circuit elements within the physical structure of the antenna itself. These elements, working together with the antenna, could be used to manipulate the field and current distribution locally such that the desired transient or steady state response is obtained. In other words, we can engineer the impulse response of the antenna for the desired output. The specifications can include the required driving impedance, pulse response, radiation pattern or frequency spectrum. Thus, by using integrated electronic components, one can *dynamically* manipulate the operation and basic design parameters of the antenna. Many of the original antenna parameters have an inherent assumption of a single-feed radiating element. However, for example, if multiple feeds with various phase/amplitude of excitation are used, the antenna impedance parameter, for example, would need to be changed to an impedance matrix. The design space provided by this new breed of radiating elements, where the antennas are merged with electronics (Antentronics), is no longer limited by the constraints of the traditional antenna structures.

In this particular design, we use integrated transistors to perform antenna switching in time-domain to obtain narrow pulses. As shown in the Antentronic structure of Fig. 3.6, CMOS switches are integrated on the folded slot dipole antenna structure. The impulse

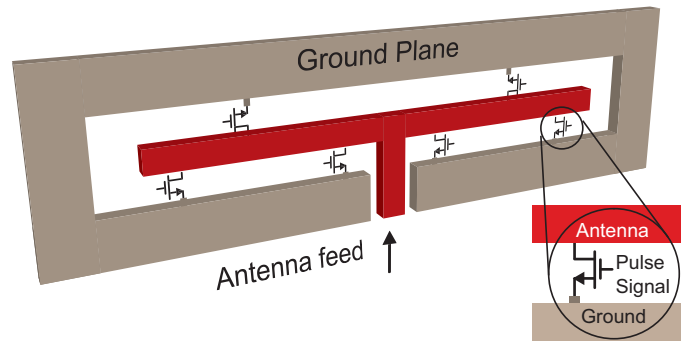


Figure 3.6: Antentronic structure with folded slot dipole antenna and synthesizable impulse response.

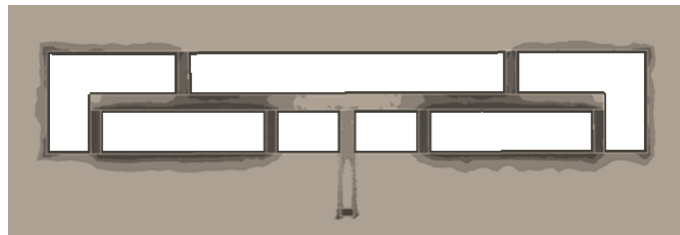


Figure 3.7: Current distribution (darker regions on the metal indicate higher current density) in the non-radiating mode (at specific point in cycle). Switch elements short the current to ground at multiple points and disturb the fields thus eliminating radiation.

response of the antenna is artificially altered to obtain the desired time-domain response. Here, the impulse response synthesis is aimed to extend the BW of the antenna but it could as well be employed to obtain spectral or pattern properties. These switches turn the antenna on/off depending on their conductivity.

Placement of distributed switches as opposed to a single large switch is advantageous in both the radiating mode and non-radiating mode. In the former, the switch capacitances are distributed along the antenna structure (similar to a distributed amplifier) and therefore their parasitic effects are mitigated. In other words, in the radiating mode, the capacitance of the switches does not limit the antenna transient response in the same way as a single large switch in which case the $R_{ON}C_{OFF}$ product determines the response.

The other aspect of the pulse generation is turning off the pulse and removing the energy from a “resonant” system. This happens in the transition to the non-radiating mode. Contrary to conventional techniques, as previously described, the antenna actually participates in this transition. In other words, the antenna is actively turned off as opposed to terminating the signal down the chain and waiting for the signal energy to dissipate in the antenna (a “passive” approach). Secondly, this turning off is achieved much more efficiently by distributing the transistors across the physical dimensions of the antenna. By reducing

the Q and providing shunt paths at multiple points (Fig. 3.7) the ON to OFF transition speeds up. With a single large switch, the antenna structure will have a “pulse-tail” due to the energy being dissipated over several cycles. At the same time, once the switches turn on and shunt the energy away from the antenna, the impedance (and hence efficiency) will drop significantly and the delivered PA power will also be considerably reduced further helping the ON to OFF transition. Together, these effects lead to a faster turn off for the antenna as well as a faster turn on due to smaller parasitics arising from smaller switch elements. Simulations of this “distributed antenna” show 21dB of static on/off ratio (combined with the PA, a larger on/off ratio can be realized). The equivalent on-impedance of the switches is 60Ω with 15fF (at center frequency) whereas in the off mode the shunt resistance goes up to around 600Ω (taking into account the large signal swing on the antenna structure).

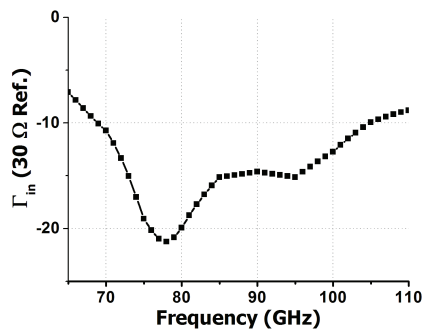


Figure 3.8: Simulated input impedance of the antenna in transmit mode.

Due to non-ideal drive strength provided to the gates of the NMOS switches, the large voltage swings in the antenna can couple to the gates of the NMOS switches. This will modulate the conductance of the devices and by that reduce the antenna efficiency in the transmit mode. For +15dBm of power on the antenna, this resistance drops from 600Ω to 300Ω affecting the radiation efficiency negatively. Input impedance simulation is shown in Fig. 3.8.

Another issue regarding the switching response is that in this structure, it is important for pulse signals to arrive on switching CMOS devices simultaneously. Signal distribution needs to take into account all routing parasitic loadings such that this condition is met. In addition to that, since there are large voltage swings, it is important to inhibit direct signal coupling from these feed-lines to the antenna structure. The large ground plane around the slot antenna is beneficial in this regard and can be used to reduce this direct coupling.

This antentronic structure is used in the voltage switching transmitter described in section 4.6.

3.2.2 Dual-Loop Antentronic Structure

Core Antenna Element

In the second antentronic design we have used a dual-loop antenna on a ground plane. The loop antennas are resonant and are realized in the top metal layer. The ground is placed underneath the chip. The radiation will be directed only to the top hemisphere and the sensitivity to the external support substrate is eliminated. This is an advantage for implementation of low-cost large-scale arrays.

To explain the radiation pattern and properties of a loop antenna, we start by a small elementary loop (Circumference (C) $\ll \lambda$). This antenna acts as a dual to an infinitesimal dipole [56]. This magnetic dipole will radiate in the plane of the loop and have a null at broadside ($\theta = 0$) (Fig. 3.9). For the imaging array, a broadside gain is required which a small loop is incapable to provide. As the loop size increases, there is considerable phase shift of the current around the loop. At $C \simeq \lambda$, the phase shift around half of the loop is $\sim \pi$. At this point, the two half-circles act similar to two electrical dipoles that are driven in phase (because of the phase shift) and are separated by a distance close to a diameter (D). This is shown in Fig. 3.10. The radiation pattern of this structure would then be the combination of the two dipoles and hence there will be radiation on broadside. Also, the total pattern is dependent on the distance between these two equivalent dipoles and is different from a simple dipole pattern. In case of a fully circular loop antenna, this effective distance is the diameter.

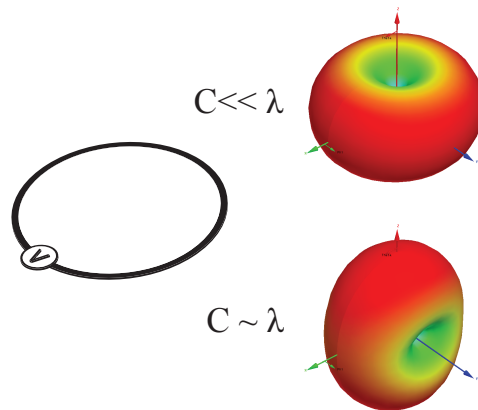


Figure 3.9: Loop antenna together with radiation pattern in cases of small loop (circumference $\ll \lambda$) and resonant loop (circumference $\sim \lambda$).

Next is the back-side reflector. A loop antenna will still be vulnerable not only to surface-wave modes in the substrate but also to the material on which the chip resides. Adding a reflector underneath the antenna will make the pattern unidirectional. Fig. 3.11 shows the directivity of a resonant loop on a ground plane that is 4λ on the side (there is no silicon

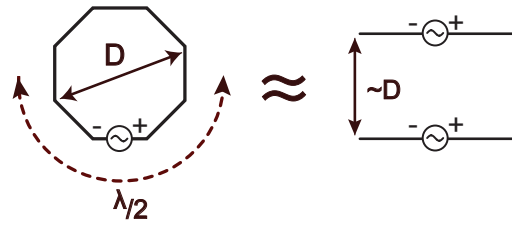


Figure 3.10: Resonant loop antenna equivalent to two half-wavelength dipoles.

substrate in this simulation). Substantial directivity results from this reflector. When the height is $\frac{\lambda}{2}$, images of these dipoles radiate an opposing signal and cancellation takes place on broadside (hence the directivity drops). The ground plane also affects the input impedance of the antenna (and hence potentially efficiency) through mutual impedances from the elements and the image components. As the distance is reduced ($H \ll \lambda$), the antenna input resistance drops significantly and it becomes impractical as an efficient radiator.

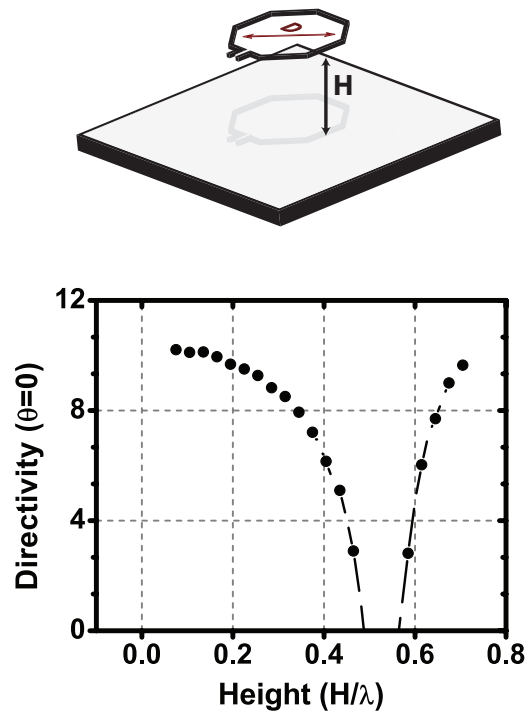


Figure 3.11: Loop antenna on ground plane with simulated broadside gain with varying the distance between the loop and the reflector. The ground size is assumed to be 4λ on each side.

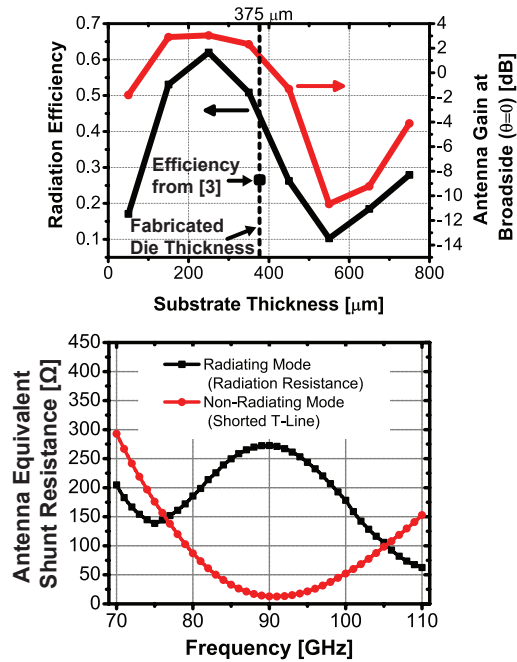


Figure 3.12: Radiation efficiency, broadside gain and input resistance of loop antenna with silicon substrate on top of ground reflector.

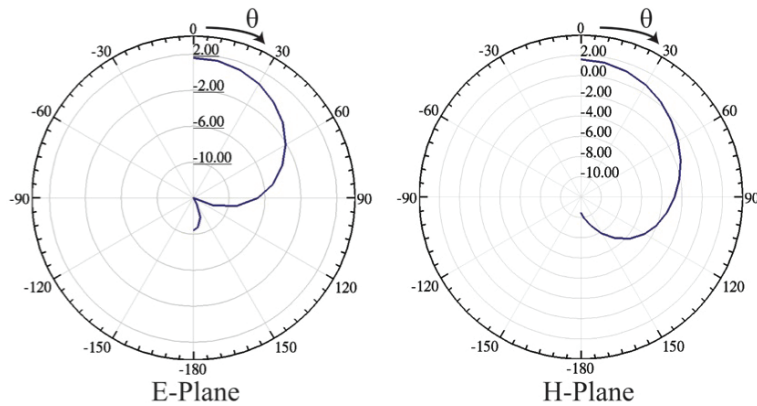


Figure 3.13: Simulated E and H-plane pattern of the loop antenna on silicon substrate ($375\mu\text{m}$) with finite reflector.

Fig. 3.12 shows the antenna broadside gain and radiation efficiency with the addition of silicon substrate. This is the simulation for the core antenna and does not include connection and matching losses. Also, Fig. 3.13 shows the E and H-plane patterns with the substrate thickness of $375\mu\text{m}$. It is clear from gain and efficiency simulations that there exists an

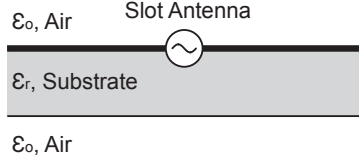


Figure 3.14: Slot antenna on a dielectric substrate.

optimal substrate height for which the efficiency is maximized. This optimum is close to $\frac{\lambda_g}{4}$, where λ_g is the guided wavelength in the substrate. This is to be expected since the dipole (effectively a current source) provides the maximum power when it is presented with an “open”. The approximate short circuit realized by the ground plane is transformed to an open with the quarter-wave substrate. The other interesting point is that compared to a slot antenna realized in [11], the efficiency is somewhat improved. This is partly due to the use of ground plane underneath the substrate. A slot antenna sitting on a substrate of permittivity equal to ϵ_r and thickness d will have a large portion of the power absorbed through the backside (some of which is turned into substrate modes). The ratio of the front to back radiated power from an slot antenna (voltage source) (Fig. 3.14), assuming no dissipation, can be approximated as:

$$\frac{P_{front}}{P_{back}} = \frac{Z_{back}}{Z_{front}} = \frac{Z_{back}}{Z_{air}} \quad (3.2)$$

In the special case of quarter-wave substrate we have:

$$\frac{P_{front}}{P_{back}} = \frac{Z_{sub}^2}{Z_{air}} = \frac{1}{\epsilon_r} \quad (3.3)$$

The grounded loop is unidirectional and does not suffer from the same effect and also, as discussed previously, does not excite TE₀ modes (as does a dipole with ungrounded substrate) and the TM₀ is not excited as strongly as in the slot for larger thickness substrates [42]. There is, however, another effect that helps the loop antenna be more efficient compared to the slot. Since the loop effectively acts as two dipoles spaced apart, one can optimize the dimensions and spacings to cancel specific surface-wave modes. This would be similar to a dual-slot structure as proposed in [50]. The difference here is that there is only a single feed and the structure automatically acts like two dipoles. The disadvantage is that the dominant TM₀ mode is primarily launched off the end of the dipole and the best cancellation is achieved when using end-to-end dipoles [50]. However, with the implemented $375\mu\text{m}$ substrate and at higher frequencies approaching 100 GHz, other non-TM₀ modes are also activated and cancellation can focus on these higher modes. In fact, for a dipole on a grounded substrate with medium thickness, a large portion of the power is lost to the

TE0 mode. TE0 is propagating broadside to the dipole and hence it can be canceled by using a dual-dipole structure (or equivalently an appropriately sized loop). This (partial) cancellation can only be applied to a single mode and if the thickness is larger such that other modes are dominating (e.g. TM1 or TE1), it will not be as effective. For cancellation, the lateral dimension of the loop needs to be close to one half of the guided wavelength of that particular mode. Here, we are using a octagonal loop and if better cancellation is required, the horizontal and vertical dimensions could be designed to be slightly different (elliptical shape). The other important point is that the size and position of the ground plane underneath the chip plays an important part in obtaining the increased efficiency.

Antenna Switching Network

As was briefly discussed before, the goal of the antentronic structures designed here are to provide pulse switching functionality in the antenna. In Chapter 4 we will discuss pulse generation mechanisms and the system design associated with initiating ps pulses. In this chapter, antenna network designs that incorporate switching and can generate and transmit high-frequency pulses are addressed.

In this second design, to accomplish switching, a dual-loop antenna is implemented. Two, symmetric, co-centric, independently driven loops are realized. The loops are driven by a current multiplexer (MUX) in this combined antenna-electronic (antentronic) structure (Fig. 3.15). The circuit details of the MUX and current switching scheme are available in Chapter 4.

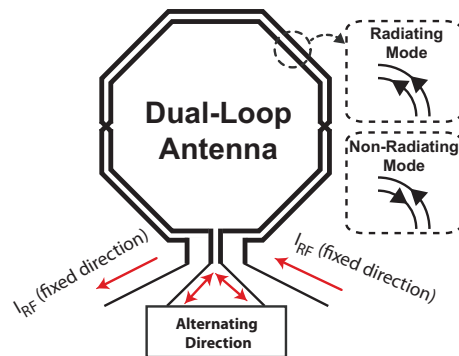


Figure 3.15: Dual-loop antentronic structure.

Through the use of the current MUX, the AC current of the outer loop antenna is kept at a single polarity while the inner loop's current is phase-inverted using a double-balanced Gilbert between radiating (Rad) and non-radiating (NR) modes. In the Rad mode, the two loops carry the exact same current and hence the dual loop antenna radiates. In the NR mode, the two loops have opposite polarity currents and the antenna turns into a transmission line driven with opposite polarities from the two sides (Fig. 3.16). This method

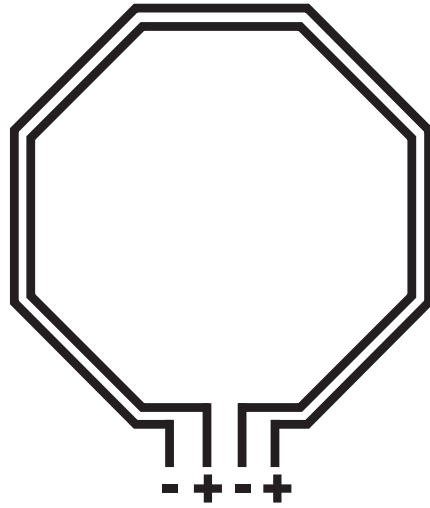


Figure 3.16: Dual loop antenna under non-radiating drive conditions. This resembles a transmission line.

of realizing a pulse is fundamentally different than conventional techniques in that instead of turning the pulser off by shorting the output of the device, we are essentially employing an *active cancellation* scheme. If designed properly, this active cancellation can take place faster than turning off the element. It will also be more effective since the propagating pulse gets canceled both in the near-field and far-field. In the near-field, since it is now a transmission line, the stored energy close to the element is reduced as compared to a large loop where the magnetic flux is not canceled. In the far-field, we can see this as two collocated antennas with alternate polarity where the residual transmitted pulse energy cancels. Simulations show the antenna efficiency dropping by almost two orders of magnitude from Rad to NR mode. The core efficiency goes from 47% to 0.5%.

In the NR mode, the input impedance seen into the network will be low (limited by losses). Fig. 3.12 shows the equivalent shunt input resistance of the dual-loop structure in Rad and NR modes (neglecting conductor losses). The impedance in the NR mode could be slightly changed with varying the distance between the antennas (which changes $Z_{0,diff}$) but remains low due to the specific topology and drive. The low impedance reduces the delivered power to the structure. This reduction in power improves the total ON/OFF ratio beyond the near-field/far-field cancellation effect.

3.3 Wideband Antennas

Wideband antennas are essential for the receiver section of the array. There are limits on the ability on removing antenna imperfections by post-processing and hence antenna bandwidth plays a key role in array performance. Larger bandwidth often comes with the price of lower efficiency. In addition to efficiency, spatial response of the antenna can suffer when wider bandwidths are required.

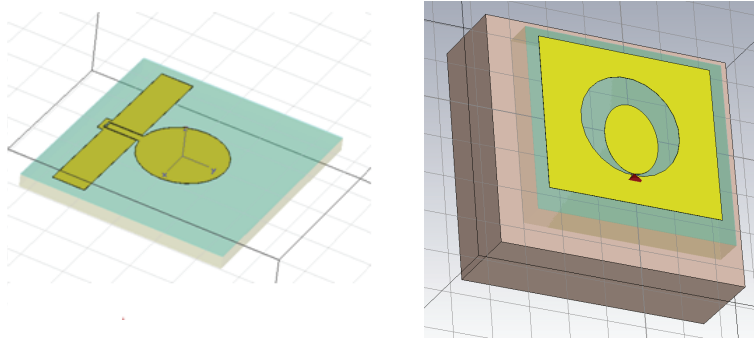


Figure 3.17: Some wideband antenna designs. A disc antenna (left) as well as a tapered slot loop antenna (right) are shown. The antenna on right could also be categorized in the coplanar path family.

Some examples of investigated antennas in the 94 GHz band are shown in Fig. 3.17. A disc antenna provides a very large bandwidth. Multiple disc elements can be placed in different sides of the common ground plane. The disc combines traveling and standing waves to achieve a large bandwidth. It, however, requires a large metal area and may not be suitable for arrays. Depending on the ground size used, the pattern and efficiency may also suffer in an on-chip version of this antenna. Another interesting option is the slot loop antenna (or equivalently a coplanar patch). It can be designed as a ring or as a square. It is basically the complementary version to a loop antenna. The slot loop works similar to a dual-slot antenna (described in previous sections) and can have a large efficiency on the silicon substrate (possibility for TM₀ mode cancellation). However, the bandwidth (close to 10 %) is not adequate for our applications. A wider gap can provide larger bandwidths at the cost of other complications. A tapered slot loop can provide the bandwidth as well as the efficiency. This is shown in Fig. 3.17. This antenna can provide a -10dB bandwidth from 65 GHz to 115 GHz. One problem with this antenna is the large metal area used in the disc as well as the ground plane. Density requirements necessitate using metal grids that can ultimately decrease conductivity and reduce efficiency.

For the TUSI project a tapered loop antenna with ground reflector was adopted. This is essentially a microstrip dipole antenna as previously described. The tapering increases the antenna bandwidth. The structure of the antenna for the 94 GHz band is shown in Fig. 3.18.

The antenna radiation pattern (antenna gain) is shown in Fig. 3.19. This antenna is used in the transceiver chip.

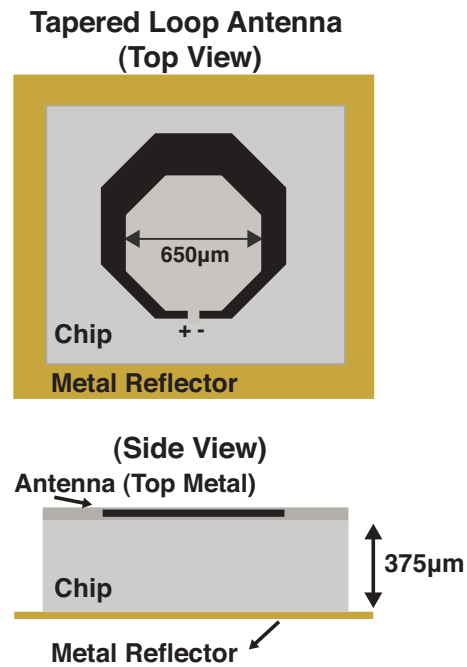


Figure 3.18: Tapered loop antenna on silicon substrate.

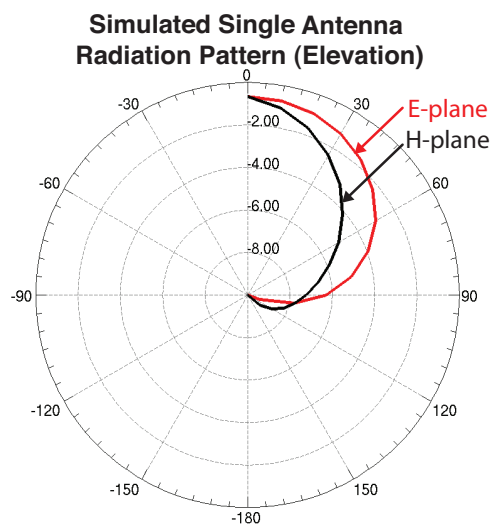


Figure 3.19: Single-element tapered loop antenna pattern. Antenna gain (dB) is shown in both E and H planes.

Chapter 4

TUSI Transmitters

In this section we will describe the design and measurements of two transmitter chips in a $0.13\ \mu\text{m}$ SiGe BiCMOS process. The two chips (Voltage Switching or VS and Current Switching or CS) have many similar components but differ in some critical parts. First the antentronic structures are fundamentally different (see chapter 3). The VS version uses a folded slot dipole antenna whereas the CS transmitter utilizes a dual-loop structure. As a result, pulse driver circuits see different loads and use different designs to provide adequate rise/fall times. Secondly, the power amplifier drives a differential antenna in one design and a single ended antenna in the other. This leads to different common-mode concerns with stability and otherwise. Moreover, the CS design incorporates additional functionality such as power tuning capability as well as a significantly higher PRF. In this chapter, the common areas of designs are addressed at the beginning of the chapter. Differences as well as experimental results are presented towards the end.

4.1 Process Overview

A $0.13\ \mu\text{m}$ SiGe BiCMOS process is used in this design. The technology node, particularly the frequency response, used for our prototype has been well documented in the literature (e.g. by Garcia et al. [57]).

Another process figure-of-merit for high-frequency designs is the breakdown voltage of the technology. This is especially important in design of circuit blocks with a large output power (e.g. drivers or the power amplifier). As will be described later, the breakdown voltage plays an important part in determining the maximum power per device. We can then combine multiple device outputs to provide a larger total power. However, this combining has limits in terms of losses as well as bandwidth and therefore a larger total transmit power will eventually dictate a larger per element power.

The background details regarding breakdown voltage in MOS and bipolar devices is provided in [58]. In a bipolar device, the two main metrics are collector-emitter breakdown

voltage with the base being open (BV_{CEO}) and the collector-base breakdown voltage with an open emitter (BV_{CBO}).

In the normal operation region of the device, the CB junction is reverse biased. Hence, the BV_{CBO} value reflects the reverse junction breakdown of the CB junction (in a common-base configuration). The common-emitter characteristic breakdown is a little more complex. This is because the generated electron-hole pairs in the avalanche process play a role in the base current. More specifically, the holes, in a npn device, contribute to the base current. This will lead to an amplified avalanche effect. Details of the analysis are provided in [58]. It is shown that the BV_{CEO} is substantially lower than BV_{CBO} . In this process $BV_{CBO}=5.5$ V and $BV_{CEO}=1.6$ V.

An important result from the breakdown analysis is that if the base impedance is reduced in the CE configuration, then the reversed current that flows on the base does not lead to a large voltage increase and effectively a larger CE breakdown voltage can be obtained. The simulated BV_{CEO} of a $5\mu\text{m}$ device is shown in Fig. 4.1. Biasing circuits that provide a lower “drive” impedance on the base are preferred to obtain a larger breakdown voltage.

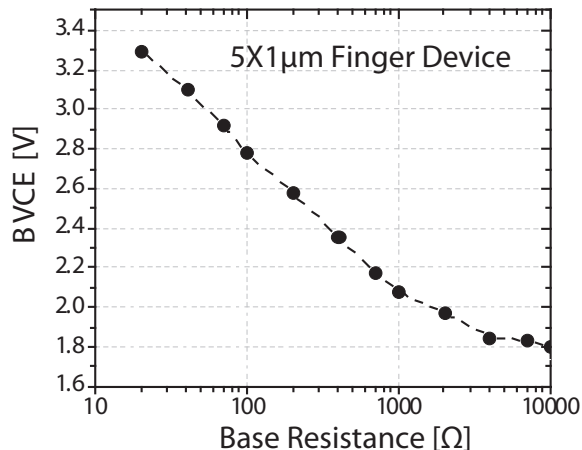


Figure 4.1: Collector-Emitter breakdown voltage of a $5\mu\text{m}$ device.

Another significant concern for design of high-frequency circuits in silicon is the passive losses incurred on the chip. In chapter 5 we will describe specific concerns related to the design of extremely wideband distributed amplifiers. Due to the specific available substrate, oxide and metal thicknesses, a microstrip line is preferable to a coplanar waveguide (CPW) in the SiGe BiCMOS process. For differential structures a dual microstrip line with or without side-walls can be employed. The dispersion shape as well as the losses are within acceptable limits for designs up to a few hundred gigahertz. Details of the transmission line losses in this process are outlined in the literature [57] [59].

4.2 Transmitter Architecture

The TUSI system aims for a programmable pulse generation from 350 ps to 25 ps with larger than 15 dBm of equivalent peak CW power at 90 GHz carrier. Obtaining a variable pulse width down to 25 ps results in bandwidths in excess of 40GHz around the carrier and this is a significant challenge. In time-domain, this translates to pulse widths that are only a few cycles long. The challenge is further exacerbated by the peak power and programmability requirements on the pulse.

To generate a narrow pulse on the mm-wave carrier, two distinct objectives need to be addressed. First we need to generate an accurate, narrow, and programmable pulse (or edges), and second we need a mechanism to modulate the carrier envelope.

4.2.1 Pulse Modulating the RF Carrier

Assuming a pulse (or edge) is created by the high speed baseband circuits, a switching function is needed to modulate the carrier. If this switch is placed early on in the path (Fig. 4.2a), then all stages following it will need to have the full bandwidth (BW) of the pulse modulated carrier signal which is in excess of 40 GHz in this design. The actual 3dB BW will have to be even larger to maintain constant group delay and pulse integrity especially if multiple stages are cascaded. This BW requirement will severely limit the attainable power and gain in the amplification stages. Distributed amplifiers can provide a large BW [60] [61] [62] [63] but are often limited in terms of their gain and maximum deliverable power. On the other hand if the switch is placed towards the end of the chain (Fig. 4.2b), power handling, loss, and BW of the switch driver are the limiting factors, especially if a series type switch is used. We push switching elements down the chain towards the antenna to preserve pulse integrity. Series switches are avoided due to their undesirable loss/BW tradeoff. The switching functionality is embedded in both the antenna (using the Antentronic architectures described chapter 3) as well as the final stage of the PA (as current commutation).

4.2.2 Hybrid Switching

Conceptually, to generate short pulses, one needs to have two complementary delayed edges and combine them to obtain a pulse. The pulse generation can be thought of as an ANDing of these two complementary signals. Generating a pulse early in the driver stages (Fig. 4.3a) will lead to amplitude reduction due to the limited BW of the stages that the pulse must travel through. Thus, the ANDing of the edges should take place as close to the PA and antenna as possible in order to preserve pulse integrity. One option is to combine the two edges as late as possible to generate a pulse and then send this pulse directly to either the PA or antenna (Fig. 4.3b). The other option is to use both the antenna and PA to realize the AND function by sending one edge to the PA and the other to the antenna

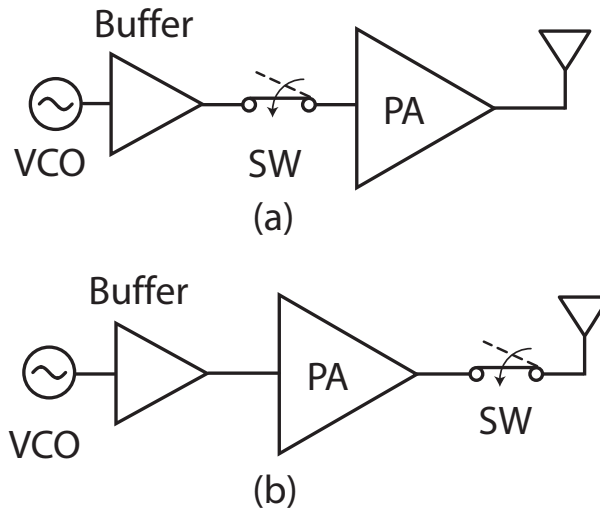


Figure 4.2: Conceptual switching options for pulse modulating the carrier. Switching early in the RF path (a) or towards the end of the path (b).

(Fig. 4.3c). The latter technique is the hybrid method proposed in this paper. In this approach, a full baseband pulse is never formed in the driving stages. The pulse-modulated carrier is formed by a collaboration between the PA and the antenna.

Intuitively, it is much more difficult to successively turn on and off a system with memory (i.e. capacitors and inductors) in a short period of time (relative to system time constants) than it is to perform two independent transitions. Generally speaking, this is due to the memory of the system; when you want to quickly perform two successive switches, during the second switch you must first counteract the initial momentum placed into the system. This requires more work/energy than a single switch. Hence, hybrid switching should outperform independent switching for narrow pulse generation.

One way to partially see the advantage is by considering the effects of driving the switching circuits of the PA or antenna. The switching circuits of the PA/antenna present a capacitive load to the driving stage. This driving stage will see some series resistance and inductance when driving this capacitor (Fig. 4.4). For independent switching, a narrow pulse signal is passed to this series RLC network whereas in hybrid switching, only a rising/falling edge (i.e. a step) is delivered. Due to the higher bandwidth of the narrow pulse signal, the output, which is the switching signal, will have more severe ringing and is thus potentially more prone to false pulse generation. Fig. 4.4 shows a plot of the responses of a series RLC tank to a step, a 25ps pulse, and a 15ps pulse, all of which have unit amplitude. For a given damping factor, the step function exhibits less ringing. Therefore, for a given capacitive load and designing for no false pulses, hybrid switching can meet specs with a smaller damping factor (i.e. higher parasitic inductance). In Fig. 4.4, we see that when the step response is on the onset of giving a false pulse (defined as 50% of final value), the

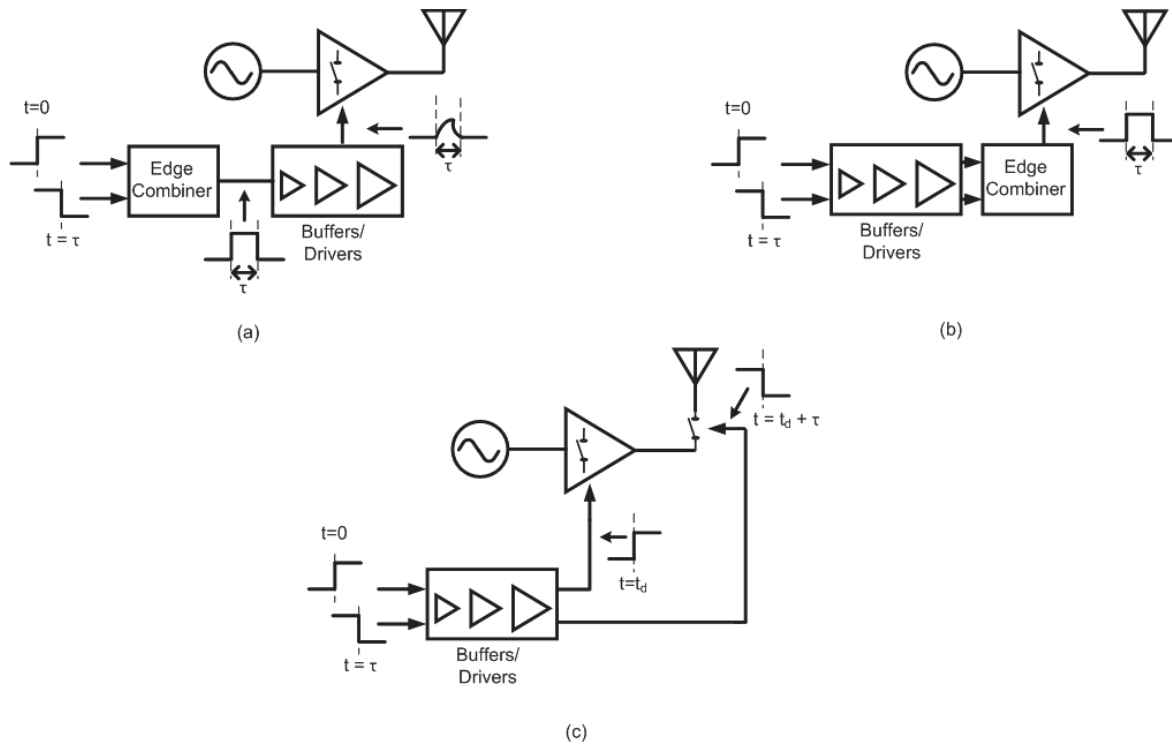


Figure 4.3: Pulse generation options. (a) Edge combining (ANDing) early in the chain leads to amplitude loss; (b) Edge combining as close to PA/Antenna in order to retain pulse integrity; (c) Hybrid switching.

pulse response gives a relatively strong false pulse. Secondly, the independent switching performance degrades at lower pulse widths as seen by the reduced amplitude of the main pulse and increased ratio of the false pulse energy to main pulse energy in the 15ps case versus the 25ps case. Such behavior leads to an echo problem and thus limits the detection sensitivity of the receiver.

To summarize, the hybrid switching technique has three main advantages. It pushes the ANDing functionality and pulse formation to the end of the chain and by that pushes the limits of pulse width and fidelity. It also avoids pushing a single element in the chain into ON/OFF states rapidly. Thirdly, it provides a better control over the programmability of the pulse width since here two edges are controlled independently. Being able to control the pulse width with finer steps will also help achieve the limit on the pulse width.

There is a subtle issue about the pulse generation. Since the RF path has a delay from the PA to the antenna, this will change the obtained pulse width and needs to be taken into account.

In order to obtain ultra-short pulses with maximum flexibility, the TUSI TX chip is designed with hybrid switching, independent switching and continuous-wave (CW) modes.

The block diagram of the TUSI TX chip is given in Fig. 4.5. Details of the individual blocks are described in the following section. Here, the antenna switching network is shown by a conceptual diagram.

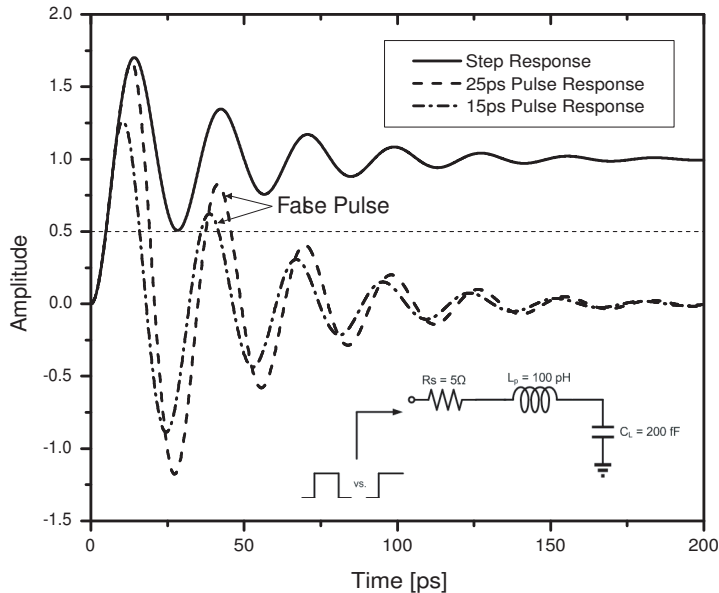


Figure 4.4: Input equivalent circuit used to model the load of the driver stages with step and pulse response of the PA/Antenna switching sections.

4.3 High-Speed Timing Circuitry

Given such a small pulse width (PW) requirement, emitter-coupled logic (ECL) [64] was chosen for the pulse generation, mode selection circuitry, and high-speed buffers that are used to drive the switching circuits of the PA and antenna. The high f_t of the bipolar devices enables ECL to provide high-speed, low-jitter operation at the cost of higher power consumption compared to CMOS. In our design, we use minimum emitter length devices to minimize power, $800 \mu\text{A}$ ($1.3 \text{ mA}/\mu\text{m}$) tail current to achieve the maximum f_t , and 500Ω loads to obtain 400 mV of output swing (Fig. 4.6).

At the start of the timing chain, the transmitter receives a clock signal (up to 3.45GHz for CS and 1.6GHz for VS) from an external source and uses that to set the pulse repetition frequency (PRF). Due to signal dispersion and losses from elements on board and especially bond-wires, a high speed clock receiver is required to regenerate signal edges and to maintain clock signal integrity. This block is placed close to the receiver pads. The high speed receiver drives the pulse-generation (PG) circuitry. The PG block generates and distributes pulses

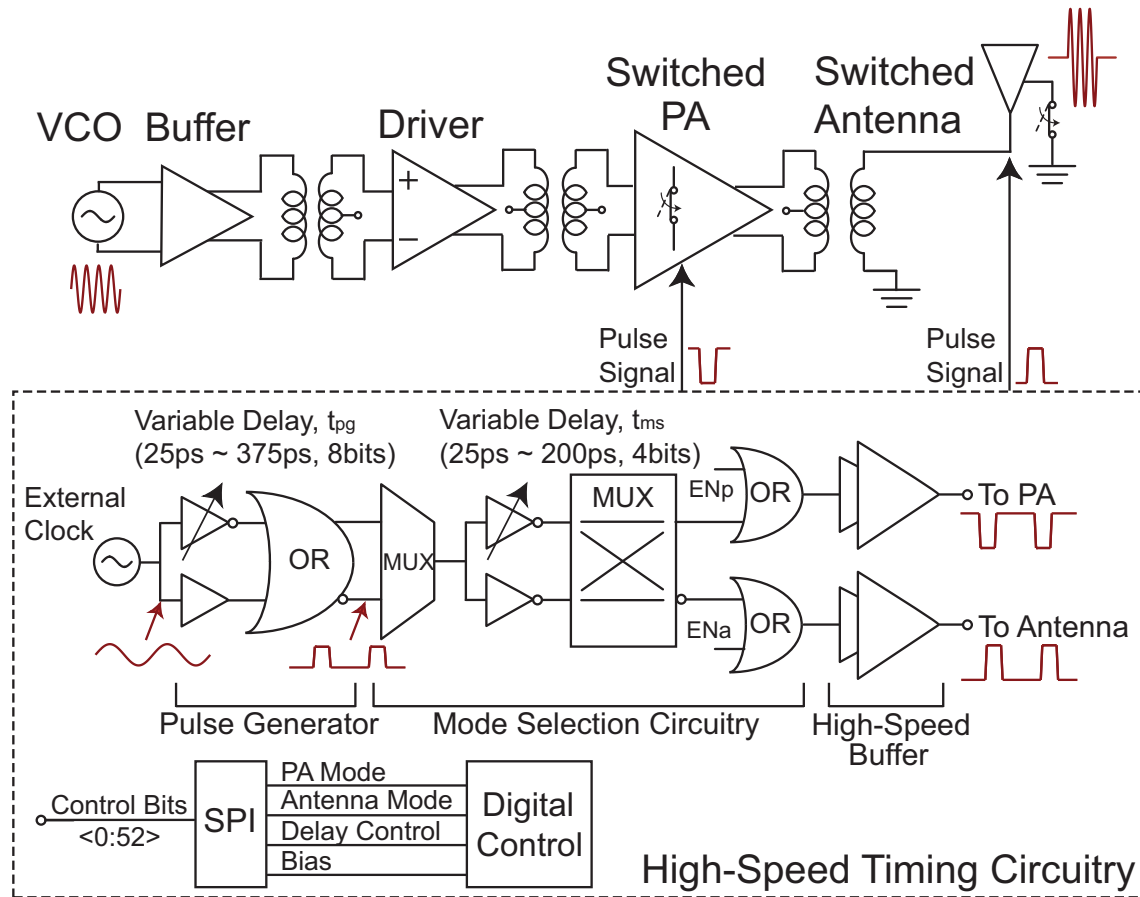


Figure 4.5: System block diagram of the TUSI transmitter.

or edges depending on the operation setting described above. Various sections of the timing circuitry will be described in the following sections.

4.3.1 Mode Selection Circuitry

The timing circuitry allows for several modes of operation (Fig. 4.7). The modes are: independent switching of the PA (Modes 5-6), independent switching of the antenna (Modes 7-8), hybrid switching (Modes 1-4), and no switching (Mode 9). In modes 1-2, the antenna switches OFF->ON (while PA is ON) to initiate the pulse, then the PA switches ON->OFF to end the pulse. The difference between these modes is the duty cycle of each switching signal. In mode 1, the antenna turn-on time is t_{pg} while in mode 2, the turn-off time is t_{pg} . In modes 3 and 4, the switching sequence is reversed; the PA switches OFF->ON (while antenna is ON) to initiate the pulse, then the antenna switches ON->OFF to end the pulse. In mode 5, notches are generated by turning off the PA during the short pulse width (t_{pg}).

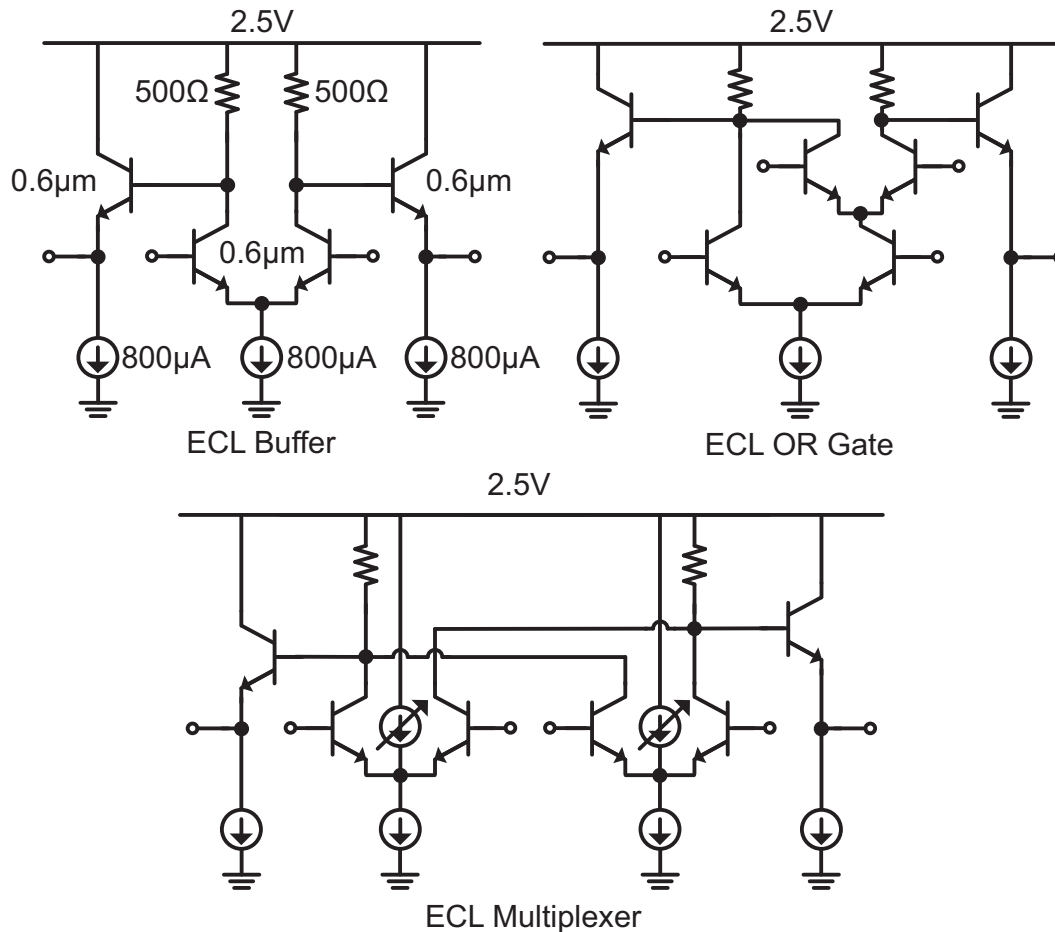


Figure 4.6: Schematic of ECL logic gates used in design of High-Speed Timing Circuitry.

In mode 6, pulses are generated by turning on the PA during the short pulse width (t_{pg}). In modes 7-8, the same operations are done by turning on/off the antenna. In mode 9, both the antenna and the PA are always on and not switched, enabling continuous transmission.

A block diagram of the mode selection circuitry is shown in Fig. 4.8. The OR gates located closest to the PA/antenna drivers allow for static enabling of the PA and/or antenna via two enable signals (ENp and ENa). In hybrid switching modes, both enable signals are low and variable-delay inverters are used to generate a delay difference between the critical transitions of the PA and ANT signals. This delay difference, t_{ms} , is digitally controlled from 25 to 200 ps and determines the output pulse width in hybrid switching modes. The combinations of the select signal of the two multiplexers generate 4 distinct modes of operation (Modes 1-4). For example, if instead of driving a “Direct” signal to the second MUX as in Fig. 4.8, we drive a “Switched” signal, the two output signals would be interchanged and we would change from Mode 1 to Mode 3. Similarly, by setting the first

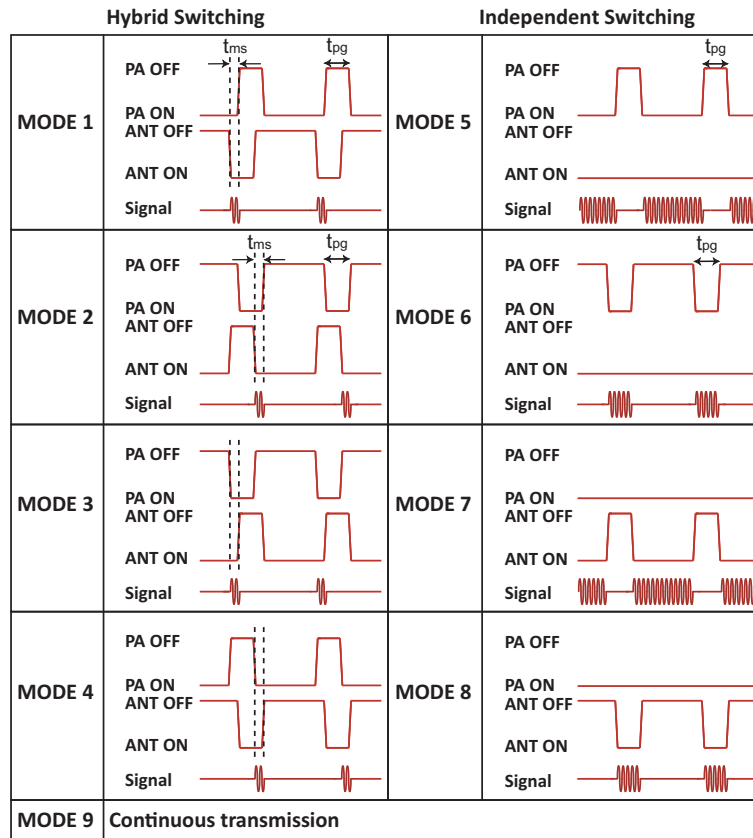


Figure 4.7: Operation modes of TUSI transmitter.

MUXs control signal to “Notch” instead of “Pulse”, we can alter between Modes 2 and 4.

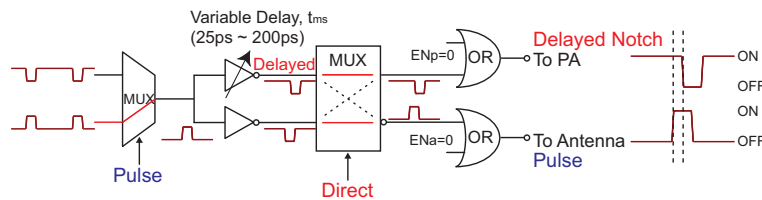


Figure 4.8: Block diagram of Mode Selection circuitry (programmed for Mode 1).

In independent switching modes, one of the enable signals (ENp or ENa) is held high, t_{ms} is set to zero, and the t_{pg} of the pulse generator determines the output pulse width. The pulse generator architecture is described next.

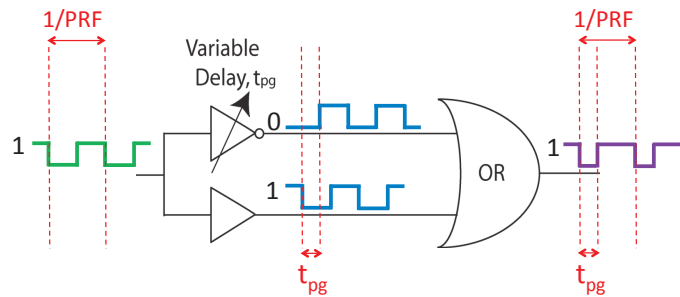


Figure 4.9: Simplified schematic of pulse generator.

4.3.2 Pulse Generator

A simplified schematic of the pulse generator is shown in Fig. 4.9. A clock signal is passed to two parallel paths, one having an inverting delay element, and both are fed into a two input OR gate. Pulse generation is triggered by a falling edge on the input clock. Thus, the pulse repetition frequency (PRF) is equal to frequency of the input clock. The PW is equal to the difference in delays between the two paths. Since the PW is determined by a difference in path delays as opposed to an absolute path delay, this architecture theoretically allows arbitrarily small PWs to be generated and is not limited by the gate delay of the process.

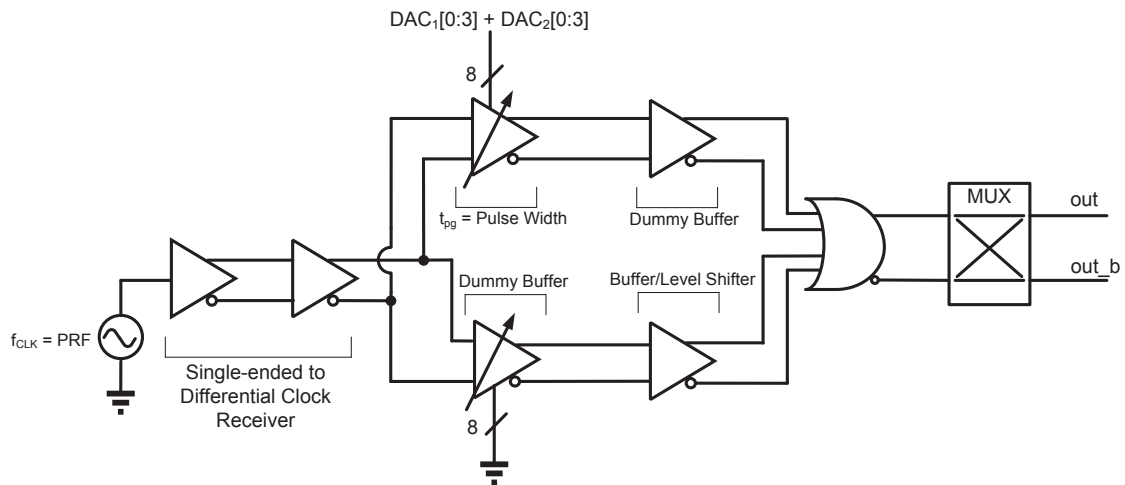


Figure 4.10: Detailed block diagram of pulse generator.

A detailed block diagram of the entire pulse generator is shown in Fig. 4.10. A single-ended to differential buffer is placed at the input to receive and sharpen the input clock signal

(from pads). The sharpened clock signal is passed to two parallel variable-delay differential buffer stages. Each buffer stage is comprised of two four-bit programmable-delay buffers. The buffer on the lower path serves as a dummy buffer (all 8 bits are tied to ground). This ensures that any difference in the path delays is primarily due to the programmable delay of the upper path. Without this dummy buffer, there would be a static difference between path delays for the case when all programmable bits are 0, thereby limiting the minimum achievable PW. Simulations show that without the dummy buffer, the minimum attainable PW is limited to ~ 35 ps. A buffer stage is placed before the OR gate to perform level shifting (emitter follower of ECL cell is removed) on one of the inputs to the OR gate. An output MUX is added at the output and acts as a butterfly switch for providing either a pulse or notch.

4.3.3 Programmable Delay

Two delay differences (t_{pg}, t_{ms}) are digitally controlled by switch capacitors. There are many options for obtaining variable delay in a buffer cell. Variable capacitive load was chosen for the ease of implementation (other methods, such as variable tail current, require external circuitry in order to maintain a constant V_{swing}). For a fixed tail current and load resistance, the ECL buffer delay is approximately

$$t_d = \frac{C_{var} \times V_{swing} \times \ln 2}{I_{tail}} \quad (4.1)$$

where C_{var} is the variable capacitive load, I_{tail} is the tail current, and $V_{swing} = I_{tail} \times R_L$. Thus, C_{var} provides a direct (and linear) handle on the delay of the buffer cell which translates into a direct (and linear) handle on the PW. A plot of the attainable pulse widths at the end of the chain (output of PA driver) is shown in Fig. 4.11. Although the pulse generator can deliver pulses below 25ps, as discussed in the previous section, these pulses do not propagate to the end of the chain due to the bandwidth limitation of the path. This effect is depicted in Fig. 4.12.

Although the simplicity of controlling buffer delay via variable capacitive loading makes it appealing for design, it does have some drawbacks. The first issue arises during hybrid pulse generation. In these modes, the pulse signal is passed to two parallel paths with a programmable delay difference, t_{ms} (Fig. 4.13). Ideally, each edge of the input pulse generates two edges whose transition times differ by t_{ms} . However, when $t_{pg}/3 < t_{ms}$, the second edge becomes susceptible to timing errors. Since t_{ms} is obtained by altering the output slope of one of the buffers, when $t_{pg}/3 < t_{ms}$, the output of the variable delay buffer will not settle properly in the time between rising and falling edges (given by the input pulse width, t_{pg}). Thus, the second edge no longer generates output edges with transition times differing by t_{ms} . Modes 2 and 4 suffer from this issue when $t_{pg}/3 < t_{ms}$. In our system this does not cause a problem since for narrow pulses we can easily ensure that t_{pg} is

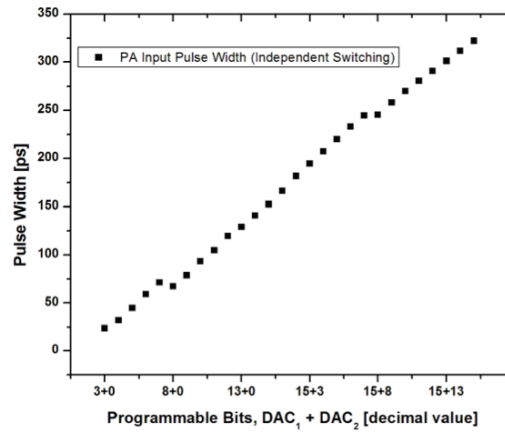


Figure 4.11: Plot of simulated pulse widths at the output of PA driver.

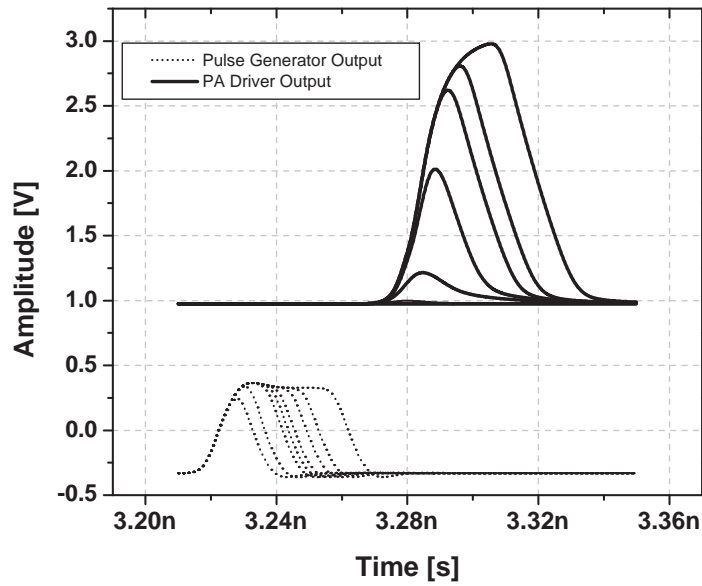


Figure 4.12: Transient simulation comparing pulse generator output signal to the signal arriving at the PA driver.

large enough compared to the pulse width. Large pulses could be generated in independent switching modes where this is not an issue altogether.

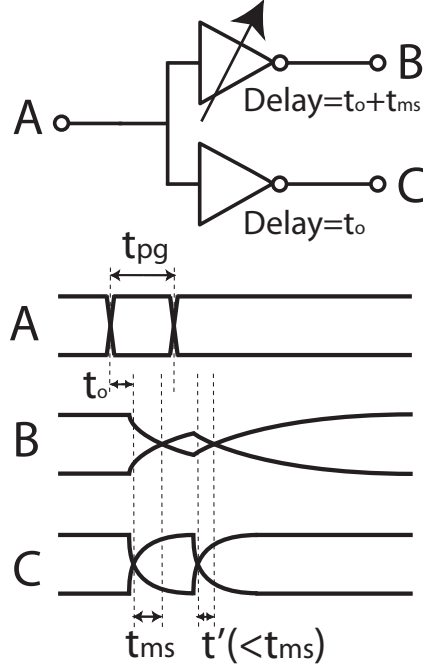


Figure 4.13: Illustration of timing uncertainty that occurs with specific settings of t_{pg} and t_{ms} .

4.3.4 Impact of Jitter in Pulse Generation

The integrity of the time at which a pulse is generated is crucial for high-resolution imagers. Determining the time of flight (TOF) of a transmitted pulse is crucial in any radar system. Assuming a perfect receiver (i.e., no jitter on sampling clock), any timing uncertainty during pulse generation will result in uncertainty in determining time of flight. Thus, for the TUSI system, the jitter performance of the pulse generator is of significant importance.

For the variable capacitor topology used, changes in delay are obtained by altering the slope of the output waveform. Since output jitter is inversely related to the slope at the zero crossing [65], large delay values (shallow slope) result in poor jitter performance. The standard deviation of the timing error (jitter) was derived for an ECL gate and is given by

$$\sigma_t = \sqrt{\frac{2kTC_{var}}{I_{tail}^2} + \frac{qV_{swing}C_{var}}{2I_{tail}^2} + \frac{qr_bC_{var}}{3I_{tail}}} \quad (4.2)$$

where k is Boltzmann's constant, T is temperature (in Kelvin), q is the elementary charge, and r_b is the effective base resistance of the bipolar device. From 4.2, we see that increasing the PW (increasing C_{var}) will result in an increase in timing uncertainty. A plot of expected

σ_t vs. PW is shown in Fig. 4.14. For this plot, σ_t is the simulated rms-jitter of a single variable-delay cell.

However, since the primary requirement for low jitter pulses is in the short pulse regime, this technique provides adequate accuracy. For longer pulses, due to lower BW, resolution is compromised in favor of signal level and jitter tolerances are higher.

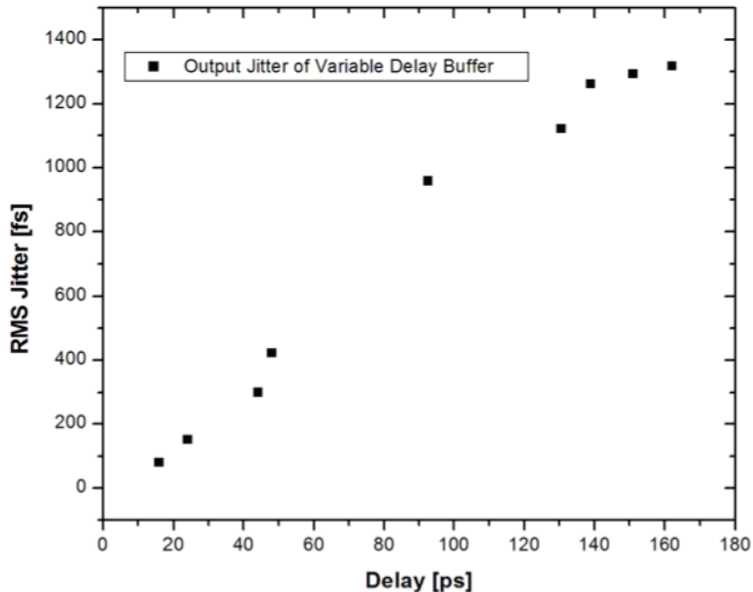


Figure 4.14: Simulated period jitter (rms) of variable delay buffer.

4.3.5 High-Speed Output Buffer

For a short pulse generation, driving the large capacitances from the PA or the antenna network poses a great challenge. In the case of the CS scheme, the PA presents a capacitive load close to 200 fF while the antenna network has a load closer to 100 fF. In the VS scheme, the antenna will present a larger load of 200 fF due to using CMOS devices as the switching elements. The bandwidth of the system needs to support adequately short rise time (t_r). For generation of pulses in the order of 50 ps or less, t_r of 10-20 ps is required for the drivers. In terms of small signal BW, this translates to $BW \simeq \frac{0.7}{t_r}$. For example, for a 20 ps rise-time, 35 GHz of bandwidth is required. At the same time, for the large-signal swing, adequate current has to be provided to drive the 200 fF load as well as any other parasitic capacitances. For 1V swing and 20 ps rise time, an average current of 10-15 mA needs to be provided to the capacitor. To meet the fast transition time as well as the large output current requirements, an emitter follower structure is employed (Fig. 4.15).

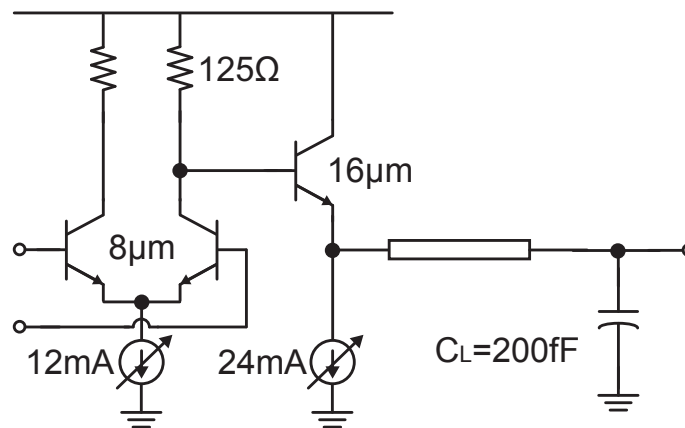


Figure 4.15: Circuit schematic of the high-speed buffer driving the PA and the antenna network.

Swing Requirements

Depending on the switching scheme, different swing requirements are to be met. For example, in the CS scheme where the current is steered between two different states using bipolar devices (the scheme is described later in this chapter), a smaller swing is required compared to the VS scheme where CMOS conductance is modulated to perform RF switching. In the case of CMOS devices, the capacitive load is larger and a larger swing is also required. This reduces the ON/OFF ratio of this scheme.

For the VS scheme, a 1.4-1.5 V maximum output swing is deliverable to turn the load ON/OFF. Emitter followers are used as mentioned previously. When driving the antenna CMOS switches, an optional negative voltage (-0.4 V) could be applied at the emitter of the tail current source in order to drive the CMOS gates closer to 0V and thus reduce the CMOS leakage current.

In the CS scheme, the swing is again programmable and can be as high as 1.4 V. However, in measurements, often a smaller swing is selected to reduce transition times. The minimum swing to have an acceptable ON/OFF ratio is about 0.4-0.6 V.

High-Speed Challenges

For high speed and low jitter switching, fast rise/fall times are desired. For our design, we aimed for a nominal rise/fall time of 40ps (0-to-90% time) for the maximum swing setting. For smaller swings this can be reduced (through bias settings and changing drive amplitude of the pulse). Due to the minimum-sized stages used in the preceding blocks, intermediate buffers are needed. Each successive stage's current is sized up by a factor of 4 and the device sizes are scaled accordingly. The final differential amplifier and emitter follower are shown

in Fig. 4.15. The amplifier stage uses 8 μm emitter length device, 125 Ω load resistors, and the nominal tail current of 12 mA (1.5 mA/ μm), resulting in an output swing as large as 1.5 V. The emitter follower of the output buffer uses 24 mA (1.5mA/ μm) current and its emitter length is 16 μm . The currents are programmable by a 4 bit digital-to-analog converter (DAC).

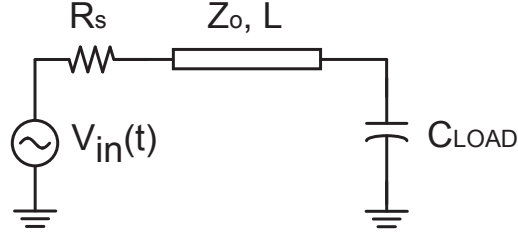


Figure 4.16: Simplified circuit model for buffer circuit including the trace.

While very fast, the emitter follower may have stability problems when driving a capacitive load. Depending on operating conditions, the output of the follower can appear inductive [58]. Combined with a capacitive load, the output signal will then be susceptible to ringing. To make matters worse, the physical distance between the output buffer and switching circuits results in a non-negligible parasitic inductance being placed in series with the capacitive load (Fig. 4.16) which makes the output of the follower appear more inductive. Although ringing in the response can reduce the rise/fall time, large peaking can make the transistors fall into the high-current region and increase the output jitter. In addition, large ringing can cause the succeeding stage to repeatedly switch on/off before settling, making the switching time ambiguous by generating false pulses or echoes.

At the end of the buffer, there is a finite length of a trace that connects to the PA or the antenna network. The line is terminated by a dominantly capacitive load. The model is shown in Fig. 4.16. In the case of the matched driver (i.e. $R_s = Z_0$), for an applied input step, the output capacitor voltage charges to the final value of $V_m = V_{source}$ with:

$$V_L(t) = V_m \{1 - e^{-(t-t_d)/\tau}\} \quad t \geq t_d \quad (4.3)$$

where $t_d = l/v$ with v the wave velocity in the line and $\tau = Z_0 C_L$. If the driver is not matched to the line, there will be additional reflections from the source side. In the extreme case of ideal driver ($R_s = 0\Omega$), capacitive load and a lossless line, the output voltage starts to toggle between $2V_m$ and 0 with a period of $2t_d$. This is due to the total reflection at the source with $\Gamma_s = -1$. This situation is equivalent to the severely underdamped response of a second order circuit.

The case we are interested is not an ideal step response but rather the response to a step with a finite rise time. We will use an exponential step (i.e. $V_{in}(t) = V_0(1 - e^{-t/\tau_0})$) to

model the input finite rise time. With this input voltage the output of a capacitively loaded transmission line follows:

$$V_L(t) = V_0 + V_0 \left(\frac{-\tau}{\tau - \tau_0} \right) \{ e^{-(t-t_d)/\tau} \} + V_0 \left(\frac{\tau_0}{\tau - \tau_0} \right) \{ e^{-(t-t_d)/\tau_0} \} \quad t \geq t_d, \quad \tau \neq \tau_0 \quad (4.4)$$

This simplifies back to (4.3) when $\tau_0 = 0$. It also simplifies to a single time-constant exponentially approaching step when one of τ or τ_0 dominate the other. Fig. 4.17 shows the results.

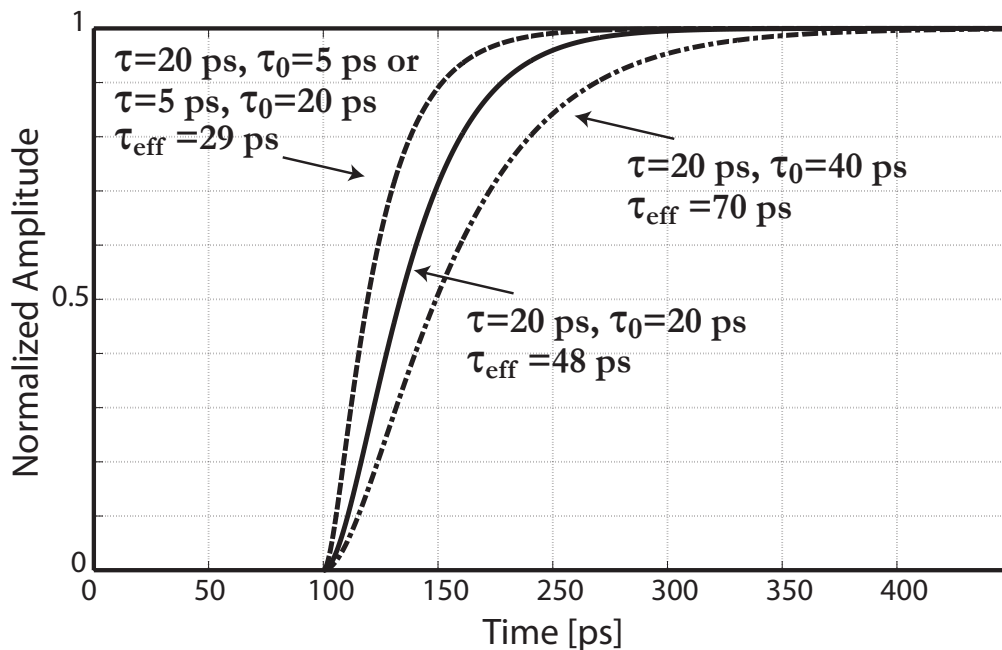


Figure 4.17: Response of a capacitively terminated transmission line to an exponentially approaching step function. Line delay (t_d) is 100ps. The transmission line is assumed matched on the source side.

Due to limitations of the ECL follower driver response in small and large signal scenarios, matching the driver side will pose challenges. Also, to decrease the step rise time, we can design the circuit to be slightly underdamped. To avoid excessive ringing that might lead to “echo” pulses, care must be taken in the design. To simplify the analysis we observe that for practical drivers in this technology process ($f_t=230\text{GHz}$), the rise time is typically much longer than the delay incurred in the trace transmission line leading to the antenna or the PA and we can essentially approximate the line with the equivalent inductance (i.e. $L \simeq Z_0 t_d$ [47]). In that case, we are looking at the response of a second order RLC circuit to a ramp or exponentially approaching step input. For this network, a smaller driver

impedance leads to larger Q factor and hence faster rise time but with more severe ringing. Due to the design of the switching network, we primarily care about a fast rise time as long as the ringing or more importantly the undershoot voltage value remains higher than a specific “threshold”. Fig. 4.18(a) shows the response of the second-order system to an exponentially approaching step input. Due to a non-ideal drive (exponential input signal), the output signal can experience ringing prior to reaching its final value. This is in contrast to a step-response. Care must be taken in the analysis due to this effect. Fig. 4.18(b) shows the contours of constant undershoot (in percent from final value) in a system with a fixed capacitor load (200 fF) and with an exponential input ($t_r=22$ ps) under different trace inductance and source resistance values. Given that a 15% undershoot value is acceptable, this figure can be used to select acceptable L/R regions. As shown in Fig. 4.18(c), the time it takes to reach 80% of the final value generally increases with increasing the damping factor or source resistance.

4.4 Quadrature Voltage Controlled Oscillator (QVCO)

A 94-GHz clock source is generated by an on-chip VCO using a differential Colpitts architecture, which is commonly used for millimeter-wave applications [66] [67]¹. Compared to the cross-coupled VCO, the differential Colpitts VCO offers better phase-noise performance and inherent buffering which mitigates the large capacitive loading from the output buffer. Fig. 4.19 shows the simplified schematic of the Colpitts VCO. Neglecting C_{bc} , the impedance seen into the base terminal is

$$Z_{in0}(s) = \frac{g_m}{s^2 C_1 C_2} + \frac{1}{s C_{in0}} \quad (4.5)$$

where

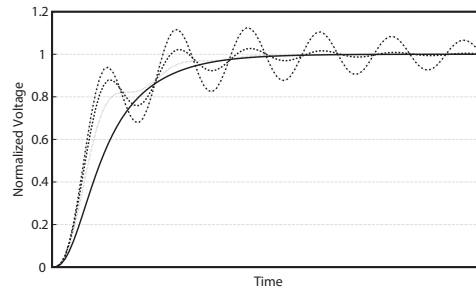
$$C_{in0} = \frac{C_1 C_2}{C_1 + C_2} \quad (4.6)$$

Note that C_1 includes both C_{be} and any external capacitance. However, the existence of C_{bc} not only increases the input capacitance but also reduces the amount of negative resistance generated. This is because the current that flows through C_{bc} is not sensed by the transistor and is “wasted”. The input impedance with C_{bc} is

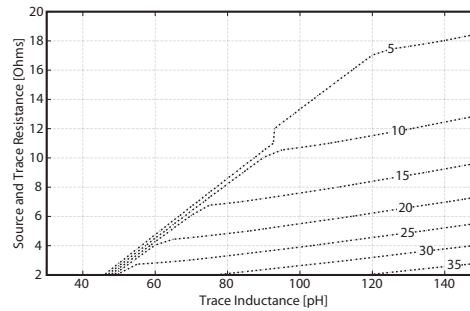
$$Z_{in}(s) = Z_{in0}(s) \frac{1}{\left(1 + \frac{C_{bc}}{C_{in0}} + \frac{g_m C_{bc}}{s C_1 C_2}\right)} \quad (4.7)$$

With $C_{bc} \sim 0.5 C_{in0}$, the negative resistance is lowered by a factor of 2.2 at 94 GHz. The situation is worse if the outputs of the VCO are taken from the collectors since Miller effect increases the effective capacitance.

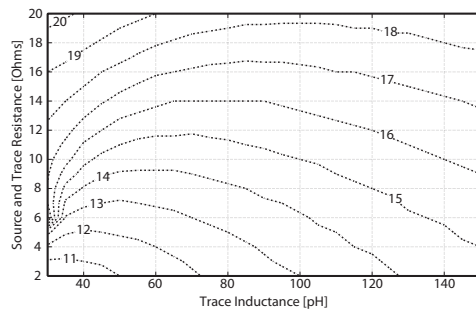
¹Many thanks to Jun-Chau Chien for collaboration in the TUSI project and in the design of VCO



(a) Response of a second order circuit to an exponential step input.



(b) Contours of undershoot value (in percent from final value) in the case of driver circuit. The input is an exponential step with time constant τ_0 .



(c) Contours of 80% rise-time values (in ps) for the driver circuit.

Figure 4.18: Design curves for pulse driver circuit parameters. Input is an exponential step with a rise-time (10-90%) of 22ps.

To improve negative resistance, transformer-coupling [68] is utilized by cross-coupling the differential inductors at the bases and the collectors, as shown in Fig. 4.20. Such a technique has two benefits: (1) the arrangement of the coupling leads to additional positive

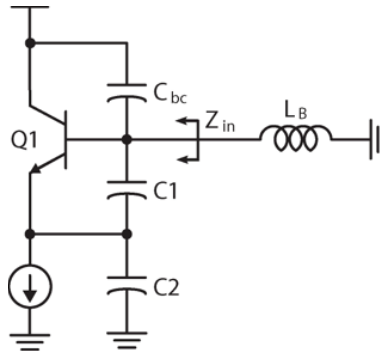


Figure 4.19: Simplified schematic of the Colpitts oscillator.

feedback which increases the loop gain while the capacitance between the two windings acts as a negative Miller capacitance; (2) the mutual inductance improves the area efficiency and facilitates the implementation and layout of coupled VCO for quadrature generation.

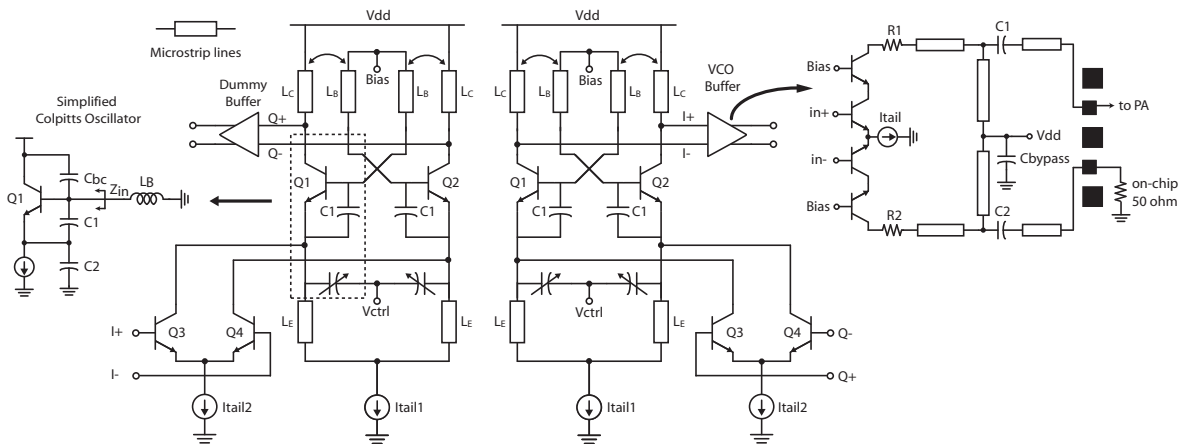


Figure 4.20: Complete schematic of QVCO.

To enable future integration with the receiver, quadrature outputs are generated even though only one differential output is required in the transmitter. Anti-phase coupling is realized by sensing the oscillation signals from the collectors of one VCO and feeding these signals to the emitters of the other [69]. More details on the design of the oscillator are available in [11]. Frequency tuning is achieved by introducing dual-oxide AMOS varactors in the capacitive feedback network. With a channel length of $0.85 \mu\text{m}$, the varactors show a $\frac{C_{max}}{C_{min}}$ ratio of 2.7 and Q_{min} of 1.2 at 94 GHz. The overall tank quality factor varies from 3.1 to 10.6 over the tuning range.

Fig. 4.21 shows the schematic of the cascode buffers with the output matching network. The outputs of the buffers are matched to 50Ω with microstrip lines. Two resistors are

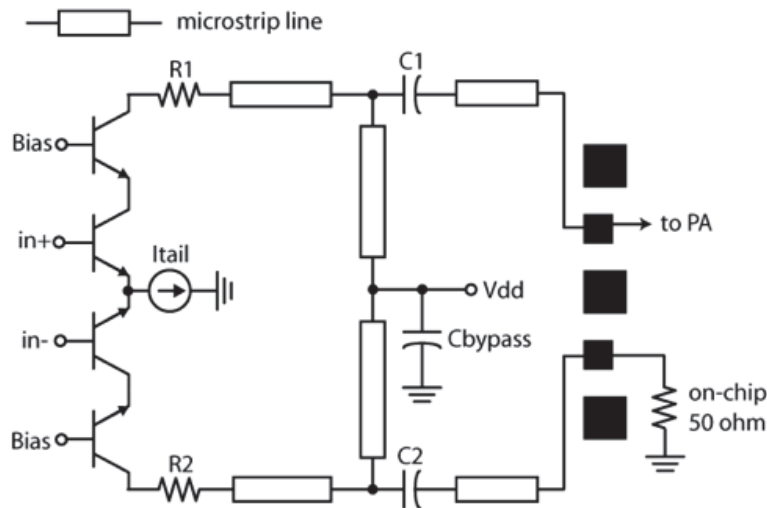


Figure 4.21: Schematic of the output buffer.

added at the collectors to improve stability and bandwidth. Simulations show that the buffer is able to deliver more than 0 dBm of output power with a tail current of 8 mA and a bandwidth of 20 GHz.

Note that Q_{VCO} is susceptible to injection-pulling from both the PA and antenna due to its low Q_{tank} . Fortunately, the ground of microstrip lines provides first-order shielding from the substrate. Nevertheless, attention must be paid in the layout to prevent coupling through the ground plane.

Fig. 4.22 shows the measured tuning curve of a separate VCO test circuit. The measured frequency tuning range of the VCO is 74 - 89 GHz while delivering an average output power of -4 dBm. The measured phase noise at 74 GHz (with minimum $Q_{varactor}$) is -104 dBc/Hz at 10 MHz offset, as shown in Fig. 4.22. Note that in the measurement, bimodal oscillation is not observed. However, the oscillation frequency is lower compared to simulation. This is mainly due to the modeling errors in the varactors. In the system measurement, the bias current of the coupling transistors are changed using the integrated programmable 4-bit DACs to increase the oscillation frequency such that it matches the design of PA and antenna. This problem with the modeling is addressed in the design of the final VCO/PLL in the transceiver (chapter 6).

4.5 Switched Power Amplifier

The power amplifier needs to provide $\sim +20$ dBm of saturated output power and is therefore handling large voltages/currents. Technology breakdown voltages for the bipolar device in 0.13 μ m SiGe process are: $BV_{CEO} = 1.6V$ and $BV_{CBO} = 5.5V$. BV_{CEO} is the collector to emitter breakdown voltage when the base node is open [58]. As long as the impedance

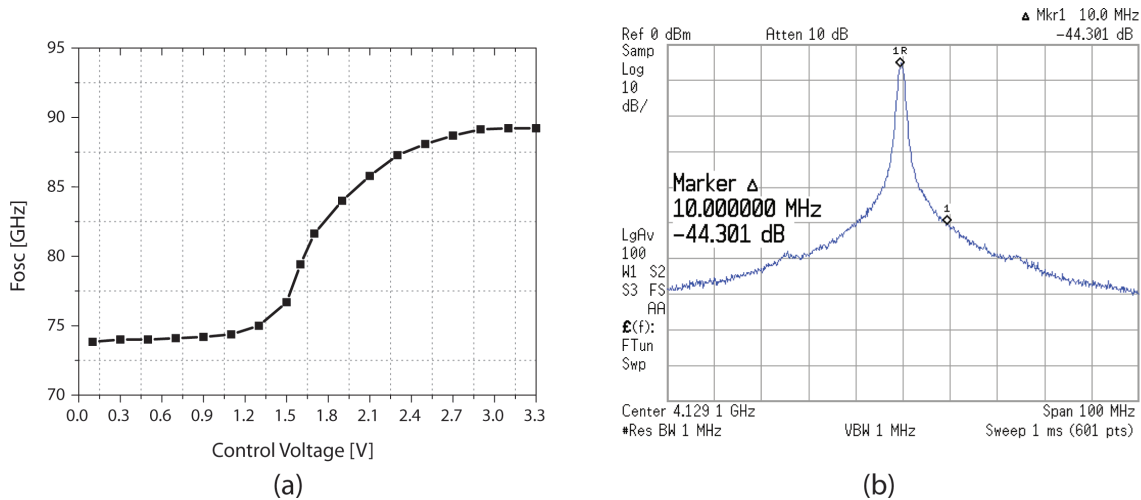


Figure 4.22: Measured VCO tuning range as well as the measured spectrum at 74 GHz.

that is driving the base node is small, we can have a voltage swing between collector and emitter that is considerably higher than BV_{CEO} . Fig. 4.1 shows the simulated BV_{CEO} with varying the base resistance. The DC bias network should accommodate the required low impedance at the base of the device.

Often when designing a high-frequency “linear” type power amplifier, to first-order, the figure of merit for the the device (in terms of output power) is $\frac{I_{DC}}{C} \times V_{BR}$ in which C is the device capacitance and V_{BR} is the break-down voltage. A bipolar device in a scaled technology typically operates with 1-1.5mA/ μm (of emitter length) with a capacitance of approximately $\sim 10\text{fF}/\mu\text{m}$. So the first part of the figure of merit described above, the current-to-capacitance ratio, is approximately $\frac{0.1\text{mA}}{\text{fF}}$. In a scaled CMOS node (e.g. 65nm) the current density in a general purpose (GP) process could be as high as 0.4mA/ μm (channel width) with the capacitance of $\sim 1.5\text{fF}/\mu\text{m}$ leading to a ratio of $\sim \frac{0.25\text{mA}}{\text{fF}}$. Therefore, in terms of current, CMOS has an approximate advantage by a factor of 2.5. In terms of break-down voltage, which is the second part of the figure of merit, bipolar devices could outperform CMOS by a factor of 2-3 depending on the specific topology and bias conditions. All in all, bipolar devices are not at a great advantage point in terms of high power mm-wave signal generation except that high voltage operation reduces sensitivity to supply network. Therefore, as is the case for CMOS mm-wave PAs, it is necessary to employ power-combining schemes to achieve the desired power levels in this design.

The PA is implemented as a two-stage differential-cascode transformer-coupled amplifier as shown in Fig. 4.23. Design details are available in the following publications [11] [70]. Differential design reduces the amplifiers sensitivity to the surrounding ground plane and reduces the need for bypass capacitors. This is especially important because the size of bypass capacitors is limited at mm-wave frequencies due to self-resonance. Also, differential

design offers more predictable performance by confining high frequency signal loops.

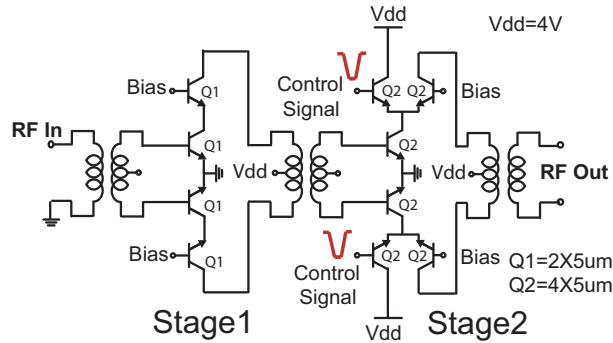


Figure 4.23: Two-stage transformer-coupled power amplifier.

The first stage uses $5(\text{finger}) \times (2\mu\text{m})$ transistors and the second stage uses $5(\text{finger}) \times (4\mu\text{m})$ transistors. The design consumes about 72 mA of current from a 4V power supply. It achieves a $P_{-1\text{dB}}$ of 14.4 dBm with P_{sat} varying between 15-17dBm across 80-90GHz. The peak measured PAE is 9.25% at 89 GHz. To enable switching, the PA uses a Gilbert current commutation topology. As shown in Fig. 4.23, the control pulse signal from the high speed timing circuitry is applied to the auxiliary path of the switching stage. The cascode transistor on the main path is biased at 2.4V. The control signal voltage on the gate of the auxiliary path provides a pulse signal with a maximum transition from 1.8V to 3V. Due to the differential nature of the switching pair, when the control signal is at its lowest level of 1.8V, the RF signal is routed towards the output stage and on-chip antenna. When the pulse control signal is at its highest value of 3V, the RF signal is diverted towards the auxiliary path and away from the antenna. To reduce possible transient glitches to the output, the base of the dummy path is controlled by the pulse and the main path's base is biased and bypassed. Since the base of the switching device in the PA presents a large capacitive load to the ECL driver, potential instability concerns exist and were met with placement of small loss on the base side. This was further explained in section 4.3.5.

Stability analysis is very important in the power amplifier design. Since antenna impedance can be a function of its environment or, in this case, the switching state, the load impedance of the power amplifier can effectively change and cause stability concerns. In the VS transmitter, the antenna impedance changes as the switches in the antentronic structure go from the OFF state to the ON state (chapter 3). In the CS transmitter, the MUX in the antentronics structure shields the PA to some extent and a smaller impedance variation is presented to the PA. Nevertheless, due to presence of residual variations, stability under a variable load should be addressed. Fig. 4.24 shows the PA stability factor for various loads.

When we are dealing with the design of a differential power amplifier, not only does differential stability have to be considered, but common mode stability is an important factor as well. Transformers suppress the common mode signal by nature. However, the

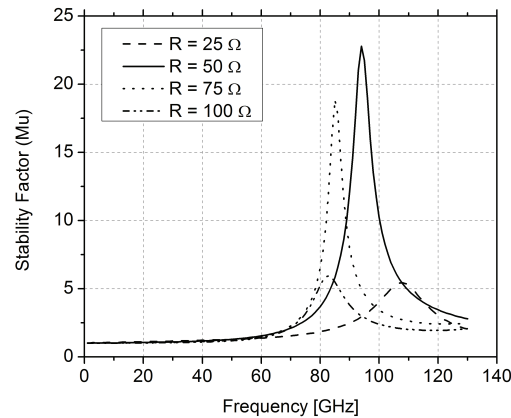


Figure 4.24: Stability factor of the PA under different presented loads to the output.

amount of suppression in common mode is usually a function of frequency and may not be adequate. To reduce the stability concerns of the CM loop, a small resistor can be placed in the CM path (bias). This does not affect the differential-mode gain and hence can ensure stability without AC performance loss.

4.6 Voltage Switching Transmitter

In this section, a voltage-switching 90 GHz pulsed transmitter will be described. Voltage switching refers to using an array of shunt CMOS switches in the antenna structure to abruptly turn the transmission off. A pulse voltage is applied to the gates of these CMOS devices simultaneously to activate them. The term voltage-switching is primarily used to distinguish this mechanism to that of the current-switching (or current-steering) system described in section 4.7. In reality, the actual output power of the device is modulated using this switching scheme.

4.6.1 Voltage Switching Antentronic Network

This transmitter uses the antenna structure described in section 3.2.1. Therefore, both the antenna and the PA have built-in switching functionality and the hybrid switching technique described before can be utilized. The block diagram of the system is shown in Fig. 4.5.

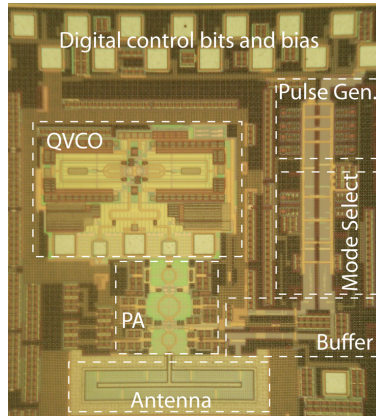


Figure 4.25: Chip micrograph of the TUSI VS transmitter.

4.6.2 Experimental Results

The transmitter is realized in a $0.13\mu\text{m}$ SiGe BiCMOS process with $f_t=230\text{GHz}$ for bipolar HBT devices. The die photo is shown in Fig. 4.25. The chip occupies a small footprint of $1\text{mm}\times 1.2\text{mm}$ (including on-chip antenna). TX measurements are done using a W-band horn antenna, a downconverter (D.C.) (5dB ripple up to 33GHz), a 50GHz spectrum analyzer, and a triggered oscilloscope. Initial measurements were done using the Agilent 86100C sampling oscilloscope with the 20GHz head [71]. We later used the 70GHz head (option 040) to obtain better measurements. A chip-on-board assembly setup is used to avoid the need of RF or DC probes, thus providing an appropriate radiation environment for the antenna. Fig. 4.26 shows the measurement setup. The horn antenna is situated at a distance of 15 cm above the chip plane. A clock generator is used to provide the pulse repetition frequency (PRF) for the chip as well as to trigger the oscilloscope.

Even with the high frequency sampling oscilloscope, three main factors limit measurements for very narrow pulses: (1) The limited BW of the external downconverter (D.C.) (up to 33GHz); (2) a free-running VCO (not locked to the PRF or the D.C. LO); and (3) the LO feed-through of the external D.C. mixer to the IF output. These effects in combination make measurements of narrow pulses/notches very challenging. Moreover, the signal is down-converted to an IF of 15 GHz, where it still has a carrier, which is detected by the sampling oscilloscope. These limitations cause pulse amplitude reduction and ringing. Fig. 4.27 shows a simulation of an “ideal” 90GHz 25ps (50%-50%) pulse going through an otherwise ideal down-converter that is bandlimited to 33GHz. The results are superimposed on the PRF window as would be the case for the sampling oscilloscope. The obtained pulse has considerable ringing, confirming that even in the absence of any other system imperfections, the down-converter’s limited bandwidth will result in waveform artifacts. The LO frequency selection plays an important role in the amplitude and shape of the ringing terms. The curves are simulated for the setting used to obtain the pulses depicted

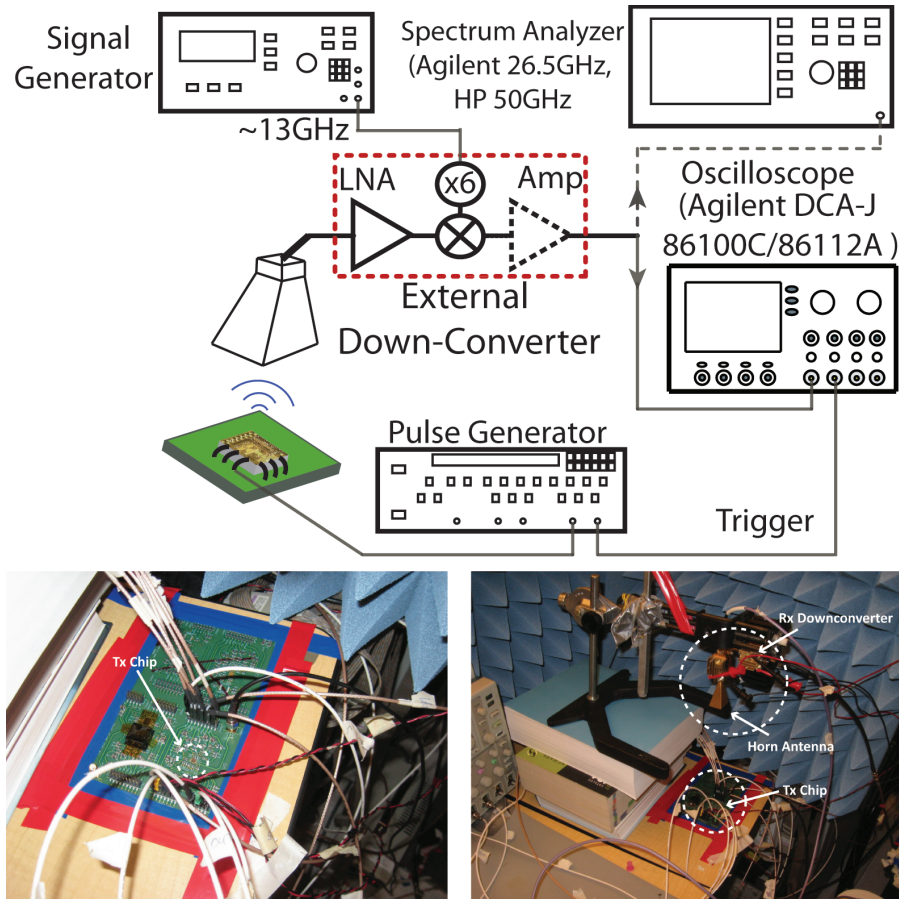


Figure 4.26: Transmitter measurement setup.

in Fig. 4.28. Since the frequency of the carrier is not an integer multiple of the trigger frequency (the PRF), the initial phase of the received sinusoid is random, leading to the pulse envelope being “filled-up” by the carrier signals. With the integration of the PLL this issue will be resolved. Measurements of the transceiver including the PLL are presented in chapter 6.

Fig. 4.28 shows the measurements of 33ps pulse in Mode 6 (Independent PA switching). This is an order of magnitude improvement over previously reported pulse widths obtained in silicon [72]. As predicted by the discussed simulations, the bandwidth limitation in the measurement equipment results in ringing in the measured pulse. Fig. 4.29 shows the spectrum and time domain response of the selected modes including a plot of obtained pulse-widths for the independent modes. Spectral measurements are used to pinpoint the impairments in the systems in terms of frequency response or leakage tones.

As explained in Fig. 4.26, the down-converter uses a frequency multiplier ($\times 6$) on the LO side to eliminate the need for a high-power mm-wave signal generator. The fundamental tone

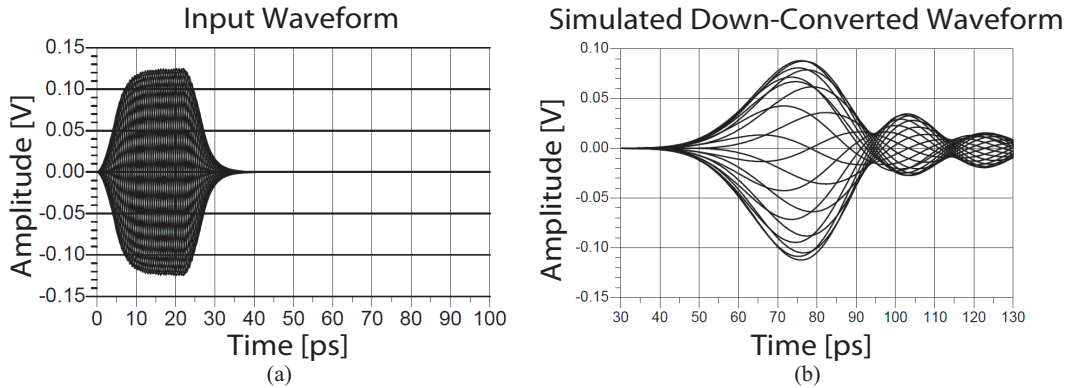


Figure 4.27: Ideal unlocked RF pulse (left) and simulated effect of non-ideal down-converter (right). Bandwidth limitation is the only non-ideality considered here.

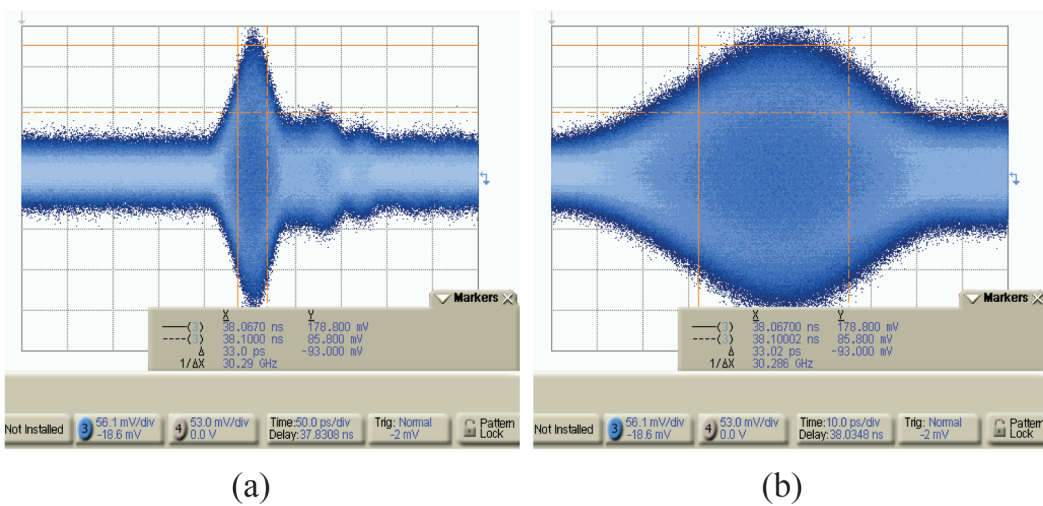


Figure 4.28: Measured pulse in mode 6 with (a) 50ps/div and (b) 10ps/div. Ringing due to equipment BW limitation is visible in (a).

provided to the multiplier is around 13 GHz. The feedthrough of this tone and its second harmonic are clearly visible in the spectrum measurements, which limits our minimum detectable signal. The downconverted RF signal is also visible in spectrum measurements. The spectrum exhibits a large peak at the equivalent of RF frequency since this spectrum is from a notch mode (Mode 5) in which the carrier is present at output most of the time.

Fig. 4.30 shows another version of the spectrum measurements for a hybrid switching mode. Here a 50 ps hybrid mode setting (Mode 3) is programmed. Fig. 4.31 shows the measured pulse in the hybrid switching Mode 1. An impressive pulse width of 26 ps is measured in this case. Here, the ringing amplitude is larger than previous measurements.

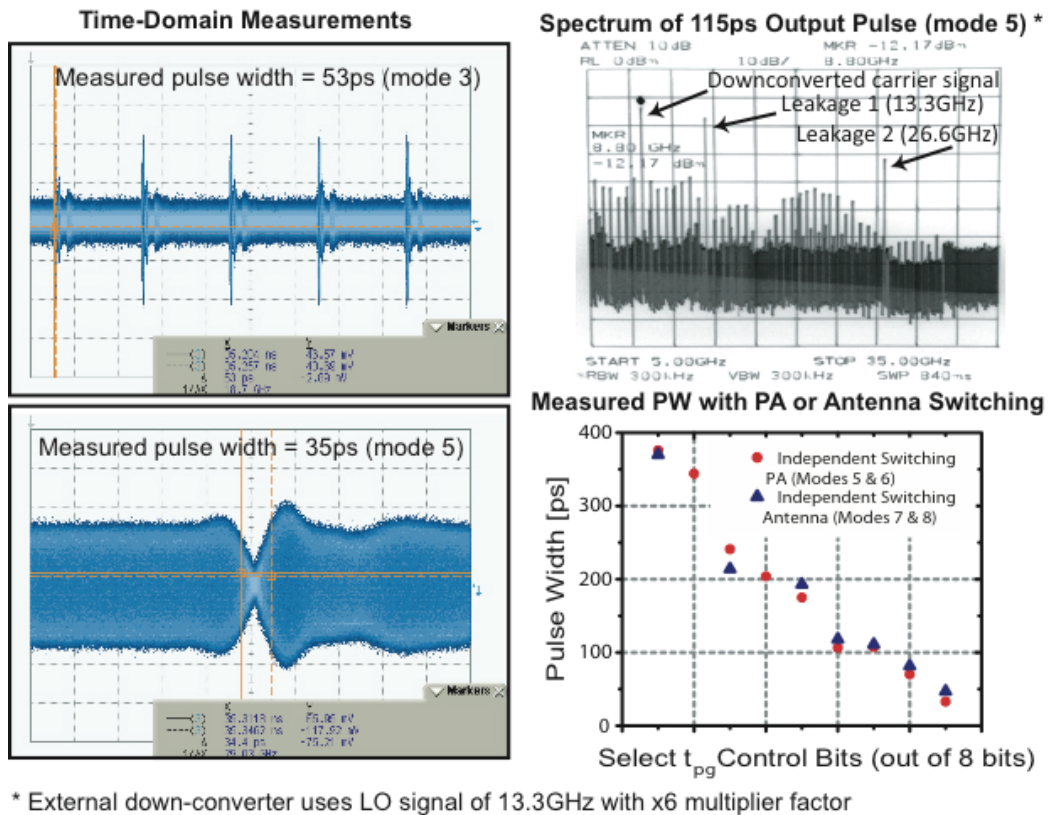


Figure 4.29: Time domain measurements of mode 3 (53ps) and mode 5 (35ps) with frequency domain measurements of (mode 5). Selected measured pulse-widths for independent switching (modes 5-8).

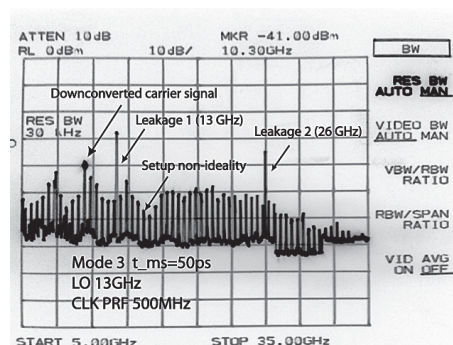


Figure 4.30: Spectrum measurements of positive pulse in mode 3 (hybrid).

This is partly due to the measurement artifacts (exacerbated by the larger BW imposed by this shorter pulse) and partly due to the transmitter getting close to its performance

boundaries. Obtained results from hybrid mode measurements are shown in Fig. 4.32.

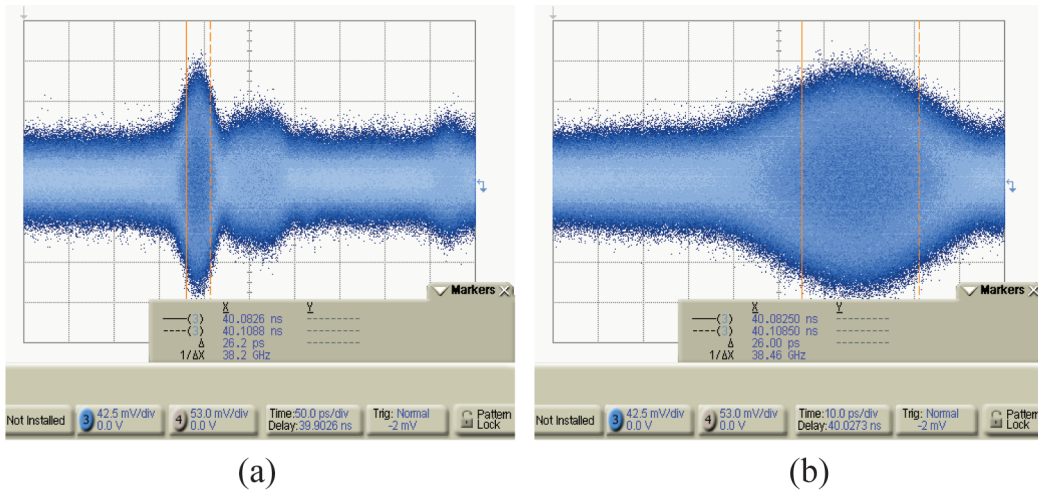


Figure 4.31: Time domain measurements from Mode 1 (hybrid) with (a) 50ps/div and (b)10ps/div settings. Measurement shows pulse width of 26ps (50% – 50%).

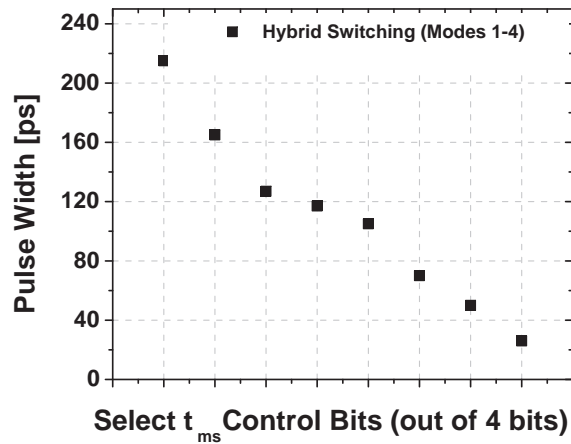


Figure 4.32: Obtained results from hybrid mode measurements (for select control bits (t_{ms}))

Measurements were also performed using reflections from an angled metallic reflector as shown in Fig. 4.33 (to observe polarization, spatial dispersion, and pulse accuracy). The distance on each side of this reflector was 20 cm and the receiver antenna position was adjusted to observe measurable positioning accuracy. As before, the oscilloscope is set to

the infinite persistence mode. After the first set of data was taken, the target position was moved. This way, both targets show up on the screen. Without any post-processing, two pulses were distinguishable with a time difference down to 23 ps. Frequency tone measurements were successfully performed up to a distance of 1.2 m. The noise floor of the oscilloscope (arising from noise and also LO leakage of the D.C.) limited pulse measurements at that distance.

Table 4.1 summarizes the measurement results for the TUSI VS transmitter.

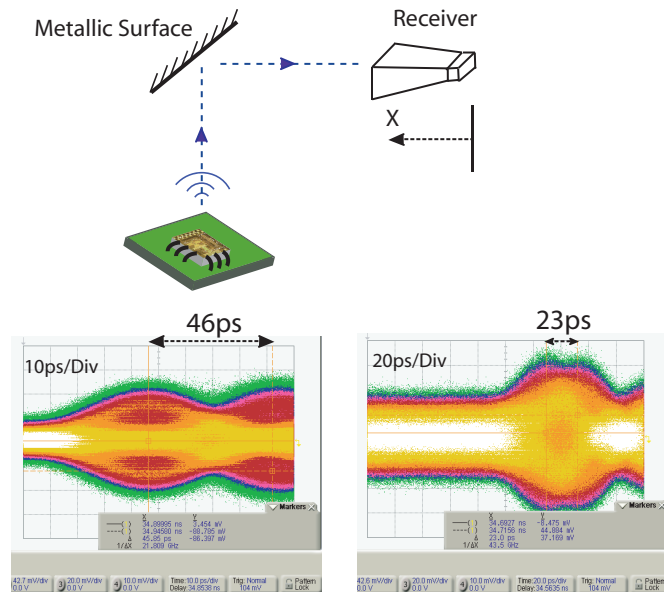


Figure 4.33: Measurements with a metallic reflector and infinite setting on oscilloscope. With $\Delta X = 13.8$ mm ($c \times 46$ ps) the pulses are distinguishable (left). With $\Delta X = 6.9$ mm ($c \times 23$ ps) they are close to the accuracy limits and barely distinguishable (right).

Table 4.1: Measurements Summary for TUSI VS

Performance Summary	
Technology	0.13-μm SiGe BiCMOS
Die Area	1.0 x 1.2 mm²
PA	
Gain	12.25 dB (at 90GHz)
Psat	17.2 dBm (4V)
VCO	
Frequency	74 – 89 GHz (18.4%)
Phase Noise	-104 dBc/Hz at 10-MHz (at 74GHz)
Measured TX Performance*	
Independent Switching	Mode 5. 35ps
	Mode 6. 33ps
	Mode 7. 47ps
Hybrid Switching	Mode 1. 26ps
	Mode 3. 53ps
Power Consumption	
PA	270 mW
VCO (w/ buffer)	180 mW
ECL Timing Circuits	252 mW
Total Power (including bias)	739 mW

* Measurement limited by equipment bandwidth. New measurements obtained compared to [49] using higher BW oscilloscope and on wider range of chip settings

4.7 Current Switching Transmitter

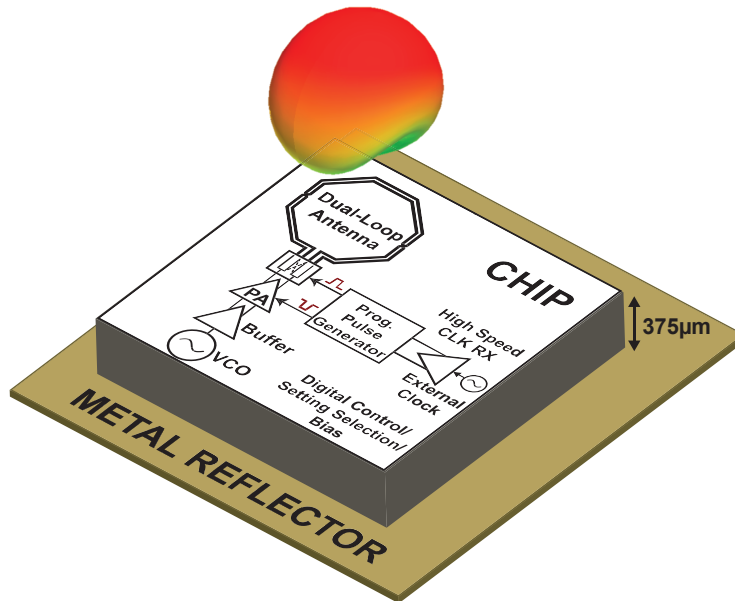


Figure 4.34: System level block diagram of the current-switched transmitter with dual-loop antenna.

The second transmitter uses a current-switching (CS) architecture and the dual-loop antenna described in chapter 3. The CS transmitter block diagram is shown in Fig. 4.34. It consists of a high speed clock receiver, pulse generation (PG)/distribution circuitry, VCO, VCO buffer, LO distribution, power amplifier (PA), the switching dual-loop antenna with the current multiplexer, digital control and biasing totaling to approximately 6000 transistors.

4.7.1 Power Tuning Capability

Power tuning capability is an important functionality for an array imager. Once the object is placed in the near-field of the array (and not the antenna itself), the signal levels reaching the target from various sources are going to be widely different. Power tuning allows for first order correction of this effect. It also allows for solving some of the “near-far” or ambiguous range problems especially in cases where the PRF is very high. The system may be used in scenarios where the losses are considerably smaller or only primary returns are of interest. In such cases, the power levels can be tuned lower to avoid undesirable reflections outside the main PRF as well as to avoid saturating the receiver.

In this chip, power control functionality is implemented in both the PA as well as the current multiplexer in the antenna. In the PA, power tuning takes the form of bias control through a 4-bit DAC. In the antenna network, the bias current of the secondary side of the transformer is tunable to allow for variable power generation (Fig. 4.35). Large changes in the current will lead to undesirable frequency response impairments and are avoided. The output power can be modulated by as much as 18dB using a combination of settings.

4.7.2 Current Switching Scheme

As mentioned previously, the switching functionality (for pulse or edge generation on the carrier frequency) takes place both in the antenna and the PA. In the CS version, to accomplish switching, a dual-loop antenna is implemented. The specifics of the antenna design are described in chapter 3. Two symmetric, co-centric, independently driven loops are realized. The loops are driven by a current multiplexer (MUX) in this combined antenna-electronic (Antentronic) structure (Fig. 4.35). The PA provides the signal to the current MUX through a single-turn wideband transformer. The secondary side of that transformer uses a programmable current source that can vary the transmit power.

To obtain the full benefit of the energy cancellation scheme, the currents in the two loops have to have the same amplitude with 180° phase difference. Therefore, symmetry and matching of the two paths is extremely important. In addition, when switching occurs, the current in the outer loop should be stable and unaltered. Fig. 4.35 shows the circuit diagram. The MUX uses two additional devices in the outer-loop side where the collectors are tied to the switching elements. This topology ensures that under both switching levels of the input, this branch carries the same amplitude as the inner loop.

4.7.3 Experimental Results

Pulse Measurements

The TX was fabricated in a $0.13\mu\text{m}$ SiGe BiCMOS process [73]. The die photo is shown in Fig. 4.36. The chip has footprint of $1.45 \times 1.2 \text{ mm}^2$. A chip-on-board assembly was used to measure the chip. The measurement setup is very similar to that of the VS version. A W-band horn antenna, a 3-33 GHz downconverter, Agilent E4440 spectrum analyzer, and Agilent Infiniium 86100C triggered oscilloscope with 70 GHz sampling head were used for performing measurements. The chip is placed 25 cm away and directly underneath the receiver horn antenna for direct pulse measurements.

Spectral measurements are shown in Fig. 4.37. As expected, the measured frequency spectrum resembles that of a triangular pulse. In this figure, the sinc^2 function is superimposed on top of the spectral measurements to provide a comparison to the ideal spectrum

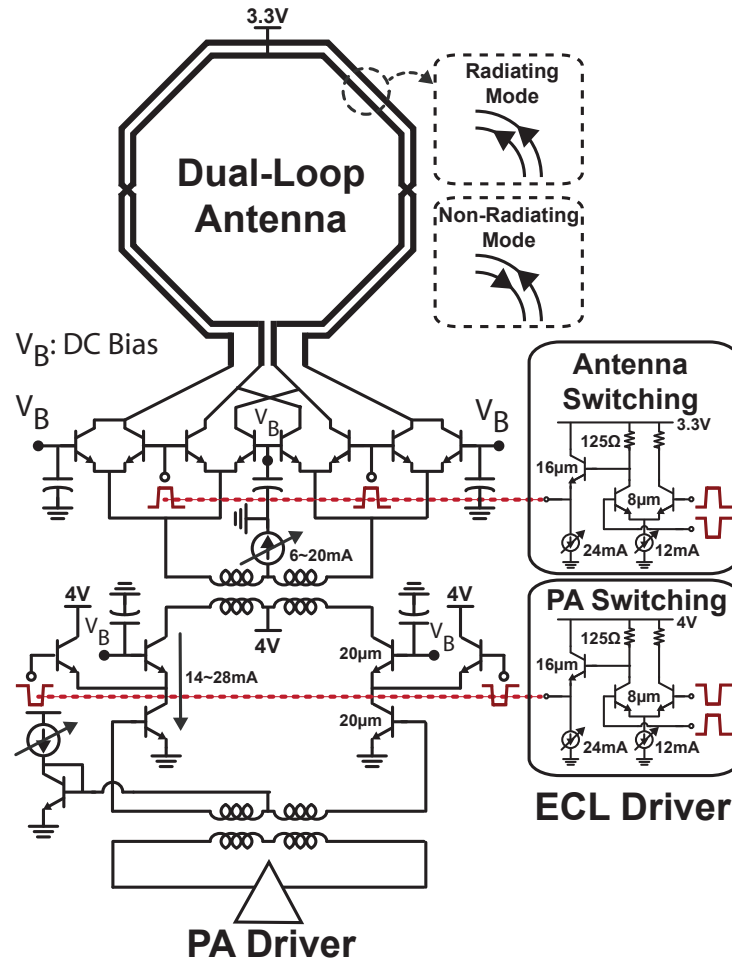


Figure 4.35: Circuit schematics of the current switching scheme transmitter.

of a triangular pulse. Measurements are performed close to the bandwidth limitations of the external downconverter.

The down-converter uses a $\times 6$ multiplier on its LO side. Spectrum measurements in Fig. 4.37 reveal the LO feedthrough of the downconverter and the second harmonic. This increases the signal level floor in the time-domain measurements. As seen by the same figure, a 13 GHz tuning range for the center frequency of the transmitted pulse is obtained. The center frequency varies between 77 GHz and 90 GHz.

Pulse widths down to 46ps have been measured. A transmit power of +10dBm is calculated from far-field power measurements. Measured half power beam-width is ± 39 degrees on H-plane. Fig. 4.38 also shows the achieved PRF of 3.45GHz as well as a plot of a selection of measured pulse-widths.

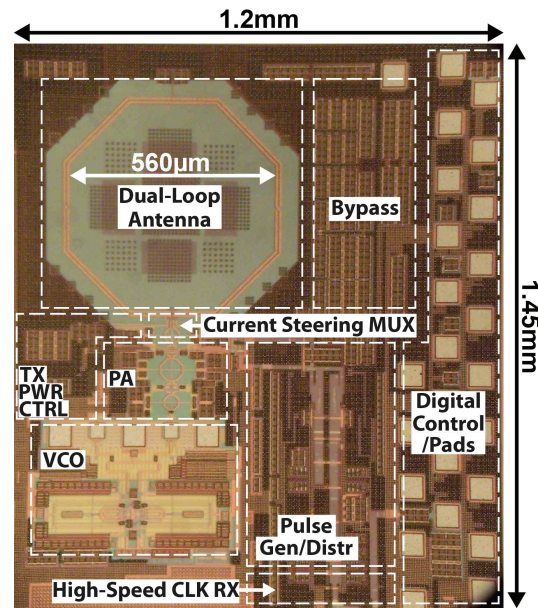


Figure 4.36: TUSI CS chip micrograph.

Radar Measurements

A bi-static setup was used to measure the performance of the transmitter. In this setup, the pulses are reflected from an external surface and the reflections are then measured using the downconverter. Time spacing of 46ps between distinguishable pulses was measured. This equates to a spatial resolution of 13.8 mm in air (~ 7.9 mm in human body). Direct measurements (in which the TX points towards the RX horn antenna with no reflections) were able to determine 10ps pulse position changes (translating to 3mm displacement in air and ~ 1.5 mm in human body). This accuracy is limited by the jitter accumulation in the TX clock path and gets worst for longer pulses.

Measurements were also performed for the detection of multiple (more than two) targets in which successive pulses are measured in the receiver. The setup is shown in Fig. 4.39. A corrugated metallic surface is used to present the radar with multiple close-by targets. As seen in the figure, 4 targets were distinguished within an area of 6cm. Table 4.2 summarizes transmitter performance.

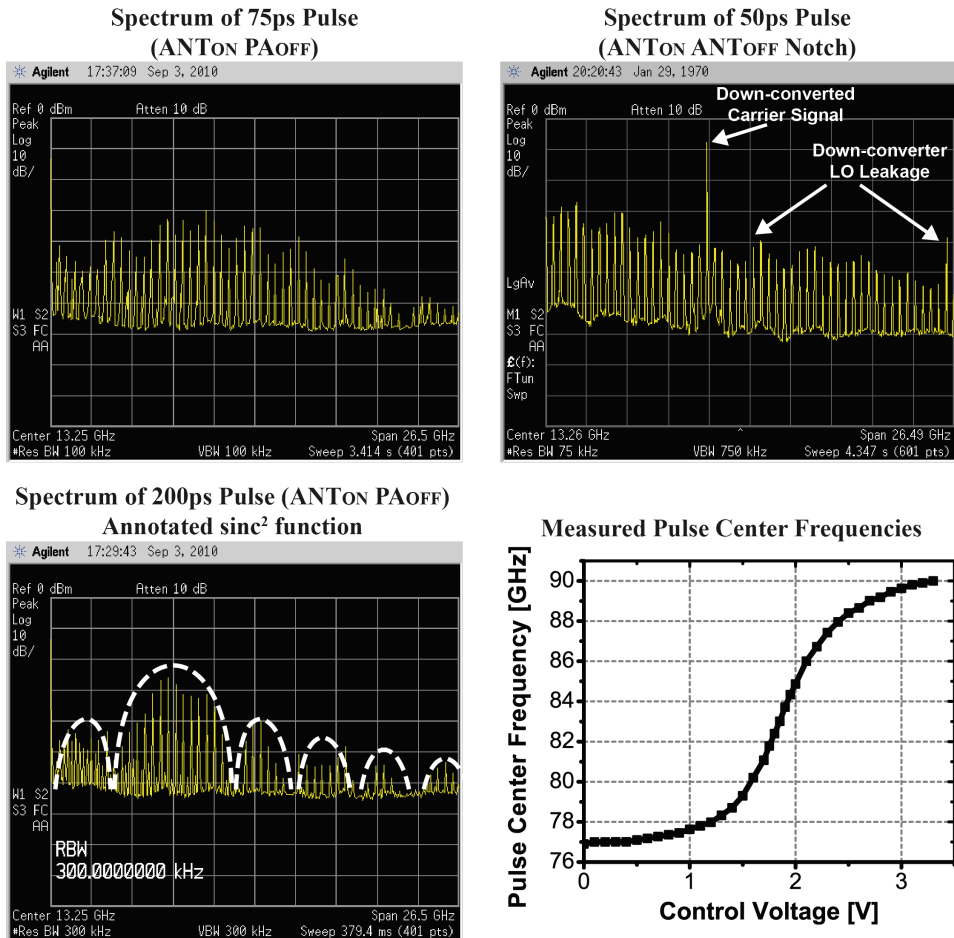


Figure 4.37: Measured PSD of various pulses and pulse center-frequency tunability. Spectrum measurements of the down-converted signal are from DC to 26.5GHz.

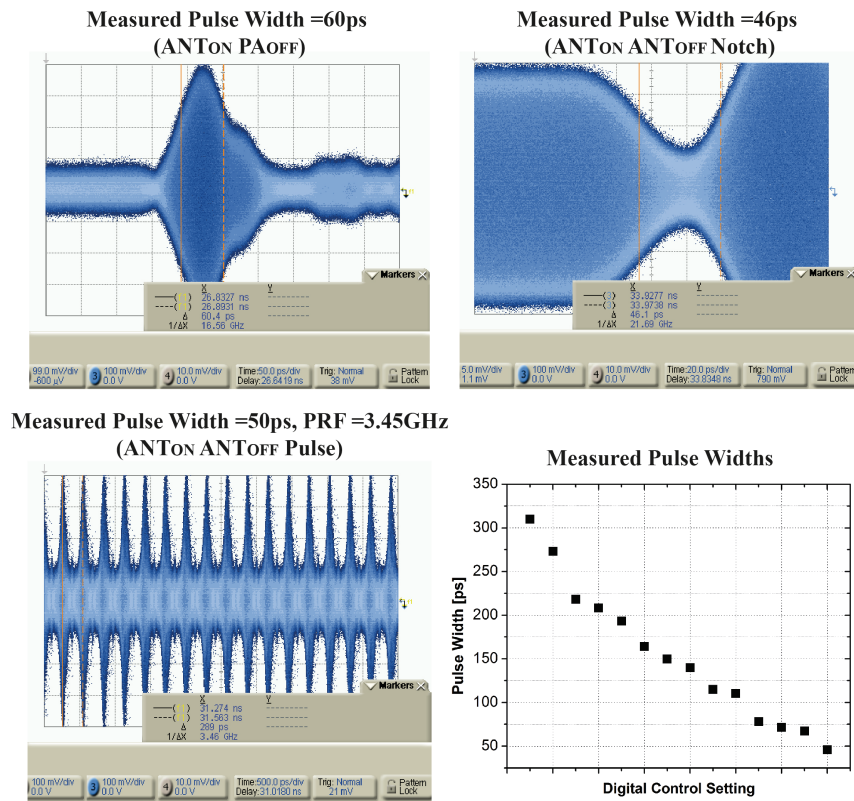


Figure 4.38: Time-domain measurements for different settings.

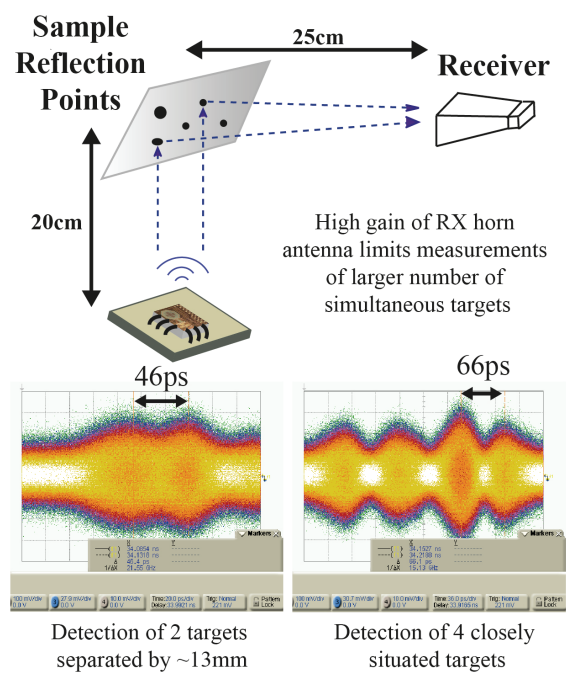


Figure 4.39: Bistatic reflection measurements.

Table 4.2: Measurements Summary for TUSI CS

Technology		0.13 μm SiGe BiCMOS
Area		1.45 x 1.2 mm^2
PA (standalone)	Gain	15 dB (at 89 GHz)
	P _{-1dB}	14.4 dBm
	P _{sat}	17 dBm
	DC Power	306 mW (small signal)
VCO (in system)	Tuning Range	77 – 90 GHz
	DC Power	115 mW
ECL	Programmable PW	350 – 25 ps (simulated)
	DC Power	360 mW
Antenna	-3 dB Beamwidth (H-plane)	$\pm 39^\circ$
	Gain	1.6 dB (simulated)
	DC Power	46mW
System Meas.	PW	46 -310 ps
	Max PRF	3.45 GHz
	CLK spur	-46 dBc
	P _{out,max}	+10 dBm
	P _{out} Tunability	18 dB
	Min Δt_{pulse} (Bi-Static)	46 ps

Chapter 5

Receiver

5.1 94GHz Receiver Overview

As previously mentioned, the first frequency band to be implemented was chosen to be the W-band. The 94 GHz transceiver is implemented in a SiGe BiCMOS process technology. Unlike the external down-converters used in Chapter 4, here we utilize a direct-conversion receiver. Fig. 5.1 shows the block diagram of the receiver. The external receiver had a block diagram shown in Fig. 5.2.

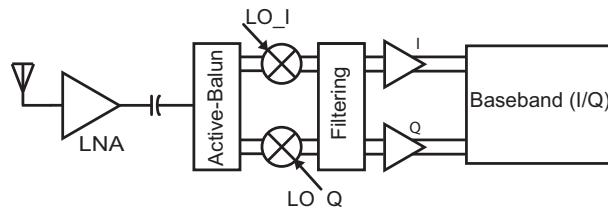


Figure 5.1: Block diagram of the direct-conversion TUSI receiver chain.

Direct-conversion receiver requires quadrature signal generation on the LO side as well as separate I/Q paths for conversion.

Direct conversion suffers from some well-known issues that are documented in the literature [74] [75]. For brevity, only a brief description of some of the issues will be provided here.

5.1.1 Distortion

In band distortion results from various sources in a direct-conversion receiver. Harmonics of the input signal fall in the baseband spectrum and can directly couple through the mixer to cause distortion. As an example, the second harmonic of the RX signal at the LNA output is amplitude rectified and can go through the mixer and show up as a distortion

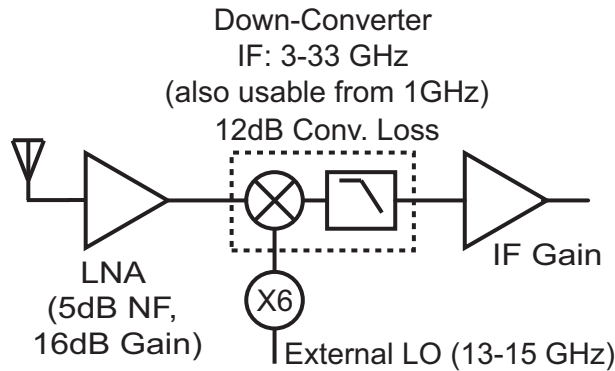


Figure 5.2: Block diagram of the external down-converter used to characterize the transmitter chips.

element in baseband. This signal occupies a large bandwidth (typically close to 2X the RF BW). Since this is an imaging device, these distortion products, even when small compared to the main signal, can cause severe limitations by generating false echoes. As previously illustrated, a large reflection (usually from the skin) is accompanied by smaller echoes from internal organs. Since the group delay response of the system is not flat with frequency, the distortion products of the first reflection can coincide (in time) with later echoes rendering detection impossible. In a data modulation system, the tolerance to phase error is set by the complexity of the modulation and spectral utilization. In simple modulation schemes (e.g. QPSK), this tolerance is relatively relaxed and is often compensated for using phase rotation. In an imaging system, however, this can be much more complicated and the tolerances smaller due to the unpredictability and large dynamic range of reflections.

In practice the AC coupling capacitor between the LNA and the mixer plays an important part in reducing feedthrough of this component. The coupling capacitor can however deteriorate the extremely wideband impedance matching between the LNA and the mixer. Also, other harmonics of the RF signal can mix with overtones of the LO to generate signals close to baseband. This issue is not dominant in our system due to the high frequency in use. Third and higher harmonics are extremely small due to excessive losses at 270 GHz and upwards. Also, balanced circuits are used where possible to mitigate even harmonics to first order.

The other distortion component is due to self-mixing of the interfering signals falling into the desired baseband spectrum. In addition, interferers can generate baseband components due to IM2 products in the mixer. In the case of the 94 GHz receiver, close interferers are not present due to sparsity of the spectrum usage as well as small range of operation of the imager. Also, the frequency selectivity of the antenna mitigates the effect of other interfering signals at much lower frequencies.

One of the most important issues is the DC offset problem due to self-mixing and leakage. This offset can be large enough to overwhelm the receiver. Care must be taken to remove

or mitigate DC offset in the imager. As will be explained later, the mixer incorporates fine-tuning control settings to calibrate and to first order remove the offset.

In our system, the RF bandwidth is quite large compared to the center frequency. The generated pulses have components that span as much as 40 GHz or more around the 94 GHz carrier. This will also add to the problem of RF feedthrough since the 94 GHz is difficult to fully filter at the corner of baseband frequency. Quantitatively, the baseband cutoff is selected to be larger than 30 GHz for each of the quadrature components.

5.2 Wideband Amplification

Fundamentally, the two critical circuit functionalities that are required for a pulsed-imager are wideband, low noise/ low distortion gain, and narrow pulse generation and control. In this section, the design constraints related to wideband amplification in silicon will be discussed. Two CMOS implementations that explore the fundamental boundaries of broadband amplification will be presented. Finally, a SiGe distributed amplifier with gain-bandwidth product in excess of 1.5 THz will be discussed.

5.2.1 Distributed Amplifiers

Wideband circuits find applications in various fields such as high speed links, broadband radio transceivers, high frequency instrumentation circuitry, high resolution radar and imaging systems. With the scaling of CMOS technology and transistor cutoff frequencies in excess of 100GHz, considerable research effort is invested in CMOS broadband circuits. CMOS is a low cost (high volume) alternative to III-V technologies that are currently the main option for millimeter-wave components. CMOS technology provides many advantages as flexibility in number and topology of active devices and disadvantages mainly related to the passive components on the lossy substrate. Mainstream CMOS technology does not provide additional options for RF and microwave circuits and this results in excessive conductive (series) and dielectric (shunt) losses in passive components. Also, lower intrinsic gain from the devices decreases the margin for modeling errors and requires careful prediction of device characteristics.

Distributed amplifiers (DAs) provide a large bandwidth in a given process with low sensitivity to mismatches and modeling deficiencies and therefore are a prime solution for extremely wideband amplification [76] [77] [78]. The operation of a DA relies on the operation of a synthesized transmission line formed by external inductive elements with the parasitic capacitances from active devices. The addition of signal currents on the low impedance drain line leads to a relatively low gain, albeit a large bandwidth. Numerous CMOS and silicon based DAs in various forms have been reported [79] [80] [81]. More recent distributed amplifiers have been published in the recent year [82] [63].

The rationale behind the DA will be shown using an approximate analysis. Starting from

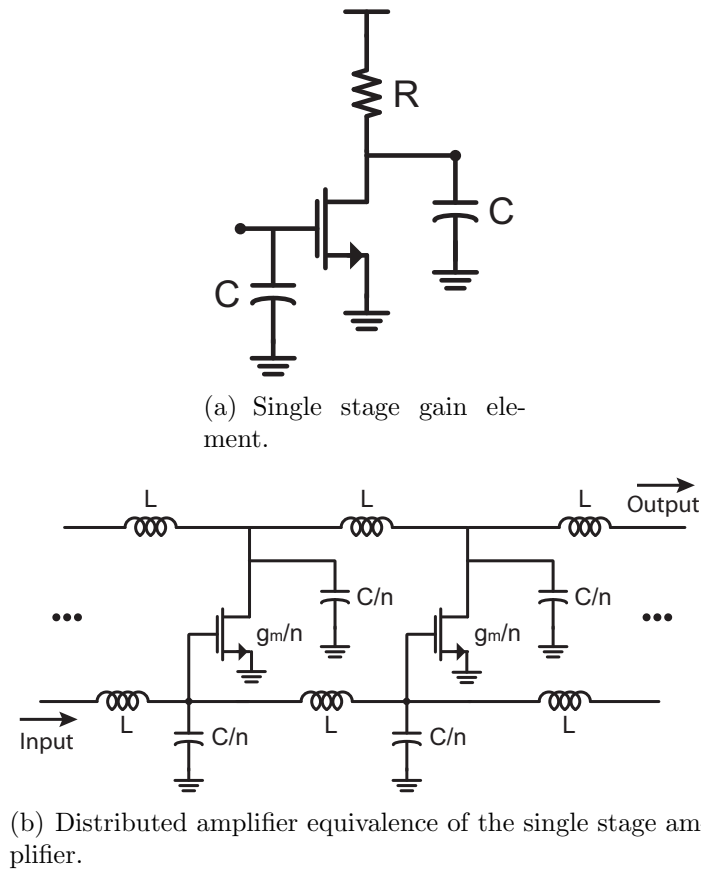


Figure 5.3: Comparing a single stage common-source to a distributed amplifier.

a simple single stage amplifier shown in Fig. 5.2.1 the gain, bandwidth and gain-bandwidth product will be:

$$\begin{aligned}
 G &= g_m R_L \\
 BW &= \frac{1}{2\pi R_L C} \\
 G \times BW &= \frac{g_m}{2\pi C} \simeq f_t
 \end{aligned} \tag{5.1}$$

Here we have assumed that the load capacitance of the element is C which could be the next stage loading or the output capacitance of the stage. Now if we divide this device to n smaller devices each device will have transconductance of $\frac{g_m}{n}$ and a capacitance of $\frac{C}{n}$ (leading to the same cutoff frequency). Inserting inductance elements between these capacitances leads to a distributed amplifier with the following gain and bandwidth parameters (section 5.2.4).

$$\begin{aligned}
G &= \frac{1}{2}n\frac{g_m}{n}Z_0 = \frac{1}{2}n\frac{g_m}{n}\sqrt{\frac{L}{\frac{C}{n}}} \\
BW &= \frac{2}{2\pi\sqrt{L\frac{C}{n}}} \\
G \times BW &= \frac{ng_m}{2\pi C}
\end{aligned} \tag{5.2}$$

Hence using this simplified analysis we see that the DA has a potential to provide a much larger gain-bandwidth (GBW) product. The real picture will need to include parasitics associated with fragmenting the device as well as line loss and parasitic capacitance [78].

Increasing the gain of CMOS DAs usually comes at the heaviest trade-off in bandwidth. Other amplifiers that achieve high bandwidths can only provide high gains at the lower part of the available spectrum (based on the speed of the process) and this is mainly through multi-resonant and compensation techniques that often provide non-predictable peaking and droops and hence severe group delay variations. In the DA regime, previous techniques to increase the gain rely on cascading elements and sections [83] [61] and this, without provisions, can significantly reduce the bandwidth through the introduction of multiple poles or cause of unpredictable effects in gain and bandwidth.

5.2.2 Passive Design

One of the main issues with the design of a CMOS DA is the inductive elements in the synthesized transmission lines. Traditionally, these elements have been realized using spiral inductors in CMOS circuits. On chip spiral inductors are very common and by using the right geometry, the quality factor (Q) and the self-resonant frequency (SRF) could be optimized for the required inductance value. However, when small inductance values are required, spirals introduce a problem. In a DA, the spiral inductor (an inherently one-port element) is driven from opposite sides as a two-port element (the signal comes in from one side and exits the other side). This further complicates the picture with the additional leads dependent on the size of the spiral as well as the way it would fit with the rest of the DA. Given that the spirals do not lend themselves to an accurate scalable model, it is rather difficult to design using these elements unless several iterations between the full-wave electromagnetic (EM) solver and the circuit analyzer are performed. Even then, the position of nearby elements and unavoidable ground planes in the final layout could alter the inductance value and/or change the self-resonance frequency.

For large bandwidth DAs, the desired value of the inductor for the line is quite small. This is because the ratio of inductance to capacitance being constant (and proportional to the required characteristic impedance of the line), the product determines the bandwidth of

the synthesized line beyond which the operation of the DA is not possible. The required inductance value is in the 40-140pH range where the lower side corresponds to the inductances in the M -derived sections.

Transmission lines are used as inductive elements in the synthesized sections of a DA to remedy some of the issues for high frequency operation. CMOS transmission lines can be relatively accurately modeled and they are inherently two-port devices. Our approach has been to use a set of measurement based data to calibrate the loss parameters in the EM simulator and to verify this with measurements from various passive components (transmission lines, inductors and transformers). This allows for dependable data from the EM solver and makes accurate modeling of transmission lines possible. Note that we require transmission lines of high characteristic impedance (Z_0). One can calculate the equivalent impedance by modeling the section between two transistors by a Π -model and with a short line approximation

$$Z_{in} = jZ_0 \times \tan(\beta l_{seg}) \simeq jZ_0 \times \beta l_{seg} \simeq jZ_0 \times (2\pi f)t_d \quad (5.3)$$

where t_d is the delay in the section. Eq. 5.3 shows that the inductance of the segment can be approximated by $Z_0 t_d$. Similarly, the total capacitance of the section can be approximated by $t_d Y_0$, half of this on each side of the model. A transmission line with extra parasitic capacitances loading it periodically has an approximate characteristic impedance of

$$Z_{final} = \sqrt{\frac{L_{line}}{C_{line} + C_{par}}} = \frac{Z_0}{\sqrt{1 + \frac{C_{par}}{C_{line}}}} \quad (5.4)$$

Replacing C_{line} by $\frac{t_d}{Z_0}$ and replacing t_d by $\frac{l_{seg}}{v_{line}}$ with v_{line} being the wave velocity in the line, one can derive the required segment length similar to [84] as

$$l_{seg} = \frac{C_{par} \times v_{line} \times Z_0}{\left[\frac{Z_0}{Z_{final}}\right]^2 - 1} \quad (5.5)$$

On the other hand the maximum attainable bandwidth of the line depends on the length of the segment. Quantitatively, the bandwidth of a synthesized line ($\omega_c = \frac{2}{\sqrt{LC}}$) is given by $\frac{2v_{line}}{l_{seg}}$. Fig. 5.4 illustrates the required initial Z_0 for a desired cutoff frequency given realistic loading from 90nm CMOS devices. Impedances in excess of 85Ω will be required to achieve amplifiers with bandwidth larger than 70-80GHz. This is a challenge in a scaled standard CMOS process with low resistivity silicon substrate, thin oxide stack and no thick metal option.

For the realization of high impedance transmission lines, the coplanar waveguide (CPW) structure is the best option among conventional topologies especially in a digital CMOS process where the stack height does not suffice for design of high impedance microstrip lines. Also, the grounded CPW may produce higher losses at high frequencies [85] as well

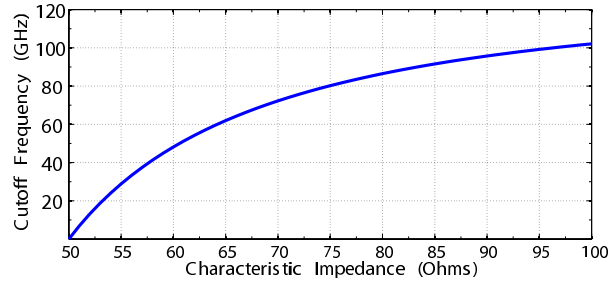


Figure 5.4: Line bandwidth limitation due to the initial Z_0 of transmission line.

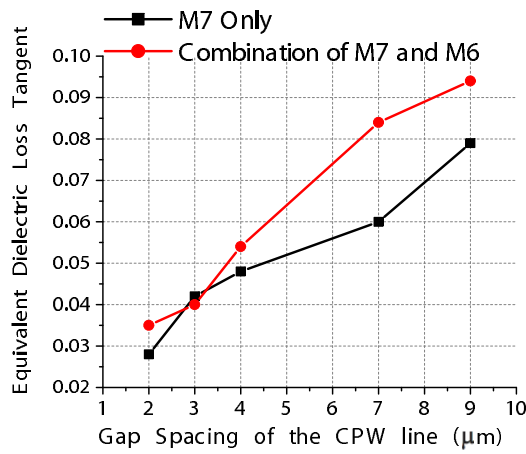


Figure 5.5: Equivalent loss tangent of the dielectric (to model the shunt losses) for various gap spacings and two different CPW structures (M7 only and M6-M7 combination).

as having a lower overall characteristic impedance. In an integrated CPW, the impedance increases by decreasing the ratio $\frac{W}{W+2G}$ where W is the width of the signal conductor and G

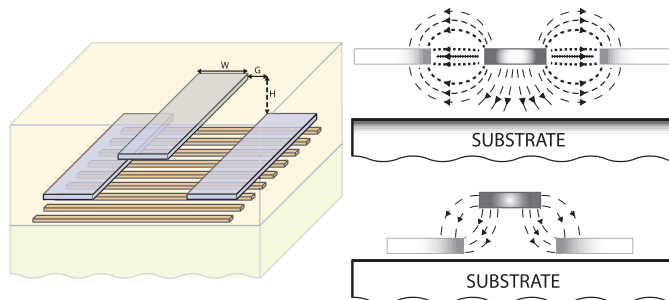


Figure 5.6: Conceptual layout of shielded elevated CPW (left) and illustration exaggerating current flow and electric field in both cases (right).

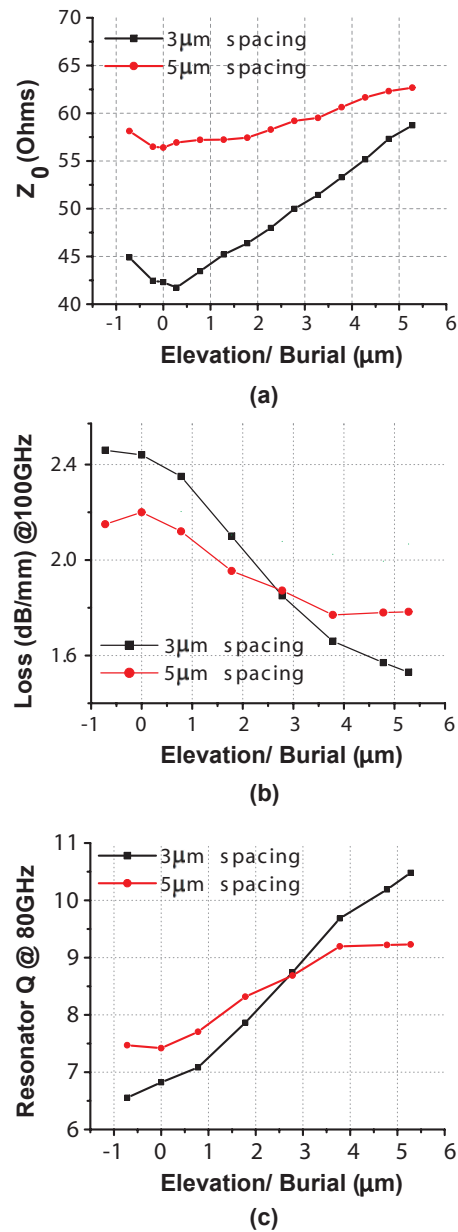


Figure 5.7: HFSS simulations of (a) Z_0 , (b) loss and (c) resonator Q for various elevation and burial conditions and two lateral gap spacings (G).

is the gap spacing. Increasing G will increase the shunt losses. Fig. 5.5 shows the increase in the equivalent loss tangent of the dielectric as a function of the gap spacing for two

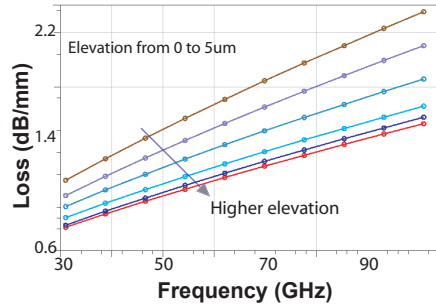


Figure 5.8: HFSS simulations of loss for various elevations with respect to frequency.

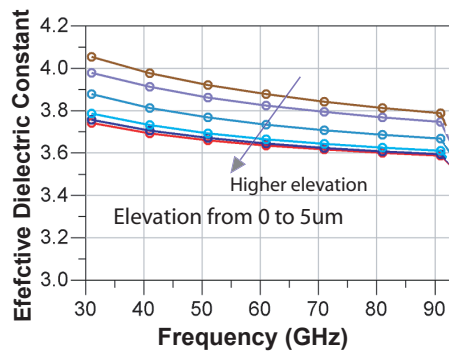


Figure 5.9: HFSS simulations of effective dielectric constant for various elevations in terms of frequency.

different CPW structures. The data is from measurements in the 90nm standard CMOS process used in the study. The Z_0 of the lines covers the range from 32Ω to 65Ω based on the choice of line width and spacing. It is interesting to note that moving away from very small gap spacings can actually slightly decrease the conductive losses since it reduces field concentration in the sides of the signal conductor and can result in a more even current distribution. However, the overall losses increase sharply (since G increases) for a large gap spacing required for impedances larger than 70Ω . Decreasing W will also have some effect in increasing line impedance (depending on the W/G ratio) however this will result in increase in conductive losses through the signal conductor.

To provide the higher impedance while minimizing the increase in losses, we propose using elevated CPW transmission lines in the DA structure. Here, the ground conductors will be lowered with respect to the signal conductor. This way, the lateral capacitance of the line is reduced (by preventing a “face to face” structure) and also the physical distance of the line to ground is increased. This will lead to higher impedance CPW lines. Lower loss is achieved as more fields are “captured” by the ground line and also the current is more evenly

spread across the signal conductor. To further reduce losses, shielding metal filaments could be added underneath the transmission line. This would add shunt capacitances without altering the inductance very much and hence “slow down” the wave and require shorter length lines for the same inductance. Line elevation on the other hand, leads to minor decrease in the effective dielectric constant of the line. Although the two effects are opposite, the final wave-velocity could remain lower than a conventional CPW with the right choice of elevation and spacing. Fig. 5.6 shows the conceptual layout of a shielded elevated CPW (E-CPW) together with an illustration that exaggerates current flow and electric field in both cases for clarification.

To confirm the effect of elevation, simulations were performed in Ansoft HFSS using the 90nm CMOS process stack properties and loss parameters. The signal line is assumed to stay on the top metal line and the ground line is shifted for comparison purposes. In these simulations, the ground line thickness is not changed. This could be achieved in practice by stacking two metal lines together. Fig. 5.7 shows the simulated characteristic impedance, loss and resonator Q of the elevated CPW lines (with no shielding filaments). Resonator Q is plotted to show that even accounting for the change in wave velocity elevation, the overall equivalent losses for a given phase shift reduces. In Fig. 5.8 the simulated losses for various elevations are shown in terms of frequency.

Fig. 5.9 shows the effective dielectric constant derived from transmission parameters of the elevated line. As predicted above, line elevation does actually increase the effective wave velocity. This is explained in part due to the fringing fields over several stack materials in the oxide layers similar to a microstrip line with air interface.

To verify the simulations, a set of test structures were fabricated and measured. Fig. 5.10 shows the measurements of characteristic impedance and also loss (dB/mm) for a wide-gap elevated CPW both with and without the shielding elements [62]. Measurements are accompanied by simulations from HFSS (for the non-shielded case). A relatively close match between simulations and measurements are obtained, with the possible explanation for the minor discrepancy arising from the effective dielectric constant of the oxide layers employed in HFSS. Metal filaments reduce losses at the cost of reduced impedance. This will lead to a lower inductance and hence a longer required line, but at the same time lower losses and hence a tradeoff.

To verify the extent to which elevation can reduce losses, a further step was taken with using filaments on poly and M1 layer with the signal line being on the aluminium capping layer. This is an extreme case of elevation and shielding. Fig. 5.11 shows the measurements of the loss of this transmission line. This shows that the losses are considerably lower in the proposed case and therefore these transmission lines can even be used in cases where moderate impedances are required. The effective permittivity related to the elevated line reduces from 4.3 to 3.7 at frequencies close to 40GHz.

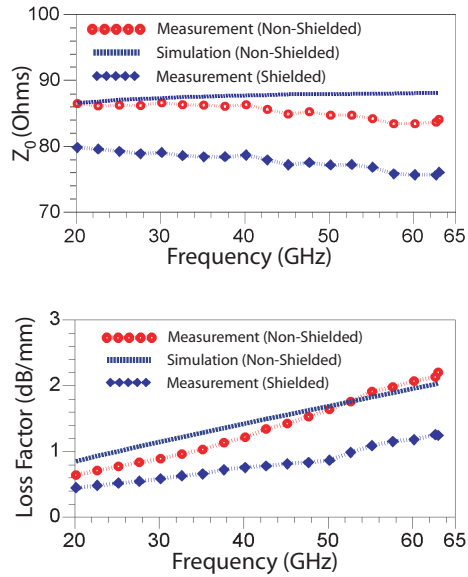


Figure 5.10: Measurements of Z_0 (*top*) and loss (*bottom*) of a wide-gap E-CPW line both with and without shielding filaments.

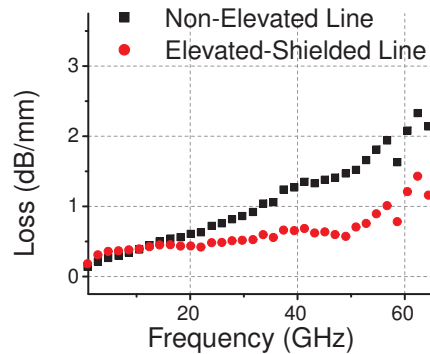


Figure 5.11: Measurements of a shielded elevated CPW with M1 and poly as filaments.

5.2.3 Active Element Design

An issue concerning DA design in CMOS is the selection of the optimum device size. Conventional microwave DA design does not provide optimal design when applied to CMOS. This is because in CMOS technology, device sizing and exact topological layout (number of fingers) are free parameters and should be exploited for DA design.

In order to exploit the extra degrees of freedom available in CMOS technology, we exported first-order scalable device model parameters into MATLAB. Together with CPW line models that were extracted from measurements, a parametric expression for the gain

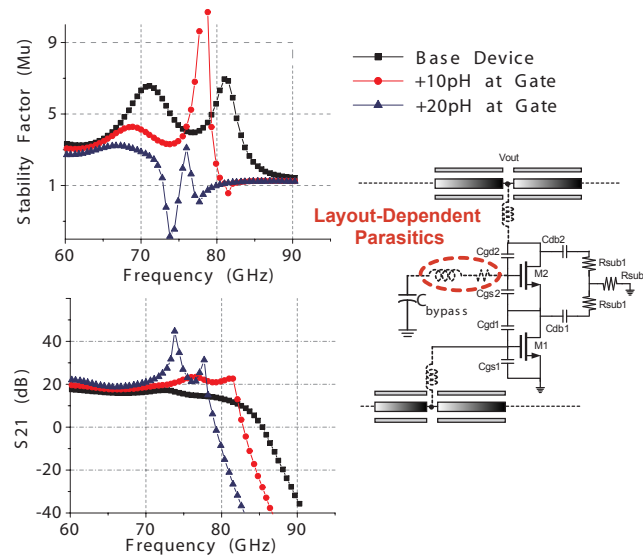


Figure 5.12: Schematic diagram (*right*) of cascode device with important parasitic elements shown. Stability factor (*top*) and S_{21} (*bottom*) for a sample DA incorporating the cascode element is also shown. Gate parasitic inductance is varied.

of a conventional DA stage was derived and optimized for a given bandwidth. The gain function has different sensitivities to various design parameters in our design space. For a well chosen line Z_0 and finger width, it is seen that device topology, number of devices (in the DA) and number of fingers are the three most significant factors determining the gain. For device topology, cascodes and common-source topologies are the two main candidates. The cascode device provides a higher stable gain at lower frequencies limited by a pole at approximately half the device cutoff frequency. Extra care must be taken with the cascode device since parasitics can easily lead to instability and/or undesired gain peaking. HFSS co-simulations to capture extra layout-dependent parasitics were performed to ensure predictable performance by the cascode elements. Fig. 5.12 shows the schematic diagram of a cascode device with the important parasitics annotated. To emphasize the significance of small parasitic components, the effect of added inductance on the gate is illustrated both on the stability factor and S_{21} of a sample DA incorporating these cascode elements. To circumvent these issues, apart from careful modeling using EM simulators, DA-friendly layout topologies for the cascode device (with appropriate rotation and aspect ratio for the input/output transmission lines and capacitances) were used with appropriate local bypass capacitors close to the gate. Any interconnect section was designed to have the least added inductance.

Once the device topology is chosen (cascode versus common source), MATLAB simulations of the overall gain and bandwidth were performed with number of devices and number

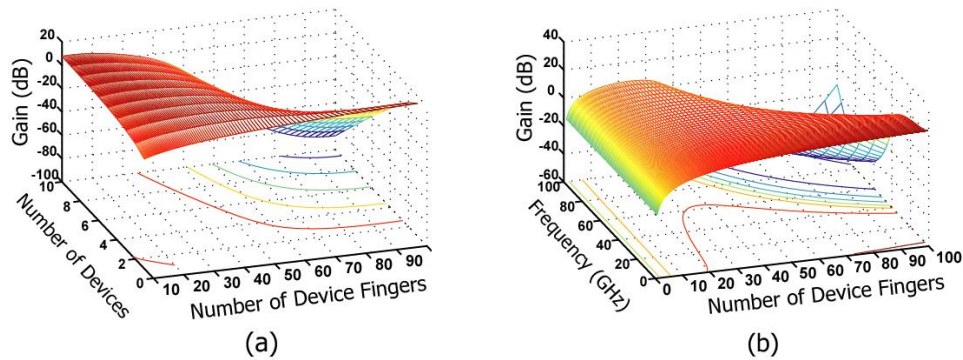


Figure 5.13: Simulated DA gain (a) at 60GHz with varying device size/ number of devices and (b) for 6 devices with varying frequency/ device sizes.

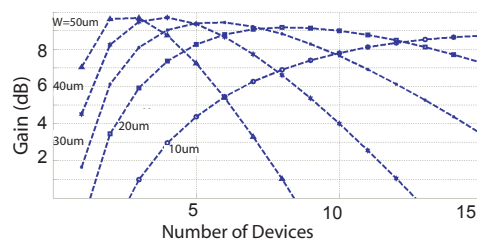


Figure 5.14: Simulated DA gain vs. number of devices for various number of fingers.

of device fingers as the two main variables. Fig. 5.13 shows the simulated gain of the DA with finger width of $1.2\mu\text{m}$. As these device models are approximate scalable models, after this step, the closest transistors to the candidates are chosen from the custom library and accurate, in-house models were used in the optimization process. It is observed that the gain is not constant for a constant “total width” of all devices. Nor is the sensitivity to each parameter constant for different decompositions of total width. Larger devices provide more gain with the same total width but are more sensitive to various parameters as well as showing more parasitics in their structure. Fig. 5.14 shows the 2D version of the simulations for clarity. For our design, the $40\mu\text{m}$ cascode device provides the optimal gain with 3-5 devices (depending on the required BW).

This procedure will select the optimal topology for maximizing gain and/or bandwidth. There are additional figures of merit for a DA (output power, noise figure, efficiency, etc.) that could be used for the optimization process with the appropriate functions.

5.2.4 Distributed Amplifier with Internal Feedback

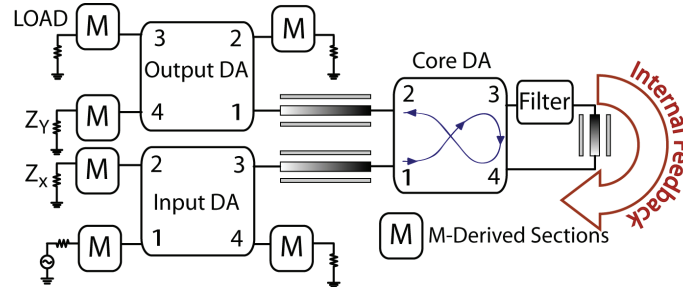


Figure 5.15: Proposed DA architecture for improved gain-bandwidth product.

The circuit topology of the internal feedback distributed amplifier (FBDA) is shown in Fig. 5.15. It consists of three separate DA stages denoted by input, core and output stages. The input and output stages are conventional DA stages with appropriate terminations. The core stage has only two terminations with the other two connected together with the means of a delay element (filter). The feedback uses the output of a conventional DA to feed its input on the terminating gate side. This way, without using any extra devices (leading to larger area and power), the gain would be proportional to A^2 where A is the gain of the core stage when operated as a DA. The input and output stages ensure that the reflected wave from the feedback does not end up in the input causing a large VSWR. The essential role of the input stage is to have an acceptable S_{11} across the band of interest. Also, it can reduce the overall noise figure if the gain is sufficient. Similarly, the output stage serves the purpose of delivering the current to the load with an acceptable VSWR. The output devices could be biased separately to optimize the power delivery performance of the system while the previous stages provide the necessary gain. From the above discussion, the minimum number of stages in the input and output stages is limited by S_{11} and S_{22} constraints.

Distributed Amplifier Gain Overview

Qualitatively, the core operates in the following manner: The gained up signal appearing at the drain side is traveling on a transmission line with impedance Z_0 . The gate impedance being the same, this signal can be fed back to the gate side allowing it to once again experience the gain stages this time to the left termination where it is finally terminated by the gate line of the output stage. The filter cleans up the unwanted signals that are out of band and helps in biasing the circuit once the gate and drain lines are connected.

The fact that the signal experiences a gain proportional to the square of what it would otherwise see using the same number of active and passive elements is the key advantage of this design. To analyze the operation of the circuit we will start by calculating the forward and reverse gains of a regular distributed amplifier stage.

The forward gain of a DA is the consequence of equi-phase signals summing up as they travel on the gate and drain line. Analytically the equation for I_d on the drain side would be:

$$I_d = \frac{1}{2} \{ I_1 e^{-j(n-1)\beta_d} + I_2 e^{-j(n-2)\beta_d} + \dots + I_n e^{-j(n-n)\beta_d} \} \quad (5.6)$$

The term β refers to the phase delay of one inductive element. It can be shown that this term can be approximated by $\frac{2f}{f_c}$ where f_c is the cut-off frequency of the line. In our analysis we will neglect the effect of losses when the expressions for gain and noise figure are being derived. Exact equations are derived elsewhere [78] [86]. Here, we aim at showing the proposed concept in a clear cut manner as a proof of concept. Simulation results without neglecting the losses are presented in later sections.

The voltage wave traveling down the gate line also experiences a phase delay from the gate transmission line. For the gate voltages on each of the devices we have:

$$V_1 = V_{in}, V_2 = V_{in} e^{-j(1)\beta_g}, \dots, V_n = V_{in} e^{-j(n-1)\beta_g} \quad (5.7)$$

The relationship between I_p and V_p is from the conductance of the device which is assumed to be independent of frequency ($I_p = g_m V_p$).

With the assumption of lossless propagation, the magnitude of all the gate voltage and therefore currents at the drain would be equal. If we combine previous equations and simplify the geometric series we obtain:

$$I_d = \frac{1}{2} g_m V_{in} e^{-j(n-1)\beta_d} \left\{ \frac{1 - e^{-jn(\beta_g - \beta_d)}}{1 - e^{-j(\beta_g - \beta_d)}} \right\} \quad (5.8)$$

The $\frac{1}{2}$ factor in this equation is a result of current division effect at the drain side. The gate and drain phase delay being equal, the expression for the magnitude of the output current to input voltage gain reduces to:

$$A_F = \frac{I_o}{V_{in}} = \frac{ng_m}{2} \quad (5.9)$$

Also, for the available gain from Vs to the load we have:

$$G_A = \frac{n^2 g_m^2 Z_g Z_d}{4} \quad (5.10)$$

The reverse gain (input at the left side and output from the left side of the drain line) can also be derived in a similar manner. To derive the reverse gain we sum the output current at the reverse side of what we just calculated. Signals will not add in phase.

$$I_{dR} = \frac{1}{2} g_m V_{in} \left\{ \frac{1 - e^{-jn(\beta_g + \beta_d)}}{1 - e^{-j(\beta_g + \beta_d)}} \right\} \propto \frac{\sin(n\beta)}{\sin \beta} \quad (5.11)$$

This equation shows that the gain has a $\frac{\sin(nx)}{\sin(x)}$ form that is well known as a periodic Sinc function and is encountered in digital filter design. Fig. 5.16 shows MATLAB simulations for reverse gain expression for different number of stage DAs.

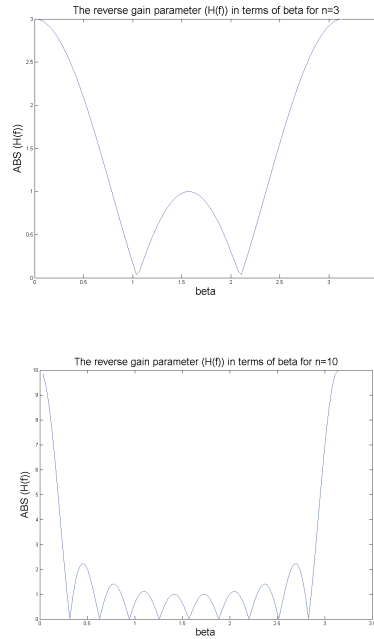


Figure 5.16: MATLAB simulations of reverse gain in hypothetical 3 and 10 stage DAs.

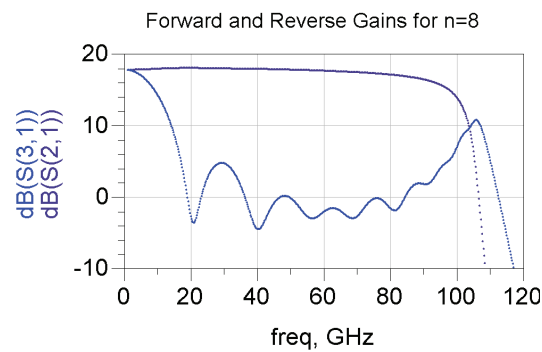


Figure 5.17: Simulations for forward and reverse gain of 8-stage DA.

The “bath-tub” shape is confirmed by simulations of a DA as shown in Fig. 5.17. This figure compares the forward and reverse gain of a DA that is designed using level 1 SPICE

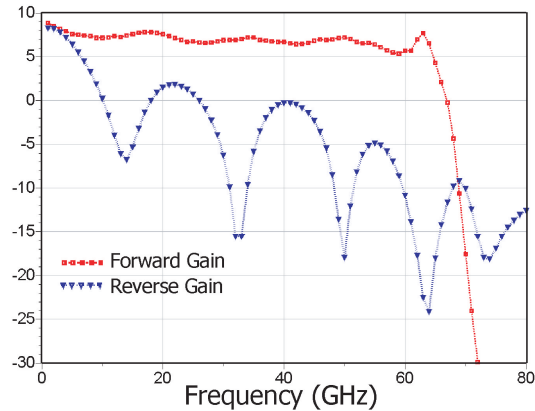


Figure 5.18: DA forward and reverse gain simulation using realistic device and transmission line models.

models. The strange shape on the high frequency side is due to operation close to the cut-off frequency of the transmission line. Another important point is that the reverse gain experiences a dip proportional to n and therefore the forward-to-reverse gain ratio is proportional to n^2 in pass band. Also, it can be seen from Fig. 5.17 that the effect diminishes when $\beta = \pi$. Fig. 5.18 shows the forward-reverse gain simulation with a more realistic device and transmission line mode that incorporates losses.

In the core stage of the feedback DA (FBDA), the gain from port 1 to port 3 is the forward gain. If we assume that the number of devices in the core stage is sufficient, then there will not be any gain from port 1 to port 2 (reverse gain) at this point. Also, the FBDA is AC coupled for biasing purposes, and therefore, the low frequency failure of the reverse gain rejection is not of interest. Now, the signal is gained up at port 3 and after going through a delay and filtering element it gets to port 4. From port 4 to port 2, the signal experiences a similar gain to that of port 1 to port 3 which is the forward gain of the DA. From port 4 to port 3 there will not be any gain as discussed in this section and the signal does not “fall back” onto itself. Therefore, the loop is stable and the signal goes through the loop only twice as desired.

The fact that the signal does not fall back onto itself deserves more attention at this point. The quantitative picture using forward and reverse gain was explained. Qualitatively, different currents coming to port 3 from port 4 do not add up in phase and in fact cancel out as they get to port 3 (reverse gain). This is similar to a set of vectors in the plane adding up and being equidistance on a circle. The addition would result in zero or a small residual vector.

A heuristic analysis of the signal at port 2 due to the initial signal at port 1 of the core section when the feedback is in place could be done by taking into account all the current and voltage terms simultaneously. This analysis could be useful for cases where n (number of devices in the core section) is not too large as to neglect the reverse gain. Here, we

will assume that the delay in the gate and the drain lines are equal. Also, the analysis is for mid-band where the delay of transmission lines is not very small to make the circuit a lumped circuit with feedback. The current at the output of port 2 is given by:

$$I_d = \frac{1}{2}\{I_1 + I_2e^{-j\beta} + \dots + I_n e^{-j(n-1)\beta}\} \quad (5.12)$$

The other half of the current flows to port 3 which then finds its way to port 4. Let us calculate the equivalent total voltage at the gate of the transistors due to the input voltage and the current flowing back. Once these voltages are calculated, using the above equation we can find the equivalent gain of the stage. Notice that this gain takes all forward and reverse gains into account at the same time. For the gate voltages we have:

$$\begin{aligned} V_k = & V_{in}e^{-j(k-1)\beta} - \frac{Z_0}{2}I_1e^{-j[(n-1)\beta+\beta+(n-k)\beta]} - \\ & \frac{Z_0}{2}I_2e^{-j[(n-2)\beta+\beta+(n-k)\beta]} - \dots - \frac{Z_0}{2}I_n e^{-j[\beta+(n-k)\beta]} \end{aligned} \quad (5.13)$$

Using this representation and the fact that the current and voltage of the device are related through the transconductances, we have a set of mutually coupled equations for currents I_1 to I_n that could be solved simultaneously. In order for the analysis to be realistic, line characteristics should also be taken into account. The results will then resemble that of Fig. 5.17.

Noise Analysis

For large bandwidth pre-amplifiers for optical communication modules or for other receivers requiring high sensitivity across the band, noise performance is critical. As another example, a wideband imaging system employing a DA at its front-end requires very good sensitivity to suppress the background and to provide acceptable contrast. Also, for such systems the linearity of the receiver is very important as usually there happens to be an overwhelming jamming signal (e.g. initial reflections from the skin). Noise figure of DA has been analyzed using matrix and other methods in the literature [86] [87]. The analysis here follows a similar methodology.

In general, as shown in later sections, the noise figure of distributed amplifiers has a bathtub shape. This is because at very low frequencies the electrical lengths of the lines are very short and the gate termination resistance contributes to the output noise as much as the source resistance and therefore minimum theoretical noise figure without even considering active elements is 3dB. Of course with active elements contributing, noise figure passes the 5 or 6dB marks. At higher frequencies the noise figure rises as the line starts to exhibit larger impedances than that of Z_0 and the cut-off frequency is approached.

A very interesting observation in the noise analysis of distributed amplifiers is the fact that the noise figure reduces with the addition of devices. Adding new stages will add

linearly to the output signal. The noise however adds in power at the drain side and results in improved SNR as the number of devices is increased.

To calculate the noise figure of the system, once again, we partition the overall FBDA into three sections; the input stage, core stage and the output stage. We will assume here that the impedance is chosen to be Z_0 globally and hence we can take advantage of cascaded noise calculations once the noise figure of each section is derived. If the impedances are changed (as they would be in practice), available power calculations for noise should be performed.

Noise of a Conventional DA

First, we will calculate the noise figure of a general DA. In this derivation we will use Van der Ziel noise sources at the gate and the drain of the amplifier. Also, as the correlation between the two current noise generators is complex and the noise drives real impedance, we have neglected the correlation (approximation). Similarly a voltage noise source could be used at the gate to neglect the correlation and also to include the noise from any poly resistance as well as the NQS element [88].

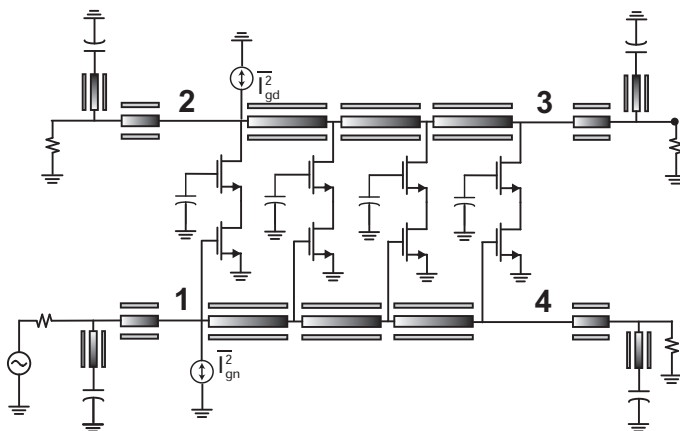


Figure 5.19: Calculated normalized gain for a conceptual feedback DA.

The reference figure in our noise analysis is Fig. 5.19 with noise current source added in the gate and drain of the devices (here just shown for the first stage). The contribution from the gate terminating resistance is by a reverse gain factor to the output while the output noise from the source resistance is through a forward gain element. Also, the drain terminating resistance contributes noise directly to output (kTB) from a simple voltage divider. The two main noise contributions come from the active devices themselves. The following equations govern the values of the noise elements in MOS devices:

$$\begin{aligned}\bar{i}_g^2 &= \frac{4KT B \omega^2 C_{gs}^2 \delta}{5g_m} \\ \bar{i}_d^2 &= 4kTB\gamma g_m\end{aligned}\quad (5.14)$$

Noise from the gate of the k^{th} MOS device gets to the output port (port 3) by two mechanisms. First by forward gain of the succeeding stages and second by reverse gain due to earlier stages. The resulting current is added and the result noise powers of all devices then summed to find the cumulative effect. The current due to the k^{th} device at the output from forward amplification is:

$$\begin{aligned}I_{out}(k) &= \frac{g_m i_{gk} Z_0}{4} \{e^{-j[(n-k)+0]\beta} + e^{-j[(n-k-1)+1]\beta} + \dots + e^{-j[(n-n)+(n-k)\beta]}\} \\ &= \frac{g_m i_{gk} Z_0}{4} (n - k + 1) e^{-j(n-k)\beta}\end{aligned}\quad (5.15)$$

The extra 0.5 factor comes from the current division on the gate side where only half of the current flows towards the right. As can be seen the combination of right-going gate current and that of right-going drain current adds in phase and produces a factor at the output proportional to the number of devices to the right of the k^{th} device. The output current from the k^{th} device at the output due to the reverse gain is as follows:

$$I_{out,R}(k) = \frac{g_m i_{gk} Z_0}{4} e^{-jn\beta} \frac{\sin(k-1)\beta}{\sin(\beta)}\quad (5.16)$$

The vector sum of the two components of the currents at the output would be:

$$\begin{aligned}I_{out-tot}(k)^2 &= \frac{(g_m i_{gk} Z_0)^2}{16} \\ &\{(n - k + 1)^2 + \left[\frac{\sin((k-1)\beta)}{\sin(\beta)}\right]^2 + 2(n - k + 1) \frac{\sin((k-1)\beta)}{\sin(\beta)} \cos(k\beta)\}\end{aligned}\quad (5.17)$$

If the number of devices n is large then the first term dominates and the other terms due to the reverse gain and the interaction can be neglected for simplification. The expression in the second large parentheses is denoted by $g(n, k, \beta)$. The output current due to the drain noise of the k^{th} device is $I_d/2$ from a simple current divider directing only half of the current to the load.

The total noise figure of a conventional DA adds up to be:

$$\begin{aligned}F &= 1 + \frac{G_F}{G_R} + \frac{1}{G_F} + \frac{Z_0 \omega^2 C_{gs}^2 \delta \sum_{k=1}^n g(n, k, \beta)}{n^2 g_m} + \frac{4\gamma}{n g_m Z_0} \\ &= 1 + \left[\frac{\sin(n\beta)}{n \sin(\beta)}\right]^2 + \frac{4}{n^2 g_m^2 Z_0^2} + \frac{Z_0 \omega^2 C_{gs}^2 \delta \sum_{k=1}^n g(n, k, \beta)}{n^2 g_m} + \frac{4\gamma}{n g_m Z_0}\end{aligned}\quad (5.18)$$

where G_F represents the forward gain of the amplifier from port 1 to 3 (Fig. 5.19) and G_R is the reverse gain (from port 1 to port 2). The second term is due to the gate termination resistance and can be reduced by increasing number of stages. This term starts to show up itself at the lower and higher bounds of the frequency response. The third term is due to the termination drain resistance and is generally negligible due to being at the output and getting divided by the gain. The fourth term is due to the gate noise of the MOS devices [87]. If we assume the number of devices is large, then we can approximate $g(n, k, \beta)$ by $(n - k + 1)$. With this approximation the sum term in the fourth term in the noise figure can be approximated by $n(n + 1)(2n + 1)/6$. From this expression the fourth term would be proportional to n while the fifth is inversely proportional to n . This dictates an optimum number of devices for the given device parameters at certain operating frequencies. Fig. 5.20 illustrates the analytically derived noise figure for various number of devices in a DA. The data is based on a 90nm process with f_t of 100 GHz and with a line cut-off frequency of 100 GHz.

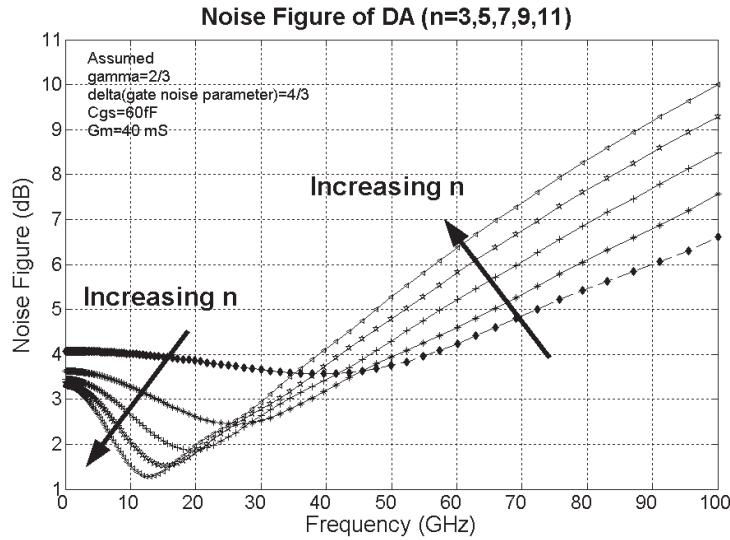


Figure 5.20: Theoretical NF of DA in terms of frequency for different n .

Noise of a the Feedback DA

The noise figure of the whole FBDA is just the cascade of the three noise figures. The NF of the input and output stages have been determined in the previous section. Here the NF of the core stage will be derived. Without loss of generality, some simplifications are made to facilitate the analytical derivation of noise figure for this structure. As previously mentioned, in case of unequal impedances, extra precautions must be made. For this analysis Fig. 5.19 is used with the difference of ports 3 and 4 being connected by a transmission line

of electrical length equal to β and the termination components of these ports taken away.

The noise due to the gate of the devices has four components based on the direction of the flow of the current in the gate and the drain lines. Also, there is another component from the gate current going directly to the output which could be neglected as it is smaller compared to other terms that are gained up. If the direction of current in the gate-line is the opposite to that of the drain-line, then the gain factor would be similar to the reverse gain explained in previous sections. The reverse gain rejection is large enough for large n values and therefore we will neglect such terms in the calculation of the noise figure of the core stage. This forward/reverse ratio is only large for mid-band frequencies. At the higher frequency range, where the DA ceases to provide gain, the ratio drops. The ratio is also small for the very low frequencies but these frequencies are filtered out in the feedback DA.

The first term is from the gate and drain components of the signal going to leftwards to port 2. In other words, the gate current noise of the k^{th} device has a forward gain path to the output by directly traveling to the left (port 2 is the output in the FBDA). The drain components related to this will also add in phase. This component can be described as

$$\begin{aligned} I_{1,out}(k) &= \frac{g_m i_{gk} Z_0}{4} \{e^{-j[\beta+(k-2)\beta]} + e^{-j[2\beta+(k-3)\beta]} + \dots + e^{-j[(k-1)\beta]}\} \\ &= \frac{g_m i_{gk} Z_0}{4} k e^{-j(k-1)\beta} \end{aligned} \quad (5.19)$$

The other component of the gate current initiates a wave that travels towards the feedback port (port 3) of this stage. This component can be expressed as:

$$I_2(k) = \frac{g_m i_{gk} Z_0}{4} \{e^{-j[(n-k)\beta]}\} (n - k + 1) \quad (5.20)$$

This component is fed back to port four (through the feedback connection) and once again travels through the devices experiencing the gain. The output current in port 2 is now:

$$I_{2,out}(k) = \frac{g_m^2 i_{gk} Z_0^2}{8} \{e^{-j[(n-k+1)\beta]}\} (n - k + 1) n \{e^{-j[(n-1)\beta]}\} \quad (5.21)$$

Here a delay of β is assumed through the feedback path (as has been the case in prior analysis as well). The two terms need to be added taking into account their phases and the fact that one has a negative sign with respect to the other from experiencing the negative gain twice. With the extra n in the latter, we can safely neglect the effect of former as a first order approximation.

The forward voltage to current gain factor of the core stage with feedback is

$$G_{mT} = \frac{I_{oF}}{V_{in}} = \frac{1}{2} Z_0 \left[\frac{ng_m}{2} \right]^2 \quad (5.22)$$

With the above definition, summing currents we have:

$$I_{d,g}^2 = Z_0^2 \frac{G_{mT}^2}{n} \bar{i}_g^2 \sum_{k=1}^n (n - k + 1)^2 \quad (5.23)$$

The drain noise of the devices could also contribute to the output noise. This can be calculated by noticing that the main contribution would be from the current flowing towards the feedback port and going through the gain stages. This results in a gained current at the output.

$$I_{d,d}^2 = \left(\frac{1}{4}\right)^2 n \bar{i}_d^2 (g_m n Z_0)^2 = \frac{\left(\frac{G_{mT}}{g_m/2}\right)^2 \bar{i}_d^2}{n} \quad (5.24)$$

The noise figure of this stage ends up being

$$F = 1 + \frac{Z_0 \omega^2 C_{gs}^2 \delta n (n+1) (2n+1)}{30 n^2 g_m} + \frac{4\gamma}{n g_m Z_0} \quad (5.25)$$

The analysis shows that the essential elements of the noise figure in the core element of the FBDA is very similar to a conventional DA (it is just that the gain is experienced two folds). The equation above slightly underestimates the noise figure due to all the approximations made in this section. These approximations assumed large gain (or number of DA stages). The benefit of using FBDA is the elimination of the gate and drain termination resistances. The elimination of the gate termination reduces the slope of the rise in noise figure close to the higher frequency portion of the band. In general though, the noise figure of FBDA could be actually slightly higher than DA because of the following two reasons. First, in the core we have several less dominating noise paths that we neglected in the calculations. Second, the overall noise figure is from a cascade of three blocks and this could generate more noise if not properly designed. On the other hand, with careful design of the input stage DA, one can improve the noise performance of a FBDA.

As proof of concept and for verification purposes, a FBDA is designed with level 1 MOS models with main device parameters similar to a 90nm process with $f_t=100\text{GHz}$. A conventional DA using same number of devices and biased with the same DC current is also designed to comparative purposes. Fig. 5.21 and Fig. 5.22 illustrate the gain and noise figure of the designed DA and FBDA respectively. The two cases use the same number of stages. The gate losses are assumed to be mainly from NQS resistance and partly (1/3) from poly resistance. The inductor Q is assumed to be 10 at 10 GHz for the gate and drain lines. In comparison, the FBDA achieves a very large gain with comparable noise performance. The FBDA has been designed for maximizing the gain (rather than NF) while in practice the input stage could be used to improve the NF.

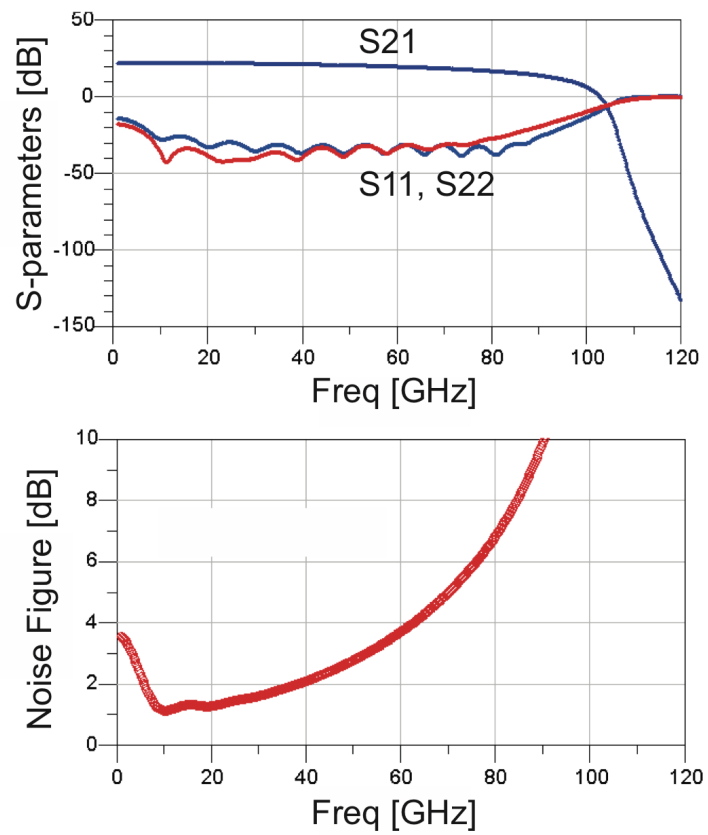


Figure 5.21: Simulated gain and NF of a 16 stage DA in level 1 MOS models.

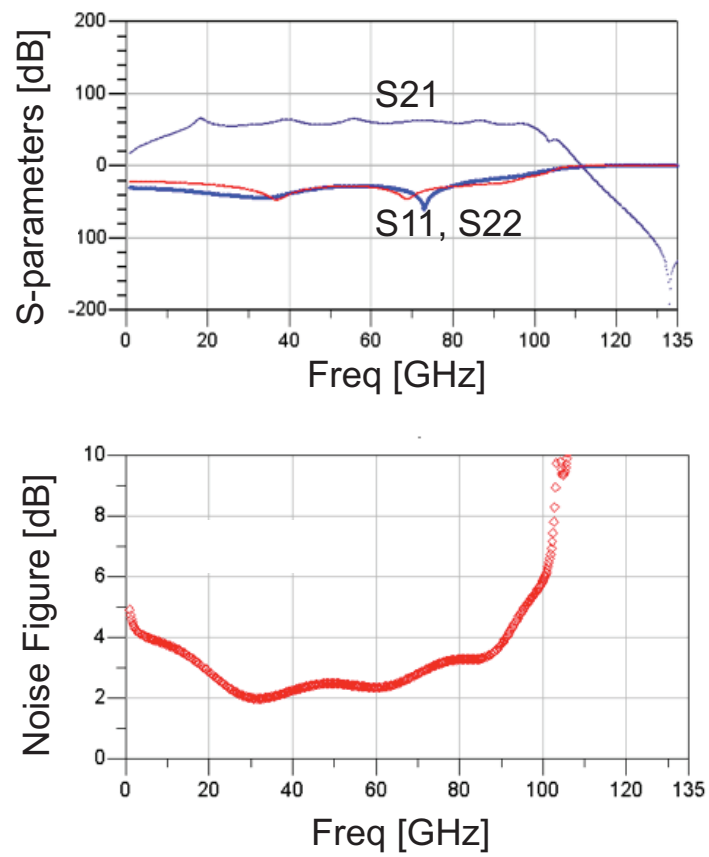


Figure 5.22: Simulated gain and NF of a FBDA with the same number of stages.

Measurement Results

The FBDA uses $40\mu\text{m}$ cascode device that were selected based on the mechanisms introduced in section 5.2.3. The cascodes are biased with 0.6 V (gate-bias) to draw 7 mA from the 1.2 V supply.

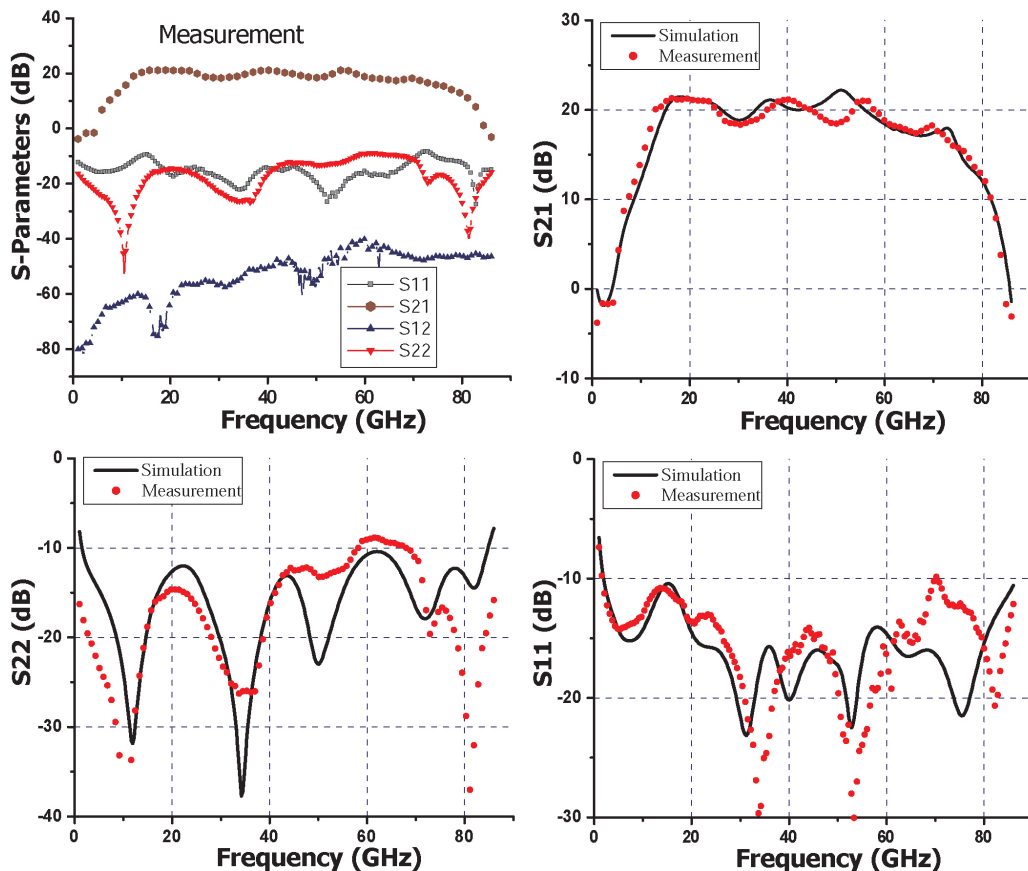


Figure 5.23: Measured s-parameters of the FBDA.

Measurements are performed using on-wafer probing. Fig. 5.23 shows the s-parameters of the amplifier. As illustrated, the return loss stays better than 9 dB in the band and the reverse isolation (S_{12}) is better than 40dB. This is due to the use of cascode elements and also the topology of the amplifier. The output 1dB compression point varies between 3.7dBm and 0.3dBm and the noise figure between 5.2 and 6 dB in the 15-45 GHz band (Fig. 5.24). Noise figure is measured using a noise source and noise meter for frequencies below 26GHz (Method 1). After this frequency, measurements are done using the Gain method using an external amplifier and a spectrum analyzer (Method2). The discrepancy between the two methods at 25 GHz is 0.5dB. The chip is fabricated in a 90nm digital CMOS process (no extra RF options) with power consumption of 84mW and chip area of

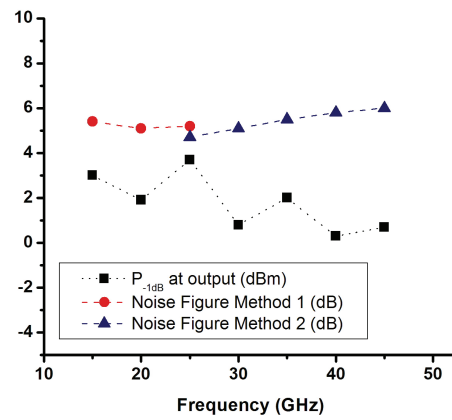


Figure 5.24: Measured noise and P_{1dB} of the FBDA.

1.5 mm by 0.79 mm. The performance comparison to previous CMOS DAs is given in Table 5.1 and the chip micrograph is shown in Fig. 5.25.

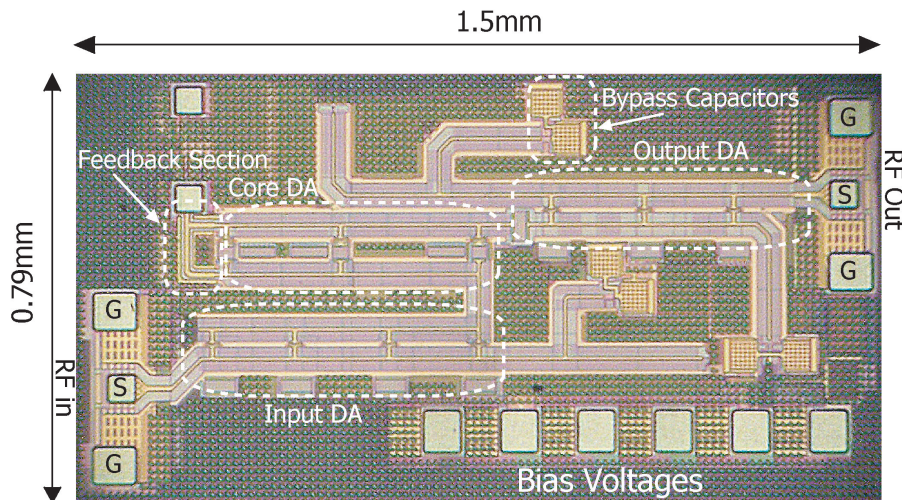


Figure 5.25: Chip micrograph of the FBDA.

Table 5.1: Performance comparison of the FBDA to the prior state of the art.

Ref.	Lui, ISSCC 2005	Kim, ISSCC 2004	Tsai, ISSCC 2005	Shimegatsu, ISSCC 05	Moez, ISSCC 2007	Chien, ISSCC 2007	This work (ISSCC 08)
Tech.	1P9M-90nm RF-CMOS	0.12 μ m SOI-CMOS	1P9M 90nm CMOS	0.18 μ m CMOS	0.13 μ m CMOS	0.18 μ m CMOS	1P7M 90nm CMOS
GBW (GHz)	190	320	157	62	136	394	660
S21(dB)	7.4	11	7	4	9.8	20	19
BW (GHz)	80	90	70	39	43.9	39.4	74
NF (dB)	N/A	4.8-6.2 <18 GHz	6-6.9 <25 GHz	N/A	2.5-7.5 <40 GHz	8-9.4* <18 GHz	5.2-6 <45 GHz
S11/S22 (dB)	<-10/<-8	<-7/<-5	<-7/<-12	<-10/<-10	<-14/<-8	<-10/<-20	<-9.5/<-9
P _{-1dB} (dBm)	6-8*	12	10	N/A	N/A	6.5	3.7 @25 GHz
P _{diss} (mW)	120	210	122	140	103	250	84
V _{dd} (V)	2.4	2.5	N/A	N/A	N/A	N/A	1.2
Area (mm ²)	0.72	1.28	0.72	3.3	1.5	2.24	1.19
f _t (GHz)	N/A	196	160	51	N/A	50	100
1000.FOM	NF N/A	34.7	18	NF, P _{-1dB} N/A	P _{-1dB} N/A	19	51

$$FOM = \left(\frac{GBW}{f_t} \right) \left(\frac{P_{-1dB}}{P_{DC} \cdot F_{N,avg}} \right), F_N = \text{Noise Factor}$$

*From the ISSCC presentation

5.2.5 Tapered-Cascaded Multi-Stage Distributed Amplifier

The multi-stage DA has the advantage of having an extra degree of freedom in the choice of internal idle termination impedances. For a cascade of single stage DAs, [89] suggests open terminations for maximizing the gain. However, this comes with the cost of limitations on the BW from destructive combinations of forward and backward traveling signals and limits the number of stages. This results in poor input and output return losses. Also, the idle termination technique cannot be used on the input and output lines since the return loss is not acceptable.

We propose tapering the impedance of the line segments which can also be used on the interface sections if the tapering coefficients are not too large as to cause undesired matching properties for the amplifier. In this tapering, the line segment impedances is tapered starting from the load impedance of 50Ω and increased by \sqrt{K} (the tapering coefficient) per stage. This could be achieved by the change in line lengths (and hence in the inductances) or by varying the spacing/height of the elevated-CPW to change the Z_0 of the transmission lines. In this work, we have employed the former technique for better predictability of the equivalent section impedance change. The active elements are kept identical in size and hence reflections occur. This is in contrast with previous work on downsizing both the active and passive impedance for improved bandwidth [79] or output power [90] [63].

To analyze a tapered synthesized line one can follow several approaches. To begin we predict the impedance of a uniform line based on the $ABCD$ matrix of cascaded two-port sections [47]. The image impedance can be calculated as follows

$$Z_{in} = \frac{DZ_L + B}{CZ_L + A} \quad (5.26)$$

Here, the $ABCD$ parameters are used to describe the input impedance of a T-section (Fig. 5.26). When an infinite number of sections are added, the input impedance becomes

$$Z_{in} = \sqrt{\frac{B}{C}} = \sqrt{(Z_1 + Z_2)Z_3 + Z_1Z_2} = \sqrt{\frac{L}{C}} \sqrt{1 - \left(\frac{\omega}{\omega_c}\right)^2} \quad (5.27)$$

Eq. 5.27 shows the impedance seen at the input of many identical cascaded T-sections with $\omega_c = \frac{2}{\sqrt{LC}}$. It can be seen that if the frequency is much lower than the cutoff frequency of the line (a good assumption with high impedance transmission lines), then the familiar $\sqrt{\frac{L}{C}}$ is a good approximation.

One can observe the tapered transmission line in a DA as a loaded transmission line in which the loading is not periodic. Therefore, if (by approximation) the capacitive loadings are absorbed into the specific sections, assuming a uniform absorption, leads to $Z_i = \sqrt{\frac{L}{C + \frac{C_p}{\Delta X_i}}}$ for impedance of each section. Taking a step further, one may neglect the line capacitance with respect to the parasitic capacitance of active elements, a relatively good approximation for high- Z_0 lines. Effectively, the line segment acts as an inductor and one may in fact assume equal length cascaded lines with non-equal impedances. In the limit of short segments, this would lead to conventional continuous impedance tapering (as in [47]) applied for matching purposes. The internal reflections are approximated as

$$\Delta\Gamma = \frac{Z + \Delta Z - Z}{Z + \Delta Z + Z} \simeq \frac{\Delta Z}{2Z} \quad (5.28)$$

If the segments are not so small for the continuous approximations [47] to hold, one can sum all the reflections in (5.28). This will lead to

$$\Gamma(\theta) = \sum_1^N e^{-j2\beta X_i} \frac{\Delta Z}{2Z} \quad (5.29)$$

Here, X_i is the sum of the length of all the lines before the i^{th} reflection. In this work, as briefly discussed above, a multiplicative taper is assumed with the approximation of the length of the line being scaled by K and the impedance by $\alpha = \sqrt{K}$. Quantitatively, for this tapering profile, $Z_{m+1} = \sqrt{K}Z_m = \alpha Z_m = (1 + \delta)Z_m$ and one may obtain the following

$$\Gamma(\theta) = \sum_1^N e^{-j2\beta X_1[\sum_0^{i-1} K^i]} \times \frac{Z_0\alpha^i - Z_0\alpha^{i-1}}{Z_0\alpha^i + Z_0\alpha^{i-1}} \quad (5.30)$$

The final fraction in (5.30) is equal to $\frac{\delta}{2}$ for small α values close to unity. With this tapering profile, all the reflections are of equal size ($\frac{\delta}{2}$). Taking advantage of the propagation loss effect ($e^{-2\alpha L}$), one can allow the impedances further away from the termination to be more abruptly increased since their effect is somewhat mitigated by the loss. With the above equations and knowing device and transmission line frequency characteristics and also the VSWR tolerances, one can obtain the tapering coefficient.

From a different perspective, one can combine a linear incremental or exponential taper with filter synthesis methods [91] to achieve the impedance profile desirable for higher bandwidths in spite of gain roll-off of active or passive elements. From a filter theory perspective, once poles are brought closer to the imaginary axis in the synthesis process, the peaking will increase which can potentially cancel droops of active elements. Another factor that needs to be taken into account is the line cutoff frequency (or equivalently the low-pass filter cutoff) that is obtained once the tapering and other techniques are applied. Large segments of the line tend to dominate the cutoff frequency and this causes problems especially for more aggressive tapering coefficients.

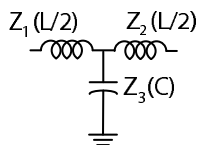


Figure 5.26: T-section of a synthesized transmission line.

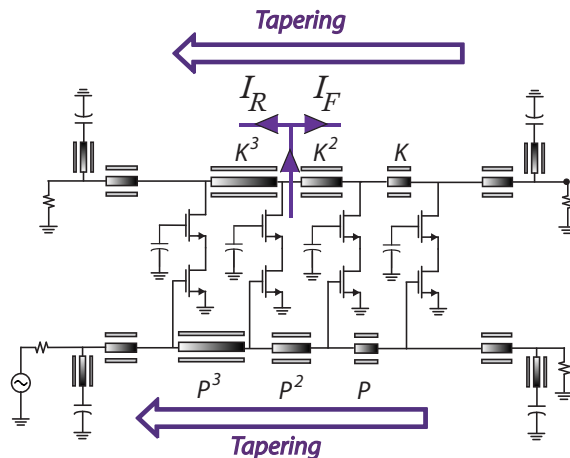


Figure 5.27: Schematic diagram of a tapered distributed amplifier.

In a distributed amplifier the current injection on the drain lines favors one direction for tapering rather than the other. One would desire that currents be directed towards the final load rather than the idle termination and this would for example favor an “up-tapering” from the load impedance on the drain line towards the idle termination. Fig. 5.27 shows the schematic diagram of a distributed amplifier with tapering from the load. It is important to notice that apart from the reflections seen by the load impedance (determining the S_{22} of the amplifier), internally, each of the active elements’ current is divided unequally (I_R and I_F in the figure) at the drain and also is reflected multiple times before getting terminated on either side. The current divisions at the drain can be formulated using the section impedances as follows

$$I_{F,i} = \frac{Z_i}{Z_{i+1} + Z_i} I_i = \frac{1 - \Gamma_i}{2} I_i \quad (5.31)$$

$$I_{R,i} = \frac{Z_{i+1}}{Z_{i+1} + Z_i} I_i = \frac{1 + \Gamma_i}{2} I_i \quad (5.32)$$

where the index i increases towards the load to the right. Each of the $I_{F,i}$ and $I_{R,i}$ components are reflected at all the intersections to left and right before terminating at the load.

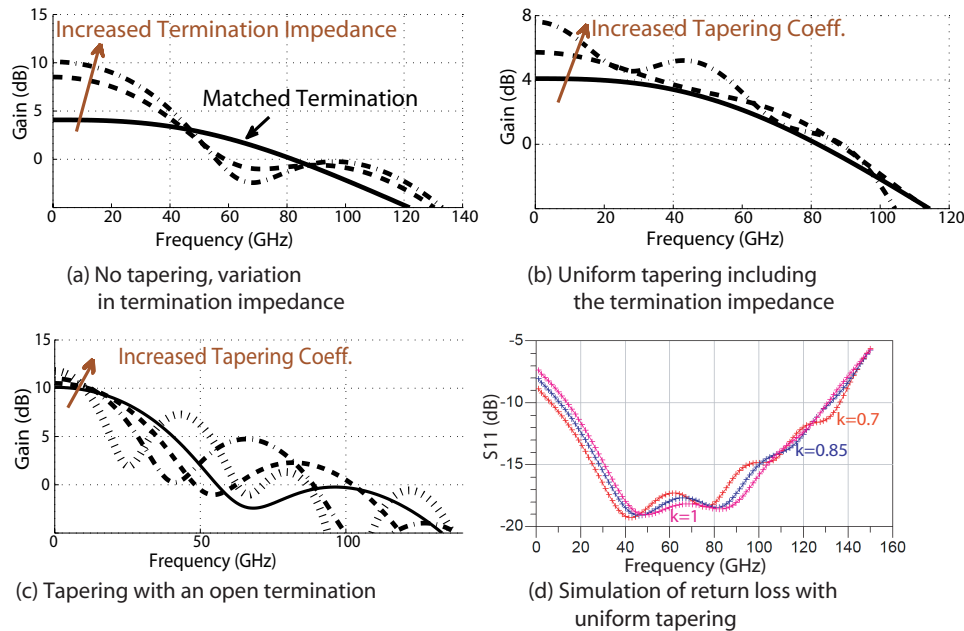


Figure 5.28: Simulated gain with (a) no tapering, (b) uniform tapering, (c) tapering with an open termination and also (d) the return loss from uniform tapering.

To calculate the reflections, we need to keep track of the phase of the combinational signals. The Γ functions can be approximated by $\frac{\delta}{2}$ or rather by the continuous approximation assuming smooth impedance transitions.

Fig. 5.28 illustrates the effect of tapering by first order simulations. Here, a simple model of DA neglecting higher order reflections, that are mitigated by loss, has been utilized to provide intuitive results. Fig. 5.28a shows the effect of having no tapering and terminating a DA drain line with various large loads. Fig. 5.28b uses multiplicative tapering all the way to the terminating resistance (with the same coefficient). Fig. 5.28c uses an open termination with various tapering coefficients for the segments. Here, there is a relatively large mismatch between the last segment and the termination element. In Fig. 5.28d, first order circuit simulations have been used for the case of uniform multiplicative tapering (as in part b of this figure) to verify possible application in the input and output stages. As observable in this figure, tapering provides means of extending the gain while having control over different local gain variations (similar to conditions with pole/zero placement).

Measurement Results

A tapered cascaded multi-stage distributed amplifier was designed and implemented in a 90nm 1P7M digital CMOS process with no additional RF options. The native NMOS device has a post-layout $f_T = 100$ GHz. The schematic of the T-CMSDA is shown in

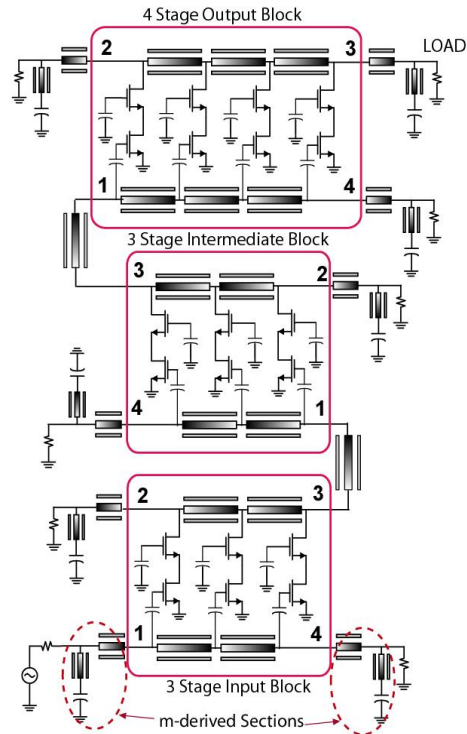


Figure 5.29: Schematics of the T-CMSDA.

Fig. 5.29. Multi-stage amplifiers allow for extra degrees of freedom in the choice of internal and external termination impedances. The amplifier is tested in a 50Ω environment. The chip micrograph is shown in Fig. 5.31. The chip consumes an area of 1.15mm by 1.5mm .

A large series capacitor is used to AC-couple the input of the amplifier. M-derived matching sections are used to improve matching to required impedances. Intermediate terminations as well as the input and output terminations are tapered (with different coefficients) according to the descriptions in previous section. To achieve the required frequency response, both line segment sizes and input capacitance at the intersection nodes are varied. Series capacitors are used to control the equivalent input capacitance.

Measurements were taken directly using wafer probes. All the pads and parasitics are included in the design and hence the measurements. The measured s -parameters are shown in Fig. 5.30. The amplifier has an average pass-band gain of 14dB with a 3dB bandwidth of 73.5GHz . The zero-dB bandwidth of the amplifier is at 83.5GHz . The S_{11} and S_{22} of the T-CMSDA stay below -9dB up to 77GHz and 94GHz , respectively. The GBW product of this amplifier is 370GHz . The zero-dB gain-bandwidth is 419GHz . The output referred 1-dB compression point is shown in Fig. 5.32. The output power remains higher than -0.2dBm up to 60GHz . Fig. 5.33 demonstrates the group delay of the T-CMSDA in the frequency band of interest. At low frequencies the group delay variations are due to the AC

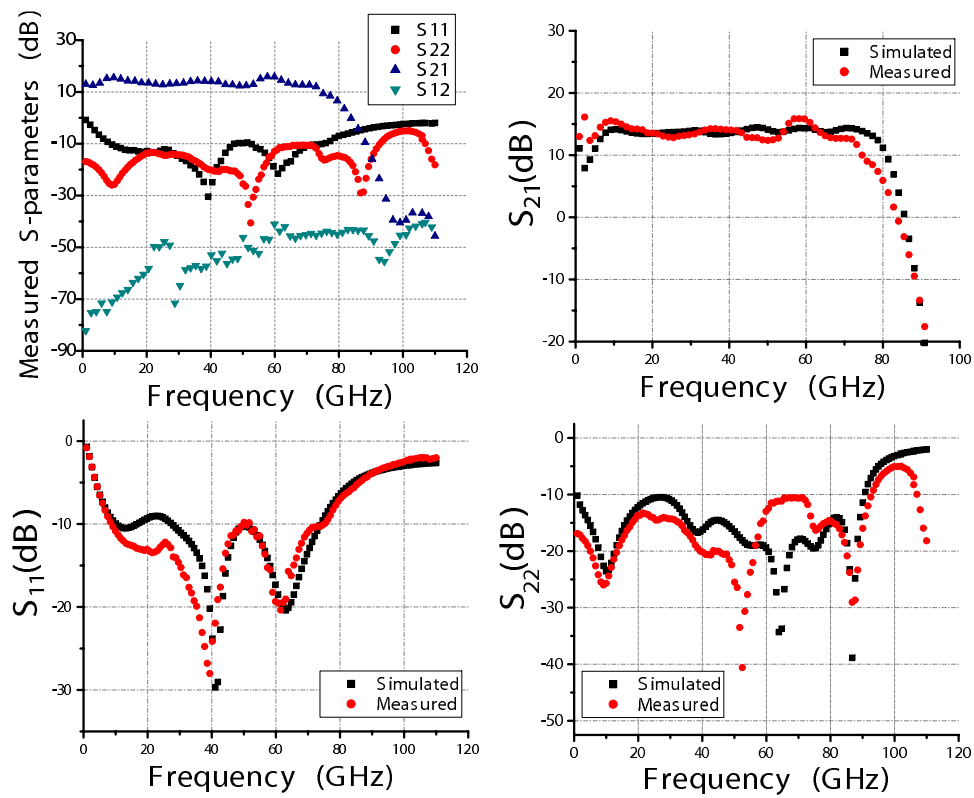


Figure 5.30: S-parameter simulation and measurements of the T-CMSDA.

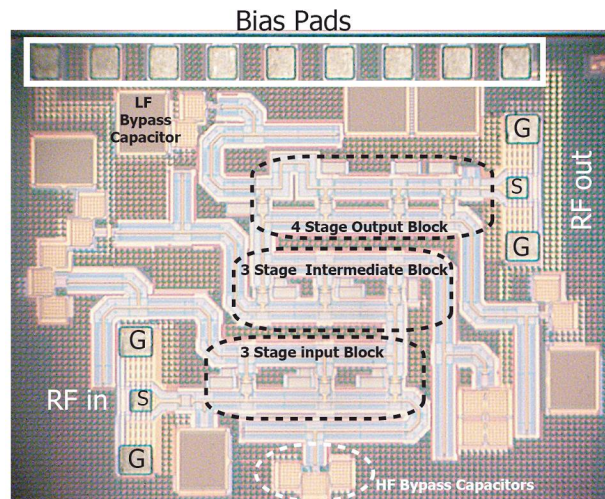


Figure 5.31: Chip micrograph of the of the T-CMSDA.

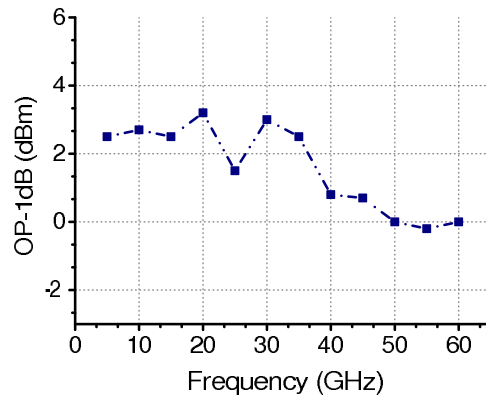


Figure 5.32: Output compression point measurements of the T-CMSDA.

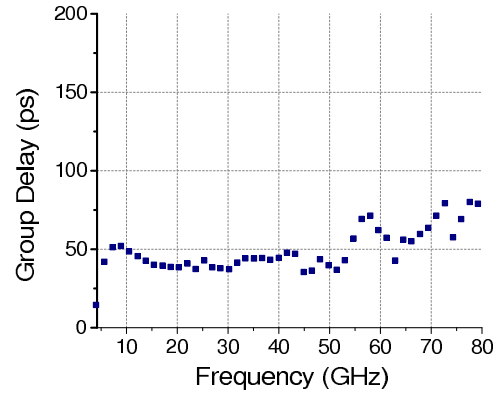


Figure 5.33: Measured group delay of the amplifier in frequency band of interest.

coupling capacitor used. The amplifier draws 70mA from a 1.2V supply. Comparison to other published CMOS DAs is given in Table 5.2.

Table 5.2: Comparison table for the tapered DA (RFIC 2008, [92])

Ref.	This Work	[80]	[83]	[93]	[61]	[79]	[81]
Process	90nm CMOS	90nm RF CMOS	90nm CMOS	130nm CMOS	0.18 μ m CMOS	0.18 μ m SiGe CMOS Only	0.12 μ m SOI CMOS
GBW (GHz)	370	190	157	136	394	61	144
S21 (dB)	14	7.4	7	9.8	20	7.8	4
BW (GHz)	73.5	80	70	43.9	39.4	25	91
S11/S22 (dB)	-9/-9	-10/-8	-7/-12	-14/-8	-10/-20	-10/-10	-7/-7
OP1dB (dBm)	3.2 @20GHz	6-8	10	N/A	6.5	4.2	9
Vdd (V)	1.2	2.4	N/A	N/A	N/A	1.8	2.6
Power (mW)	84	120	122	103	250	54	90
Area (mm^2)	1.72	0.72	1.28	1.5	2.24	1.32	0.8

5.2.6 SiGe BiCMOS Distributed Amplifier for the Imager Front-End

Design Overview

Based on the experiences gained from previous designs, an extremely wideband amplifier for the imager front-end was conceptualized. Contrary to previous designs, this amplifier uses a 130 nm SiGe BiCMOS process (same process described in chapter 4). In this process, the cutoff frequency of the core device is 230 GHz. Even with the addition of connections the cutoff frequency is still high and close to 200 GHz. The 3 dB corner frequency we aim for in the front-end is 130 GHz. Therefore the $\frac{f_{-3dB}}{f_t}$ ratio is smaller than what was realized in the previously described CMOS DAs (where we achieved $\frac{f_{-3dB}}{f_t} \simeq 0.74$). In addition to that, potentially, the core device gain is higher in the bipolar devices than what was available in the 90 nm CMOS process. Therefore, the architecture design of the DA has to take into account these opportunities as well challenges that will be described later in order to lead to an optimal solution. A three-stage cascaded DA with uniform gate and drain lines is utilized here.

Some of the key challenges will be briefly described here. First, we will briefly describe the transmission line used for obtaining a high characteristic impedance. We will then describe the design of the active element and challenges related to maintaining a flat response in terms of gain and input impedance. We will then describe one of the key challenges that is rarely addressed in DA design and that is with providing biasing to the collector lines. Often, DA designs use biasing either through series resistors or through external chokes. The former requires raising the supply voltage to compensate for the voltage drop across the resistor and is costly in terms of power consumption. It also caused potential BW issues, especially for larger BW DAs, due to the parasitic capacitance of the resistors. External chokes are not suitable for integrated system. Here we describe an integrated inductive choke that provides the required BW.

Passive Element Design

As previously mentioned, the design of a wideband DA requires a simultaneously low-loss and high-impedance line. Often times, the two requirements are contradictory due to the conductive losses of the line as outlined in previous sections.

Since the dielectric stack height is larger in this process compared to the CMOS process, a microstrip line was used as the element in the synthesized line. Details of the process stack and line losses can be found in the literature [57] [59]. The transmission parameters are shown in Table 5.3.

Table 5.3: Transmission line parameters.

Line Width	$2\mu m$
Z_0	82Ω
Loss (@70 GHz)	0.69dB/mm
$\epsilon_{\text{eff}}@100\text{GHz}$	3.9

Active Element Design

The target bandwidth of the amplifier is 130 GHz with the cutoff frequency of the device at 200 GHz. The cutoff frequency, however, only determines the current driven response of the device whereas in a realistic drive scenario other poles come into play. For the distributed amplifier, placing optimal matching networks to obtain the maximum stable gain (MSG) is not an option. The design is extremely wideband and assumes a dominantly capacitive input impedance. As outlined by Beyer et al. [78], the input and output resistances of the device ultimately limit the DA performance. Therefore, device topology optimization has to go further than optimizing the core f_t and must, for example, maximize the total input *impedance* at the input port.

Compared to the CMOS designs, the base resistance (r_b) plays a detrimental role in determining the bandwidth. As an example, a single finger $1.5\mu m$ device with 2 mA bias current has an input profile of $C_{\text{ser}} = 20\text{fF}$ and $R_{\text{ser}} = 43\Omega@80\text{GHz}$. If we double the device size and current, we end up with $C_{\text{ser}} = 42\text{fF}$ and $R_{\text{ser}} = 24\Omega@80\text{GHz}$ which has a slightly worst ω_{ser} but nevertheless shows the tradeoff between capacitance and resistance at the input. It is clear that with these values obtaining 130 GHz of bandwidth is impossible since at and around 110 GHz we will basically have a dominantly resistive input.

In order to reduce the effect of r_b , a series input capacitor can be used. Prasad [94] has proposed using series capacitors for MESFET DAs to obtain a gain-bandwidth tradeoff as well as for larger input power handling. Here, we primarily focus on reducing bandwidth reduction effects from series base resistance. Bandwidth is obtained at the expense of higher power consumption as will be described next.

In an advanced bipolar device, the emitter contacts are placed directly on top of the emitter to reduce the parasitic resistance on the emitter side. This forces the base contacts to be on the sides of the emitter. Typically, two base contacts are placed on either side of the device [58] [95]. With this topology, there is going to be internal base resistance under the emitter. With the current in the structure, there is going to be a voltage drop across the resistance and this is more severe towards the center of the emitter (longer path). This leads to current crowding. For a structure with two-sided base contacts we can approximate this component of the base resistance as:

$$R_B = \frac{R_{sh} W_e}{12L_e} \quad (5.33)$$

In this process, the emitter width is fixed to $0.27 \mu\text{m}$. The length and number of emitters are scalable. Given that with a fixed current density the input capacitance scales linearly with the emitter length, the input series-RC pole will more or less stay constant. As shown in Fig. 5.34, we will now place a series capacitor with the input. To explore the trade-offs inherent in scaling, we will assume the device is scaled by a factor of m and $C_s = kC_\pi$. With these scaling factors we have:

$$\begin{aligned} C_{in} &= \frac{mk}{m+k} C_{\pi 0} \\ g_{m,eff} &= \frac{mk}{m+k} g_{m0} \\ I_{dc} &= mI_0 \\ \omega_{ser} &= \frac{1}{2\pi \frac{r}{m} C_{in}} = \frac{k+m}{k} \omega_{ser,0} \end{aligned} \quad (5.34)$$

Here, the subscript zero represents default values prior to device scaling. We keep the input capacitance the same by first scaling up the device ($m > 1$), and then adding the appropriately sized C_s such that $C_{in} = C_{\pi,0}$. From the above equation, this leads to the condition $mk=m+k$. With that, $\omega_{ser} = m\omega_{ser,0}$ with m times the DC current consumption. So with $m \times I_{DC}$, we have higher input series pole but the same effective transconductance. Ultimately, this sizing is limited by the fixed parasitics (e.g. trace capacitances) associated with the device.

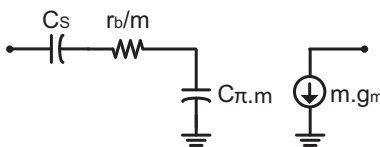


Figure 5.34: Single gain stage equivalent circuit.

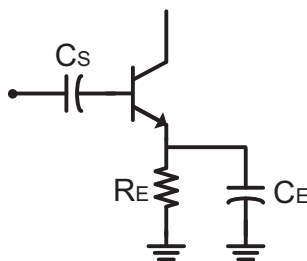


Figure 5.35: Common emitter stage with degeneration and series capacitor.

In our design, this technique alleviates the problem but does not completely solve it especially closer to the cutoff frequency of the amplifier. Emitter degeneration and capacitor-peaking (Fig. 5.35) are used to further extend the usable bandwidth of the device. Resistive

degeneration stabilizes the bias and operation of the element as well as providing yet another way of achieving lower input load with the sacrifice of DC current. A shunt capacitor provides a zero in the transfer function ($\frac{1}{2\pi R_E C_E}$ and a pole) and enhances the high frequency response of the element. Details are available in the literature (e.g. [58]).

Emitter capacitor-peaking increases the potential for instability. Emitter capacitor together with C_π resembles a Clapp oscillator. Therefore, the quest to reduce the series part of the input resistance can ultimately lead to a negative value (at least in some parts of the frequency spectrum) and that will lead to instability.

To ensure stability, we can confine the real part of the input impedance to be positive over the entire frequency range. Taking Fig. 5.35 as reference, we would like the input impedance after C_s to have a positive real part (C_s will not affect the series representation of real part). This impedance has two poles and a zero.

$$\begin{aligned}\omega_{P1} &= \frac{1}{r_\pi C_\pi} \\ \omega_{P2} &= \frac{1}{R_E C_E} \\ \omega_z &= \frac{1}{(R_E // \frac{r_\pi}{\beta+1})(C_E + C_\pi)} = \frac{1}{R_X C_X}\end{aligned}\quad (5.35)$$

After some simplification, the condition for a positive real part translates to:

$$\begin{aligned}1 + \frac{\omega^2}{\omega_z \omega_{p1}} + \frac{\omega^2}{\omega_z \omega_{p2}} - \frac{\omega^2}{\omega_{p1} \omega_{p2}} &\geq 0 \quad \text{or} \\ (R_X C_X)(r_\pi C_\pi + R_E C_E) - (r_\pi C_\pi)(R_E C_E) &\geq \frac{-1}{\omega^2}\end{aligned}\quad (5.36)$$

This inequality puts a bound on the selection of the transfer function zero ($\frac{1}{2\pi R_E C_E}$). Small values of this zero (e.g. large C_E) will lead to negative input resistance. In reality some of that is compensated by losses as well as the r_b of the device.

If we assume the operation frequency of the device to be much higher than $\frac{\omega_t}{\beta}$ and closer to ω_t , we can simplify the impedance to a form of:

$$\begin{aligned}Z(s) &= \frac{R(1 + \frac{s}{\omega_z})}{sC_\pi(1 + \frac{s}{\omega_p})} \\ \omega_P &= \frac{1}{R_E C_E} \\ \omega_z &= \frac{1 + g_m R_E}{R_E(C_\pi + C_E)}\end{aligned}\quad (5.37)$$

In the impedance representation above, in order for the real part to be positive, the pole has to come after the zero. This translates to:

$$R_E C_E \leq \frac{C_\pi}{g_m} \quad (5.38)$$

Or equivalently, $\omega_E \geq \omega_t$ where $\omega_E = \frac{1}{R_E C_E}$. This condition is conservative since it does not include base losses. In reality a slightly lower ω_E can be safely targeted. In this design $R_E = 20\Omega$ and $C_E = 45fF$ were chosen.

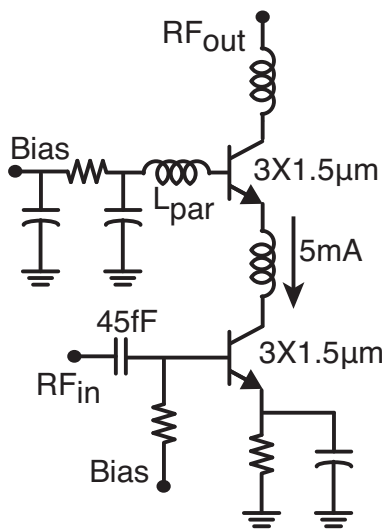


Figure 5.36: DA cascode gain element.

The overall details of the gain element is shown in Fig. 5.36. In addition to the bandwidth extension techniques discussed, a series inductor is also placed in the cascode stage to further improve the bandwidth [96]. Again, the size of this inductance will affect the input impedance and can potentially cause oscillation if it is selected to be large.

Simulations for the input loading of the gain stage is shown in Fig. 5.37.

Biasing

Biasing a wideband distributed amplifier poses several challenges. An RF choke is required to provide biasing to the drain of the devices. For the base, a large series resistor can be used since the current is substantially lower. Conventional ways of biasing the drain/collector lines include external bias-tees or series resistors. Both of these methods are ineffective once the DA is integrated into a larger system that uses the chip supply. Series resistor biasing is inefficient in terms of power and may cause reliability (narrow resistors) or bandwidth limitations (wide resistors). In addition to that, cascaded DA topologies further

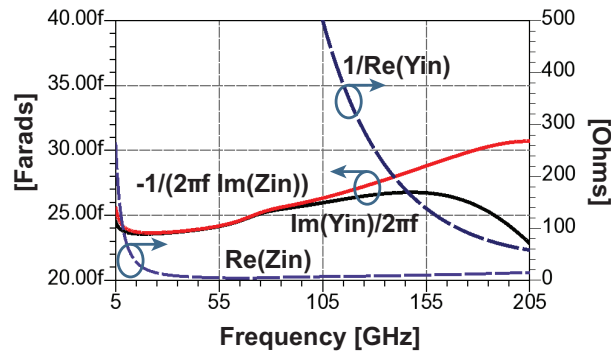


Figure 5.37: Input loading from the DA gain element.

complicate biasing due to need for several bias points as well as AC connection between “drain” lines of previous stages with the “gate” line.

Here, we are not interested in providing DC or close to DC gain in the amplifier. The final system will cover bands that will go down to a few GHz or so. We therefore design a cascaded RF choke using a spiral inductor in the core. The design uses a short high impedance line followed by a 4 turn square spiral ($1.5 \mu\text{m}$ spacing between turns). The structure is connected to the bias point by a longer trace that provides the low end inductance. The architecture is shown in Fig. 5.38. The first segment of the high-Z line together with part of the choke provide adequate inductance for the highest frequencies. The multi-turn choke provides the inductance at mid-frequencies. It is designed to have a high self-resonance frequency. Capacitor C_1 is a local and relatively small capacitor (400 fF) that ensures that the high frequency response is decoupled from the second high-Z line and whatever comes after that. Capacitor C_2 is a large capacitor that isolates the cascaded-choke structure from biasing circuits. For bringing the response down to lower frequencies, we could replace the second high-Z line (between capacitors) with another larger multi-turn inductor. In this design, the current structure provides the required response with only a single multi-turn inductor.

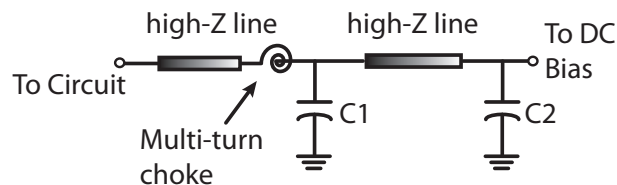


Figure 5.38: Cascaded choke architecture.

Amplifier Characteristics and Measurement Results

A three-stage cascaded DA is realized in $0.13 \mu\text{m}$ SiGe BiCMOS process. Each stage uses 5 gain elements. The cacode stages are biased at 5 mA each and hence the total current in the active part of the DA is 75 mA.

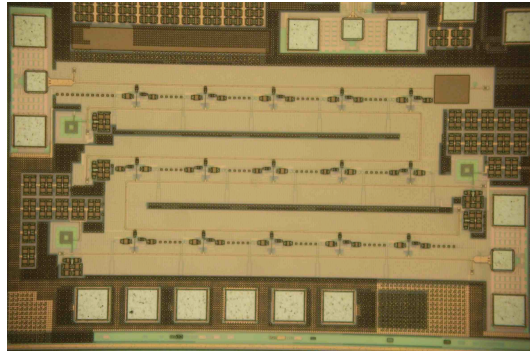


Figure 5.39: Chip micrograph of the SiGe distributed amplifier.

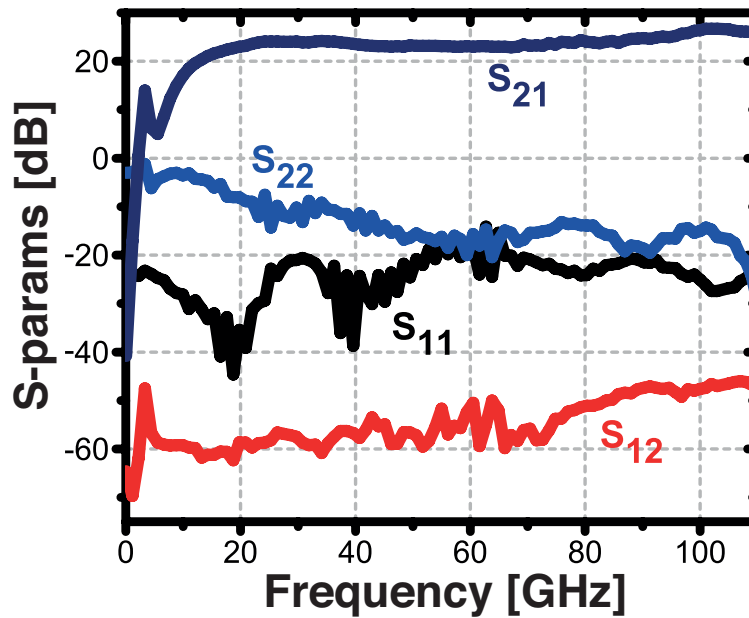


Figure 5.40: S-parameter measurements of the SiGe DA.

The chip micrograph is shown in Fig. 5.39. S-parameter measurements are shown in Fig. 5.40. The amplifier provides an average gain of 24dB and larger than 110GHz of bandwidth (measurements limited by the VNA). Measurements include all pads/parasitics

and no de-embedding is done on the DA. The DA has a measured GBW product in excess of 1.5THz. Simulations show that a BW of 125 GHz is expected. Further measurements beyond the W-band will assess operation to 125 GHz (leading to a GBW product in excess of 1.7 THz).

5.3 Wideband Conversion to Differential

Single-ended to differential conversion over a very wide frequency bandwidth is very challenging. This is especially true if the phase response (or group-delay characteristic) of the block is important as is the situation here. A passive balun structure could be designed to provide 20-30% relative bandwidth but that usually comes at the price of poor phase response close to the frequency edges. In addition to that, we target an imager that uses a wide frequency range with multiple-carriers. The LNA will be a common block that is re-used between the stages. This necessitates a broadband conversion.

Here, the conversion to differential is performed through an active balun. The whole structure is very similar to a “micro-mixer” proposed by Gilbert in [97]. The design uses common-base topology on one path and a combination of a diode and common-collector on the other path to generate opposite polarities of the signal.

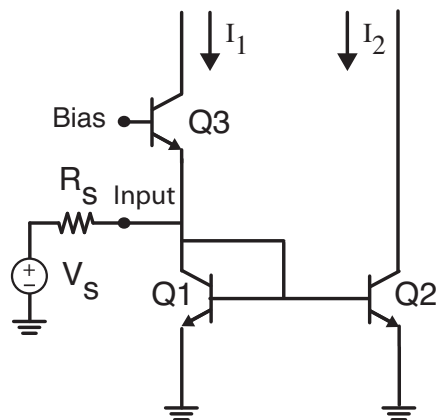


Figure 5.41: Active balun circuit schematic.

We will first analyze the active balun circuit independently. Fig. 5.41 shows the basics of the circuit. At low frequencies and with the assumption of a high current gain we have:

$$I_1 = I_{in} Z_{in} g_m = I_2 \quad (5.39)$$

in which $Z_{in} \simeq \frac{1}{2g_m}$. With a finite device current gain, input impedance will be slightly smaller. Nevertheless, the interesting point of this architecture is that irrespective of the input impedance (which will change slightly with a non-ideal current gain), the output

currents of the two branches will be enforced by the voltage of the common input node modulating the base-emitter of Q2 and Q3. There is some resiliency to variations of these differential currents with the addition of non-idealities. However, even with this topology, once the frequency of operation becomes comparable to f_t (as is the case here with $f_{operation,max} \simeq 0.6f_t$), non-idealities will show themselves. In addition to that, adding series resistance for the base and emitter junctions of the devices will add to the mismatch. There will also be a slight mismatch between the AC and DC copy ratios. Because of the DC current mismatch between branches (base current error), the g_m values will be different and that will lead to a differential error on the output currents. To mitigate this problem and reduce errors, a resistor is added in series with the base of the diode connected device (Q1). This can reduce the DC current mismatch to the first order. In the AC response, adding the resistor will add a pole and zero to the impedance of this node. This could be used to equalize the response of the LNA-mixer interface section. The pole is located at $\frac{\beta C_\pi}{2\pi g_m(\beta+1)} \simeq f_t$ and the zero at $\frac{1}{2\pi C_{pi}(R_B/r_\pi)}$. This zero is common between the paths and will not affect the differential response error of the active balun.

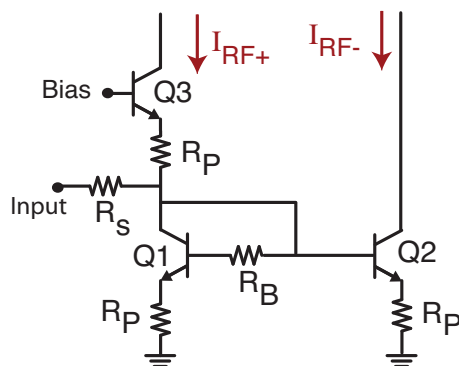


Figure 5.42: Circuit schematic of the active balun with degeneration and base resistances.

In addition to this, a degeneration resistor is also included in the circuit to linearize the input impedance. The tradeoff is in the noise penalty associated with this. Under large swings the input impedance of the stage gets modulated. Adding a fixed resistor will help by reducing the variation at the expense of higher noise from these elements. In general, wideband “non-resonant” input matching comes with the cost of higher noise in the first place but this will also add to that. Fig. 5.42 shows the schematic of the circuit with the additional resistances.

Fig. 5.43 shows the phase of the current components I_1 and I_2 together with the phase error. The error is less than 6° up to 130 GHz. The other issue is with the magnitude of the currents. The magnitude error before correction is shown in Fig. 5.44. This shows the magnitude of the currents as well as the error percentage. This error exceeds 10% for frequencies close to 110 GHz. The resulting CMRR is shown in Fig. 5.45. This needs to be improved and is partially corrected through another leverage point with the control voltage

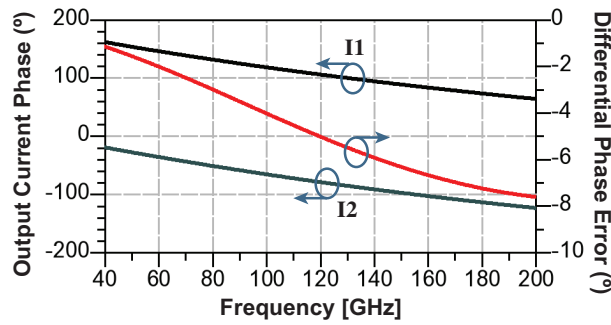


Figure 5.43: Phase response of the active balun. Both single-ended phase and phase error on differential output is shown.

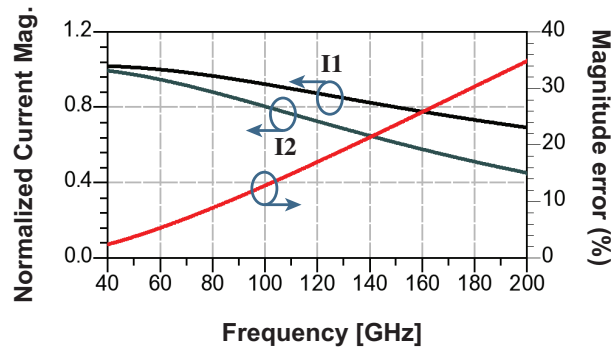


Figure 5.44: Initial magnitude response and error in magnitude of the active balun circuit. Current buffer's control voltage (Fig. 5.46) can be used to further reduce the error.

of the current buffer that comes after the active-balun (and by that the collector-emitter voltages of transistor Q2).

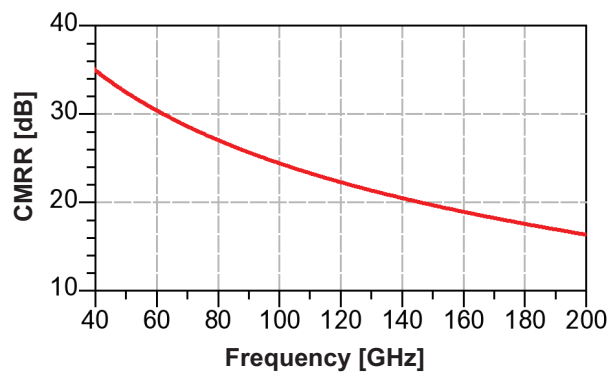


Figure 5.45: CMRR of the core active balun circuit.

5.4 Mixer Design

A fully-balanced active mixer is used in the direct-conversion receiver. The schematic of the mixer with the active balun is shown in Fig. 5.46.

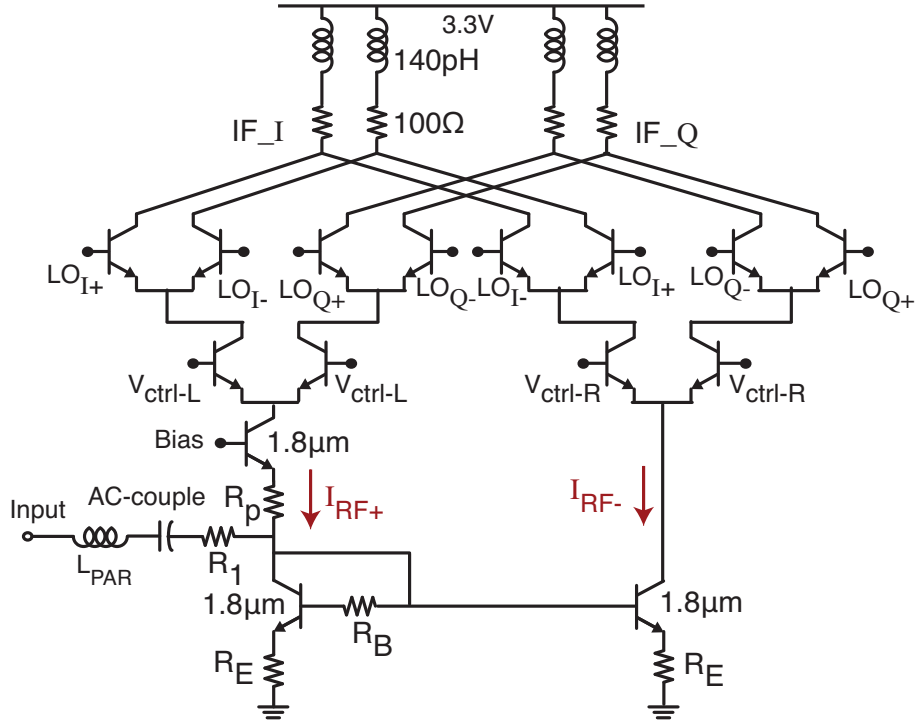


Figure 5.46: Overall schematic of the active balun, current buffer and quadrature mixer.

5.4.1 Overview of Challenges

We will start by a fully-balanced current commuting mixer core (a balanced Gilbert cell). First order analysis of the dominant distortion terms in the mixer are provided. First, we will ignore the effect of offset voltage in the mixer quad core and assume perfect LO swing. I_1 and I_2 are the currents feeding into the quad core (RF and DC). We will define CM and DIFF currents as follows:

$$\begin{aligned} I_{diff} &= I_1 - I_2 \\ I_{cm} &= \frac{I_1 + I_2}{2} \end{aligned} \quad (5.40)$$

We will then have:

$$\begin{aligned}
I_{out_1} &= I_1 \times P_1(t) + I_2 \times P_2(t) = \\
&I_1 \times \left(0.5 + \frac{2}{\pi} \cos(\omega_{LO}t)\right) + I_2 \times \left(0.5 - \frac{2}{\pi} \cos(\omega_{LO}t)\right) = \\
&\quad I_{cm} + \frac{2}{\pi} \cos(\omega_{LO}t) I_{diff} \\
I_{out_2} &= I_2 \times P_1(t) + I_1 \times P_2(t) = \\
&I_2 \times \left(0.5 + \frac{2}{\pi} \cos(\omega_{LO}t)\right) + I_1 \times \left(0.5 - \frac{2}{\pi} \cos(\omega_{LO}t)\right) = \\
&\quad I_{cm} - \frac{2}{\pi} \cos(\omega_{LO}t) I_{diff}
\end{aligned} \tag{5.41}$$

where we have only taken the first harmonics into account. The common-mode RF current goes through without frequency conversion as a common-mode in baseband. Assuming the baseband circuits provide filtering and some common-mode rejection, this component will be non-significant.

Next is the differential component. To first order we could describe this component as:

$$I_{diff} = I_{RF-diff} + \Delta I_{dc} + \Delta I_{IM2} \tag{5.42}$$

where ΔI_{dc} is the difference in DC currents of the two paths and ΔI_{IM2} is the residual difference of the IM2 components generated at the input RF stage. This results from any imbalances or mismatch in the gm-stage of the mixer. Notice that these components will be up/down-converted by the LO term in equation (5.41).

The DC current mismatch term will generate an LO feedthrough to output (even though we have a fully-balanced topology). This large LO signal at output has the potential to overwhelm the baseband gain stages. As an example, with 2 mA of DC current and a 5% mismatch between the paths, a 100 μ A of “LO” current will find its way to the output. For a 100 Ω load, the LO swing will be about 10 mV which is quite substantial.

Two partial solutions are designed into the mixer. First, to provide first order cancellation of the DC term, the control voltage at the base of the current buffers shown in Fig. 5.46 is designed to be programmable. This voltage allows for changing the collector-emitter voltage of the bottom devices and by that allows a small change in the DC current of the paths. In the real implementation, a 4 bit DAC is used to control these voltages. Also, an override option is provided to allow external control by a finer step.

Secondly, an LO trap can be placed on the output of the mixer. This is in the form of a narrow microstrip line that minimizes the loading on the baseband signal itself. Given that the cutoff frequency of the baseband signal (30 GHz) is relatively close to the 94 GHz LO, this is not an easy task.

In addition to the mentioned error terms, the mismatch between the mixer core devices also generates additional error components. Without the mismatch between these devices,

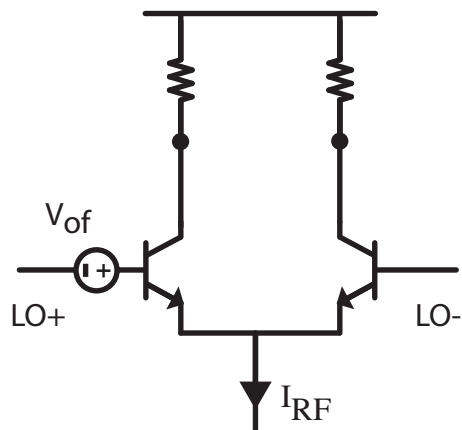


Figure 5.47: Mixer core with input offset voltage.

IM2 components, for example, do not fall in the baseband since the common-mode part of IM2 will remain in common-mode and the differential mode is up-converted as described above. However, the mismatch component changes this [98] [99]. The offset voltage (shown in Fig. 5.47) will generate duty cycle distortion. This is due to the movement of the switching threshold on one of the devices. To reduce this effect, a large swing or a pulse input has to be applied which are both are very costly with a 94 GHz LO. The feedthrough is to first order related to the ratio of the offset voltage to the peak swing. There is another mechanism through which distortion is generated and that is by periodic charging of the capacitance at the common emitter node [98]. This effect is reduced by the addition of the current buffers and hence the capacitance of the common node in this topology.

5.4.2 Mixer Core Design

The mixer core has separate I and Q fully-balanced cells that take the input from the current buffers. The left and right current buffers reduce the loading on the lower devices and by that reduce bandwidth limitations. They also divide the current in two paths providing the input for both the I and Q balanced Gilbert cells. At the same time, the voltage at the base could be adjusted to both change the DC currents of the two paths (to calibrate for LO feedthrough) as well as to maximize differential component (by slight change in gain of the two polarities). The current buffer transistors also reduce the LO to RF feedthrough by adding an intermediate layer. This feedthrough component will manifest itself as an undesired transmitted tone retransmitted from the receiver. Time-varying effects can modulate this tone and turn into an undesirable echo in the system.

A series 300 fF MIM capacitor AC couples the LNA and mixer. It is sized to both provide adequate matching at the bands of interest as well as not to introduce a large parasitic capacitance to ground. The mixer core uses device size of 1.4 μm . The current

dividers are 1 μm devices. Bypass capacitor banks are placed at the base of the current mirror to reduce chance of oscillation and noise pick up. A combination of MOS capacitors (400 fF) and MIM capacitors (130 fF) are placed in close vicinity of each of these control voltages. Each branch of the active balun carries a nominal current of 2 mA.

As shown in the schematic, the mixer uses shunt-peaking in the form of high-Z meandered microstrip line. Using shunt-peaking is a delicate compromise between the amplitude response and the phase/group-delay response. Excessive use of shunt peaking will result in large changes in the group delay of the amplifier and is avoided. If we define m as the ratio of the inductive corner frequency to the capacitive corner frequency or $\frac{R^2C}{L}$, it is shown that $m \simeq 1.41$ leads to the best bandwidth extension [77]. On the other hand, $m \simeq 3.1$ will result in the best group delay response. Here, $m \simeq 2.2$ is chosen as a tradeoff which leads to a bandwidth extension ratio of 1.7. This leads to a line inductance that is close to 140 pH in series with a 100 Ω resistive load.

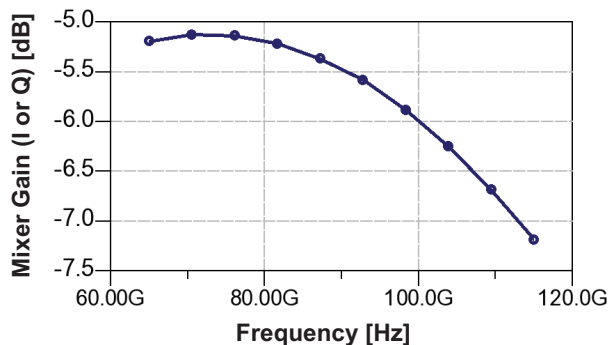


Figure 5.48: Mixer gain referred to RF frequency (voltage gain with 100 Ω load or 200 Ω differentially).

The simulation results for the extracted version of the mixer using Hicuml2 models is shown in Fig. 5.48. The gain varies between -5 and -7 dB across the 65 GHz to 115 GHz band. Using other model cards, we obtain a smaller gain variation. The Hicuml2 has been seen to show slightly pessimistic results (compared to measurements) in the past results. The standard model provides a larger gain by 1-1.5 dB.

5.4.3 LO Buffer

As discussed in previous chapters, frequency locking remains an important issue in obtaining the required resolution. To achieve this using the central integrated PLL, a distribution network is required. This necessitates LO dividers to distribute the PLL signal to the TX and RX side as well as the divider in the loop. Instead of using Wilkinson dividers for this purpose, this design incorporates transformer/amplifier combination for signal division.

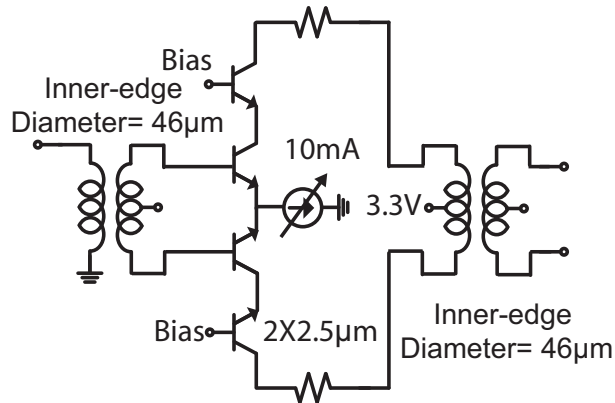


Figure 5.49: The schematic of the single-ended to differential LO buffer.

A combination of a balun, differential cascode amplifier, and a transformer is designed to obtain the required conversion¹. The schematic is shown in Fig. 5.49.

One of the key challenges with this design is the common-mode response of the balun. To reduce the common-mode levels, a capacitor (in series with some resistance to de-Q the network) is placed on the center tap of the secondary (balanced) side. The right value of capacitance (and hence the impedance of that node) reduces the common-mode response of the structure significantly.

To ensure the stability of the buffer as well as to provide an adequate bandwidth on the output match, loss (in form of $12\ \Omega$ resistance) is placed in series with the collectors of the devices.

The single-ended input is matched to $50\ \Omega$ and the differential output is matched to $100\ \Omega$. The buffer gain is 10dB, OP_{-1dB} is 2dBm, and P_{sat} is 5.4dBm at 94GHz. Simulation of the output power level of the buffer is shown in Fig. 5.50.

5.5 Baseband Amplification

The quadrature signals are delivered to the baseband amplifiers. The budget on the input capacitance of the buffer is extremely tight. As mentioned previously, excessive capacitance will lead to bandwidth limitation and group delay distortion. At the same time, the baseband amplifier needs to drive internal/external $50\ \Omega$ lines as well as to provide some gain. A 3dB bandwidth of 25GHz or more is required on each of the quadrature signals in order to meet range resolution limitations from group delay distortion.

The input stage is designed as a differential follower to decrease input capacitive loading

¹Thanks to Stefano Dal-Toso for assistance in the initial design and to Shinwon Kang for modifications.

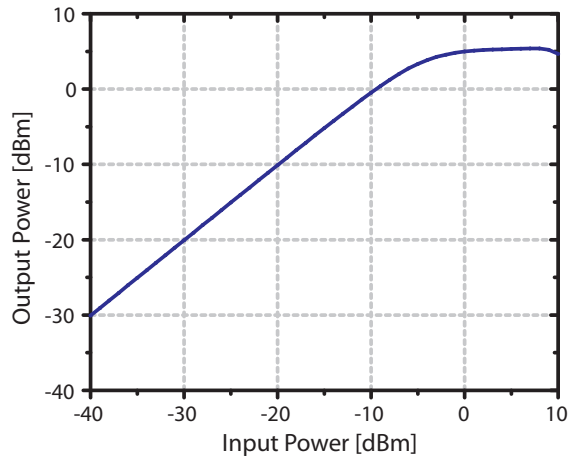


Figure 5.50: LO buffer output power.

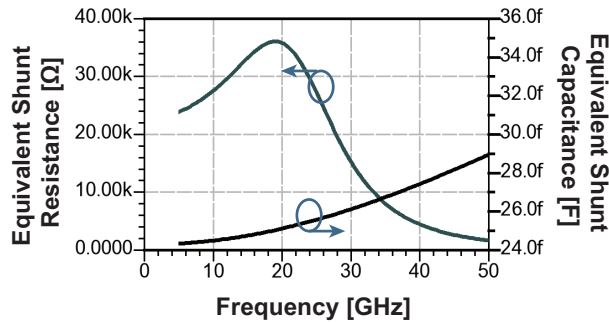


Figure 5.51: Input impedance at the input of the mixer.

to the mixer. This is important since the mixer output is a critical node with a relatively larger capacitance and therefore we avoid extracting gain from the first buffer stage that loads this node. In addition to that, LO feedthrough (by mechanisms described earlier) exists on this node and a buffer stage has a larger swing tolerance and its frequency response can suppress the undesirable 94 GHz signal. On the other hand, stability becomes a major concern with the follower stage and needs to be balanced against bandwidth considerations. The first stage is biased at 6 mA (device biased at 1 mA/ μm). The equivalent input shunt resistance ($\frac{1}{\text{Re}(Y_{in})}$) and shunt capacitance ($\frac{\text{Im}(Y_{in})}{\omega}$) are shown in Fig. 5.51.

The second stage is designed for gain. It uses larger devices ($2 \times 2.5\mu\text{m}$) in a differential cascode topology. Some emitter degeneration is used to reduce input capacitive loading as well for linearity concerns. This stage has a nominal bias current of 14 mA which places the transistor bias at 1.4 mA/ μm .

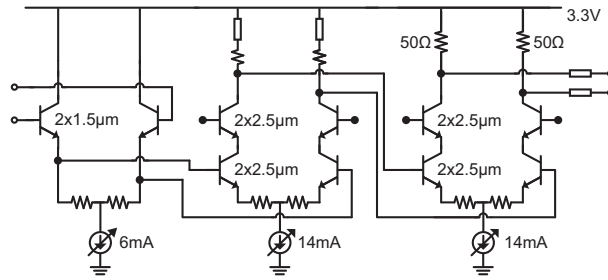


Figure 5.52: The schematic of the baseband buffer.

A third stage is used to drive the $50\ \Omega$ line. Another differential cascode structure is used. The device bias and sizing is similar to that of stage 2. Shunt-peaking is not utilized. Fig. 5.52 shows the schematic of the three stage buffer.

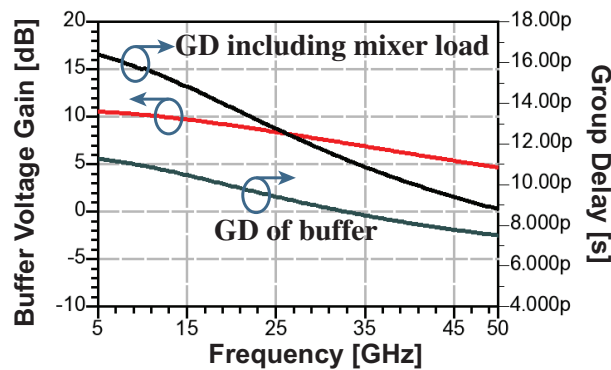


Figure 5.53: Baseband buffer voltage gain and group delay simulations.

Fig. 5.53 shows the simulation results of the baseband buffer. This includes EM simulation (Ansoft HFSS) of all traces and interconnects. As can be seen here, the results show that the design has not been very aggressive on using extensive shunt-peaking. In other words, the total balance is tilted towards a better group delay response as compared to gain flatness. It is also observable that the node at the mixer load consumes a large portion of the total bandwidth budget. The tradeoff there is between gain, group delay/BW, noise, and linearity.

Chapter 6

Integrated 94GHz Radar Transceiver

6.1 System Overview

In the next step towards the integration of a large-scale array imager, a single-chip mm-wave radar transceiver has been designed in a $0.13\mu\text{m}$ SiGe BiCMOS process. The system schematic is shown in Fig. 6.1. The annotated chip micrograph is shown in Fig. 6.2. As discussed in chapter 2, this is a pulsed-based radar with the ability to move the transmit (or receive) window using a high-resolution delay-locked loop. The center frequency is nominally at 94GHz and is programmable from 87-97GHz. This adds to the resolution capabilities of the system as described previously.

The chip integrates antennas for the transmitter and the receiver, LNA, wideband active balun, quadrature mixers and IF gain stages, a 94 GHz PLL, variable width pulse generators, 1.47 GHz DLL, PA, and quadrature LO distribution. An external 2.94 GHz reference is divided to provide the nominal PRF for the pulser and a reference for the DLL and PLL. Therefore, the carrier frequency is locked to the PRF and integration produces coherent pulses. To realize beamforming, the pulse envelope can be shifted in fine time increments through embedded interpolation of DLL phases. For perfect phase combination, this will also require a mm-wave phase shifter that would shift the phase at 94 GHz. This part is not implemented in this chip.

6.1.1 Receiver Architecture

The receiver section uses the blocks and components described in chapters 5 and 3. The antenna element is a tapered loop antenna with a back-side reflector. The two TX and RX antennas are placed on the two ends of the chip to reduce mutual coupling. Also, the back-side ground reflector of the two antennas could be separated to provide more isolation. The loop diameter is $650\mu\text{m}$ and the width is tapered from $20\mu\text{m}$ on the input side to $70\mu\text{m}$ on the outer side.

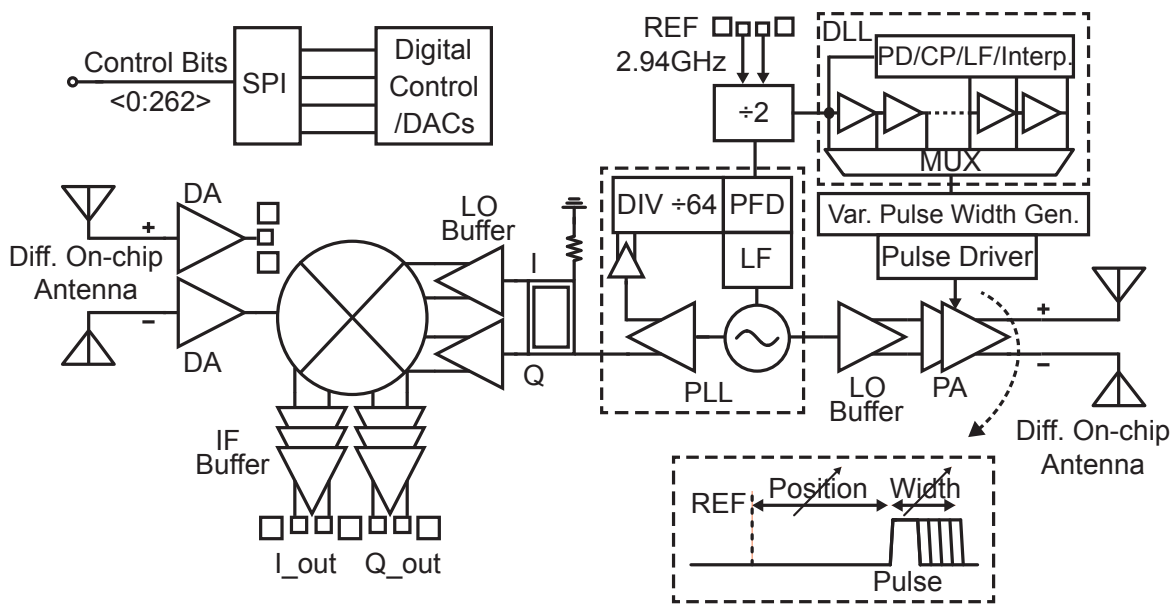


Figure 6.1: Block-level schematics of a pulsed-based radar transceiver.

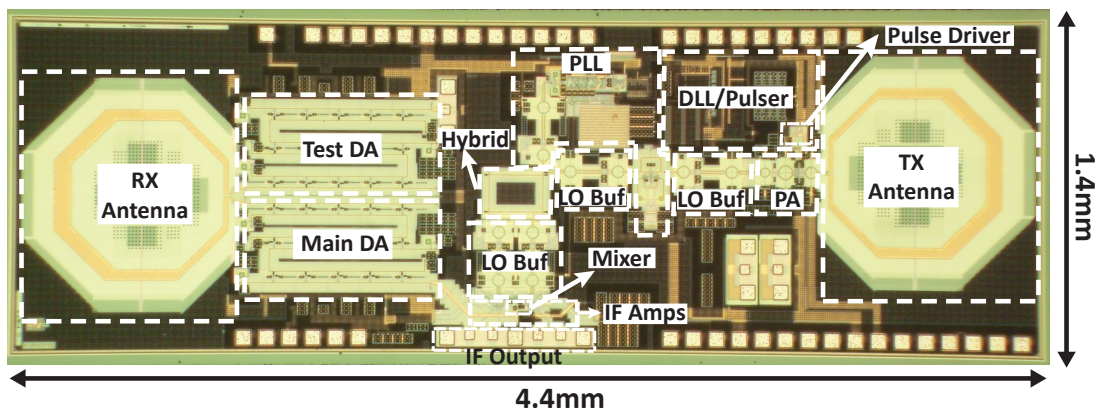


Figure 6.2: Chip micrograph of the radar transceiver.

In this chip, two identical distributed amplifier (DA) blocks are integrated. The differential antenna feeds each of the single-ended amplifiers symmetrically. The top DA feeds a separate RF pad structure for testing and debugging of the antenna, the amplifier, and the interface. The bottom DA feeds the single-ended to differential active balun and mixer. The interface is a 50Ω transmission line.

In the LNA (cascaded 3-stage DA), the biasing of the first stage is separated from the two other stages. The first stage bias DAC provides a range between $800\ \mu\text{A}$ to $6.8\ \text{mA}$ by

4-bits. This feeds 5 cascode gain elements in the first DA stage. A current multiplication of 1 to 6 exists for each of the cascodes. Nominal bias current for the cascode elements is close to 5mA. The second and third DA stages share a common current source which is programmable by two 4-bit DACs. By slightly changing (lowering) the current of the first stage, we can reduce the noise figure of the LNA.

The active balun circuit topology is described in chapter 5. A current buffer separates the active balun from the quadrature mixer core circuit. The output of the mixer is inductively loaded for shunt-peaking and drives the IF amplifiers. The I and Q outputs are placed on the left and right of the mixer and each drive a separate gain path in the IF. Two single-ended to differential LO buffers are used to provide the required signal swing to the mixer LO port. The layout is completely balanced with dummy transmission lines placed to provide complete symmetry to the mixer and IF stages. The mixer nominal current is set to 2 mA for each of the branches and is programmable by a 4-bit DAC (0.4 mA- 3.4 mA per branch).

The IF gain stages are design to provide 11dB of gain across a 26GHz BW. The final stage drives the 50 Ω pads and off-chip components (RF probes in this case). The current consumption of the three stages of the IF amplifier is 6 mA, 14 mA, and 14 mA all from a 3.3 V supply voltage. The current DACs provide a range of 1.2 mA-10.2 mA for the first stage and 2.8 mA-23.8 mA for the second and third stages.

The output of the IF gain stages drive the RF differential Ground-Signal-Signal-Ground (GSSG) pads. The I/Q pad structures share a ground pad. The layout of the mixer and baseband stages is shown in Fig. 6.3. In this figure the I/Q LO signals arrive from the top and are routed to the mixer core. Meandered microstrip lines that provide shunt-peaking are placed on the top of the mixer and move to the left and right for the I and Q components. The IF stages also use shunt-peaking and that is also implemented as meandered lines. Symmetrical routing provides the I/Q signals to the RF pads. Simultaneous measurements of the I/Q components will require custom-built RF probes.

6.1.2 Transmitter Architecture

In the transmit path of this chip, an independent switching scheme is employed. Future designs can incorporate hybrid-switching for shorter pulse generation (demonstrated in Chapter 4). Here, the switching functionality is solely realized in the power amplifier stage. The power amplifier has been redesigned from the previous design introduced in Chapter 4 to allow improved ON/OFF ratio (reduced leakage) as well as faster transition. The block diagram of the TX section is shown in Fig. 6.1.

The VCO has a differential output. One of the outputs feeds the RX as well as the divider and the PLL loop. The other output is used to drive the TX side. An identical LO buffer to the one on the RX path is used to convert the signal to differential. The PA receives a differential CW signal in the TX band of 87-97GHz. PA schematic is shown in Fig. 6.4. The first stage of the PA is a cascode implementation and provides the required gain and power for the second stage. Switching takes place in the second stage. Current switching or

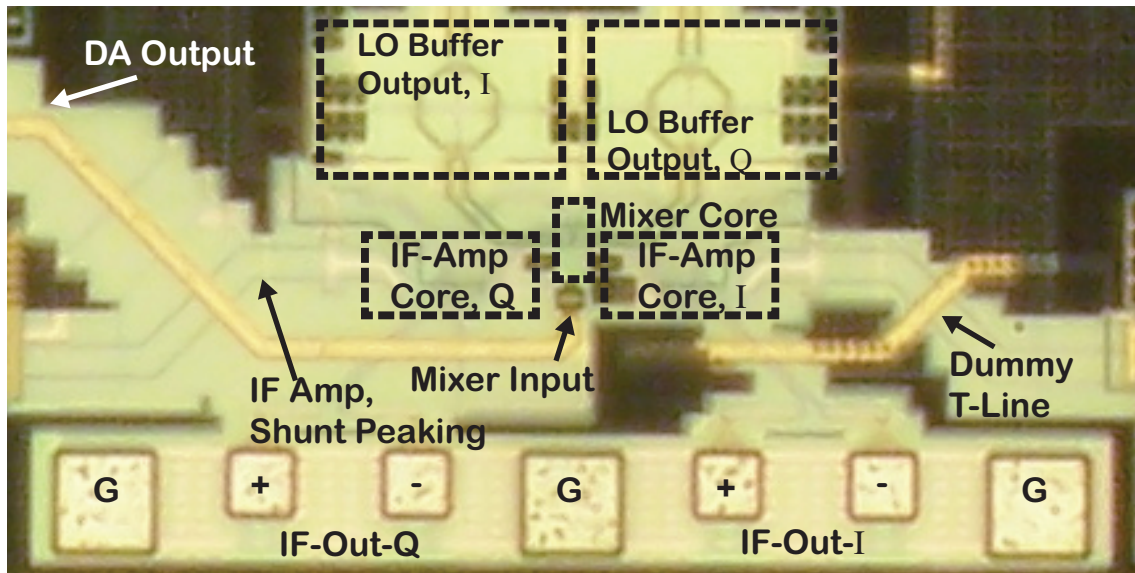


Figure 6.3: Layout of the baseband section.

steering is realized by a Gilbert structure. Contrary to the implementation described in the standalone transmitters of Chapter 4, in this design the pulse input waveform is completely differential. This could potentially increase feedthrough of the pulse (base to collector) from the signal side of the Gilbert quad to the output. To circumvent this problem, transistors in the signal side are stacked to increase isolation.

The PA driver provides the differential pulse signal to the PA switching transistors. The schematic is shown in Fig. 6.5. A differential buffer with a programmable bias is used to generate sharp edges for switching. Series resistor loss is placed on the path to increase stability. The PA uses large devices in the second stage and hence the capacitance in the common-emitter node of the differential Gilbert stage is quite large (in excess of 100 fF). In the switching event, the PA current commutation stage acts similar to a follower loaded with a large capacitance. Stability will be a major concern for this stage. To stabilize the stage we need to increase the real part of the impedance seen from the base. The series losses provide the extra part.

6.1.3 Frequency Generation and Distribution

The 94GHz LO signal is generated by the PLL and distributed into both TX and RX. The block diagram of the PLL is shown in Fig. 6.6. The PLL uses similar loop elements and characteristics to a previous design with Shinwon Kang reported in [100]. There are some modifications: here, the VCO uses a coarse two-bit digital control to set the band and hence a locking range of 10GHz is obtained. The loop uses the LO buffer introduced in Chapter

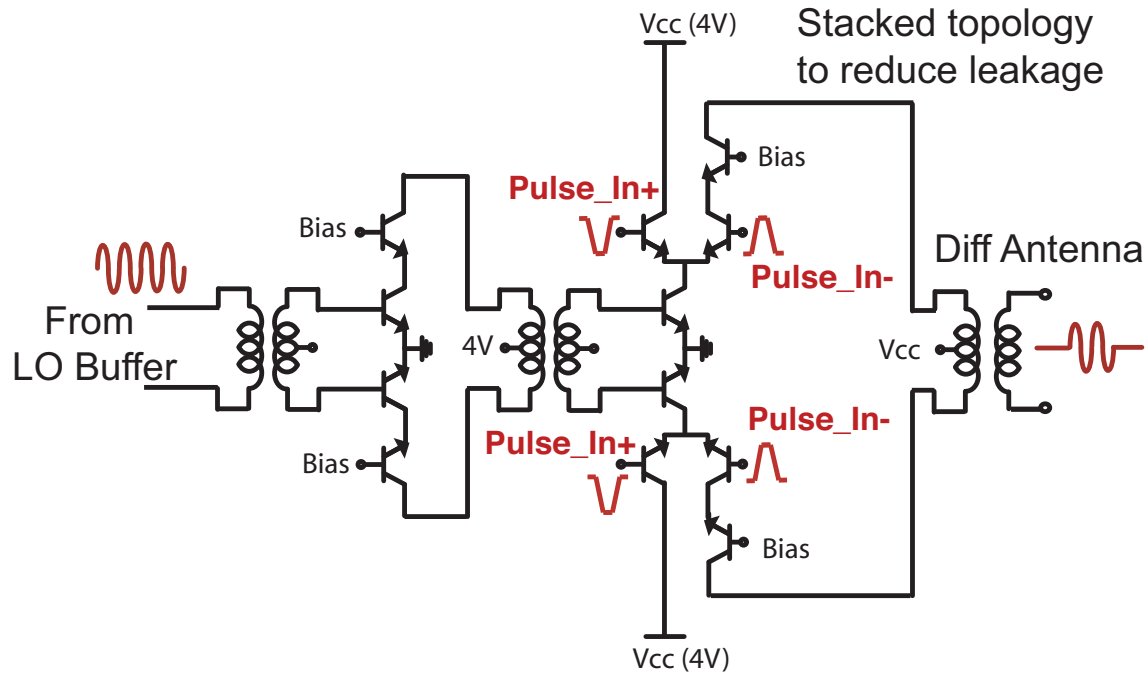


Figure 6.4: Two-stage transformer-coupled power amplifier.

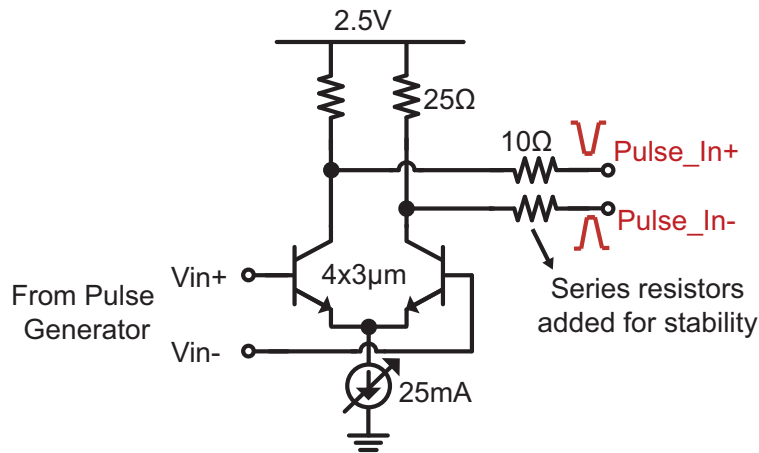


Figure 6.5: PA pulse driver schematics.

5 to feed the divider. A fundamental-mode differential cascode Colpitts VCO is chosen for the PLL. If designed properly, a Colpitts oscillator can provide better performance in terms of oscillation frequency and noise compared to a cross-coupled design (Chapter 4).

One single-ended VCO output is amplified by the LO buffer and feeds the PA. The other VCO output is converted to differential signals by the LO buffer to drive the PLL

divider and the RX. Quadrature signals are required for the receiver. Initial transmitter designs incorporated a quadrature VCO for this purpose (Chapter 4). However, here we use a passive phase-shift approach. This is to ensure low phase noise and locking across the large PLL locking range. A quadrature hybrid, made of four microstrip lines and four MIM capacitors, is used. These quadrature signals are then connected to single-to-differential LO buffers which drive the mixer with a power of +3dBm. The VCO frequency can be varied by changing two coarse bits and by adjusting the VCO bias current.

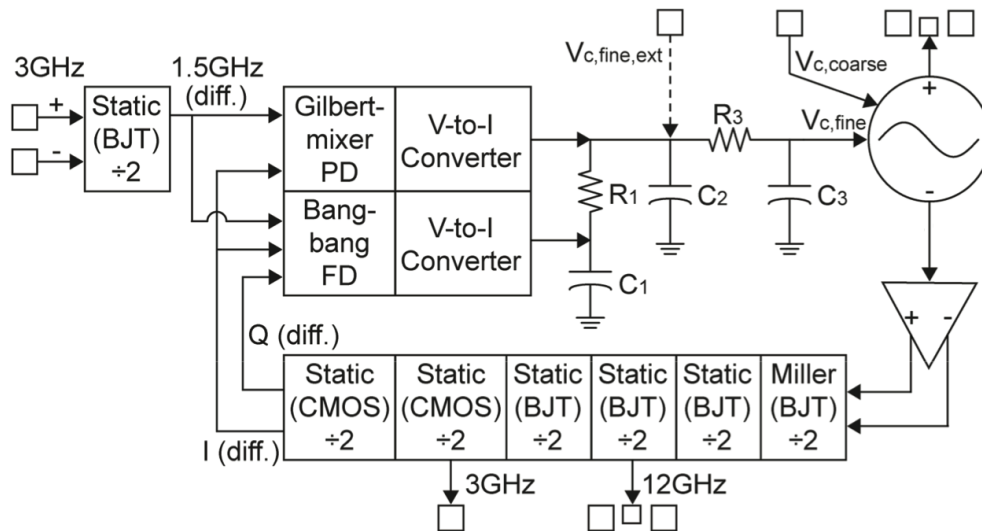


Figure 6.6: PLL architecture implemented in TUSI transceiver.

6.1.4 Pulse Position and Width Programming

In order to be able to perform pulse beamforming, an accurate control of the pulse position is required. The pulse width controls the tradeoff between depth and resolution with narrower pulses providing better resolution at the cost of lower penetration depth. In order to generate a programmable pulse position, a delay-locked loop is incorporated in the transceiver. All of the transceiver modules in the array receive a synchronized common clock and would then impose the correct delay for beamforming.

The DLL here uses ideas from previous work of Toifl et al. [101] and is similar in architecture to the design introduced by Steven Callender et al. in [102]. The basic idea behind the design is similar to that introduced in Chapter 2. If the elements in the delay chain are reduced from N to $N-1$ cells, the delay per cell changes from $\frac{T}{N}$ to $\frac{T}{N-1}$. We can implement the change in the number of delay stages by choosing a different delay point in the chain in the feedback loop of the DLL. In addition to that, phase-interpolation could

also take place by giving different weights to the phase-detector outputs of various delay points in the chain. The schematic of the idea is shown in the conceptual drawing of Fig. 6.7. Here, a 9 stage delay chain is implemented with the last three taps used in the interpolation (interpolation between $9/8$, $8/7$ and $9/7$). Each output tap will see an adjustment in the delay depending on the interpolation steps. Outputs that are further down the chain will experience a larger change due to the accumulation. To provide coarse timing, a MUX element selects which of the taps are routed to the very output.

As described in [102], the phase adjustment introduced here results in non-linear delay steps. Assuming we have N stages in the path, the delay from the m -th tap is $m\frac{T}{N}$. This is not a linear function from N 's viewpoint. We therefore need to design for finer steps to meet the maximum delay error requirements. A calibration procedure may be required for the chip to derive the mapping table for the obtainable delay steps.

The DLL receives the divided down 1.47GHz reference signal as its input. This reference signal is the common synchronous clock for the array and also feeds the PLL. Therefore, the PLL and DLL are locked together making coherent pulse transmission possible. With this DLL topology, the resolution of the phase steps is dictated by the resolution of the current DACs used to perform the interpolation. A 5-bit thermometer coded DAC is used, resulting in a theoretical worst-case phase step of less than 3ps.

The output of the DLL block feeds a variable pulse generator. The pulse generator is similar to that of Chapter 4. Together with the DLL, the baseband blocks can produce variable pulse width and position for the array.

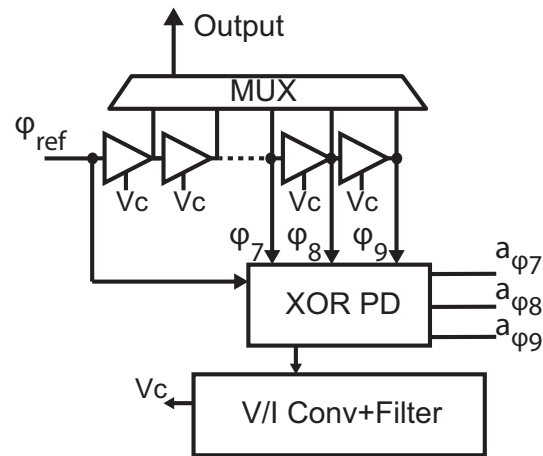


Figure 6.7: Conceptual schematics of delay-locked loop (DLL).

6.2 End-to-End Measurement Results

The chip is fabricated in the $0.13 \mu\text{m}$ SiGe BiCMOS process described in previous sections. The die thickness is set to the standard $375 \mu\text{m}$.

The chip is bonded to a PCB board using a chip-on-board (COB) assembly method. The antenna ground reflector is implemented as soft-gold ground plane underneath the chip. Here, a single continuous plane is used to cover the entire chip from the back. To reduce direct coupling between the TX and RX, the grounds could be separated (cut in the middle). The chip has two high-frequency interfaces: the input references clock which is at 2.94GHz (nominal) and the IF output which is probed. The SPI control is programmed through a scan-chain with an external FPGA board. In total, the chip has 262 bits of control to set bias currents, select TX and RX settings, and to set the pulse width, delay and polarity. A separate power board is designed to provide supply voltages and current references. The RF board is shown in Fig. 6.8.

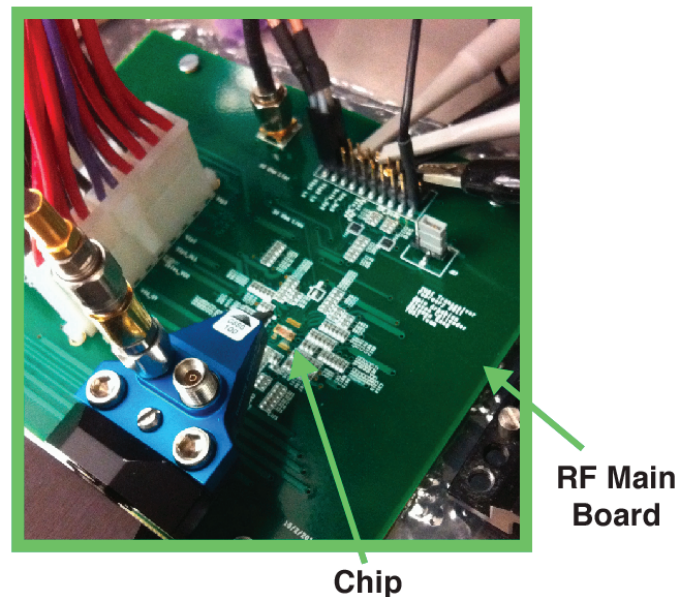


Figure 6.8: Picture of the RF board in the measurement setup.

To perform transceiver testing a sampling oscilloscope (Agilent 86100C with 70 GHz head) was used as the primary testing module. The clock frequency (3 GHz) was provided by the Agilent 8267D. A different signal source (Agilent 4438C) triggers the oscilloscope. The two sources are locked together using the 10 MHz references. This is not an ideally stable locking scheme and has some negative impact on the system measurements. The measurement setup is shown in Fig. 6.9.

The locking range of the PLL is measured by two methods. In the first method, the

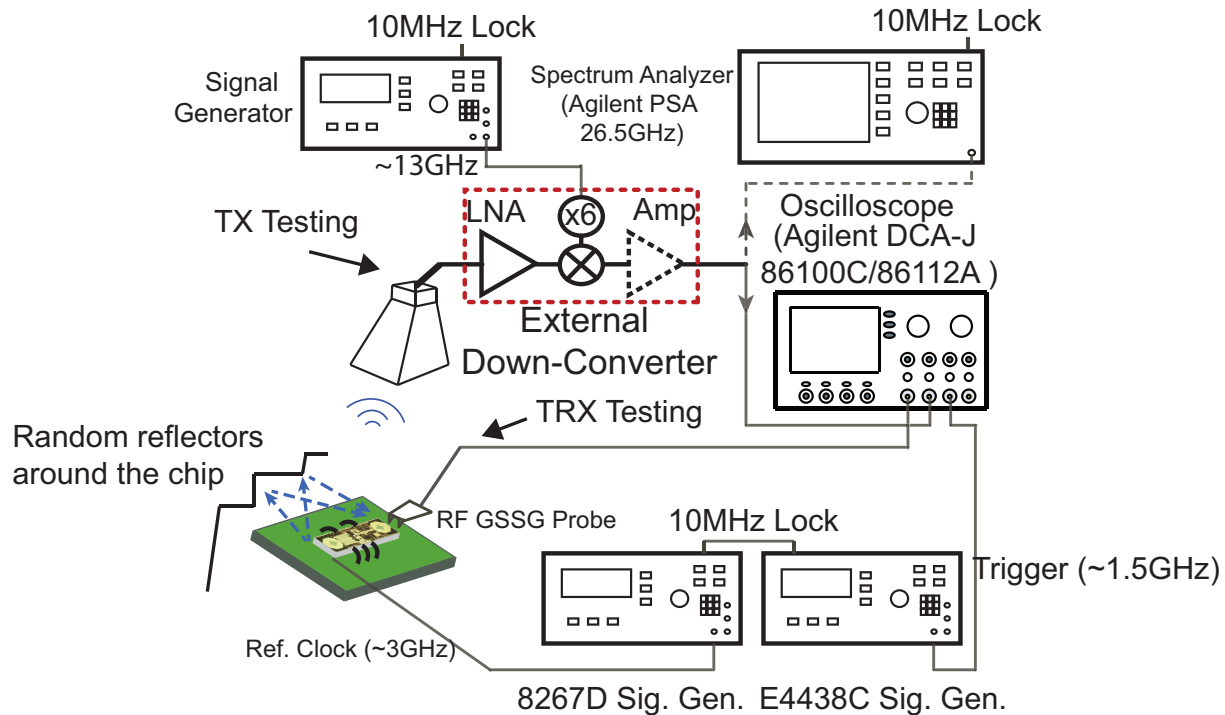


Figure 6.9: Measurement setup for testing the radar transceiver. Both TRX and separate TX measurement setups are shown. The TX is measured by an external down-converter.

transmitted signal is intercepted directly using an external down-converter. Both pulse and continuous-wave outputs could be used to measure the center frequency of the PLL. When the frequency falls outside the PLL locking range, several changes are observable. First, in the frequency domain, the tones show a pulling behavior. This is shown in Fig. 6.10. Here, the transceiver output (end-to-end measurements from the whole system including the TX and RX) is measured in the pulse mode. Zoom-in version of the spectrum is illustrated around 4.42 GHz (PRF is set to 1.47 GHz). The spectrum is shown for both locked and non-locked versions.

In addition to that, depending on the pulling frequency offset, there will be some distortion. The spectral lines of non-locked case in Fig. 6.10 are separated by approximately 20 MHz. This separation is not a fixed factor and will change by changing VCO range, reference frequency and etc. Depending on the relative values of this frequency spacing, pulse width, observation time (total integration time), and time-of-flight, this distortion can lead to a detrimental effect on both range and phase accuracy of the system. Longer integration windows can partially mitigate this issue with averaging out the beat-frequency effect.

The other way by which a non-locked system is observable is through time-domain measurements (both direct TX and also TRX). If the frequency is not stable, received echoes will not arrive in a perfectly coherent manner and some instability is observed in the

time-domain. Fig. 6.11 shows an example of the case where the PLL is not locked.

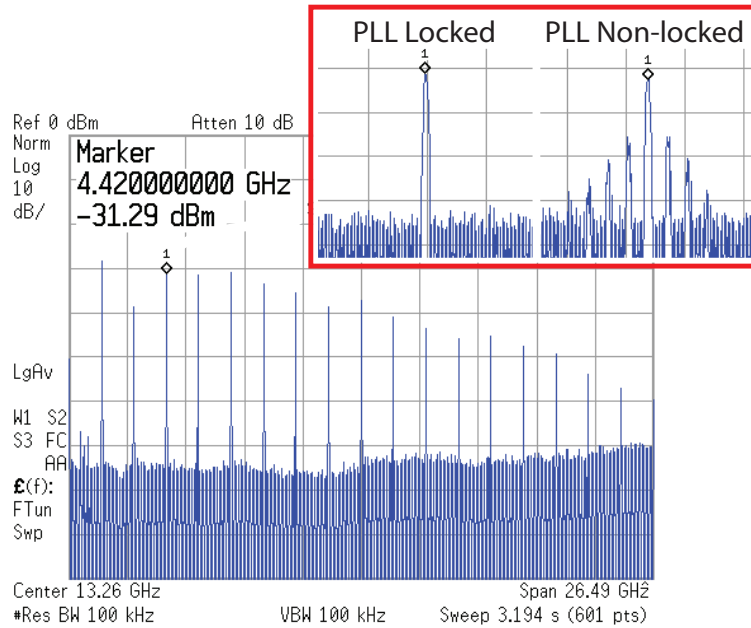


Figure 6.10: Frequency spectrum measurements of the transceiver together with zoomed in versions for locked and non-locked PLL conditions.

Measurements of the lock-range of the PLL are shown in Fig. 6.12. The VCO frequency band is selected by 2-bit digital control string which introduces extra fixed varactors to the tank. To adjust the frequency of the VCO and allow increase in locking, VCO bias current can also be slightly modified from nominal. The PLL locks from 87.3 GHz to 97.2 GHz.

Once the PLL is locked and the carrier frequency is a multiple of the PRF, coherent integration of pulses is made possible. To observe the output, infinite persistence mode is utilized in the sampling oscilloscope. In addition to persistence, averaging can also be performed by the oscilloscope to reduce amplitude noise. Fig. 6.13 shows measurements of the received pulses in TRX settings. Both averaged and non-averaged versions are shown with a PRF being set to 1.47 GHz. The measured pulse width can go down to around 36 ps (50 % to 50 %) at which point some ringing starts to occur.

Next, the DLL measurements are performed. As before, these are end-to-end measurements of the whole system in which the DLL control settings are adjusted to observe the received waveform timing position. As previously mentioned, the DLL provides a non-linear delay step because of the tuning mechanism used. The DLL provides both fine and coarse delay tuning. Coarse steps yield programmability over an entire pulse repetition interval (PRI). First the whole range of DLL tuning is measured. This is obtained by sequentially looking through the MUX outputs (from various delay stages) and performing interpolation

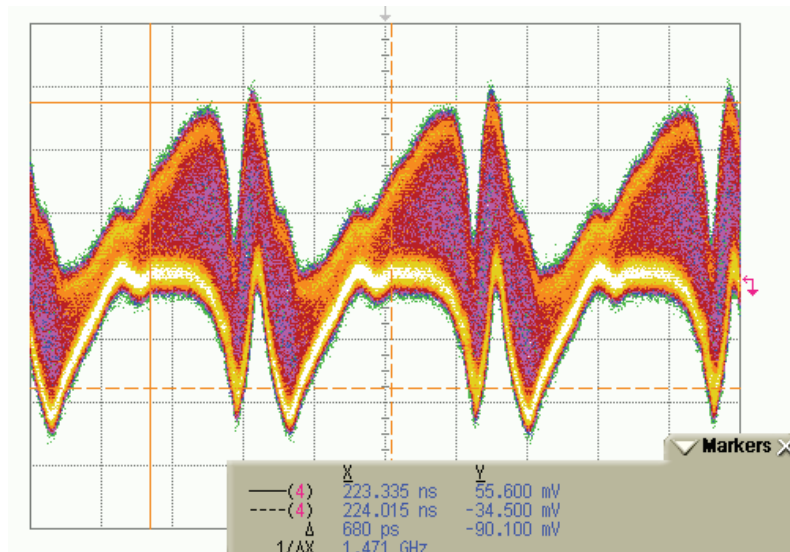


Figure 6.11: Time-domain measurements with a non-locked PLL. The distortion is due to the addition of non-coherent pulses. Due to the high PRF used in this chip, some level of coherency is still observable.

on the last three taps. By going through all steps, we can derive the total delay variation possible on the chip. First, we selected 8 different points that provide uniform coarse delay over the whole PRI. This is shown in Fig. 6.14. Next, we looked at specific points on the tuning curve to observe the fine tuning capability. This is best observed in the last three tap outputs (stages 7, 8 and 9). Looking at these three stages from the MUX, we can tune the delay on a 238 ps span with an average step of 2.28 ps. Max step is 6.8ps and corresponds to the jump from one stage output to the next. Redundant and duplicate points are removed from these measurements. Results of fine step tuning is shown in Fig. 6.15.

Separate measurements were also performed on the transmitter. An external down-converter using a horn antenna is used to intercept the transmitted pulses. The down-converter is similarly setup to measurements of Chapter 4. To observe coherent pulses, one needs to lock the PRF, center frequency, and also the down-converter frequency. It is also mandatory that these frequencies be multiples of the PRF since we would like to observe pulses that are superimposed on the top of the previous ones. An example of frequency selections could be setting PRF to 1.5 GHz, center frequency to 96 GHz, and down-converter LO to 90GHz. This way, the down-converted signal has a center frequency that is a multiple of the PRF ($6 \text{ GHz} = 4 \times 1.5 \text{ GHz}$). However, since the pulse BW is greater than 6GHz, aliasing is expected. These measurements are shown in Fig. 6.16. Since frequencies are locked, we can use averaging to reduce noise levels on the pulse. The observable output signal jitter is larger here compared to TRX measurements. This is because, the down-converter as shown in Fig. 6.9 uses a multiply by 6 on the LO side and hence for a 90

GHz LO, a 15GHz signal has to be fed to the system. This 15 GHz signal also needs to be frequency locked to the reference signal provided to the chip (3 GHz here) as well as the oscilloscope trigger (1.5 GHz). Locking is maintained by the 10 MHz references from the two sources. However, at such high frequencies, the 10 MHz reference does not provide adequate phase/frequency stability and hence the observed signal quality drops.

Table 6.1 shows the summary of transceiver performance.

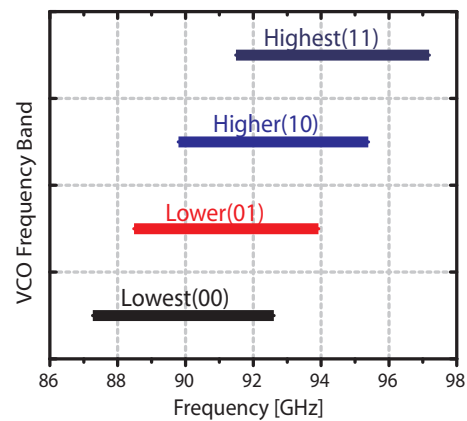


Figure 6.12: PLL locking range based on various VCO frequency bands. Fixed capacitors are switched in and out using a two-bit control.

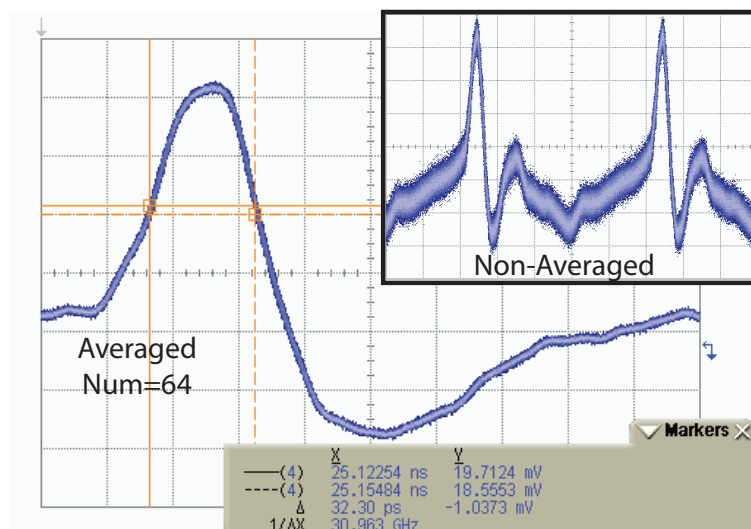


Figure 6.13: Measurements of TRX pulse output waveform both with and without averaging. The pulse goes through the entire TX and RX chains.

Table 6.1: Summary of radar transceiver performance.

94GHz Pulsed Radar Transceiver Performance Summary		
Technology		0.13 μm SiGe BiCMOS
Area		4.4mm X 1.4mm
System Specifications	PRF (nominal)	1.47 GHz (94 GHz/64)
	PRF Range	1.364 -1.519 GHz
	Min. Pulse Width	36ps (50%-50%)
	Total TRX RMS Jitter	1.2ps
	BB I/Q BW	26 GHz
	Total DC Power (including biasing)	Continuous (1.9W*), Duty-Cycle TX at 20% (1.4W)
LNA (Distributed Amplifier)	BW (-3dB)	15-110 GHz (**), (up to 125 GHz in sim)
	Gain	24 dB
	DC Power	75mA (3.3V)
I-Q Mixer/ Broadband Single to Diff.	Amp/ Phase error (Single-ended to diff)	70-110GHz: 1.2dB/ 5-7° (sim)
	Conversion loss	5-7dB (70-120GHz) (sim)
	DC Power	4mA (3.3V)
Antenna	Type	Tapered loop antenna with metal reflector
	Peak Broadside Gain	-0.5dB (sim)
PLL/ LO distribution	Lock Range (in System)	87.3- 97.2GHz
	PLL Phase Noise (@95GHz) (measured in standalone)	-92.5dBc/Hz@100KHz, -102dBc/Hz@1MHz
	CLK Spur	<-60dBc
	DC Power	152mA (3.3V), 23mA (2.5V), 17mA (1.2V)
DLL/ Pulser/ Pulse Driver	Max coarse delay	Full PRI (680ps)
	Max delay span for fine steps	238ps (avg. step=2.28ps, max step=6.8ps)
	DC Power	94mA (2.5V)
PA	Gain, P _{-1dB} , P _{sat}	16dB, +8dBm, +13dBm
	DC Power	80mA (4V)

PLL and DLL measured performance are from end-to-end system measurements.

* Pulse measurements performed with no duty cycling of DC power

** Measurements limited by VNA frequency range

*** All DC currents of major blocks are programmable by combination of 4-bit DACs

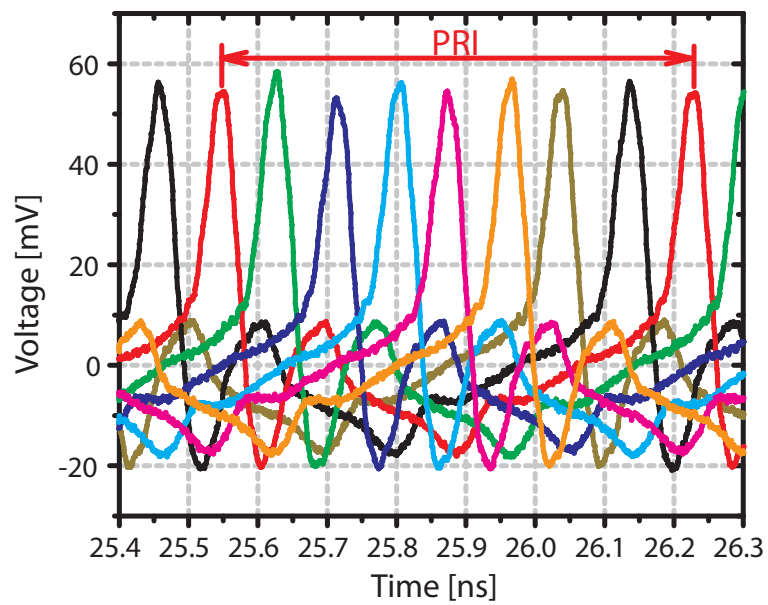


Figure 6.14: Pulse delay profile in coarse setting. The pulse position is adjustable across the whole PRI. All measurements are end-to-end with the entire transmit and receive chains included.

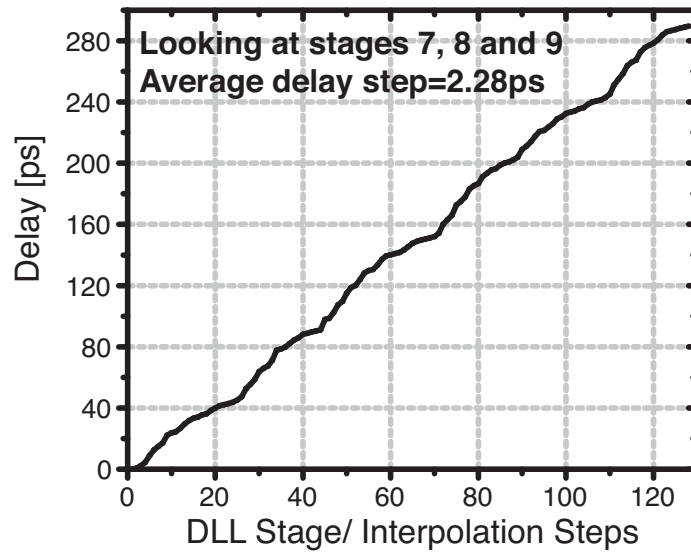


Figure 6.15: Measured pulse delay (referenced to initial setting) for outputs of stages 7,8 and 9. Fine tuning is obtained by interpolation between stages 7, 8 and 9 in the loop of the DLL.

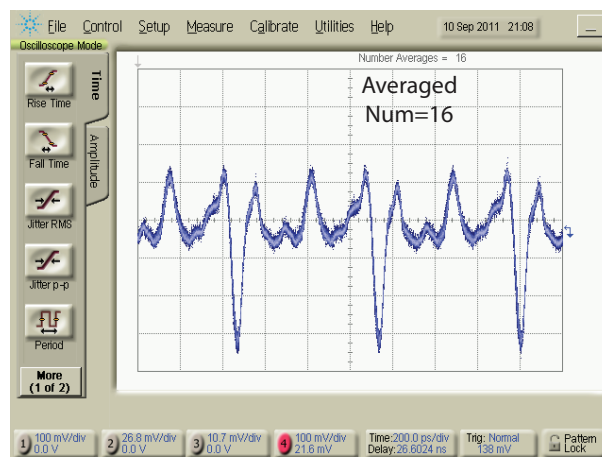


Figure 6.16: Intercepted pulse from the transmitter. An external down-converter is used to observe this pulse.

Chapter 7

Conclusion

In this work, we have presented the system and circuit level design of an integrated mm-wave pulsed-based imager in silicon technology. The presented work concluded with a highly integrated transceiver covering 70-110 GHz with sub-50ps pulses as well as coherent carrier frequency tuning (from PLL). Besides diagnostic imaging, other applications for this system include PPM modulated high speed data communication, short-range (board to board) wireless link with greater than 30 Gbps data rate, non-medical imaging (e.g. security applications), and intelligent surfaces with 3D imaging capabilities.

In chapter 2 we addressed design challenges related to a pulsed-based array imager. Details of radar signal processing and arrays basics can be found in the literature ([103, 37, 13, 24]). Practical limitations on obtaining mm-level preprocessed resolution were addressed. The importance of synchronization and clock accuracy was also examined.

In chapter 3 we discussed challenges and opportunities regarding integrated antennas. Efficiency enhancement as well as various switching techniques were addressed. Antentronics was introduced as a way of manipulating the current and field distribution of the antenna to synthesize the desired impulse response. Two antentronic structures were introduced. Loop antenna on a grounded substrate was proposed as a potential candidate for high efficiency on-chip antennas.

In chapter 4 we investigated challenges, techniques and methodologies related to generation and control of short mm-wave pulses. Hybrid switching was introduced for shorter pulse generation. Two different transmitter architectures together with measurement results were presented. Design tradeoff and specifics related to the transmitter chain including the high-speed pulser, VCO, and the PA were examined. In the measurements, preprocessed resolution capabilities of the imager were assessed and measured. Multiple-target detection capability was shown.

In chapter 5 we described the design of the receiver sections. This included wideband amplification, active balun, wideband mixers, baseband stages, and LO distribution. Fundamental limitations in attaining a large gain-bandwidth product were examined and several wideband techniques proposed. Distributed amplifiers with internal feedback, tapering in

line segments, and gain stage optimization were introduced. A multi-stage biasing choke was proposed for operation beyond 120 GHz. The SiGe amplifier provided a GBW product better than 1.5 THz.

In chapter 6 we introduced an integrated pixel-scalable transceiver in a SiGe BiCMOS process. The transceiver achieves a very high level of integration compared to the state-of-the-art 94GHz silicon systems and includes the transmit/receiver chain as well as a 87-97 GHz PLL and a DLL for pulse positioning/beamforming. System-level measurement results were presented. The TRX is able to provide coarse mm-wave pulse positioning in the entire PRI as well as fine pulse positioning in a smaller range with 2.28 ps of average time step. This shows beam-steering functionality for the array.

Future directions include integration of multiple transceivers and assessment of array resolution. In addition, power reduction in various blocks is another key component of future designs. Larger frequency span transceivers will need to be addressed to broaden the application span for TUSI. Assessment of imaging capabilities *In Vitro* as well as with animals are among future steps for this modality.

On the signal processing side, there is a need to co-develop signal conditioning techniques with the hardware architecture for the imager. Optimal processing distribution between the single element transceivers and the central processor is important for power reduction as well as improved signal acquisition. Various super-resolution algorithms can be applied to the designed hardware. Given the availability of center-frequency tuning as well as pulse position/width selection, both time and frequency domain algorithms can be used.

In the area of antenna design and antentronics there are many interesting architectures and applications that could be pursued. As an example, the TUSI folded slot antentronic structure described in Chapter 3 can be modified to provide beam-steering capability. We can program the timing elements such that the pulse signals to the CMOS switches arrive in different times. With that, the spatial response of the transmitted pulse (or the “pulse pattern”) can be altered dynamically. This could also be done on multiple antentronic elements in an array (or an antentronic array). Future integrated antennas in the mm-wave to THz frequencies are envisioned to be designed differently than what they are now. Electronic elements will play a larger role in dynamically changing the antenna characteristics and in synthesizing the desired impulse response.

Bibliography

- [1] E. C. Fear, S. C. Hagness, P. M. Meaney, M. Okoniewski, and M. A. Stuchly, “Enhancing Breast Tumor Detection with Near-field Imaging,” *IEEE Microwave Magazine*, vol. 3, pp. 48–56, Mar. 2002.
- [2] A. Babakhani, X. Guan, A. Komijani, A. Natarajan, and A. Hajimiri, “A 77-GHz Phased-Array Transceiver With On-Chip Antennas in Silicon: Receiver and Antennas,” *IEEE Journal of Solid-State Circuits*, vol. 41, no. 12, pp. 2795–2806, 2006.
- [3] C. Marcu *et al.*, “A 90 nm CMOS Low-Power 60 GHz Transceiver With Integrated Baseband Circuitry,” *IEEE Journal of Solid-State Circuits*, vol. 44, no. 12, pp. 3434–3447, 2009.
- [4] M. Tabesh, J. Chen, C. Marcu, L. Kong, S. Kang, A. Niknejad, and E. Alon, “A 65 nm CMOS 4-Element Sub-34 mW/Element 60 GHz Phased-Array Transceiver,” *Solid-State Circuits, IEEE Journal of*, vol. 46, no. 12, pp. 3018–3032, 2011.
- [5] B. Floyd, S. Reynolds, U. Pfeiffer, T. Zwick, T. Beukema, and B. Gaucher, “SiGe bipolar transceiver circuits operating at 60 GHz,” *IEEE Journal of Solid-State Circuits*, vol. 40, no. 1, pp. 156–167, 2005.
- [6] B. Razavi, “CMOS transceivers for the 60-GHz band,” *IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, 2006.
- [7] T. Mitomo, N. Ono, H. Hoshino, Y. Yoshihara, O. Watanabe, and I. Seto, “A 77 GHz 90 nm CMOS Transceiver for FMCW Radar Applications,” *IEEE Journal of Solid-State Circuits*, vol. 45, no. 4, pp. 928–937, 2010.
- [8] C. H. Doan, S. Emami, A. Niknejad, and R. W. Brodersen, “Millimeter-wave CMOS design,” *IEEE Journal of Solid-State Circuits*, vol. 40, pp. 144–155, Jan. 2005.
- [9] K. R. Foster and E. A. Cheever, “Microwave radiometry in biomedicine: A reappraisal,” *Bioelectromagnetics*, vol. 13, no. 6, pp. 567–579, 1992.

- [10] S. Chaudhary, R. Mishra, and A. Swarup, "Dielectric properties of normal and malignant human breast tissues at radiowave and microwave frequencies," *Indian journal of Biochem. Biophys.*, 1984.
- [11] A. Arbabian, S. Callender, S. Kang, B. Afshar, J.-C. Chien, and A. Niknejad, "A 90 GHz Hybrid Switching Pulsed-Transmitter for Medical Imaging," *IEEE Journal of Solid-State Circuits*, vol. 45, pp. 2667–2681, Dec. 2010.
- [12] S. Davis, B. Van Veen, S. Hagness, and F. Kelcz, "Breast Tumor Characterization Based on Ultrawideband Microwave Backscatter," *IEEE Transactions on Biomedical Engineering*, vol. 55, pp. 237–246, Jan. 2008.
- [13] M. Skolnik, *Introduction to Radar Systems*. McGraw-Hill Science/Engineering/Math, 3 ed., Dec. 2002.
- [14] M. Richards, *Fundamentals of Radar Signal Processing*. McGraw-Hill, 1 ed., June 2005.
- [15] A. Arbabian and A. Niknejad, "Medical Imaging in Microwave Frequencies." BWRC Poster presentation 2007.
- [16] A. Arbabian and A. Niknejad, "Time-Domain Ultra-Wideband Synthetic Imager." BWRC Poster presentation 2009.
- [17] H. K. Weir, M. J. Thun, B. F. Hankey, L. A. G. Ries, H. L. Howe, P. A. Wingo, A. Jemal, E. Ward, R. N. Anderson, and B. K. Edwards, "Annual Report to the Nation on the Status of Cancer, 1975-2000, Featuring the Uses of Surveillance Data for Cancer Prevention and Control," *Journal of the National Cancer Institute*, vol. 95, no. 17, pp. 1276–1299, 2003.
- [18] S. J. Nass, I. C. Henderson, and Institute of Medicine (U.S.). Committee on Technologies for the Early Detection of Breast Cancer, *Mammography and Beyond*. developing technologies for the early detection of breast cancer, Natl. Academy Pr., 2001.
- [19] A. Surowiec, S. Stuchly, J. Barr, and A. Swarup, "Dielectric properties of breast carcinoma and the surrounding tissues," *IEEE Transactions on Biomedical Engineering*, vol. 35, no. 4, pp. 257–263, 1988.
- [20] B. Lindelöf and M. Hedblad, "Accuracy in the clinical diagnosis and pattern of malignant melanoma at a dermatological clinic," *The Journal of dermatology*, 1994.
- [21] A. W. Kopf, M. Mintzis, and R. S. Bart, "Diagnostic Accuracy in Malignant Melanoma," *Archives of Dermatology*, vol. 111, p. 1291, Oct. 1975.

- [22] A. Lightstone, A. Kopf, and L. Garfinkel, "Diagnostic Accuracy— A New Approach to Its Evaluation: Results in Basal Cell Epitheliomas," *Archives of Dermatology*, vol. 91, p. 497, May 1965.
- [23] F. Rampen and P. Rumke, "Referral Pattern and Accuracy of Clinical-Diagnosis of Cutaneous Melanoma," *Acta Dermato-Venereologica*, vol. 68, no. 1, pp. 61–64, 1988.
- [24] M. I. Skolnik, *Radar Handbook*. McGraw-Hill Professional, Nov. 2007.
- [25] L. Borcea, G. Papanicolaou, and C. Tsogka, "Theory and applications of time reversal and interferometric imaging," *Inverse Problems*, vol. 19, no. 6, pp. S139–S164, 2003.
- [26] A. Devaney, "A filtered backpropagation algorithm for diffraction tomography," *Ultrasonic Imaging*, vol. 4, pp. 336–350, Oct. 1982.
- [27] C. Gabriel, S. Gabriel, and E. Corthout, "The dielectric properties of biological tissues: I. Literature survey," *Physics in Medicine and Biology*, vol. 41, pp. 2231–2249, Jan. 1999.
- [28] S. Gabriel, R. Lau, and C. Gabriel, "The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz," *Physics in Medicine and Biology*, vol. 41, p. 2251, 1996.
- [29] S. Gabriel, R. Lau, and C. Gabriel, "The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues," *Physics in Medicine and Biology*, vol. 41, p. 2271, 1996.
- [30] C. Gabriel, "Compilation of the Dielectric Properties of Body Tissues at RF and Microwave Frequencies.," *AFOSR/NL Report*.
- [31] F. Homburger and W. H. Fishman, *The Physiopathology of Cancer*. Hoeber Harper.
- [32] A. Campbell and D. Land, "Dielectric properties of female human breast tissue measured in vitro at 3.2 GHz," *Physics in Medicine and Biology*, vol. 37, p. 193, 1992.
- [33] R. C. González and R. E. Woods, *Digital Image Processing*. Addison-Wesley, 2002.
- [34] C. A. Balanis, *Antenna Theory. analysis and design*, Wiley-Interscience, 2005.
- [35] M. Hussain, "Principles of space-time array processing for ultrawide-band impulse radar and radio communications," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 3, pp. 393–403, 2002.
- [36] M. Hussain, "Ultra-Wideband Impulse Radar-An Overview of the Principles," *IEEE Aerospace and Electronic Systems Magazine*, vol. 13, no. 9, pp. 9–14, 1998.

- [37] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Radar-Sonar Signal Processing and Gaussian Signals in Noise*. Wiley-Interscience, May 2003.
- [38] T. Miki, H. Yamaguchi, and Y. Nagaki, "An Accurate Wide-Band Automatic Waveform Analyzer," *IEEE Transactions on Instrumentation and Measurement*, vol. 26, pp. 279–291, Dec. 1977.
- [39] A. Karalis, J. Joannopoulos, and M. Soljacic, "Efficient wireless non-radiative mid-range energy transfer," *Annals of Physics*, vol. 323, pp. 34–48, Jan. 2008.
- [40] Barakat, "On the design of 60 GHz integrated antennas on 0.13 um SOI technology," in *2007 IEEE International SOI Conference*, pp. 117–118, 2007.
- [41] Y. Zhang, "Antenna-on-Chip and Antenna-in-Package Solutions to Highly Integrated Millimeter-Wave Devices for Wireless Communications," *IEEE Transactions on Antennas and Propagation*, 2009.
- [42] N. Alexopoulos, P. Katehi, and D. Rutledge, "Substrate Optimization for Integrated Circuit Antennas," *IEEE Transactions on Microwave Theory and Techniques*, vol. 31, no. 7, pp. 550–557, 1983.
- [43] D. Pozar, "Considerations for millimeter wave printed antennas," *IEEE Transactions on Antennas and Propagation*, 1983.
- [44] M. Abdel-Aziz, H. Ghali, H. Ragaie, H. Haddara, E. Larique, B. Guillon, and P. Pons, "Design, implementation and measurement of 26.6 GHz patch antenna using MEMS technology," *IEEE Antennas and Propagation Society International Symposium*, vol. 1, pp. 399–402, 2003.
- [45] Rutledge, D., "Imaging antenna arrays," *IEEE Transactions on Antennas and Propagation*, vol. 30, no. 4, pp. 535–540, 1982.
- [46] M. Nezhad Ahamdi and S. Safavi-Naeini, "On-chip antennas for 24, 60, and 77GHz single package transceivers on low resistivity silicon substrate," *IEEE Antennas and Propagation Society International Symposium*, pp. 5059–5062, 2007.
- [47] D. M. Pozar, *Microwave Engineering*. Wiley, 2005.
- [48] D. B. Rutledge, D. P. Neikirk, and D. P. Kasilingam, "Integrated Circuit Antennas," *Infrared and Millimeter Waves*, vol. 10, pp. 1–90, 1983.
- [49] R. E. Collin, *Field Theory of Guided Waves*. Oxford University Press, USA, Sept. 1996.

- [50] S. Wentworth, R. Rogers, J. Heston, D. P. Neikirk, and T. Itoh, "Millimeter wave twin slot antennas on layered substrates," *International journal of infrared and millimeter waves*, 1990.
- [51] D. Thiel, *Switched Parasitic Antennas for Cellular Communications*, Artech House. Inc., 2001.
- [52] A. Babakhani, et al., "Transmitter Architectures Based on Near-Field Direct Antenna Modulation," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 12, pp. 2674–2692, 2008.
- [53] L. Pringle, P. Harms, S. Blalock, G. Kiesel, E. Kuster, P. Friederich, R. Prado, J. Morris, and G. Smith, "A reconfigurable aperture antenna based on switched links between electrically small metallic patches," *IEEE Transactions on Antennas and Propagation*, vol. 52, pp. 1434 – 1445, june 2004.
- [54] V. Fusco and Q. Chen, "Direct-signal modulation using a silicon microstrip patch antenna," *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 6, pp. 1025–1028, 1999.
- [55] J.-C. Ke, C.-W. Ling, and S.-J. Chung, "Implementation of a multi-beam switched parasitic antenna for wireless applications," *IEEE Antennas and Propagation Society International Symposium*, pp. 3368–3371, 2007.
- [56] C. A. Balanis, *Antenna Theory: Analysis and Design, 3rd Edition*. Wiley-Interscience, 3 ed., Apr. 2005.
- [57] P. Garcia, A. Chantre, S. Pruvost, P. Chevalier, S. Nicolson, D. Roy, S. Voinigescu, and C. Garnier, "Will BiCMOS stay competitive for mmW applications ?," in *IEEE Custom Integrated Circuits Conference*, pp. 387–394, 2008.
- [58] P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*. Wiley, 5 ed., Jan. 2009.
- [59] S. Montusclat, F. Gianesello, D. Gloria, and S. Tedjini, "Silicon integrated antenna developments up to 80 GHz for millimeter wave wireless links," in *The European Conference on Wireless Technology.*, pp. 237–240, 2005.
- [60] A. Arbabian and A. Niknejad, "A broadband distributed amplifier with internal feedback providing 660GHz GBW in 90nm CMOS," *IEEE International Solid-State Circuits Conference*, 2008.
- [61] J.-C. Chien and L.-H. Lu, "40-Gb/s High-Gain Distributed Amplifiers With Cascaded Gain Stages in 0.18- μ m CMOS," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 12, pp. 2715–2725, 2007.

- [62] A. Arbabian and A. M. Niknejad, "Design of a CMOS Tapered Cascaded Multistage Distributed Amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 4, pp. 938–947, 2009.
- [63] J. Chen and A. M. Niknejad, "Design and Analysis of a Stage-Scaled Distributed Power Amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 5, pp. 1274–1283, 2011.
- [64] M. Alioto and G. Palumbo, *Model and Design of Bipolar and MOS Current-Mode Logic*. CML, ECL and SCL digital circuits, Kluwer Academic Pub, Dec. 2005.
- [65] J. McNeill, "Jitter in ring oscillators," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 6, pp. 870–879, 1997.
- [66] N. Pohl, H.-M. Rein, T. Musch, K. Aufinger, and J. Hausner, "SiGe Bipolar VCO With Ultra-Wide Tuning Range at 80 GHz Center Frequency," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 10, pp. 2655–2662, 2009.
- [67] H. Li, H.-M. Rein, T. Suttorp, and J. Bock, "Fully integrated SiGe VCOs with powerful output buffer for 77-GHz automotive Radar systems and applications around 100 GHz," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 10, pp. 1650–1658, 2004.
- [68] D. Baek, T. Song, E. Yoon, and S. Hong, "8-GHz CMOS quadrature VCO using transformer-based LC tank," *IEEE Microwave and Wireless Components Letters*, vol. 13, no. 10, pp. 446–448, 2003.
- [69] B. Razavi, *Design of Integrated Circuits for Optical Communications*. McGraw-Hill Science/Engineering/Math, 2003.
- [70] B. Afshar, *Millimeter-Wave Circuits for 60GHz and Beyond*. PhD thesis, EECS Department, University of California, Berkeley, Aug. 2010.
- [71] A. Arbabian, B. Afshar, J.-C. Chien, S. Kang, S. Callender, E. Adabi, S. Toso, R. Pillard, D. Gloria, and A. Niknejad, "A 90GHz-carrier 30GHz-bandwidth hybrid switching transmitter with integrated antenna," in *IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pp. 420–421, 2010.
- [72] V. Jain, F. Tzeng, L. Zhou, and P. Heydari, "A single-chip dual-band 22-to-29GHz/77-to-81GHz BiCMOS transceiver for automotive radars," in *IEEE International Solid-State Circuits Conference - Digest of Technical Papers*, pp. 308–309, 2009.
- [73] A. Arbabian, S. Kang, S. Callender, B. Afshar, J.-C. Chien, and A. M. Niknejad, "A 90GHz pulsed-transmitter with near-field/far-field energy cancellation using a dual-loop antenna," in *IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, pp. 1–4, June 2011.

- [74] B. Razavi, "Design considerations for direct-conversion receivers," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 44, no. 6, pp. 428–435, 1997.
- [75] A. Abidi, "Direct-Conversion Radio Transceivers for Digital Communications," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1399–1410, 1995.
- [76] Ginzton, "Distributed Amplification," in *Proceedings of the IRE*, pp. 956–969, 1948.
- [77] T. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge University Press, 2003.
- [78] J. Beyer, S. Prasad, R. Becker, J. Nordman, and G. Hohenwarter, "MESFET Distributed Amplifier Design Guidelines," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, no. 3, pp. 268–275, 1984.
- [79] A. Yazdi, D. Lin, and P. Heydari, "A 1.8V three-stage 25GHz 3dB-BW differential non-uniform downsized distributed amplifier," in *IEEE International Solid-State Circuits Conference*, pp. 156–590, 2005.
- [80] R.-C. Liu, T.-P. Wang, L.-H. Lu, H. Wang, S.-H. Wang, and C.-P. Chao, "An 80ghz travelling-wave amplifier in a 90nm cmos technology," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, pp. 154–590, 2005.
- [81] J.-O. Plouchart, J. Kim, N. Zamdmer, L.-H. Lu, M. Sherony, Y. Tan, R. Groves, R. Trzcinski, M. Talbi, and Ray, "A 4-91 GHz distributed amplifier in a standard 0.12 m SOI CMOS microprocessor technology," in *IEEE Custom Integrated Circuits Conference*, pp. 159–162, 2003.
- [82] A. Jahanian and P. Heydari, "A CMOS distributed amplifier with active input balun using GBW and linearity enhancing techniques," *IEEE Radio Frequency Integrated Circuits Symposium*, 2011.
- [83] M.-D. Tsai, H. Wang, J.-F. Kuan, and C.-S. Chang, "A 70ghz cascaded multi-stage distributed amplifier in 90nm cmos technology," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, pp. 402–606, 2005.
- [84] B. Kleveland, C. Diaz, D. Vook, L. Madden, T. Lee, and S. Wong, "Exploiting CMOS reverse interconnect scaling in multi-gigahertz amplifier and oscillator design," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 10, pp. 1480–1488, 2001.
- [85] A. M. Niknejad, *Electromagnetics for High-Speed Analog and Digital Communication Circuits*. Cambridge Univ Press, 2007.

- [86] K. Niclas, W. Wilser, T. Kritzer, and R. Pereira, "On Theory and Performance of Solid-State Microwave Distributed Amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 31, no. 6, pp. 447–456, 1983.
- [87] C. Aitchison, "The Intrinsic Noise Figure of the MESFET Distributed Amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 33, no. 6, pp. 460–466, 1985.
- [88] M. Pospieszalski, "On the Measurement of Noise Parameters of Microwave Two-Ports," *IEEE Transactions on Microwave Theory and Techniques*, vol. 34, no. 4, pp. 456–458, 1986.
- [89] J. Liang and C. Aitchison, "Gain performance of cascade of single stage distributed amplifiers," *Electronics Letters*, vol. 31, no. 15, pp. 1260–1261, 1995.
- [90] G. D. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. Wiley-Interscience, July 2005.
- [91] Y. Zhu and H. Wu, "Distributed amplifiers with non-uniform filtering structures," *IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, 2006.
- [92] A. Arbabian and A. M. Niknejad, "A tapered cascaded multi-stage distributed amplifier with 370GHz GBW in 90nm CMOS," *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 57–60, 2008.
- [93] K. Moez and M. Elmasry, "A 10dB 44GHz Loss-Compensated CMOS Distributed Amplifier," in *IEEE International Solid-State Circuits Conference*, pp. 548–621, 2007.
- [94] S. Prasad, J. Beyer, and I.-S. Chang, "Power-bandwidth considerations in the design of MESFET distributed amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, no. 7, pp. 1117–1123, 1988.
- [95] S. M. Sze and K. K. Ng, *Physics of semiconductor devices*. Wiley-Blackwell, 2007.
- [96] C. Zhang, D. Huang, and D. Lou, "Optimization of cascode cmos low noise amplifier using inter-stage matching network," in *IEEE Conference on Electron Devices and Solid-State Circuits*, pp. 465–468, 2003.
- [97] B. Gilbert, "The micromixer: a highly linear variant of the gilbert mixer using a bisymmetric class-ab input stage," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 1412–1423, Sept. 1997.
- [98] D. Manstretta, M. Brandolini, and F. Svelto, "Second-order intermodulation mechanisms in CMOS downconverters," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 394–406, Mar. 2003.

- [99] H. Darabi and A. Abidi, "Noise in RF-CMOS mixers: a simple physical model," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 1, pp. 15–25, 2000.
- [100] S. Kang, J.-C. Chien, and A. M. Niknejad, "A 100GHz Phase-Locked Loop in 0.13 μm SiGe BiCMOS process," *IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, pp. 1–4, Mar. 2011.
- [101] T. Toifl, C. Menolfi, P. Buchmann, M. Kossel, T. Morf, R. Reutemann, M. Ruegg, M. L. Schmatz, and J. Weiss, "A 0.94-ps-RMS-jitter 0.016-mm² 2.5-GHz multiphase generator PLL with 360° digitally programmable phase shift for 10-Gb/s serial links," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2700–2712.
- [102] S. Callender and A. M. Niknejad, "A Phase-Adjustable Delay-Locked Loop Utilizing Embedded Phase Interpolation," *IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2011.
- [103] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Optimum Array Processing*. Wiley-Interscience, Apr. 2002.