

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Millimeter-Wave/Sub-Terahertz CMOS Transceivers for High-Speed Wireless Communications

### Permalink

<https://escholarship.org/uc/item/2d018435>

### Author

Kang, Shinwon

### Publication Date

2014

Peer reviewed|Thesis/dissertation

**Millimeter-Wave/Sub-Terahertz CMOS Transceivers  
for High-Speed Wireless Communications**

by

Shinwon Kang

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

in

Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ali M. Niknejad, Chair  
Professor Robert G. Meyer  
Professor Paul K. Wright

Spring 2014

**Millimeter-Wave/Sub-Terahertz CMOS Transceivers  
for High-Speed Wireless Communications**

Copyright 2014  
by  
Shinwon Kang

## Abstract

Millimeter-Wave/Sub-Terahertz CMOS Transceivers  
for High-Speed Wireless Communications

by

Shinwon Kang

Doctor of Philosophy in Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Ali M. Niknejad, Chair

Millimeter-wave and sub-terahertz frequency bands are available for wideband applications such as high data-rate communication systems. As the respective wavelength is on the order of a millimeter, a compact on-chip antenna can be designed, thereby reducing the overall form factor and obviating expensive off-chip packaging. However, the channel propagation loss increases significantly with the frequency. Although CMOS technology is prevalent in digital processing and data communication, CMOS devices are lossy and inefficient at such high frequencies. Thus, it is challenging to implement an efficient and wideband transceiver at sub-terahertz frequencies using CMOS technology.

The aim of this dissertation is to demonstrate sub-terahertz wireless links for high-speed chip-to-chip communication in CMOS. First, transceiver architectures and building blocks are discussed to address the challenges and limitations of the CMOS process. Two fully integrated CMOS transceivers, a 260 GHz OOK transceiver and a 240 GHz QPSK/BPSK transceiver, are then demonstrated using on-chip antennas. Frequency multiplication and mixer-first design are employed to operate beyond the cut-off frequency. In the QPSK modulation, a maximum data rate of 16 Gbps is realized with an energy efficiency of 30 pJ/bit.

These demonstrations show that millimeter-wave/sub-terahertz wireless communication can be a promising solution for high-speed chip-to-chip communication. Improvements in the energy efficiency and silicon area of these wireless links can result in replacing or complementing existing wired links.

To my wife Hyesun.

# Contents

<b>Contents</b>	<b>ii</b>
<b>List of Figures</b>	<b>v</b>
<b>List of Tables</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Organization . . . . .	3
<b>2 Background Studies</b>	<b>4</b>
2.1 Wireless Communication . . . . .	4
2.1.1 Link Budget . . . . .	4
2.1.2 Modulation and BER . . . . .	7
2.2 CMOS Technology . . . . .	9
2.2.1 Active Device . . . . .	9
2.2.2 Passive Device . . . . .	10
<b>I Building Blocks for High-Speed Wireless Communications</b>	<b>12</b>
<b>3 Frequency Generation</b>	<b>13</b>
3.1 <i>LC</i> VCO . . . . .	13
3.2 65 GHz <i>LC</i> VCO . . . . .	17
3.3 Varactor-less <i>LC</i> VCO . . . . .	19
3.4 100 GHz Active-Varactor VCO . . . . .	22
3.5 Injection-Locked Loop . . . . .	29
3.6 Frequency Multiplication . . . . .	37
3.6.1 <i>N</i> -push Multiplier . . . . .	37
3.6.2 <i>N</i> -push Oscillator . . . . .	42
3.6.3 Harmonic Matching . . . . .	43
3.6.4 Up-Converting Mixer . . . . .	44
3.7 Frequency Synthesizer . . . . .	45

3.7.1	PLL with a Fundamental VCO . . . . .	45
3.7.2	PLL with an $N$ -push VCO . . . . .	45
3.7.3	PLL and a Frequency Multiplier . . . . .	47
3.7.4	PLL and an Injection-Locked Oscillator . . . . .	47
3.7.5	Design Considerations . . . . .	48
3.8	LO Distribution . . . . .	48
<b>4</b>	<b>Transmitter</b>	<b>50</b>
4.1	Conventional Transmitter Architecture . . . . .	50
4.2	Sub-Terahertz Transmitter Architecture . . . . .	51
4.3	I/Q Generation . . . . .	53
4.3.1	Quadrature Oscillator . . . . .	53
4.3.2	Quadrature Hybrid . . . . .	54
4.3.3	Delayed Line . . . . .	57
4.4	Modulator . . . . .	57
4.4.1	OOK . . . . .	58
4.4.2	QPSK and BPSK . . . . .	58
4.5	80 GHz Power Amplifier . . . . .	59
4.6	240 GHz Frequency Tripler . . . . .	64
<b>5</b>	<b>Receiver</b>	<b>67</b>
5.1	Conventional Receiver Architecture . . . . .	67
5.2	Sub-Terahertz Receiver Architectures . . . . .	68
5.2.1	Diode . . . . .	68
5.2.2	Self-Mixer . . . . .	69
5.2.3	Sub-harmonic Mixer . . . . .	69
5.2.4	Heterodyne Conversion Mixer . . . . .	70
5.2.5	Direct Conversion Mixer . . . . .	72
5.3	OOK Demodulator . . . . .	72
5.4	Leakage-Free Receiver Design . . . . .	75
<b>6</b>	<b>Baseband Circuits</b>	<b>77</b>
6.1	High-Speed Logic Gates . . . . .	77
6.2	PRBS Generator . . . . .	80
6.3	Baseband Amplifier . . . . .	83
6.4	Operational Amplifier . . . . .	86
<b>II</b>	<b>Fully Integrated Wireless CMOS Transceivers</b>	<b>89</b>
<b>7</b>	<b>260 GHz Wireless OOK Transceiver</b>	<b>90</b>
7.1	System Overview . . . . .	90

7.2	Transmitter . . . . .	91
7.3	Receiver . . . . .	93
7.4	Fabrication . . . . .	94
7.5	Measurement Results . . . . .	94
<b>8</b>	<b>240 GHz QPSK/BPSK Transceiver</b>	<b>99</b>
8.1	System Overview . . . . .	99
8.2	Transmitter . . . . .	101
8.3	Receiver . . . . .	104
8.4	Fabrication . . . . .	105
8.5	Measurement Results . . . . .	106
<b>9</b>	<b>Advanced Transceivers</b>	<b>118</b>
9.1	Higher-Order Modulation . . . . .	118
9.2	Higher Carrier Frequency . . . . .	119
9.3	Phased Array . . . . .	120
9.4	Multiple Channels . . . . .	120
9.5	Lens . . . . .	121
<b>10</b>	<b>Conclusion</b>	<b>123</b>
	<b>Bibliography</b>	<b>125</b>



# List of Figures

2.1	Atmospheric attenuation. . . . .	5
2.2	BER curves of selected modulations (C-OOK: coherent OOK, NC-OOK: non-coherent OOK). . . . .	7
3.1	Schematics of <i>LC</i> VCOs. . . . .	14
3.2	Schematic of a passive varactor. . . . .	17
3.3	Quality factors of passive varactors (worst-case, from extraction and post-layout simulation). . . . .	17
3.4	Schematic of the 65 GHz <i>LC</i> VCO. . . . .	19
3.5	Simulation results of the 65 GHz VCO (a) frequency (b) output power (c) phase noise. . . . .	20
3.6	65 GHz VCO (a) microphotograph (b) measured frequency. . . . .	21
3.7	A loss-less transformer. . . . .	21
3.8	Concept of the active-varactor VCO. . . . .	23
3.9	Schematics of the proposed active-varactor VCO (a) fully off (b) fully on. . . . .	24
3.10	Schematic of the 100 GHz active-varactor VCO. . . . .	25
3.11	Microphotograph of the 100 GHz active-varactor VCO. . . . .	26
3.12	Measurement setup. . . . .	27
3.13	Measurement results of the 100 GHz active-varactor VCO (a) frequency (b) power dissipation (c) output spectrum. . . . .	28
3.14	Architectures of mutual coupling (a) uni-directional coupling (b) bi-directional coupling using buffers (c) bi-directional capacitive coupling (this work). . . . .	29
3.15	Phase diagram of injection locking in the bi-directional coupling. . . . .	30
3.16	Schematic of the injection-locked loop. . . . .	31
3.17	Layout of the injection-locked loop. . . . .	32
3.18	Design of passive devices (a) transformer (b) transmission-line-based capacitive coupling. . . . .	32
3.19	Microphotograph of the injection-locked loop. . . . .	33
3.20	Measurement results of the injection-locked loop (a) frequency (b) power dissipation (c) output spectrum. . . . .	34
3.21	Comparison of the 100 GHz active-varactor VCO and the injection-locked loop (a) frequency (b) power dissipation (c) output spectrum (d) phase noise. . . . .	35

3.22	Multi-phase waveforms. . . . .	36
3.23	An ideal non-linear block. . . . .	37
3.24	Operation of $N$ -push multiplier. . . . .	38
3.25	$g_0$ and $g_N$ of $N$ -push multipliers (a) Push-push ( $N=2$ ) (b) Triple-push ( $N=3$ ) (c) Quadruple-push ( $N=4$ ). . . . .	40
3.26	Schematic of the 26.6 GHz push-push doubler. . . . .	41
3.27	Simulation results of the 26.6 GHz push-push doubler (a) output currents (b) efficiency (matches to $ g_N /g_0$ ). . . . .	42
3.28	Operation of harmonic matching. . . . .	43
3.29	Operation of up-converting mixer. . . . .	44
3.30	Schematic of the 240 GHz frequency tripler. . . . .	44
3.31	Frequency synthesizer architectures (a) PLL with a fundamental VCO (b) PLL with an $N$ -push VCO (c) PLL with a frequency multiplier (d) PLL with an injection-locked oscillator. . . . .	46
4.1	Conventional transmitter architecture for I/Q modulations. . . . .	50
4.2	Sub-terahertz transmitter architecture. . . . .	51
4.3	Sub-terahertz transmitter architectures (a) with doubler (b) with tripler (c) with quadrupler. . . . .	52
4.4	Operation of a quadrature oscillator. . . . .	54
4.5	Quadrature hybrids (a) transmission-line-based (b) lumped-element-based (c) transformer-based (d) differential input/output. . . . .	55
4.6	Simulation results of the 94 GHz quadrature hybrid (a) S-parameters (b) amplitude/phase mismatches. . . . .	56
4.7	Delayed lines (a) $90^\circ$ delay (b) $180^\circ$ delay. . . . .	57
4.8	Schematic of QPSK modulator. . . . .	59
4.9	Simulation results of the QPSK modulator (LO power at the hybrid input) (a) output power (b) power dissipation. . . . .	60
4.10	Schematic of the class-E switching power amplifier . . . . .	61
4.11	Simulation results of the 80 GHz power amplifier. . . . .	62
4.12	Inter-stage matching transformer for 80 GHz amplifiers (not to scale). . . . .	62
4.13	Schematic of the 80 GHz amplifiers of the 240 GHz QPSK/BPSK transmitter. . . . .	63
4.14	Schematic of the 240 GHz frequency tripler. . . . .	64
4.15	Layout design of the 240 GHz frequency tripler. . . . .	65
4.16	Simulation results of the 240 GHz transmitter and frequency tripler. . . . .	66
5.1	Conventional receiver architecture for I/Q modulations. . . . .	67
5.2	Receiver architecture with a diode detector. . . . .	68
5.3	Receiver architecture with a self-mixer. . . . .	69
5.4	Receiver architecture with a sub-harmonic mixer. . . . .	69
5.5	Receiver architecture with a heterodyne conversion mixer. . . . .	70
5.6	Schematic of the 260 GHz heterodyne mixer. . . . .	71

5.7	Simulation results of the 260 GHz heterodyne mixer (a) conversion gain (b) double sideband noise figure. . . . .	71
5.8	Receiver architecture with a direct-conversion mixer. . . . .	72
5.9	Operation of an OOK demodulator (a) ON state (b) OFF state. . . . .	73
5.10	An ideal non-linear block with sub-threshold current. . . . .	73
5.11	Comparison of $I_{DC,ON}$ and $I_{DC,OFF}$ . . . . .	75
6.1	Buffer and inverter (a) symbol (b) CML. . . . .	78
6.2	NAND/AND and NOR/OR (a) symbol (b) CML. . . . .	78
6.3	XNOR/XOR (a) symbol (b) CML. . . . .	79
6.4	Multiplexer (a) symbol (b) CML (c) current-steering multiplexer. . . . .	79
6.5	Latch (a) symbol (b) CML (c) an alternative representation. . . . .	79
6.6	PRBS generator. . . . .	81
6.7	PRBS distributor for OOK modulation. . . . .	81
6.8	PRBS generator for QPSK/BPSK/continuous-wave modes. . . . .	82
6.9	Output buffer of the PRBS generator (a) schematic (b) equivalent model. . . . .	83
6.10	Schematic of the baseband amplifier. . . . .	84
6.11	Simulation results of the baseband amplifier (a) gain (b) noise figure. . . . .	85
6.12	Schematic of the operational transconductance amplifier. . . . .	86
6.13	Schematic of the biasing circuit. . . . .	87
6.14	Stacked devices (a) device symbol shown in schematics (b) real implementation. . . . .	87
6.15	Layout of the operational transconductance amplifier. . . . .	88
6.16	Simulation results of the OTA (a) gain and frequency (b) gain and input common-mode voltage. . . . .	88
7.1	Architecture of the 260 GHz OOK transceiver. . . . .	91
7.2	Block diagram of the 260 GHz transmitter. . . . .	92
7.3	Block diagram of the 260 GHz receiver. . . . .	93
7.4	Microphotograph of the 260 GHz OOK transceiver. . . . .	94
7.5	Measurement setup for the transmitter. . . . .	95
7.6	Measurement setup for the V-band leakage. . . . .	95
7.7	Measured frequency spectra of OOK modulation signals (a) 2 Gbps (b) 6 Gbps (c) 10 Gbps (d) 14 Gbps. . . . .	96
7.8	Measurement setup for the transmitter and receiver. . . . .	97
7.9	Measured on/off states (on-off ratio is 10 dB). . . . .	98
8.1	Architecture of the 240 GHz QPSK/BPSK transceiver. . . . .	100
8.2	Block diagram of the 240 GHz transmitter. . . . .	102
8.3	QPSK constellation points preserved by a frequency tripler. . . . .	103
8.4	Layout design of the transmitter. . . . .	103
8.5	Block diagram of the 240 GHz receiver. . . . .	104
8.6	Layout design of the receiver. . . . .	106

8.7	Microphotographs of the transmitter and receiver. . . . .	107
8.8	Measurement setup for the transmitter. . . . .	108
8.9	Measured output power of continuous-wave signals. . . . .	109
8.10	Measured frequency spectrum of 4 GSps modulation signal ( $f_{null,1} = 4$ GHz, $f_{beat} = 31.5$ MHz). . . . .	110
8.11	Measured frequency spectrum of 6 GSps modulation signal ( $f_{null,1} = 6$ GHz, $f_{beat} = 47.2$ MHz). . . . .	110
8.12	Measured frequency spectrum of 8 GSps modulation signal ( $f_{null,1} = 8$ GHz, $f_{beat} = 63.0$ MHz). . . . .	111
8.13	Measured transmitter eye diagrams (a) 4 GSps (b) 6 GSps (c) 8 GSps. . . . .	111
8.14	Measurement setup for the transmitter and receiver. . . . .	112
8.15	Measured output power of continuous-wave signals. . . . .	113
8.16	Measured I and Q output powers (a) varying transmitter frequency (b) varying receiver frequency. . . . .	114
8.17	Measured receiver eye diagrams (a) 4 Gbps BPSK (b) 6 Gbps BPSK (c) 8 Gbps BPSK (d) 9 Gbps BPSK. . . . .	114
8.18	Measured receiver output spectra (a) 8 Gbps QPSK (b) 16 Gbps QPSK. . . . .	115
8.19	Measured receiver eye diagrams (a) 8 Gbps QPSK (b) 12 Gbps QPSK (c) 16 Gbps QPSK. . . . .	115
8.20	BER bathtub curves of BPSK modulation. . . . .	116
8.21	BER bathtub curves of QPSK modulation. . . . .	116
9.1	Constellation points (a) QPSK (b) 16-QAM. . . . .	119
9.2	Phased array. . . . .	120
9.3	Multi-channel communication. . . . .	121
9.4	Ray diagrams (a) without lens (b) with lenses. . . . .	122

# List of Tables

3.1	Circuit parameters of the 100 GHz active-varactor VCO. . . . .	26
3.2	VCO performance summary and comparison . . . . .	27
3.3	Comparison of bi-directionally injection-locked loops. . . . .	36
3.4	Comparison of the $N$ -push multipliers . . . . .	41
3.5	Comparison of the frequency synthesizer architectures . . . . .	47
4.1	Simulation results of the 80 GHz class-E switching power amplifier . . . . .	60
7.1	Link parameters for the 260 GHz OOK transceiver . . . . .	92
7.2	Summary of the 260 GHz OOK transceiver . . . . .	97
8.1	Link parameters for the 240 GHz QPSK transceiver . . . . .	101
8.2	Power breakdown of the 240 GHz transmitter . . . . .	108
8.3	Comparison of sub-terahertz transmitters . . . . .	109
8.4	Power breakdown of the 240 GHz receiver . . . . .	113
8.5	Comparison of sub-terahertz receivers . . . . .	117
8.6	Comparison of sub-terahertz transceivers . . . . .	117

## Acknowledgments

It is a great pleasure to acknowledge those who have helped me write this dissertation thesis and complete the Ph.D. degree during my time at Berkeley.

First of all, I would like to thank my research advisor, professor Ali M. Niknejad, for providing sufficient guidance and advice. He has encouraged me to come up with new ideas and design circuits and systems. His confidence and patience are the primary reasons that I could successfully demonstrate a sub-terahertz wireless transceiver.

I would also like to thank professor Robert G. Meyer for reading this dissertation and M.S. degree report, professor Paul K. Wright for agreeing to be on my dissertation committee and qualifying exam committee, and professor Jan M. Rabaey for serving as the chair of my qualifying exam committee. Their feedbacks and comments have kept me in the right direction.

I thank professor Elad Alon as I took many of his circuit courses and learned theoretical knowledge and practical skills. I also thank many other professors, Borivoje Nikolic, Jaijeet Roychowdhury, Kannan Ramchandran, and Martin White, for valuable lectures.

I especially thank Siva Thyagarajan for working on two sub-terahertz transceivers together. I cannot forget the moment we saw the first eye diagram. I also thank Jungdong for working on an on-off keying transceiver.

I thank TUSI members, Amin Arbabian, Jun-Chau Chien, Steven Callender, Bagher Afshar, and Ehsan Adabi. It was a pleasure to work with them as a team member. I also thank research group members, Cristian Marcu, Sahar Tabesh, Jiashu Chen, Lu Ye, Sriramkumar Venugopalan, Andrew Townley, Paul Swirhun, Greg Lacaille, Lucas Calderin, and Nai-Chung Kuo.

I thank the Berkeley Wireless Research Center (BWRC). My research would have not gone without BWRC faculty, staff, and sponsors. The BWRC laboratory and other facilities are amazing. I also thank BWRC students and friends, Namseog, Jihoon, Kwangmo, Kyoo-hyun, Jaehwa, Jaeduk, Jesse, Lingkai, Yue, Chintan, Charles, Yida, Wen, Thura, Travis, Ping-Chen, Dan, Matt, John, Seyong, and many others.

I acknowledge Samsung scholarship for supporting me for five years. I also acknowledge Qualcomm Inc. for allowing me to intern twice. It was a great opportunity to broaden my horizons. I also acknowledge support under NSF Grant ECCS-1201755 and SRC/TxACE 1836.082, and TSMC and STMicroelectronics for donating the fabrication support.

Finally, and most importantly, none of this would have been possible without support, encouragement, patience, and love of my wife Hyesun Yeom. I dedicate this dissertation to Hyesun.

# Chapter 1

## Introduction

### 1.1 Introduction

The term millimeter-wave generally refers to the frequency range from 30 GHz to 300 GHz because the corresponding wavelength is between 10 mm and 1 mm, and the term sub-terahertz refers to the frequencies below 1 THz. The millimeter-wave/sub-terahertz frequency range includes molecular vibration frequencies and absorption bands of the atmosphere and common materials and is therefore important for fundamental applications. This frequency range is also useful for technological applications, such as imagers and radars. Imaging at these frequencies can be used to detect concealed weapons or explosives and reveal features in skin, teeth, and solid organs. Radars are used to measure the distance or speed of aircrafts, ships, automotive vehicles, and many other objects.

There are many wide and unallocated frequency bands in the millimeter-wave/sub-terahertz frequencies that are attractive for wideband applications. The low frequency bands (below 10 GHz) are narrowly divided and have been legally appropriated for various existing applications, making it difficult to accommodate the wide bandwidth requirements of emerging applications. For example, a 10 Gbps BPSK communication system requires a bandwidth of approximately 20 GHz around the carrier frequency [1], [2]. A pulse width of 50 ps in a pulse-based radar results in a bandwidth of 20 GHz [3].

Millimeter-wave/sub-terahertz frequencies have not been fully explored in integrated circuits and systems. The 60 GHz frequency band has been utilized and standardized for data communications. Many IC prototypes have been demonstrated at the frequencies between 30 GHz and 100 GHz for radars, imagers, and communications. However, frequencies above 100 GHz have recently been investigated and demonstrated. As the operating frequency increases, it becomes increasingly difficult to design and implement efficient and robust integrated circuits because of the limited performance of devices, the process variations, and inaccurate measurements and device models.

Advances in process technologies are beneficial for the millimeter-wave/sub-terahertz frequencies. As the technology has been scaled down, the cut-off frequencies ( $f_T$ ,  $f_{max}$ ) have

increased, resulting in higher operation frequencies. In particular, very high  $f_{max}$  values over 1 THz have been achieved for III-V semiconductors [4], and  $f_{max}$  values above 500 GHz have been achieved for SiGe process technologies [5]. However, these technologies are too expensive to be used ubiquitously in consumer devices. For decades, CMOS technology has dominated digital and mixed-signal ICs because of its good switching performance and low cost. CMOS technology also offers advantages for millimeter-wave/sub-terahertz circuits and systems. Processors or memories can be integrated on a chip, and the received data can be quickly processed and stored. In addition, high-frequency circuits can be digitally controlled and calibrated.

Despite these advantages, many challenges and limitations to CMOS technology should still be overcome to realize millimeter-wave/sub-terahertz circuits and systems. The cut-off frequencies ( $f_T$ ,  $f_{max}$ ) that are typically between 200 GHz and 300 GHz are relatively lower than for other technologies [6]. The breakdown voltages are so low that the supply voltage and AC voltage swings should be restricted for reliability. In addition, passive devices suffer from high losses and low quality factors. As the technology is scaled down, the thickness and the elevation of the metals also shrink, thereby increasing losses in transmission lines and inductors. Frequency multiplication and power combining have been proposed to overcome these challenges, but the resulting conversion gain and efficiency are still not sufficiently high.

Wireless communication requires antennas to convert electric signals to electromagnetic waves, and vice versa. For a fixed antenna gain, the antenna size (area) is proportional to the square of the wavelength. The wavelength at sub-terahertz frequencies is on the order of 1 mm, such that antennas can be integrated on the same chip as other transceiver circuits, thereby reducing the overall form factor and obviating expensive off-chip packaging. However, this on-chip antenna solution is mitigated by the high losses of passive devices in CMOS technology. The radiation efficiency and antenna gain are degraded by the low resistivity ( $\approx 10 \Omega\text{cm}$ ) and high (relative) permittivity ( $\approx 12$ ) of the silicon substrate. The antenna bandwidth should also be sufficiently wide to achieve a high data rate ( $> 10$  Gbps). Thus, it is difficult to achieve high gain and wide bandwidth simultaneously, and a novel on-chip antenna design is required.

Electronic devices have multiple chips on a printed circuit board (PCB). A large amount of data can be exchanged between chips, for example, between two processor chips or between a processor chip and a memory chip. As devices require more functions and applications, more chips are used on a small PCB, and the interconnections among chips become more complicated. It can be very challenging to design a PCB because of the numerous interconnections and high data rates. To complement a wired link and obtain more flexibility, wireless chip-to-chip communication has been developed. A high data rate requires a wide bandwidth; thus, the millimeter-wave/sub-terahertz frequency can be used for chip-to-chip communication. Moreover, if a wireless transceiver is implemented with the CMOS technology, then it is even more pronounced. The energy efficiency and the silicon area are important factors for complementing a wired link. In this dissertation, millimeter-wave/sub-terahertz circuit techniques are discussed, and two sub-terahertz CMOS transceivers are



demonstrated. The aim of this dissertation is also to develop novel ideas and designs. At the circuit level, novel circuits are designed and verified to overcome the limitations of the CMOS process. At the architecture level, various transmitter and receiver architectures are analyzed.

## 1.2 Organization

The dissertation is organized as follows. Chapter 2 present preliminary knowledge of the wireless communication. The advantages and challenges of CMOS technology are also discussed.

In Part I, building blocks are demonstrated for sub-terahertz high-speed wireless communications. Chapter 3 discusses frequency generation circuits, such as oscillator, frequency multiplier, injection-locked loop, and phase-locked loop. Chapter 4 describes transmitter architectures and circuits. Quadrature phase generation, data modulation, and power amplification are discussed. Chapter 5 describes receiver architectures and circuits. Several mixer-first architectures are discussed. In addition, a mixer design and a demodulation are presented. Chapter 6 describes baseband circuits, such as PRBS generator, baseband amplifier, and operational amplifier.

In Part II, integrated wireless CMOS transceivers are demonstrated. Chapter 7 shows a 260 GHz OOK transceiver and Chapter 8 demonstrates a 240 GHz QPSK/BPSK transceiver. System designs and transceiver measurements are presented. Chapter 9 discusses advanced architectures, which may be good resources for future research.

Finally, Chapter 10 summarizes and concludes this dissertation.

# Chapter 2

## Background Studies

This chapter presents preliminary knowledge to better understand this dissertation. Communication theories and practical equations are described for designing wireless transceivers. Finally, the limitations of the CMOS technology are reviewed for the millimeter-wave/sub-terahertz circuits and systems.

### 2.1 Wireless Communication

In wireless communications, the required transmit power and receiver noise figure are determined by the channel loss, the signal bandwidth, and the signal-to-noise ratio (SNR). The link budget and the bit error rate are analyzed in this section. All the equations presented in this section can be obtained and derived from [7]–[9].

#### 2.1.1 Link Budget

Transmitters emit modulated signal to the air by an antenna. When the output power is  $P_{TX}$ , the effective isotropic radiated power (EIRP) can be larger than  $P_{TX}$  because the transmit signal is not radiated isotropically but directed by an antenna. EIRP is simply calculated by

$$EIRP_{TX} = P_{TX} \cdot G_{A,TX} \quad (2.1)$$

where the transmitter antenna gain ( $G_{A,TX}$ ) is given by

$$G_{A,TX} = E_A \cdot D \quad (2.2)$$

where  $E_A$  is the antenna efficiency and  $D$  is the directivity.

At millimeter-wave/sub-terahertz frequencies, the channel loss ( $L_{channel}$ ) mainly includes the path loss ( $L_{path}$ ) and atmospheric loss ( $L_{atm}$ ), neglecting other interferences.

$$L_{channel} = L_{path} \cdot L_{atm} \quad (2.3)$$

The path loss is given by the Friis equation, which is

$$L_{path} = \left( \frac{\lambda}{4\pi d} \right)^2 \quad (2.4)$$

$$L_{path,dB} = 20 \log_{10} \left( \frac{\lambda}{4\pi d} \right) \quad (2.5)$$

where  $d$  is the distance and  $\lambda$  is the wavelength of the radiated signal.

The atmospheric loss is given by

$$L_{atm,dB} = -\alpha_{atm}(f) \cdot d \quad (2.6)$$

$$L_{atm} = 10^{-\frac{1}{10}\alpha_{atm}(f) \cdot d} \quad (2.7)$$

where  $d$  is the distance and  $\alpha_{atm}(f)$  is the measured values of the atmospheric attenuation.  $\alpha_{atm}(f)$  is plotted in Fig. 2.1 using [10] (P=1 atm, T=288.15 °K, precipitable water vapor=2 cm). Because the loss is relevant to absorption of molecules, the atmospheric loss is tabularized, not represented as an equation.

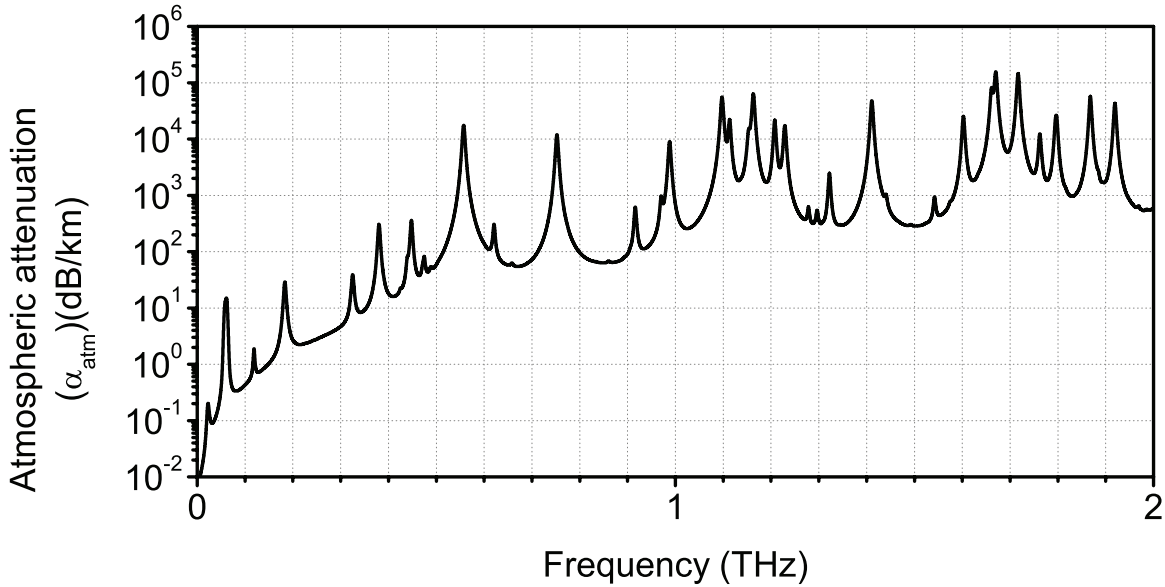


Figure 2.1: Atmospheric attenuation.

The path loss is proportional to the square of the distance (Eq. 2.4), but the atmospheric loss varies exponentially with the distance (Eq. 2.7). Thus, the path loss dominates the overall channel loss in a short distance, but the atmospheric loss dominates in a long distance. For example, at 250 GHz, two losses are same when the distance is approximately 63 km. At 550 GHz, two losses are equal when the distance is approximately 40 m.

In the receiver, the signal power at the antenna output is defined as the received power ( $P_{RX}$ ), which is given by

$$P_{RX} = EIRP_{TX} \cdot L_{channel} \cdot G_{A,RX} = P_{TX} \cdot G_{A,TX} \cdot L_{channel} \cdot G_{A,RX} \quad (2.8)$$

where  $G_{A,RX}$  is the gain of the receiver antenna. The noise power is simply assumed by

$$N_{IN} = k \cdot T \cdot W \quad (2.9)$$

where  $k$  is the Boltzmann constant,  $T$  is the temperature, and  $W$  is the bandwidth. The SNR is then given by

$$SNR_{IN} = \frac{P_{TX} \cdot G_{A,TX} \cdot L_{channel} \cdot G_{A,RX}}{k \cdot T \cdot W} \quad (2.10)$$

The receiver amplifies, down-converts, and then amplifies again. It is assumed that the overall gain of the receiver blocks is  $G_{RX}$  and the output noise generated by only the receiver blocks is  $N_{RX}$ . The noise factor is then given by

$$F = \frac{SNR_{IN}}{SNR_{OUT}} = \frac{S_{IN}}{S_{OUT}} \cdot \frac{N_{OUT}}{N_{IN}} = \frac{1}{G_{RX}} \cdot \frac{G_{RX}N_{IN} + N_{RX}}{N_{IN}} = 1 + \frac{N_{RX}}{G_{RX}N_{IN}} \quad (2.11)$$

Accordingly, the SNR at the receiver output is

$$SNR_{OUT} = SNR_{IN} \cdot \frac{1}{F} = \frac{P_{TX} \cdot G_{A,TX} \cdot L_{channel} \cdot G_{A,RX}}{k \cdot T \cdot W \cdot F} \quad (2.12)$$

From this equation, many parameters affects the received SNR; thus, the link budget should be carefully planned and designed.

Typically, communication systems specify the required SNR for the system performance because the bit error rate (BER) is determined by the received SNR. In such real systems, the received SNR ( $SNR_{OUT}$ ) should be higher than the required SNR ( $SNR_{reqd}$ ), introducing the link margin ( $LM$ ).

$$SNR_{OUT} = SNR_{reqd} \cdot LM \quad (2.13)$$

Therefore, the link margin is given by

$$LM = \frac{P_{TX} \cdot G_{A,TX} \cdot L_{channel} \cdot G_{A,RX}}{SNR_{reqd} \cdot k \cdot T \cdot W \cdot F} \quad (2.14)$$

In practice, the implementation losses should be taken into account and the link margin should be properly estimated.

### 2.1.2 Modulation and BER

The bit error rate (BER) depends on the modulation and the received SNR. Fig. 2.2 shows BER curves of some popular modulations.  $E_b$  is the bit energy, written as

$$E_b = \frac{S_{OUT}}{R} = \frac{P_{RX} \cdot G_{RX}}{R} \quad (2.15)$$

where  $R$  is the bit rate.  $N_o$  is the noise power spectral density, given by

$$N_o = \frac{N_{OUT}}{W} \quad (2.16)$$

where  $W$  is the bandwidth. The BER is then a function of  $E_b/N_o$ , as plotted in Fig. 2.2.

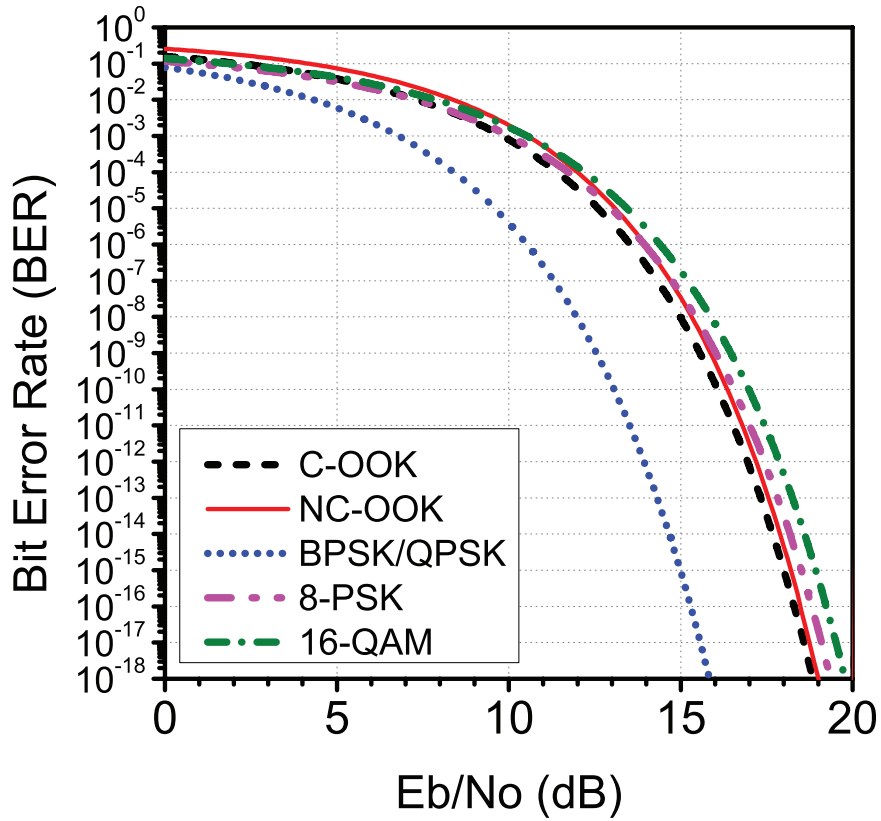


Figure 2.2: BER curves of selected modulations (C-OOK: coherent OOK, NC-OOK: non-coherent OOK).

In this section,  $P_E$  is defined as probability of the symbol error and BER is given by

$$BER \approx \frac{P_E}{\log_2 M} \quad (2.17)$$

where  $M$  is the size of alphabet per symbol and  $\log_2 M$  is the number of bits per symbol.

For coherent OOK (on-off keying) modulation, there is one bit per symbol, and BER is equal to  $P_E$ , expressed as

$$BER_{C-OOK} = P_E = Q\left(\sqrt{\frac{E_b}{N_o}}\right) \quad (2.18)$$

For non-coherent OOK modulation, error probabilities are given by [7].  $P_{E,S}$  is the error probability of a high bit being received as a low bit, as

$$P_{E,S} \approx 1 - Q\left(\sqrt{\frac{4E_b}{N_o}}, \sqrt{2 + \frac{E_b}{N_o}}\right) \quad (2.19)$$

where  $Q(\cdot, \cdot)$  is the Marcum Q function ( $M=1$ ).  $P_{E,M}$  is the error probability of a low bit being received as a high bit, as

$$P_{E,M} \approx \exp\left(-\frac{1}{2}\sqrt{2 + \frac{E_b}{N_o}}\right) \quad (2.20)$$

If high bits and low bits are equally distributed, the overall probability is

$$BER_{NC-OOK} = P_E \approx \frac{1}{2}(P_{e,S} + P_{e,M}) \quad (2.21)$$

The BER curves of the coherent and non-coherent OOK modulations are plotted in Fig. 2.2, showing that the coherent OOK has lower BER than the non-coherent OOK.

The BER of the coherent FSK (frequency-shift keying) is

$$BER_{C-FSK} = P_E = Q\left(\sqrt{\frac{E_b}{N_o}}\right) \quad (2.22)$$

and the BER of the non-coherent FSK is

$$BER_{NC-FSK} = P_E = \exp\left(-\frac{E_b}{2N_o}\right) \quad (2.23)$$

For the coherent BPSK (binary phase-shift keying), the BER is

$$BER_{BPSK} = P_E = Q\left(\sqrt{\frac{2E_b}{N_o}}\right) \quad (2.24)$$

and the BER of the non-coherent DBPSK (differential BPSK) is

$$BER_{DBPSK} = P_E = \exp\left(-\frac{E_b}{N_o}\right) \quad (2.25)$$

For  $M$ -PSK, the probability of the symbol error is

$$P_E \approx 2Q \left( \sqrt{\frac{2 \log_2 M \cdot E_b}{N_o}} \cdot \sin \frac{\pi}{M} \right) \quad (2.26)$$

and the BER is

$$BER_{M-PSK} = \frac{P_E}{\log_2 M} \approx \frac{2}{\log_2 M} Q \left( \sqrt{\frac{2 \log_2 M \cdot E_b}{N_o}} \cdot \sin \frac{\pi}{M} \right) \quad (2.27)$$

Specifically, error probabilities of QPSK (quadrature phase-shift keying) are

$$P_E \approx 2Q \left( \sqrt{\frac{2E_b}{N_o}} \right) \quad (2.28)$$

$$BER_{QPSK} = \frac{P_E}{\log_2 4} \approx Q \left( \sqrt{\frac{2E_b}{N_o}} \right) \quad (2.29)$$

From this, the BER of QPSK is the same as the BER of BPSK for the given  $E_b/N_o$ , as shown in Fig. 2.2. Besides, BPSK and QPSK show lower BER than other modulations.

Finally, for  $M$ -QAM (quadrature amplitude modulation) modulations, if  $\log_2 M$  is even, the error probabilities are

$$P_E \approx 4 \left( 1 - \frac{1}{\sqrt{M}} \right) Q \left( \sqrt{\frac{3 \log_2 M}{M-1} \cdot \frac{E_b}{N_o}} \right) \quad (2.30)$$

$$BER_{M-QAM} = \frac{P_E}{\log_2 M} \approx \frac{4}{\log_2 M} \left( 1 - \frac{1}{\sqrt{M}} \right) Q \left( \sqrt{\frac{3 \log_2 M}{M-1} \cdot \frac{E_b}{N_o}} \right) \quad (2.31)$$

In Fig. 2.2, BERs of 8-PSK and 16-QAM are also plotted.

## 2.2 CMOS Technology

CMOS technology offers numerous advantages, as discussed in Chapter 1. This section reviews the disadvantages and limitations of CMOS technology for millimeter-wave/sub-terahertz circuits and systems. The details are discussed in [11]–[13].

### 2.2.1 Active Device

The high-frequency performance of an active device can be represented in terms of the unity-gain frequency ( $f_T$ ) and the maximum oscillation frequency ( $f_{max}$ ). As the frequency

increases, the transistor gain decreases because of parasitic capacitors and resistors. The  $f_T$  is given by [12]

$$f_T = \frac{\omega_T}{2\pi} = \frac{1}{2\pi} \frac{g_m}{C_{gs} + C_{gd}} \quad (2.32)$$

where  $g_m$  is the transconductance of the transistor,  $C_{gs}$  is the gate-source capacitance, and  $C_{gd}$  is the gate-drain capacitance. In addition, the  $f_{max}$  is estimated by [11], [12]

$$f_{max} \approx \frac{f_T}{2\sqrt{g_m R_g (C_{gd}/C_{gg}) + g_{ds} (R_g + r_{ch} + R_s)}} \quad (2.33)$$

where  $R_g$ ,  $r_{ch}$ , and  $R_s$  are the gate resistance, the channel resistance, and the source resistance, respectively.  $C_{gd}$ ,  $C_{gg}$ , and  $g_{ds}$  are the gate-drain capacitance, the gate capacitance, and the drain-source conductance, respectively. Unlike  $f_T$ ,  $f_{max}$  is strongly dependent on the gate resistance ( $R_g$ ) and the device layout, and  $f_{max}$  is therefore difficult to estimate. By definition, no current gain and no power gain can be achieved beyond  $f_T$  and  $f_{max}$ , respectively. Thus, if the signal frequency exceeds the cut-off frequencies, the amplifiers become lossy, and there is no way to boost the signal power in the given process.

Active devices suffer from breakdown issues [13]. The gate oxide of the short-channel CMOS devices is thin  $< 10$  nm; thus, an applied electric field can cause oxide breakdown. For reliable operations, this breakdown should be prevented by limiting the voltage swing. The supply voltage in a typical bulk CMOS process is equal to or lower than 1 V, which limits the voltage swing of AC signals. Thus, the maximum output power is also limited. High output power can be realized by lowering the impedance, but the resulting increase in the impedance transformation ratio increases the insertion loss of the matching network [12], [13].

Device modeling presents another challenge. Increases in the frequency up to 100 GHz or higher make it difficult to accurately measure and characterize the devices. Measurement is limited by (de-)embedding and calibration. It is increasingly challenging to characterize all of the effects (parasitic resistance/inductance/capacitance, large-signal performance, nonlinearities, noise, etc.) at such high frequencies, which results in inaccurate simulations.

### 2.2.2 Passive Device

The low resistivity of the substrate ( $\approx 10 \Omega\text{cm}$ ) in CMOS processes creates losses in passive devices such as inductors, transformers, and on-chip antennas [13]. This substrate conductivity causes unwanted coupling. The smaller distance between the metal layers and the substrate causes capacitive coupling and displacement currents. In addition, eddy currents are induced in the substrate, which reduces the quality factor of spiral inductors and other passive devices. Noise in a digital circuit can affect adjacent analog circuits through the substrate. Moreover, substrate contacts on the die can form a return path to ground; thus, the substrate contacts and the return path should be properly modeled in a simulation. Well layers are often employed below the passive devices to reduce the conductivity of the substrate,



At millimeter-wave/sub-terahertz frequencies, the wavelength is on the order of 1 mm; therefore, passive devices occupy a (relatively) small area. An on-chip antenna is an attractive solution for wireless transceivers. However, it is well known that wire impedance increases with frequency because of the skin effect. Passive devices have poor quality factors at high frequencies. In addition, passive devices are more sensitive to variations and mismatches, which increases system errors. For example, the measured performance of a quadrature hybrid may be different from simulation results, and an I/Q phase mismatch can cause a high error rate.

## Part I

# Building Blocks for High-Speed Wireless Communications

# Chapter 3

## Frequency Generation

This chapter discusses challenges, designs, and implementations of frequency generation circuits. In millimeter-wave and sub-terahertz wireless systems, a carrier frequency is required to transmit and receive data over the air. For better system performance, the carrier frequency should be efficiently generated, and its phase noise should be minimized. First, a millimeter-wave voltage-controlled oscillator (VCO) is described to present the issues and the trade-offs in VCO design. Varactor-less VCOs and a 100 GHz active-varactor VCO are then demonstrated to resolve the issues of passive varactors and to reduce power dissipation and phase noise. By employing the active-varactor VCO, a bi-directionally injection-locked loop is also implemented and measured. Moreover, to generate a frequency above cut-off frequencies of CMOS devices, frequency multiplication is discussed. The analysis and design of a frequency doubler, tripler, and quadrupler are presented. In addition, architectures of millimeter-wave and sub-terahertz frequency synthesizers are summarized and discussed. Finally, the LO (local oscillator) distribution is considered.

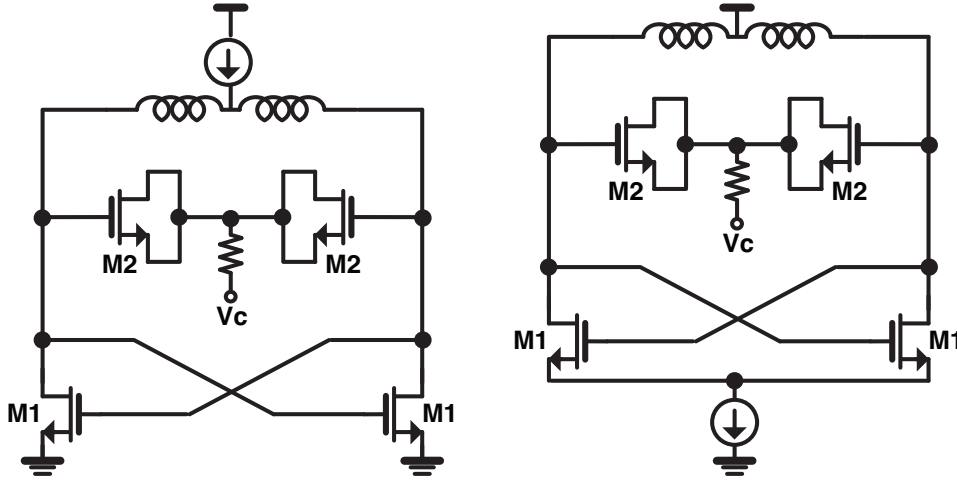
### 3.1 *LC* VCO

The *LC* VCO is widely used in wireless systems because it achieves lower phase noise than other oscillator types, such as ring oscillators, although it requires an inductor, which leads to a large area. Fig. 3.1 shows the schematics of conventional *LC* VCOs in CMOS processes. The oscillation frequency is determined by inductance,  $L_p$ , and capacitance,  $C_p$ , in loss-less parallel *LC* oscillators.

$$\omega = \frac{1}{\sqrt{L_p C_p}} \quad (3.1)$$

However, resistive losses physically exist in passive devices, and the quality factor is defined by

$$Q_p = \frac{R_p}{X_p} \quad (3.2)$$

Figure 3.1: Schematics of  $LC$  VCOs.

where  $R_p$  is the parallel resistance and  $X_p$  is the parallel reactance. Thus, a higher quality factor means lower loss, and a loss-less inductor or capacitor has an infinite quality factor ( $Q = \infty$ ) because  $R_p = \infty$ . The inductor  $Q$  ( $Q_L$ ) and the capacitor  $Q$  ( $Q_C$ ) are, respectively, given by

$$Q_L = \frac{R_{Lp}}{\omega L_p} \quad (3.3)$$

$$Q_C = \omega C_p R_{Cp} \quad (3.4)$$

which lead to

$$\frac{1}{Q_L} + \frac{1}{Q_C} = \frac{\omega L_p}{R_{Lp}} + \frac{1}{\omega C_p R_{Cp}} = \omega L_p \left( \frac{1}{R_{Lp}} + \frac{1}{R_{Cp}} \right) = \frac{\omega L_p}{R_{tank}} = \frac{1}{Q_{tank}} \quad (3.5)$$

$$\frac{1}{Q_L} + \frac{1}{Q_C} = \frac{1}{Q_{tank}} \quad (3.6)$$

where  $R_{tank}$  represents the overall parallel loss of the  $LC$  tank and  $Q_{tank}$  represents the overall quality factor of the tank. The loss of the  $LC$  tank should be compensated by an active circuit to start up and sustain the oscillation. As shown in Fig. 3.1, cross-coupled transistors (M1) create a negative conductance by positive feedback, and if the output resistance is ignored, the parallel conductance is given by

$$G_{p,active} = -\frac{g_{m,M1}}{2} \quad (3.7)$$

The start-up condition is

$$|G_{p,active}| > \frac{1}{R_{tank}} \quad (3.8)$$

which shows

$$\frac{g_{m,M1}}{2} > \frac{1}{R_{tank}} = \frac{1}{\omega L_p Q_{tank}} = \sqrt{\frac{C_p}{L_p}} \cdot \frac{1}{Q_{tank}} \quad (3.9)$$

From Eq. 3.9, the required  $g_m$  is determined by the  $Q_{tank}$  and  $LC$  values ( $C_p$ ,  $L_p$ ). In  $LC$  VCO, power dissipation and power efficiency are strongly related to the  $g_{m,M1}$  of the transistor (M1); therefore,  $Q_{tank}$  should be maximized for the given process, and small  $C_p$  should be chosen to relax the start-up condition and reduce power dissipation. However, the core transistor (M1), varactor (M2), output buffer, and interconnection wires contribute tank capacitance, so reducing  $C_p$  is limited. Therefore, to minimize power dissipation of VCO, the  $LC$  tank should achieve high  $Q_{tank}$  and small parasitic capacitance.

In VCO, the oscillation frequency can be tuned by a control voltage. This frequency tuning is necessary because the center frequency cannot be exactly predicted and simulated. A frequency shift is caused by PVT (process, voltage, and temperature) variations, inaccuracy of device models, and unexpected EM coupling. Particularly at millimeter-wave/sub-terahertz frequencies, it is difficult to capture and simulate all of the effects accurately. Moreover, the VCO is typically integrated with a phase-locked loop (PLL) that requires frequency tunability. In addition, some applications, such as a frequency-modulated continuous wave (FMCW) radar and a VCO-based analog-to-digital converter (ADC), require frequency modulation. For these reasons, the oscillation frequency should be able to be controlled, and the tuning range should be sufficiently wide to meet system specifications. In the CMOS process, MOS capacitors are typically used for fine tuning. Basically, the channel capacitance of MOS transistors can be tuned by changing the gate voltage and the source/drain voltage. The MOS transistor (M2) in Fig. 3.1 acts as variable capacitor or variable reactor, thus called a varactor. The gate voltages are fixed, and the other source/drain voltage is controlled by  $V_C$ . As  $V_C$  increases, the channel disappears, the capacitance decreases, and the VCO frequency increases. Conversely, as  $V_C$  decreases, the channel appears, the capacitance increases, and the VCO frequency decreases.

From Eq. 3.1, the VCO frequency is given by

$$\omega = \frac{1}{\sqrt{L_p(C_{var} + C_{par})}} \quad (3.10)$$

where  $C_{var}$  is capacitance of varactor (M2) and  $C_{par}$  is parasitic capacitance caused by M1, output buffer, and other electric coupling. The varactor capacitance ( $C_{var}$ ) can be changed as

$$C_{min} \leq C_{var} \leq C_{max} \quad (3.11)$$

The VCO frequency is then given by

$$\omega_{min} = \frac{1}{\sqrt{L_p(C_{par} + C_{max})}} \leq \omega \leq \frac{1}{\sqrt{L_p(C_{par} + C_{min})}} = \omega_{max} \quad (3.12)$$

which gives the tuning range as

$$TR = \frac{\omega_{max} - \omega_{min}}{\omega_{center}} = 2 \cdot \frac{\omega_{max} - \omega_{min}}{\omega_{max} + \omega_{min}} = 2 \cdot \frac{\sqrt{\frac{C_{max} + C_{par}}{C_{min} + C_{par}}} - 1}{\sqrt{\frac{C_{max} + C_{par}}{C_{min} + C_{par}}} + 1} < 2 \cdot \frac{\sqrt{\frac{C_{max}}{C_{min}}} - 1}{\sqrt{\frac{C_{max}}{C_{min}}} + 1} \quad (3.13)$$

where  $\omega_{center}$  is the average of  $\omega_{max}$  and  $\omega_{min}$ . From Eq. 3.13, to achieve a wide tuning range,  $C_{par}$  should be minimized, and  $C_{max}/C_{min}$  should be maximized because  $\frac{\sqrt{x-1}}{\sqrt{x+1}}$  is a monotonically increasing function.

One important parameter of VCO is  $K_{VCO}$ , the slope of the tuning curve, which is defined by

$$K_{VCO} = \left. \frac{d\omega(V)}{dV} \right|_{\omega=\omega_{center}} \quad (3.14)$$

At the center frequency,  $\omega(V)$  can be modeled as a linear function of a tuning voltage and the slope is  $K_{VCO}$ . This  $K_{VCO}$  value is important in some applications, such as the phase-locked loop and the FMCW radar. If  $K_{VCO}$  is too large, the phase noise is more affected by noise from the  $V_C$  node. If  $K_{VCO}$  is too small, the frequency range which can be changed by  $V_C$  is narrow. In the phase-locked loop, for example,  $K_{VCO}$  appears in the dynamic equations, hence it can affect the loop stability and the settling time. Typically, it is hard to achieve a linear tuning curve over the entire range of  $V_C$  because varactor capacitance does not linearly change. This non-linear behavior may affect the system performance. For instance, the PLL is locked in a partial range of  $V_C$  and the frequency modulation of the FMCW radar may be degraded.

Finally, one of the most important VCO parameters is phase noise. Phase noise of LO frequency degrades the error rate of the communication system and accuracy of the radar/imager system. Typically, the VCO phase noise can be high-pass-filtered by a PLL; thus, phase noise values at offset frequencies between 1 MHz and 1 GHz are of interest, depending on system specifications. According to Leeson's analysis [14], the phase noise is given by

$$\mathcal{L}(\Delta\omega) = 10 \log \left[ \frac{2kT}{P_S} \left( \frac{\omega_o}{2Q_{tank}\Delta\omega} \right)^2 \right] \quad (3.15)$$

where  $k$  is the Boltzmann constant,  $T$  is the temperature,  $P_S$  is signal power,  $\omega_o$  is oscillation frequency,  $Q_{tank}$  is the quality factor of  $LC$  tank, and  $\Delta\omega$  is an offset frequency. In Eq. 3.15, if a center frequency and temperature are given, then two variables,  $P_S$  and  $Q_{tank}$ , affect the phase noise. From this equation,  $P_S$  and  $Q_{tank}$  both should be made larger to reduce the phase noise.

VCO has important factors, such as power efficiency, tuning range, and phase noise. Based on the above analyses, increasing  $Q_{tank}$  improves the power efficiency and phase noise, but not the tuning range because  $C_{par}$  increases  $Q_C$  but degrades the tuning range. Thus, it is critical to improve both the  $Q_{tank}$  and the tuning range. However,  $Q_L$  and  $Q_C$  are strongly dependent of oscillation frequency and the process technology, and it is more

challenging to achieve high quality factors at millimeter-wave/sub-terahertz frequencies. In addition, inductance and capacitance values are small (50 pH and 50 pF gives 100 GHz) and the tuning range is more sensitive to parasitic capacitance. Therefore, novel designs and techniques are needed, and varactor-less VCO and active-varactor VCO will be demonstrated in this chapter.

### 3.2 65 GHz LC VCO

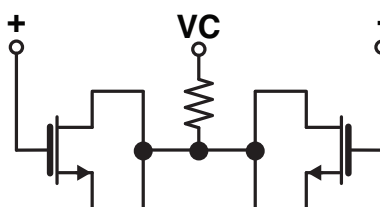


Figure 3.2: Schematic of a passive varactor.

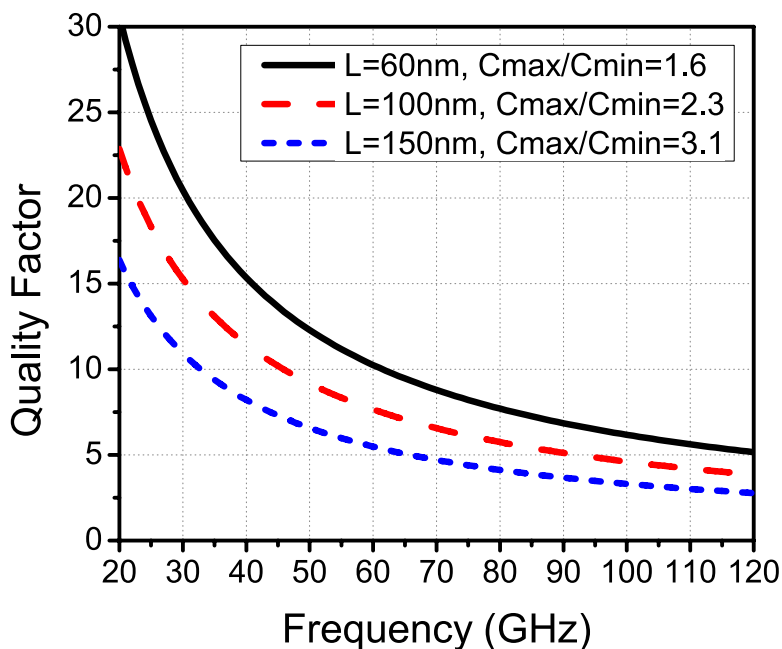


Figure 3.3: Quality factors of passive varactors (worst-case, from extraction and post-layout simulation).

From the previous section,  $Q_{tank}$  determines the power efficiency and phase noise. Thus, both  $Q_L$  and  $Q_C$  should be kept high. First, inductors can be designed with a spiral inductor

or a transmission line. In the CMOS process, unless a special process option is employed, the quality factors of inductors are typically between 10 and 20 at millimeter-wave frequencies. However,  $Q_C$  shows quite a different behavior. The passive varactor, which is widely employed in VCOs, is shown in Fig. 3.2 and the varactor  $Q$  is given by

$$Q_{var} = \frac{1}{\omega R_S C_{var}} \quad (3.16)$$

where  $R_S$  is the series loss in the varactor. The series loss is used in this equation because the varactor loss is dominated by gate resistance and channel resistance that are in series with the varactor. Moreover, the passive varactors are laid out, extracted, and simulated with 65 nm CMOS process. Fig. 3.3 shows that the varactor  $Q$  decreases with the frequency, as expected from Eq. 3.16. If the parasitic resistance and inductance are taken into consideration, the quality factor can be much lower. When the frequency is lower than approximately 30 GHz,  $Q_{var}$  is higher than  $Q_L$ . On the contrary,  $Q_{var}$  is lower than  $Q_L$  at frequencies above 60 GHz because  $Q_{var}$  is inversely proportional to  $\omega$  but  $Q_L$  does not vary much with frequency. In addition, Fig. 3.3 also demonstrates a trade-off between  $Q_{var}$  and the  $C_{max}/C_{min}$  ratio. As the channel length ( $L$ ) increases,  $Q_{var}$  decreases, but  $C_{max}/C_{min}$  increases. First, a channel length of 150 nm is chosen to achieve a target tuning range of 10%, then at frequencies above 60 GHz,  $Q_{var}$  is less than 6, which is much lower than typical  $Q_L$ . Thus,  $Q_{tank}$  is dominated by the varactor  $Q$ .

$$\frac{1}{Q_{tank}} = \frac{1}{Q_C} + \frac{1}{Q_L} \approx \frac{1}{Q_{var}} \quad (Q_C \approx Q_{var} \ll Q_L) \quad (3.17)$$

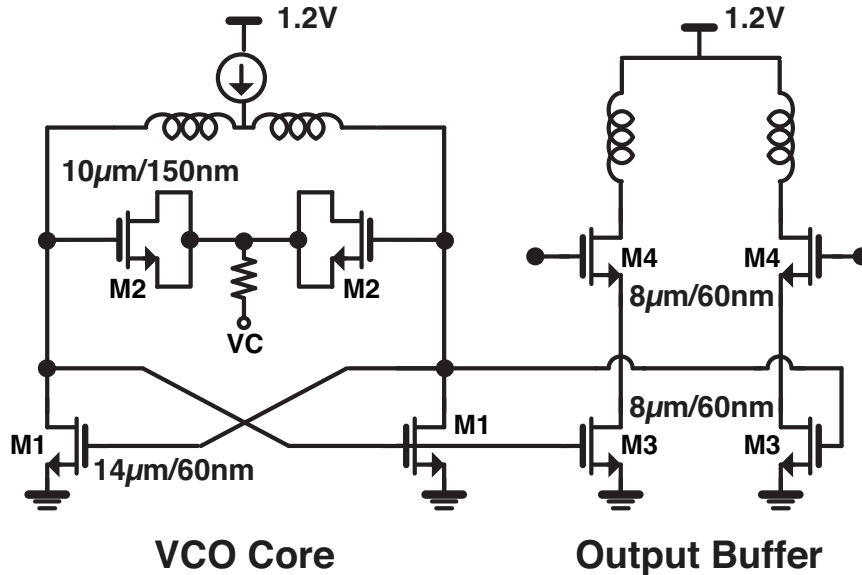
As a result,  $Q_{tank}$  is approximately  $Q_{var}$  in a conventional millimeter-wave  $LC$  VCO.

In low-frequency (< 30 GHz) VCOs, switched capacitors are widely used for coarse frequency tuning. However, switched capacitors are not used in this VCO because the series resistive loss and parasitic capacitance are too high to be used in millimeter-wave VCOs in 65 nm CMOS.

With the passive varactor, a 65 GHz VCO is designed in 65 nm CMOS. The schematic is shown in Fig. 3.4. In addition to the VCO core, the output buffer is also shown. Although a low-noise and efficient VCO is achieved, it is useless without an output buffer. Because the output buffer loads the VCO core, the tuning range is degraded, and  $Q_{tank}$  can also be degraded depending on the quality factor of the input impedance of the buffer. Therefore, the buffer design should be simultaneously performed with the core design.

In the VCO core, an NMOS cross-coupled pair is used to create negative conductance. Its channel length is minimum and its width is 14  $\mu\text{m}$ . The supply voltage is 1.2 V and a PMOS device is used for the current source at the top. A spiral inductor is used for the tank inductor, and the inductance is approximately 150 pH and  $Q_L$  is 12 at 65 GHz. The passive varactor is used for frequency tuning, and a channel length of 150 nm and a width of 10  $\mu\text{m}$  are chosen for the target tuning range of 10%. In the buffer design, a cascode amplifier is employed to reduce the Miller effect in common source amplifiers and to reduce the load-pulling effect. The device size is roughly half the core NMOS size. All of the parameters are



Figure 3.4: Schematic of the 65 GHz *LC* VCO.

tuned and optimized in simulation to achieve a tuning range of 10% and an output power of 1 dBm. Fig. 3.5 shows the simulation results of the VCO. The output frequency ranges from 61.1 to 67.0 GHz and the tuning range is 9.2%. The worst-case output power is 0.9 dBm at 67 GHz. The phase noise at 1 MHz offset is less than  $-87$  dBc/Hz. The DC power consumed by the VCO core and the output buffer is approximately 30 mW.

The 65 GHz VCO is fabricated in 65 nm CMOS with a 260 GHz OOK transceiver, which is demonstrated in Chapter 7. Fig. 3.6(a) shows the microphotograph of the VCO, and the size is  $130 \mu\text{m} \times 130 \mu\text{m}$ , dominated by the spiral inductor. The locations of the VCO core, the output buffer, and the current source are marked. The output frequency can be determined by measuring a leakage tone radiated by matching transformers of the output amplifier chain. The measured frequency ranges from 61.3 to 66.1 GHz (7.54%), as shown in Fig. 3.6(b).

### 3.3 Varactor-less *LC* VCO

Because a passive varactor degrades the VCO performance, as explained in the previous section, designing a VCO without a passive varactor (variable capacitor) is attractive and has been proposed by several research groups [15]–[23], which is called varactor-less VCO. Removing the passive varactor increases the  $Q_{\text{tank}}$  and breaks the stringent trade-off between the  $Q_{\text{tank}}$  and the tuning range. However, a varactor-less VCO still has some different trade-offs among efficiency, tuning range, and phase noise, depending on the tuning method.

There are several methods to tune the oscillation frequency without a varactor. A trivial

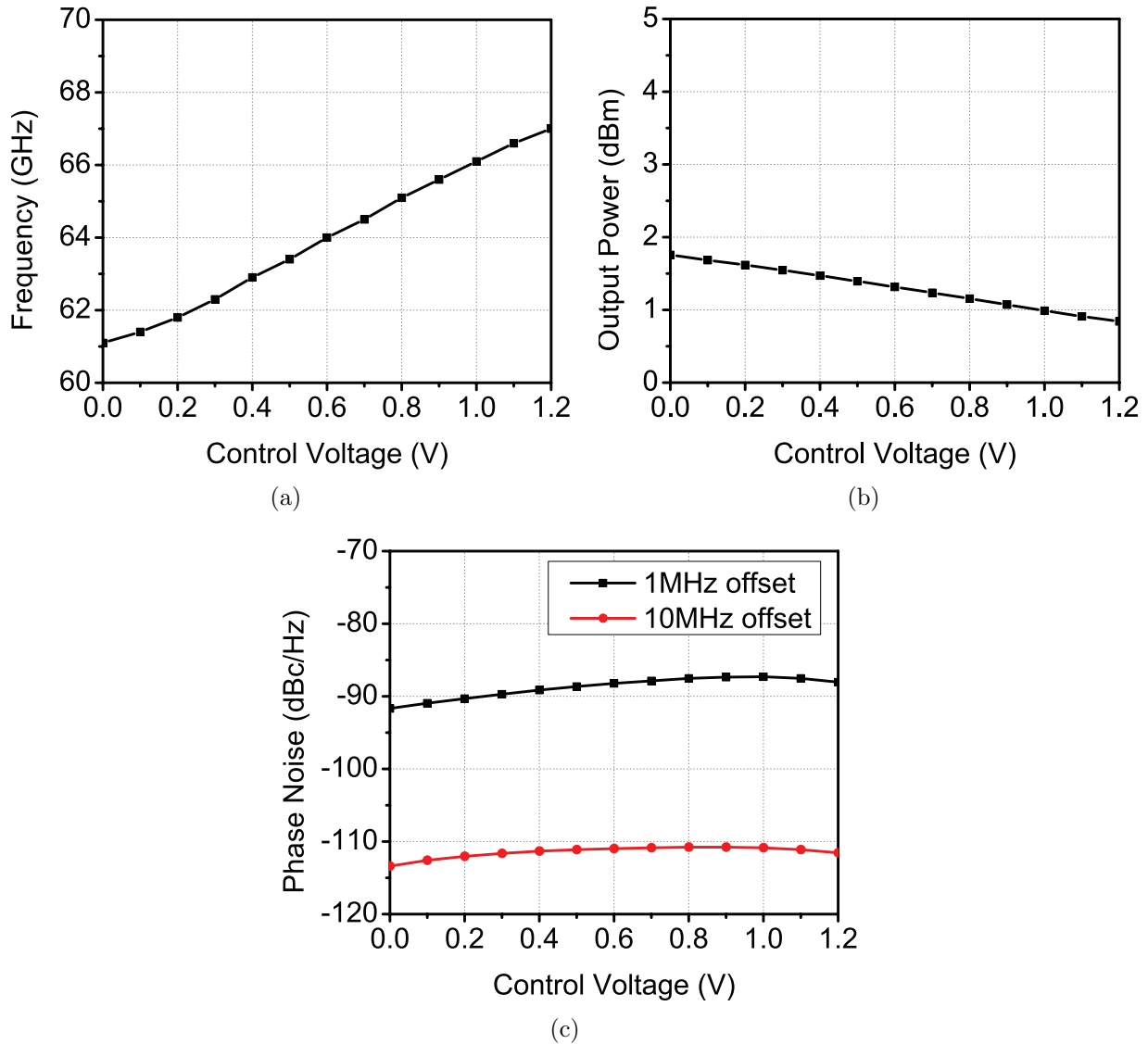


Figure 3.5: Simulation results of the 65 GHz VCO (a) frequency (b) output power (c) phase noise.

way is to change the DC bias conditions. If bias voltages and currents are changed, device capacitances are accordingly changed, such that the frequency can be tuned. Typically, the bias current can be digitally controlled by a current source; therefore, most oscillators use this method to adjust the biasing point or frequency range. Although the oscillation frequency is properly tuned and the tuning range is wide, this technique can significantly change the device operating points, hence the output power, output impedance, and phase noise are notably varied with the tuning voltage. Most likely, the start-up condition may not be satisfied when devices are turned into triode or cut-off mode.

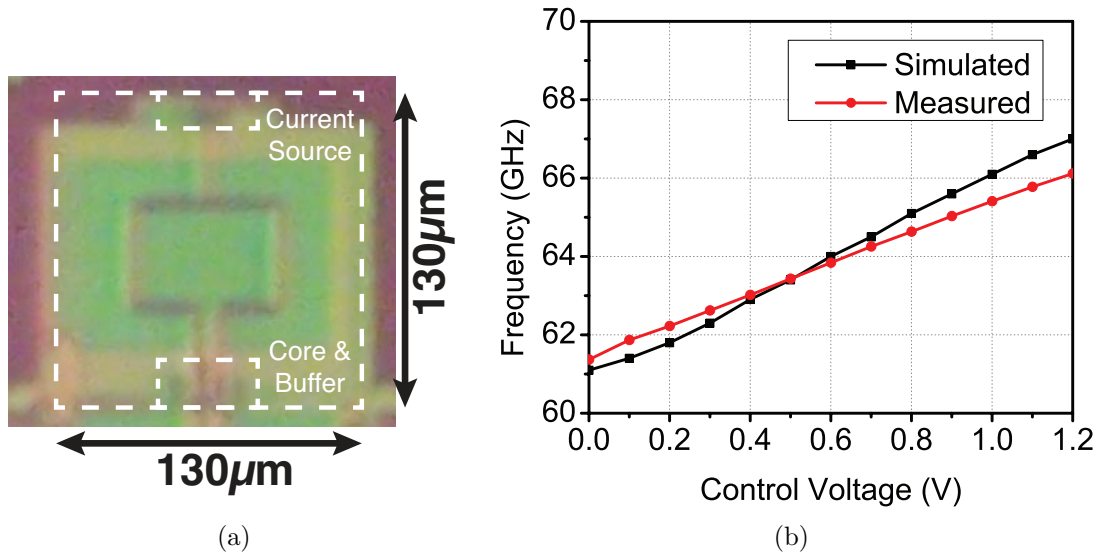


Figure 3.6: 65 GHz VCO (a) microphotograph (b) measured frequency.

The next method is to tune the inductance, not the capacitance. As switched capacitors are used for coarse tuning in low-frequency oscillators, switched inductors can be used. As demonstrated in [15]–[17], the inductor is designed together with MOS switches, and the inductance can be changed by controlling switches inside the inductor. At millimeter-wave frequencies, the MOS switch is not close to an ideal switch, and it has high resistive loss and parasitic capacitance; therefore, it is challenging to achieve high  $Q_L$  and wide tuning range.

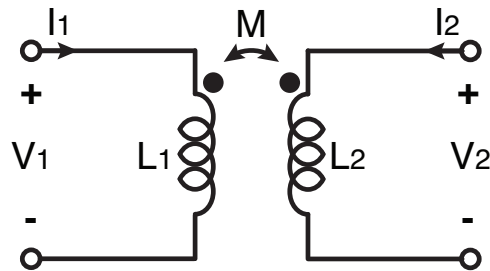


Figure 3.7: A loss-less transformer.

At such high frequencies, using a transformer instead of a MOS switch is presented in [18]–[20]. This method uses a characteristic property of transformer. A loss-less transformer is illustrated in Fig. 3.7 and its Z-parameters are given by

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} j\omega L_1 & j\omega M \\ j\omega M & j\omega L_2 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} \quad (3.18)$$

Here, if  $I_1$  and  $I_2$  have a relationship as

$$I_2 = \alpha I_1 \quad (3.19)$$

$$-\alpha_M \leq \alpha \leq \alpha_M \quad (3.20)$$

where  $\alpha$  is assumed to be a real number for simplicity and  $\alpha_M$  is the maximum. The input impedance seen from the primary side is

$$V_1 = j\omega(L_1 + \alpha M)I_1 \quad (3.21)$$

$$\frac{V_1}{I_1} = j\omega(L_1 + \alpha M) \quad (3.22)$$

Thus, the effective inductance of the primary side is  $L_1 + \alpha M$ , which can be tuned by changing  $\alpha = I_2/I_1$ . The tuning range is given by

$$TR = 2 \cdot \frac{\frac{1}{\sqrt{(L_1 - \alpha_M M)C}} - \frac{1}{\sqrt{(L_1 + \alpha_M M)C}}}{\frac{1}{\sqrt{(L_1 - \alpha_M M)C}} + \frac{1}{\sqrt{(L_1 + \alpha_M M)C}}} \quad (3.23)$$

where  $C$  is tank capacitance ( $\omega = 1/\sqrt{L_1 C}$ ). If  $L_1 \gg \alpha_M M$  is assumed, then the tuning range becomes

$$TR \approx \frac{\alpha_M M}{L_1} \quad (3.24)$$

From Eq. 3.24, to increase the tuning range, the coupling coefficient of the transformer and the  $I_2/I_1$  ratio should be maximized. The relationship (Eq. 3.19) can be implemented by a buffer, which takes  $V_1$  or  $I_1$  and drives the secondary side. The detailed implementations are found in [18]–[20]. [18] achieves a tuning range of 24% at 26 GHz, [19] achieves 67% at 5 GHz, and [20] achieves 24% at 3 GHz.

The last method uses the inductive degeneration [21], [22]. With a degeneration inductor ( $L_{deg}$ ), the input impedance is a function of the  $g_m$  of the transistor and  $L_{deg}$ . By changing the  $g_m$  by bias voltages or currents, the input impedance (both real and imaginary parts) changes; therefore, the oscillation frequency changes without a varactor. [21] achieves a tuning range of 31% at 4.5 GHz and [22] achieves 3% at a fundamental frequency of 100 GHz.

### 3.4 100 GHz Active-Varactor VCO

This section proposes and demonstrates another novel architecture, active-varactor VCO, published in [24]. The active-varactor VCO does not use a passive varactor, similar to the varactor-less VCO, but it employs an active varactor (variable capacitor). Whereas the passive varactor is lossy, the active varactor provides power gain, as well as capacitive frequency tuning.

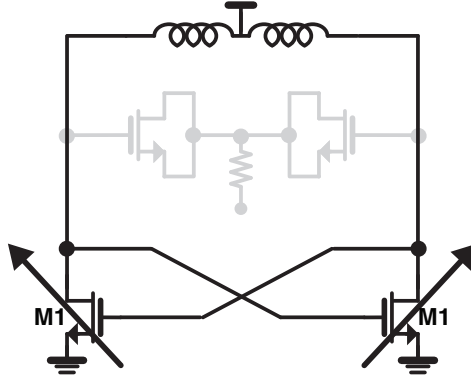


Figure 3.8: Concept of the active-varactor VCO.

Fig. 3.8 illustrates the concept of the active-varactor VCO. The passive varactor is removed, and the device size can then be changed by turning some devices on or off. First, let the minimum device size that can sustain oscillation be  $W_{min}$  and the corresponding tank capacitance be  $C_{min}$ . The oscillation frequency is then given by

$$\omega_{max} = \frac{1}{\sqrt{LC_{min}}} \quad (3.25)$$

Next, increase the device size to infinity (not possible in practice), and the corresponding tank capacitance ( $C_{max}$ ) should be infinity and the oscillation frequency is zero.

$$\omega_{min} = \frac{1}{\sqrt{LC_{max}}} \approx 0 \quad (3.26)$$

The tuning range is

$$TR = 2 \cdot \frac{\omega_{max} - \omega_{min}}{\omega_{max} + \omega_{min}} = 2 \cdot \frac{\omega_{max} - 0}{\omega_{max} + 0} = 2 \quad (3.27)$$

which shows that the active-varactor VCO can achieve a tuning range of 200 % ideally. In real circuits, it is impossible to realize such a wide tuning range because the device size is finite and  $C_{max}/C_{min}$  ratio is limited by parasitic capacitance. The beauty of this method is that the lossy component is only an inductor; therefore,  $Q$  is roughly  $10 \sim 20$  in a typical CMOS process at millimeter-wave/sub-terahertz frequencies. The frequency tuning is determined by active devices, which have negative conductance and do not degrade  $Q_{tank}$ .

Fig. 3.9 shows the real circuit to explain the concept. The gates of M1 and M2 are AC-coupled, and the gate voltage (VC) of M2 can be changed. The VC tunes the capacitance of M2, and it can change the negative conductance and the oscillator frequency at the same time. Two extreme cases are illustrated in Fig. 3.9.

In Fig. 3.9(a), VC is low and M2 devices are completely turned off, such that the capacitance is  $C = C_1 + C_{2,off}$ , and the frequency is  $\omega = \frac{1}{\sqrt{L(C_1 + C_{2,off})}}$ . The conductance is  $g_m = g_{m1}$  and the DC current is  $I = I_1$ . On the contrary, in Fig. 3.9(b), VC is high and

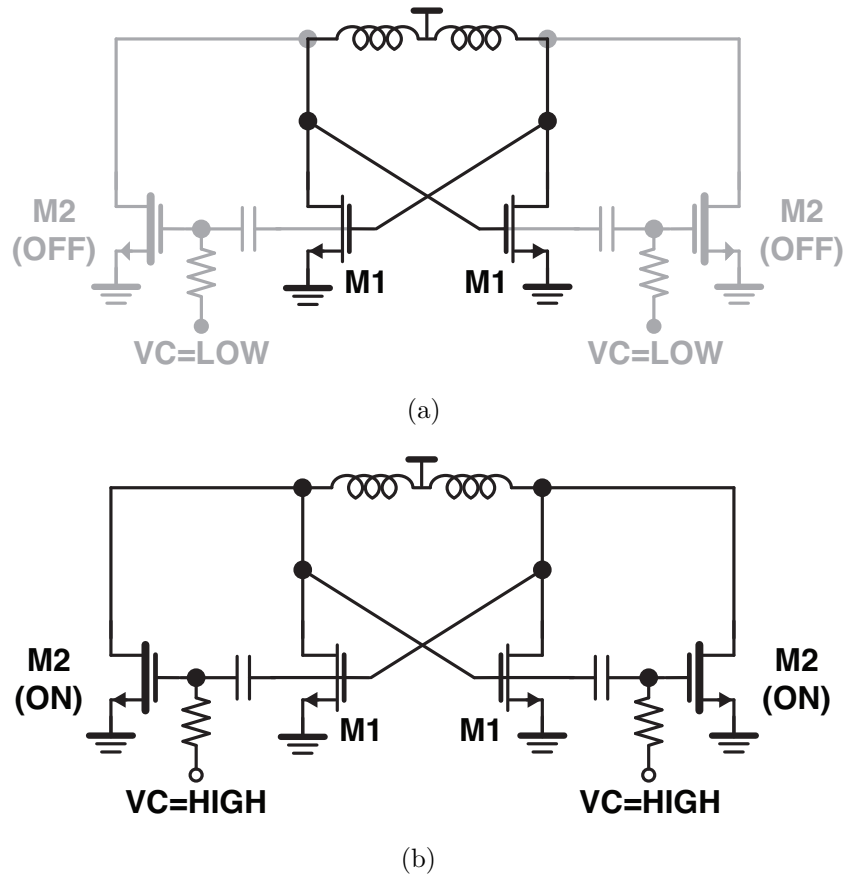


Figure 3.9: Schematics of the proposed active-varactor VCO (a) fully off (b) fully on.

M2 devices are fully turned on, such that the capacitance is increased to  $C = C_1 + C_2$ , and the frequency is lowered to  $\omega = \frac{1}{\sqrt{L(C_1+C_2)}}$ .  $g_m$  is increased to  $g_{m1} + g_{m2}$ . Accordingly, the DC current is also increased to  $I = I_1 + I_2$ . Thus, if VC is tuned continuously, the oscillator frequency is tuned with a conventional passive varactor.

Regarding the phase noise, when M2 is off, the oscillator operates in the current-limited regime. When M2 is on, the oscillator operates in the voltage-limited regime. However, the phase noise variation over the tuning range is not significant ( $< 5$  dB).

A 100 GHz fundamental active-varactor VCO is schematically shown in Fig. 3.10. Two active-varactor oscillators are coupled by a transformer. Each oscillator has two cross-coupled pairs (a main pair and an auxiliary pair), and the control voltages (VC1 and VC2) change the gate voltages of the auxiliary pairs (M3 and M4), respectively. Without a conventional passive varactor, the oscillator frequency can be changed by tuning the control voltage. Increasing the control voltage increases the device capacitance, hence the frequency decreases. The auxiliary pair has the negative conductance; thus, the overall quality factor ( $Q_{tank}$ ) is the quality factor of the inductor ( $Q_L$ ), which is higher than the quality factor of a passive

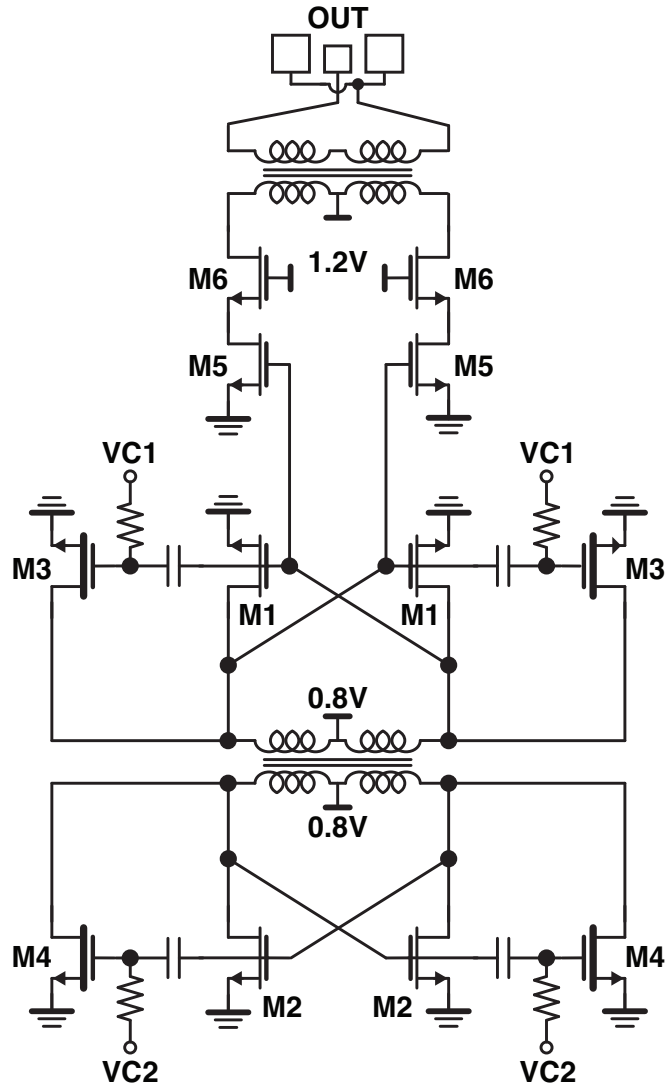


Figure 3.10: Schematic of the 100 GHz active-varactor VCO.

varactor at 100 GHz, leading to higher power efficiency and lower phase noise. Table 3.1 presents the sizes and types of the NMOS devices (M1~M6). The standard- $V_{th}$  devices are used for M3 and M4 to adjust the center of the frequency tuning curve.

In addition, a transformer is employed to couple two oscillators. In this case, using a transformer has several advantages. First, by using two different control voltages (VC1 and VC2) for the primary side and the secondary side, the currents of the two sides are independently controlled; thus, the inductive frequency tuning happens as well as the capacitive tuning. Second, the magnetic flux of a transformer is generally lower than that of a single inductor due to two windings; thus, any unwanted magnetic coupling can be reduced. Lastly, it enables bi-directional injection locking, which will be discussed in the next section.

The cascode amplifier (M5, M6) is an output buffer, and its output is matched to  $50\ \Omega$  using another transformer. The output buffer is only in the primary side; therefore, the sizes of M3 and M4 are designed to be different.

Table 3.1: Circuit parameters of the 100 GHz active-varactor VCO.

Main pair	M1	6 $\mu\text{m}/60\ \text{nm}$	Low $V_{th}$
	M2	6 $\mu\text{m}/60\ \text{nm}$	Low $V_{th}$
Auxiliary pair	M3	8 $\mu\text{m}/60\ \text{nm}$	Standard $V_{th}$
	M4	12 $\mu\text{m}/60\ \text{nm}$	Standard $V_{th}$
Output buffer	M5	4 $\mu\text{m}/60\ \text{nm}$	Low $V_{th}$
	M6	4 $\mu\text{m}/60\ \text{nm}$	Low $V_{th}$

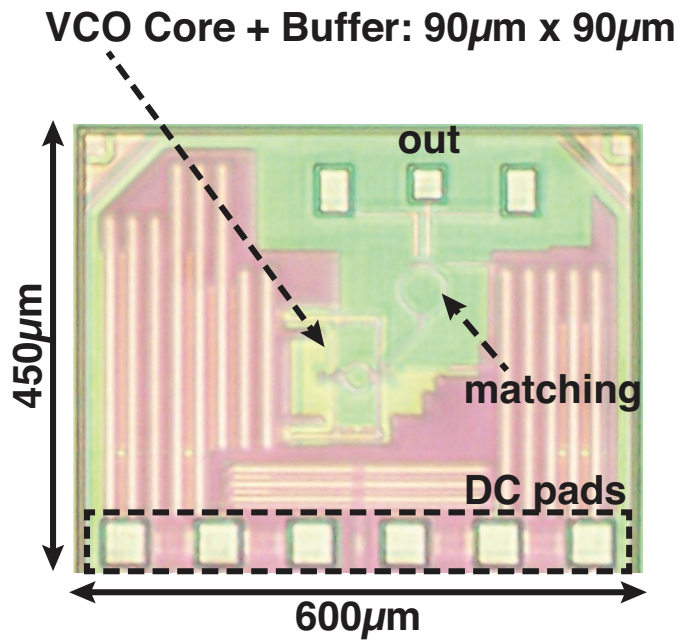


Figure 3.11: Microphotograph of the 100 GHz active-varactor VCO.

The 100 GHz active-varactor VCO is fabricated in 65 nm bulk CMOS. The die photo is shown in Fig. 3.11. The actual area of the VCO tank core and its output buffer is  $90\ \mu\text{m} \times 45\ \mu\text{m}$ . As illustrated in Fig. 3.12, a W-band probe, a DC probe, a down-converter, a spectrum analyzer and a power meter were used to perform the measurements. The measurement results are shown in Fig. 3.13. By sweeping two control voltages independently, the output frequency, output power, phase noise, and power dissipation are measured. VC1 and VC2 are asymmetric because the output buffer is only connected to the primary side, and the sizes of auxiliary pairs are designed to be different. Table 3.2 summarizes the performance and compares with other VCOs.



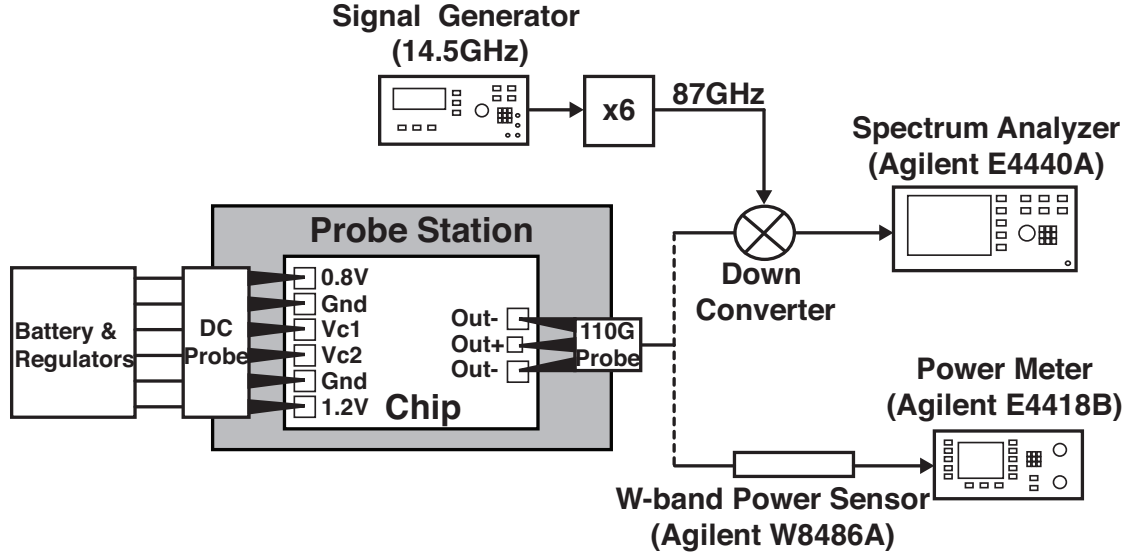
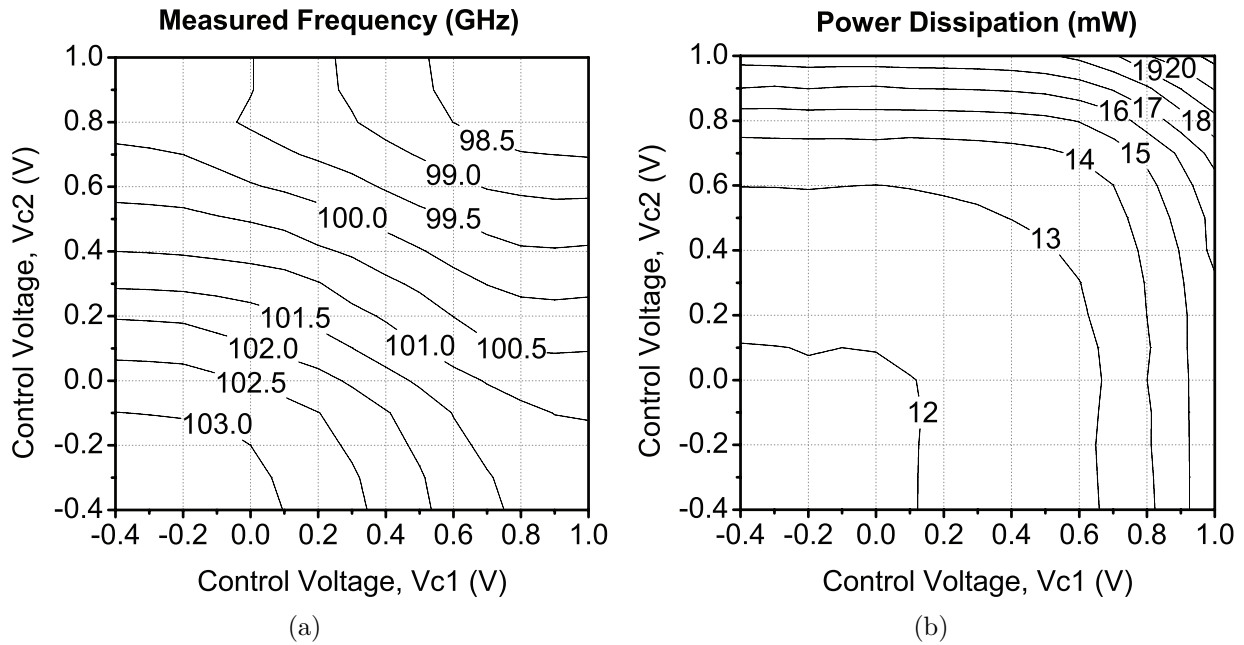


Figure 3.12: Measurement setup.

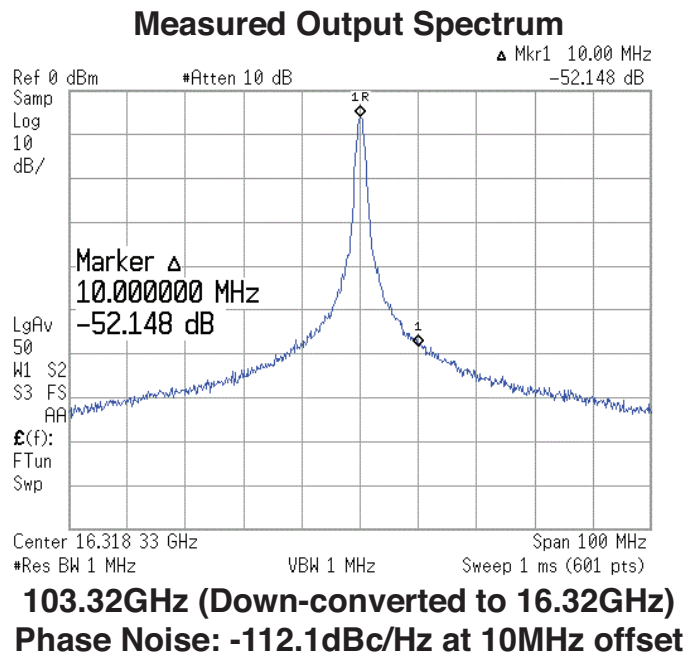
Table 3.2: VCO performance summary and comparison

	This work [24]	Tsai [25]	Laskin [26] <sup>(3)</sup>	Volkaerts [27]
Technology	65 nm CMOS	90 nm CMOS	65 nm CMOS	65 nm CMOS
Frequency (GHz)	98 ~ 103.3	90 ~ 92.5	88.3 ~ 91.3	113.5 ~ 122.5
Tuning Range (TR) (%)	5.2	2.74	3.34	7.8
Phase Noise @ 10 MHz (dBc/Hz)	-112.1	-107	-115 <sup>(4)</sup>	-104 <sup>(4)</sup>
Differential Output Power (dBm)	-5 ~ -2	-12 ~ -20	-4 ~ 3	-28 ~ -14
Power Consumption (mW)	12 ~ 21	14 ~ 87.2	57.6 <sup>(5)</sup>	5.6 <sup>(5)</sup>
Supply Voltage (V)	0.8, 1.2	0.6 ~ 1.45	1.2	1
Control Voltage (V)	-0.4 ~ 1.0	0.6 ~ 1.45	-1.2 ~ 1.8	0 ~ 2
Core Area (mm <sup>2</sup> )	90×45	620×550	150×170	105×65
$FoM^{(1)}$ (dBc/Hz)	178.6 <sup>(1)</sup>	155.8	176.5	156.9
$FoM_T^{(2)}$ (dBc/Hz)	172.9 <sup>(1)</sup>	144.5	167.0	154.7
<p>(1) <math>FoM = \left(\frac{f_{osc}}{f_{off}}\right)^2 \cdot \frac{1}{L\{f_{off}\}} \cdot \frac{P_{out}}{P_{diss}}</math></p> <p>(2) <math>FoM_T = \left(\frac{f_{osc}}{f_{off}}\right)^2 \cdot \frac{1}{L\{f_{off}\}} \cdot \frac{P_{out}}{P_{diss}} \cdot \left(\frac{TR(\%)}{10}\right)^2</math></p> <p>(3) Quadrature VCO</p> <p>(4) Estimated from 1 MHz-offset phase noise</p> <p>(5) Buffer power is not included.</p>				



(a)

(b)



(c)

Figure 3.13: Measurement results of the 100 GHz active-varactor VCO (a) frequency (b) power dissipation (c) output spectrum.

### 3.5 Injection-Locked Loop

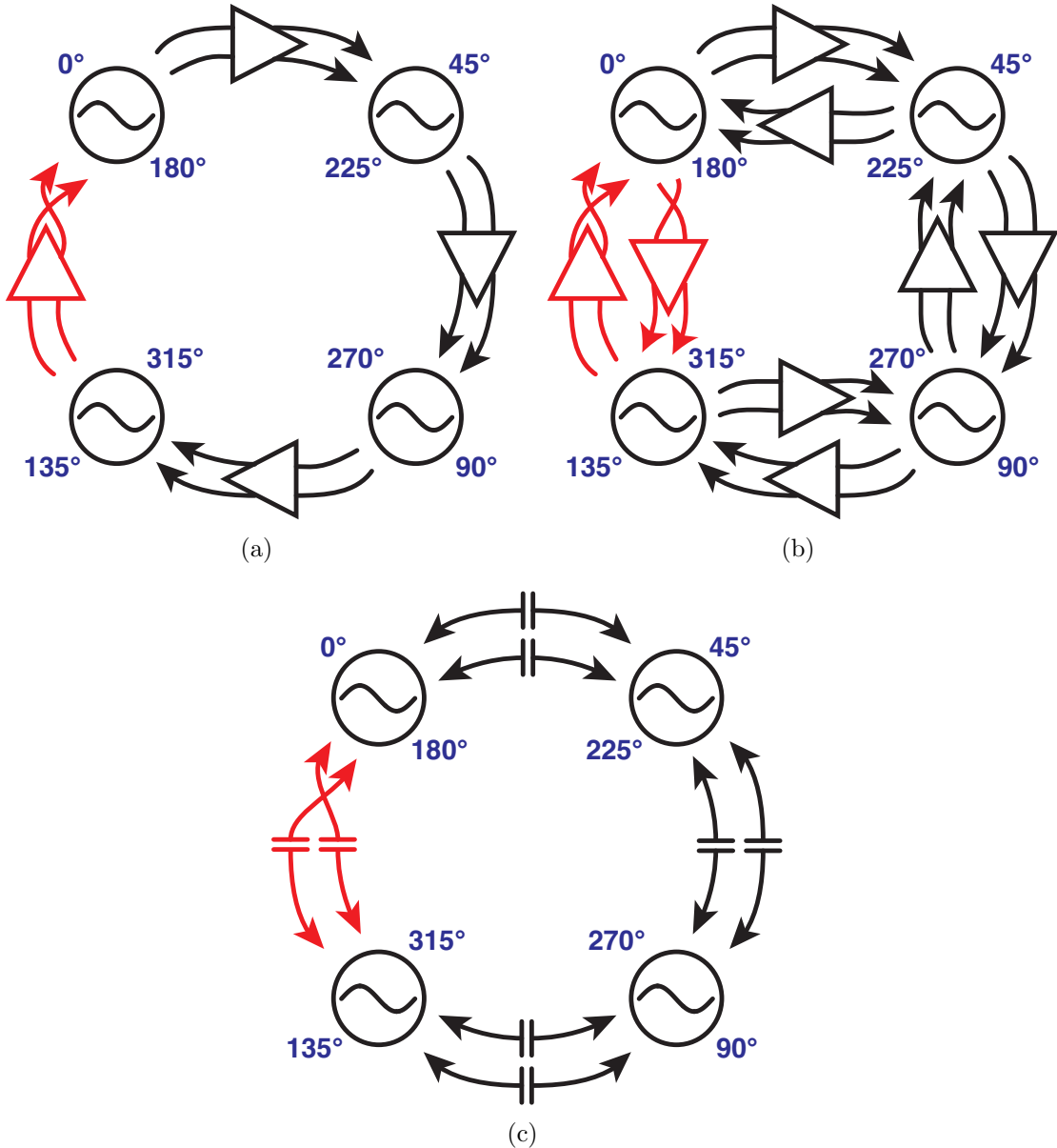


Figure 3.14: Architectures of mutual coupling (a) uni-directional coupling (b) bi-directional coupling using buffers (c) bi-directional capacitive coupling (this work).

This section proposes and demonstrates a capacitively coupled bi-directionally injection-locked loop, published in [24]. It is well known that if multiple oscillators are coupled properly, their phase noise can be reduced [28]. In addition, multiple phases can be created

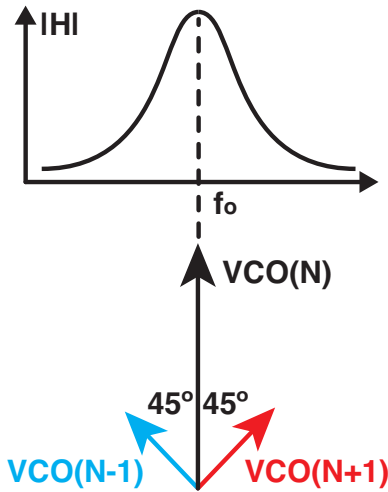


Figure 3.15: Phase diagram of injection locking in the bi-directional coupling.

and used for many applications ( $N$ -push frequency multiplication, I/Q up/down-conversion, and phase rotating/interpolating). Many oscillator outputs can enable various system architectures and give higher output power if they are combined. The drawbacks of this design are larger area, higher power dissipation, and more complicated routing. However, each oscillator size is relatively small at 100 GHz, and DC current can be reduced by using small-size transistors and the active-varactor tuning scheme proposed in the previous section. Furthermore, the proposed simple capacitive coupling solves the complex routing issues associated with locking the VCOs.

Many coupling topologies have been proposed such as cyclic or non-cyclic and uni-directional or bi-directional [28]. Fig. 3.14 shows three architectures to couple four oscillators in the cyclic loop. One connection (marked in red) is cross-coupled for out-of-phase locking to generate eight phases ( $45^\circ$ ). Fig. 3.14(a) shows a uni-directional coupling and Fig. 3.14(b) shows a bi-directional coupling using buffers. The proposed coupling architecture is a capacitively coupled bi-directionally injection-locked loop, as presented in Fig. 3.14(c).

First, the bi-directional injection-locking is chosen for the following reasons. The oscillator signal and two injected signals from adjacent oscillators are vector-summed as shown in Fig. 3.15, and because the two injected signals have opposite phases ( $\pm 45^\circ$ ), the direction of the summed signal is ideally the same as that of the oscillator signal. Thus, there is no frequency shift, and the oscillators keep operating at the  $LC$  resonance point, which is the point of the maximum swing and gain. The beauty of the bi-directional coupling is its use with the active-varactor VCO because the quality factor is high and the tank impedance is higher at the resonance frequency. The uni-directional coupling has a frequency shift, mitigating the merits achieved from high  $Q_{tank}$  and high impedance. Therefore, the bi-directional injection-locking fits well with the active-varactor VCO. The full schematic is shown in Fig. 3.16.

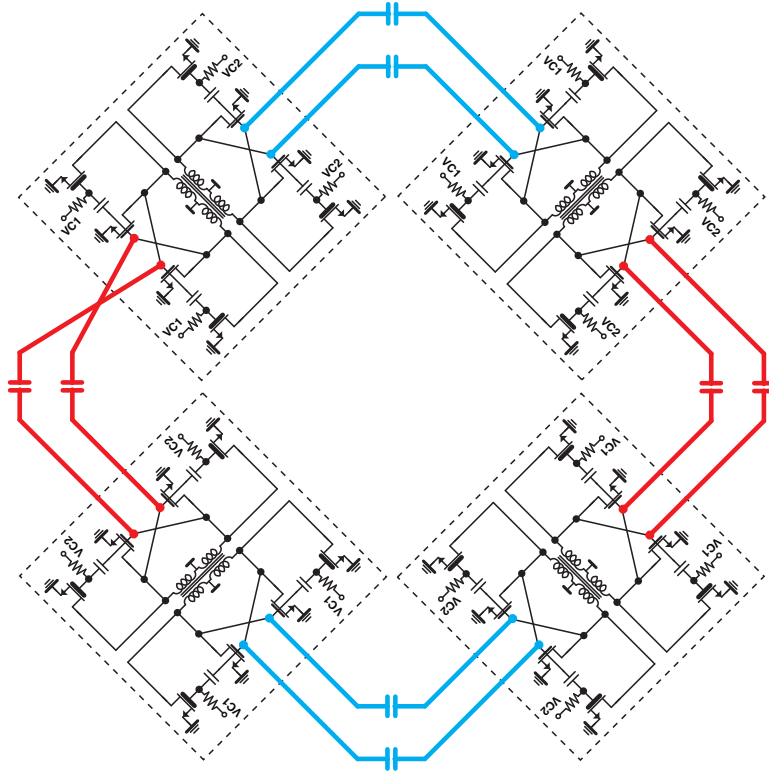


Figure 3.16: Schematic of the injection-locked loop.

Second, as shown in Fig. 3.14(b), typical bi-directional coupling requires eight buffers, causing more power dissipation. To remove the buffers, the transmission-line-based capacitive coupling is proposed, as shown in Fig. 3.14(c) and Fig. 3.17. Oscillators are coupled and injection-locked simply through a capacitor. The capacitor itself is bi-directional, noise-less, and does not dissipate power. Similar techniques were reported at frequencies lower than 11 GHz, but applied to only QVCO ( $90^\circ$ ) [29].

Third, a problem with reported capacitive coupling in three or more oscillators is that all of the oscillators are connected through capacitors; therefore, an oscillator signal may affect all of the other oscillators, potentially negating the effect of bi-directional coupling. Therefore, the transformer of the VCO plays a role to enable and ensure proper locking, because its coupling factor ( $\approx 0.65$ ) gives some attenuation between the two sides, preventing undesired coupling, as illustrated in Fig. 3.18(a). The oscillators are not in phase in this design; therefore, the capacitor is seen by the oscillator, and the capacitor lowers the frequency. Thus, the coupling capacitance should be minimized to reduce the frequency shift. However, if the coupling capacitance is too small, the injection-locking is so weak that the oscillators cannot be locked properly. In this design, 5 fF is selected to balance this trade-off.

The layout of the injection-locked loop is illustrated in Fig. 3.17. Four VCOs and four transformers are placed at the corners of a square. Coplanar striplines are extended from

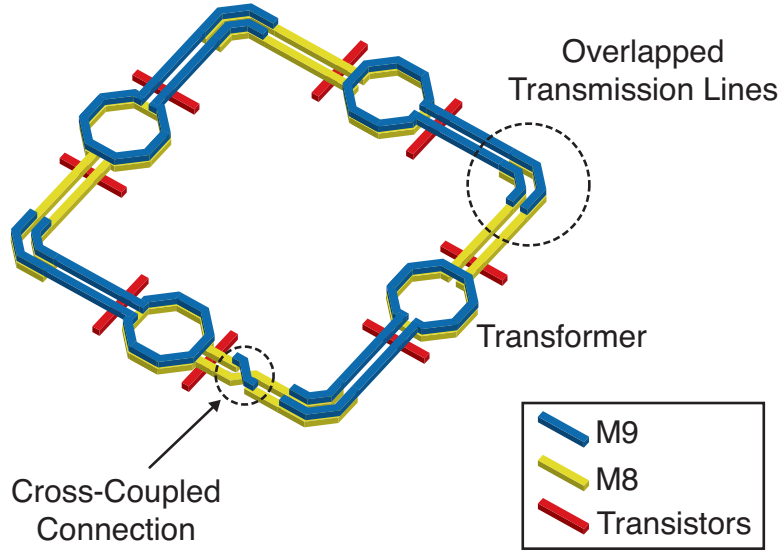


Figure 3.17: Layout of the injection-locked loop.

the transformers. M8 is extended from the primary side of an oscillator, and M9 is extended from the secondary side of the other oscillator. They are overlapped in the middle over a length of  $14\ \mu\text{m}$ , while the length of each extended line is approximately  $70\ \mu\text{m}$ . Effectively the total reactance is approximately  $5\ \text{fF}$ , and the electrical delay is adjusted to be  $360^\circ$ . The input phase and the output phase of the coupling path (M8 lines, overlapped regions, and the M9 lines) should be the same to enable bi-directional injection-locking, as shown in Fig. 3.18(b).

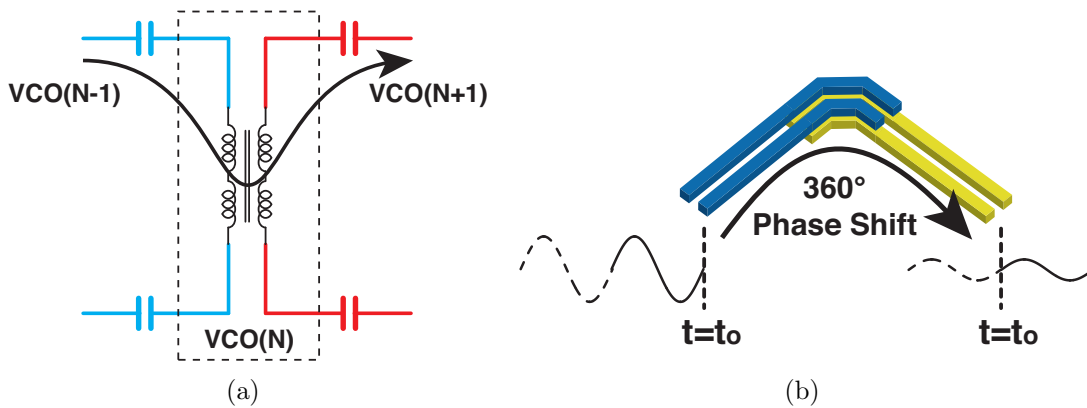


Figure 3.18: Design of passive devices (a) transformer (b) transmission-line-based capacitive coupling.

The injection-locked loop is fabricated in  $65\ \text{nm}$  CMOS. The die photo is shown in Fig. 3.19. The actual area of the injection-locked VCO loop is  $320\ \mu\text{m} \times 320\ \mu\text{m}$ . A W-band

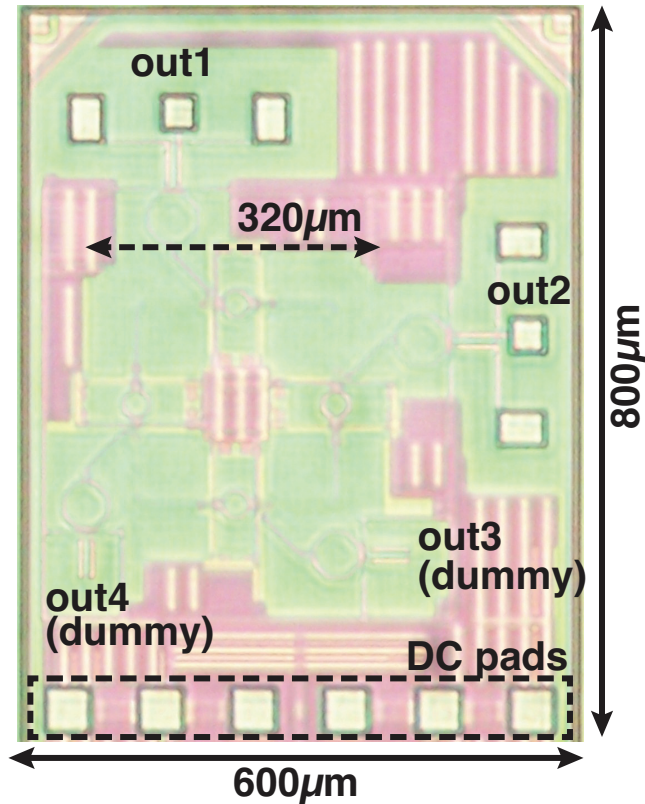


Figure 3.19: Microphotograph of the injection-locked loop.

probe, a DC probe, a down-converter, a spectrum analyzer and a power meter were used to perform the measurements, as illustrated in Fig. 3.12. The measured output frequency, output power, phase noise, and power dissipation are shown in Fig. 3.20.

Fig. 3.21 shows the comparison between the single 100 GHz active-varactor VCO and the injection-locked loop. The frequency of the injection-locked loop is decreased because of added coupling capacitance; thus, the frequency ranges from 91.7 to 95.5 GHz, and the tuning range is 4.1 %. Because of the capacitive coupling, the DC current is increased by four times without any additional power dissipation. The realized phase noise is approximately  $-118.8$  dBc/Hz at 10 MHz offset. In the case of the injection-locked loop, 6 dB ( $= 10 \log 4$ ) of phase noise improvement is clearly visible for offset frequencies between 2 MHz and 100 MHz. Additionally, the RMS jitter integrated from 2 MHz to 1 GHz is reduced from 77 fs to 33 fs. This technique should also improve phase noise at lower offsets, but due to measurement limitations and oscillator drift, the improvement is visible above 2 MHz. The simulated multiple phases are demonstrated in Fig. 3.22. The measurement results match well with theory and simulation, demonstrating that injection-locking works properly. The performance summary is in Table 3.3. The phase noise realized at 94 GHz is similar to the phase noise of the 20 GHz injection-locked loop [28].

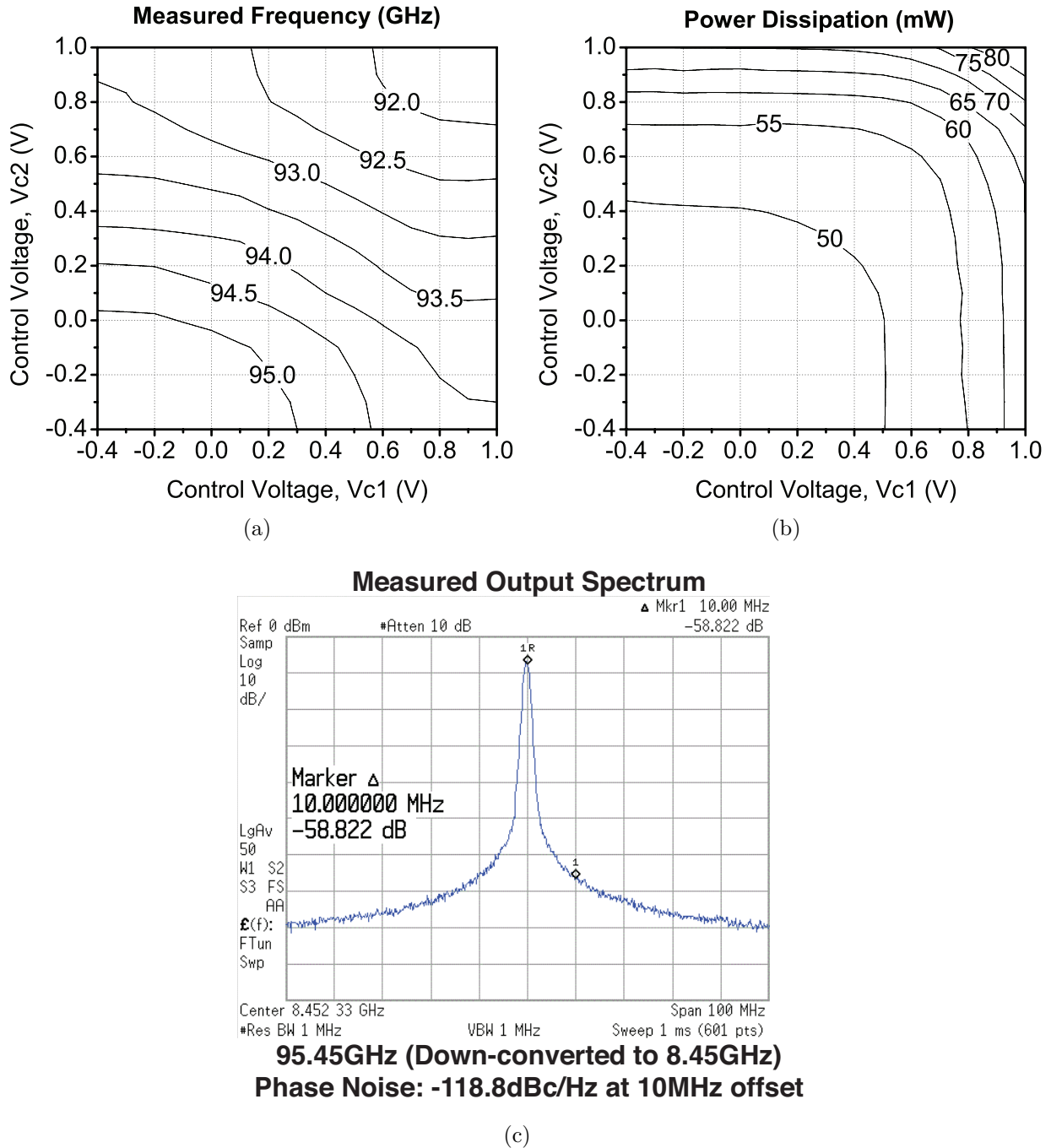


Figure 3.20: Measurement results of the injection-locked loop (a) frequency (b) power dissipation (c) output spectrum.



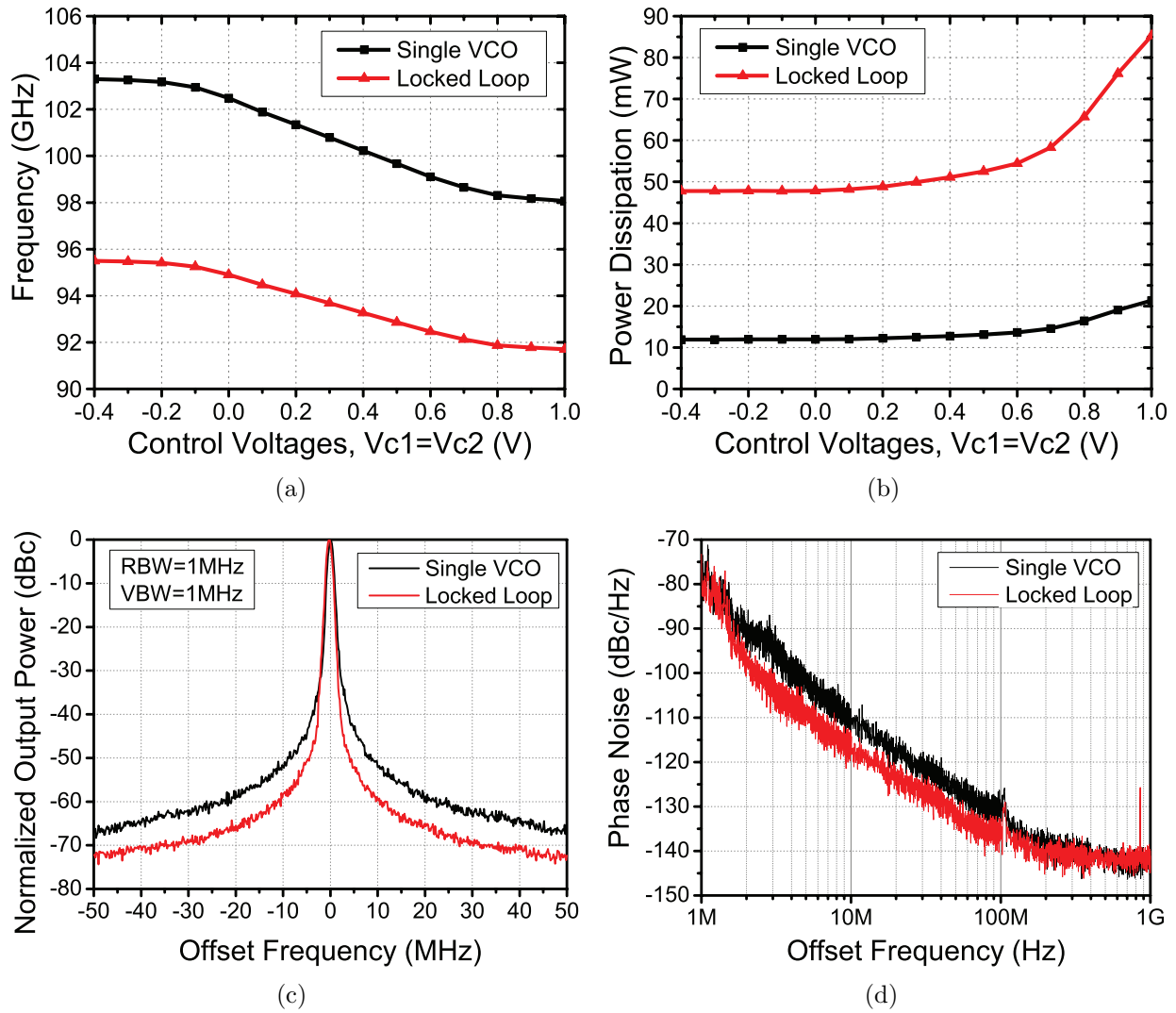


Figure 3.21: Comparison of the 100 GHz active-varactor VCO and the injection-locked loop (a) frequency (b) power dissipation (c) output spectrum (d) phase noise.

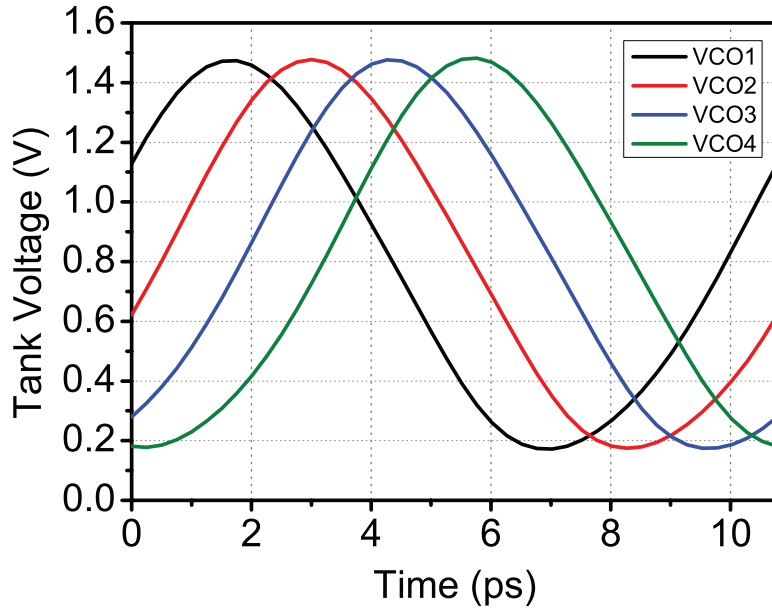


Figure 3.22: Multi-phase waveforms.

Table 3.3: Comparison of bi-directionally injection-locked loops.

	This Work [24]	Hekmat [28]
Technology	65 nm CMOS	90 nm CMOS
Coupling	Bi-directional Injection	Bi-directional Injection
Coupling Method	Transmission-Line-Based Capacitors	Buffers
Number of Coupled Oscillators	4	4
Frequency (GHz)	91.7 ~ 95.5	19 ~ 21
Tuning Range (TR) (%)	4.1	10
Phase Noise @ 10MHz (dBc/Hz)	-118.8	-121 <sup>(1)</sup>
Phase Error (°)	< 1	< 1
Differential Output Power (dBm)	1 ~ 4	-
Power Consumption (mW)	48 ~ 85	42.8
Supply Voltage (V)	0.8, 1.2	1
Control Voltage (V)	-0.4 ~ 1	-
Area ( $mm^2$ )	320×320	500×400
(1) Estimated from 1 MHz-offset phase noise		

## 3.6 Frequency Multiplication

In an  $LC$  VCO, the losses of the inductor and capacitor should be cancelled out by the negative conductance of the MOS devices. However, if the oscillation frequency is higher than  $f_{max}$  of the device, then the fundamental oscillation does not occur. As such, frequencies above  $f_{max}$  cannot be generated directly by a fundamental oscillator. For an oscillation frequency is slightly less than  $f_{max}$ , the oscillation is not always guaranteed because of PVT variations. Therefore, frequency multiplication is required to make a higher frequency using lower frequencies. Note that a frequency higher than  $f_{max}$  can be generated by a multiplier but not amplified by an amplifier; therefore, conversion gain and power efficiency are important factors of frequency multipliers. There are several methods to generate a sub-terahertz frequency as described below.

### 3.6.1 $N$ -push Multiplier

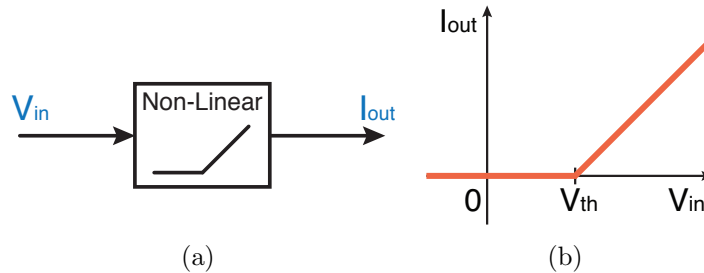


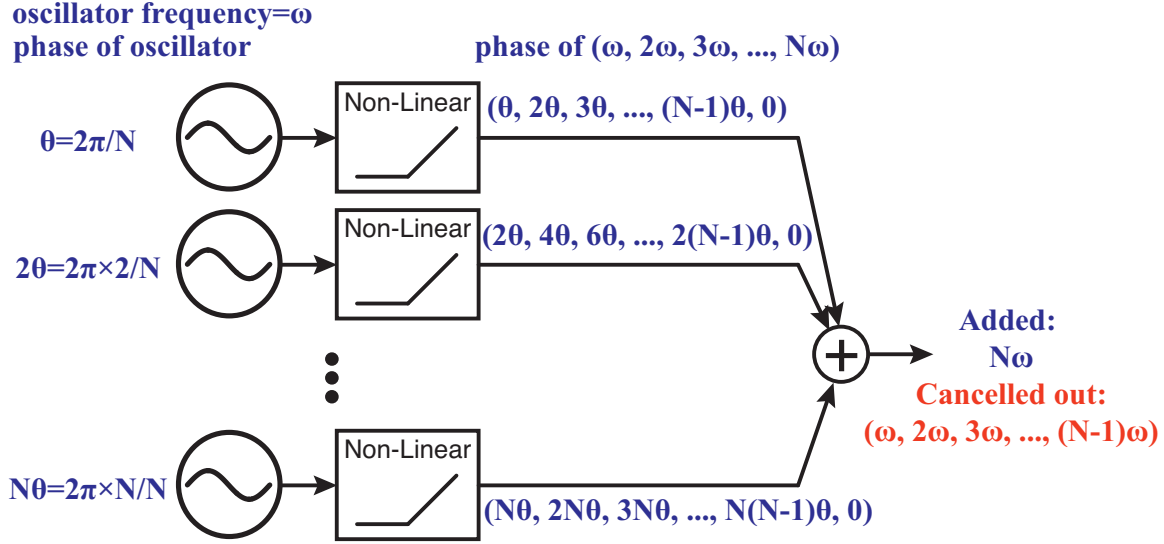
Figure 3.23: An ideal non-linear block.

The first method is  $N$ -push multiplier ( $N = 2, 3, 4, \dots$ ), which creates a frequency of  $Nf_{in}$ . Conventionally, it is called push-push, triple-push, and quadruple-push if  $N$  is 2, 3, and 4, respectively. Basically, it uses a non-linear circuit which generates harmonics. For instance, the input frequency is  $\omega$ , and the output has integer-multiple frequencies ( $\omega, 2\omega, 3\omega, 4\omega, \dots$ ). Fig. 3.23 shows a typical non-linear model in CMOS circuit design. The function in Fig. 3.23(b) is given by

$$I_{out} = \begin{cases} g_m(V_{in} - V_{th}), & (V_{in} > V_{th}) \\ 0, & (otherwise) \end{cases} \quad (3.28)$$

where  $V_{th}$  is a threshold voltage. This non-linear block simply models the  $I-V$  characteristic of the NMOS device.

Fig. 3.24 illustrates the operation of  $N$ -push multiplier. The multiplier consists of  $N$  number of non-linear blocks and an adder. As input signals,  $N$  number of synchronized oscillators are driving non-linear blocks as shown. The oscillators have the same frequency

Figure 3.24: Operation of  $N$ -push multiplier.

( $\omega$ ) but have progressive phase distributions ( $\theta, 2\theta, \dots, N\theta$ ). Note that  $\theta = 2\pi/N$ . The oscillator outputs are given by

$$\begin{aligned} V_1(t) &= V_{in} e^{j\omega t} e^{j\theta} \\ V_2(t) &= V_{in} e^{j\omega t} e^{j2\theta} \\ &\vdots \\ V_N(t) &= V_{in} e^{j\omega t} e^{jN\theta} \end{aligned} \quad (3.29)$$

After each oscillator output goes through a non-linear block, the input-output relationship is

$$I(t)/V_{in} = g_0 + g_1 e^{j\omega t} + g_2 e^{j2\omega t} + g_3 e^{j3\omega t} + \dots + g_N e^{jN\omega t} \quad (3.30)$$

where  $g_0, g_1, \dots, g_N$  are Fourier coefficients calculated from given  $g_m, V_b, V_{th}$ , and  $V_{in}$ . Accordingly the outputs become

$$\begin{aligned} I_1(t)/V_{in} &= g_0 + g_1 e^{j\omega t} e^{j\theta} + g_2 e^{j2\omega t} e^{j2\theta} + g_3 e^{j3\omega t} e^{j3\theta} + \dots + g_N e^{jN\omega t} e^{jN\theta} \\ I_2(t)/V_{in} &= g_0 + g_1 e^{j\omega t} e^{j2\theta} + g_2 e^{j2\omega t} e^{j4\theta} + g_3 e^{j3\omega t} e^{j6\theta} + \dots + g_N e^{jN\omega t} e^{j2N\theta} \\ &\vdots \\ I_N(t)/V_{in} &= g_0 + g_1 e^{j\omega t} e^{jN\theta} + g_2 e^{j2\omega t} e^{j2N\theta} + g_3 e^{j3\omega t} e^{j3N\theta} + \dots + g_N e^{jN\omega t} e^{jN^2\theta} \end{aligned} \quad (3.31)$$

All the outputs are then summed, using this equation:

$$\sum_{k=1}^N e^{jk\theta} = 0, \quad \left( \theta = \frac{2\pi}{N} \right) \quad (3.32)$$

and the final output current is given by

$$I_{out} = \sum_{k=1}^N I_k(t) = Ng_0V_{in} + Ng_NV_{in}e^{jN\omega t} \quad (3.33)$$

which concludes that the  $N$ -th harmonic term is created and other harmonics are cancelled out, obviating additional filtering for undesired harmonics. The conversion gain (the ratio of the  $N$ -th harmonic to the fundamental) can be calculated by

$$G_{conv} = \frac{|I_{out}(t)|}{N|V_k(t)|} = \frac{Ng_NV_{in}}{NV_{in}} = g_N \quad (3.34)$$

The  $g_N$  can be calculated using Fourier series and Eq. 3.28.

$$g_N = \frac{2g_m}{\pi} \left( \frac{\sin\left(\frac{N+1}{2}\alpha\right)}{2(N+1)} + \frac{\sin\left(\frac{N-1}{2}\alpha\right)}{2(N-1)} - \frac{\cos\left(\frac{1}{2}\alpha\right)\sin\left(\frac{N}{2}\alpha\right)}{N} \right) \quad (3.35)$$

where  $\alpha$  is a conduction angle, which is defined by

$$\cos\left(\frac{\alpha}{2}\right) = \frac{V_{th} - V_b}{V_{in}} \quad (3.36)$$

In addition, the  $g_0$  is calculated by

$$g_0 = \frac{g_m}{\pi} \left( \sin\left(\frac{\alpha}{2}\right) - \frac{\alpha}{2} \cos\left(\frac{\alpha}{2}\right) \right) \quad (3.37)$$

where  $\alpha$  is a conduction angle, which is defined by Equation 3.36. Using this  $g_0$ , the DC current is given by

$$I_{DC} = Ng_0V_{in} = \frac{Ng_mV_{in}}{\pi} \left( \sin\left(\frac{\alpha}{2}\right) - \frac{\alpha}{2} \cos\left(\frac{\alpha}{2}\right) \right) \quad (3.38)$$

To achieve a maximum  $g_N$  from Equation 3.35, the derivative of  $g_N$  is given by

$$\frac{dg_N}{d\alpha} = \frac{g_mV_{in}}{\pi N} \sin\left(\frac{\alpha}{2}\right) \sin\left(\frac{N\alpha}{2}\right) \quad (3.39)$$

and set to zero as

$$\frac{dg_N}{d\alpha} = \frac{g_mV_{in}}{\pi N} \sin\left(\frac{\alpha}{2}\right) \sin\left(\frac{N\alpha}{2}\right) = 0 \quad (3.40)$$

The solution is given by

$$\alpha = 0, \frac{2\pi}{N}, \frac{4\pi}{N}, \dots \quad (3.41)$$

The first non-trivial solution is given by

$$\alpha_1 = \frac{2\pi}{N} \quad (3.42)$$

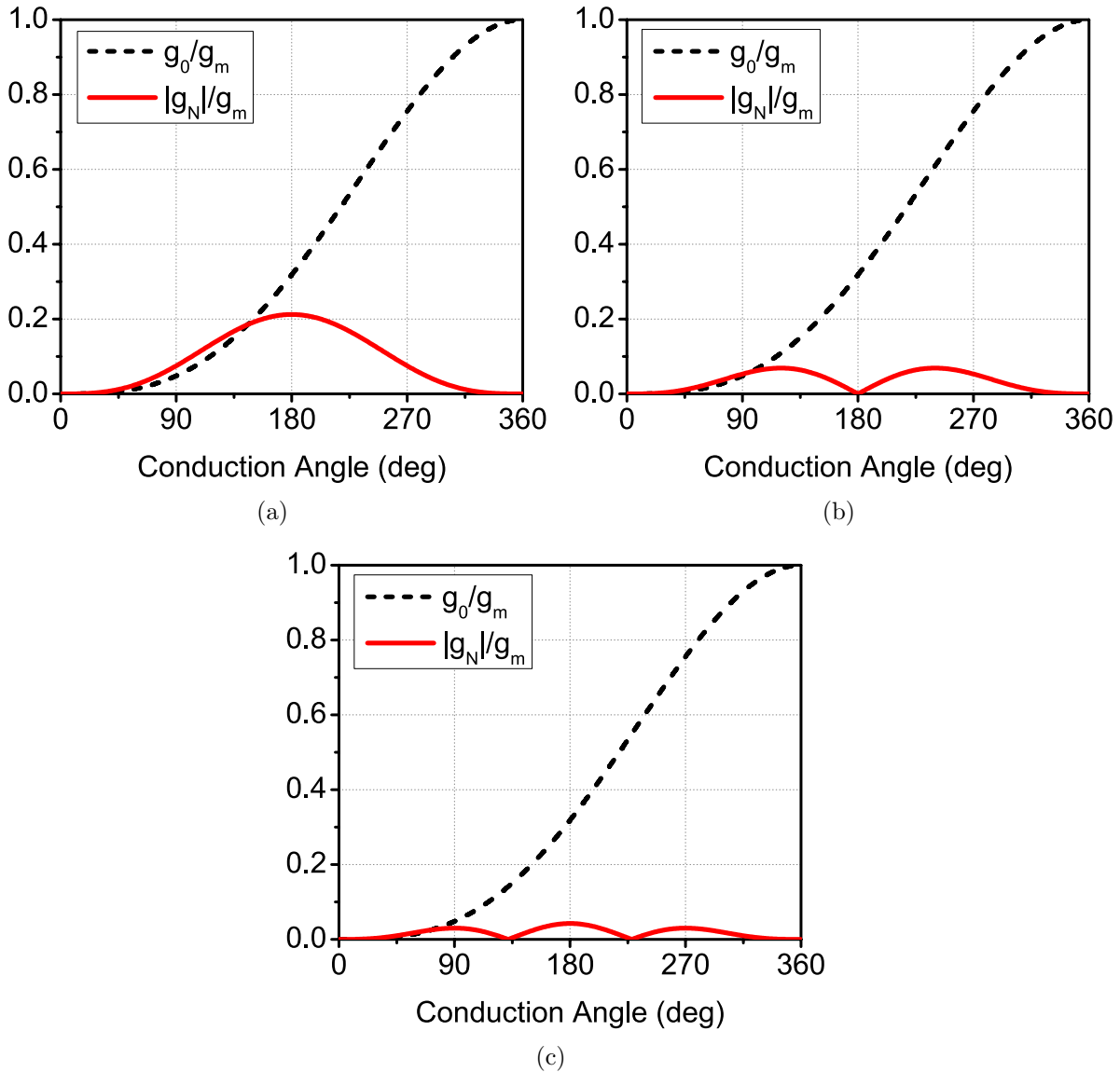


Figure 3.25:  $g_0$  and  $g_N$  of  $N$ -push multipliers (a) Push-push ( $N=2$ ) (b) Triple-push ( $N=3$ ) (c) Quadruple-push ( $N=4$ ).

and the corresponding bias voltage is

$$V_{b,1} = V_{th} - V_{in} \cos\left(\frac{\pi}{N}\right) \quad (3.43)$$

The push-push doubler is widely used because the two phases ( $0^\circ$ ,  $180^\circ$ ) are naturally achieved by a differential signal and because the conversion gain and efficiency are higher than triple-push and quadruple-push multipliers (Table 3.4). In contrast, triple-push or

Table 3.4: Comparison of the  $N$ -push multipliers

Architecture	$N$	Conduction angle ( $\alpha$ )	$ g_N /g_m$	$g_0/g_m$	$ g_N /g_0$
Push-push	2	180°	0.21	0.32	0.67
Triple-push	3	120°	0.069	0.11	0.64
		240°	0.069	0.61	0.11
Quadruple-push	4	90°	0.030	0.048	0.62
		180°	0.042	0.32	0.13
		270°	0.030	0.75	0.040

quadruple-push multipliers have a higher multiplication ratio than doubler, but it requires multiple (3 or 4) phases and increases complexity. The achievable conversion gain or efficiency is also lower. In the above analysis, amplitude mismatch and phase mismatch are not taken into account but any mismatch can degrade the gain and efficiency; thus, increasing  $N$  makes it harder to achieve theoretical performance. Sometimes, two cascaded push-push doublers are used to quadruple a frequency rather than a quadruple-push multiplier, depending on process performance and system specifications.

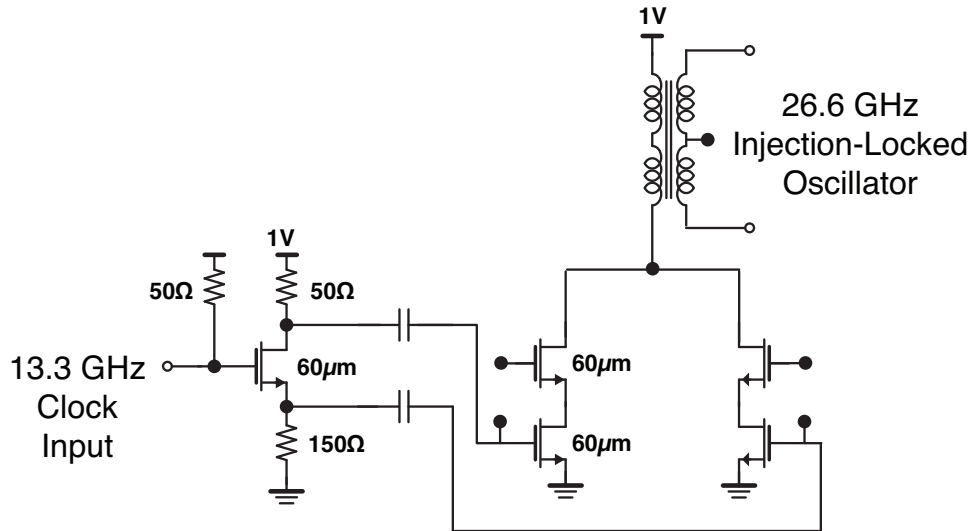


Figure 3.26: Schematic of the 26.6 GHz push-push doubler.

As shown in Fig. 3.26, a 26.6 GHz push-push doubler is designed in 65 nm CMOS. First, a 13.3 GHz external clock drives 50  $\Omega$ -matched active balun, which is self-biased and inductorless. Two output voltages (source and drain) have the same amplitude and a 180° phase difference. At low frequencies,  $R_S$  (source degeneration resistor) and  $R_D$  (drain resistor) should have same resistance to obtain the same amplitude, but  $R_S$  is made larger than  $R_D$  because of lowered capacitance impedance at 13.3 GHz. Next, these two phases drive a doubler shown on the right side. AC-coupling capacitors make gate-biasing independent

of the previous stage and the bias voltage can be chosen to maximize the conversion gain or efficiency. The cascode topology increases the output impedance, boosting the gain and efficiency. Finally, the 26.6 GHz doubler output is single-ended; thus, a passive balun is employed to make a differential signal. Simulation results are presented in Fig. 3.27. When  $V_b$  is 330 mV, the output current is maximized and the efficiency is 0.68, which is close to  $g_N/g_0$  in Table 3.4. The bias voltage is set to a slightly lower voltage, 300 mV, considering both efficiency and output current. The input power of the front-end active balun is 0 dBm and the output power is  $-10$  dBm including the output matching network. The active balun consumes 2.6 mW and the doubler consumes 4.2 mW from 1 V supply. The 26.6 GHz doubler is fabricated in 65 nm CMOS with a 240 GHz QPSK transceiver, which is demonstrated in Chapter 8.

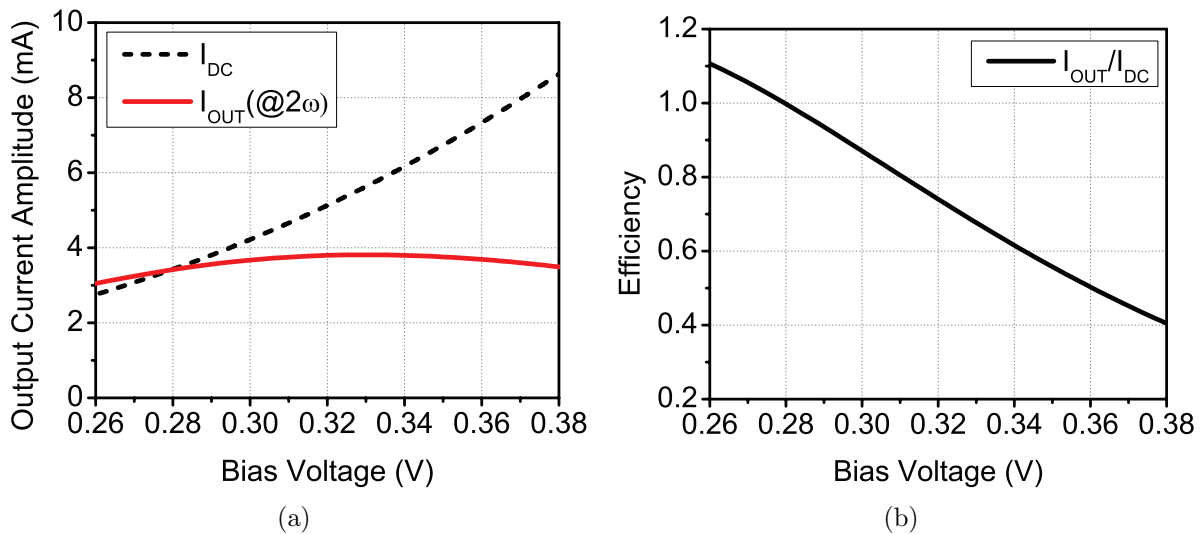


Figure 3.27: Simulation results of the 26.6 GHz push-push doubler (a) output currents (b) efficiency (matches to  $|g_N|/g_0$ ).

### 3.6.2 $N$ -push Oscillator

The  $N$ -push multiplier needs multiple-phase oscillators, as introduced in the previous subsection. The oscillator is often combined with the linear block because the oscillator output has harmonics by device non-linearity. For example, a differential oscillator has the second harmonic at a common mode node by the push-push operation. The second harmonic can be taken out using a matching network; therefore, it is called a push-push oscillator. [30]–[32] demonstrates a triple-push oscillator, which has three oscillators connected to one common node where the third harmonic is found. However, in the case of the triple-push oscillator, undesired modes (in-phase mode) should be prevented because three oscillators may have the same phase by improper injection-locking or PVT variations. The  $N$ -push



oscillators typically suffer from low output power due to low  $|g_N|/g_m$  value of the  $N$ -push operation.

### 3.6.3 Harmonic Matching

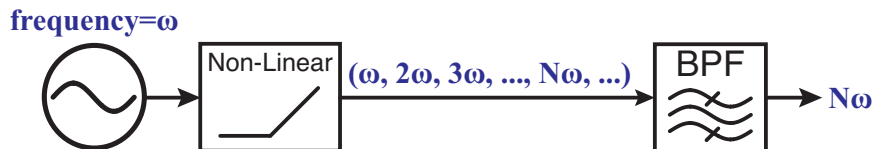


Figure 3.28: Operation of harmonic matching.

The main drawback of  $N$ -push multiplication is that  $N$  number of oscillators are required when  $N$  is higher than 2. Moreover, all of the oscillators have the same frequency, the same amplitude, and the equally spaced phases. The occupied area is also large. Therefore, the harmonic matching method proposes using just one branch, as illustrated in Fig. 3.28. Basically, the output of a non-linear block has harmonics presented in Eq. 3.30 and all of the harmonics, except one desired harmonic, can then be rejected by a band-pass filter (BPF). Fig. 3.28 shows that only the  $N$ -th harmonic tone passes through, which is the basic operation of the harmonic matching method. The conversion gain, efficiency, and output current can be derived from equations of the  $N$ -push operation. The filtered output current is given by

$$I_{out} = g_N V_{in} e^{jN\omega t} \quad (3.44)$$

where  $g_N$  is the conversion gain, which is the same as the gain of the  $N$ -push operation (Eq. 3.35). The DC current is given by

$$I_{DC} = g_0 V_{in} \quad (3.45)$$

where  $g_0$  is expressed in Eq. 3.37. The conversion efficiency ( $g_N/g_0$ ) is also the same as that of the  $N$ -push operation.

A drawback of the harmonic matching method is that undesired harmonics cannot be perfectly filtered out. For instance, the fundamental output tone is usually much stronger than the higher harmonics, hence the strong tone may cause the following stage, unless it is properly filtered out. If a filter stage is inserted, then the insertion loss degrades the conversion gain. Therefore, the harmonic matching method should be selected when the undesired harmonics do not cause a problem.

In practice, many differential frequency triplers employ this method. In fully differential amplifiers, the differential output has fundamental and odd-harmonic tones because even harmonics appear as common mode at the output. In addition, an output matching network designed at the third harmonic rejects the fundamental tone. If more rejection is needed, a notch filter, which attenuates the fundamental tone, can be used. Compared to the triple-push multiplier, this tripler is much simpler, but the strong fundamental tone is an issue.

### 3.6.4 Up-Converting Mixer

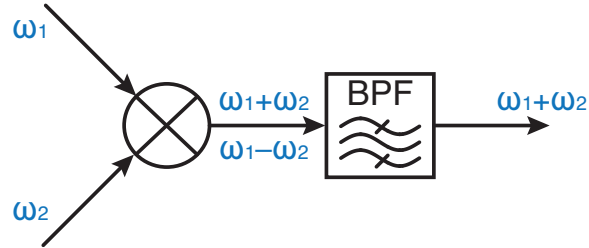


Figure 3.29: Operation of up-converting mixer.

An up-converting mixer is a simple way to add two frequencies. As illustrated in Fig. 3.29, a higher frequency ( $\omega_1 + \omega_2$ ) can be obtained using two frequencies ( $\omega_1, \omega_2$ ). The mixer can be combined with a frequency multiplier. If  $\omega_2 = 2\omega_1$ , for instance, then the output frequency is  $3\omega_1$ . It becomes a frequency tripler.

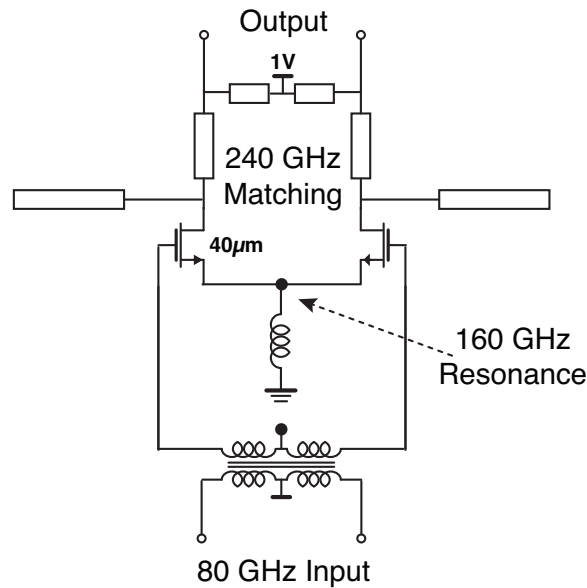


Figure 3.30: Schematic of the 240 GHz frequency tripler.

Fig. 3.30 shows a 240 GHz frequency tripler, which is integrated with a 240 GHz QPSK transceiver demonstrated in Chapter 8. This frequency tripler uses two methods, harmonic matching and up-converting mixer. Basically, the input frequency is 80 GHz ( $\omega$ ), and the output is matched at 240 GHz ( $3\omega$ ); thus, the third harmonic, 240 GHz, is the output frequency. The conversion gain is limited by  $g_N/g_0$  of the non-linear device, but mixer operation is added to boost the conversion gain and output power. The tail inductor at the common source node resonates out parasitic capacitance at 160 GHz ( $2\omega$ ), and the second harmonic

is raised, making the transistor operate as a mixer. 80 GHz ( $\omega$ ) comes to the gate and 160 GHz ( $2\omega$ ) is at the source, and then 240 GHz ( $\omega + 2\omega$ ) is present at the drain. By using two methods in one stage, additional insertion loss is precluded, and the conversion gain and output power can be boosted. The details and results of this frequency tripler will be demonstrated in Chapter 4.

## 3.7 Frequency Synthesizer

This section discusses popular architectures of frequency synthesizers, included in [33]. The frequency synthesizer is used to reduce phase noise and to synchronize oscillator phases. The frequency of an oscillator alone drifts over time and temperature; thus, a frequency synthesizer is required to prevent frequency drift. In communication systems, phase modulations require phase-locking. In phased-array systems, multiple chips cannot coherently operate without a frequency synthesizer. This section introduces frequency synthesizer architectures for millimeter-wave/sub-terahertz systems. Fig. 3.31 illustrates four architectures which are widely used. In these architectures, each synthesizer accepts the same input reference frequency and produces the same output frequency. Here, the output frequency is assumed to be, for example, 120 GHz for simplicity of calculation because 120 is a multiple of 2, 3 and 4.

### 3.7.1 PLL with a Fundamental VCO

The first architecture uses a PLL that employs a fundamental-frequency VCO [34]–[37]. As shown in Fig. 3.31(a), the VCO output and the first-stage divider input both operate at 120 GHz. This is the only architecture that requires a 120 GHz frequency divider; thus, the high-frequency divider design is critical. The fundamental VCO incurs design challenges that arise from the low gain in the transistors and the low Q in the varactors, which limits the tuning range at the high frequency of 120 GHz. For this design, the achievement of a high  $Q_{tank}$ , high swing, and low phase noise is challenging, but the VCO can be very small in size because of the low tank inductance ( $\approx 40$  pH). In addition, the output frequency cannot exceed  $f_{max}$ .

### 3.7.2 PLL with an $N$ -push VCO

The second architecture uses an  $N$ -push VCO instead of a fundamental VCO in the PLL [30], [36], [38]. Fig. 3.31(b) illustrates the case of  $N=3$ . One advantage of this architecture is that the first-stage divider need not operate at 120 GHz; therefore, the power dissipation of the divider chain can be reduced. Another advantage is that the VCO operates at a lower frequency (60 or 40 GHz). Thus, the transistors have higher gain and the varactors have higher Q compared to the previous architecture, and the VCO design can be relaxed depending on the factor of  $N$ . The output frequency can range up to  $N \cdot f_{max}$ . However,

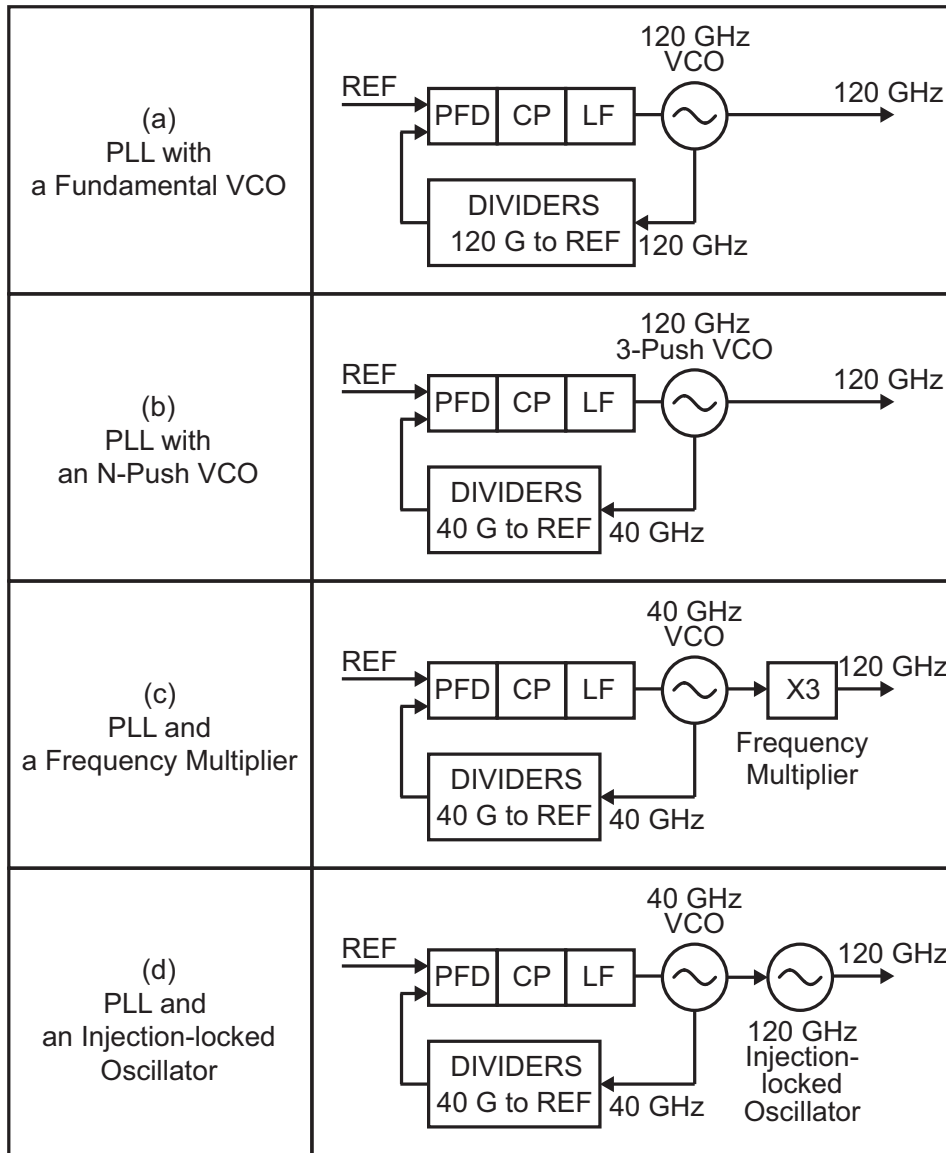


Figure 3.31: Frequency synthesizer architectures (a) PLL with a fundamental VCO (b) PLL with an  $N$ -push VCO (c) PLL with a frequency multiplier (d) PLL with an injection-locked oscillator.

$N$ -push VCOs suffer from low output power because the output power relies on the non-linearity of the devices. If  $N$  is 2, two phases can be easily obtained from a differential signal, but if  $N$  is 3 or more, more phases or more oscillators are required, leading to increased power consumption and possibly more complex routings. Amplitude and phase mismatches also reduce the output power. Thus, more buffer stages are needed to generate the desired output power.

### 3.7.3 PLL and a Frequency Multiplier

The third architecture uses a low-frequency PLL and an additional frequency multiplier [39], [40]. Here a frequency multiplier is defined as a non-oscillating block (not inside the PLL) that generates an output frequency that is a multiple of the input frequency. The multiplication ratio ( $N$ ) may be 2, 3, or higher; in Fig. 3.31(c), the multiplication ratio is 3. As the ratio increases, the conversion gain generally decreases, the output power decreases, and the required input power increases. Similar to the previous architecture, the maximum output frequency is  $N \cdot f_{max}$ . To obtain high output power, the input power (the VCO output power) should be even higher. In addition, the strong fundamental tone of the VCO can leak through the multiplier and degrade the mixer or system performance; thus, the unwanted tones should be properly filtered out. On the other hand, one considerable advantage of this architecture is that the PLL is designed at a lower frequency.

### 3.7.4 PLL and an Injection-Locked Oscillator

The last architecture uses a low-frequency PLL and an injection-locked oscillator [40], [41]. As illustrated in Fig. 3.31(d), this architecture is similar to the previous one, but it requires a 120 GHz oscillator and uses the injection-locking technique, which is widely used to improve phase noise. The output frequency is lower than  $f_{max}$  and another frequency multiplier is needed to generate a frequency above  $f_{max}$ . The oscillator should have a wide locking range to ensure injection locking despite PVT variations. The oscillator does not require a varactor for fine tuning but should have some switched capacitors to compensate for the frequency shifts caused by such variations. If the oscillator fails to be locked by the low-frequency oscillator, then it will exhibit pulling effects and contaminate the spectrum, leading to system misbehavior. Therefore, more design margins should be included to guarantee injection locking. To obtain a wide locking range requires low Q and strong injection [42], but a low Q increases power dissipation and phase noise. Moreover, the input injection signal is generated as a harmonic of the low-frequency VCO; thus, the VCO output power should be high, as in the third architecture.

Table 3.5: Comparison of the frequency synthesizer architectures

Architecture	Required Blocks	Advantages	Disadvantages
(a) PLL with a Fundamental VCO	Fundamental VCO, High-frequency divider	Low complexity Small area	High-frequency divider Low varactor Q/tuning range
(b) PLL with an $N$ -push VCO	$N$ -push VCO	Low division ratio Wide tuning range	Low output power Many oscillators ( $N > 2$ )
(c) PLL and a Frequency Multiplier	Low-frequency VCO A Frequency multiplier	Low division ratio Wide tuning range	Low output power Output harmonics
(d) PLL and an Injection-Locked OSC	Low-frequency VCO An injection-locked OSC	Low division ratio Better phase noise	Injection pulling issue Narrow locking range

### 3.7.5 Design Considerations

Table 3.5 summarizes the synthesizer architectures discussed above. The architecture should be carefully chosen depending on system requirements. The designer should first consider which blocks require the LO signal, how much output power and phase noise are desirable for the blocks, and whether the blocks require multiple phases, as in the case of an I/Q mixer. The designer should also consider the process technology (device characteristics, transmission-line performance, etc.). Next, the chip floorplan should be considered, including the PLL location, the distance between the VCO and the mixers/buffers, and the number of necessary routings. The designer cannot know or anticipate every parameter during initial design, but the above information is useful for architecture selection.

## 3.8 LO Distribution

This section introduces LO distribution, included in [33]. The importance of the LO distribution is often neglected in millimeter-wave/sub-terahertz systems. Even if the LO signal is generated well, if it is not properly delivered, the overall system performance will be degraded. Therefore, the LO distribution network should be carefully considered, even during the initial stages of the design of a millimeter-wave/sub-terahertz transceiver. A VCO is typically placed far from other transmitter/receiver amplifiers to avoid coupling or pulling issues [42]. Accordingly, the length of the routing line increases and becomes comparable to a significant amount of the wavelength in millimeter-wave frequencies. Such long lines can cause significant attenuation and phase shift. Moreover, because the VCO drives transmitter, receiver, and PLL divider, the signal distribution should be included in the analysis.

In millimeter-wave/sub-terahertz system design, the chip floorplan can significantly affect the layout and design of the sub-blocks. In particular, the LO distribution part is typically implemented after other parts (transmitter, receiver, and LO generation) are designed because of the dependence on the VCO output power, the power required by the transmitter or receiver, and the number of blocks that require the LO signal. Sometimes, LO distribution blocks may consume more DC power than is permitted by the power budget to satisfy the power requirements. Thus, it is important to know the performance of the LO distribution components and to apply it into the high-level system design. The following components are briefly introduced.

1. Transmission line: The types of lines that are typically implemented on silicon are the microstrip lines, coplanar waveguides, and coplanar striplines. Each transmission line is characterized by four parameters ( $Z$ ,  $\lambda$ ,  $Q_L$ ,  $Q_C$ ) that depend on the geometry [43], but in a high-level design, it is sufficient to know the characteristic impedance ( $Z$ ) and the line loss ( $\gamma$ ). From the chip floorplan, the length of the transmission line can be estimated, and the loss should be properly taken into account.

2. Power divider: It is useful to split one LO signal and to deliver the LO to many other blocks. For example, the Wilkinson divider is used to split the LO signal in phase and to isolate two outputs in the phased-array system [44]. The insertion loss is roughly 1dB (technology and frequency dependent) at millimeter-wave frequencies below 100 GHz.
3. Passive balun: Often employed for the conversion from a single-ended signal to a differential signal or vice versa. The transformer can be made small with a single turn. However, the transformer should be properly matched or loaded to balance the differential output. Typical insertion loss is roughly 1dB (technology and frequency dependent) at millimeter-wave frequencies below 100 GHz.
4. Active balun: Used to amplify the input signal and to convert it to the differential output. There is no loss, but it requires DC power, and linearity can be an issue. The input LO signal is a large signal; hence the second harmonic and the third harmonic can be generated through the active balun. If these harmonics degrade the mixer performance or the overall system performance, then harmonics should be attenuated, and an additional filter may be required.
5. Quadrature hybrid: Commonly used in systems that require quadrature up/down conversion. There is a trade-off between using a quadrature VCO and using a quadrature hybrid, and the trade-off is described well in [41]. For hybrids on silicon, transmission lines and lumped capacitors are typically used for area reduction, and lumped-transformer-based hybrids are even smaller [44]. Chapter 4 discusses details of the quadrature hybrid.

# Chapter 4

## Transmitter

This chapter discusses transmitter architectures and building blocks for sub-terahertz communications. The transmitter should properly modulate the carrier and generate high output power. Because the carrier frequency is higher than  $f_{max}$ , it is lossy to use an amplifier at the transmitter front-end, contrary to typical RF transmitters. First, this chapter compares three ways (doubler, tripler, or quadrupler) to multiply the carrier frequency at the front-end. Moreover, the transmitter requires the data modulation and the wideband signal path in addition to the frequency generation. Thus, quadrature phase generation, data modulation, and wideband amplifier design are discussed. Next, a power amplifier and a frequency tripler are demonstrated for a 240 GHz wideband QPSK/BPSK transceiver.

### 4.1 Conventional Transmitter Architecture

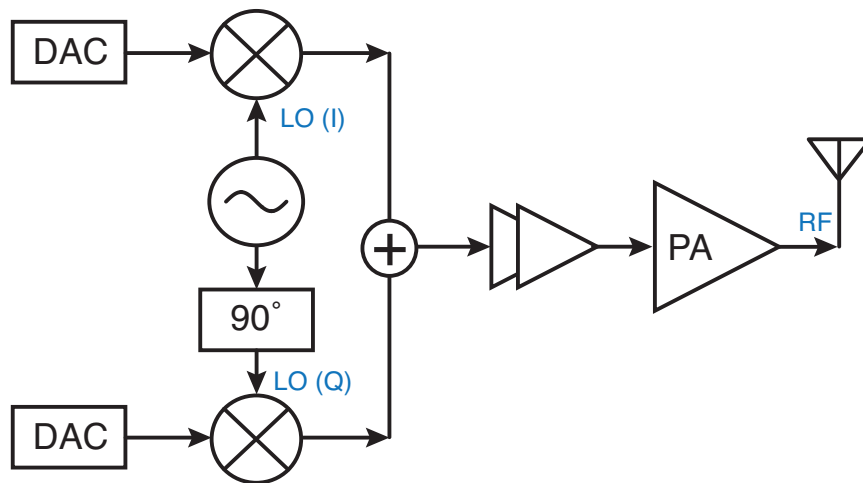


Figure 4.1: Conventional transmitter architecture for I/Q modulations.



Fig. 4.1 illustrates a typical transmitter architecture for I/Q modulations. An oscillator generates an LO signal, and the I and Q phases are then created. Two (I and Q) mixers up-convert the baseband (I and Q) data to the carrier frequency. Next, they are summed, amplified, and radiated by an antenna. The last-stage power amplifier (PA) should efficiently amplify the modulated signal and maximize the output power. Therefore, output power, efficiency, power dissipation, and linearity are important metrics.

## 4.2 Sub-Terahertz Transmitter Architecture

The sub-terahertz transmitter should have a different architecture. The process technology has finite unity-gain frequency ( $f_T$ ) and maximum oscillation frequency ( $f_{max}$ ). To communicate beyond the cut-off frequencies, amplifiers have no gain at such a high frequency, a different architecture should be proposed. It is difficult to generate a sub-terahertz carrier frequency, and it is more challenging to accurately modulate the carrier.

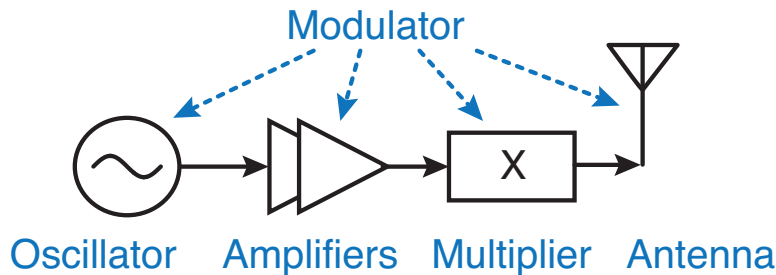


Figure 4.2: Sub-terahertz transmitter architecture.

A general architecture is shown in Fig. 4.2. A frequency multiplier should be placed immediately before the antenna. Note that once the sub-terahertz frequency is generated, no block can amplify it. The output power is determined by the output power of the amplifiers and the conversion gain of the frequency multiplier. High output power can be achieved by employing a power amplifier as typical transmitters. The conversion gains of frequency multipliers are generally lower than 0 dB; therefore, it is necessary to reduce the loss of the multiplier. The loss is highly dependent on the multiplication ratio; thus, the multiplication ratio should be carefully selected considering the process technology and the circuit limitations.

The modulation is also important but not included in Fig. 4.2. How to perform modulation is a main issue in the transmitter design. If the modulation is performed after the frequency multiplication, then the insertion loss should be minimized. Conversely, if the modulation is performed before the multiplication, then the frequency multiplication can affect the modulation. In the two transceivers (the 260 GHz OOK transceiver and the 240 GHz QPSK/BPSK transceiver), the modulations are performed before the frequency multiplication to reduce the insertion loss due to the modulator. As the modulated signal has a wide

bandwidth, the amplifiers and the frequency multiplier should be designed to have a wide bandwidth.

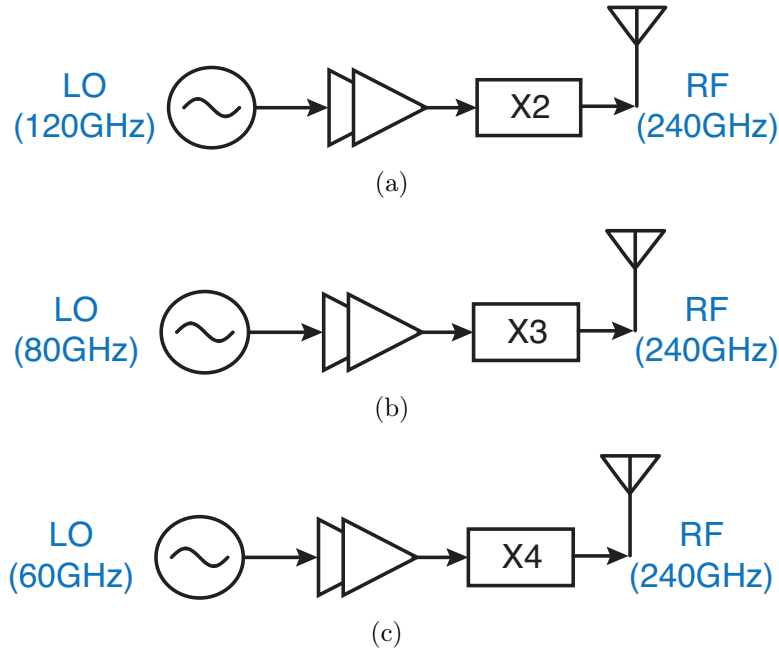


Figure 4.3: Sub-terahertz transmitter architectures (a) with doubler (b) with tripler (c) with quadrupler.

Fig. 4.3 shows three typical ways to generate a sub-terahertz frequency, for example, 240 GHz. In Fig. 4.3(a), a 120 GHz oscillator, amplifiers and a frequency doubler are used. In Fig. 4.3(b), an 80 GHz oscillator, amplifiers, and a frequency tripler are used. Finally, in Fig. 4.3(c), a 60 GHz oscillator, amplifiers, and a frequency multiplier are used. Modulators should also be taken into account. To fairly compare the three architectures, four parts should be separately examined: oscillator, amplifier, multiplier, and modulator.

First, as the frequency increases, millimeter-wave oscillators suffer from a low quality factor, low power efficiency, high phase noise and narrow tuning range, as analyzed in Chapter 3. To achieve higher efficiency, a lower-frequency oscillator is preferred, but most power is dissipated by other amplifiers and frequency multiplier. Thus, the phase noise and the tuning range should be considered first in the oscillator part.

Second, the millimeter-wave CMOS amplifiers are introduced in [13]. It is well known that the output power and efficiency generally drop as the frequency increases. In addition, if modulated signals go through amplifiers, then a wide bandwidth is required for high-speed communications. For advanced modulations, such as QAM, the linearity and back-off efficiency are important factors to consider.

Third, as the multiplication ratio increases, the conversion gain typically decreases, as analyzed in Chapter 3. In  $N$ -push multipliers, conversion gain and efficiency are presented

in Fig. 3.25. Because the input power is very large ( $> 10$  dBm), the power dissipation is comparable to the power amplifier. The conversion gain, power dissipation, and bandwidth are mainly considered.

The last one is the modulation. It is hard to give a general rule because there are many modulations and various implementations. In Chapter 7 and Chapter 8, the OOK and QPSK/BPSK modulations are discussed for sub-terahertz transceivers.

The transmitter EIRP is one of the important metrics in the transmitter design. In these sub-terahertz transmitter architectures, the EIRP is given by

$$EIRP_{TX} = P_{out,PA} \cdot L_{conv,X} \cdot G_A \quad (4.1)$$

where  $P_{out,PA}$ ,  $L_{conv,X}$ , and  $G_A$  are the output power of the last-stage amplifier (PA), the conversion loss of the frequency multiplier, and the antenna gain, respectively. In addition, insertion losses of matching networks can degrade the  $EIRP_{TX}$ . Modulation method can also affect the  $EIRP_{TX}$ . For example, if a lossy modulator is placed between the PA and the multiplier, the  $EIRP_{TX}$  is reduced by the loss of the modulator.

## 4.3 I/Q Generation

In QPSK and QAM modulations, quadrature phases (I and Q) are necessary to generate constellation points. Any amplitude mismatch or phase mismatch can degrade the transmitter EVM (error vector magnitude), increasing the bit error rate. Therefore, it is important to accurately generate I and Q phases to reduce the EVM and BER in high-speed communication systems. In low-frequency ( $\omega < 30$  GHz) systems, a frequency divider is widely used to generate I and Q phases because an oscillator frequency of  $2\omega$  can be efficiently achieved. However, at sub-terahertz frequencies ( $\omega > 100$  GHz); thus,  $2\omega$  is inefficient or impossible to generate. This section investigates three ways to create I/Q phases at sub-terahertz frequencies.

### 4.3.1 Quadrature Oscillator

The first method is to use a quadrature oscillator. Basically, two oscillators are injection-locked by each other, as illustrated in Fig. 4.4. In Chapter 3, it is introduced that if multiple oscillators are properly coupled, multiple phases can be created. Likewise, two oscillators are coupled through a buffer, which provides a  $90^\circ$  phase shift. The phases of the first differential oscillator are  $0^\circ$  and  $180^\circ$ . The  $90^\circ$  phase, shifted by the buffer, is injected to the other oscillator, the phases of which are  $90^\circ$  and  $270^\circ$ . The  $270^\circ$  phase becomes  $360^\circ$  by the buffer, but the connection is crossed, such that  $180^\circ$  is injected to the first oscillator. Thus, the two oscillators are mutually locked.

In practice, there are many non-idealities. The buffer delays are not exactly  $90^\circ$ , causing a phase shift and a frequency shift. As such, the phase difference deviates from  $90^\circ$  and the oscillators do not operate at the natural frequency. If the phase shift is too large or

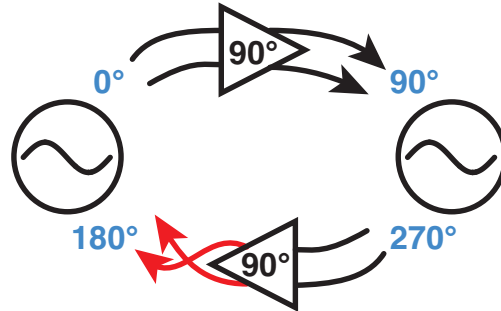


Figure 4.4: Operation of a quadrature oscillator.

the frequency is out of the locking range, then the two oscillators are not fully locked and pull each other. As a result, stable I and Q phases cannot be achieved. In addition, device mismatches can cause the phase/frequency shift. At high frequencies, where the device size is small and the values of L and C are small, there can be variations and imbalances, leading to amplitude and phase mismatches. The layout should be carefully designed to reduce the mismatch. Moreover, non-linearity of device capacitors and parasitics worsens I and Q mismatches.

Typically, a quadrature oscillator suffers from narrow tuning range and low output power. Two injection buffers are connected to each oscillator and one additional output buffer is required. Thus, the varactor size is limited and the tuning range gets narrower. To achieve a wide locking range, the quality factor of the tank is designed to be low. Some amount of the output power is used by the injection buffer, leading to lower output power. Finally, frequencies higher than  $f_{max}$  cannot be generated; therefore, an additional frequency multiplier is required to generate a higher frequency than  $f_{max}$ . The additional multiplier typically adds to the amplitude and phase mismatches. Moreover, the quadrature phase ( $90^\circ$ ) can be lost depending on the multiplication ratio.

### 4.3.2 Quadrature Hybrid

The passive quadrature hybrids can generate I and Q phases. Basically, it has four ports: P1(input), P2(output), P3(output), and P4(isolated). The two outputs (P2 and P3) have  $90^\circ$  phase difference, and the isolated port has no output power when the four ports are properly matched. Because the hybrid is a passive device, there is no frequency limit like active devices, but there is an insertion loss in addition to splitting the input power.

Fig. 4.5 shows four different designs of quadrature hybrids. First, Fig. 4.5(a) uses quarter-wave ( $\lambda/4$ ) transmission lines. The characteristic impedances are noted in the Fig. 4.5(a). By the even-odd mode analysis [45], the function of the quadrature hybrid is verified. When the port impedances are  $Z_o = 50 \Omega$ , two transmission lines have a characteristic impedance of  $50 \Omega$  and the other two transmission lines have an impedance of  $\frac{50}{\sqrt{2}} = 35.4 \Omega$ . The design and analysis are mathematically performed; thus, it is straightforward to implement the

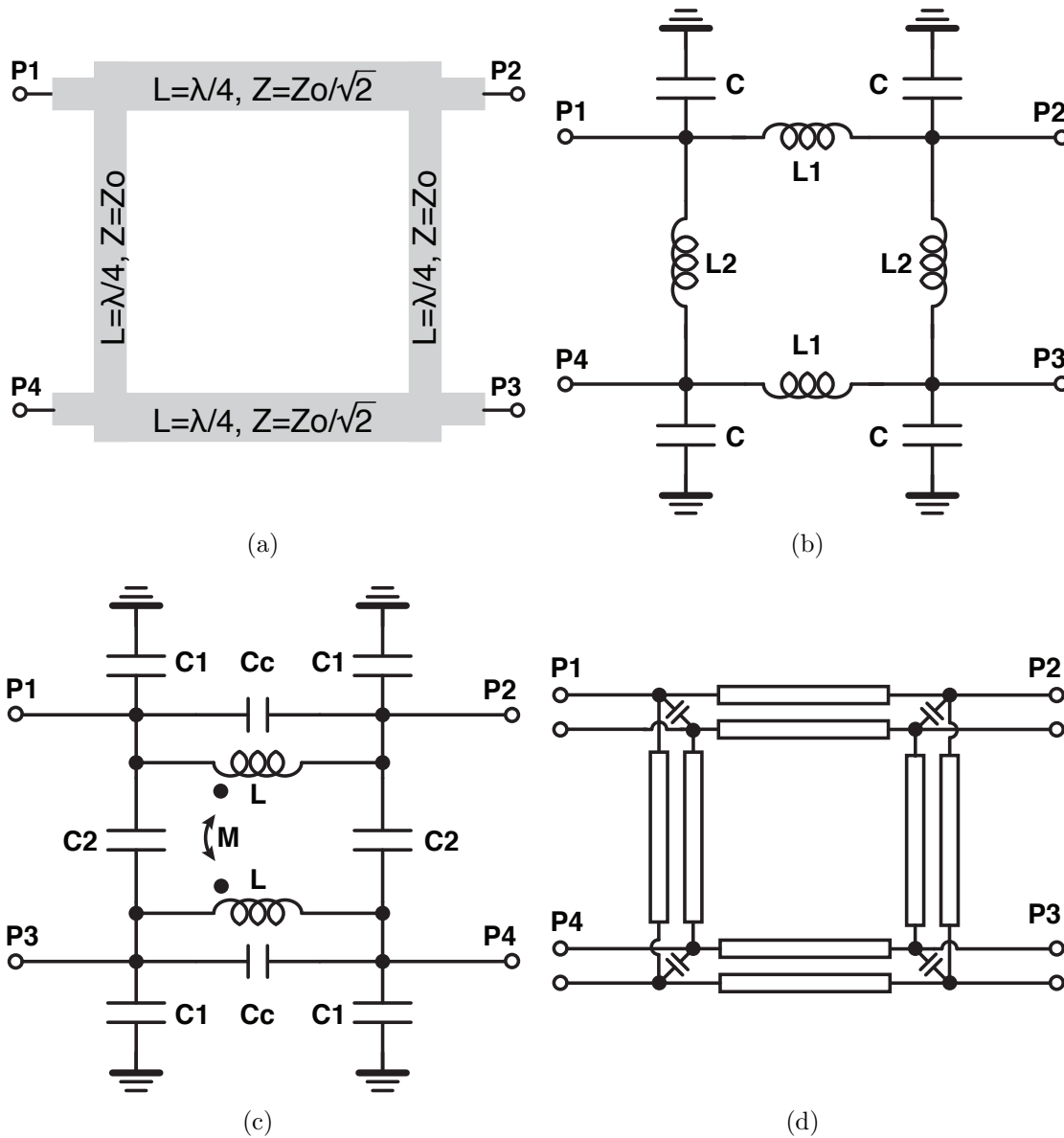


Figure 4.5: Quadrature hybrids (a) transmission-line-based (b) lumped-element-based (c) transformer-based (d) differential input/output.

hybrid. However, the problem is that the transmission-line-based hybrids occupy large area. The lengths of transmission lines are  $\lambda/4$  and the size is very large. Meandering the lines reduces the size, but hybrids are still bulky and lossy.

To reduce the size, lumped components (inductors and capacitors) can be employed instead of the transmission lines. Fig. 4.5(b) shows one of various designs. Basically, a transmission line can be modeled as a  $\Pi$ -model (C-L-C), neglecting resistive losses. The

transmission line ( $Z_o/\sqrt{2}$ ) between P1 and P2 can be modeled using  $L_1$  and  $C_1$ , expressed as

$$L_1 = \frac{Z_o}{\sqrt{2} \cdot \omega} \quad (4.2)$$

$$C_1 = \frac{\sqrt{2}}{Z_o \cdot \omega} \quad (4.3)$$

The other transmission line ( $Z_o$ ) can be modeled using  $L_2$  and  $C_2$ , expressed as

$$L_2 = \frac{Z_o}{\omega} \quad (4.4)$$

$$C_2 = \frac{1}{Z_o \cdot \omega} \quad (4.5)$$

The capacitors can be summed as

$$C = C_1 + C_2 = \frac{1 + \sqrt{2}}{Z_o \cdot \omega} \quad (4.6)$$

In real circuits, inductors can be implemented with transmission lines shorter than  $\lambda/4$ . A 94 GHz quadrature hybrid [46] is implemented using microstrip lines and MIM-capacitors. The size is  $162 \mu\text{m} \times 237 \mu\text{m}$ , while  $\lambda/4$  is approximately  $400 \mu\text{m}$  in silicon. The simulation results are shown in Fig. 4.6.

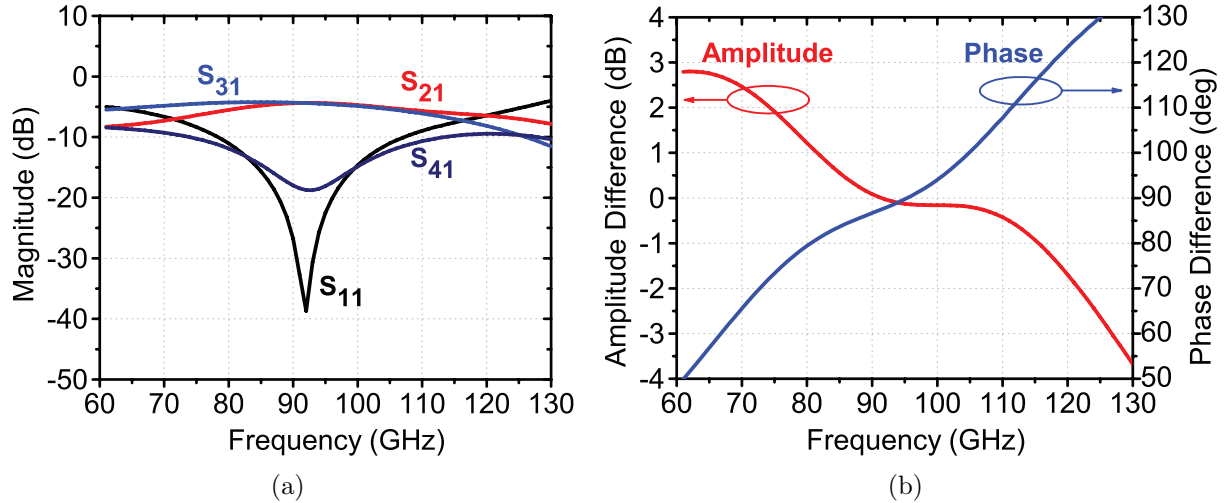


Figure 4.6: Simulation results of the 94 GHz quadrature hybrid (a) S-parameters (b) amplitude/phase mismatches.

To further reduce the size, a transformer-based quadrature hybrid is reported in [44], and the schematic is illustrated in Fig. 4.5(c). The diameter of the transformer is  $36 \mu\text{m}$  and the hybrid size is  $0.002\lambda^2$ , while  $\lambda/4$  is approximately  $625 \mu\text{m}$  in silicon [44].

Fig. 4.5(d) shows a differential quadrature hybrid. The differential operation is commonly used in many circuits, such as differential oscillators, amplifiers, and balanced mixers. The beauty of the differential hybrid is that it is straightforward to connect the hybrid with other differential circuits. Compared to a single-ended hybrid, the insertion loss of a differential hybrid itself is generally higher, but the single-to-differential and differential-to-single operations can be removed; thus, the overall loss of the differential hybrid can be smaller than that of the single-ended hybrid. The size of the 80 GHz differential hybrid is  $210\ \mu\text{m} \times 270\ \mu\text{m}$ , and the simulated amplitude mismatch is  $\pm 1\ \text{dB}$  and the phase mismatch is  $\pm 3^\circ$  in  $80 \pm 5\ \text{GHz}$  [1].

### 4.3.3 Delayed Line

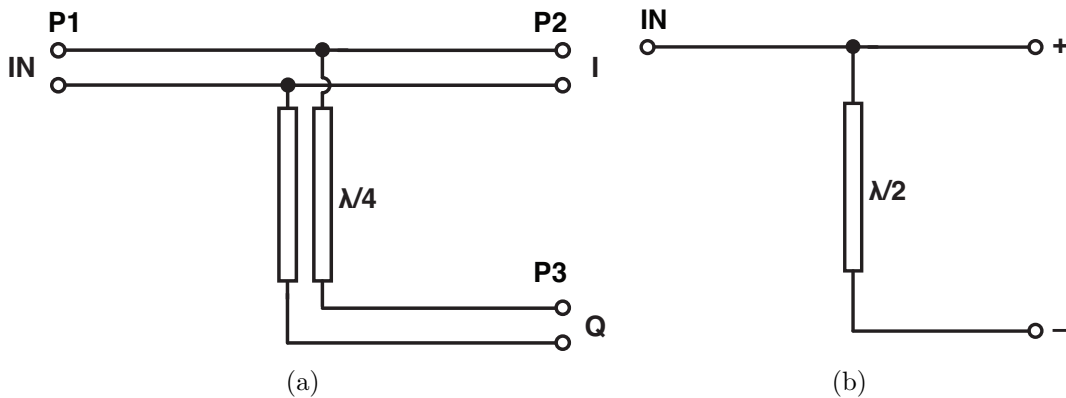


Figure 4.7: Delayed lines (a)  $90^\circ$  delay (b)  $180^\circ$  delay.

A delayed line (Fig. 4.7(a)) can be employed to generate I and Q quadrature phases, whereas Fig. 4.7(b) shows a way to generate a differential signal using a delayed line. Although the delay is well designed in simulation, the actual delay may have delay error because the line delay is sensitive to the variations of width, thickness, and conductivity. In addition, an amplitude mismatch occurs due to the loss of the quarter-wave line. As a result, amplitude and phase mismatches exist, depending on the frequency and process technology. The isolation between two outputs can be a problem because the two outputs are connected with quarter-wave transmission lines. However, the insertion loss is relatively lower than the quadrature hybrid at sub-terahertz frequencies. In some applications, the amplitude or phase mismatch is not a significant issue, allowing the use of this delayed line.

## 4.4 Modulator

Modulation is one of the important parts in transmitters. Data communication is not possible without modulation and demodulation, and no information is delivered over the

air. As presented in Chapter 2, there are many modulation schemes in wireless communication systems: OOK, FSK, PSK, QAM, etc. In this section, some modulation schemes are examined for millimeter-wave/sub-terahertz transmitters.

#### 4.4.1 OOK

OOK is a simple, one-bit modulation and performed by turning the transmitter output signal on or off. A receiver can detect the amplitude of the received signal and determines whether the bit is ON or OFF. Both coherent and non-coherent detections are possible.

In real IC and CMOS technology, it is challenging to completely turn off the transmitter output signal in a few cycles. The *LC* tank stores reactive energy, and it takes time for the reactive energy to be attenuated. The time is roughly proportional to the quality factor of the tank. Because a low quality factor causes high loss and low output power, it is not practical to lower the quality factor to reduce the turn-off time. In addition, a finite on-off ratio degrades the link performance (dynamic range and BER) because the two signals (ON and OFF) become indistinct.

There are two typical ways to turn on/off AC signals: voltage switching and current switching. In the voltage switching, a MOS switch is placed in shunt. When the switch is on, two nodes are shorted, such that AC signal is terminated by the low impedance. When the switch is off, the switch is open and the signal is passing through. In the current switching, a common node of two switches is connected to the input, and the other node of each switch is connected to different loads, as performed in a differential pair. By turning on a switch and off the other switch, the input can be connected to the transmitter output or a dummy load. These methods are used in various applications [3], [47]. For high-speed communications, the 260 GHz OOK transceiver (Chapter 7) uses both voltage and current switching [48]. To improve the on-off ratio, a current-mode switch is used before a power amplifier, and a voltage-mode switch is used after the power amplifier.

#### 4.4.2 QPSK and BPSK

QPSK and BPSK are widely used in real communication systems. They have lower BER than other modulations for a given SNR, as shown in Fig. 2.2. The implementation is not complicated. A phase rotator can be employed to change the phase of the AC signal. Because the amplitude is fixed, the linearity specification is relaxed, and the phase noise or phase distortion is a main issue.

A 80 GHz QPSK modulator is designed for a 240 GHz QPSK/BPSK transceiver, demonstrated in Chapter 8. It can also be used for an 80 GHz QPSK transmitter. Fig. 4.8 shows the schematic of the QPSK modulator. Basically, it consists of two double-balanced mixers (I and Q), and the two outputs are summed at the output. The baseband ports (BB-I and BB-Q) are driven by an on-chip PRBS generator, which can support three operation modes (QPSK, BPSK, and a continuous-wave mode). The LO ports (LO-I and LO-Q) are driven



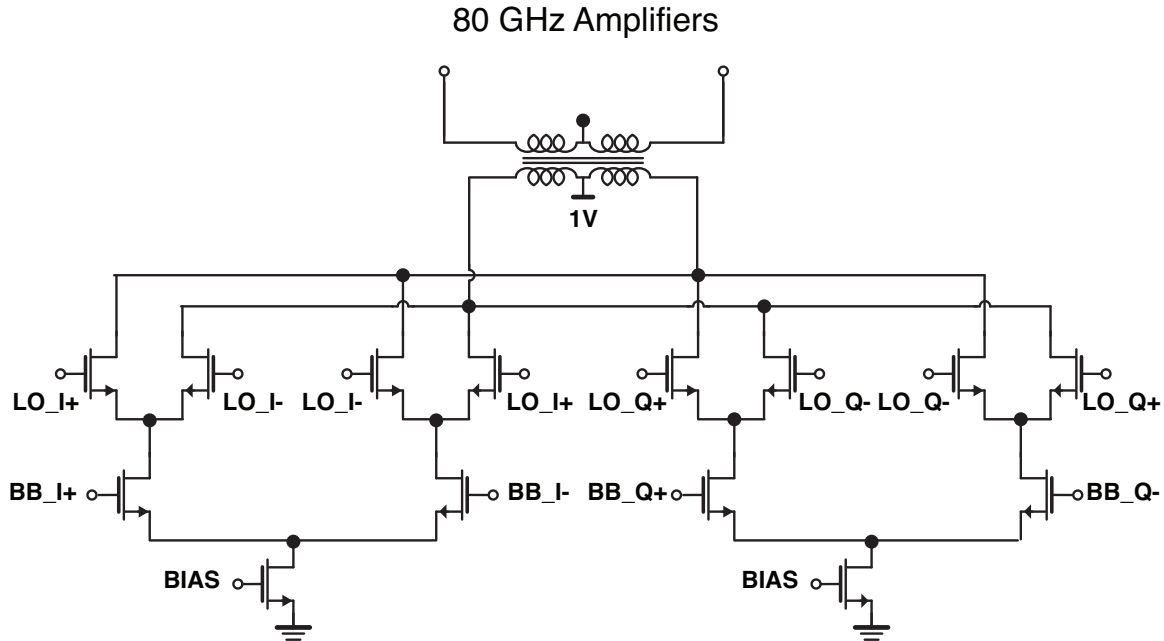


Figure 4.8: Schematic of QPSK modulator.

by a differential quadrature hybrid. The output port is transformer-coupled and connected to 80 GHz amplifiers.

In the modulator design, mismatches should be minimized to generate precise constellation points. In real layout, asymmetrical routings and unbalanced connections are unavoidable. Thus, the layout of the transistors and routing lines should be carefully designed using post-layout simulations and EM simulations to minimize amplitude/phase error. Moreover, sharp transition of the baseband data is required to achieve a high data rate and to reduce noisy (rising/falling) regions. The method of sharpening the rising/falling edges is discussed in Chapter 6. Consequently, the modulator can operate up to 40 Gbps in the QPSK mode.

The 80 GHz QPSK modulator is simulated using a differential hybrid and a PRBS generator [1]. While the baseband data are fixed, the input port is at the hybrid input and the output port is on the secondary side of the output transformer. Fig. 4.9 presents the output power and power dissipation of the modulator, which vary with the LO power. When the LO power is 0 dBm, the output power is  $-7$  dBm and the power dissipation is 7.3 mW; therefore, driving amplifiers and a power amplifier should achieve a power gain of over 20 dB to maximize the output power.

## 4.5 80 GHz Power Amplifier

In sub-terahertz transmitters, the transmitter output power is determined by the PA output power and the conversion gain of the frequency multiplier. Thus, the PA output

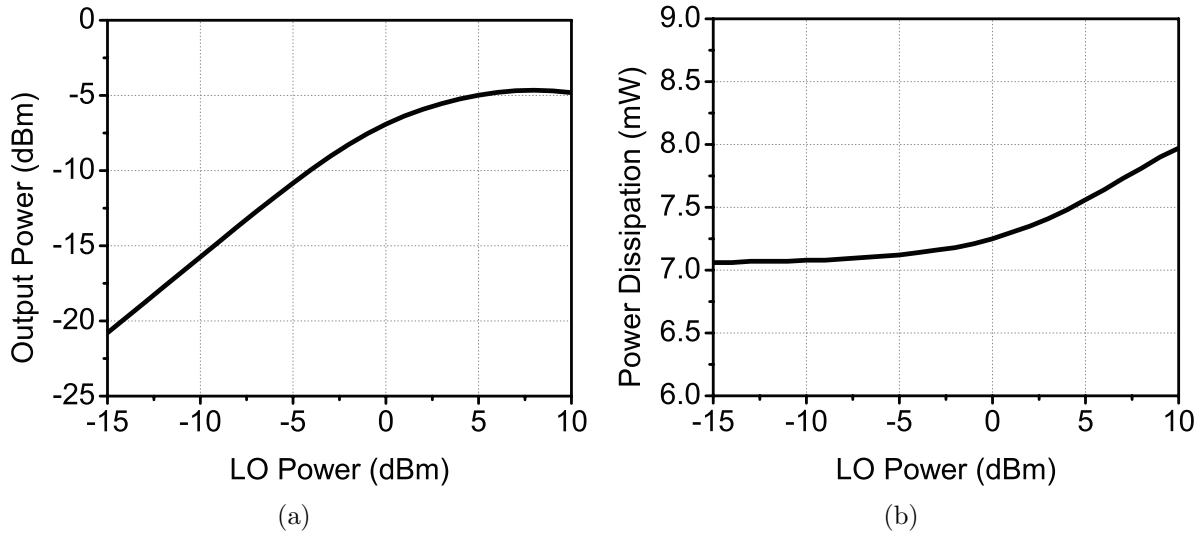


Figure 4.9: Simulation results of the QPSK modulator (LO power at the hybrid input) (a) output power (b) power dissipation.

power should be maximized, while achieving high efficiency. Fig. 4.10 shows the schematic of the power amplifier. The class-E switching PA is selected to efficiently generate an output power above 13 dBm in 65 nm CMOS.

Table 4.1: Simulation results of the 80 GHz class-E switching power amplifier

Device Size ( $\mu\text{m}$ )	$R_{out,p}$ ( $\Omega$ )	$L_{out,p}$ (pH)	$Y_{in,p}$ (mS)	$C_{in,p}$ (fF)	$P_{in}$ (dB)	Gain (dB)	$P_{out}$ (dBm)	$I_{avg}$ (mA)	DE (%)	PAE (%)
60	53.1	64.3	14	75.8	2.42	6.26	8.67	13.7	53.9	41.4
80	47.5	71.9	20.5	99.7	4.09	5.72	9.81	16.8	57	41.7
100	35.9	54.4	24.7	126	4.94	5.94	10.8	21.6	56.2	41.8
120	30.3	48.1	30.4	149	5.8	5.82	11.6	25.8	56.3	41.6
140	25.0	43.1	35.7	173	6.49	5.82	12.3	30.6	55.6	41.0
160	23.0	31.9	38.6	203	6.83	6.03	12.9	34.5	56.5	42.0

First, the device size is selected to be 100  $\mu\text{m}$ , considering the output power and efficiency. Table 4.1 shows the simulation results for a single-ended class-E switching power amplifier using a single NMOS device (low threshold voltage). The input signal is an ideal sinusoidal wave with an amplitude of 0.5 V ( $1 V_{p-p}$ ). A load-pull analysis is performed for each device size to determine the optimum load impedance ( $R_{out,p}$  and  $L_{out,p}$ ) that produces the highest output power. The load impedance is then used to obtain the input impedance of the amplifier in a large-signal periodic simulation. The input power is given by the input voltage amplitude (0.5 V) and  $Y_{in,p}$ . Thus, the gain can be calculated as  $P_{out} - P_{in}$ . The power

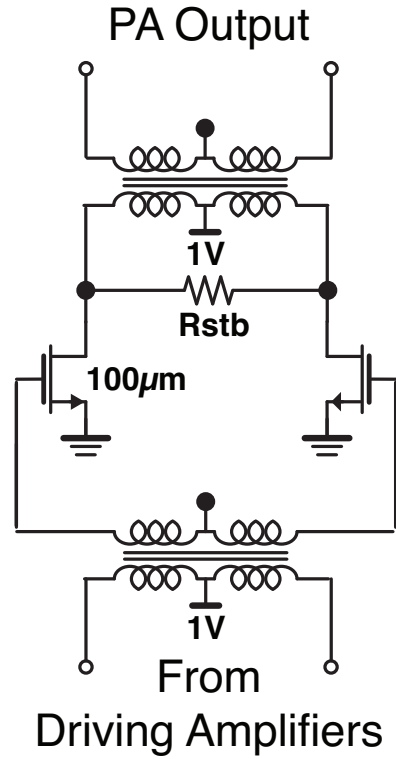


Figure 4.10: Schematic of the class-E switching power amplifier

dissipation is obtained by averaging the drain current, and the drain efficiency (DE) and the power added efficiency (PAE) are then calculated. The table shows that the output power increases with the device size but that large devices have low output impedances compared to the input impedance of the frequency tripler. If the output impedance decreases, then the insertion loss of the output matching network increases, degrading the overall transmit output power. For a device size of  $100\mu\text{m}$ , the differential amplifier achieves an output power of 13.8 dBm (which is 3 dB higher than 10.8 dBm), a gain of 6 dB, and a PAE of 41 % without input and output matching networks.

The 80 GHz switching PA is simulated using driving amplifiers. Fig. 4.11 presents the drain voltage and the drain current over a period and shows the switching operation. The class-E operation is not clearly visible at millimeter-wave frequencies because the input voltage is not switching fully. When driving amplifiers are used, the output power is not significantly different from the output power shown in Table 4.1.

The three-stage driving amplifiers and the power amplifier are illustrated in Fig. 4.13. The device sizes of the driving amplifiers are scaled down from that of the power amplifier. The first-stage driving amplifier operates as a linear class-A amplifier to achieve a high gain. A wide bandwidth is obtained by the inter-stage impedance matching using loosely coupled transformers, as shown in Fig. 4.12. Parallel resistors ( $R_{stb}$ ) are added to improve the

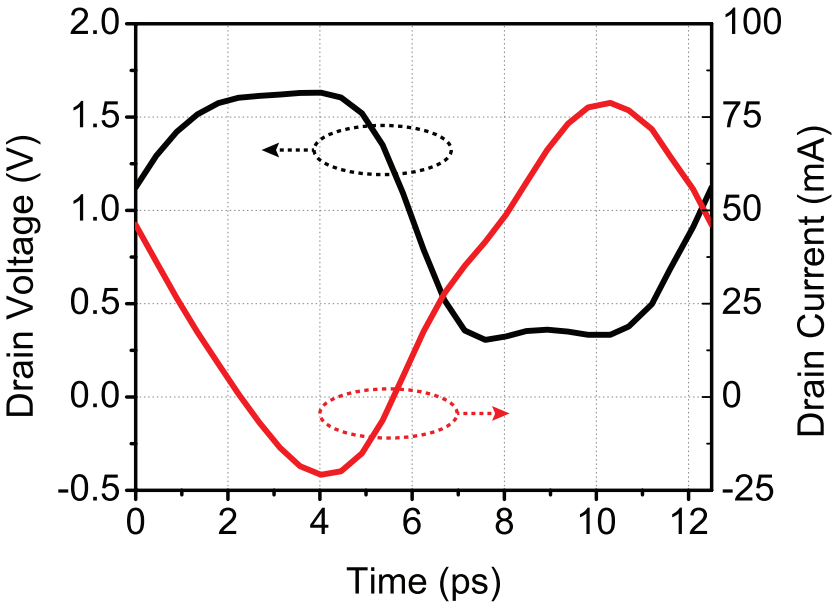


Figure 4.11: Simulation results of the 80 GHz power amplifier.

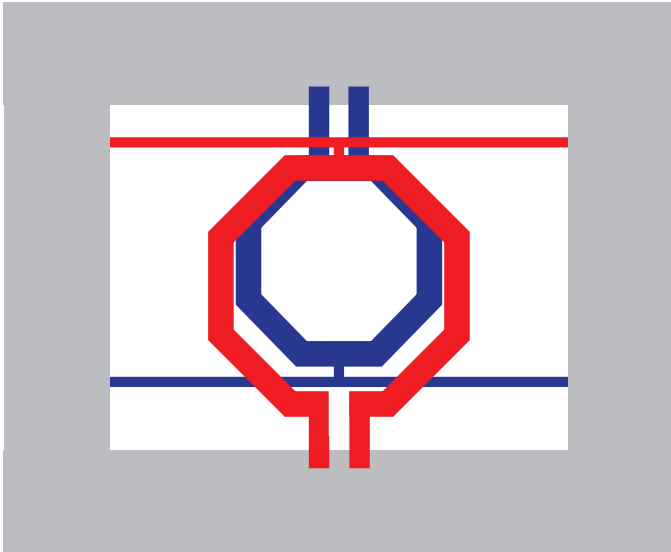


Figure 4.12: Inter-stage matching transformer for 80 GHz amplifiers (not to scale).

bandwidth and the stability. The amplifier chain is stable in the small-signal and large-signal simulations.

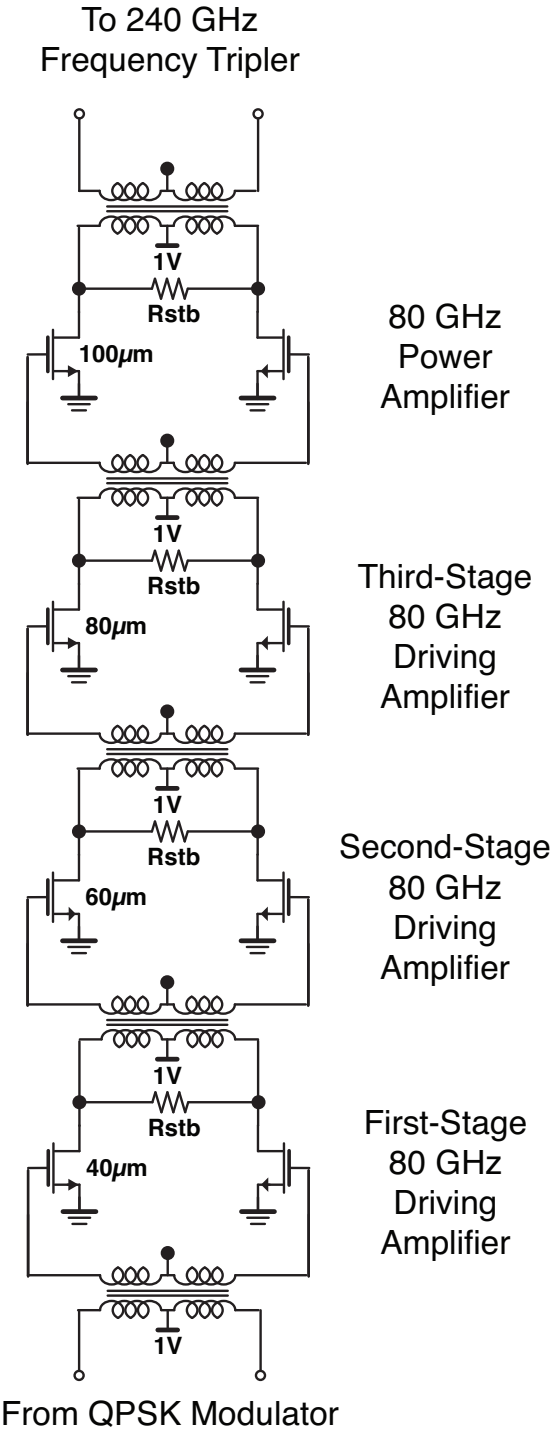


Figure 4.13: Schematic of the 80 GHz amplifiers of the 240 GHz QPSK/BPSK transmitter.

## 4.6 240 GHz Frequency Tripler

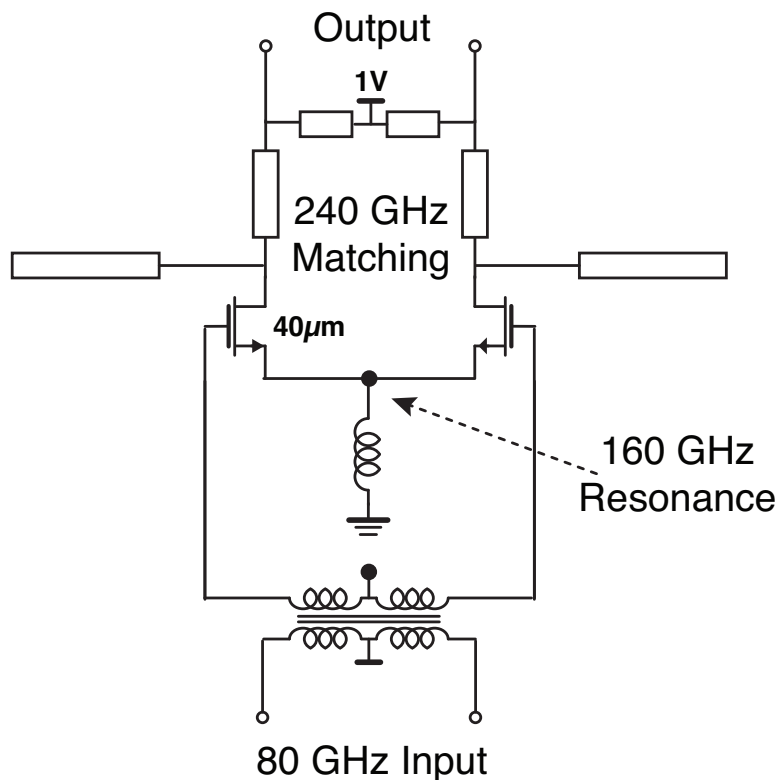


Figure 4.14: Schematic of the 240 GHz frequency tripler.

A 240 GHz wideband frequency tripler is designed for a 240 GHz QPSK/BPSK transceiver, which is demonstrated in Chapter 8. As discussed in Chapter 3, the frequency tripler simultaneously uses two frequency-multiplication methods: harmonic matching and an up-converting mixer. The schematic of the frequency tripler is shown in Fig. 4.14. The 80 GHz input is fed by the 80 GHz switching power amplifier described in the previous section. The PA has a low output impedance because of the large device size, but the tripler has a high input impedance because of the small size ( $40\ \mu\text{m}$ ); thus, in the transformer design, the transformer ratio is approximately 1:2, and the coupling coefficient is approximately 0.6. The transistors are biased to enhance the third harmonic tone. The tail inductor is attached to the common source node to resonate the second harmonic (160 GHz), and the transistors act as an up-converting mixer. The tripler output is matched at the third harmonic (240 GHz) using microstrip lines. The layout design is shown in Fig. 4.15. The red rectangles represent the NMOS transistors, and the green line represents the tail inductor (microstrip line). The blue line represents the supply line that is connected at the center, and the black lines represent the microstrip lines. The top two lines are connected to the differential antenna.

The output power (at the PA drains) is plotted in Fig. 4.16, and the saturated output power is approximately 13.5 dBm. The loss of the input matching network is 1.5 dB, and the transmitter output power (at the antenna input) is approximately 1 dBm in simulation. The conversion loss, including the input matching and the output matching, is 12.5 dB, and the conversion loss of the core tripler is approximately 9 ~ 10 dB. The undesired fundamental tone and the 160 GHz tone are rejected by the output matching network and the 240 GHz on-chip antenna. The power dissipation of the amplifiers and the tripler is 170 mW for a modulator output power of 0 dBm.

This frequency tripler is used for high data rates and should therefore have a wide bandwidth. Transient simulations are performed using a PRBS generator, a differential hybrid, a modulator, amplifiers and this tripler. Thus, in simulation, a data rate of up to 24 Gbps is achieved in the QPSK modulation.

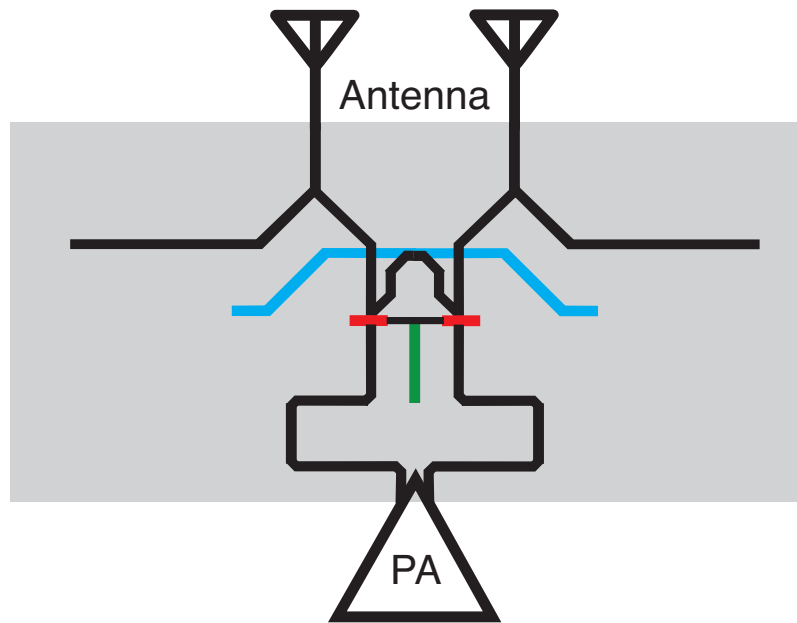


Figure 4.15: Layout design of the 240 GHz frequency tripler.

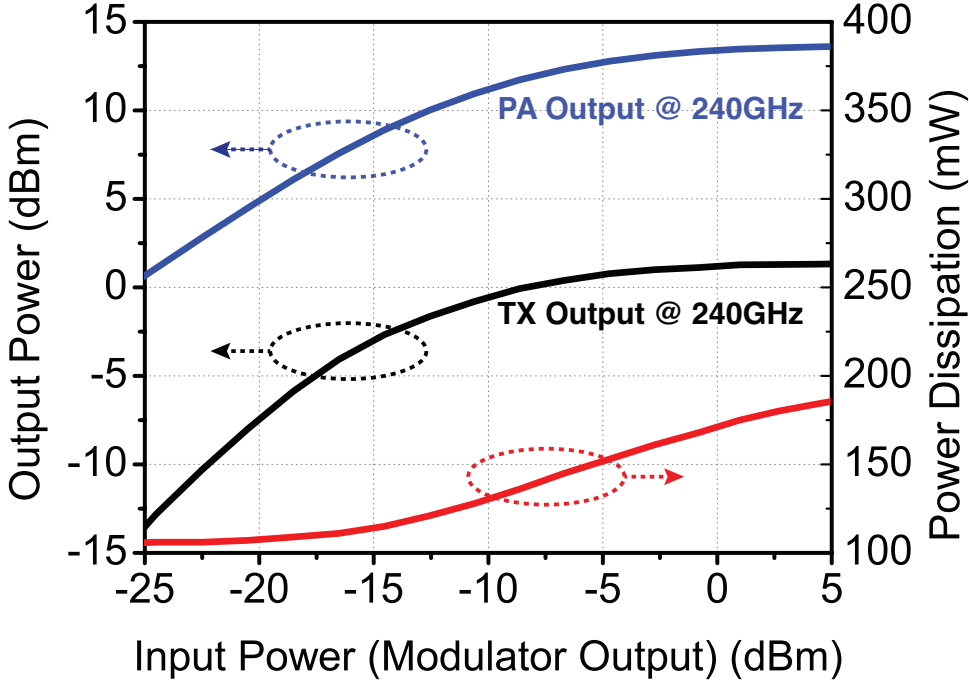


Figure 4.16: Simulation results of the 240 GHz transmitter and frequency tripler.



# Chapter 5

## Receiver

This chapter discusses receiver architectures and building blocks for sub-terahertz communications. The receiver should achieve low noise figure, high conversion gain/efficiency, and wide bandwidth. If the carrier frequency is higher than  $f_{max}$ , then no amplification can be achieved. Thus, a low-noise amplifier, conventionally employed by RF receivers, cannot be used in sub-terahertz receivers. Various mixer-first receiver architectures are discussed to find an optimal way to down-convert the RF signal. As a building block, a 260 GHz heterodyne down-converting mixer is demonstrated. In addition, receiver demodulation is discussed and analyzed. As receivers deal with very small input power, any noise or leakage from other circuits can impact SNR and BER. This chapter thereby provides design considerations to avoid the noise/leakage issues.

### 5.1 Conventional Receiver Architecture

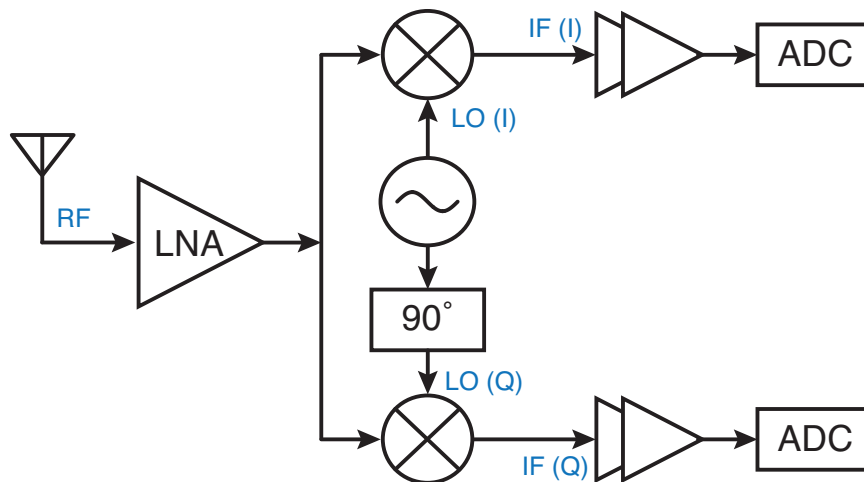


Figure 5.1: Conventional receiver architecture for I/Q modulations.

Fig. 5.1 illustrates a typical architecture of I/Q receivers. The RF input signal is first amplified by an low-noise amplifier (LNA) to minimize the overall noise figure, and the LNA output is then down-converted to I and Q mixers. The mixers use I and Q phases of the LO signal generated by an oscillator. The down-converted IF (intermediate frequency) signals are amplified and then converted by ADCs.

## 5.2 Sub-Terahertz Receiver Architectures

In typical receivers, an LNA is used right at the front-end to minimize the noise figure. However, an amplifier is lossy if the carrier frequency is higher than  $f_T/f_{max}$  and therefore not effective. Thus, down-converting should be first performed at the receiver front-end, which is typically called a mixer-first architecture. The down-converting is also lossy, and the conversion loss should be low. The minimum detectable signal power is limited by the conversion loss. The significantly high channel loss also results in low input power and SNR. The bandwidth of the receiver path should be sufficiently wide to down-convert the modulated signal. Several receiver architectures are discussed below.

### 5.2.1 Diode

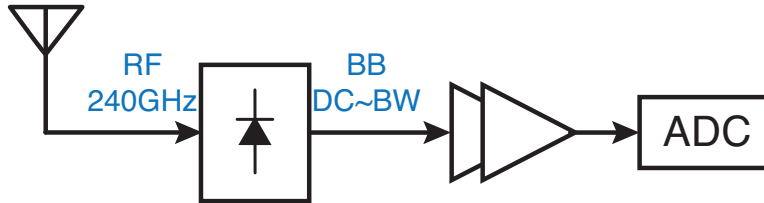


Figure 5.2: Receiver architecture with a diode detector.

Diodes are widely used in optical receivers. [49], [50] have proposed using a Schottky barrier diode to detect millimeter-wave signals. [49] defines the cut-off frequency of the Schottky diode as follows:

$$f_{cutoff} = \frac{1}{2\pi R_s C_o} \quad (5.1)$$

where  $R_s$  is the series resistance, and  $C_o$  is the zero-bias junction capacitance plus the parasitic capacitance. The cut-off frequencies of the diodes go up to 2 THz in 130 nm CMOS process, and the cut-off frequencies can go beyond 3 THz in advanced CMOS processes. [50] has demonstrated a 280 GHz detector using a Schottky diode. The amplitude modulation is used with a modulation frequency of 1 MHz. For a received power of  $0.33 \mu\text{W}$  ( $= -35 \text{ dBm}$ ), the output RMS voltage is 11.9 mV. The noise equivalent power (NEP) is estimated at  $36 \text{ pW}/\sqrt{\text{Hz}}$ , which is dominated by the flicker noise. For multi-Gbps communications, the responsivity should be improved and a wideband design is required.

### 5.2.2 Self-Mixer

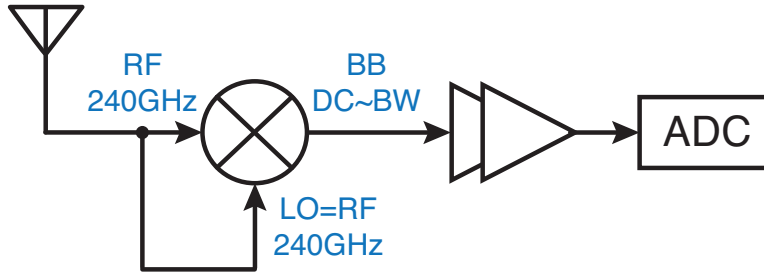


Figure 5.3: Receiver architecture with a self-mixer.

The next option is to self-mix the input signal as shown in Fig. 5.3. As for the diode, the self-mixer consists of one or two transistors. In addition, an LO signal is not necessary; thus, the design is simple, the power dissipation is low, and a small area is occupied. The self-mixer has disadvantages as well as the merits. First, the self-mixer requires a high signal power. The self-mixing results in an output swing that is proportional to the square of the amplitude. At sub-terahertz frequencies, a high transmit power and a high antenna gain are required because of the high channel loss. Otherwise, the output power is too small to detect. In addition, the self-mixer can be used only for ASK (amplitude-shift keying) or OOK (on-off keying) modulation, and other phase modulations cannot be used.

### 5.2.3 Sub-harmonic Mixer

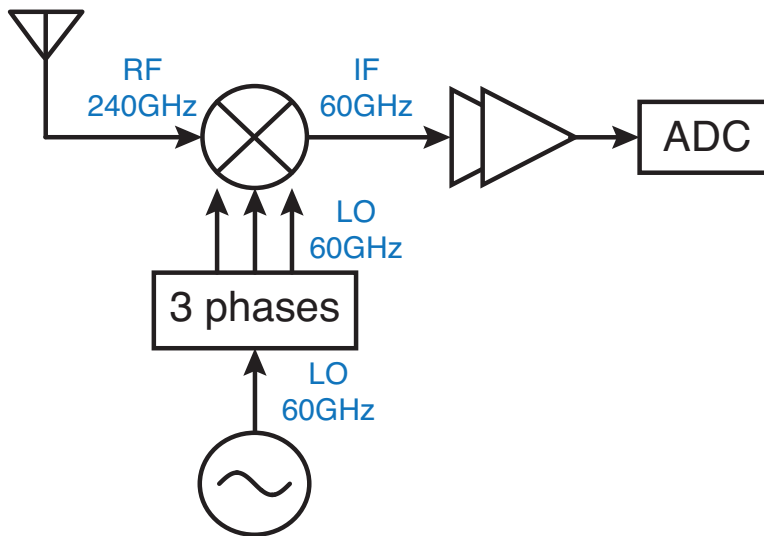


Figure 5.4: Receiver architecture with a sub-harmonic mixer.

Fig. 5.4 shows a sub-harmonic mixer. The sub-harmonic mixer inherently has  $N$ -push operations. When the mixer takes three different phases, the third-harmonic tone is generated inside by the non-linear operations, and the input RF signal is then down-converted. The sub-harmonic mixer does not require a direct 180 GHz LO signal; thus, the LO generation is relaxed. However, the conversion gain of the  $N$ -push operation and the conversion gain of the mixer should be both taken into account. Analysis of the  $N$ -push operation shows that the conversion gain is low for  $N$  above two. The mixer is also lossy when the RF frequency exceeds  $f_T/f_{max}$ . Thus, the overall conversion gain of the sub-harmonic mixer is relatively low with a reasonable LO power. The noise figure is accordingly high. Another problem is created by the LO-to-IF feedthrough. If the LO frequency is the same as the IF frequency, as shown in Fig. 5.4, then leakage can be problematic. Finally, the device size is small at such high frequencies, and the mismatches degrade the  $N$ -push operation and the mixer performance.

#### 5.2.4 Heterodyne Conversion Mixer

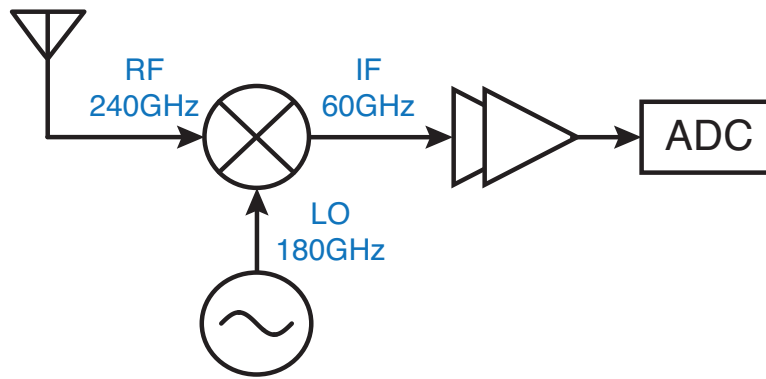


Figure 5.5: Receiver architecture with a heterodyne conversion mixer.

Fig. 5.5 shows a heterodyne mixer. The RF frequency is 240 GHz, and the LO frequency is 180 GHz. Thus, the IF frequency is 60 GHz. Unlike the sub-harmonic mixer, the mixer does not need to generate a higher harmonic tone. Thus, the high LO power can produce a high conversion gain and a low noise figure. However, the disadvantage is that a high-frequency LO is required. An output power above 0 dBm is typically required to achieve an acceptable conversion gain and noise figure. In CMOS, it is challenging to efficiently generate such a high output power at 180 GHz. A frequency doubler or tripler can be employed with a 90 GHz or 60 GHz power amplifier, respectively. In addition, the IF frequency should be appropriate for the signal bandwidth. If the bandwidth is too wide, efficient IF amplifiers are difficult to design and implement.

A 260 GHz heterodyne mixer is designed and fabricated in 65 nm CMOS with a 260 GHz OOK transceiver, which is demonstrated in Chapter 7. Fig. 5.6 shows that a double-balanced

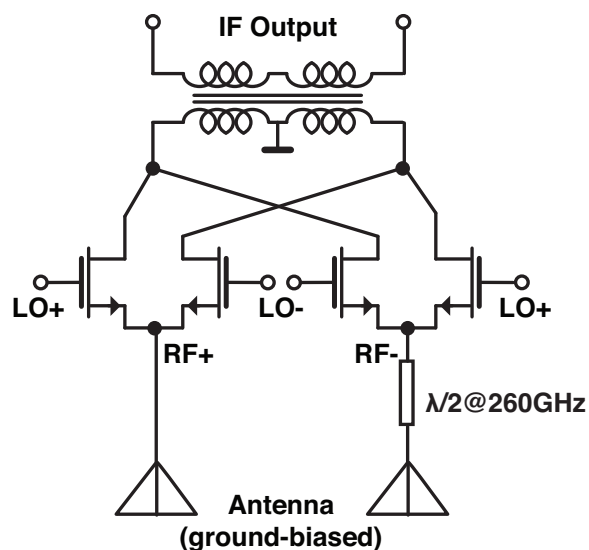


Figure 5.6: Schematic of the 260 GHz heterodyne mixer.

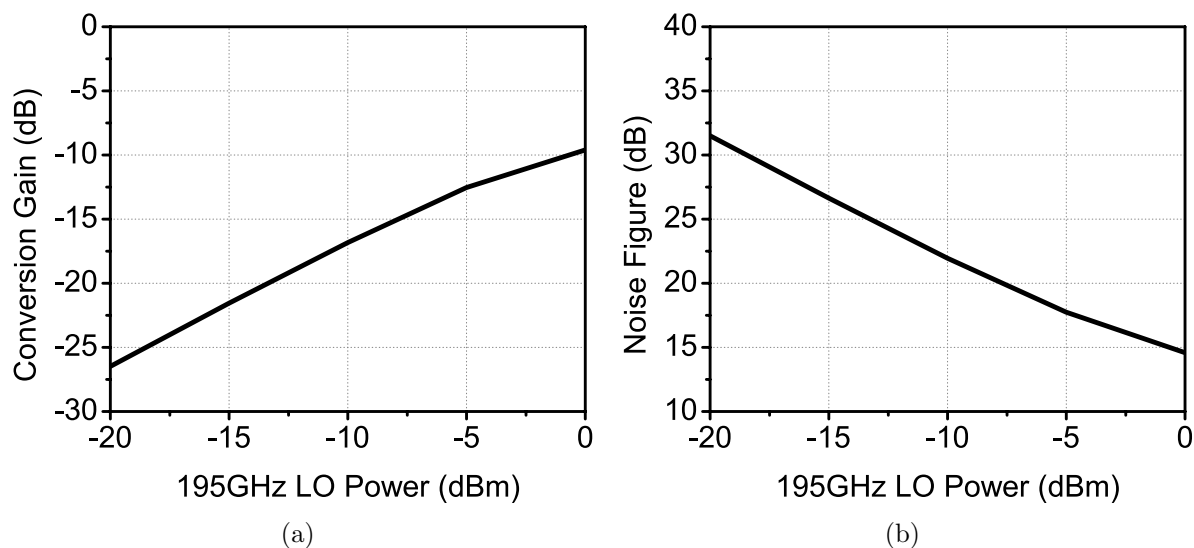


Figure 5.7: Simulation results of the 260 GHz heterodyne mixer (a) conversion gain (b) double sideband noise figure.

mixer is employed. The RF frequency is 260 GHz, and a differential input is realized using a  $\lambda/2$  delayed line. The receiver IF frequency is 65 GHz. The LO frequency is 195 GHz, which is generated by tripling a 65 GHz signal. If the LO has a parasitic 65 GHz tone, then the tone may degrade or contaminate the IF output; therefore, it should be properly attenuated or filtered out. The mixer uses class-C biasing. The source and drain voltages are 0 and 1 V, respectively. The gate is biased at 300 mV, around the threshold voltage, to improve

the conversion gain and the noise figure. Each device has a channel length of 60 nm and a total width of 10  $\mu\text{m}$ . The device size is limited by the high RF frequency. The LO power essentially determines the mixer performance. Fig. 5.7 shows that for an LO power of 0 dBm, the conversion gain is  $-10$  dB, and the (double sideband) noise figure is 15 dB.

### 5.2.5 Direct Conversion Mixer

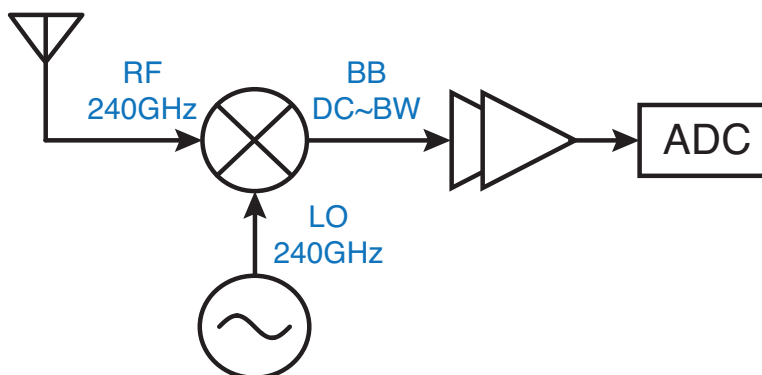


Figure 5.8: Receiver architecture with a direct-conversion mixer.

Fig. 5.8 shows a direct conversion mixer, that is widely used for I/Q modulations such as BPSK, QPSK, and QAM. Only one path (I or Q) is shown for simplicity. This direct conversion requires a higher LO frequency than the heterodyne mixer, but the IF design is relaxed. Because the IF frequency is zero, baseband amplifiers can be employed at the mixer output to improve the power efficiency. The increased LO frequency can be addressed by increasing the multiplication ratio of the frequency multiplier or increasing the oscillator frequency with a fixed multiplication ratio. The mixer performance (the conversion gain and the noise figure) depends on the LO power, as in the heterodyne mixer. In addition, this architecture does not suffer from LO-to-IF leakage.

A 240 GHz direct conversion mixer is designed and implemented in CMOS for the 240 GHz QPSK/BPSK transceiver. The designs and results are demonstrated in [2]. For an LO power of  $-3$  dBm, the conversion gain is  $-3$  dB and the (double sideband) noise figure is 9 dB.

## 5.3 OOK Demodulator

The non-coherent OOK modulation requires an envelope detector to extract the digital data. A passive non-linear diode is usually employed, but its size is limited by the operation frequency. Moreover, the output current is typically low; therefore, accumulating or averaging is required to achieve the desired SNR. As such, it is not appropriate for high-speed communications. In CMOS technology, the  $N$ -push operation can be used to detect the input amplitude. The output is not the  $N$ -th harmonic but the baseband (around DC)

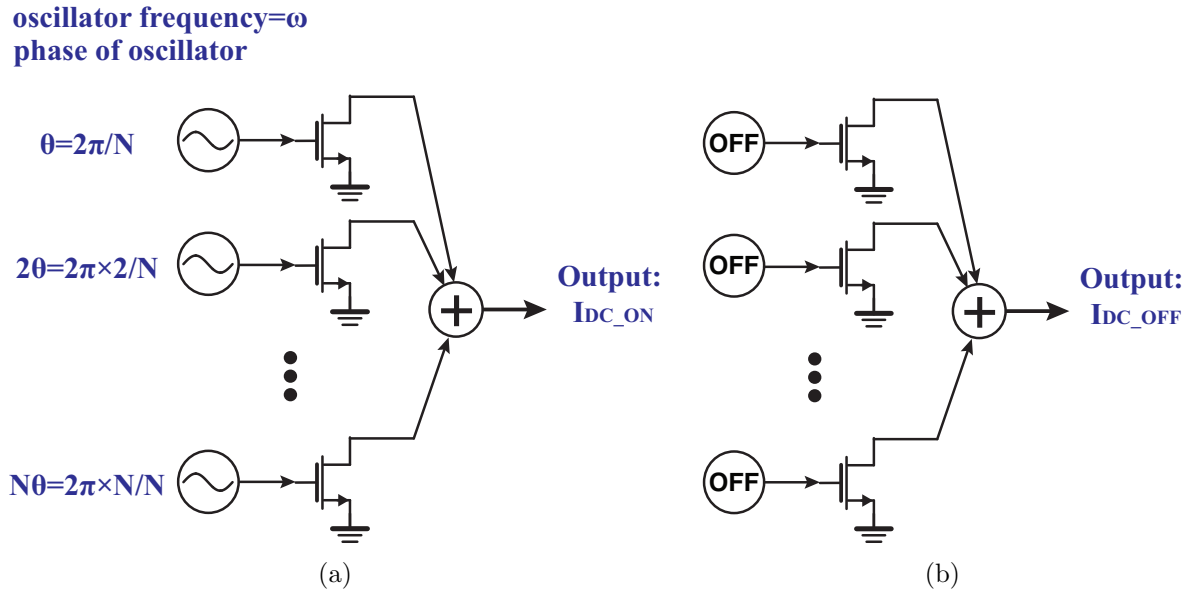


Figure 5.9: Operation of an OOK demodulator (a) ON state (b) OFF state.

current, varying with the input amplitude, the device gain, and the biasing point. Because of the OOK modulation, the input has two states (ON and OFF), and the output current has two values ( $I_{DC,ON}$  and  $I_{DC,OFF}$ ), analyzed in this section. Fig. 5.9 illustrates the OOK demodulation using the  $N$ -push operation. Here, the input phases are not necessarily equally distributed because the output is the baseband current.

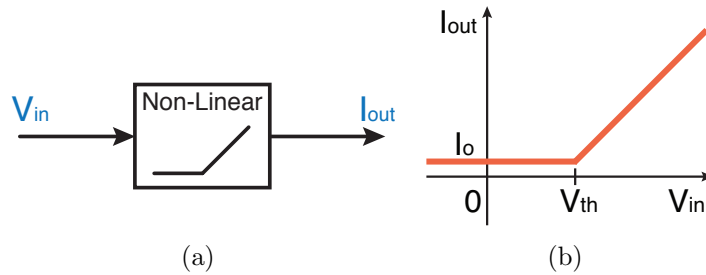


Figure 5.10: An ideal non-linear block with sub-threshold current.

$$I_{out} = \begin{cases} I_o + g_m(V_{in} - V_{th}), & (V_{in} > V_{th}) \\ I_o, & (otherwise) \end{cases} \quad (5.2)$$

The non-linear block is modeled similarly to Equation 3.28. The only difference is the presence of the sub-threshold current ( $I_o$ ). As presented in Equation 5.2,  $I_o$  is added for all

of the input voltages. To analyze  $I_{DC,ON}$  and  $I_{DC,OFF}$ , two cases are considered. The first case is where the bias voltage is lower than the threshold voltage, as shown in Equation 5.3.

$$V_b < V_{th} \quad (5.3)$$

$I_{DC,ON}$  can be calculated using Equation 3.37 and is given by

$$I_{DC,ON} = I_o + Ng_0V_{in} = I_o + \frac{Ng_mV_{in}}{\pi} \left( \sin\left(\frac{\alpha}{2}\right) - \frac{\alpha}{2} \cos\left(\frac{\alpha}{2}\right) \right) \quad (5.4)$$

Because the bias voltage is lower than the threshold voltage, the off-state output current is given by

$$I_{DC,OFF} = I_o \quad (5.5)$$

The difference between  $I_{DC,ON}$  and  $I_{DC,OFF}$  is given by

$$I_{DC,ON} - I_{DC,OFF} = \frac{Ng_mV_{in}}{\pi} \left( \sin\left(\frac{\alpha}{2}\right) - \frac{\alpha}{2} \cos\left(\frac{\alpha}{2}\right) \right) \quad (5.6)$$

The ratio between  $I_{DC,ON}$  and  $I_{DC,OFF}$  is given by

$$\frac{I_{DC,ON}}{I_{DC,OFF}} = 1 + \frac{Ng_mV_{in}}{\pi I_o} \left( \sin\left(\frac{\alpha}{2}\right) - \frac{\alpha}{2} \cos\left(\frac{\alpha}{2}\right) \right) \quad (5.7)$$

Conversely, if the bias voltage is higher than the threshold voltage as shown in Equation 5.8,  $I_{DC,ON}$  and  $I_{DC,OFF}$  can be calculated similarly.

$$V_b > V_{th} \quad (5.8)$$

$I_{DC,ON}$  is the same as Equation 5.4.

$$I_{DC,ON} = I_o + Ng_0V_{in} = I_o + \frac{Ng_mV_{in}}{\pi} \left( \sin\left(\frac{\alpha}{2}\right) - \frac{\alpha}{2} \cos\left(\frac{\alpha}{2}\right) \right) \quad (5.9)$$

When the input signal is off, the output current is dependent of the bias point.

$$I_{DC,OFF} = I_o + Ng_m(V_b - V_{th}) = I_o - Ng_mV_{in} \cos\left(\frac{\alpha}{2}\right) \quad (5.10)$$

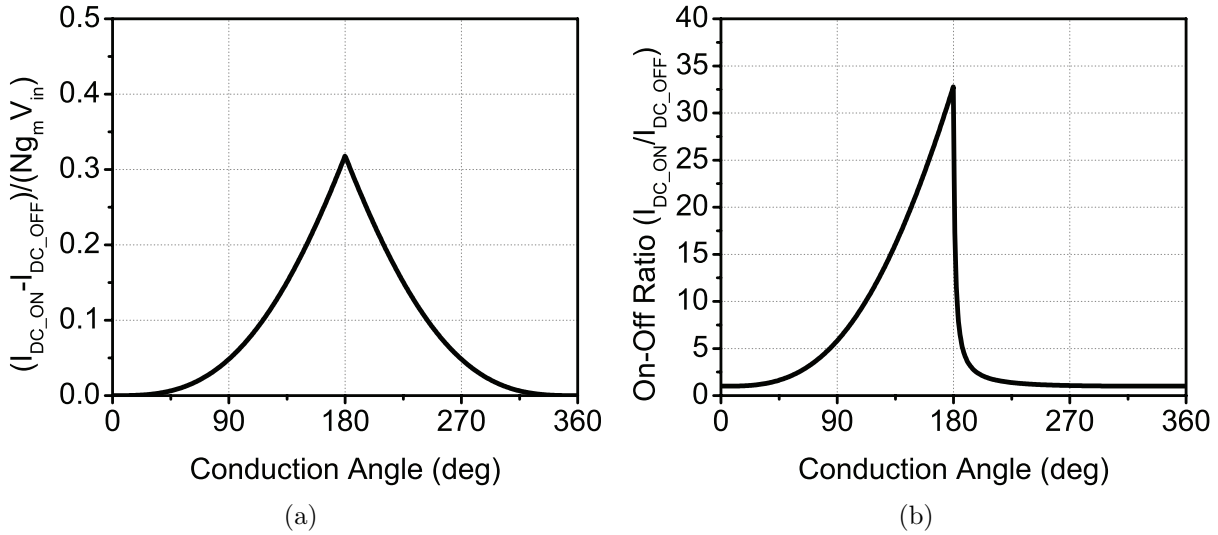
The difference between  $I_{DC,ON}$  and  $I_{DC,OFF}$  is given by

$$I_{DC,ON} - I_{DC,OFF} = \frac{Ng_mV_{in}}{\pi} \left( \sin\left(\frac{\alpha}{2}\right) + \left(\pi - \frac{\alpha}{2}\right) \cos\left(\frac{\alpha}{2}\right) \right) \quad (5.11)$$

If the sub-threshold current is assumed to be negligible (compared to the input signal) as

$$\frac{I_o}{Ng_mV_{in}} \ll 1 \quad (5.12)$$



Figure 5.11: Comparison of  $I_{DC,ON}$  and  $I_{DC,OFF}$ .

then the ratio between  $I_{DC,ON}$  and  $I_{DC,OFF}$  can be approximated by

$$\frac{I_{DC,ON}}{I_{DC,OFF}} \approx \frac{1}{-\pi \cos\left(\frac{\alpha}{2}\right)} \left( \sin\left(\frac{\alpha}{2}\right) - \frac{\alpha}{2} \cos\left(\frac{\alpha}{2}\right) \right) = \frac{1}{\pi} \left( \frac{\alpha}{2} - \tan\left(\frac{\alpha}{2}\right) \right) \quad (5.13)$$

The sub-threshold current is assumed to be as small as

$$\frac{I_o}{Ng_m V_{in}} = \frac{1}{100} \quad (5.14)$$

then the difference and the ratio can be plotted, as shown in Fig. 5.11. Unlike the  $N$ -push frequency multiplier, the optimum point does not depend on  $N$ , which represents the number of non-linear blocks. In the demodulator,  $V_b$  should be  $V_{th}$  because  $180^\circ$  of the conduction angle gives the highest on-off difference.

## 5.4 Leakage-Free Receiver Design

In receivers, the input power is typically lower than  $-40$  dBm due to the high channel loss. The input signal can easily be contaminated by other leakages or noises. Therefore, the receiver architecture and building blocks should be carefully planned, implemented, and verified. It is analogous to sensitive analog circuits surrounded by noisy digital circuits. The analog circuits should be guarded by substrate contacts and separate supply lines. VCO pulling is another example. If two different VCOs are coupled through the substrate or parasitics, the VCO frequencies are not stable and can have many side tones.

Here are some guidelines for millimeter-wave/sub-terahertz receivers:

1. Single-ended circuits: Single-ended circuits are affected by common-mode noise, which is from supply/ground lines, substrate, or any adjacent circuits. Because the receiver is for high-speed digital data communication, digital circuits can be integrated close to the receiver. It is well known that differential circuits are not sensitive to common-mode noise; therefore, fully differential circuits are preferred. For single-ended circuits, the return path should be clearly defined, and bypassing the supply and ground should be performed well to provide a return path and to reduce noise.
2. Large-signal blocks: Oscillators and amplifiers are usually large-signal blocks because the output power is significantly higher than the receiver input power. The output power of the power amplifiers can be higher than 10 dBm. The high-power signal can leak to the signal path of the receiver, thereby degrading SNR and the dynamic range. For example, the receiver IF frequency is 100 GHz, and the oscillator frequency is also 100 GHz, then the large-signal output of the oscillator can contaminate the small-signal receiver signal, unless the oscillator signal is blocked well. Thus, large-signal blocks should be placed far from the receiver signal path and the leakage should be properly eliminated.
3. Harmonic leakages: In sub-terahertz transceivers, frequency multipliers are often employed and undesired harmonics should be properly attenuated. For example, the output of a frequency tripler can have a strong fundamental tone and a second-harmonic tone, as well as the desired third-harmonic tone. The fundamental and second-harmonic tones can affect following circuits and the receiver signal path through leakage. If the receiver IF frequency is 100 GHz, then a 150 GHz oscillator and a frequency doubler is preferred to generate 300 GHz rather than a 100 GHz oscillator and a frequency tripler.

# Chapter 6

## Baseband Circuits

This chapter describes baseband circuits for high-speed applications. A high data rate requires wide bandwidth and low noise, which make it challenging to design baseband circuits. As clock speed increases to 10 GHz, it is not trivial to ensure that digital logic gates operate correctly. Moreover, it is difficult to achieve efficient circuits at such a high clock rate. In some millimeter-wave/sub-terahertz systems, the overall performance (efficiency, data rate, dynamic range, etc.) can be limited by baseband circuits such as ADC, DAC, and baseband amplifiers. First, this chapter introduces current-mode logic gates and then demonstrate PRBS (pseudo-random bit sequence) generators using the logic gates. Second, a broadband amplifier is presented for high-speed communications. Lastly, an operational transconductance amplifier is discussed.

### 6.1 High-Speed Logic Gates

The logic gates are NOT, AND, OR, XOR, multiplexer, latch, flip-flop, and so on, as well known. In CMOS technology, the logic gates can be implemented without static DC current, leading to efficient digital circuit design. However, power dissipated by capacitive switching and short-circuit current significantly increases with data rate or clock speed. In addition, the speed or bandwidth is bounded by ON-resistance and parasitic capacitance of devices. For sub-terahertz CMOS transceivers, current-mode logic (CML) circuitry is employed because it can reach higher data rate. Although CML requires static DC current, it is basically a differential circuit, immune to common-mode noise, and achieves a wide bandwidth to prevent baseband circuits from bounding the overall data rate. More importantly, CML generates less supply noise.

Fig. 6.1 to Fig. 6.5 show symbols and schematics of CML gates. Fig. 6.1 shows buffer and inverter. CML circuitry is fully differential; thus, an inverting stage can be implemented by simply swapping input nodes or output nodes. Thus, buffer and inverter have different symbols but the circuit schematics are the same except input/output connections. Because this is a digital gate, the output has two states  $(V_{DD}, V_{DD} - I_B R_D)$  and  $(V_{DD} - I_B R_D, V_{DD})$  and

the peak-to-peak swing is  $2I_B R_D$ . The static power consumption is  $V_{DD} I_B$ . Fig. 6.2(b) shows NAND, AND, NOR, and OR gates, depending on the input and output connections. Fig. 6.3 shows an XNOR and XOR gates. The multiplexer has two possible circuits. Fig. 6.4(b) is similar to Fig. 6.3(b), and Fig. 6.4(c) is a current-steering multiplexer used when the supply voltage is low, while requiring one more current source. The current-steering scheme can be applied to AND, OR, and XOR gates. Finally, Fig. 6.5 shows a D-latch. A D-flip-flop can be implemented by cascading two latches.

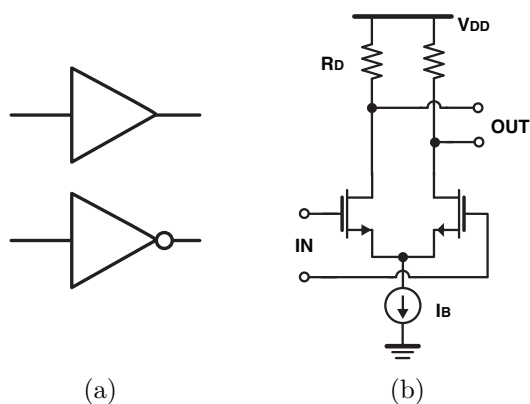


Figure 6.1: Buffer and inverter (a) symbol (b) CML.

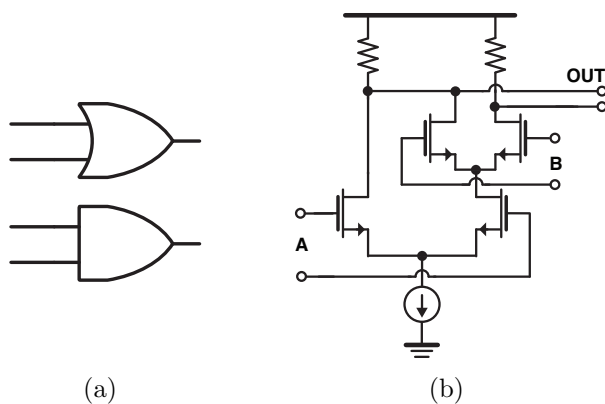


Figure 6.2: NAND/AND and NOR/OR (a) symbol (b) CML.

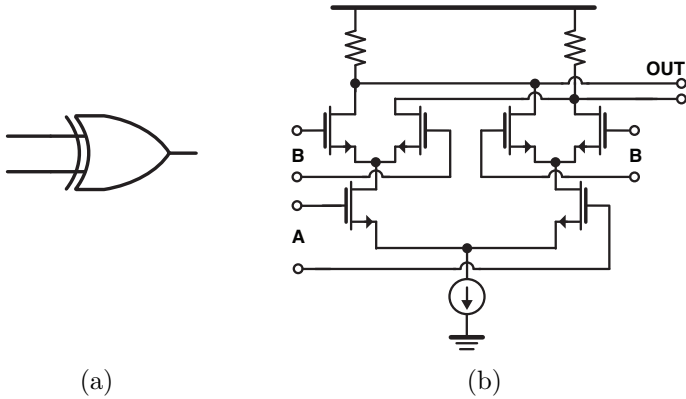


Figure 6.3: XNOR/XOR (a) symbol (b) CML.

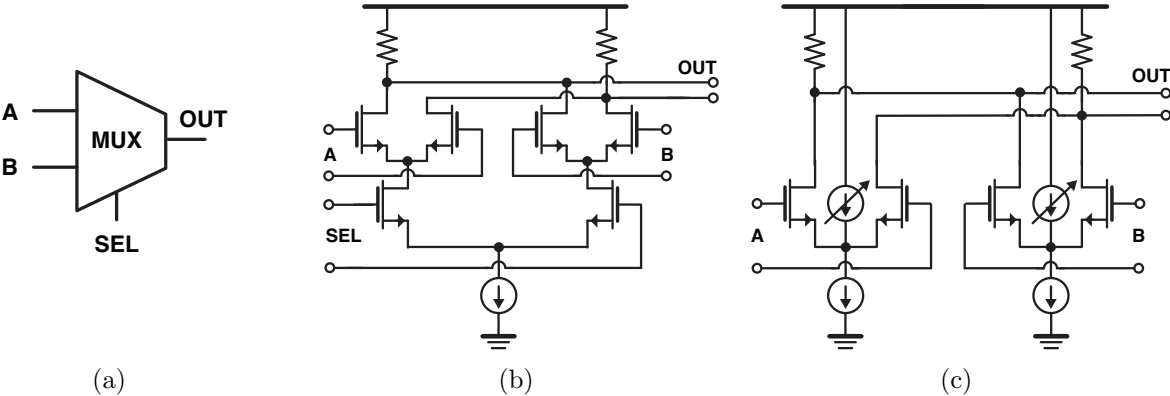


Figure 6.4: Multiplexer (a) symbol (b) CML (c) current-steering multiplexer.

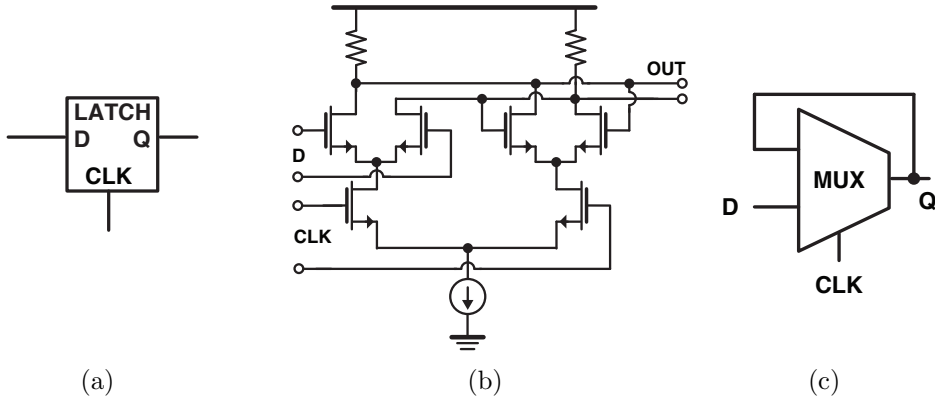


Figure 6.5: Latch (a) symbol (b) CML (c) an alternative representation.

## 6.2 PRBS Generator

In real communication systems, a transmitter sends application data which may be a video/audio stream, file data, or any system control bits. Thus, the transmitter sometimes sends designated bit patterns, such as preamble or packet header, but transmits arbitrary bits most of the time. To test and measure a wireless transceiver, integrating all of the layers up to the application layer is time-consuming and not efficient. Therefore, for the transmit data, a PRBS is widely used to test most of the bit transition patterns. The length ( $N$ ) of the PRBS bit pattern is chosen as  $N = 2^k - 1$ , where  $k$  is the number of bits in the linear feedback shift register. The PRBS is typically generated by a shift register and an XOR gate, and  $-1$  is for removing the all-zero state. For example,  $k = 7$ ,  $N = 127$ , and the polynomial is  $x^7 + x^6 + 1$ , which mathematically characterizes the shift register. The  $N$  number of bits are continuously repeating, and it is periodic and pseudo-random. The frequency spectrum of PRBS looks like a SINC function, but it is composed of discrete frequency tones because of its periodicity. The null locations are determined by symbol rate ( $R_{sym}$ ), and the beat frequency between two discrete tones depends on symbol rate and  $N$  of PRBS, as follows:

$$f_{null,1} = R_{sym} \quad (6.1)$$

which shows that the first null is placed at  $R_{sym}$ , and other nulls are at multiples of  $R_{sym}$ .

$$f_{beat} = \frac{R_{sym}}{N} = \frac{R_{sym}}{2^k - 1} \quad (6.2)$$

which shows that as  $N$  grows,  $f_{beat}$  decreases and the bit pattern is closer to random pattern.

The  $2^7 - 1$  PRBS generator is shown in Fig. 6.6, and the interleaved architecture is based on [51]. Each lane is composed of seven latches and one XOR gate. Two lanes are multiplexed at the end. The target data rate is 20 Gbps, and the corresponding clock frequency is 10 GHz. The CML gates, introduced in the previous section, are employed for high data-rate operations. If all of the latch outputs are zero, then PRBS is not initiated; therefore, a start-up switch should be designed to prevent the all-zero state, as shown in Fig. 6.6.

This PRBS generator is employed for a 260 GHz OOK transceiver, which is demonstrated in Chapter 7. The OOK modulator requires digital bits (0 V and 1 V), and the CML-to-CMOS conversion block is used, as shown in Fig. 6.7. The CMOS inverter is self-biased with a large resistor, and the biasing point is at the metastable point when the input AC signal is not present. Two OOK modulator switches are at the input and output of a power amplifier to increase the on-off ratio; thus, the driver (Fig. 6.7) has two outputs with a delay, which is matched to a delay of the power amplifier. The inverter sizes are properly designed, such that the fan-out is kept as low as two to prevent the rising/falling edges from being unsharpened.

Fig. 6.8 shows a PRBS generator for QPSK (two-bit modulation). Two sequences are generated by separate PRBS blocks. Note that the two start-up switches are differently

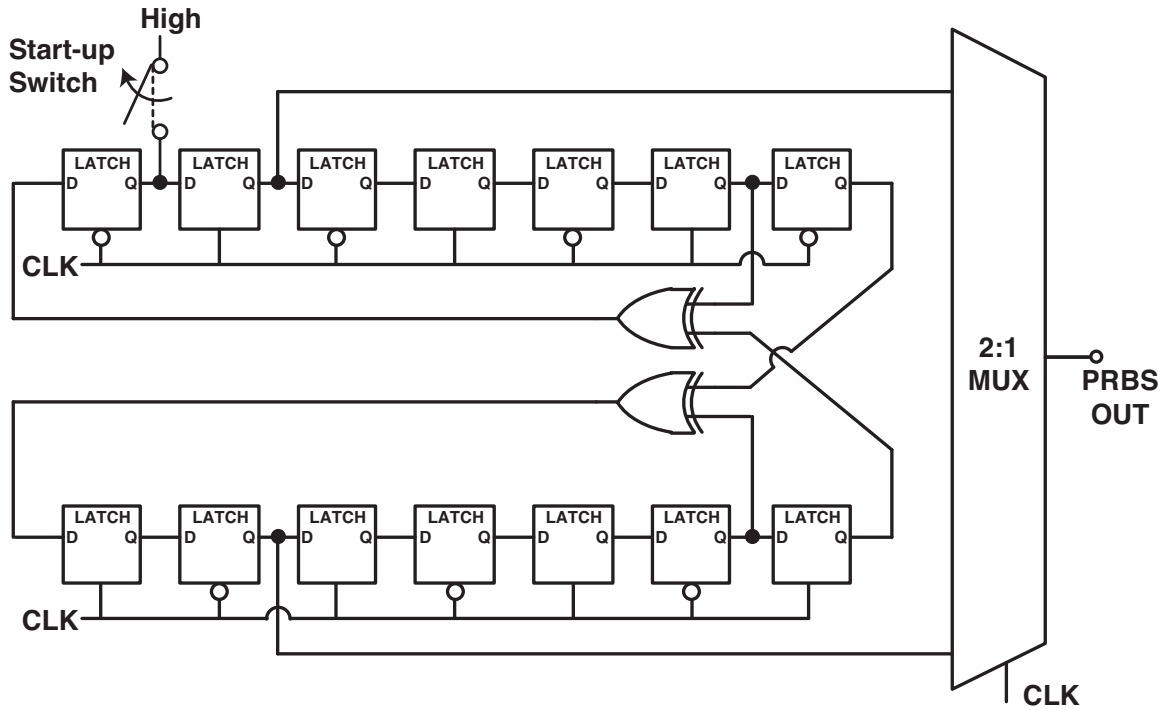


Figure 6.6: PRBS generator.

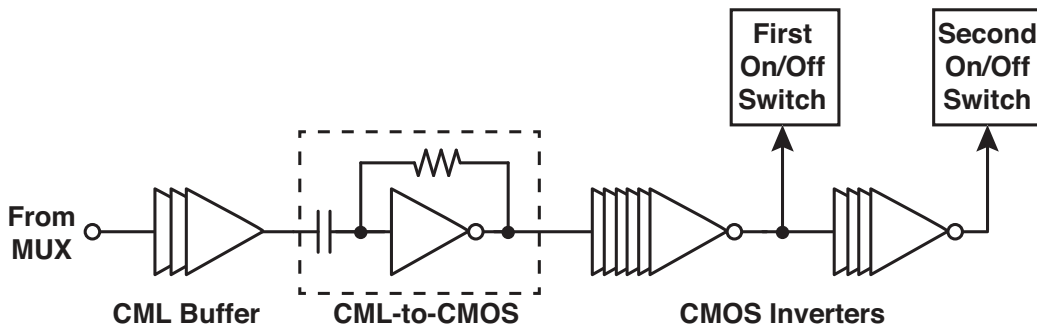


Figure 6.7: PRBS distributor for OOK modulation.

located in order that all of the transitions among the four states are equally distributed in the case of QPSK. To enable BPSK mode and continuous-wave mode, one multiplexer stage is added at the end. In the QPSK mode, the outputs have four states: (0, 0), (0, 1), (1, 0), and (1, 1). If the BPSK mode is selected, two multiplier outputs are identical, and only the (0, 0) and (1, 1) states are present. If two PRBS blocks are not initiated, then all of the latches have zero outputs, leading to continuous-wave mode or no modulation. This PRBS generator operates up to 40 Gbps in the QPSK mode and is integrated in a 240 GHz QPSK/BPSK transceiver, which is demonstrated in Chapter 8.

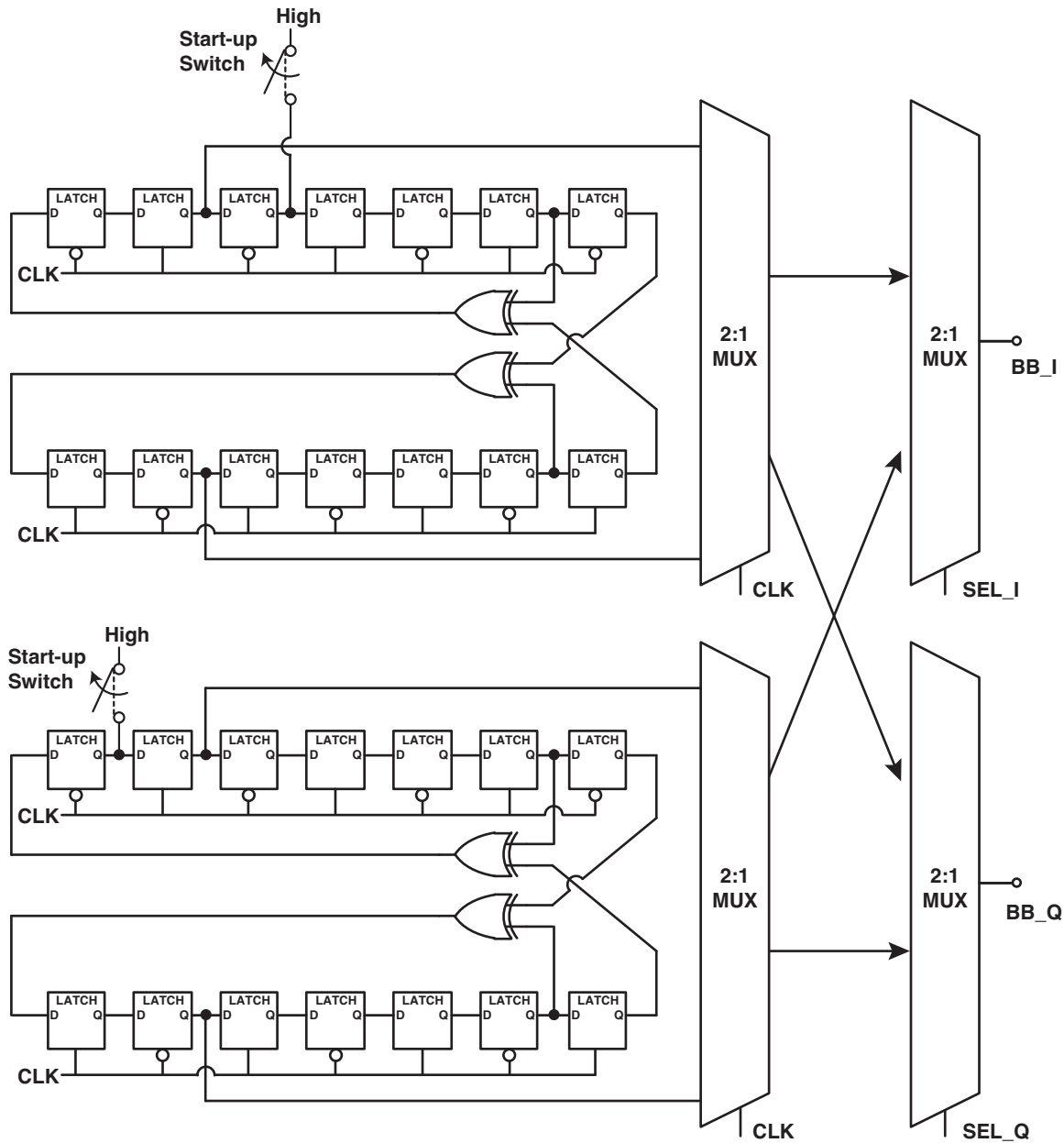


Figure 6.8: PRBS generator for QPSK/BPSK/continuous-wave modes.

In high-speed communications, the rise/fall time of PRBS output can affect SNR and BER. To reduce bit errors, the rise/fall time should be minimized and the rising/falling edge should be sharpened. Because the QPSK modulator is a Gilbert mixer, as shown in Fig. 4.8. the load is capacitive. In Fig. 6.9(b), without inductance  $L$ , the time constant is given by  $R_D(C_D + C_L)$ , where  $R_D$  is a drain load resistance of the output buffer,  $C_D$  is a drain parasitic capacitance of the output buffer, and  $C_L$  is an input capacitance of the



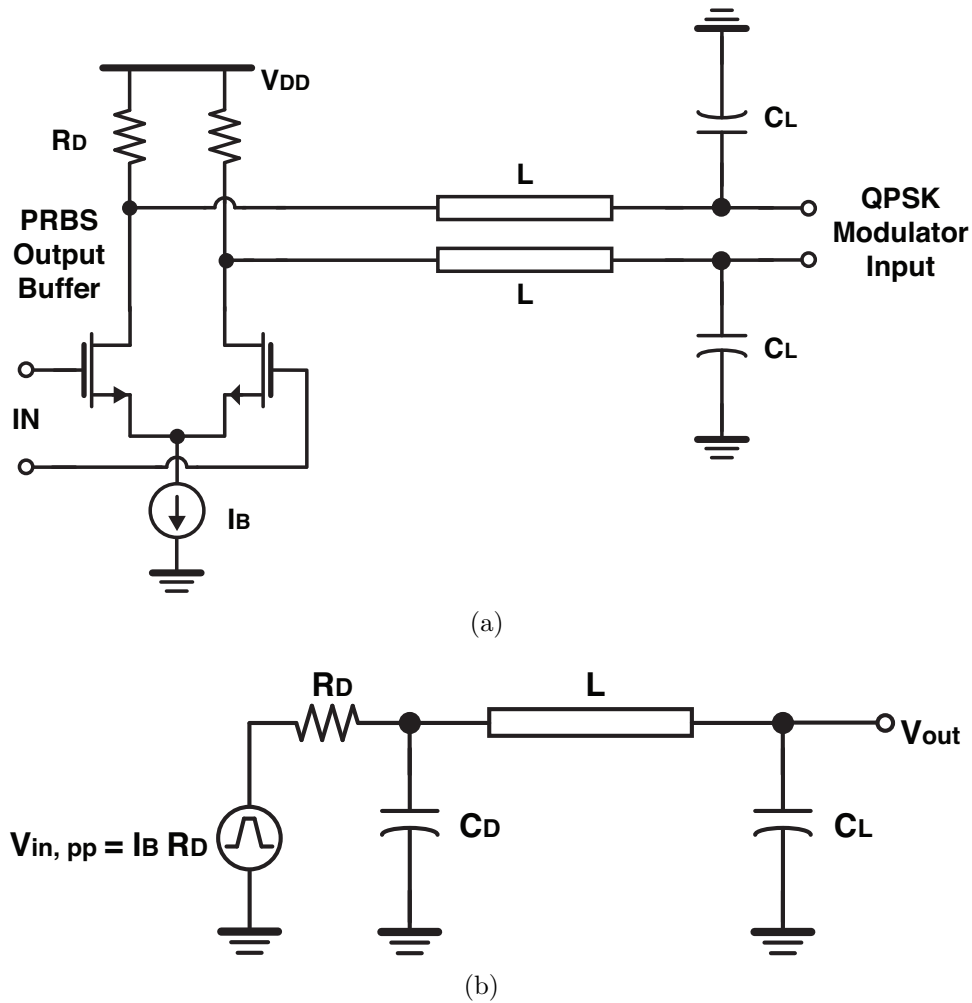


Figure 6.9: Output buffer of the PRBS generator (a) schematic (b) equivalent model.

QPSK modulator. By adding inductance  $L$ , the effective time constant can be reduced but parasitic resonance (ringing) occurs. The transfer function is given in Equation 6.3. In the real design,  $L$  is adjusted to the extent that the ringing does not significantly affect the QPSK modulator.

$$\frac{V_{out}}{V_{in}} = \frac{1}{1 + sR(C_D + C_L) + s^2LC_L + s^3RLC_D C_L} \quad (6.3)$$

### 6.3 Baseband Amplifier

A baseband amplifier is employed to amplify the received signal down-converted by a direct-conversion mixer in a 240 GHz QPSK/BPSK transceiver, which is demonstrated in

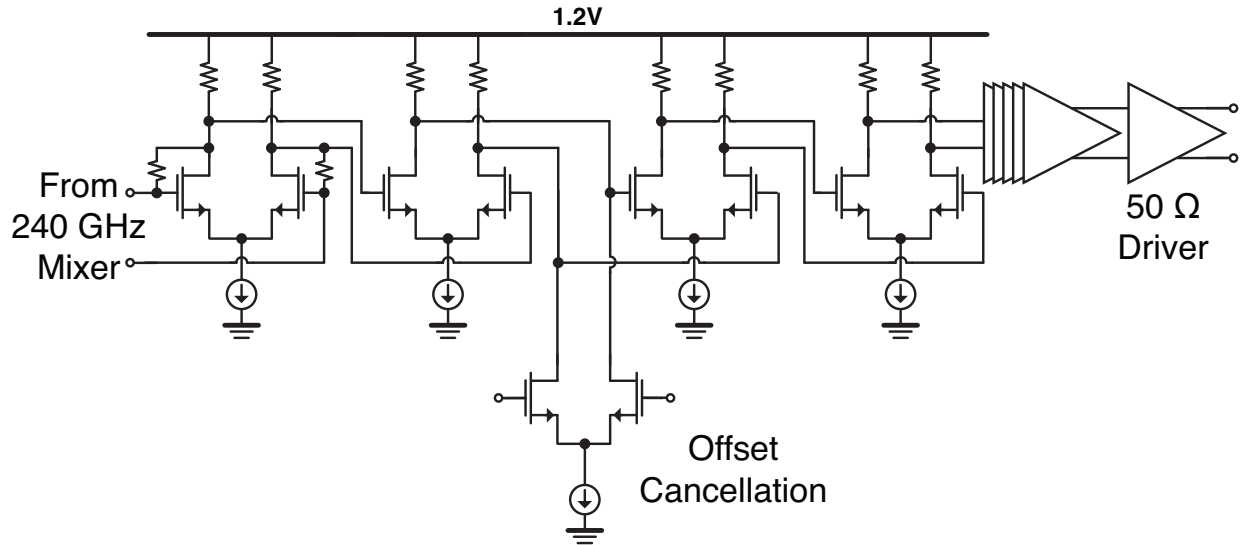


Figure 6.10: Schematic of the baseband amplifier.

Chapter 8. The baseband amplifier is required to have a high gain, low noise figure, wide bandwidth, and low power dissipation. The input power is determined by the transmit power, channel loss, and mixer conversion gain. Accordingly, the required gain is approximately 40 dB for reasonable output voltage swing on the oscilloscope. The noise figure should be minimized and the 3 dB bandwidth should be wider than 7 GHz to achieve a data rate of 20 Gbps in the QPSK modulation (or 10 Gbps in BPSK). The schematic is shown in Fig. 6.10. The baseband amplifier consists of nine differential pairs, and the last output amplifier is to drive two 50  $\Omega$  loads (differentially 100  $\Omega$ ) for measurements. Device sizes and tail currents are taper-designed to optimize the performance. In post-layout simulations, the baseband amplifier achieves a gain of 36 dB, as presented in Fig. 6.11(a). The noise figure is less than 4 dB over the bandwidth, as shown in Fig. 6.11(b). The power consumption is 40 mW with a supply voltage of 1.2 V.

The offset voltage of the differential amplifier originates from the mismatches of the two matched input devices and load resistors. The device mismatch results from variations in the threshold voltage ( $V_{th}$ ) and the current factor ( $\beta$ ). Typically, the variation in the threshold voltage dominates, and the variation is given by [52]

$$\sigma_{\Delta V_{th}} = \frac{A_{V_{th}}}{\sqrt{WL}} \quad (6.4)$$

where  $A_{V_{th}}$  is the Pelgrom constant for the threshold voltage, which is dependent on the process technology. According to [53],  $A_{V_{th}}$  is approximately 3.0 mV  $\mu\text{m}$  in 65 nm CMOS. The variation or offset voltage is inversely proportional to the square root of the channel area. In other words, to reduce the offset voltage, the channel length (L) and width (W) should be enlarged. In addition, if many amplifiers are cascaded, then the overall input-referred offset

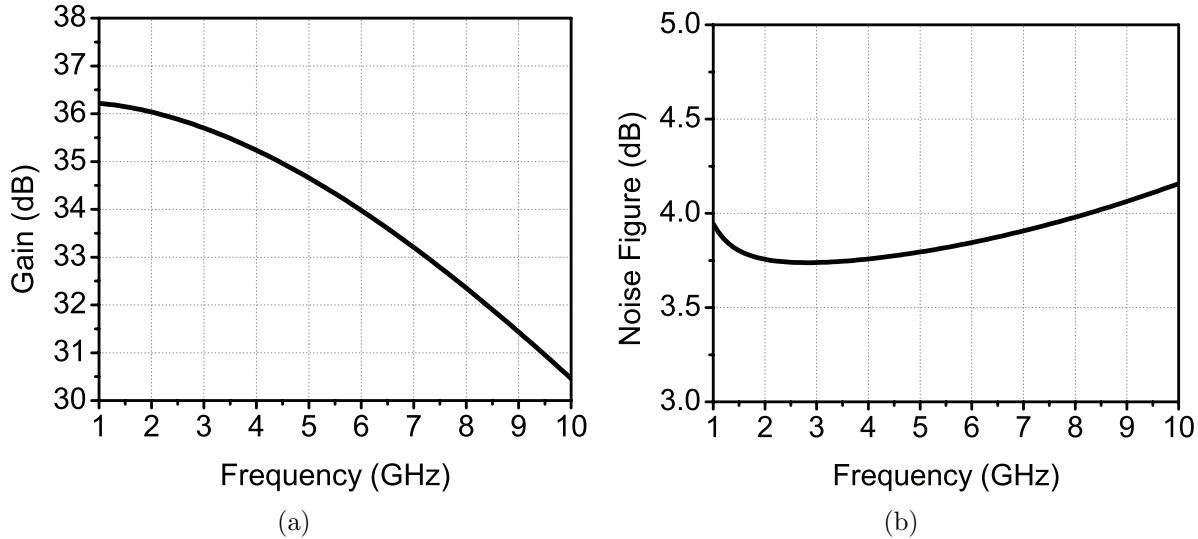


Figure 6.11: Simulation results of the baseband amplifier (a) gain (b) noise figure.

voltage is given by

$$V_{os} = V_{os,1} + \frac{V_{os,2}}{A_1} + \frac{V_{os,3}}{A_1 A_2} + \frac{V_{os,4}}{A_1 A_2 A_3} + \dots \quad (6.5)$$

where  $V_{os,n}$  and  $A_n$  are the input-referred offset voltage and the voltage gain of the  $n$ -th device, respectively. Thus, the first stage should have high gain and low offset voltage.

The offset cancellation is performed manually by a differential pair, as shown in Fig. 6.10. In simulation, the offset cancellation circuit is designed to correct the worst-case offset. If the size of the offset cancellation circuit is too large, the gain, bandwidth, and power consumption are degraded; thus, some design iterations are needed.

The overall noise figure is dominated by the first few stages in  $n$  cascaded (impedance-matched) amplifiers because the noise factor is given by

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \dots \quad (6.6)$$

where  $F_n$  and  $G_n$  are the noise factor and power gain of the  $n$ -th device, respectively. Although the impedances are not matched in this case, the gains of the first and second stages should be large to reduce both the offset voltage and noise figure. Basically, low noise figure can be achieved by increasing gain or  $g_m$  of the input devices, thereby increasing bias current. This is a trade-off between noise figure and power dissipation.

It is well known that there is a trade-off between gain and bandwidth. The gain is  $g_m R_L$  and the bandwidth is roughly  $1/R_L C_{in}$ ; thus, the gain-bandwidth product is given by  $g_m/C_{in}$ . Shunt peaking is a way to extend the gain-bandwidth product. According to [54], a shunt peaking can extend the bandwidth up to 85%, for a given gain. However, the inductor size is so large that it is not acceptable in some applications. Thus, active inductors can be employed without the area overhead.

## 6.4 Operational Amplifier

Operational amplifier is a key component for various applications including communication systems [44]. Its high input impedance and high gain enable plenty of circuit functions. Typically used with negative feedback, two input nodes are virtually short, and the input current into the amplifier is almost zero. This section demonstrates an operational transconductance amplifier (OTA), which is integrated in a 60 GHz phased-array transceiver [44]. The required gain is over 20 dB and the 3 dB bandwidth should be less than 1 MHz with an output loading of 3 pF. Additionally, the input common-mode range includes both rails (0 V  $\sim$  1.2 V).

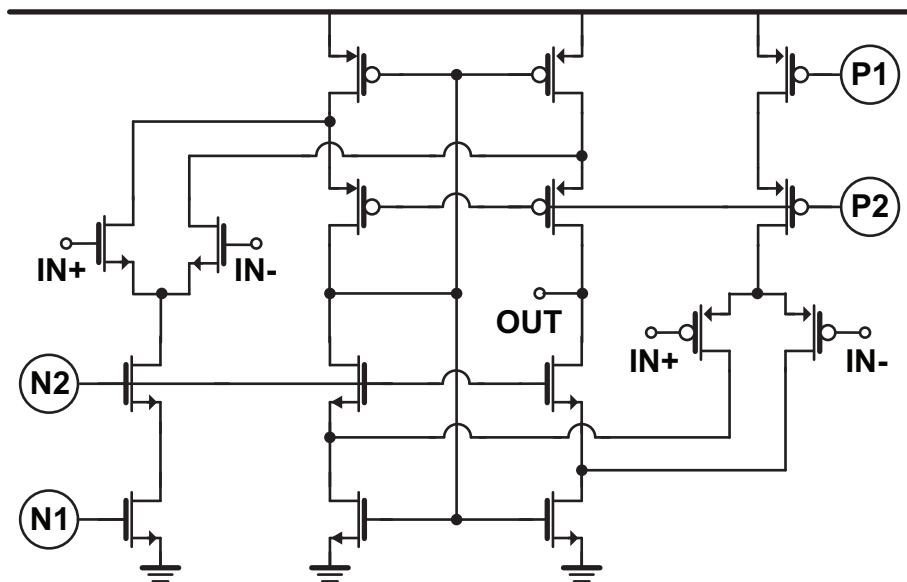


Figure 6.12: Schematic of the operational transconductance amplifier.

An OTA is designed and implemented in 65 nm CMOS, and schematically shown in Fig. 6.12. Basically, a folded-cascode amplifier is chosen to achieve high gain and wide input common-mode range. Because the input common-mode voltage ranges from 0 to 1.2 V, both NMOS and PMOS devices are used as a  $g_m$  device. For a single-ended output, mirroring is employed. The output impedance is high by the cascode structure. The bias voltages (P1, P2, N1, N2) are driven by a biasing circuit shown in Fig. 6.13. For all of the devices, two 250  $\mu\text{m}$  transistors are connected in series, and the summed channel length is 500  $\mu\text{m}$ , as illustrated in Fig. 6.14.

The OTA and biasing circuit occupy an area of 60  $\mu\text{m} \times 44 \mu\text{m}$ , as shown in Fig. 6.15. They are surrounded by a supply rail, which consists of MOS dummy capacitors and substrate contacts. Thus, it attenuates supply noise and substrate noise from adjacent circuits to some extent. The simulation results are presented in Fig. 6.16. The low-frequency gain is 31 dB and the 3 dB bandwidth is 160 kHz with an output loading of 3 pF. Over the entire input

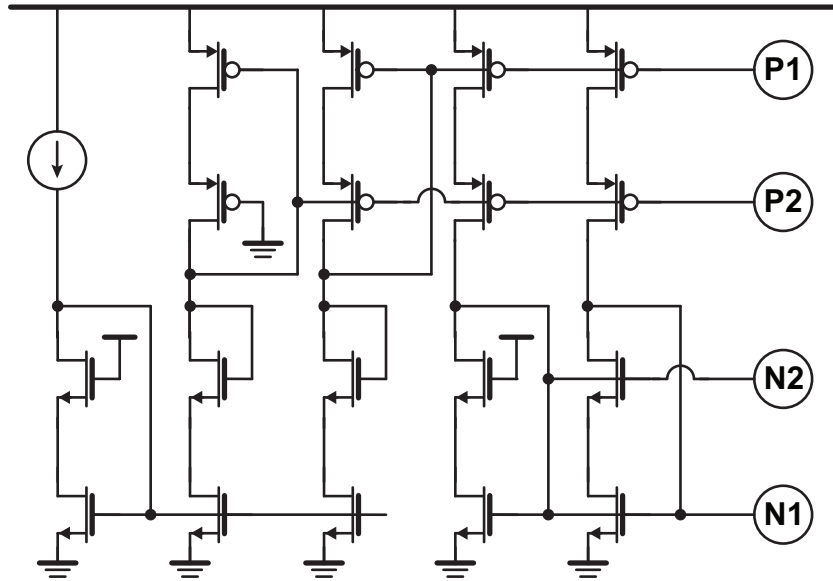


Figure 6.13: Schematic of the biasing circuit.

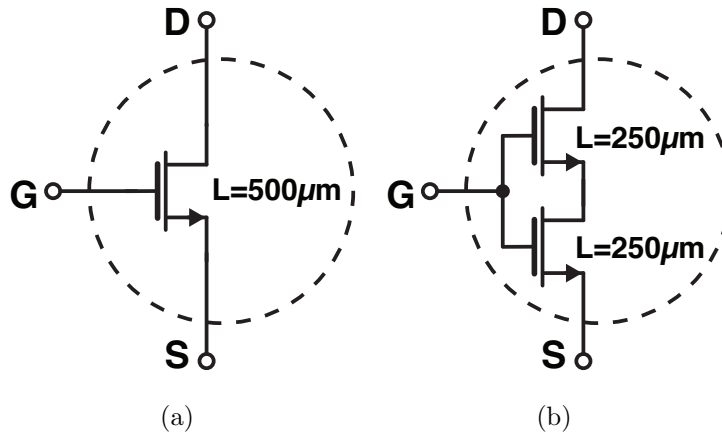


Figure 6.14: Stacked devices (a) device symbol shown in schematics (b) real implementation.

common-mode range, the gain is higher than 23.5 dB, and it is almost flat between 0.3 V and 0.9 V. For the use of feedback, the phase margin is approximately  $60^\circ$  with a unity-gain feedback. The OTA and biasing circuit draw  $90 \mu\text{A}$  from 1.2 V supply.

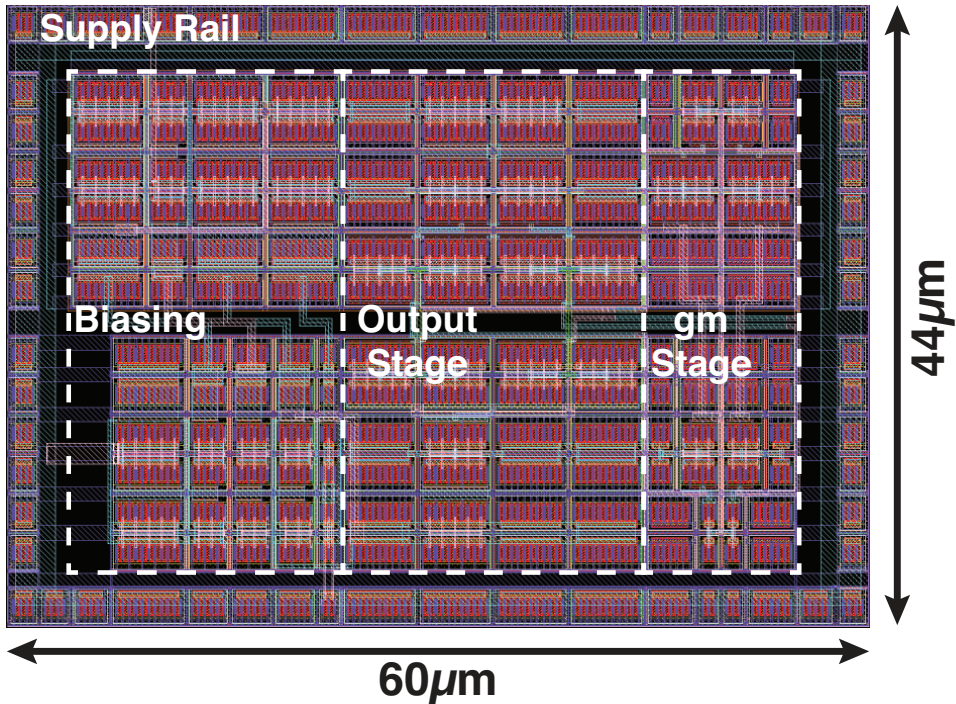


Figure 6.15: Layout of the operational transconductance amplifier.

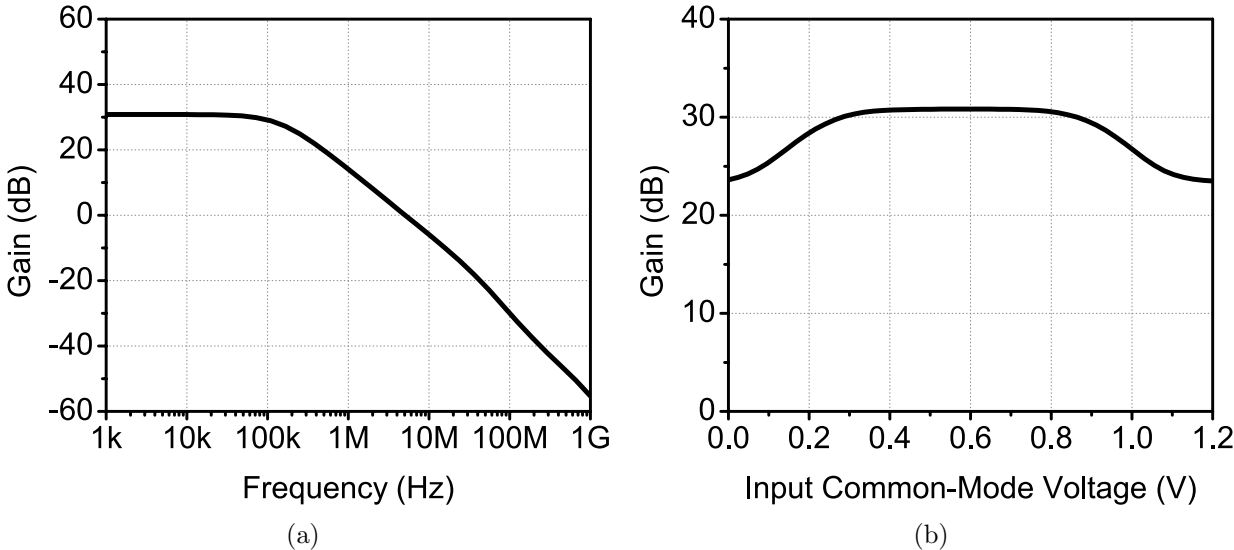


Figure 6.16: Simulation results of the OTA (a) gain and frequency (b) gain and input common-mode voltage.

## Part II

# Fully Integrated Wireless CMOS Transceivers

## Chapter 7

# 260 GHz Wireless OOK Transceiver

Part II demonstrates sub-terahertz wireless CMOS transceivers for chip-to-chip communication. This chapter describes a 260 GHz OOK transceiver implemented in 65 nm CMOS. In the transmitter, a 65 GHz VCO, amplifiers, a quadrupler, and an on-chip antenna are employed to generate 260 GHz carrier frequency. An on-off switch is designed with a 65 GHz class- $D^{-1}$  switching PA for OOK modulation and one more switch is added before the PA to improve the on-off ratio. For baseband data, a PRBS generator and distributor are designed to demonstrate 20 Gbps communication. In the receiver, the mixer-first architecture is chosen and the frequencies of LO and IF are 195 and 65 GHz, respectively. The 195 GHz LO is generated by a 65 GHz VCO and a frequency tripler. IF amplifiers are designed to achieve high gain, low noise figure, and wide bandwidth. The on-off demodulation is performed by an envelope detector, and a  $50\Omega$  driver is placed at the end. This work is performed in collaboration with Siva Thyagarajan and Jungdong Park [48].

### 7.1 System Overview

To demonstrate a sub-terahertz wireless chip-to-chip communication, a 260 GHz OOK transceiver is designed in 65 nm bulk CMOS process. The system block diagram is shown in Fig. 7.1. The transmitter is on the left side, and the receiver is on the right side. The transceiver requires DC bias voltages and one external clock for generating baseband data. The received data can be measured from a baseband amplifier.

In CMOS technology, the output power and noise figure are limited by active devices. The quality factor or insertion loss of passive devices is also limited by back-end process design. To overcome the challenges and increase the output power, phased-array structure is beneficial while sacrificing the silicon area. Here in the transceiver, two identical paths are used to boost the output power and EIRP, which increases the SNR or communication range. The EIRP can be boosted by the antenna gain. The on-chip antenna has a trade-off between gain and bandwidth; thus, it is challenging to achieve high EIRP and high data rate simultaneously.



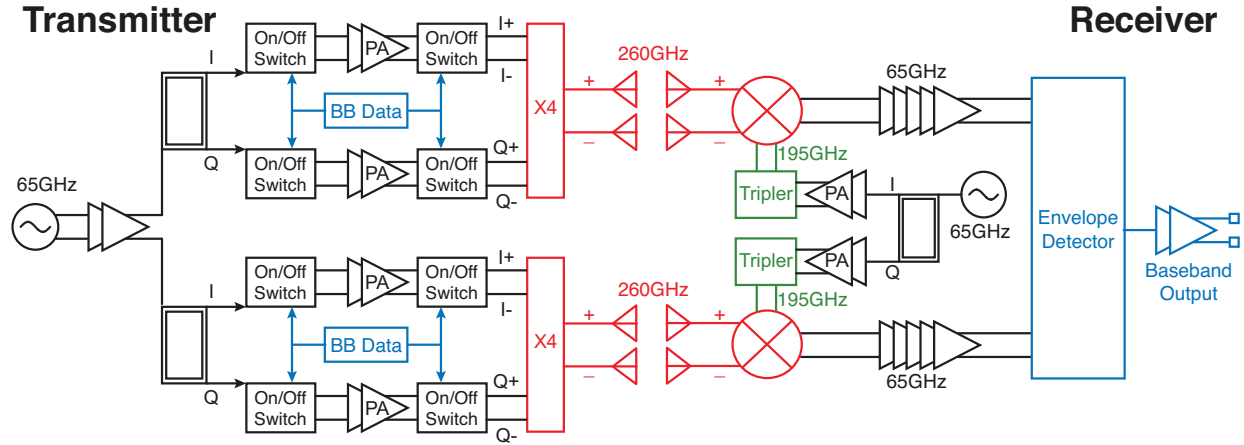


Figure 7.1: Architecture of the 260 GHz OOK transceiver.

In this system, the non-coherent OOK modulation is chosen to obviate the frequency synchronization between transmitter and receiver, simplifying the system design, although the required SNR is higher to achieve the same BER. The coherent operation is not necessary, and the (on-off) amplitude ratio is more important than the phase noise or phase distortion in system and circuit design.

The link-budget parameters are shown in Table 7.1. The transmitter power and receiver performance are based on preliminary simulation results. The path loss is from the Friis equation and the atmospheric loss is from the atmospheric attenuation shown in Fig. 2.1. The temperature is assumed to be 290 °K. From the link-budget analysis, the receiver output SNR is approximately 17 dB without margin, for a data rate of 6 Gbps and a range of 1.5 cm. The maximum achievable BER is  $4.3 \times 10^{-12}$ . All of the values in this table are not from measurements and the link margin is not taken into account.

Without the use of a lens or any optical devices, the achievable communication range with CMOS at 260 GHz is on the order of 1 cm, depending on the data rate and the target BER. Thus, the transceiver is suitable for short-range wireless links. If lenses are available, the communication range can be extended to over 1 m, depending on the lens gain.

## 7.2 Transmitter

The full block diagram (Fig. 7.1) is so complicated that it is necessary to simplify the transmitter architecture, as illustrated in Fig. 7.2. Basically, the carrier frequency of 260 GHz is generated by a 65 GHz LO and a quadruple-push frequency multiplier. The multiplier requires four phases ( $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ ,  $270^\circ$ ). The transmit power is determined by the output power of the power amplifier and the conversion gain of the multiplier. The 65 GHz VCO output is amplified by several stages to maximize the output power. The OOK modulation is performed by on-off switches distributed at the input and output of the power amplifier.

Table 7.1: Link parameters for the 260 GHz OOK transceiver

Quantity	Symbol	Value
Transmitter Output Power	$P_{TX}$	0.5 dBm
Transmitter Antenna Gain	$G_{A, TX}$	4.5 dBi
EIRP	$EIRP_{TX}$	5 dBm
Carrier Frequency	$f$	260 GHz
Wavelength	$\lambda$	1.15 mm
Link Range	$d$	1.5 cm
Path Loss	$L_{path}$	-44.3 dB
Atmospheric Loss	$L_{atm}$	$-6.1 \times 10^{-5}$ dB
Channel Loss	$L_{channel}$	-44.3 dB
Receiver Antenna Gain	$G_{A, RX}$	2 dBi
Receiver Input Power	$P_{RX}$	-37.3 dBm
Temperature	$T$	290 °K
Data Rate (OOK)	$R$	6 Gbps
Double Sideband Bandwidth	$W$	12 GHz
Double Sideband Noise Figure	$NF$	19 dB
Signal to Noise Ratio	$SNR_{OUT}$	16.9 dB
Bit Error Rate	$BER$	$4.3 \times 10^{-12}$

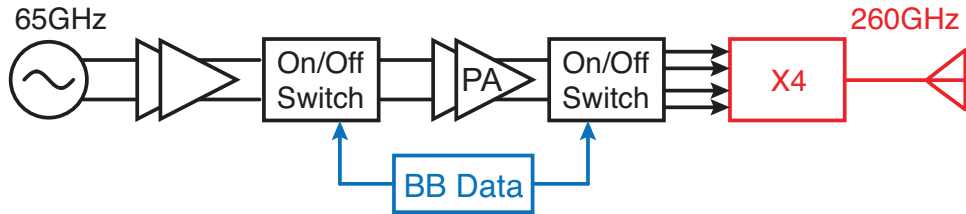


Figure 7.2: Block diagram of the 260 GHz transmitter.

It is hard to completely turn on/off the  $LC$  resonant tank for high-speed data communication; thus, two on-off switches are employed to increase the on-off ratio. Because the finite on-off ratio degrades SNR of the receiver output and BER of the system, more switches are required, but insertion loss increases. The two switch signals should be properly synchronized; otherwise, the delay mismatch between the PA delay and the baseband data delay can degrade the transmitter SNR.

As shown in Fig. 7.1, the transmitter has two identical paths to increase the output power and EIRP, thereby extending the communication range. The 65 GHz quadrature phases are generated by a 65 GHz VCO (Fig. 3.4) and the quadrature hybrid. Two phases (I and Q) are amplified and modulated by a class- $D^{-1}$  switching power amplifier and MOS switches. The on-off switches are driven by the PRBS generator and distributor, as shown in Fig. 6.6

and Fig. 6.7. Finally, the quadrature-push multiplier drives the on-chip antenna. The circuit schematics and details are presented in [48]. The PA output power is 13 dBm and the output power is approximately 0~0.5 dBm. The power dissipated by the transmitter is 688 mW, without power consumption of the PRBS generator.

### 7.3 Receiver

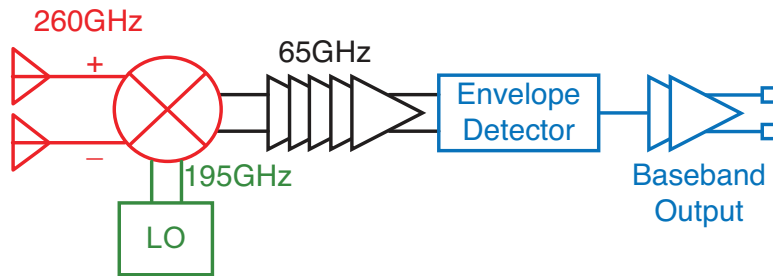


Figure 7.3: Block diagram of the 260 GHz receiver.

The simplified receiver design is illustrated in Fig. 7.3. Because the carrier frequency is higher than cut-off frequencies, the mixer-first architecture is chosen. The received RF signal is first down-converted, and the LO and IF frequencies are selected as 195 and 65 GHz, respectively. If the LO output has a parasitic 65 GHz tone, then it can potentially ruin the OOK demodulation due to the LO-to-IF leakage. The 65 GHz IF amplifiers should have high gain, wide bandwidth, and low noise figure to increase the receiver output SNR. The envelope detector is an OOK demodulator, detecting whether the incoming signal exists or not. Finally, the detector output is buffered by broadband amplifiers and a  $50\ \Omega$  driver.

As shown in Fig. 7.1, the receiver has also two identical paths to increase the output power and SNR. The 260 GHz heterodyne mixer is shown in Fig. 5.6. The LO frequency is generated by a 195 GHz frequency tripler. Two input phases of the tripler are different by  $90^\circ$ ; thus, the output phases are different by  $270^\circ$  ( $-90^\circ$ ), conserving quadrature phases. The 65 GHz IF amplifiers utilize loosely coupled transformers to achieve wide bandwidth. The operation of the envelope detector is described in Fig. 5.9. A  $50\ \Omega$  driver is then used to deliver output signals to an external instrument. As in the transmitter, the 65 GHz quadrature phases are generated by a 65 GHz VCO (Fig. 3.4) and a quadrature hybrid. A class- $D^{-1}$  switching power amplifier is used to maximize the input power of the lossy frequency tripler. The circuit schematics and details are presented in [48]. The power dissipation of the receiver is 485 mW.

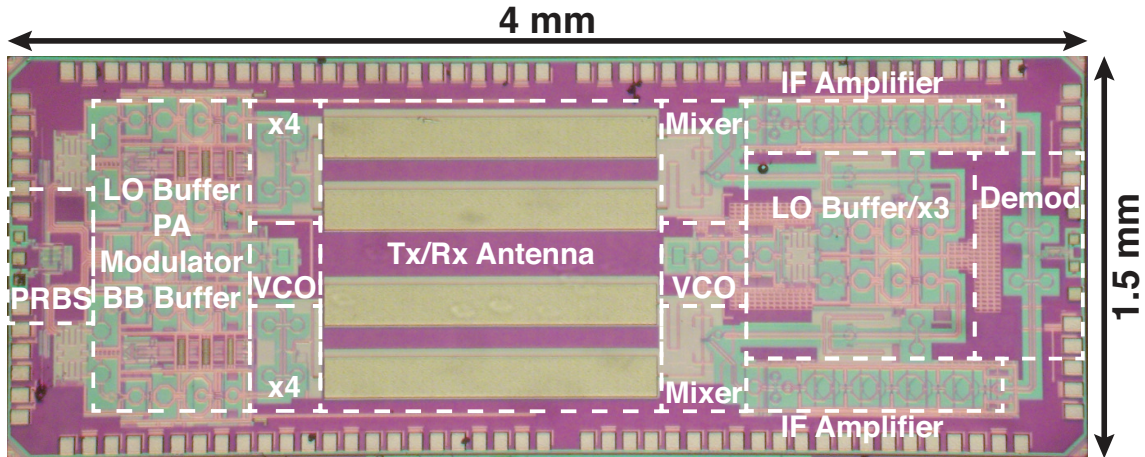


Figure 7.4: Microphotograph of the 260 GHz OOK transceiver.

## 7.4 Fabrication

The 260 GHz OOK transceiver is fully integrated in 65 nm CMOS without special process options. From simulation of a single-gate-contact device ( $L = 60$  nm,  $W = 1$   $\mu$ m), the unity-gain frequency ( $f_T$ ) is approximately 190 GHz and the maximum oscillation frequency ( $f_{max}$ ) is approximately 250 GHz, depending on the device layout. The top two metal layers are thick, and they can be used in passive devices, such as transformer and transmission line. The die photo is shown in Fig. 7.4, and the size is 4 mm  $\times$  1.5 mm. In the center, two differential antennas (four transmission lines) are placed. The transmitter is on the left side. The VCO and the PRBS generator are shown, and buffers, amplifiers, and OOK modulators are in between. The frequency quadrupler is attached to the antenna. The receiver is on the opposite side. The mixers are attached to the antenna. The VCO, LO buffer, and a tripler are in the middle, and the IF amplifiers deliver the received signal from the mixer to a demodulator, which is placed at the right end. The antennas are connected by both the transmitter (left) and the receiver (right). When the chip operates in the transmitter mode, the receiver part is turned off, and the antenna-transmitter matching is performed with the receiver off. Conversely, when the chip operates in the receiver mode, the transmitter part is turned off and the antenna-receiver matching is performed with the transmitter off. The PCB is designed to support both transmitter and receiver modes; therefore, the bypass mode can be enabled to test direct data communication not through the air.

## 7.5 Measurement Results

The 260 GHz OOK transceiver can be measured in various ways. Fig. 7.5, Fig. 7.6, and Fig. 7.8 show three measurement setups. First, the 260 GHz transmit signal can be captured with a WR3.4 horn antenna, as shown in Fig. 7.5. The power meter and sensor, attached to

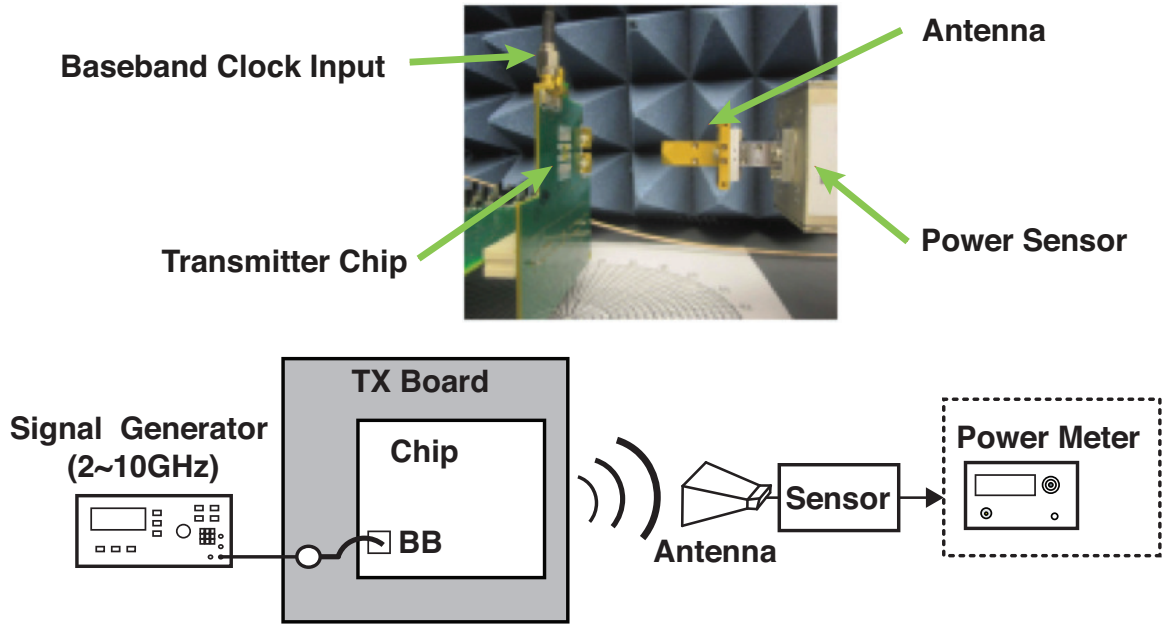


Figure 7.5: Measurement setup for the transmitter.

the antenna, can measure the radiated transmit power. The measured EIRP is 5 dBm, and the output power is approximately 0.5 dBm. The measured antenna pattern and the beam width are presented in [48].

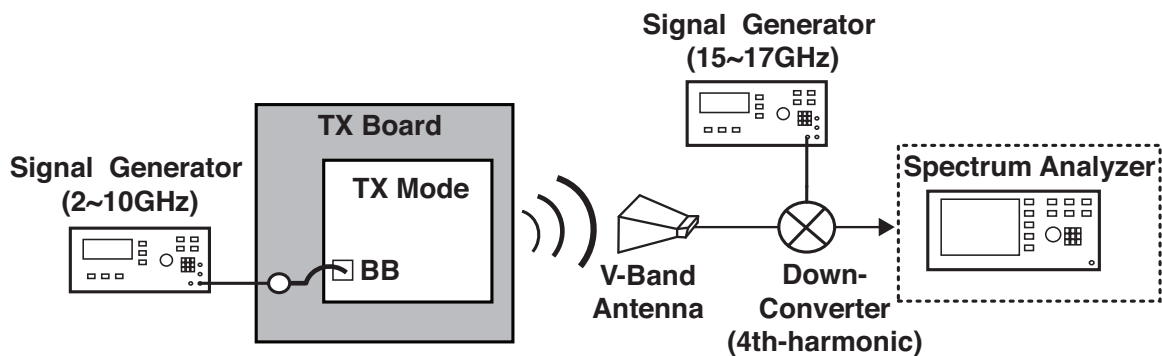


Figure 7.6: Measurement setup for the V-band leakage.

Fig. 7.6 shows an indirect way to measure the V-band LO signal and modulated signal, before it is up-converted by the frequency quadrupler. The V-band horn antenna captures the leakage signal radiated from 65 GHz transformers and any EM structures. The frequency of the 65 GHz VCO can be measured and shown in Fig. 3.6. The measured frequency ranges from 61 GHz to 67 GHz. The modulated signals are shown in Fig. 7.7. The data rate can be varied by changing the clock frequency of the PRBS generator. The center tone is stronger

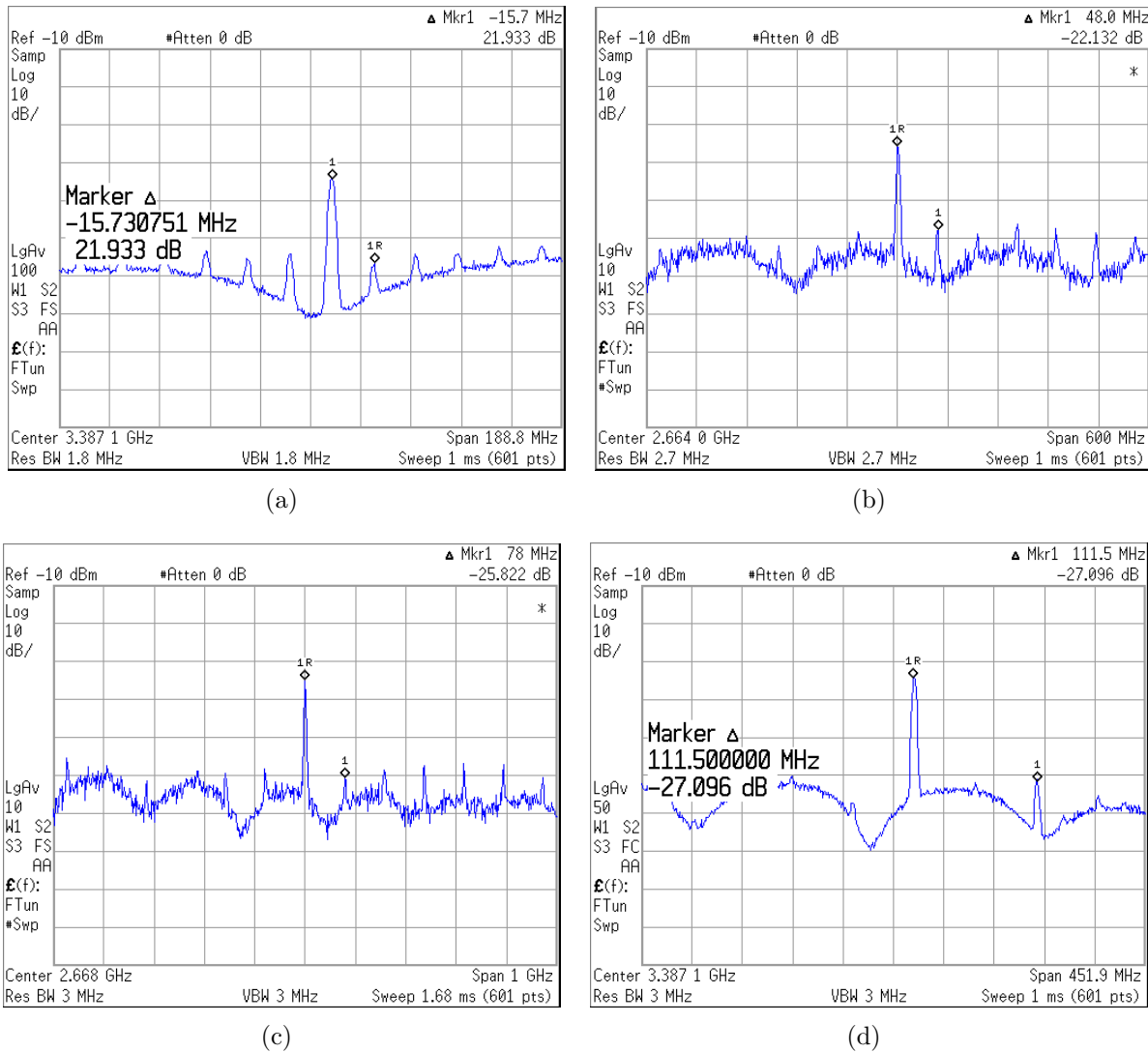


Figure 7.7: Measured frequency spectra of OOK modulation signals (a) 2 Gbps (b) 6 Gbps (c) 10 Gbps (d) 14 Gbps.

than the side tones because of the finite on-off ratio. The beat frequency is given by Eq. 6.2 and matches the measured spectrum. The maximum data rate of 14 Gbps can be observed as shown in Fig. 7.7(d).

The wireless link (transmitter and receiver) can be measured as shown in Fig. 7.8. With a continuous-wave mode, the on-off ratio of 10 dB is measured. The receiver SNR is 10 dB without modulation. The transceiver performance is summarized in Table 7.2. Although data communication is not fully demonstrated, it shows the feasibility of the sub-terahertz wireless link in CMOS technology. In the next chapter, a 240 GHz 16 Gbps QPSK modulation is successfully demonstrated in 65 nm CMOS.

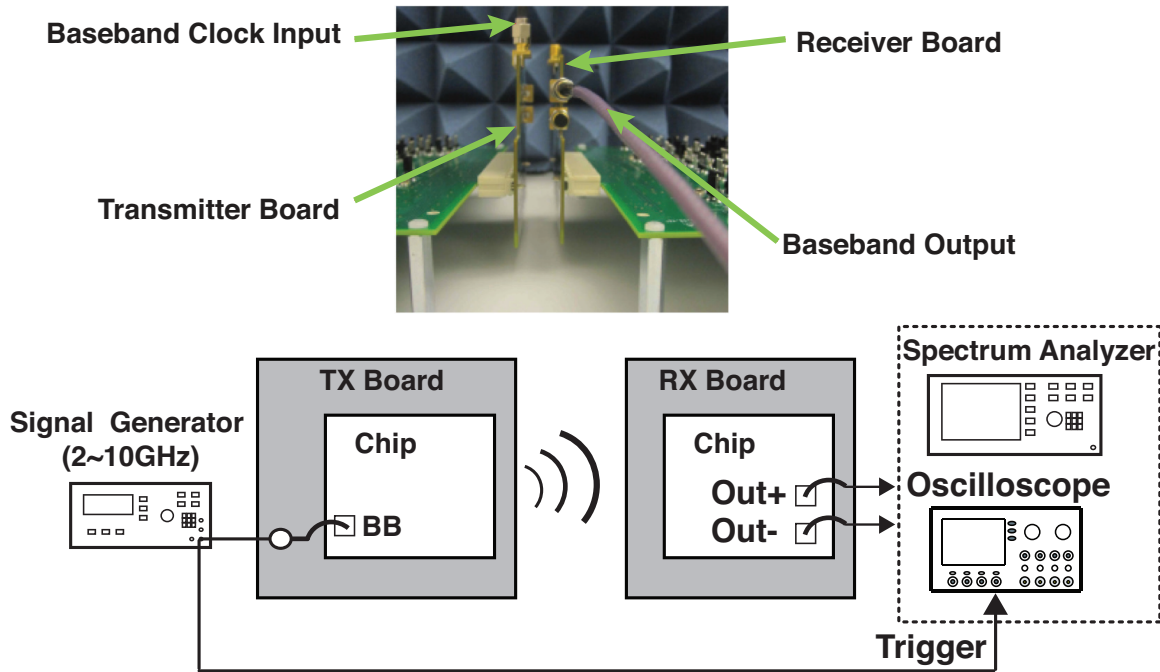


Figure 7.8: Measurement setup for the transmitter and receiver.

Table 7.2: Summary of the 260 GHz OOK transceiver

Technology	65 nm CMOS
Modulation	OOK (on-off keying)
Frequency (GHz)	260
Pout (dBm)	0.5
EIRP (dBm)	5
Receiver Gain (dB)	17
Receiver (DSB) Noise Figure (dB)	19
Power Consumption (mW)	1173 (TX: 688, RX:485)
Area ( $mm^2$ )	6
Antenna	on-chip
Antenna gain (dB)	4.5
Range (cm)	1 ~ 4

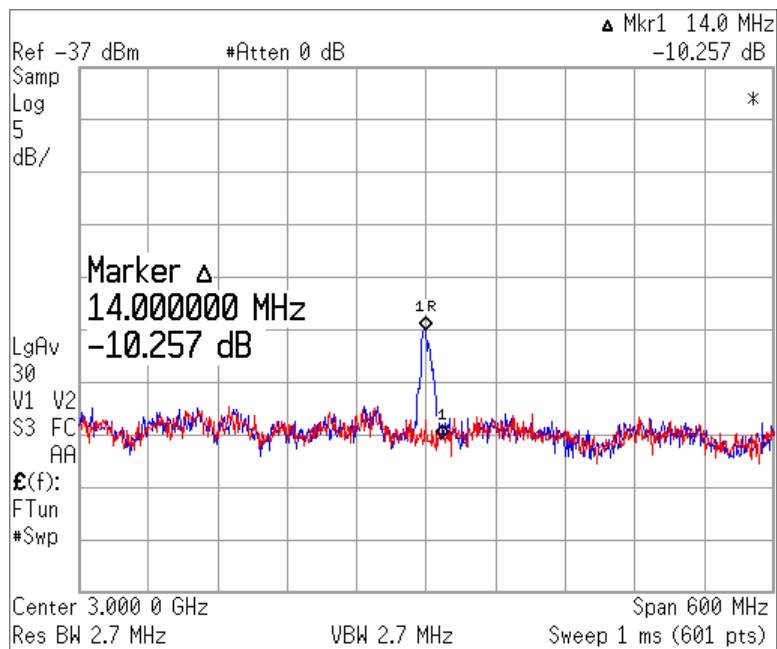


Figure 7.9: Measured on/off states (on-off ratio is 10 dB).



## Chapter 8

# 240 GHz QPSK/BPSK Transceiver

This chapter describes a 240 GHz QPSK/BPSK transceiver implemented in 65 nm CMOS. In the transmitter, an 80 GHz injection-locked oscillator, amplifiers, a frequency tripler, and an on-chip antenna are implemented to generate a 240 GHz carrier frequency. A differential hybrid and a QPSK modulator are used to generate QPSK/BPSK modulation signals, and the frequency tripler, which is at the end of the transmitter, multiplies the LO frequency, while conserving the QPSK/BPSK modulation. An on-chip PRBS generator is designed and there are three operation modes (QPSK, BPSK, and a continuous-wave mode). In the receiver, the mixer-first architecture and direct conversion are chosen to achieve low noise figure and high conversion gain. The 240 GHz LO is generated by an 80 GHz injection-locked oscillator, amplifiers, and a tripler. The mixer output goes through baseband amplifiers that have high gain, low noise figure, and wide bandwidth. The transceiver achieves a maximum data rate of 16 Gbps and an energy efficiency of 30 pJ/bit with a communication range of 1.5 cm. This work is performed in collaboration with Siva Thyagarajan [1], [2].

### 8.1 System Overview

A 240 GHz QPSK/BPSK transceiver is implemented in 65 nm CMOS process for high-speed wireless communications. The system block diagram is shown in Fig. 8.1. The transmitter is on the left side, and the receiver is on the right side. The transceiver requires DC bias voltages, a 13.3 GHz LO clock to synchronize the frequency and phase, and a baseband clock to generate PRBS data. The received data (I and Q) can be measured from the baseband amplifiers. Digital circuits are also included for digital control and calibration.

The transmitter and receiver both have one differential on-chip antenna each. For a phased-array system, multiple chips can be used to increase the output power, the EIRP, and the SNR. The transmitter and receiver chips can be synchronized using a 13.3 GHz LO clock. Compared to the previous OOK transceiver, the size of the on-chip antenna is significantly reduced to lower the overall form factor.

The transceiver uses coherent QPSK modulation to increase the spectrum efficiency. For

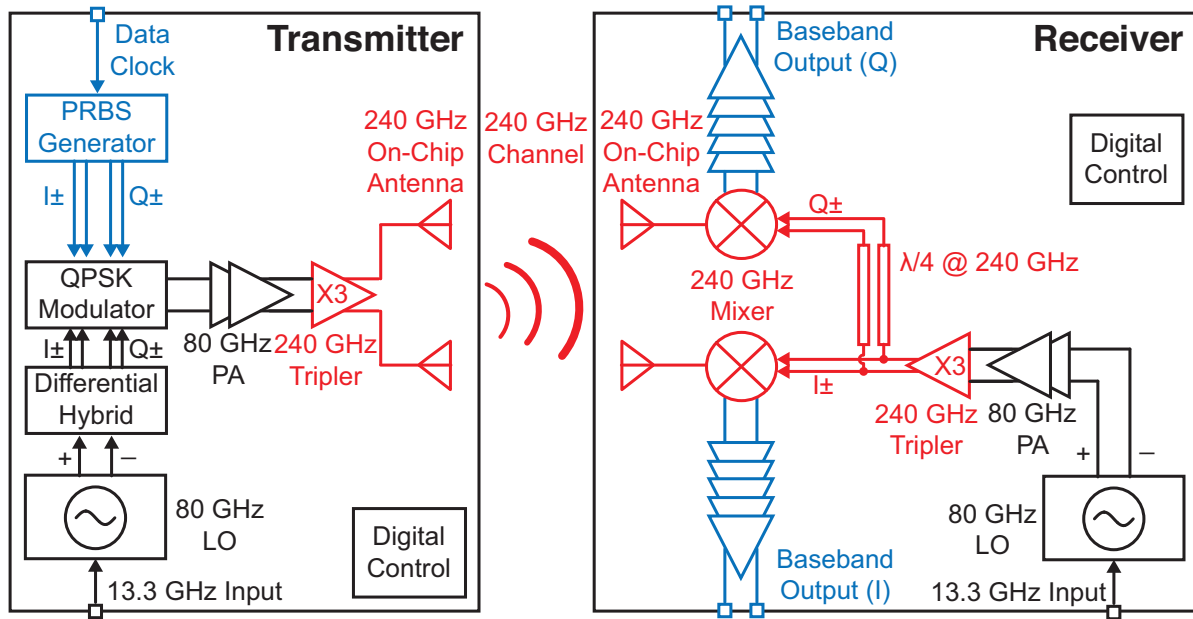


Figure 8.1: Architecture of the 240 GHz QPSK/BPSK transceiver.

a given bandwidth, the data rate of QPSK is twice that of OOK. Frequency and phase synchronization circuits are required, but real chips typically have a frequency synthesizer; therefore, the use of a synchronization block does not present a great burden. In addition, in high data-rate systems, changing the phase of the carrier is easier than turning the carrier on/off because it is difficult to completely turn off the carrier in a short time. Moreover, the required  $E_b/N_o$  is lower than any other modulation for a given BER, as shown in Fig. 2.2. The phase noise is one of the most important metrics in phase modulations. The transmitter has three operation modes (QPSK, BPSK, and a continuous-wave mode) to test and measure variable cases.

The link-budget parameters are shown in Table 8.1. The transmitter power and receiver performance are based on preliminary simulation results. The path loss is calculated using the Friis equation, and the atmospheric loss is obtained from the atmospheric attenuation shown in Fig. 2.1. The temperature is assumed to be 290 °K. From the link-budget analysis, the receiver output SNR is approximately 15.6 dB without margin for a data rate of 16 Gbps (8 GSps, QPSK) and a range of 1.5 cm. The maximum achievable BER is  $8.0 \times 10^{-10}$ . All of the values in this table are not from measurements and the link margin is not taken into account.

Without the use of a lens or any (quasi-)optical devices, the achievable communication range with CMOS at 240 GHz is on the order of 1 cm, depending on the data rate and the target BER. Thus, the transceiver is suitable for short-range wireless links. If lenses are available, the communication range can be extended to over 1 m, depending on lens size.

Table 8.1: Link parameters for the 240 GHz QPSK transceiver

Quantity	Symbol	Value
Transmitter Output Power	$P_{TX}$	0 dBm
Transmitter Antenna Gain	$G_{A,TX}$	1.5 dBi
EIRP	$EIRP_{TX}$	1.5 dBm
Carrier Frequency	$f$	240 GHz
Wavelength	$\lambda$	1.25 mm
Link Range	$d$	1.5 cm
Path Loss	$L_{path}$	-43.6 dB
Atmospheric Loss	$L_{atm}$	$-5.2 \times 10^{-5}$ dB
Channel Loss	$L_{channel}$	-43.6 dB
Receiver Antenna Gain	$G_{A,RX}$	0.7 dBi
Receiver Input Power	$P_{RX}$	-41.4 dBm
Temperature	$T$	290 °K
Data Rate (QPSK)	$R$	16 Gbps
Double Sideband Bandwidth	$W$	16 GHz
Double Sideband Noise Figure	$NF$	15 dB
Signal to Noise Ratio	$SNR_{OUT}$	15.6 dB
Bit Error Rate	$BER$	$8.0 \times 10^{-10}$

## 8.2 Transmitter

The transmitter architecture is illustrated in Fig. 8.2. The transmitter generates a carrier frequency of 240 GHz using an 80 GHz LO and a frequency tripler. The 80 GHz LO part consists of a 13.3 GHz input clock and two injection-locked oscillators. The frequency of the input signal is doubled by the 26.6 GHz doubler, and the output is injected to a 26.6 GHz injection-locked oscillator. The 26.6 GHz oscillator output is frequency-tripled and then injected to an 80 GHz injection-locked oscillator. The two oscillators should be simultaneously injection-locked to generate a low-noise and stable LO signal. The LO should be synchronized to an external clock because the transceiver uses coherent phase modulations. An 80 GHz phase-locked loop can be employed here. Next, a differential hybrid generates I/Q quadrature phases, which is for QPSK modulation. Compared to a single-ended hybrid, the insertion loss is higher, but the single-to-differential or differential-to-single operation is eliminated, and the overall loss can be reduced, because the 80 GHz oscillator and modulator are fully differential circuits. The QPSK/BPSK modulation is performed at 80 GHz because tripling the frequency preserves the constellation points of the QPSK and BPSK modulations, as illustrated in Fig. 8.3. Thus, the 80 GHz modulator can operate more efficiently than the 240 GHz modulator. The phase noise or EVM is tripled by the frequency tripler, but this drawback can be alleviated by the use of an 80 GHz low-noise oscillator and other

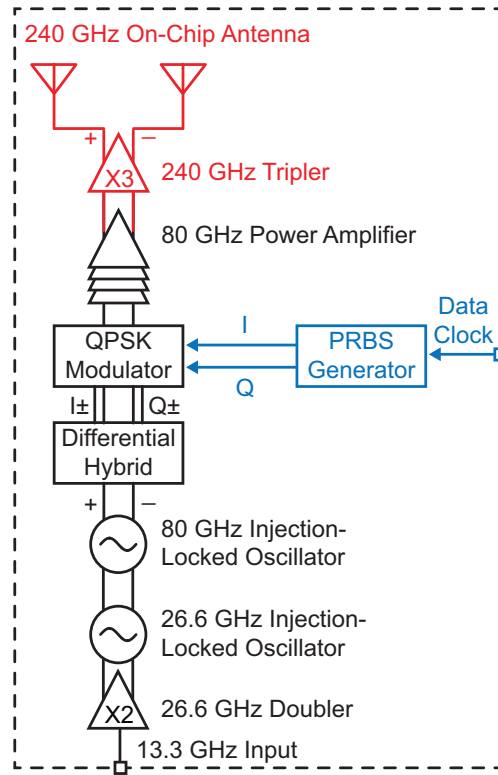


Figure 8.2: Block diagram of the 240 GHz transmitter.

blocks. A QPSK modulator acts as a double-balanced up-converting mixer and the baseband port is driven by a PRBS generator. An on-chip PRBS generator is designed to support three operation modes (QPSK, BPSK, and a continuous-wave mode) and operates up to 40 Gbps of a data rate in the QPSK mode (20 Gbps in BPSK). The modulated signal is sent to the driving amplifiers, a class-E switching power amplifier, and a frequency tripler. This signal path should have a wide bandwidth to support high data rates. For the 20 Gbps QPSK, the null-to-null bandwidth is 20 GHz, and the amplifiers and matching networks should have a bandwidth of 20 GHz, whereas the center frequency is 80 GHz. The tripler output drives a 240 GHz on-chip slotted loop antenna.

The transmitter blocks should be carefully designed to minimize the transmitter EVM. There are several design considerations for the EVM. First, the LO phase noise is an important factor for coherent phase modulations. In the transmitter, the oscillator should achieve a low phase noise, particularly at the high offset frequencies (typically between 1 MHz and 1 GHz), because phase noise at low offset frequencies ( $< 1$  MHz) can be cancelled out because a receiver can track slow change of the carrier phase. In contrast, the phase noise at high offset frequencies ( $> 1$  GHz) is dominated by thermal noise and thus cannot be cancelled, which degrades the transmitter EVM and the receiver SNR. Second, the data pattern affects the constellation points. In high-speed communications, the rise/fall time is not negligible

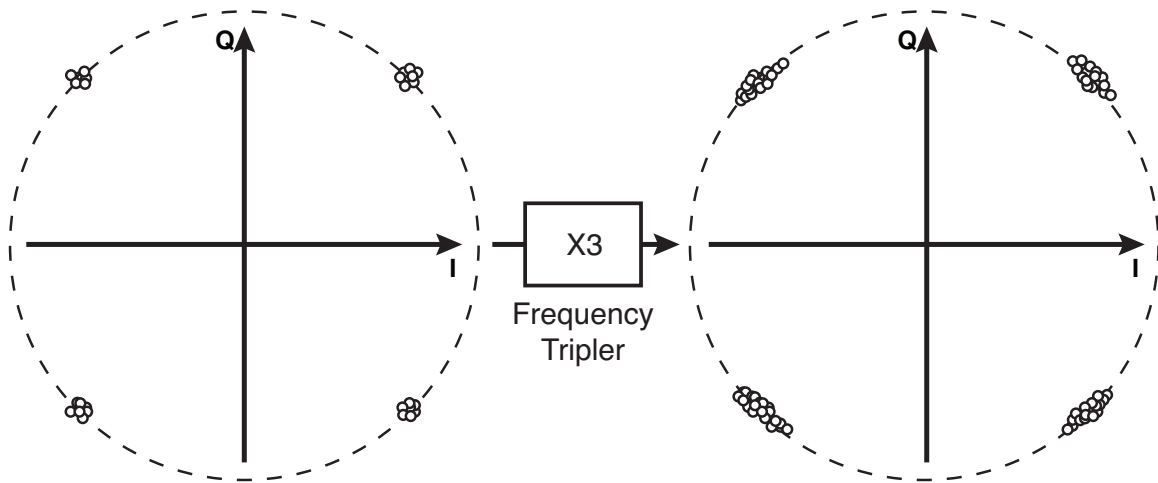


Figure 8.3: QPSK constellation points preserved by a frequency tripler.

compared to the symbol time; thus, the constellation points can be affected by previous states or transitions. Third, mismatches in the modulator result in asymmetric constellation points that increase the transmitter EVM. The 80 GHz driving and power amplifiers can also cause phase (AM-to-PM or PM-to-PM) distortions. Therefore, wideband amplifiers and antenna should be designed. Lastly, the phase errors that occur before the frequency tripler are scaled up three times, which spreads the constellation points, as illustrated in Fig. 8.3.

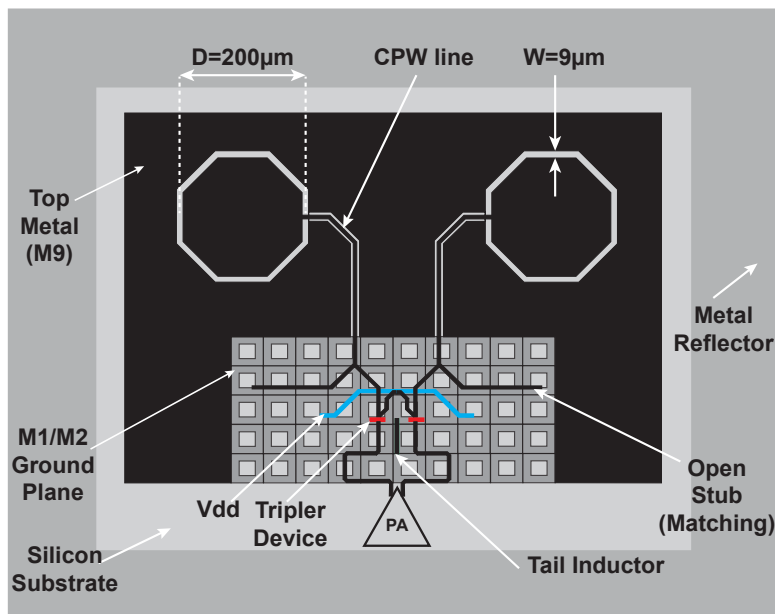


Figure 8.4: Layout design of the transmitter.

Fig. 8.4 shows a layout design of the transmitter front-end. The black part represents the top metal layer and the ground plane for the on-chip slotted loop antenna (two large octagons). The frequency tripler and its input/output matching networks are also shown. The red dots represent the tripler transistors, the black lines represent the microstrip lines, and the blue line represents the supply line. There is a microstrip-to-CPS conversion between the tripler output and the antenna. The tripler output has a fundamental tone (80 GHz) and a second-harmonic tone (160 GHz), which should be attenuated by the matching network. In simulation, the radiated powers at 80 and 160 GHz are negligible.

### 8.3 Receiver

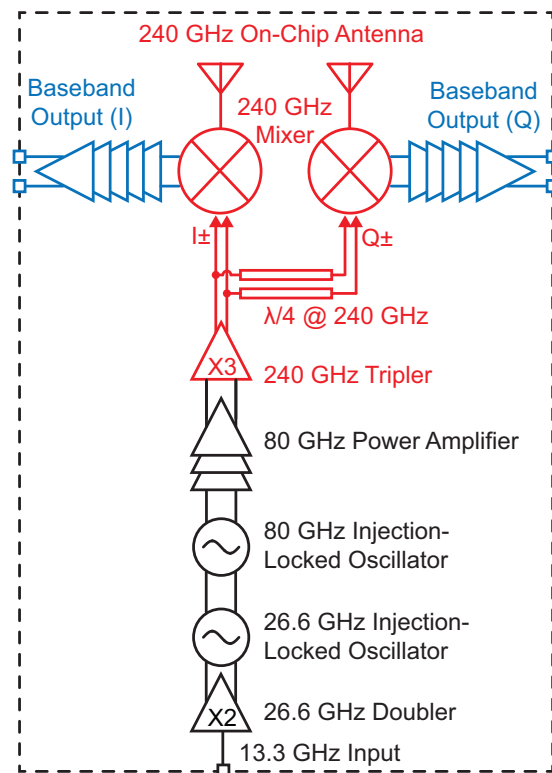


Figure 8.5: Block diagram of the 240 GHz receiver.

The receiver architecture is illustrated in Fig. 8.5. The mixer-first architecture is chosen such that a down-converting mixer is at the front-end of the receiver. A low-noise amplifier is typically the first block after the antenna but the carrier frequency (240 GHz) is higher than  $f_T/f_{max}$ ; thus, it is not practical to use an amplifier at such a high frequency. Accordingly, the overall noise figure is higher than the conversion loss of the mixer. A direct-conversion mixer is used, and the LO frequency is also 240 GHz, thereby enabling power-efficient amplification

of the baseband signal. The 240 GHz LO signal is generated in a similar way as for the transmitter. Two oscillators are injection-locked to the 13.3 GHz input clock and the 80 GHz oscillator drives the driving amplifiers, a power amplifier, and then a frequency tripler. Because the amplifiers are used to amplify the LO signal, a wide bandwidth is not required as the transmitter. The I/Q quadrature phases are generated by a differential quarter-wave delayed line. Typically, a quadrature hybrid could be implemented in this case, but the insertion loss would be too high at 240 GHz. Moreover, the amplitude and phase mismatches of the delayed line are comparable to those of the hybrid because the delayed line design is much simpler and shorter at 240 GHz. The tripler has 80 and 160 GHz leakages, which are rejected by the delay lines and do not affect the mixer performance. The mixer output has a broadband signal. The baseband amplifiers and a  $50\ \Omega$  driver boost the voltage swing of the received signal to properly capture and measure the output.

The receiver should be designed to achieve a low noise figure. There are several design considerations for the noise figure. The first receiver block dominates the noise figure; therefore, it is critical to reduce the conversion loss and the noise figure of the mixer, which depend on the LO power. Thus, the 240 GHz LO power should be high. Any parasitic tones (80 and 160 GHz) other than 240 GHz can also degrade the mixer performance. The differential RF signal generation at the antenna increases the input loss, which should be minimized to reduce the overall noise figure. The baseband amplifiers should be designed to have low noise figure, high gain, wide bandwidth, and low power dissipation. A tapered design with nine stages improves the baseband amplifier power efficiency while maintaining low noise figure. Therefore, the receiver design involves a careful co-optimization among the antenna, the mixer and the baseband amplifiers.

Fig. 8.6 shows the layout design of the receiver front-end. The on-chip antenna is the same as the transmitter antenna except for the front-end interconnection with the mixer. A single-to-differential conversion is needed to drive the double-balanced mixer. Multiple cuts are performed in the ground plane to reject the common mode signal. The red dots represent mixers, and the baseband amplifiers are placed on both the left and right sides. The frequency tripler drives the differential quarter-wave delayed lines, which are coplanar striplines. The long lines and the transformers are matched at 240 GHz and attenuate the 80 and 160 GHz tones from the frequency tripler.

## 8.4 Fabrication

The 240 GHz QPSK/BPSK transceiver chips are fabricated in 65 nm CMOS without any special process options. From the simulation of a single-gate-contact device ( $L = 60\ \text{nm}$ ,  $W = 1\ \mu\text{m}$ ), the unity-gain frequency ( $f_T$ ) is approximately 190 GHz and the maximum oscillation frequency ( $f_{max}$ ) is approximately 250 GHz. The die photos are shown in Fig. 8.7. The size of each chip is  $2\ \text{mm} \times 1\ \text{mm}$ , and the antenna size is  $800\ \mu\text{m} \times 500\ \mu\text{m}$ , including the ground plane. Chip-on-board packaging is used to provide supply voltages and clocks, and the packaging and the gold reflector on the PCB are taken into account in on-chip antenna sim-

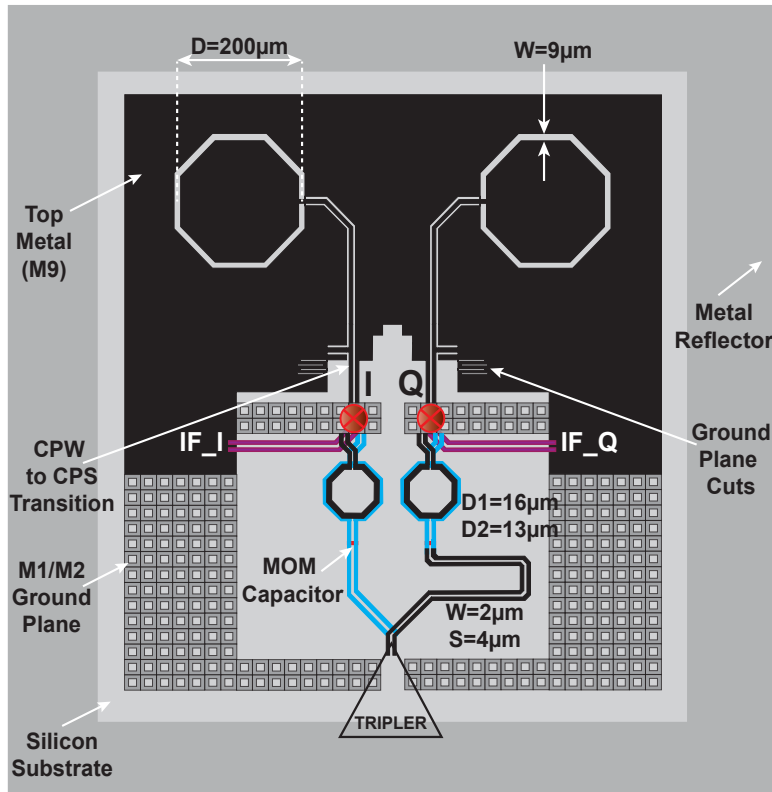


Figure 8.6: Layout design of the receiver.

ulations. The slotted-loop antennas and all of the EM structures are visible. The 13.3 GHz input clock comes through a bond wire and PCB transmission lines, which are lossy at the frequency. Thus, the lengths are minimized to reduce the loss. The 26.6 GHz doubler and two injection-locked oscillators (26.6 GHz and 80 GHz) are used in both the transmitter and receiver. In the transmitter, a differential hybrid, a modulator, 80 GHz amplifiers, and a frequency tripler are sequentially placed towards the antenna. In the receiver, 80 GHz amplifiers, a frequency tripler, delayed lines, and I/Q mixers are placed towards the antenna. The I and Q baseband amplifiers are on the left side and the right side, respectively. All of the bias voltages and control bits are changed by digital controllers.

## 8.5 Measurement Results

To measure the transmitter output power, EIRP, and modulated signals, the measurement setup for the transmitter is shown in Fig. 8.8. A WR3.4 diagonal horn antenna, a broadband down-converter, a spectrum analyzer, and an oscilloscope are used. The down-converter is a sub-harmonic passive mixer and requires 120 GHz LO signal, which is generated by a signal generator and a frequency multiplier.



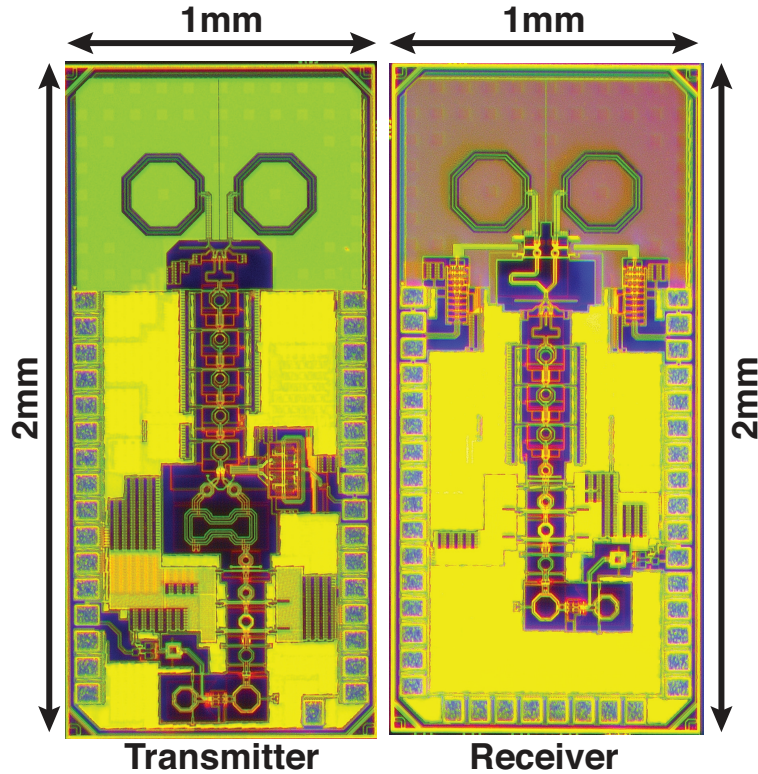


Figure 8.7: Microphotographs of the transmitter and receiver.

First, the continuous-wave mode is measured. As shown in Fig. 8.9, the output power is measured and varies with the range. The measured data follow the Friis equation well. From the measured output power, the EIRP can be calculated after compensating for the overall loss of the measurement instruments. The measured and compensated EIRP values are between 1 dBm and 2 dBm, matching well with the simulation results. The antenna radiation pattern is also measured by changing the position of the antenna, as presented in [1].

Next, in the QPSK modulation mode, the down-converted frequency spectrums are measured using a spectrum analyzer, as shown in Fig. 8.10, Fig. 8.11, and Fig. 8.12. Fig. 8.10(a) displays the SINC waveform of the 4 GSps modulation signals (8 Gbps QPSK or 4 Gbps BPSK), and Fig. 8.10(b) proves the beat frequency ( $\approx \frac{4\text{GHz}}{2^7-1} \approx 31.5\text{ MHz}$ ) matches to the modulation data rate. Fig. 8.11(a) displays the SINC waveform of the 6 GSps modulation signals (12 Gbps QPSK or 6 Gbps BPSK), and Fig. 8.11(b) shows the beat frequency ( $\approx \frac{6\text{GHz}}{2^7-1} \approx 47.2\text{ MHz}$ ). Fig. 8.12(a) displays the SINC waveform of the 8 GSps modulation signals (16 Gbps QPSK or 8 Gbps BPSK), and Fig. 8.12(b) shows the beat frequency ( $\approx \frac{8\text{GHz}}{2^7-1} \approx 63.0\text{ MHz}$ ). Additionally, the null locations verify the correct data rates. As the data rate increases, the transmit power exhibits more spread, and the SINC waveform is distorted more.

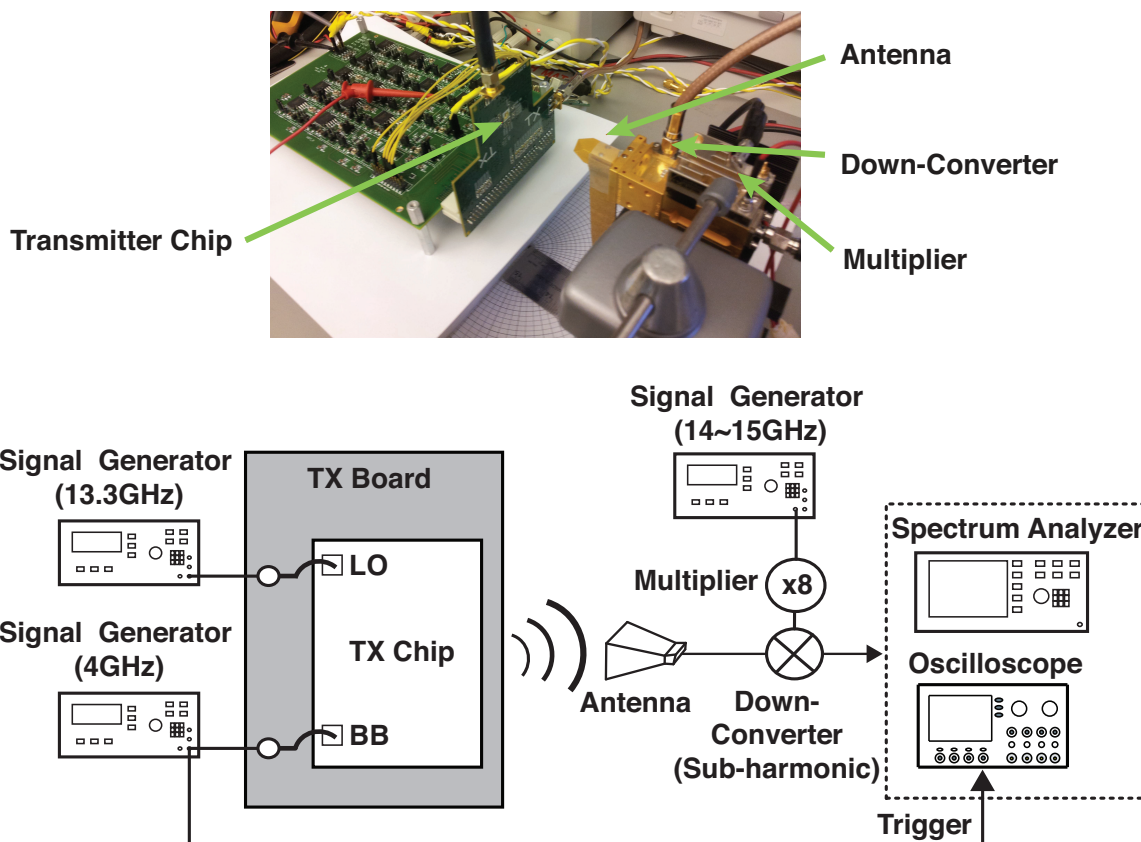


Figure 8.8: Measurement setup for the transmitter.

Fig. 8.13 shows measured time-domain eye diagrams for 4 GSps, 6 GSps, and 8 GSps modulations. The orange sinusoidal wave indicates the input data clock, confirming the 2:1 half-rate transmitter. As data rate increases, the eye opening is degraded.

Table 8.2: Power breakdown of the 240 GHz transmitter

Building block	Power dissipation (mW)	Percentage (%)
80 GHz LO Chain	41	19
80 GHz Modulator	8	4
80 GHz Amplifiers	116	53
240 GHz Tripler	45	20
Transmitter	220	100

The power dissipation of the transmitter is 220 mW without the PRBS generator, and the power breakdown is presented in Table 8.2. The performance summary and comparison are presented in Table 8.3.

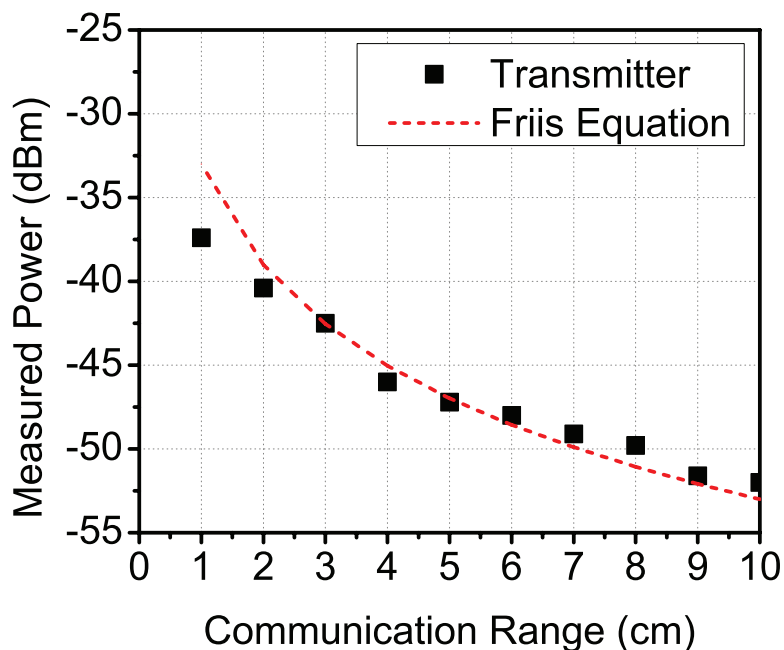


Figure 8.9: Measured output power of continuous-wave signals.

Table 8.3: Comparison of sub-terahertz transmitters

	This work [1]	Chapter 7 [48]	Wang [55]	Ojefors [56]	Han [57]	Schmalz [58]
Technology	65 nm CMOS	65 nm CMOS	32 nm SOI	130 nm SiGe	65 nm CMOS	130 nm SiGe
Modulation	QPSK	OOK	OOK	-	Pulse	-
Frequency (GHz)	240	260	210	220	260	245
Pout (dBm)	0	0.5	4.6	-1	0.5	1
EIRP (dBm)	1	5	5.1	-	15.7 (lens)	7
Pdc (mW)	220	688	240	630	800	379
Area ( $mm^2$ )	2	3	3.5	0.6	2.3	3.1
Antenna	on-chip	on-chip	on-chip	-	on-chip	on-chip
Data Rate (Gbps)	16	-	-	-	-	-
Efficiency (pJ/bit)	14	-	-	-	-	-

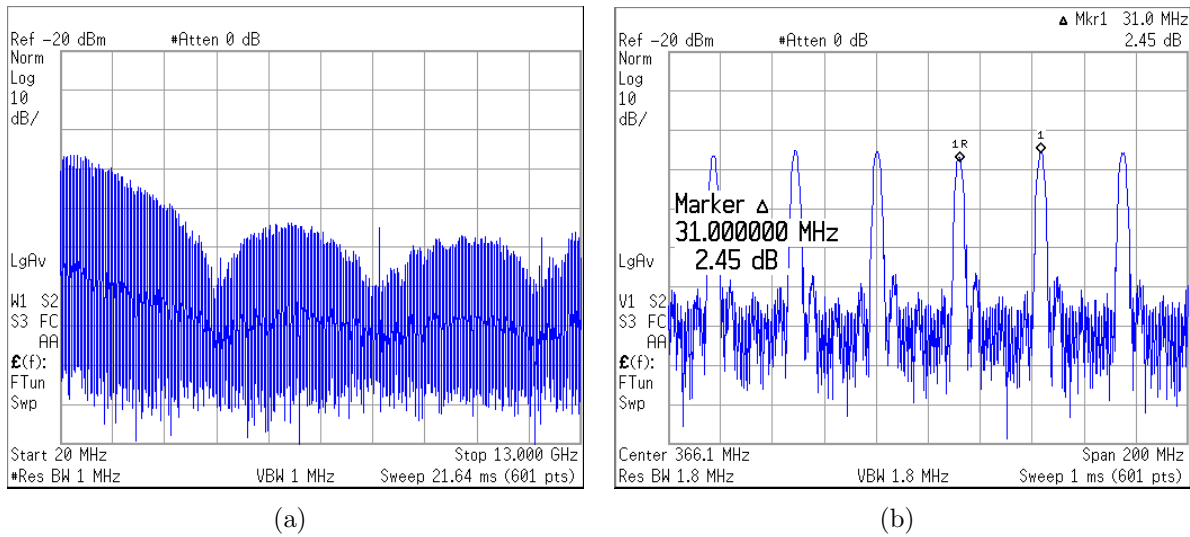


Figure 8.10: Measured frequency spectrum of 4 GSps modulation signal ( $f_{null,1} = 4$  GHz,  $f_{beat} = 31.5$  MHz).

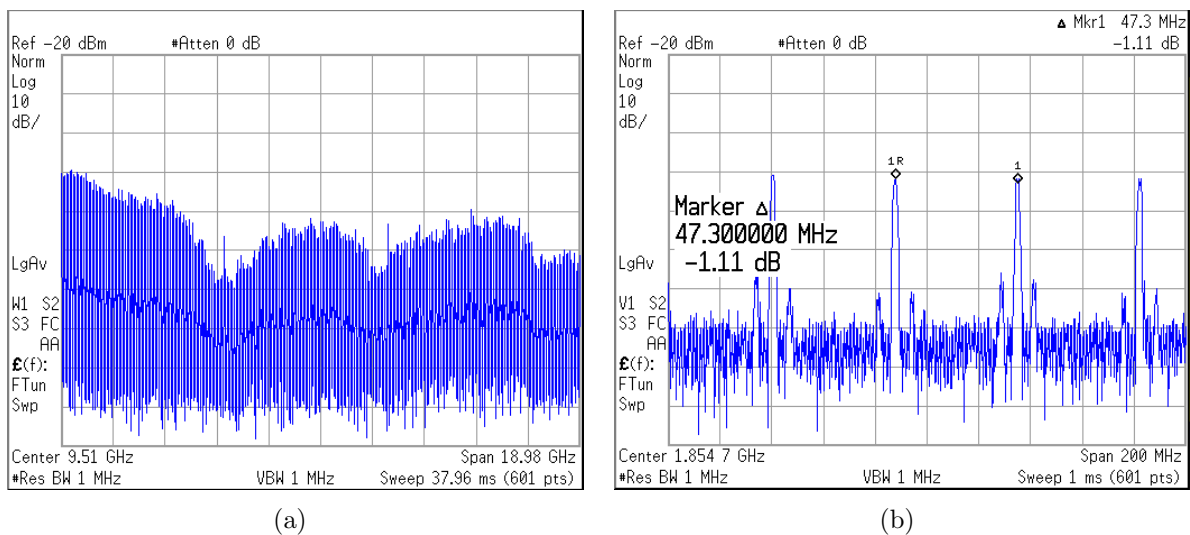


Figure 8.11: Measured frequency spectrum of 6 GSps modulation signal ( $f_{null,1} = 6$  GHz,  $f_{beat} = 47.2$  MHz).

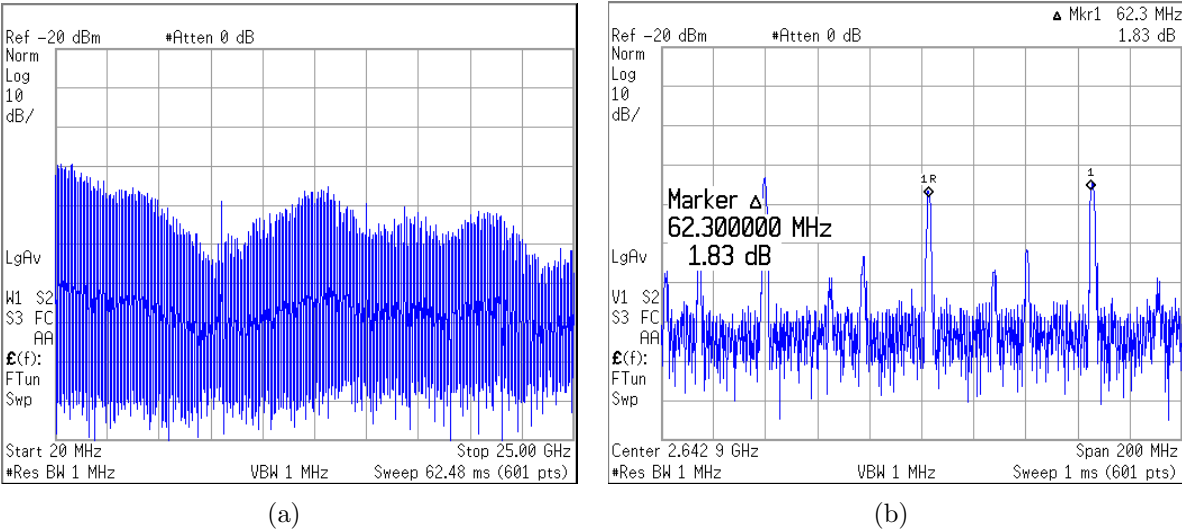


Figure 8.12: Measured frequency spectrum of 8 GSps modulation signal ( $f_{null,1} = 8\text{ GHz}$ ,  $f_{beat} = 63.0\text{ MHz}$ ).

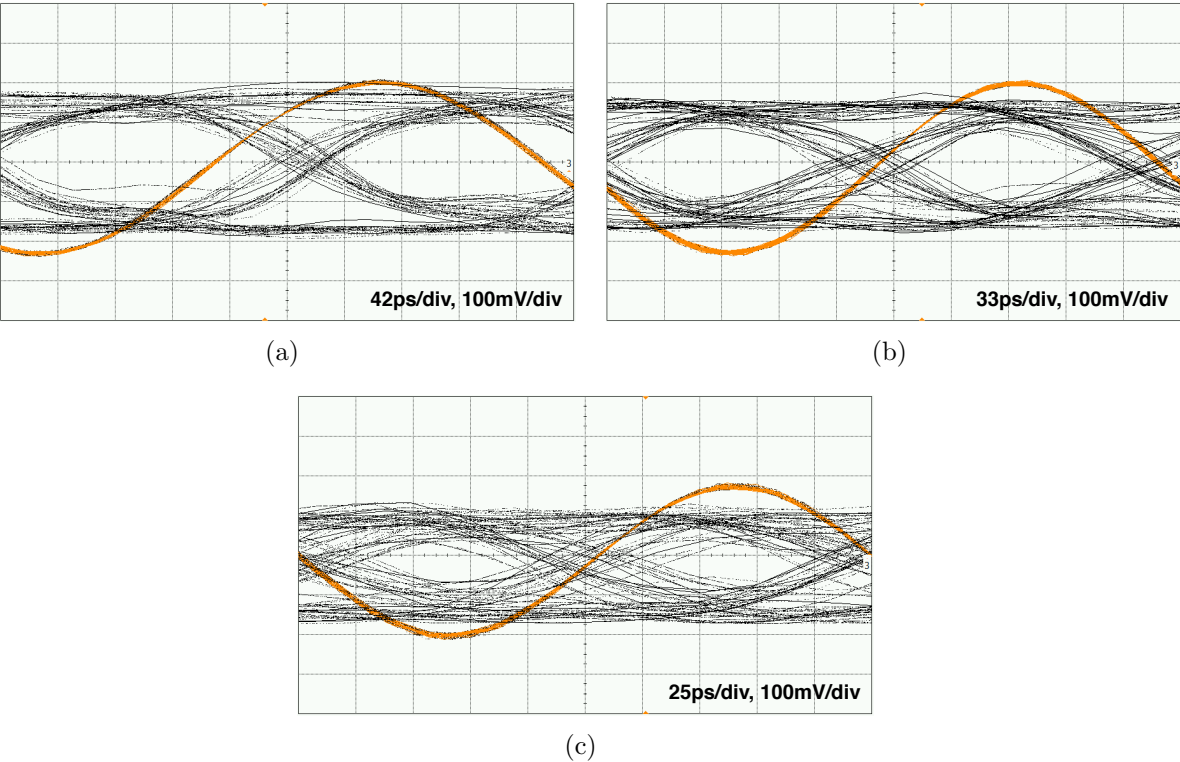


Figure 8.13: Measured transmitter eye diagrams (a) 4 GSps (b) 6 GSps (c) 8 GSps.

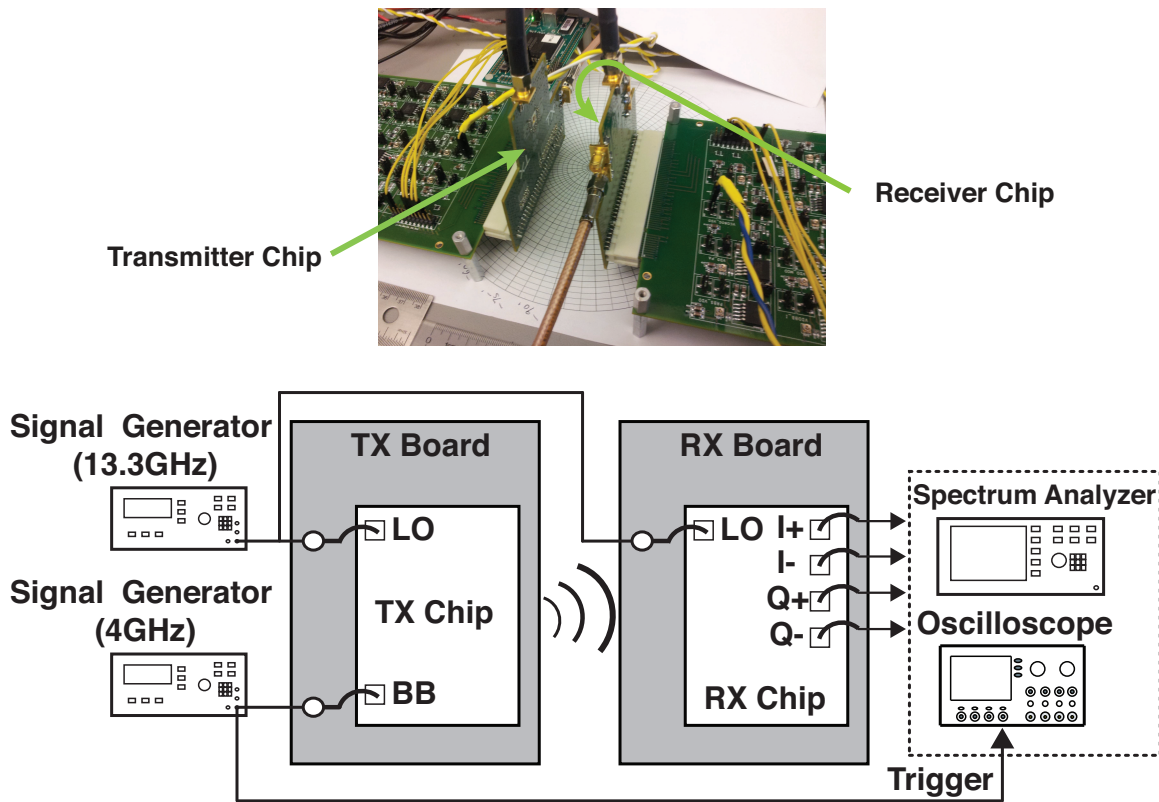


Figure 8.14: Measurement setup for the transmitter and receiver.

To measure the receiver and link performance, the measurement setup for the transceiver is shown in Fig. 8.14. Instead of using a horn antenna, the receiver chip captures transmitted signal. A spectrum analyzer and a high-speed oscilloscope are used to measure the receiver output. The 13.3 GHz input clock goes to both the transmitter and receiver to synchronize the two chips. The baseband data clock is used to trigger the oscilloscope to plot eye diagrams.

First, the continuous-wave mode is selected in the transmitter. The receiver output power is measured, varying with the range between the transmitter and receiver. The measured data are presented in Fig. 8.15. The output power matches the Friis equation. From these measurements, the receiver gain is approximately 25 dB and the (double sideband) noise figure is approximately 15 dB, matching well with the simulation results. The injection lock ranges are measured and presented in [2].

In Fig. 8.16(a), by varying the transmitter carrier frequency, the I and Q received powers are measured. The I and Q plots are symmetric, which means that the I and Q paths are well balanced. In Fig. 8.16(b), the receiver carrier frequency is changed, and the I and Q powers are measured at the receiver output. This is also symmetric.

With the BPSK modulation mode, time-domain eye diagrams are measured at the receiver output. 4 Gbps, 6 Gbps, 8 Gbps, and 9 Gbps modulations are successfully demon-

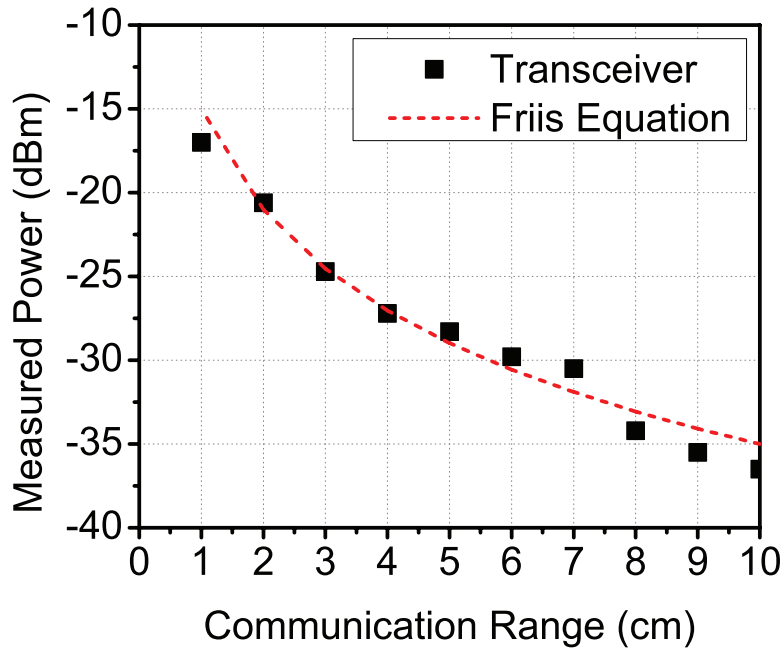


Figure 8.15: Measured output power of continuous-wave signals.

strated in Fig. 8.17. Using the data, the bathtub curves are plotted in Fig. 8.20, while the instrument limit is  $10^{-6}$ . With the 9 Gbps BPSK modulation, the BER is approximately  $10^{-5}$ .

Next, with the QPSK modulation mode, the frequency spectrums of the receiver output are measured as shown in Fig. 8.18. The eye diagrams are measured at the receiver output. 8 Gbps, 12 Gbps, and 16 Gbps QPSK modulations are shown in Fig. 8.19. As the data increases, the eye opening significantly degraded. The maximum data rate is 16 Gbps when the communication range is 1.5 cm. The BER bathtub curves are presented in Fig. 8.21. In the 16 Gbps QPSK modulation, the BER is approximately  $10^{-4}$ . The power dissipation of the receiver is 260 mW and the power breakdown is presented in Table 8.4. The summary and comparison of the receiver are presented in Table 8.5. The overall link performance is also summarized and compared with previous transceivers, as presented in Table 8.6.

Table 8.4: Power breakdown of the 240 GHz receiver

Building block	Power dissipation (mW)	Percentage (%)
80 GHz LO Chain	41	16
80 GHz Amplifiers	86	33
240 GHz Tripler	45	17
Baseband Amplifiers (I/Q)	78	30
Receiver	260	100

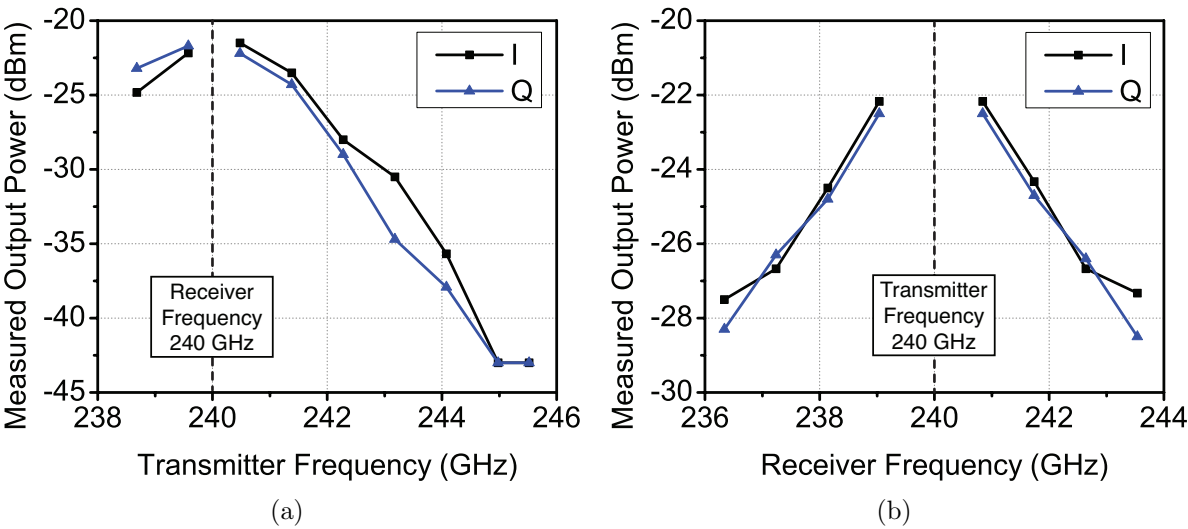


Figure 8.16: Measured I and Q output powers (a) varying transmitter frequency (b) varying receiver frequency.

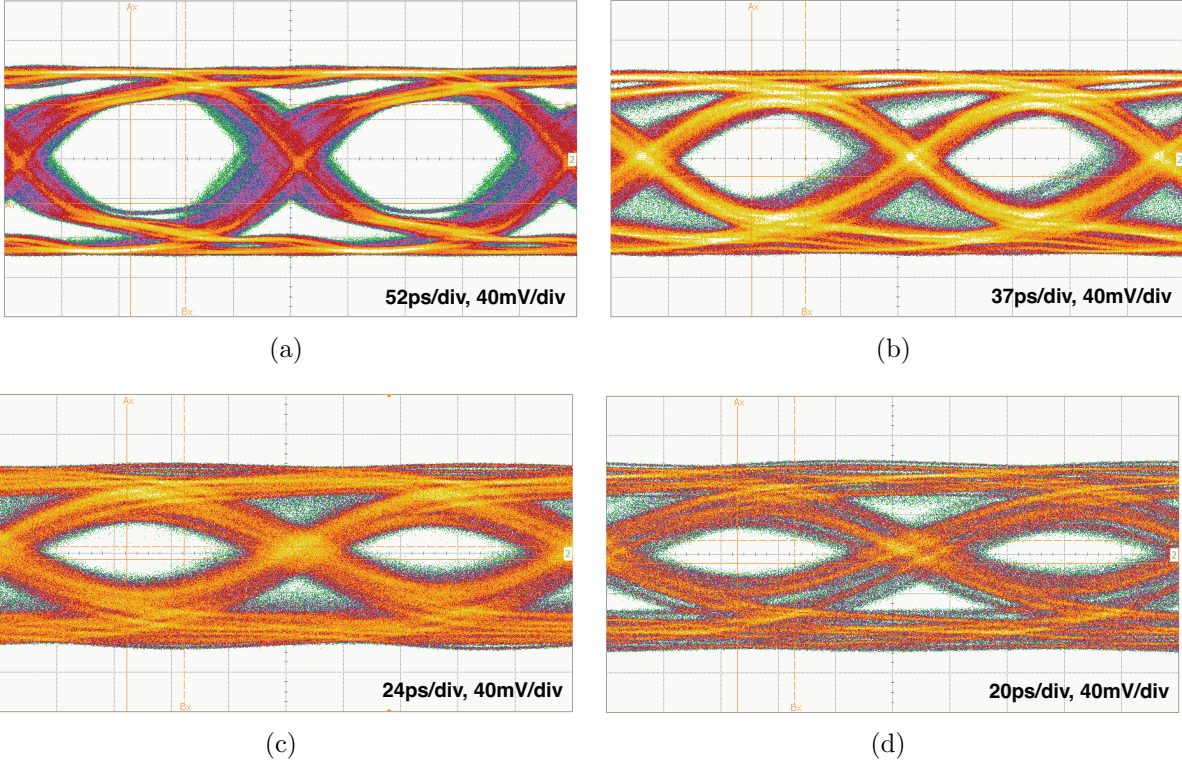


Figure 8.17: Measured receiver eye diagrams (a) 4 Gbps BPSK (b) 6 Gbps BPSK (c) 8 Gbps BPSK (d) 9 Gbps BPSK.



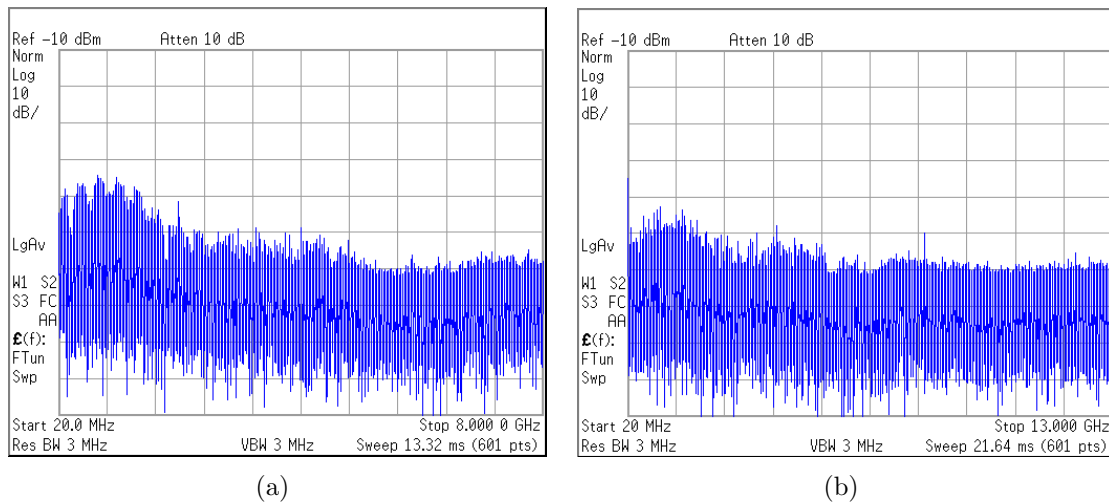


Figure 8.18: Measured receiver output spectra (a) 8 Gbps QPSK (b) 16 Gbps QPSK.

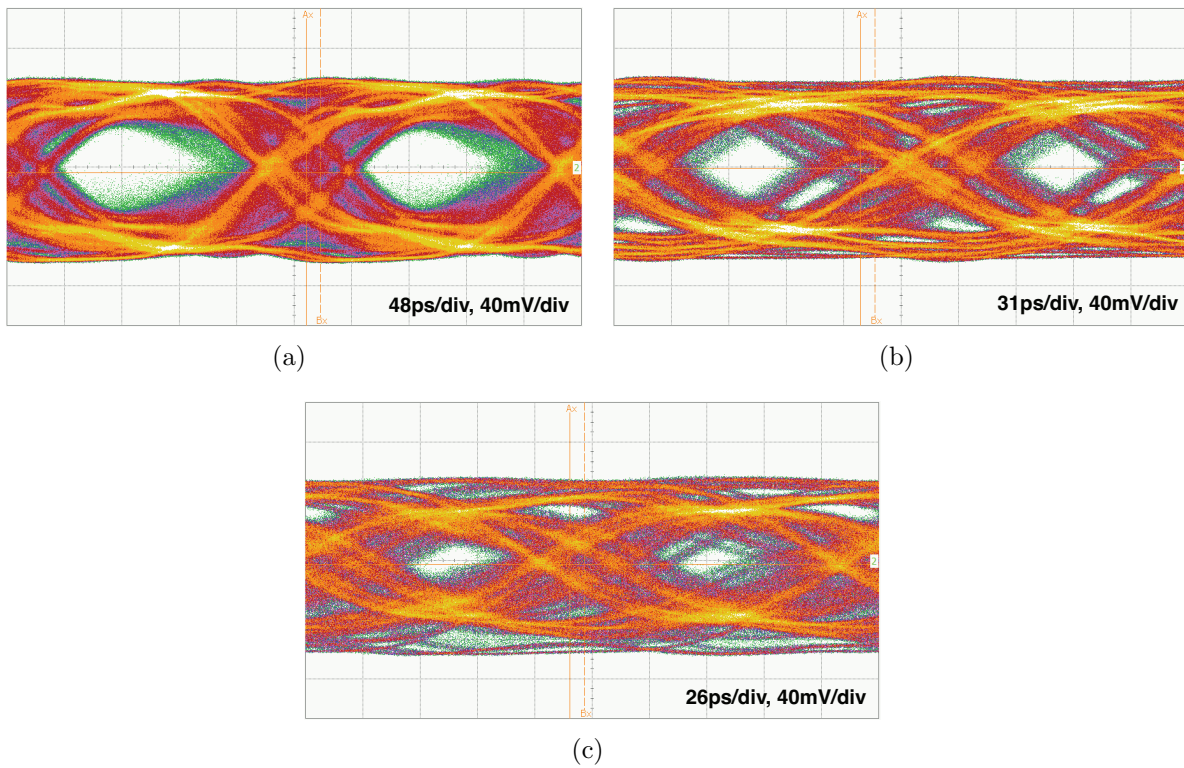


Figure 8.19: Measured receiver eye diagrams (a) 8 Gbps QPSK (b) 12 Gbps QPSK (c) 16 Gbps QPSK.

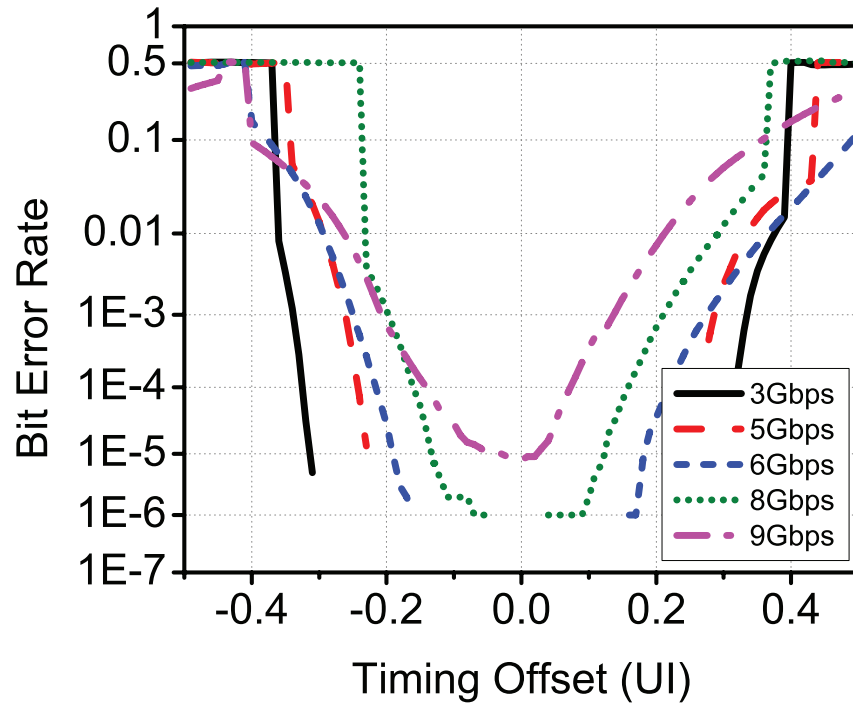


Figure 8.20: BER bathtub curves of BPSK modulation.

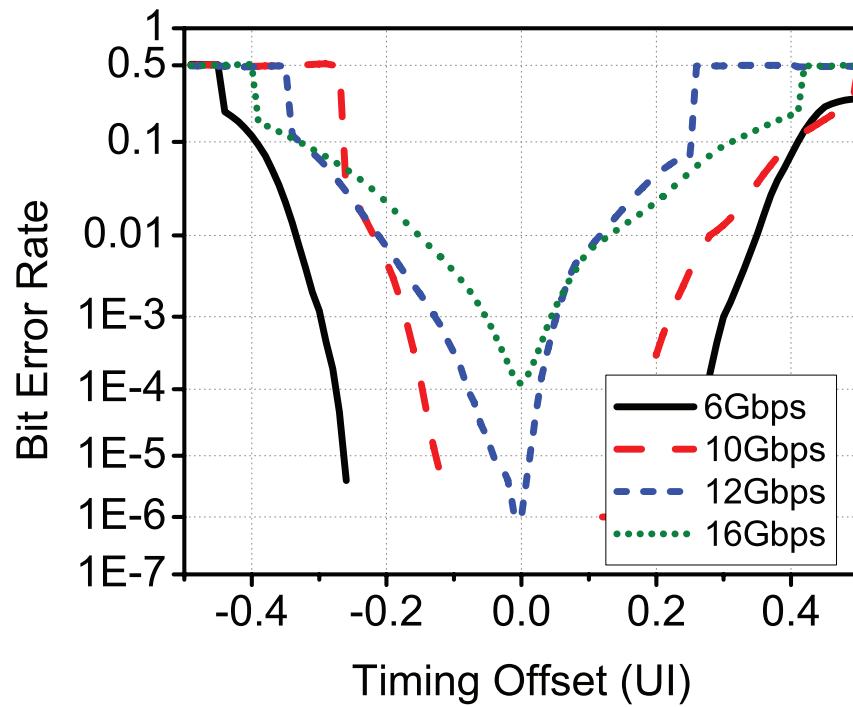


Figure 8.21: BER bathtub curves of QPSK modulation.

Table 8.5: Comparison of sub-terahertz receivers

	This work [2]	Chapter 7 [48]	Ojefors [59]	Elkhouly [60]	Guerra [61]
Technology	65 nm CMOS	65 nm CMOS	130 nm SiGe	130 nm SiGe	65 nm CMOS
Modulation	QPSK	OOK	-	I/Q	-
Frequency (GHz)	240	260	220	245	283
Gain (dB)	25	17	16	18	-6
NF (dB)	15	19	18	18	38
Pdc (mW)	260	485	216	512	97.6
Area ( $mm^2$ )	2	3	0.66	2.1	0.64
Antenna	on-chip	on-chip	-	-	-
Integration	Full	Full	LNA/Mixer	LNA/Mixer/ IF Amp/Hybrid	Mixer/LO/ IF Amp
Data Rate (Gbps)	16	-	-	-	-
Efficiency (pJ/bit)	16	-	-	-	-

Table 8.6: Comparison of sub-terahertz transceivers

	This work [1], [2]	Chapter 7 [48]	Wang [55]	Kallfass [62]	Song [63]	Jastrow [64]	Chia [65]
Technology	65 nm CMOS	65 nm CMOS	32 nm SOI	50 nm mHEMT	Photonics	Discrete	Discrete
Modulation	QPSK	OOK	OOK	QAM	ASK	QAM	BPSK
Frequency (GHz)	240	260	210	220	300	300	307
Pout (dBm)	0	0.5	4.6	1.4	-	-	-
EIRP (dBm)	1	5	5.1	-	30	-25	16.5
Pdc (mW)	480	1173	308	-	-	-	-
Area ( $mm^2$ )	4	6	4.62	3	-	-	-
Antenna	on-chip	on-chip	on-chip	off-chip +lens	on-chip +lens	off-chip +lens	off-chip
Antenna Gain	1.5 (TX) 0.7 (RX)	4.5	-	30 (lens)	34 (lens)	40	26
Range (cm)	1.5	-	-	50	50	5200	200
Data Rate (Gbps)	16	-	-	25	12.5	0.096	12.5
Efficiency (pJ/bit)	30	-	-	-	-	-	-

# Chapter 9

## Advanced Transceivers

This chapter presents advanced techniques for sub-terahertz transceivers. These techniques are widely used by low-frequency RF transceivers but are very challenging to implement in millimeter-wave/sub-terahertz transceivers yet. The 240 GHz QPSK/BPSK transceiver was successfully demonstrated in the previous chapter; thus, higher-order modulations and higher carrier frequencies are discussed here in this chapter. In addition, the phased-array system and the multiple-carrier system are discussed for high-speed wireless communications. The final section presents a lens that can increase the directivity and gain to extend the communication range.

### 9.1 Higher-Order Modulation

In Chapter 8, QPSK modulation is used to double the data rate. Likewise, higher-order modulations, such as 16-QAM or 64-QAM, can be used to increase the data rate for a given bandwidth, at the expense of the bit error rate. Fig. 9.1 illustrates the transmitter constellation points for QPSK and 16-QAM for the same peak signal power and noise power. A 16-QAM transceiver can achieve a higher data rate than QPSK but is more challenging to design. First, the bit error rate of the 16-QAM is higher than that of the QPSK. The analysis in Chapter 2 shows that the 16-QAM has a higher BER than the QPSK for the same  $E_b/N_o$ . In addition, the BER of the 16-QAM is also much higher than that of the QPSK for the same peak power and noise power, as shown in Fig. 9.1. Second, the 16-QAM requires a more stringent amplitude/phase balance of the I and Q LO signals. Third, the amplitude/phase distortions that are caused by the amplifiers and a frequency multiplier in the transmitter, significantly degrade the system performance. Pre-distortion may be employed but the performance is limited at such a high data rate. Therefore, these challenges should be addressed to use the higher-order modulations.

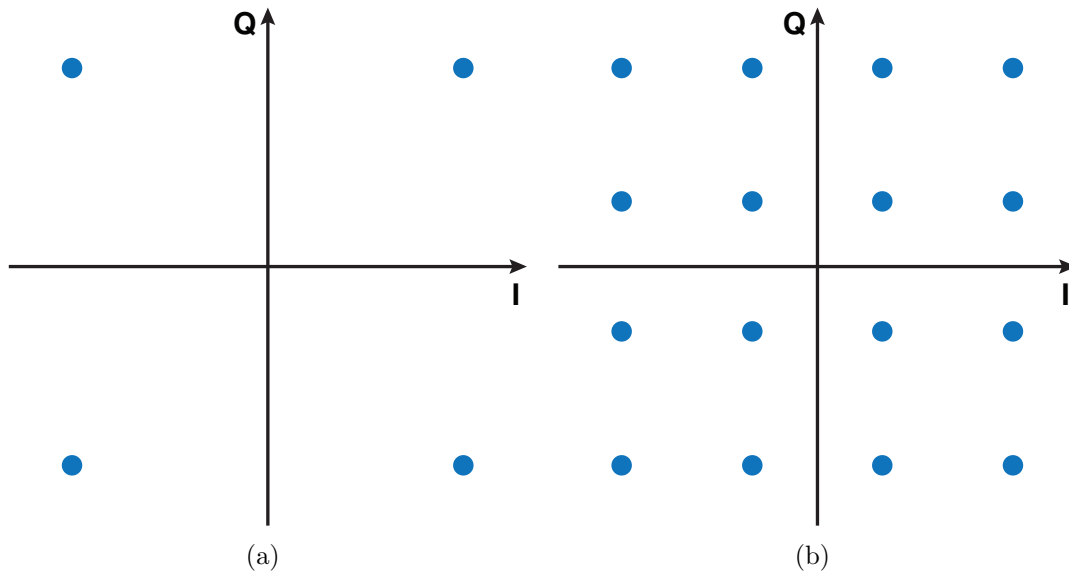


Figure 9.1: Constellation points (a) QPSK (b) 16-QAM.

## 9.2 Higher Carrier Frequency

There are advantages to increasing the carrier frequency above 240/260 GHz. As the frequency increases, the antenna size can shrink while maintaining the same antenna gain. Thus, the size of the on-chip antenna is reduced for a more compact overall form factor. The wider bandwidth can be achieved for the same quality factor at a higher carrier frequency because  $\Delta\omega = \omega/Q$ , thereby increasing the data rate. Numerous unallocated frequency bands remain above 300 GHz; thus, transceivers can utilize wide frequency bands.

However, there are many practical issues to address. Higher frequency increases the path loss, which is proportional to the square of the frequency and the distance. The distance should be reduced to maintain the same path loss. In the transmitter architecture, the multiplication ratio should be increased, producing a lower conversion gain of the frequency multiplier and a lower output power of the transmitter. At higher frequencies, passive devices are lossier, and the quality factors are lowered. Thus, a wider bandwidth may not be achieved, depending on the process technology. As such, for the same antenna gain, the antenna size may not shrink as theoretically predicted.

Therefore, the optimum carrier frequency should be found for a given process technology and system specifications. The PA output power, the conversion gain of frequency multipliers, and the noise figure of receivers should first be simulated and investigated for the frequencies of interest. Passive devices should also be considered, for example, antenna performance and transmission line losses. A link budget analysis can then be performed using the parameters to achieve the target data rate and BER.

### 9.3 Phased Array

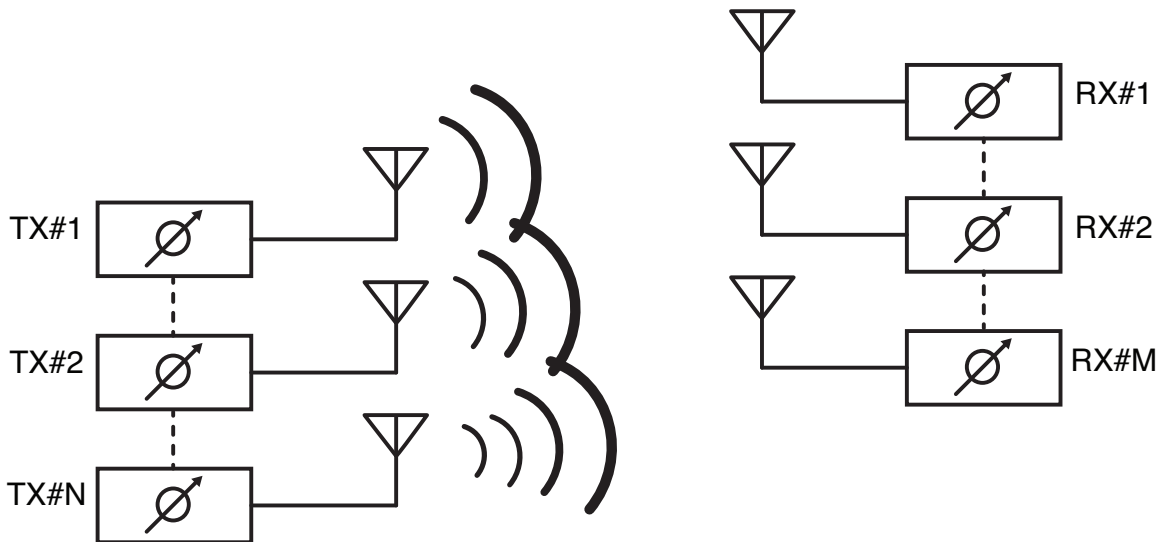


Figure 9.2: Phased array.

Fig. 9.2 shows a phased-array transmitter and a receiver ( $N \times M$ ). For  $N$  transmitters, the transmitter EIRP can be boosted by  $N^2$  because of the superposition of the electromagnetic waves. For  $M$  receivers, the receiver SNR can be boosted by  $M$  because the power of correlated signals is increased by  $M^2$ , and the power of the uncorrelated (assumed) noise is increased by  $M$ . The overall received SNR is increased by  $N^2M$ . The beam direction can be adjusted by controlling the phases of each transmitter or receiver. Thus, the directivity, the EIRP, and the received SNR are boosted, resulting in a longer communication range can be achieved and decreasing the BER. However, the phased array system does not increase the maximum data rate, and the energy efficiency (the ratio of the data rate to the power dissipation) drops by  $N$  (for the transmitter) or  $M$  (for the receiver). In addition, the phased array system requires circuits for rotating phases and synchronizing the LO signals.

### 9.4 Multiple Channels

Multiple channels or multiple carriers can be used to increase the data rate. As shown in Fig. 9.3, three carriers ( $f_1, f_2, f_3$ ) are used to triple the data rate, whereas the energy efficiency does not change. If the LO block is shared, then the efficiency improves.

The first drawback is that the receiver suffers from linearity issues. As for multi-channel TV tuners, IM2 and IM3 intermodulations can cause interference. The carrier frequencies should be properly selected to prevent interference issues, and the receiver antenna and circuits should be linear. The second drawback is that injection pulling may occur. There

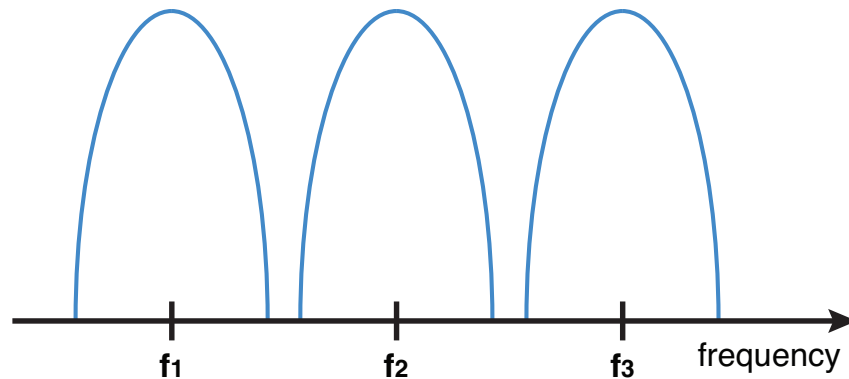


Figure 9.3: Multi-channel communication.

are three carriers and three strong signals; thus, mutual coupling through the substrate or any other coupling can cause injection pulling among the oscillators and power amplifiers.

## 9.5 Lens

It is well known that lenses refract electromagnetic waves. Light can be converged or diverged by refraction. This property can be used to extend the communication range. Fig. 9.4 illustrates how lenses can affect wave propagation. Fig. 9.4(a) shows that only one straight line can reach the receiver. However, Fig. 9.4(b) shows that plano-convex lenses converge waves, such that many other lines, as well as a straight line, can reach the receiver, thereby increasing the directivity and the EIRP depending on the lens size. The phased-array system increases the EIRP using additional transmitters and receivers, which requires more power dissipation. However, lenses can extend the communication range or increase the SNR without additional power dissipation.

The lens size determines the increment in the EIRP. To obtain a high gain, the lens size is typically large compared to the chip size, thereby increasing the overall form factor. In addition, the directivity increases; thus, if two chips are misaligned, the received signal power is degraded. Thus, the packaging should be carefully performed and verified.

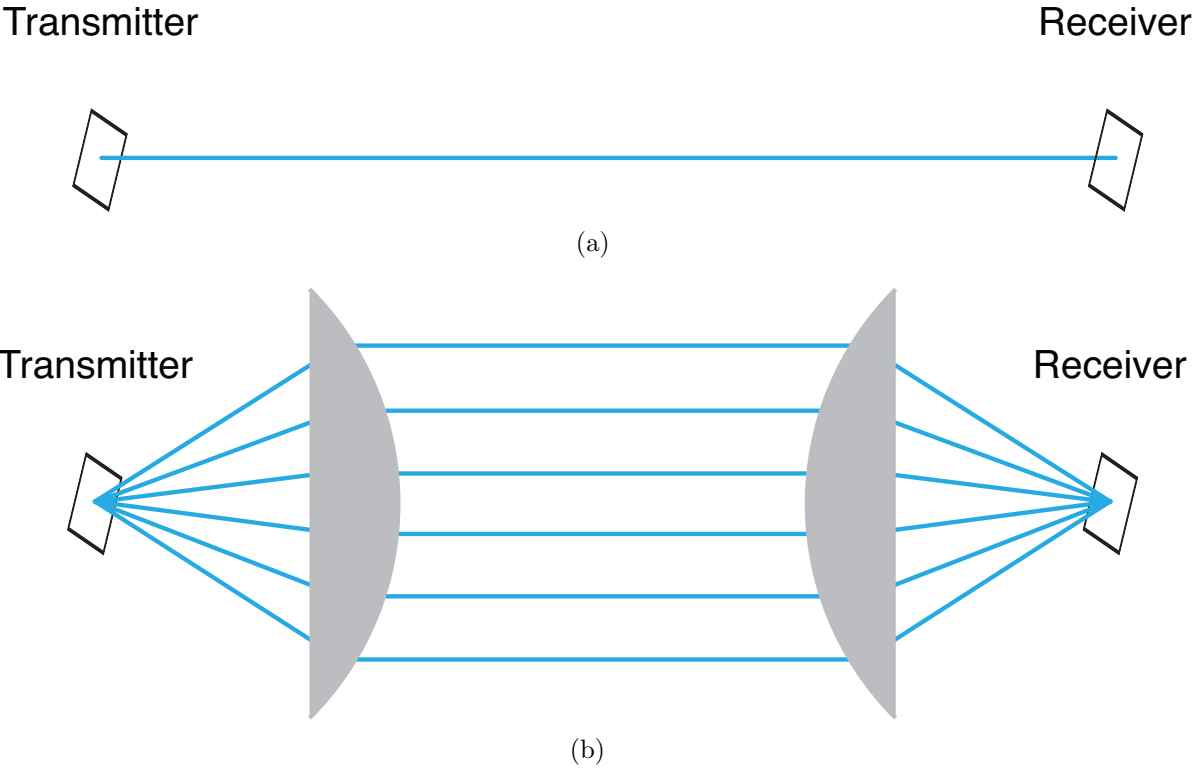


Figure 9.4: Ray diagrams (a) without lens (b) with lenses.



# Chapter 10

## Conclusion

In this dissertation, sub-terahertz wireless transceivers and building blocks are demonstrated for high-speed chip-to-chip communication. The 260 GHz OOK transceiver and 240 GHz QPSK/BPSK transceiver are fully integrated in 65 nm CMOS. In the QPSK modulation, a maximum data rate of 16 Gbps can be achieved with a range of 1.5 cm and an energy efficiency of 30 pJ/bit. The demonstration proves that sub-terahertz wideband wireless communication is feasible and practical in CMOS technology.

Chapter 2 introduced the communication theory (modulation, link budget analysis, and bit error rate calculation). The high channel loss (propagation loss and atmospheric loss) makes the sub-terahertz communication challenging. Previous sub-terahertz transceivers have been typically implemented with (quasi-)optical devices, III-V semiconductors, or SiGe, not with CMOS. For millimeter-wave/sub-terahertz circuits and systems, CMOS technology has many disadvantages and limitations, which were summarized.

Part I discussed the building blocks for millimeter-wave/sub-terahertz wireless transceivers. Transmitter and receiver architectures were also analyzed. This part should be helpful to one who focuses on specific block designs and circuit-level ideas and performance.

Chapter 3 focused on frequency generation techniques. First, a millimeter-wave VCO was discussed. To tackle the poor quality factors of passive varactors at high frequencies, an active-varactor VCO was proposed and demonstrated. Injection-locking and a bi-directionally injection-locked loop were then discussed. Frequency multiplication is an important function in sub-terahertz transceivers; thus, several ideas were explored and verified. Finally, frequency synthesizer architectures were introduced and LO distribution was also discussed.

Chapter 4 discussed a typical I/Q transmitter architecture and several architecture candidates suitable for sub-terahertz transmitters. I/Q quadrature phase generation was first described. In addition, digital data modulations were discussed and a modulator design was introduced. The transmitters should achieve high output power and wide bandwidth for high-speed communications. The driving and power amplifiers were designed and implemented. The wideband frequency tripler was designed for the 240 GHz QPSK transmitter.

Chapter 5 first investigated a typical I/Q receiver architecture and several architecture

candidates suitable for sub-terahertz receivers, similar to Chapter 4. Typically, the mixer-first architecture is employed because amplifiers have no gain above cut-off frequencies. As the first down-converting mixer, a 260 GHz double-balanced mixer was designed for the 260 GHz OOK receiver. Demodulation was also analyzed, and some techniques were presented for robust receiver design.

Chapter 6 discussed the baseband circuits used for high-speed wireless transceivers. The current-mode logic gates were investigated. Using the logic gates, PRBS generators were designed and implemented to test the digital data communication. Receiver baseband amplifiers are also important to achieve wide bandwidth, low noise figure, and high SNR. Lastly, an operational amplifier, widely used in analog/RF circuit designs, was discussed.

Part II demonstrated fully integrated sub-terahertz transceivers. It also provided link-budget analysis, system architecture, and link measurement results. Two wireless transceivers (260 GHz OOK transceiver and 240 GHz QPSK/BPSK transceiver) are described in the system level. This part is valuable for system and application designers, as well as integrated circuit designers.

Chapter 7 focused on the 260 GHz OOK transceiver. The link budget was examined and real implementation issues in the system level were discussed. The transmitter and receiver architectures were proposed, and the measurement setups and results were presented.

Chapter 8 introduced the 240 GHz QPSK/BPSK transceiver. First, the system challenges and issues were described and design challenges and considerations were presented. The transmitter and receiver architectures were proposed for high-speed coherent phase modulations. Front-end layout designs were also illustrated. Several measurement setups were shown for measuring the transmitter and receiver. Moreover, many frequency spectrums, time-domain eye diagrams, and other plots were provided. The BER bathtub curves and comparison tables were also presented. The 240 GHz QPSK/BPSK transceiver achieves a data rate of 16 Gbps in the QPSK modulation with a range of 1.5 cm and an energy efficiency of 30 pJ/bit. To date, the 240 GHz transceiver achieves the highest data rate and energy efficiency among CMOS transceivers above 200 GHz.

Chapter 9 included future directions and ideas for advanced sub-terahertz transceivers. Higher carrier frequency and better performance can be achieved with more advanced CMOS technologies. Beamforming and multiple lanes can be utilized. In addition, (quasi-)optical lenses can make the transmit EIRP higher and the communication range longer. Thus, co-design with lenses should be investigated.

Overall, the millimeter-wave/sub-terahertz transceivers and building blocks are successfully demonstrated for high-speed wireless communications in CMOS technology. The dissertation contributes ideas and techniques to the design of wideband transceivers and other relevant systems.

# Bibliography

- [1] S. Kang, S. Thyagarajan, and A. Niknejad, “A 240GHz Wideband QPSK Transmitter in 65nm CMOS,” in *Radio Frequency Integrated Circuits Symposium (RFIC), 2014 IEEE*, June 2014.
- [2] S. Thyagarajan, S. Kang, and A. Niknejad, “A 240GHz Wideband QPSK Receiver in 65nm CMOS,” in *Radio Frequency Integrated Circuits Symposium (RFIC), 2014 IEEE*, June 2014.
- [3] 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 J. Solid-State Circuits*, vol. 45, no. 12, pp. 2667–2681, Dec. 2010.
- [4] M. Urteaga, R. Pierson, P. Rowell, V. Jain, E. Lobisser, and M. Rodwell, “130nm InP DHBTs with  $f_T > 0.52\text{THz}$  and  $f_{max} > 1.1\text{THz}$ ,” in *Device Research Conference (DRC), 2011 69th Annual*, June 2011, pp. 281–282.
- [5] B. Heinemann *et al.*, “SiGe HBT technology with  $f_T/f_{max}$  of 300GHz/500GHz and 2.0 ps CML gate delay,” in *Electron Devices Meeting (IEDM), 2010 IEEE International*, Dec 2010, pp. 30.5.1–30.5.4.
- [6] M. Fujishima, “Power-efficient ultrahigh-speed CMOS wireless communications in terahertz era,” in *Electron Devices Meeting (IEDM), 2013 IEEE International*, Dec 2013, pp. 17.6.1–17.6.4.
- [7] M. Schwartz, W. R. Bennett, and S. Stein, *Communication systems and techniques*. John Wiley & Sons, 1995.
- [8] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. McGraw–Hill, New York, 2008.
- [9] B. Sklar, *Digital Communications Fundamentals and Applications*, 2nd ed. Prentice Hall, Inc, 2001.
- [10] *am* Atmospheric Model. [Online]. Available: <https://www.cfa.harvard.edu/~spaine/am/>

- [11] A. M. Niknejad *et al.*, “Compact modeling for RF and microwave integrated circuits,” in *Workshop on Compact Modeling*, 2003.
- [12] A. M. Niknejad, *Electromagnetics for High-Speed Analog and Digital Communication Circuits*. Cambridge University Press Cambridge, 2007.
- [13] A. Niknejad, D. Chowdhury, and J. Chen, “Design of CMOS Power Amplifiers,” *IEEE Trans. Microw. Theory Tech.*, vol. 60, no. 6, pp. 1784–1796, June 2012.
- [14] D. Leeson, “A simple model of feedback oscillator noise spectrum,” *Proc. IEEE*, vol. 54, no. 2, pp. 329–330, Feb. 1966.
- [15] M. Demirkan, S. Bruss, and R. Spencer, “Design of Wide Tuning-Range CMOS VCOs Using Switched Coupled-Inductors,” *IEEE J. Solid-State Circuits*, vol. 43, no. 5, pp. 1156–1163, May 2008.
- [16] M. Tiebout, “A CMOS fully integrated 1 GHz and 2 GHz dual band VCO with a voltage controlled inductor,” in *Solid-State Circuits Conference, 2002. ESSCIRC 2002. Proceedings of the 28th European*, Sept 2002, pp. 799–802.
- [17] P. Andreani, “A 1.8-GHz monolithic CMOS VCO tuned by an inductive varactor,” in *Circuits and Systems, 2001. ISCAS 2001. The 2001 IEEE International Symposium on*, vol. 4, May 2001, pp. 714–717 vol. 4.
- [18] K.-C. Kwok and J. Long, “A 23-to-29 GHz Transconductor-Tuned VCO MMIC in 0.13  $\mu\text{m}$  CMOS,” *IEEE J. Solid-State Circuits*, vol. 42, no. 12, pp. 2878–2886, Dec 2007.
- [19] G. Cusmai, M. Reposi, G. Albasini, A. Mazzanti, and F. Svelto, “A Magnetically Tuned Quadrature Oscillator,” *IEEE J. Solid-State Circuits*, vol. 42, no. 12, pp. 2870–2877, Dec 2007.
- [20] D. Pi, B.-K. Chun, and P. Heydari, “A 2.5-3.2GHz CMOS differentially-controlled continuously-tuned varactor-less LC-VCO,” in *Solid-State Circuits Conference, 2007. ASSCC '07. IEEE Asian*, Nov 2007, pp. 111–114.
- [21] Y. Chen and K. Mouthaan, “Wideband Varactorless LC VCO Using a Tunable Negative-Inductance Cell,” *IEEE Trans. Circuits Syst. I*, vol. 57, no. 10, pp. 2609–2617, Oct 2010.
- [22] P.-Y. Chiang, O. Momeni, and P. Heydari, “A 200-GHz Inductively Tuned VCO With -7-dBm Output Power in 130-nm SiGe BiCMOS,” *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 10, pp. 3666–3673, Oct 2013.
- [23] M. A. T. Sanduleanu and E. Stikvoort, “Highly linear, varactor-less, 24GHz IQ oscillator,” in *Radio Frequency integrated Circuits (RFIC) Symposium, 2005. Digest of Papers. 2005 IEEE*, June 2005, pp. 577–580.

- [24] S. Kang and A. Niknejad, "A 100GHz active-varactor VCO and a bi-directionally injection-locked loop in 65nm CMOS," in *Radio Frequency Integrated Circuits Symposium (RFIC), 2013 IEEE*, June 2013, pp. 231–234.
- [25] Z.-M. Tsai, C.-S. Lin, C. Huang, J. Chern, and H. Wang, "A Fundamental 90-GHz CMOS VCO Using New Ring-Coupled Quad," *IEEE Microw. Wireless Compon. Lett.*, vol. 17, no. 3, pp. 226–228, March 2007.
- [26] E. Laskin, M. Khanpour, R. Aroca, K. Tang, P. Garcia, and S. Voinigescu, "A 95GHz Receiver with Fundamental-Frequency VCO and Static Frequency Divider in 65nm Digital CMOS," in *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*, Feb 2008, pp. 180–605.
- [27] W. Volkaerts, M. Steyaert, and P. Reynaert, "118GHz fundamental VCO with 7.8% tuning range in 65nm CMOS," in *Radio Frequency Integrated Circuits Symposium (RFIC), 2011 IEEE*, June 2011.
- [28] M. Hekmat, D. Su, and B. Wooley, "A quadrature LO generator using bidirectionally-coupled oscillators for 60-GHz applications," in *Custom Integrated Circuits Conference (CICC), 2011 IEEE*, Sept 2011.
- [29] I.-S. Shen and C. Jou, "A X -Band Capacitor-Coupled QVCO Using Sinusoidal Current Bias Technique," *IEEE Trans. Microw. Theory Tech.*, vol. 60, no. 2, pp. 318–328, Feb 2012.
- [30] B. Çatlı and M. Hella, "Triple-Push Operation for Combined Oscillation/Divison Functionality in Millimeter-Wave Frequency Synthesizers," *IEEE J. Solid-State Circuits*, vol. 45, no. 8, pp. 1575–1589, Aug. 2010.
- [31] J. Grzyb, Y. Zhao, and U. Pfeiffer, "A 288-GHz Lens-Integrated Balanced Triple-Push Source in a 65-nm CMOS Technology," *IEEE J. Solid-State Circuits*, vol. 48, no. 7, pp. 1751–1761, July 2013.
- [32] O. Momeni and E. Afshari, "High Power Terahertz and Millimeter-Wave Oscillator Design: A Systematic Approach," *IEEE J. Solid-State Circuits*, vol. 46, no. 3, pp. 583–597, March 2011.
- [33] S. Kang, "A W-Band Phase-Locked Loop for Millimeter-Wave Applications," University of California at Berkeley, Tech. Rep., 2013.
- [34] S. Kang, J.-C. Chien, and A. Niknejad, "A 100GHz phase-locked loop in 0.13 $\mu$ m SiGe BiCMOS process," in *Radio Frequency Integrated Circuits Symposium (RFIC), 2011 IEEE*, June 2011.
- [35] J. Lee, M. Liu, and H. Wang, "A 75-GHz Phase-Locked Loop in 90-nm CMOS Technology," *IEEE J. Solid-State Circuits*, vol. 43, no. 6, pp. 1414–1426, June 2008.

- [36] S. Shahramian *et al.*, “Design of a Dual W- and D-Band PLL,” *IEEE J. Solid-State Circuits*, vol. 46, no. 5, pp. 1011–1022, May 2011.
- [37] K.-H. Tsai and S.-I. Liu, “A 43.7mW 96GHz PLL in 65nm CMOS,” in *Solid-State Circuits Conference - Digest of Technical Papers, 2009. ISSCC 2009. IEEE International*, Feb. 2009, pp. 276–277,277a.
- [38] C. Cao, Y. Ding, and K. O, “A 50-GHz Phase-Locked Loop in 0.13- $\mu$ m CMOS,” *IEEE J. Solid-State Circuits*, vol. 42, no. 8, pp. 1649–1656, Aug. 2007.
- [39] M. Abdul-Latif, M. Elsayed, and E. Sanchez-Sinencio, “A Wideband Millimeter-Wave Frequency Synthesis Architecture Using Multi-Order Harmonic-Synthesis and Variable-Push Frequency Multiplication,” *IEEE J. Solid-State Circuits*, vol. 46, no. 6, pp. 1265–1283, June 2011.
- [40] C.-C. Wang, Z. Chen, and P. Heydari, “W-Band Silicon-Based Frequency Synthesizers Using Injection-Locked and Harmonic Triplers,” *IEEE Trans. Microw. Theory Tech.*, vol. 60, no. 5, pp. 1307–1320, May 2012.
- [41] A. Musa, R. Murakami, T. Sato, W. Chaivipas, K. Okada, and A. Matsuzawa, “A Low Phase Noise Quadrature Injection Locked Frequency Synthesizer for MM-Wave Applications,” *IEEE J. Solid-State Circuits*, vol. 46, no. 11, pp. 2635–2649, Nov. 2011.
- [42] B. Razavi, “A study of injection locking and pulling in oscillators,” *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1415–1424, Sept. 2004.
- [43] C. Doan, S. Emami, A. Niknejad, and R. Brodersen, “Millimeter-wave CMOS design,” *IEEE J. Solid-State Circuits*, vol. 40, no. 1, pp. 144–155, Jan. 2005.
- [44] M. Tabesh *et al.*, “A 65 nm CMOS 4-Element Sub-34 mW/Element 60 GHz Phased-Array Transceiver,” *IEEE J. Solid-State Circuits*, vol. 46, no. 12, pp. 3018–3032, Dec. 2011.
- [45] D. M. Pozar, *Microwave engineering*, 3rd ed. John Wiley&Son In, 2005.
- [46] A. Arbabian, S. Callender, S. Kang, M. Rangwala, and A. Niknejad, “A 94 GHz mm-Wave-to-Baseband Pulsed-Radar Transceiver with Applications in Imaging and Gesture Recognition,” *IEEE J. Solid-State Circuits*, vol. 48, no. 4, pp. 1055–1071, 2013.
- [47] A. Arbabian, S. Kang, S. Callender, B. Afshar, J.-C. Chien, and A. Niknejad, “A 90GHz pulsed-transmitter with near-field/far-field energy cancellation using a dual-loop antenna,” in *Radio Frequency Integrated Circuits Symposium (RFIC), 2011 IEEE*, June 2011.
- [48] J.-D. Park, S. Kang, S. Thyagarajan, E. Alon, and A. Niknejad, “A 260 GHz fully integrated CMOS transceiver for wireless chip-to-chip communication,” in *VLSI Circuits (VLSIC), 2012 Symposium on*, June 2012, pp. 48–49.

- [49] S. Sankaran and K. O, "Schottky barrier diodes for millimeter wave detection in a foundry CMOS process," *IEEE Electron Device Lett.*, vol. 26, no. 7, pp. 492–494, July 2005.
- [50] R. Han *et al.*, "A 280-GHz Schottky Diode Detector in 130-nm Digital CMOS," *IEEE J. Solid-State Circuits*, vol. 46, no. 11, pp. 2602–2612, Nov 2011.
- [51] M. Chen and J. Notthoff, "A 3.3-V 21-Gb/s PRBS generator in AlGaAs/GaAs HBT technology," *IEEE J. Solid-State Circuits*, vol. 35, no. 9, pp. 1266–1270, Sept 2000.
- [52] M. Pelgrom, A. C. J. Duinmaijer, and A. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, Oct 1989.
- [53] L. Lewyn, T. Ytterdal, C. Wulff, and K. Martin, "Analog Circuit Design in Nanoscale CMOS Technologies," *Proc. IEEE*, vol. 97, no. 10, pp. 1687–1714, Oct 2009.
- [54] S. Mohan, M. Hershenson, S. Boyd, and T. Lee, "Bandwidth extension in CMOS with optimized on-chip inductors," *IEEE J. Solid-State Circuits*, vol. 35, no. 3, pp. 346–355, March 2000.
- [55] Z. Wang, P.-Y. Chiang, P. Nazari, C.-C. Wang, Z. Chen, and P. Heydari, "A CMOS 210-GHz Fundamental Transceiver With OOK Modulation," *IEEE J. Solid-State Circuits*, vol. 49, no. 3, pp. 564–580, March 2014.
- [56] E. Ojefors, B. Heinemann, and U. Pfeiffer, "Active 220- and 325-GHz Frequency Multiplier Chains in an SiGe HBT Technology," *IEEE Trans. Microw. Theory Tech.*, vol. 59, no. 5, pp. 1311–1318, May 2011.
- [57] R. Han and E. Afshari, "A CMOS High-Power Broadband 260-GHz Radiator Array for Spectroscopy," *IEEE J. Solid-State Circuits*, vol. 48, no. 12, pp. 3090–3104, Dec 2013.
- [58] K. Schmalz, R. Wang, J. Borngraber, W. Debski, W. Winkler, and C. Meliani, "245 GHz SiGe transmitter with integrated antenna and external PLL," in *Microwave Symposium Digest (IMS), 2013 IEEE MTT-S International*, June 2013.
- [59] E. Ojefors, B. Heinemann, and U. Pfeiffer, "Subharmonic 220- and 320-GHz SiGe HBT Receiver Front-Ends," *IEEE Trans. Microw. Theory Tech.*, vol. 60, no. 5, pp. 1397–1404, May 2012.
- [60] M. Elkhoully, Y. Mao, S. Glisic, C. Meliani, F. Ellinger, and J. Scheytt, "A 240 GHz direct conversion IQ receiver in 0.13  $\mu\text{m}$  SiGe BiCMOS technology," in *Radio Frequency Integrated Circuits Symposium (RFIC), 2013 IEEE*, June 2013, pp. 305–308.
- [61] J. Guerra, A. Siligaris, J.-F. Lampin, F. Danneville, and P. Vincent, "A 283 GHz low power heterodyne receiver with on-chip local oscillator in 65 nm CMOS process," in *Radio Frequency Integrated Circuits Symposium (RFIC), 2013 IEEE*, June 2013, pp. 301–304.

- [62] I. Kallfass *et al.*, “All Active MMIC-Based Wireless Communication at 220 GHz,” *IEEE Trans. Terahertz Sci. Tech.*, vol. 1, no. 2, pp. 477–487, Nov 2011.
- [63] H.-J. Song *et al.*, “Terahertz wireless communication link at 300 GHz,” in *Microwave Photonics (MWP), 2010 IEEE Topical Meeting on*, Oct 2010, pp. 42–45.
- [64] C. Jastrow *et al.*, “Wireless digital data transmission at 300 GHz,” *Electronics Letters*, vol. 46, no. 9, pp. 661–663, April 2010.
- [65] M. Y. W. Chia, C. Ang, B. Luo, and S. Leong, “Wideband 307 GHz transceiver system for high speed digital wireless at 12.5 Gbps,” in *Microwave Symposium Digest (MTT), 2011 IEEE MTT-S International*, June 2011.