Title
Integrated Circuits and Systems for Millimeter-Wave Frequencies

Permalink
https://escholarship.org/uc/item/3sx3g52g

Author
Mohammadnezhad, Seyed Mohammad Hossein

Publication Date
2019

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, IRVINE

Integrated Circuits and Systems for Millimeter-Wave Frequencies

DISSertation

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Electrical Engineering

by

Seyed Mohammad Hossein Mohammadnezhad

Dissertation Committee:
Professor Payam Heydari, Chair
Professor Nader Bagherzadeh
Professor Tony Givargis

2019
DEDICATION

To My Family
# TABLE OF CONTENTS

**LIST OF FIGURES**

**LIST OF TABLES**

**ACKNOWLEDGMENTS**

**CURRICULUM VITAE**

**ABSTRACT OF THE DISSERTATION**

1 A Millimeter-Wave Partially-Overlapped Beamforming-MIMO Receiver:
   Theory, Design, and Implementation  1
   1.1 Introduction to mm-Wave Wireless Communication Networks  1
   1.2 All-analog, All-digital, and Hybrid Architectures  3
   1.3 Conventional Hybrid and the Proposed Partially-Overlapped Hybrid Architectures  4
   1.4 Phase Control and Mainlobe Steering in Partially-Overlapped Hybrid Architecture  8
   1.5 Amplitude Control and Null-Steering in Partially-Overlapped Hybrid Architecture  9
      1.5.1 Null-steering with Phase Shifters  9
      1.5.2 Null-steering with VGAs  10
   1.6 Effects of Amplitude and Phase Errors on Interference Suppression  12
   1.7 The Phase-Amplitude Controlled Partially-Overlapped Hybrid Circuit Architecture  14
      1.7.1 Low Noise Amplifier (LNA)  16
      1.7.2 Phase Shifter (PS)  17
      1.7.3 Variable Gain Attenuator (VGA)  20
      1.7.4 Measurement Results of RF and RF-to-BB Channels  24
      1.7.5 Coupling between RF-Channels  25
   1.8 Measured Null-steering and Spatial Multiplexing  28
   1.9 Conclusion  33
# A Millimeter-Wave Energy-Efficient Direct-Demodulation Receiver: Theory, Design, and Implementation

## 2.1 Introduction to High-Speed mm-Wave Receivers

## 2.2 High-speed ADC Design Challenges

## 2.3 High-Order Direct-Demodulation

### 2.3.1 Current 8PSK Demodulation Techniques

### 2.3.2 Proposed 8PSK Direct-Demodulation Technique

### 2.3.3 BER of the proposed 8PSK direct-demodulation Method

## 2.4 Proposed 8PSK Direct-Demodulation Receiver

### 2.4.1 LNA circuit design

### 2.4.2 1-to-4 balun-based splitter design

### 2.4.3 Mixer-baseband circuit design

### 2.4.4 Comparator

### 2.4.5 LO Generation and Distribution Network

## 2.5 Measurement Results

## 2.6 Conclusion

## Bibliography
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>Typical hybrid system architectures: (a) sub-array, (b) full-array.</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>Conceptual representation of the hybrid system with partially-overlapped</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>clusters.</td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>(a) $M$ and $N_{RF}$ as function of $D$ for $N = 8$ elements, (b) $(M, N_{RF}, D)$</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>combinations for $N = 4$ and $N = 8$ elements.</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>Outage capacity for full-array, sub-array, and partially-overlapped array with</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>$N = 4$ and $N = 8$ elements.</td>
<td></td>
</tr>
<tr>
<td>1.6</td>
<td>(a) Array factor nulls based on phase shifter control, (b) Array factor nulls</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>based on VGA control.</td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>Effect of amplitude and phase error on weight vector.</td>
<td>13</td>
</tr>
<tr>
<td>1.8</td>
<td>Interference rejection as a function of RMS amplitude error and RMS phase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>error.</td>
<td></td>
</tr>
<tr>
<td>1.9</td>
<td>The 4-element realization of the beamforming-MIMO RX.</td>
<td>15</td>
</tr>
<tr>
<td>1.10</td>
<td>4-stage LNA.</td>
<td>17</td>
</tr>
<tr>
<td>1.11</td>
<td>(a) Layout of the on-chip balun, (b) simulated S-parameters and NF of the</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>4-stage LNA.</td>
<td></td>
</tr>
<tr>
<td>1.12</td>
<td>Quadrature Gilbert-based phase shifter.</td>
<td>18</td>
</tr>
<tr>
<td>1.13</td>
<td>Layout of the quadrature all-pass filter (QAF).</td>
<td>18</td>
</tr>
<tr>
<td>1.14</td>
<td>(a) Measured phase shifter’s phase response, (b) measured phase shifter’s RMS</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>phase error across phase states.</td>
<td></td>
</tr>
<tr>
<td>1.15</td>
<td>Measured amplitude error across phase shifter bits.</td>
<td>19</td>
</tr>
<tr>
<td>1.16</td>
<td>5-bit passive $\pi$-stage VGA.</td>
<td>20</td>
</tr>
<tr>
<td>1.17</td>
<td>Layout of one stage of VGA.</td>
<td>21</td>
</tr>
<tr>
<td>1.18</td>
<td>(a) Measured VGA steps, (b) measured VGA’s RMS gain error across attenuation</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>steps.</td>
<td></td>
</tr>
<tr>
<td>1.19</td>
<td>Measured phase error across VGA bits.</td>
<td>22</td>
</tr>
<tr>
<td>1.20</td>
<td>(a) Measured S-parameter of RF-channels (simulated in dashed), (b) simulated</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>IP1dB.</td>
<td></td>
</tr>
<tr>
<td>1.21</td>
<td>(a) Measured conversion gain and NF of RF-to-BB channel, (b) measured I/Q</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>amplitude and phase errors.</td>
<td></td>
</tr>
<tr>
<td>1.22</td>
<td>Coupling path between RF-Channels 1 and 2 of cluster 1 and RF-Channels 2 and</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>3 of cluster 1.</td>
<td></td>
</tr>
<tr>
<td>1.23</td>
<td>(a) coupling between RF-channels 2 and 3 of cluster 1, (b) coupling between</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>RF-channels 1 and 2 of cluster 1 for 4 different phase shifter settings.</td>
<td></td>
</tr>
</tbody>
</table>
1.24 (a) measured RMS gain error, and (b) measured RMS phase error due to coupling of RF-channels 2 and 3 of cluster 1, (c) measured RMS gain error, and (d) measured RMS phase error due to coupling of RF-channels 1 and 2 of cluster 1. .......................................................... 27
1.25 Measured phase scanning of each 3-element cluster. .............................................. 28
1.26 Measured spatially multiplexed array factors of two clusters steered toward 60° and 90°. .......................................................... 28
1.27 Measured array factor of each cluster for 625 VGA settings. .................................. 29
1.28 Measured signal-to-interference ratio (SIR) for different undesired incident angles. .......................................................... 29
1.29 Die micrograph. ........................................................................................................ 30
1.30 (a) RF characterization and RF-to-baseband measurement setup (b) schematic, (c) photo. .......................................................................................... 31

2.2 (a) 2D IQ signal space of 8PSK symbols partitioned by 8 LO phases, and (b) block diagram of multi-phase RF-correlator and sign-check comparators. ............... 42
2.3 (a) Equivalent baseband PAM-4 eye-diagram and 4-level normal Gaussian distribution, and (b) 8PSK bit and symbol error probabilities of the proposed direct-demodulation scheme and theory. ........................................ 45
2.4 Proposed direct-demodulation RF-to-bits 8PSK receiver architecture including front-end, 4-phase LO, mixer-baseband, and demodulator. .......................................................... 46
2.5 (a) 6-Stage common-emitter-based LNA circuit schematic, (b) simultaneous noise and power matching in the first LNA stage ($Z_{opt}$ and $Z_{in}^*$ curves). ............... 47
2.6 (a) Second single-ended LNA stage layout with MIM bypass capacitors, (b) second differential LNA stage layout with CPW-based matching networks, and (c) LNA S-parameter and NF simulation results. .......................................................... 47
2.7 (a) 4-Way splitter layout, (b) splitter S-parameter simulation results, and (c) splitter amplitude and phase error simulation results. .......................................................... 49
2.8 (a) Voltage-mode double-balanced passive mixer followed by 3-stage amplification (last amplification stage is CTLE), (b) simplified circuit of a double-balanced passive mixer, and (c) mixer input matching from 110 GHz to 140 GHz. .......................................................... 50
2.9 (a) Conversion gain of mixer-baseband for four different CTLE code settings, and (b) mixer-baseband NF. .......................................................... 52
2.10 (a) Comparator-DFF with offset calibration, and (b) comparator BER due to random noise (PSS + PNOISE simulations). .......................................................... 53
2.11 Timing diagram of the comparator-DFF operation. .......................................................... 54
2.12 (a) Tripler circuit schematic, and (b) tripler output power with and without input buffer. .................................................................................................................................... 56
2.13 (a) Low-pass phase shifter, (b) high-pass phase shifter, (c) 90° phase-shift tuning range, and (d) S-parameter simulation results. .......................................................... 57
2.14 (a) Doubler circuit schematic, and (b) doubler output power. .......................................................... 58
2.15 (a) Lower-frequency-tuned all-pass phase shifter, and (b) higher-frequency-tuned all-pass phase shifter, (c) 45° phase-shift tuning range, and (d) S-parameter simulation results. .......................................................... 59

vi
2.16 (a) LO network output saturated power for different LO center frequencies, and (b) LO network output power at different harmonics (6th, 4th, 3rd, 2nd, 1st) of the bondwired 20.83 GHz LO input.

2.17 Die micrograph of the 8PSK receiver.

2.18 (a) Wireless measurement setup schematic, and (b) photo.

2.19 Measured receiver conversion gain at the 8PSK test output.

2.20 (a) Measured DSB NF and (b) IP1dB at the 8PSK test output.

2.21 36 Gbps power spectrum at the 8PSK test output.

2.22 Wirelessly measured (a) 30 Gbps and (b) 36 Gbps eye-diagrams at the 8PSK test output.

2.23 Wirelessly measured (a) 30 Gbps and (b) 36 Gbps constellations at the 8PSK test output.

2.24 Wirelessly measured eye-diagrams of demodulated 8PSK 3-bit streams for (a) 30 Gbps and (b) 36 Gbps overall data-rates.

2.25 Variation of BER with received input power.
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Performance Summary of the 4-Element Beamforming-MIMO RX</td>
<td>32</td>
</tr>
<tr>
<td>1.2</td>
<td>Table of Comparison</td>
<td>32</td>
</tr>
<tr>
<td>2.1</td>
<td>Normalized correlation values of 8PSK symbols with eight LO phases</td>
<td>42</td>
</tr>
<tr>
<td>2.2</td>
<td>Logic table of three demodulated bits per symbol</td>
<td>42</td>
</tr>
<tr>
<td>2.3</td>
<td>Comparison table of state-of-the-art direct-demodulation receivers</td>
<td>65</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I would like to express my gratitude to my adviser for his continuous guidance and motivation. Special thanks to Huan Wang for his contribution. I would like to acknowledge STMicroelectronics, GLOBALFOUNDRIES for facilitating the chip fabrication. Finally, I would like to thank Keysight Technologies, in particular, Dave Huh for providing great assistance in test equipment.
CURRICULUM VITAE

Seyed Mohammad Hossein Mohammadnezhad

EDUCATION

Doctor of Philosophy in Electrical Engineering 2019
University of California, Irvine

Master of Science in Electrical Engineering 2018
University of California, Irvine

Bachelor of Science in Electrical Engineering 2013
Sharif University of Technology

Tehran, Iran
REFEREED JOURNAL PUBLICATIONS

A 115-135 GHz 8PSK Receiver Using Multi-Phase RF-Correlation-Based Direct-Demodulation Method
IEEE Journal of Solid State Circuits

Analysis and Design of High-Order QAM Direct Modulation Transmitter for High-Speed Point-to-Point mm-Wave Wireless Links
IEEE Journal of Solid State Circuits

A Millimeter-Wave Partially Overlapped Beamforming-MIMO Receiver: Theory, Design, and Implementation
IEEE Transactions on Microwave Theory and Techniques

A silicon-based low-power broadband transimpedance amplifier
IEEE Transactions on Circuits and Systems I: Regular Papers

Analysis and design of a wideband, balun-based, differential power splitter at mm-Wave
IEEE Transactions on Circuits and Systems II: Express Briefs

REFEREED CONFERENCE PUBLICATIONS

A Single-Channel RF-to-Bits 36Gbps 8PSK RX with Direct Demodulation in RF Domain
Custom Integrated Circuits Conference (CICC), 2019 IEEE

A 100-120GHz 20Gbps Bits-to-RF 16QAM Transmitter Using 1-bit Digital-to-Analog Interface
Custom Integrated Circuits Conference (CICC), 2019 IEEE

A 64–67GHz partially-overlapped phase-amplitude-controlled 4-element beamforming-MIMO receiver
Custom Integrated Circuits Conference (CICC), 2018 IEEE

A low-power BiCMOS 50 Gbps Gm-boosted dual-feedback transimpedance amplifier
2015 IEEE Bipolar/BiCMOS Circuits and Technology Meeting-BCTM
A broadband nonlinear lumped model for silicon IM-PATT diodes
2015 IEEE Bipolar/BiCMOS Circuits and Technology Meeting-BCTM

PATENTS

Ultra-broadband Transimpedance Amplifiers (TIA) for Optical Fiber Communications
U.S. Patent No. 20180102749A1

RF to Bits Receiver without Analog to Digital Converter (ADC) for 8-PSK Constellation
U.S. Application No. 62/806,456

Quadrature Phase Shift Keying Quadrature Amplitude Modulation Transmitter
U.S. Application No. 62/870,189

Transmitter Architecture for Generating $4^N$-QAM Constellation with no Digital-to-Analog Converters (DAC) in Signal Path Requirement
U.S. Application No. 62/712,062
There is an ever-increasing demand for higher data-rates in wireless communication networks. Data traffic in cellular networks increases by roughly 40% annually and by 7 times every 6 years. Furthermore, the number of connected devices in these networks increases by an order of magnitude every decade [1]. To accommodate for this insatiable demand, multi-antenna architectures are adopted to take advantage of the vast available spectrum at mm-wave without suffering from the higher wireless link loss at those frequencies. This thesis presents novel architectures for ultra-high speed wireless transceivers for both short-range to medium-range indoor and outdoor applications to be used in the future generations of wireless communication networks.

In the first section of this thesis, mm-wave circuit- and system-level solutions for addition of multi-user service to conventional multi-antenna phased array architectures will be introduced. The proposed architecture will enhance the link capacity, co-channel user service and hardware cost compared to conventional solutions. Theory and design of the circuits and system are detailed and comprehensive measurement results are presented verifying the system-level functionality. First section is named A Millimeter-Wave Partially-Overlapped Beamforming-MIMO Receiver: Theory, Design, and Implementation. More specifically, this section presents an analysis and design of a partially-overlapped beamforming-MIMO archi-
tecture capable of achieving higher beamforming and spatial multiplexing gains with lower number of elements compared to conventional architectures. As a proof of concept, a 4-element beamforming-MIMO receiver (RX) covering 64-67 GHz frequency band\(^1\) enabling 2-stream concurrent reception is designed and measured. By partitioning the RX elements into two clusters and partially overlapping these clusters to create two 3-element beamformers, both phased-array (coherent beamforming) as well as MIMO (spatial multiplexing) features are simultaneously acquired. 6-bit phase shifters with 360° phase control and 5-bit VGAs with 11 dB range are designed to enable steering of the two RX clusters toward two arbitrary angular locations corresponding to two users. Fabricated in a 130-nm SiGe BiCMOS process, the RX achieves a 30.15 dB maximum direct conversion gain and a 9.8 dB minimum noise figure (NF) across 548 MHz IF bandwidth. S-parameter-based array factor measurements verify spatial filtering of the interference and spatial multiplexing in this RX chip.

In the second section of this thesis, energy-efficient ultra-high speed transceiver architectures will be presented. Current high-speed transceivers rely on high-sampling-rate high-resolution power-hungry analog-to-digital converters or digital-to-analog converters at the interface of analog and digital circuitries. However, design of these backend data-converters are extremely power-hungry at very high speeds in a fully-integrated end-to-end scenario (i.e. RF-to-Bits, Bits-to-RF). Novel system-level architectures will be presented that obviate the need for such costly data converters and will significantly relax the complexity of digital signal-processing. The proposed architecture will result in orders of magnitude energy saving at ultra-high speeds. Theory, design, and measurement results of the highest-speed, highly energy-efficient fully-integrated end-to-end transceiver will be discussed in this section. Second section is named *A Millimeter-Wave Energy-Efficient Direct-Demodulation Receiver: Theory, Design, and Implementation*. More precisely, this section presents the theory, design, and implementation of an 8PSK direct-demodulation receiver based on a

\(^1\)The FCC’s newly allocated 64-71 GHz frequency band for high-speed wireless links between small cells
novel multi-phase RF-correlation concept. The output of this RF-to-bits receiver architecture is demodulated bits, obviating the need for power-hungry high-speed-resolution data converters. A single-channel 115-135-GHz receiver prototype was fabricated in a 55-nm SiGe BiCMOS process. A max conversion gain of 32 dB and a min noise figure (NF) of 10.3 dB was measured. A data-rate of 36 Gbps was wirelessly measured at 30 cm distance with the received 8PSK signal being directly demodulated on-chip at a bit-error-rate (BER) of 1e-6. The measured receiver sensitivity at this BER is -41.28 dBm. The prototype occupies $2.5 \times 3.5 \text{ mm}^2$ of die area including PADs and test circuits ($2.5 \text{ mm}^2$ active area) and consumes a total DC power of 200.25 mW.
Chapter 1


1.1 Introduction to mm-Wave Wireless Communication Networks

There is an insatiable demand for higher data rates and larger number of data consumers in mobile communication networks. Mobile data traffic is expected to increase sevenfold from 2016 reaching 48.3 Exabytes per month by 2021 with global IP video traffic making up 82% of the internet traffic by 2021 compared with 73% in 2016 [1].

To accommodate this ever-increasing demand of high data rates, three directions are pursued
in the deployment of the 5th generation wireless systems, 5G, which result in orders-of-magnitude increase in wireless capacity compared to current wireless networks: (1) \textit{mm-wave wireless communication} with technologies currently available at 26, 28, 38, 60 GHz leveraging available wide bandwidths to achieve multi-gigabit-per-second data rates (e.g., 5 Gbps by 2020 and 50 Gbps by 2024 [2]) compared to data rates achieved in 4G LTE-advanced (1 Gbps) or 3G networks (384 kbps) [3], (2) \textit{frequency reuse} leading to creation of small cells (i.e., pico or femtocells as shown in Fig. 1.1) with 10–200 m of coverage range, where intercell interference is insignificant due to the high path loss experienced by mm-wave communication, and (3) \textit{distributed base stations} (BS) with massive number of antennas (>100) providing one-tier or multi-tier high-speed wireless access to multi co-channel users [4,5].

Multi-antenna architectures are currently being adopted for both base stations and small cells to combat high path loss at mm-wave frequencies and achieve higher capacity and co-channel user service [6–9]. Increasing the number of antennas results in channel hardening and reduction of small-scale fading (less multi-path and Doppler spread), which in return simplifies baseband signal processing algorithms. Various configurations of multi-antenna architectures provide: (1) multiplexing gain to enhance link capacity through concurrent transmission of parallel data/user streams, (2) diversity gain to improve reliability of wireless

![Figure 1.1: Generic 5G network architecture (high-level topological view).](image-url)
links especially in non line-of-sight (NLOS) scenarios through transmission of copies of the same data stream, (3) antenna gain to combat path loss, integrated wide-band noise, and co-channel interference through beamforming in LOS or directed NLOS scenarios.

In low-SNR mm-wave channels, increasing the capacity is limited by both the transmitter output power and receiver integrated noise. Adopting beamforming multi-antenna architectures to transmit sharp beams with highly directional antenna gains enhances the capacity by improving the SNR, thus making it possible to employ high-order modulation schemes to achieve higher spectral efficiencies. On the other hand, in high-SNR mm-wave channels with high diversity or rank order, exploiting multiplexing gain via propagation of independent signal streams through multiple distinct paths in spatial and polarization domains can further enhance the channel capacity or multi-user service.

1.2 All-analog, All-digital, and Hybrid Architectures

An all-analog multi-antenna architecture, where beamforming weights are applied in analog domain, significantly reduces complexity and cost of baseband digital signal processing and provides wider coverage for wireless links by generating sharp transmit/receive beams. However, the mm-wave front-end of an all-analog system for applications mandating high beamforming resolution, adaptability, and multi-beam communication should satisfy stringent performance/power specifications.

On the other hand, an all-digital beamforming system enables multi-beam communication with the highest adaptability and data-rate, but demands complex power-hungry, high-speed baseband DSPs for large antenna arrays operating at mm-wave frequencies. Therefore, a hybrid architecture with both analog beamforming and digital MIMO coding is desired to reduce the complexity of the digital baseband in communication systems with massive
1.3 Conventional Hybrid and the Proposed Partially-Overlapped Hybrid Architectures

Hybrid architectures can be designed to receive (or transmit) only one dedicated data-stream per each subset of antennas (sub-array in Fig. 1.2a), or receive (or transmit) all data-streams from all antennas (full-array in Fig. 1.2b). The number of required RF chains $N_{RF}$ in a hybrid architecture is strictly lower-limited by the number of parallel data streams $K$, while beamforming gain is determined by the number of antenna elements per each cluster $M$, where each cluster is composed of dedicated RF-channels (i.e., complex weighting coefficients in Figs. 1.2a, 1.2b, and 1.3) and antennas per RF chain. A full-array, in fact, realizes the function of an all-digital architecture. The number of signal processing paths (from the digital baseband to the antenna front-end) for the sub-array in Fig. 1.2a is equal to $N$ and for full-array in Fig. 1.2b is equal to $N_{RF} \times N$, where $N$ is the total number of antennas [10]. On the other hand, beamforming gain of the sub-array is $1/N_{RF}$ of the full-array. Therefore, a trade-off exists between signal processing complexity and beamforming gain of hybrid architectures. A recent circuit implementation of a hybrid architecture was
Figure 1.3: Conceptual representation of the hybrid system with partially-overlapped clusters.

Presented in [11]. It utilizes Cartesian combining concept to enable 2-stream reception. However, this hybrid RX requires 8 splitters, 20 combiners, and 12 mixers for a 2-stream reception. These large number of signal paths introduce electromagnetic cross-talks due to many cross-overs between these paths. Therefore, this architecture is not scalable for larger number of streams and not suitable for implementation at higher frequencies.

To address the above issues, a partially-overlapped beamforming-MIMO architecture is introduced [12]. Shown in Fig. 1.3, \( N \) antennas are decomposed into \( K \) partially-overlapped clusters of \( M \) antennas with an overlapping depth of \( D \) supporting \( K \) parallel data streams. \( M \) is bounded by:

\[
\frac{N}{N_{RF}} \leq M \left[ = (N + D(N_{RF} - 1))/N_{RF} \right] \leq N
\]  

(1.1)
Overlapping the clusters reduces the complexity of RF-to-basedband signal processing paths to \( N_{RF} \times M \) from \( N_{RF} \times N \) in the case of the full-array. From (1.1), varying \( D \) from 0 to \( N \) causes the number of signal processing paths to vary from \( N \) to \( N_{RF} \times N \) and the beamforming gain improvement factor (referenced to a sub-array architecture) to vary from 1 to \( N_{RF} \). Therefore, introducing overlapping depth into the clusters of antennas and RF-channels creates a new degree of freedom that helps us reach a better compromise between the beamforming gain and the complexity of signal processing in hybrid architectures. Fig. 1.4a shows \( M \) and \( N_{RF} \) variations with respect to \( D \) in a partially overlapped hybrid for \( N = 8 \) elements, and Fig. 1.4b shows the possible combinations of \((M, N_{RF}, D)\) for \( N = 4 \) and \( N = 8 \). The behavior of multiplexing and coherent processing gain variations with respect to the overlapping depth are demonstrated in these figures.

In Fig. 1.3, the \( i \)-th complex weighting coefficient (RF-channel) in cluster \( k \) is generally defined as \( W_{i,k} = A_{ik}e^{j\phi_{ik}} \) for \( i \in \{1, ..., M\} \) and \( k \in \{1, ..., N_{RF}\} \). \( \phi_{ik} \) and \( A_{ik} \) are realized by RF phase shifters and variable gain attenuators and/or amplifiers (VGAs). As will be illustrated in Sections 1.4 and 1.5, the RF phase shifters are used for mainlobe steering of each cluster, whereas the RF VGAs enable spatial filtering of interferences from other

![Graph](image-url)

Figure 1.4: (a) \( M \) and \( N_{RF} \) as function of \( D \) for \( N = 8 \) elements, (b) \((M, N_{RF}, D)\) combinations for \( N = 4 \) and \( N = 8 \) elements.
clusters by placing the null locations of each beamforming cluster toward the directions of the interference incident angles. Also, throughout the forthcoming analysis, we assume $N_{RF} = K$ without loss of generality, as this work is primarily concerned with the RF portion of the hybrid architecture. It is noteworthy that the use of both linear amplitude and phase controls across the frequency range of interest enables this partially-overlapped hybrid architecture to achieve the same performance as a linear digital beamforming system [13,14]\(^\text{1}\).

The upper bound of channel capacity is determined by Shannon-Hartley theorem and is a function of the $SNR = P_{\text{sig}}/P_N$ where $P_{\text{sig}}$ is signal power and $P_N$ is noise power. To compare performance of these hybrid architectures, the outage capacity with respect to $SNR$ is calculated. The outage happens when the Shannon capacity falls below a certain threshold, $C_{\text{out}}$. The outage probability, $P_{\text{out}}$, considering equal power distribution across all clusters is given by [16]:

$$P_{\text{out}} = \frac{(2^{C_{\text{out}}} - 1)^K}{K! \cdot SNR_H^K}$$

\(^1\)To achieve the same performance in a hybrid with phase-only constraint, the number of parallel data streams is required to be less than half the number of the RF chains [15].
where $SNR_H$ is defined as the improvement in SNR and is equal to:

$$SNR_H = \frac{M}{K}SNR$$  \hspace{1cm} (1.3)

With the beamforming gain defined as the number of antennas per cluster, overlapping the clusters in an $N$-element hybrid allows us to allocate larger number of antennas per cluster, thereby resulting in higher beamforming gain and $SNR_H$ compared to the corresponding $N$-element sub-array. As an example, in Fig. 1.5, the outage capacity of full-array, sub-array, and partially-overlapped antenna arrays with $N = 4$ and $N = 8$ for an outage probability of 1% are compared. The partially-overlapped hybrid architecture is observed to achieve better outage capacity than the conventional sub-array counterpart.

### 1.4 Phase Control and Mainlobe Steering in Partially-Overlapped Hybrid Architecture

Phase shifters in the RF-channels of each cluster enable independent phase excitation for each RF-channel within that cluster. The array factor $AF$ of cluster $k$ is expressed as:

$$AF_{Mk} = \sum_{i=1}^{M} e^{(i-1)\times j\psi_k}$$  \hspace{1cm} (1.4)

where $\psi_k = \pi \cos \theta_k + \phi_k$. $\phi_k$ is the phase progression from one RF-channel to the next in cluster $k$, and $\theta_k$ is the incident angle from the axis of antenna array. By adjusting the phase shifters in each cluster so as to achieve relative phase difference of $\phi_k$, the array factor of cluster $k$ can be maximized toward the incident angle $\theta_k$, or in other words, cluster $k$’s mainlobe can be steered toward the incident angle $\theta_k$. As a result, the mainlobes of all clusters in this partially-overlapped hybrid architecture can be simultaneously and independently
steered toward different arbitrary angles.

1.5 Amplitude Control and Null-Steering in Partially-Overlapped Hybrid Architecture

By utilizing both VGA’s amplitude and phase shifter’s phase controls in this architecture, null locations of the array factor can be arbitrarily steered. Null-steering solely based on the VGA setting is of particular interest to enable interference suppression independently from mainlobe steering (achieved by phase shifter settings). Nevertheless, for the sake of completeness, null-steering using two mechanisms, namely, (1) phase shifter settings and (2) VGA settings with phase shifter being already preset, will be explained.

1.5.1 Null-steering with Phase Shifters

Considering an M-antenna cluster, the array factor of cluster $k$ with $\lambda/2$ spacing between antennas is written as:

$$ AF_{Mk} = \sum_{i=1}^{M} e^{(i-1)\times j\psi_k} \sum_{z=0}^{\psi_k} \sum_{i=1}^{M} z^{i-1} $$

(1.5)

where $\psi_k = \pi \cos \theta_k + \phi_k$ and $\phi_k$ is the phase progression (set by the phase shifters) from one RF-channel to the next. The power series in (1.5) is readily calculated, resulting in $AF_k = (1 - z^M)/(1 - z)$ [17]. As shown in Fig. 1.6a, this uniform distribution of VGA amplitudes and progressive phase shifter settings will result in $M - 1$ zeros (nulls) around the unity circle of the cluster’s array factor. Assuming $\psi_k$ to be the $k$-th zero on the unity circle of Fig. 1.6a, after presetting $\phi_k$ to steer the cluster’s mainlobe toward the desired angle, the null location of cluster $k$ is calculated to be at $\theta_{null,k} = \cos^{-1}(\psi_k - \phi_k)/\pi$. 

9
1.5.2 Null-steering with VGAs

Relying solely on phase shifters to achieve both mainlobe and null-steering does not provide sufficient flexibility for null control. Therefore, it is desired to delegate null-steering to VGAs and design phase shifters only for mainlobe steering purpose. The array factor of cluster $k$ with arbitrary amplitude excitation per RF-channel is given by:

$$AF_{Mk} = \sum_{i=1}^{M} A_{ik} e^{j(i-1)\times\psi_k} = \sum_{i=1}^{M} A_{ik} z^{i-1}$$ (1.6)

By allowing only conjugate pair zeros in the AF, the need for a phase shifter to realize nulls is alleviated (i.e., $(z - z_i)(z - z^*_i)$ is real, thus no phase information) at the cost of reducing the number of possible nulls to half (Fig. 1.6b). The AF of cluster $k$ with conjugate pair zeroes is [17]:

$$AF_{Mk} = \begin{cases} 
(z + 1)\Pi_{i=2}^{n}(z - z_{i-1})(z - z^*_{i-1}); & M = 2n \\
\Pi_{i=1}^{n}(z - z_i)(z - z^*_i); & M = 2n + 1 
\end{cases}$$ (1.7)

therefore, $AF_{M=2n,k} = (z + 1)AF_{M=2n-1,k}$. Moreover, the VGAs amplitude settings within a

![Image](image_url)

Figure 1.6: (a) Array factor nulls based on phase shifter control, (b) Array factor nulls based on VGA control.
cluster will be symmetric around the center element of each cluster.

In the special case of two 3-element clusters, null location can be steered arbitrarily to suppress the interference from the other cluster. The AF of cluster 1 is:

\[
AF_{31} = (z - z_1)(z - z_1^*) \\
= (z - e^{j\psi_1})(z - e^{-j\psi_1}) \\
= z^2 - 2z\cos\psi_1 + 1
\]  

(1.8)

where \( A_{21} = -2\cos\psi_1 = -2\cos(\pi \cos\theta_{\text{null},1} + \phi_1) \). \( \theta_{\text{null},1} \) is the null location of cluster 1 and \( \phi_1 \) is the required phase progression set by this cluster’s phase shifters to steer the mainlobe of cluster 1 toward the desired direction of \( \theta_{\text{main},1} \). Similarly, the AF of cluster 2 is calculated to be:

\[
AF_{32} = z^2 - 2z\cos\psi_2 + 1
\]  

(1.9)

where \( A_{22} = -2\cos\psi_2 = -2\cos(\pi \cos\theta_{\text{null},2} + \phi_2) \). \( \theta_{\text{null},2} \) represents the null location of cluster 2 and \( \phi_2 \) is set by this cluster’s phase shifters to steer the mainlobe of cluster 2 toward the desired direction of \( \theta_{\text{main},2} \). Therefore, proper adjustments of these two clusters’ phase shifters and VGAs yield \( \theta_{\text{null},1} = \theta_{\text{main},2} \) and \( \theta_{\text{null},2} = \theta_{\text{main},1} \). It also facilitates simultaneous operation of these two clusters for spatial multiplexing in addition to beamforming within each cluster.
1.6 Effects of Amplitude and Phase Errors on Interference Suppression

Non-idealities in RF phase shifters and VGAs (e.g., resolution-induced quantization error, timing jitter, and device mismatch) result in amplitude and phase errors, which degrade the effectiveness of interference suppression in the overlapped hybrid architecture. Suppose that optimum phase and amplitude excitations for each RF-channel within cluster \( k \) were derived to achieve maximum array factor in the desired direction in the presence of interference induced by simultaneous operation with other clusters. The complex weight vector containing these optimum excitations is expressed as:

\[
\vec{W}_{opt,k} = [A_{1k}e^{j\phi_{1k}}, ..., A_{Mk}e^{j\phi_{Mk}}]; k \in \{1, ..., K\} \tag{1.10}
\]

where \( A_{ik} \) and \( \phi_{ik} \) are the optimum amplitude and phase settings of the VGA and phase shifter of the RF-channel \( i \) in cluster \( k \). Accounting for amplitude and phase errors in the VGA and phase shifter, the actual weight vector will be:

\[
\vec{W}_{err,k} = [A_{1k}(1 + a_{1k})e^{j(\phi_{1k} + \delta_{1k})}, ..., A_{Mk}(1 + a_{Mk})e^{j(\phi_{Mk} + \delta_{Mk})}]; k \in \{1, ..., K\} \tag{1.11}
\]

where \( a_{ik} \) and \( \delta_{ik} \) denote the amplitude and phase errors of the VGA and phase shifter of the \( i \)-th RF-channel in cluster \( k \). Fig. 1.7 shows the beamforming weight vectors \( \vec{W}_{opt,k} \) and \( \vec{W}_{err,k} \) with an angular error of \( \theta_{err,k} \) between them. \( \vec{W}_{opt,k} \) is orthogonal to interference plane, resulting in maximum beamforming gain and interference suppression. Because of the amplitude and phase errors, \( \vec{W}_{err,k} \) is projected on both signal and interference planes by factors \( \cos(\theta_{err,k}) \) and \( \sin(\theta_{err,k}) \) [18]. As a result, for small angular errors, interference leakage (a function of \( \sin(\theta_{err,k}) \approx \theta_{err,k} \)) is more sensitive to phase/amplitude errors than the mainlobe gain (a function of \( \cos(\theta_{err,k}) \approx 1 \)). The standard deviations of \( a_{ik} \) and \( \delta_{ik} \)
for $1 \leq i \leq M$ are defined as $E[a_{ik}^2] = \sigma_{a,k}^2$ and $E[\delta_{ik}^2] = \sigma_{\delta,k}^2$ ($E[.]$ is the expected value). Assuming normalized weight vectors (i.e., $\sum_{i=1}^{M} A_{ik}^2 = 1$), the variance of error in the weight vector of cluster $k$ is calculated to be [19]:

$$
\sigma_{\theta,k}^2 = E[(\bar{W}_{opt,k} - \bar{W}_{err,k})^2] \\
= E[\sum_{i=1}^{M} A_{ik}^2 (a_{ik}^2 + \delta_{ik}^2)] \\
= \sigma_{a,k}^2 + \sigma_{\delta,k}^2
$$

(1.12)

therefore, the standard deviation of angular error in the weight vector is independent of the number of antennas per cluster. However, increasing the number of antennas facilitates suppressing larger number of interferences, thus supporting larger number of parallel data streams and improving the mainlobe gain by increasing the beamforming gain per cluster.

Fig. 1.8 shows interference rejection (i.e., $-10\log(\sigma_{\theta,k}^2)$ for normalized weight vectors assuming small angular errors $\theta_{err,k}$) as a function of RMS amplitude and phase errors. To combat interference suppression degradation due to amplitude and phase errors, high resolution VGA and phase shifters need to be employed. With $B_{amp}$-bit VGA and $B_{\phi}$-bit phase shifter per RF-channel, $2^{B_{amp}+B_{\phi}}$ weight vectors can be generated. With scalar quantization of the weight vectors, the expected value of peak-to-null-ratio $E(PNR)$ varies proportionally
with:

\[ E[PNR] \propto M \times 2^{(B_{amp} + B_\phi)} \]  

(1.13)

therefore, \( E[PNR] \) is linearly proportional to the number of elements per cluster. However, it varies exponentially with the resolution of the VGA and phase shifter, \( B_{amp} \) and \( B_\phi \), showing the importance of the VGA and phase shifter resolutions in achieving high interference suppression and high resolution beam control.

### 1.7 The Phase-Amplitude Controlled Partially-Overlapped Hybrid Circuit Architecture

Based on the idea of partially-overlapped hybrid scheme with phase/amplitude control, a 4-element beamforming-MIMO RX for \( D = 2, M = 3, N = 4, \) and \( K = 2 \) is designed (Fig. 1.9) \[?\,\,12,\,20\]. Each of the two clusters is composed of three RF-channels (i.e., three LNA-PS-VGA paths) with LNA2 and LNA3 shared between the two clusters. The phase
states of three RF phase shifters and amplitude states of three VGAs within each cluster are adjusted based on the mainlobe steering and the null-steering methods explained in Sections IV and V, respectively. The null location for each cluster is steered with VGA steps, enabling each cluster to suppress interferences at undesired incident angles located in its null space. Inter-element spacing within and between clusters in this architecture is assumed equal to λ/2 so as to avoid spatial aliasing in diversity beampatterns. Beamforming gain acquired from each cluster combined with spatial multiplexing gain from these 2 clusters improve reliability/diversity of high data-rate multi-stream links at both low and high SNR regimes. Two LNAs (LNA2-3), shared between \( U_1 \) and \( U_2 \), are followed by splitters to allow independent amplitude/phase control for these 2 clusters. The signals \( U_{1k}e^{j\theta_{1k}} \) and \( U_{2k}e^{j\theta_{2k}}, 1 \leq k \leq 4 \), from distinct incident angles \( \theta_1 \) and \( \theta_2 \) appearing on 6 RF-channels are fed to 6-bit phase shifters \( (\phi_{ij} \mid 1 \leq i \leq 2, 1 \leq j \leq 3) \) to be steered independently toward a desired angle. The phase-shifted signals are then fed to 5-bit VGAs \( (A_{ij} \mid 1 \leq i \leq 2, 1 \leq j \leq 3) \) to suppress interference due to concurrent data reception of the other cluster by the null-space of \( U_2 \) and the null-space of \( U_1 \).

Figure 1.9: The 4-element realization of the beamforming-MIMO RX.

15
steering technique explained in Section 1.5. Furthermore, any static amplitude or phase mismatch between RX’s RF-channels is compensated with high dynamic range of VGAs and phase reference control of phase shifters. The VGAs’ outputs within each cluster travel through carefully laid-out equi-length paths to a 3-to-1 combiner, where they are coherently power-combined at the desired incident angle ($U_{1(2)}e^{j\theta_{1(2)}}$ for cluster 1(2)) and suppressed at the undesired ($U_{2(1)}e^{j\theta_{2(1)}}$ for cluster 1(2)). IQ mixers down-convert the RF signals at each combining node using a shared integrated LO network to be filtered and amplified in the baseband. To characterize RF-channels, the RF signal at each cluster’s combining node is monitored at the output of a tap coupler.

1.7.1 Low Noise Amplifier (LNA)

Shown in Fig. 1.10 is the schematic of the 4-stage LNA employed in each RF-channel. The first stage is an inductively degenerated common emitter amplifier. Common emitter shows a smaller NF compared to cascode as the frequency of operation approaches $f_T$. Current density of the first stage is set to $J_{DC} \approx 0.52$ mA/$\mu$m to achieve close-to-minimum NF without compromising gain too much. The inductor $L_{deg}$ was then optimized to achieve simultaneous noise and power match. The LNA’s second stage incorporates a cascode at the same current density to achieve a higher gain. An on-chip tuned balun was designed to convert the cascode’s single-ended output (matched to 50 $\Omega$) to differential (matched to 100 $\Omega$) with minimum phase error and loss (Fig. 1.11a) [21]. Finally, two neutralized differential stages were designed for $G_{max}$ to further amplify the incoming signal. The S-parameter and NF simulations of the 4-stage LNA is depicted in Fig. 1.11b. The 4-stage LNA consumes a total DC power of 38.6 mW.
1.7.2 Phase Shifter (PS)

As shown in [22], the beam-steering resolution with a $B_\phi$-bit phase shifter is:

$$\theta_{\text{main, res}} = \sin^{-1}\left(\frac{1}{2^{B_\phi}-1}\right)$$  \hspace{1cm} (1.14)

An array utilizing 6-bit and 5-bit phase shifters results in an SNR improvement very close to the ideal SNR improvement across incident angles [22]. A 6-bit (5-bit) phase shifter results in a beam-steering resolution of 1.79° (3.58°).

Figure 1.11: (a) Layout of the on-chip balun, (b) simulated S-parameters and NF of the 4-stage LNA.
A 6-bit phase shifter (1 calibration bit to reduce RMS phase error), capable of generating 0° to 360° with 11.25° phase steps, was designed to adjust the phase of each RF-channel (Fig. 1.12). It employs an active Gilbert-based topology with a quadrature all-pass filter (QAF) whose 3D layout view is shown in Fig. 1.13 [23]. The QAF is composed of low-pass and high-pass filters to generate the differential quadrature signals, IM-IP and QM-QP, at the resonance frequency. The QAF inductor and capacitor are $L = R/\omega_0$ and $C = L/R^2$, where $R$ is the QAF characteristic impedance. The input parasitic capacitance of the Gilbert stage
Figure 1.14: (a) Measured phase shifter’s phase response, (b) measured phase shifter’s RMS phase error across phase states.

contributes to amplitude and phase errors of the QAF. The series resistor, $R_s$, lowers the network’s Q-factor as well as these amplitude and phase errors. The optimum value of $R_s$ for theoretical zero amplitude and phase errors is calculated to be $R_s = R$ [24, 25]. The tail current of each quadrature Gilbert cell is controlled by a 4-bit DAC. Moreover, two additional bits control the current steering of switch pairs M$_3$-M$_4$ and M$_5$-M$_6$ to generate positive and negative phases at the output of phase shifter. As indicated in Fig. 1.14a, the measured RF-channel phases exhibit constant group delay for different settings of phase shifter. An RMS phase error less than 2.8° was measured across the RF BW (Fig. 1.14b). The RMS amplitude error across phase shifter bits is less than 0.85 dB (Fig. 1.15). The core

Figure 1.15: Measured amplitude error across phase shifter bits.
phase shifter consumes a total DC power of 27.2 mW.

### 1.7.3 Variable Gain Attenuator (VGA)

To adjust the signal amplitude of each RF-channel, a 5-bit passive VGA was designed (Figs. 1.16 and 1.17). The VGA attenuation varies from 0.5 dB to 11.5 dB in 1-dB steps. Using Eqs. (8) and (9), an amplitude step of $A$ ($A \leq 1$) will result in a null-steering resolution of:

$$\theta_{\text{null, res}} = \cos^{-1}\left(\frac{\cos^{-1}(-A/2)}{\pi}\right) - \cos^{-1}\left(\frac{2}{3}\right)$$

for 3 elements per cluster. The null-steering resolution, $\theta_{\text{null, res}}$, is $1.78^\circ$ for 1-dB amplitude step, and is $3.97^\circ$ for 2-dB amplitude step. Therefore, the use of 1-dB-step VGA and 11.25$^\circ$-step phase shifter enables comparable resolution for the null and mainlobe steering (calculated in Section VII.B).

The attenuation stages employ $\pi$-stage topology whose shunt and series resistors, $R_P$ and

![Figure 1.16: 5-bit passive $\pi$-stage VGA.](image-url)
Figure 1.17: Layout of one stage of VGA.

Figure 1.18: (a) Measured VGA steps, (b) measured VGA’s RMS gain error across attenuation steps.

$R_S$, for different levels of attenuation $A_{dB}$ ($A_L = \sqrt{1/(10^{A_{dB}/10})}$) are derived as:

$$R_P = \frac{Z_0(1 - A_L^2)}{1 + A_L^2 - 2A_L} = \frac{Z_0}{1 - A_L}$$ (1.16)

$$R_S = \frac{(1 - A_L)R_P Z_0}{R_P - Z_0}$$ (1.17)
Figure 1.19: Measured phase error across VGA bits.

where $Z_0$ is the 50-$\Omega$ characteristic impedance of the VGA (100-$\Omega$ differential). A large shunt resistor, $R_{P,off}$, is added in series with $R_P$ and in parallel with deep N-well switches. In the VGA ON-mode, the switches turn on, $R_{P,off}$ is shorted and the resistive network $R_S-R_P$ adjusts the attenuation level. In the VGA OFF-mode, $R_{P,off}$ is placed in series with $R_P$, resulting in no attenuation from the VGA stage. The error in $A_L$ due to variations in $R_P$ and $R_S$ ($\Delta R_P$ and $\Delta R_S$) is approximately equal to:

$$A_{err} = Z_0 R_P \frac{[(Z_0 R_S) \Delta R_P - (R_P R_S + Z_0 R_S) \Delta R_S]}{(R_S R_P + R_S Z_0 + R_P Z_0)^2} \quad (1.18)$$

To reduce the effect of parasitic capacitances of the switches, the equivalent capacitance, $C$, at the VGA input and output ports are resonated out by inductors $L$ in series with $R_S$. This capacitance appears in parallel with the input and output VGA ports and is derived to be:

$$C = C_{SW} \frac{(1 + Q_P^2) Q_S^2}{(1 + Q_S^2) Q_P^2} \quad (1.19)$$

where $C_{SW}$ is the equivalent parasitic capacitance appearing in parallel with the switch, and $Q_P$ and $Q_S$ are derived to be:

$$Q_P = R_{P,off} C_{SW} \omega_0 \quad (1.20)$$
A gain stage is introduced after -2 dB stage to reduce the NF contribution of the two -4 dB stages. Assuming a loss of $L_B$ before the gain stage, a loss of $L_A$ after the gain stage, and a gain of $G_{GS}$ from the gain stage, the $IP_{1dB}$ and the noise factor referred to the VGA input are:

$$IP_{1dB,VGA} = 1/(\frac{L_B}{IP_{1dB,GS}} + \frac{G_{VGA}}{IP_{1dB,AV}})$$ (1.22)

$$F_{VGA} = L_B F_{GS} + \frac{L_B}{G_{GS}} (L_A - 1) + \frac{F_{AV} - 1}{G_{VGA}}$$ (1.23)

where $IP_{1dB,AV}$ and $F_{AV}$ are the input compression point and the noise factor of RF building blocks after the VGA (including the power combiner and mixer), respectively, and $G_{VGA} = G_{GS}/(L_B L_A)$. The effect of $IP_{1dB,GS}$ is reduced by $L_B$ and $IP_{1dB,AV}$ is thus the dominant term setting the overall linearity $IP_{1dB,VGA}$ in all-OFF mode, and is independent of the gain stage location. However, $F_{VGA}$ degrades by the absolute loss of the stages prior to the gain stage in both ON and OFF modes. Placing the gain stage after all VGA stages compared to putting it before will degrade $NF_{VGA}$ by 9.72 dB/15.7 dB and will improve $IP_{1dB,VGA}$ by 1.87 dB/10.52 dB in all-OFF/all-ON modes. Putting the gain stage after the -2 dB stage compared to putting it before all VGA stages will degrade $NF_{VGA}$ by only 2.96 dB/2.58 dB and will improve $IP_{1dB,VGA}$ by 0.79 dB/5.18 dB in all-OFF/all-ON modes.

The measured RF-channel gain for different VGA settings is shown in Fig. 1.18a. A measured RMS gain error less than 0.32 dB across the RF BW was obtained (Fig. 1.18b).
measured phase error of the VGA for all 16 states, including the cases when all stages are OFF and ON, is shown in Fig. 1.19. The RMS phase error across VGA bits is less than 1.8°. The VGA block consumes a total DC power of 9 mW.

1.7.4 Measurement Results of RF and RF-to-BB Channels

The RX S-parameters and linearity were measured from the LNA inputs down to RF test PAD1 and 2 after VGA compensation for any gain mismatch between RF-channels. The measured RX S-parameters plot in Fig. 1.20a shows a maximum gain per RF-channel of 12.3 dB at a center frequency of 65.5 GHz with 3 GHz RF BW. The worst-case IP1dB of the RF-channels vs. RF frequency is shown in Fig. 1.20b.

Fig. 1.21a shows the measured conversion gain and NF for an RF-to-BB channel comprising the entire path from the LNA down to the BB port. Upon down-conversion using the integrated LO network, a direct conversion gain of 30.15 dB was measured with a 3-dB IF BW of 548 MHz. The measured double sideband NF of a single RF-to-BB channel for minimum and maximum VGA attenuations varies from 9.8-11 dB and 11.5-12.7 dB,

![Figure 1.20: (a) Measured S-parameter of RF-channels (simulated in dashed), (b) simulated IP1dB.](image-url)
Figure 1.21: (a) Measured conversion gain and NF of RF-to-BB channel, (b) measured I/Q amplitude and phase errors.

respectively, across the IF BW (Fig. 1.21a). The IQ gain/phase mismatches were measured to be less than 0.8 dB/3.6° (Fig. 1.21b).

### 1.7.5 Coupling between RF-Channels

Intra- and inter-element couplings within and between clusters cause gain/phase errors at the two combining nodes of this RX [26]. For example, sources of coupling between three RF-channels of cluster 1 are demonstrated in Fig. 1.22 (this coupling discussion is also valid for cluster 2 due to symmetry of the RX layout). The dominant source of coupling is from the last single-ended stage of each RF-channel’s LNA to the LNA input of an adjacent RF-channel. To measure the effect of coupling on gain/phase errors of this RX, two coupling scenarios are considered. Coupling from desired RF-channel 3 of cluster 1 to RF-channel 2 of cluster 1 and coupling from desired RF-channel 1 of cluster 1 to RF-channel 2 of cluster 1.

In the first scenario, the signal at the last single-ended stage of LNA3 couples to the input of LNA2 and after getting amplified by LNA2 will pass through the undesired RF-channel 2 of cluster 1 and will appear at the combining node of cluster 1 with a gain of $G_{LNA2}G_{\phi_{12}}A_{12}$. 
Figure 1.22: Coupling path between RF-Channels 1 and 2 of cluster 1 and RF-Channels 2 and 3 of cluster 1.

Bypass rails between RF-channels of each cluster in addition to ground plane for critical passives are designed to suppress this mechanism. To measure this coupling, phase shifter $\phi_{13}$ of the desired RF-channel is set as a reference and phase shifter of the undesired path $\phi_{23}$ and $\phi_{12}$ are set to four specific values of $0^\circ$, $90^\circ$, $180^\circ$, and $270^\circ$ (Fig. 1.23). The $0^\circ$ and $180^\circ$ phase differences and the $90^\circ$ and $270^\circ$ phase differences capture maximum gain and phase errors, respectively, at the combining node. Figs. 1.24(a)-(b) show the measured

Figure 1.23: (a) coupling between RF-channels 2 and 3 of cluster 1, (b) coupling between RF-channels 1 and 2 of cluster 1 for 4 different phase shifter settings.
Figure 1.24: (a) measured RMS gain error, and (b) measured RMS phase error due to coupling of RF-channels 2 and 3 of cluster 1, (c) measured RMS gain error, and (d) measured RMS phase error due to coupling of RF-channels 1 and 2 of cluster 1.

RMS phase/gain errors at the combining node for these four different phase settings. The RMS gain and phase errors due to coupling between RF-channels 2 and 3 of cluster 1 at the combining node of cluster 1 is less than 0.56 dB and 1.4°, respectively. In the second scenario, coupling from RF-channel 1 to RF-channel 2 of cluster 1 is considered. The dominant source of coupling in this case is from the last single-ended stage of LNA1 to the input of LNA2, where the signal experiences a gain of $G_{LNA2}G_{\phi_{12}}A_{12}$ and appears at the combining node of cluster 1. Figs. 1.24(c)-(d) show the measured RMS phase/gain errors at the combining
Measured phase scanning of each 3-element cluster.

Figure 1.25: Measured phase scanning of each 3-element cluster.

Measured spatially multiplexed array factors of two clusters steered toward 60° and 90° node for four different phase settings. The RMS gain error and phase error due to coupling between RF-channel 1 and 2 of cluster 1 at the combining node of cluster 1 is less than 0.54 dB and 2.27°, respectively. Therefore, amplitude and phase errors due to coupling between RF-channels of each cluster are negligible in this RX.

Figure 1.26: Measured spatially multiplexed array factors of two clusters steered toward 60° and 90°.

1.8 Measured Null-steering and Spatial Multiplexing

Based on the phase and amplitude control techniques formulated by (1.8) and (1.9), the mainlobe and null locations of each cluster’s beam pattern can be directed toward arbitrary
Figure 1.27: Measured array factor of each cluster for 625 VGA settings.

angles. Array factor of each cluster is extracted from S-parameter measurements of each RF-channel of this RX for all phase shifter and VGA settings in accordance with Eq. (1.6).

Fig. 1.25 shows the rotational change in the location of array factor’s mainlobe for different phase shifter settings within a cluster. The half-power beamwidth increases at wide null-steering intervals due to near triggering of grating lobes in the cluster’s array factor. Fig. 1.26 shows the measured spatially multiplexed array factors of cluster 1 and cluster 2 steered toward $60^\circ$ and $90^\circ$ with the null of each cluster adjusted such that it is placed on top of the mainlobe of the other cluster.

Figure 1.28: Measured signal-to-interference ratio (SIR) for different undesired incident angles.
Fig. 1.27 shows the array factor of each cluster for 625 combinations of VGA settings, where the null location of the array factor varies in accordance with Eqs. (1.8) and (1.9) for different VGA setting. For every null location, $\theta_{\text{null}}$, multiple combinations of VGA settings $(A_{1k}, A_{2k}, A_{3k})$ for cluster $k$ can be found. Mainlobe gain degradation due to null-steering is equal to:

$$20\log\left(3 \times \frac{\max(A_{1k}, A_{2k}, A_{3k})}{2 - 2\cos(\pi\cos(\theta_{\text{null}}))}\right)$$

for a 3-element array. For a uniformly illuminated 3-element array $\theta_{\text{null}} = \cos^{-1}(2/3)$ and there is no degradation.

By having control over both the mainlobe and null locations, spatial multiplexing of multiple (in this work, two) clusters can be realized. For spatially multiplexed clusters, interference is considered to be the interference of clusters on each other during simultaneous parallel data stream reception. Assuming equal signal powers for cluster 1 and cluster 2 (signal-to-
interference-ratio (SIR) = PNR), Fig. 1.28 shows the measured signal-to-interference ratio across different incident angles of interference. In this measurement, the cluster 1’s mainlobe peak angle (i.e., the interference angle for cluster 2) is steered toward 90° and cluster 2’s mainlobe peak angle (i.e., the interference angle for cluster 1) is swept from 0° to 180° in 15° steps. To account for angle estimation errors a half-LSB phase error is included at each step. It is observed that an SIR better than 15 dB is achieved across a wide null-steering interval (22.5°-74° and 106°-157.5°). These SIR measurement results are in accordance with interference rejection contours in Fig. 9 based on the measured RMS phase and amplitude errors of phase shifter (Section VII.B) and VGA (Section VII.C).

Fig. 1.29 shows the 3.5×3mm² die micrograph of the 4-element/6-channel RX prototype fabricated in a 130-nm SiGe BiCMOS process. A two-layer FR-4 PCB which embeds the wire-bonded die was developed for measurements. The RF characterization setup for RF-
Table 1.1: Performance Summary of the 4-Element Beamforming-MIMO RX

<table>
<thead>
<tr>
<th>Architecture</th>
<th>RF Front-End Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency (GHz)</td>
</tr>
<tr>
<td>Process</td>
<td>130nm BiCMOS</td>
</tr>
<tr>
<td>f_T/ f_max (GHz)</td>
<td>200/220</td>
</tr>
<tr>
<td>Integration</td>
<td>RF, LO, Analog BB</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Phased Array</td>
<td>Phased Shifter</td>
</tr>
<tr>
<td></td>
<td>Resolution (°)</td>
</tr>
<tr>
<td>Number of RF channels</td>
<td>6</td>
</tr>
<tr>
<td>Phase Shifting</td>
<td>RF Domain</td>
</tr>
<tr>
<td></td>
<td>Gain Control (dB)</td>
</tr>
<tr>
<td>Amplitude Control</td>
<td>RF Domain</td>
</tr>
<tr>
<td></td>
<td>RMS Gain Error (dB)</td>
</tr>
<tr>
<td></td>
<td>RMS Phase Error (°)</td>
</tr>
</tbody>
</table>

Table 1.2: Table of Comparison

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase Shifting</td>
<td>TRX Phased Array</td>
<td>TRX Phased Array</td>
<td>RX Phased Array</td>
<td>RX Phased Array</td>
<td>RX Beam forming MIMO RF</td>
</tr>
<tr>
<td>Process</td>
<td>130nm BiCMOS</td>
<td>90nm BiCMOS</td>
<td>130nm BiCMOS</td>
<td>130nm BiCMOS</td>
<td>130nm BiCMOS</td>
</tr>
<tr>
<td>Number of Phased Array Channels</td>
<td>16 phased array channels per chip</td>
<td>4 phased array channels</td>
<td>4 phased array channels</td>
<td>4 phased array channels</td>
<td>RF-to-BB Channel Conversion Gain (dB)</td>
</tr>
<tr>
<td>Chip Area (mm^2)</td>
<td>15.8x10.5</td>
<td>3.4x2.1</td>
<td>2x2.7</td>
<td>5.5x5.8</td>
<td>3.5x3</td>
</tr>
<tr>
<td>Integration</td>
<td>RF, LO, BB</td>
<td>RF, LO, BB</td>
<td>RF, LO, BB</td>
<td>RF, LO, BB</td>
<td>RF, LO, BB</td>
</tr>
<tr>
<td>Frequency (GHz)</td>
<td>25-29</td>
<td>71-86</td>
<td>76-84</td>
<td>76-84</td>
<td>64-67</td>
</tr>
<tr>
<td>Channel Conversion Gain (dB)</td>
<td>8</td>
<td>26.2</td>
<td>10.1-18.9</td>
<td>30-33</td>
<td>30.15</td>
</tr>
<tr>
<td>Channel NF (dB)</td>
<td>6**</td>
<td>9-14</td>
<td>10-11</td>
<td>11.4-13</td>
<td>9.8-11**</td>
</tr>
<tr>
<td>Phase Shifter Resolution (°)</td>
<td>4.9</td>
<td>5</td>
<td>11</td>
<td>11</td>
<td>11.25</td>
</tr>
<tr>
<td>Gain Control (dB)</td>
<td>8</td>
<td>-</td>
<td>9</td>
<td>11.2</td>
<td>11</td>
</tr>
<tr>
<td>I/Q Amplitude Mismatch (°)</td>
<td>3300</td>
<td>286*</td>
<td>130**</td>
<td>1000/1200</td>
<td>528</td>
</tr>
<tr>
<td>MIMO Capability</td>
<td>Multiplexing gain</td>
<td>Number of stream</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1.2: Table of Comparison

channel measurements is shown in Fig. 1.30a. The RF-to-BB measurement setup is shown in Figs. 1.30b and 1.30c. E8361A PNA was used for S-parameter measurements and E8257D PSG signal generator along with E4448A PSA spectrum analyzer were utilized for RF-to-BB measurements. Table 1.1 summarizes the detailed RF and RF-to-BB performance of the proposed RX, and Table 1.2 provides the performance comparison with prior work. The
proposed RX supports a MIMO multiplexing gain of 2 in contrast to no multiplexing gain in conventional phased array architectures.

1.9 Conclusion

A 4-element phase-amplitude controlled receiver architecture with simultaneous beamforming and MIMO capabilities both implemented in the RF domain was presented. By partially overlapping RX elements into 2 amplitude-phase controlled clusters, the proposed RX achieves a higher MIMO multiplexing gain compared to conventional phased array or hybrid architectures with the same number elements. Analysis of amplitude control using VGA to enable null-steering independently of mainlobe steering was presented and the effect of phase/amplitude error on the achievable interference suppression was detailed. Measurement results were presented showing excellent agreement with simulation, while verifying the system-level analysis for concurrent spatial multiplexing and interference rejection.
Chapter 2


2.1 Introduction to High-Speed mm-Wave Receivers

There is a rapid increase in demand for high-speed point-to-point wireless links with data-rates comparable to wireline links for both short-range indoor and long-range outdoor communication. Enabling applications include: optical fiber replacement [27–30], high-capacity backhauls, high-speed access networks, resource management in large-scale networks [31], close-proximity wireless data transfer between mobile terminals and storage devices, wireless communication through relay nodes [32], and highly-secure modulation with long code modulation [33,34].

The abundance of available bandwidth in the mm-wave/sub-THz frequency range makes it possible to achieve the capacity of wireline links with the flexibility and low-cost of wireless
links. However, due to the degradation of active device performance at frequencies close to device $f_{\text{max}}$ [35], the operating frequency cannot be arbitrarily high. As a rule of thumb, operation at frequencies as high as $\sim f_{\text{max}}/2$ is considered to be palatable before cliff-fall degradation in the active device performance. Recent developments of advanced commercial silicon processes suggest this upper limit to be somewhere in the F-band (90-140 GHz) [36–39]. This notion implies that high spectral efficiency modulation schemes accompanied by wide bandwidth provides a more practical pathway toward tens-of-Gbps wireless transceivers.

Most importantly, although very high-speed wireless transceiver front-ends based on conventional direct-conversion [40–42] or IF-conversion [43–45] architectures have been reported recently, their inputs/outputs are still modulated baseband or IF signals. Ultra-high-speed and high-resolution data converters are thus required to (de-)modulate raw bits’ information. Based on the Nyquist criteria, the sampling rates of these data converters need to be at least 2 and 4 times the baud-rate of the modulated baseband and IF signals, respectively to avoid aliasing [46]. However, signal-to-noise-distortion-ratio (SNDR) and spurious-free-dynamic-range (SFDR) both quickly degrade with speed, leading to increasingly poor resolution. Accordingly, ultra-high-speed transceivers in the prior-art utilize expensive and bulky high-speed real-time oscilloscopes with speeds and resolutions as high as 200 GSa/s and 12 bits to demodulate their reported data-rate with an acceptable BER [47, 48]. The need for watt-level data converters and digital back-ends makes these high data-rate transceivers extremely energy-inefficient upon (de)modulation. One solution being pursued by prior works is channel-bonding [40, 49]. Nonetheless, it demands several parallel data converters at lower sampling rates, wideband LO and gain characteristics and high levels of calibration. Therefore, one can argue that using conventional architectures, designing an ultra-high data-rate wireless transceiver that also incorporates energy-efficient mixed-signal and baseband units would be quite challenging unless paradigm-shifting architecture-level solutions are explored [50–53].
As will be discussed in Section II, it is of great interest to eliminate data converters and significantly simplify the DSP to pave the way for energy-efficient ultra-high-speed wireless links. In Section III, theory and system analysis of the proposed ADC-less 8PSK direct-demodulation architecture based on a novel RF-correlation idea are detailed. The 8PSK direct-demodulation receiver along with circuit analysis of main building blocks are described in Section IV. The complete measurement results of the fabricated receiver prototype are presented in Section V, and finally, Section VI provides concluding remarks.

### 2.2 High-speed ADC Design Challenges

A fundamental trade-off exists between power dissipation ($P_{\text{diss}}$), resolution ($\text{ENOB}$), and speed ($f_s$) of data converters. This trade-off is captured in the widely used Walden figure-of-merit, $FOM = \frac{P_{\text{diss}}}{f_s2^{\text{ENOB}}}$ [54]. Although finer technology nodes improve the energy efficiency of data converters, dynamic range is severely limited due to simultaneous reduction of supply voltage. This limited dynamic range makes high-resolution (i.e. $\geq 10$-bit) ADC design at higher speeds extremely challenging, if not impossible. This has prompted state-of-the-art high-speed ADCs to focus on improving the $FOM$ by increasing energy efficiency rather than increasing SNDR [55]. However, it is insightful to study the power dissipation overhead of increasing resolution since demodulation BER is limited by the maximum achievable SNDR. According to the thermal noise requirement of an $n$-bit ADC, a lower-bound on its sampling capacitor size is calculated to be:

$$C_s = 12kT \frac{2^{2n}}{V_{\text{in,FS}}^2}$$

where $V_{\text{in,FS}}$ is the full-scale voltage at ADC input. For high-resolution ($>10$-bit) ADCs, power dissipation is dominated by thermal noise and grows proportionally with $2^{2n}$. On the other hand, for low-resolution ($<6$-bit) ADCs, it is dominated by component mismatch.
requirement and minimum capacitor size and grows proportionally with $2^n$ [56]. With simultaneous technology and supply scaling the required sampling capacitor for a given resolution increases due to reduced permissible noise levels [56], [57]. Therefore, energy efficiency of even low-resolution ADCs in finer technology nodes gets limited by more stringent thermal noise requirements.

Design of the baseband circuit is determined by the required RF bandwidth, power budget and the intended application, ranging from few Hz and micro watts for biomedical [58–61] and power-cycling [62] applications to several GHz and watts for high speed wireless communications. Additionally, ADC must maintain the required resolution across a very wide bandwidth for high-speed signal demodulation. As the parasitic capacitances become comparable with the sampling capacitor, the power-speed trade-off becomes non-linear with increasing speed resulting in continuous $FOM$ degradation [63]. Consequently, ADCs operating above $f_{FOM,\text{diff}}$ [64] will experience a drastic degradation in their resolution and energy efficiency. This $f_{FOM,\text{diff}}$, even for ADCs designed in nanoscale technologies, is in the order of only a few hundreds of MHz [64]. To achieve multi-GHz sampling rates without severely sacrificing
FOM, state-of-the-art ADCs utilize time interleaving of lower-speed sub-ADCs operating in their linear power-speed regime (see Fig. 2.1) [63]. However, still a power-hungry front-end driver with the same acquisition bandwidth as that of the overall ADC is required to drive the equivalent sampling capacitor ($C_{s,eq}$) at the ADC input. The trade-off between $kT/C_{s,eq}$ noise and bandwidth of the input driver in addition to its linearity requirements limits the available input SNDR/SFDR [65].

Furthermore, even design of the core time-interleaved ADC (i.e. sub-ADCs) for high multi-Gbps data-rate communication systems is challenging. More specifically, time-interleaved ADCs are quite sensitive to inter-channel mismatches. Gain mismatch and timing skew between channels need to be precisely calibrated to avoid aliasing and degradation of the signal integrity. Gain mismatch between channels results in an undesired signal-amplitude-dependent image in the signal spectrum. For example, to obtain an SNR of 45 dB (~8-bit resolution), a gain matching better than 0.5% is required in a 4-channel time-interleaved ADC [66]. Gain error can be detected and corrected digitally. However, the associated power overhead highly depends on the activity factor of logic gates which limits the maximum allowable number of logic gates to only a few thousands in low-resolution ADCs [67].

Furthermore, timing mismatch between channels will result in an undesired input-frequency-dependent image in the signal spectrum. To quantize a signal at frequency of $f_{in}$ with a resolution of $ENOB$, the variance of timing mismatch in an N-channel time-interleaved ADC must be less than [68]:

$$\sigma_T^2 \leq \frac{N}{N-1} \cdot \frac{2/3}{(2^{ENOB} \cdot 2\pi f_{in})^2}$$  \hspace{1cm} (2.2)

For example, for an input signal with 12-GHz bandwidth, a timing mismatch of $\sim 50 \, fs$ limits the ADC resolution to 8 bits in a 4-channel time-interleaved ADC. Achieving such a low timing mismatch is quite challenging even in current technologies and its correction requires
precisely controllable analog delay lines or high-order digital filters [69]. Additionally, each channel often occupies a large area to meet the mismatch requirements of its sub-ADC, thereby mandating long multi-phase clock routings. These long routing interconnects cause signal integrity issues and increase the power dissipation of the clock network especially at higher speeds.

Time-interleaved ADCs with sampling rates as high as 64 GSa/s have been reported in prior-art. However, due to the challenges discussed above, they suffer from low-resolution (e.g. 5.95 ENOB at Nyquist frequency) and high power dissipation (e.g. ~950 mW) [70]. Importantly, the reported state-of-the-art FOMs often leave out power dissipation of certain parts of a complete ADC including the input S/H amplifier or driver, clock generation/distribution network, reference generation/calibration (gray blocks in Fig. 2.1) [71]. Therefore, ADC remains one of the most power-hungry and challenging blocks in a high data-rate communication system.

2.3 High-Order Direct-Demodulation

High-speed receivers requiring no ADCs and achieving data-rates as high as 16 Gbps have been reported in the prior-art [41, 72]. However, these direct-demodulation architectures only support low-order modulation schemes (OOK, BPSK, and QPSK). To further increase the data-rate, these low spectral efficiency architectures need to be designed with much wider bandwidths at very high carrier frequencies (close to $f_{max}$), resulting in poor wireless link quality due to limited receiver sensitivity and transmitter output power at such high frequencies. This motivates the design of direct-demodulation architectures for high-order modulations to increase the spectral efficiency and achievable data-rate for a given bandwidth. As a natural extension following the already reported direct-demodulation architectures up to QPSK, this work presents a direct-demodulation of 8PSK constellation...
based on multi-phase RF-correlation technique which is amenable to mm-wave frequencies.

2.3.1 Current 8PSK Demodulation Techniques

The most popular way currently for 8PSK demodulation is the arctangent (ATAN) technique in parallel with a digital PLL (DPLL) that calculates the distance between complex baseband envelope and constellation points in the IQ signal space [73]. ATAN function has a complex digital implementation and requires large lookup tables (LUTs) and memory blocks to distinguish and scale phases; all being extremely challenging to implement at ultra-high speeds targeted in our work. A similar demodulation method to ATAN is the Costas-loop technique [74], where the received signal is fed to four matched filters controlled by 180°, 225°, 270°, 315° phase-shifted outputs of a reference Costas-loop. The correct received symbol is detected by finding the maximum amplitude at the output of four matched filters over one symbol period. This is a 5-level amplitude decision and requires a power-hungry high-speed ADC for our target data-rate to detect the received symbol corresponding to the maximum amplitude level.

Another digital 8PSK demodulation technique is based on complex number method [75], where IQ baseband signal is sampled with a high sampling rate processor and converted to a complex number. Afterwards, the sampled received complex number gets multiplied with a delayed conjugate version of itself and the product is then fed to a complex 8-phase slicer to detect the received symbol. Implementation of this 8PSK demodulation technique again requires ultra-high-speed processor and high levels of field programmable gate array (FPGA) resources.

Cross-correlation technique is also proposed in [76], where the 8PSK-modulated signal is cross-correlated with 180°, 225°, 270°, 315° phase-shifted versions of itself to detect eight separate angles. Realization of this technique is extremely challenging at ultra-high input
frequencies and data-rates, as the correlated signal at the output of each cross-correlator is at twice the input frequency and is composed of positive (in-phase) and negative (180° out-of-phase) terms. A sign-slicer is used to separate positive and negative parts. The absolute value of both parts is calculated and the peak absolute value among cross-correlator outputs identifies the correct symbol. The absolute values of cross-correlated angles need to be updated, compared and recorded in LUTs every half cycle of the carrier, which requires significant amount of hardware resources.

2.3.2 Proposed 8PSK Direct-Demodulation Technique

To obviate the need for ultra-high-speed high-resolution ADCs and complicated digital demodulation computations with high-level of FPGA resources, a simple 8PSK direct-demodulation technique is presented in this work. Fig. 2.2 shows the proposed multi-phase RF-correlation-based technique for direct-demodulation of 8PSK symbols. The 2D IQ RF signal space is partitioned by four differential LO phases into eight angular subsections (Fig. 2.2a). Phase references of RF carrier and LO are offset by 22.5° to maximize the Euclidean distance between RF symbols and the LO boundaries, thus maximizing the error tolerance in detecting the symbols. A multi-phase RF-correlator is utilized for 8PSK symbol detection (Fig. 2.2b). By downconverting symbols with four differential 45° phase-shifted mixers followed by low-pass filtering, only the sign of RF-correlated signals \( C_{\text{in},k=1,2,3,4} \) is needed to determine the received symbols. Table 2.1 shows the normalized correlation values of 8PSK symbols with LO phases. Assuming Gray-coding, the 3 bits per symbol are readily extracted from the retimed output of RF correlators with simple XOR logic gates. These 3 demodulated bits are related to \( C_{\text{out},k=1,2,3,4} \) with simple Boolean expressions: \( B_2 = C_{\text{out},1} \), \( B_1 = C_{\text{out},3} \), and \( B_0 = C_{\text{out},2} \oplus C_{\text{out},4} \) (Table 2.2).
Figure 2.2: (a) 2D IQ signal space of 8PSK symbols partitioned by 8 LO phases, and (b) block diagram of multi-phase RF-correlator and sign-check comparators.

Table 2.1: Normalized correlation values of 8PSK symbols with eight LO phases.

<table>
<thead>
<tr>
<th>Symbols</th>
<th>LO Phases</th>
<th>LO0°</th>
<th>LO180°</th>
<th>LO45°</th>
<th>LO225°</th>
<th>LO90°</th>
<th>LO270°</th>
<th>LO135°</th>
<th>LO315°</th>
</tr>
</thead>
<tbody>
<tr>
<td>S₁</td>
<td></td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.38</td>
<td>-0.38</td>
<td>-0.38</td>
<td>+0.38</td>
</tr>
<tr>
<td>S₂</td>
<td></td>
<td>+0.38</td>
<td>-0.38</td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.38</td>
<td>-0.38</td>
</tr>
<tr>
<td>S₃</td>
<td></td>
<td>-0.38</td>
<td>+0.38</td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.92</td>
</tr>
<tr>
<td>S₄</td>
<td></td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.38</td>
<td>+0.38</td>
<td>+0.38</td>
<td>-0.38</td>
<td>+0.92</td>
<td>-0.92</td>
</tr>
<tr>
<td>S₅</td>
<td></td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.38</td>
<td>+0.38</td>
<td>+0.38</td>
<td>-0.38</td>
</tr>
<tr>
<td>S₆</td>
<td></td>
<td>-0.38</td>
<td>+0.38</td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.38</td>
<td>+0.38</td>
</tr>
<tr>
<td>S₇</td>
<td></td>
<td>+0.38</td>
<td>-0.38</td>
<td>-0.38</td>
<td>+0.38</td>
<td>-0.92</td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.92</td>
</tr>
<tr>
<td>S₈</td>
<td></td>
<td>+0.92</td>
<td>-0.92</td>
<td>+0.38</td>
<td>-0.38</td>
<td>-0.38</td>
<td>+0.38</td>
<td>-0.92</td>
<td>+0.92</td>
</tr>
</tbody>
</table>

Table 2.2: Logic table of three demodulated bits per symbol.

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Correlator</th>
<th>C_{out,1}</th>
<th>C_{out,3}</th>
<th>C_{out,2} \oplus C_{out,4}</th>
</tr>
</thead>
<tbody>
<tr>
<td>S₁</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>S₂</td>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>S₃</td>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>S₄</td>
<td></td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>S₅</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>S₆</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>S₇</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>S₈</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
2.3.3 BER of the proposed 8PSK direct-demodulation Method

Receiver RF 8PSK symbols, \( S_{i=1,2,...,8}(t) \), after downconversion by eight differential LO phases (LO\(_{0}\), LO\(_{180}\), LO\(_{45}\), LO\(_{225}\), LO\(_{90}\), LO\(_{270}\), LO\(_{135}\), LO\(_{315}\)) and baseband filtering, generate four parallel 4-level pulse-amplitude modulation (PAM-4) signals with unequal spacing. The equivalent PAM-4 symbol (\( S_{m=1,2,3,4,PA M} \)) for each 8PSK symbol (\( S_{i=1,2,...,8} \)) is derived in Eq. (2.3). In this equation, \( \omega_B \) is the baseband angular frequency and \( g(t) \) is a unity baseband pulse with a unity amplitude in the symbol period (\( 0 < t < T_s \)) and zero everywhere else. As shown in Fig. 2.3a, for a baseband signal energy of \( E_g \) these four levels are: 

\[-d_2 \sqrt{E_g}/2, -d_1 \sqrt{E_g}/2, +d_1 \sqrt{E_g}/2, \text{ and } +d_2 \sqrt{E_g}/2, \]

where \( d_1^2 + d_2^2 = 1 \) and \( d_2/d_1 = \cos(22.5^\circ)/\cos(67.5^\circ) \approx 2.4. \)

Modulated baseband signal energy has a normal Gaussian distribution at each level of the PAM-4 eye-diagram. The probability of receiving an error in detecting \( B_2 \) of 8PSK symbols \( S_{i=1,4,5,8}(t) \) (red regions in Fig. 2.3a) upon RF correlation with (LO\(_{0}\), LO\(_{180}\)) is equal to:

\[
P_r(E|S_{1,PA M}) = P_r(E|S_{1 \text{ or } 8}) = \int_{-\infty}^{0} P(r|S_{1 \text{ or } 8})dr = \int_{-\infty}^{0} \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(r-d_2 \sqrt{E_g}/2)^2}{N_0}} dr = Q(\sqrt{\frac{d_2^2 E_g}{N_0}})
\]

\[
S_{4,PA M} : \begin{cases} \begin{align*} &-g(t)d_2 \cos(\omega_B t), S_{i=4,5} \end{align*} \\ &-g(t)d_2 \cos(\omega_B t), S_{i=4,5} \end{cases} \begin{align*} &-g(t)d_2 \cos(\omega_B t), S_{i=4,5} \end{align*}
\]

\[S_{3,PA M} : \begin{cases} \begin{align*} &-g(t)d_1 \cos(\omega_B t), S_{i=3,6} \end{align*} \\ &-g(t)d_1 \cos(\omega_B t), S_{i=3,6} \end{cases} \begin{align*} &-g(t)d_1 \cos(\omega_B t), S_{i=3,6} \end{align*}
\]

\[S_{2,PA M} : \begin{cases} \begin{align*} &+g(t)d_1 \cos(\omega_B t), S_{i=2,7} \end{align*} \\ &+g(t)d_1 \cos(\omega_B t), S_{i=2,7} \end{cases} \begin{align*} &+g(t)d_1 \cos(\omega_B t), S_{i=2,7} \end{align*}
\]

\[S_{1,PA M} : \begin{cases} \begin{align*} &+g(t)d_2 \cos(\omega_B t), S_{i=1,8} \end{align*} \\ &+g(t)d_2 \cos(\omega_B t), S_{i=1,8} \end{cases} \begin{align*} &+g(t)d_2 \cos(\omega_B t), S_{i=1,8} \end{align*}
\]

PAM-4 Symbols:

\( \begin{align*} &\overrightarrow{1} \quad (LO_{0} \quad \overrightarrow{1} \quad LO_{180}) \\ &\overrightarrow{1} \quad (LO_{45} \quad \overrightarrow{1} \quad LO_{225}) \\ &\overrightarrow{1} \quad (LO_{90} \quad \overrightarrow{1} \quad LO_{270}) \\ &\overrightarrow{1} \quad (LO_{135} \quad \overrightarrow{1} \quad LO_{315}) \end{align*} \)

(2.3)
\[ P_r(E|S_{4,PAM}) = P_r(E|S_4 \text{ or } 5) = \int_0^\infty P(r|S_4 \text{ or } 5)dr \]
\[ = \int_0^\infty \frac{1}{\sqrt{\pi N_0}}e^{-\frac{(r+d_2\sqrt{\gamma g^2})^2}{N_0}}dr = Q\left(\sqrt{\frac{d_2^2\gamma g}{N_0}}\right) \]  

(2.5)

where \( N_0 \) and \( Q(x) \) are the noise power spectral density and the tail distribution function of normal Gaussian distribution, respectively. Similarly, the probability of receiving an error in detecting \( B_2 \) of 8PSK symbols \( S_i=2,3,6,7(t) \) (blue regions in Fig. 2.3a) upon RF correlation with \((LO_{0^\circ}, LO_{180^\circ})\) is equal to:

\[ P_r(E|S_{2,PAM}) = P_r(E|S_{3,PAM}) \]
\[ = P_r(E|S_2 \text{ or } 7) = P_r(E|S_3 \text{ or } 6) = Q\left(\sqrt{\frac{d_2^2\gamma g}{N_0}}\right) \]  

(2.6)

Based on Eqs. (2.4), (2.5) and (2.6), the bit error probability of \( B_2 \) extracted from RF correlation of 8PSK symbols \( (S_{i=1,2,\ldots,8}) \) with \((LO_{0^\circ}, LO_{180^\circ})\) is equal to:

\[ P_{B_2}(E) = \sum_{i=1}^{8} P_r(E|S_i) \]
\[ = \frac{4}{8}\left(P_r(E|S_1 \text{ or } 4 \text{ or } 5 \text{ or } 8) + P_r(E|S_2 \text{ or } 3 \text{ or } 6 \text{ or } 7)\right) \]
\[ = \frac{1}{2}\left(Q\left(\sqrt{\frac{d_2^2\gamma g}{N_0}}\right) + Q\left(\sqrt{\frac{d_2^2\gamma g}{N_0}}\right)\right) \]  

(2.7)

Similarly, the bit error probability of \( B_1 \) upon RF correlation of 8PSK symbols with \((LO_{90^\circ}, LO_{270^\circ})\) is equal to the bit error probability of \( B_2 \) in (2.7). For an average baseband signal energy per bit of \( E_{b,\text{avg}} = (2(d_1\sqrt{\gamma g/2})^2 + 2(d_2\sqrt{\gamma g/2})^2)/6 = \gamma g/6 \), the bit error probabilities of \( B_2 \)
and $B_1$ are equal to:

$$P_B(E) = P_{B_1}(E) = P_{B_1}(E)$$

$$\approx \frac{1}{2} Q\left(\sqrt{\frac{6\sigma_{b,avg}}{N_0}}\cos(22.5^\circ)\right) + \frac{1}{2} Q\left(\sqrt{\frac{6\sigma_{b,avg}}{N_0}}\sin(22.5^\circ)\right) \tag{2.8}$$

Bit error probability of $B_0$ is calculated by XORing the bit error probabilities of 8PSK symbols after RF correlations with $(LO_{45^\circ}, LO_{225^\circ})$ and $(LO_{135^\circ}, LO_{315^\circ})$, which is equal to $P_{B_0}(E) = 2P_B(E) - P_B(E)^2$. $P_{B_0}(E), P_{B_1}(E),$ and $P_{B_2}(E)$ are shown in Fig. 2.3b. The symbol error rate of the proposed 8PSK direct-demodulation is the same as the theoretical 8PSK symbol error rate. Therefore, the proposed multi-phase RF-correlation method enables direct-demodulation of 8PSK symbols without imposing any degradation in detection of those symbols. Furthermore, another advantage of this direct-demodulation method is that symbols are detected using only sign-check comparators (i.e., BPSK decision) and simple Boolean expressions, thereby requiring no multi-level amplitude decision or peak detection. This sign-check comparison (BPSK decision) reduces symbol detection sensitivity to LO phase deviation. In theory, as long as each LO phase deviation from its ideal angle remains less than 22.5°, all levels in PAM-4 eye-diagram will be non-zero for correct sign-check decision. Therefore, LO phase error in four RF-correlators can be tolerated as long as BER
2.4 Proposed 8PSK Direct-Demodulation Receiver

The proposed 8PSK direct-demodulation receiver architecture is shown in Fig. 2.4 [50, 51]. It is comprised of a wideband LNA followed by a balun-based 1-to-4 splitter. The splitter outputs provide four phase-matched differential RF signals at the RF ports of double-balanced passive mixers driven by four 45° phase-shifted differential LOs.

The 125 GHz LO distribution network generates close to 0 dBm at the center frequency after frequency multiplication by 6 (a tripler followed by a doubler) of the bondwired 20.83 GHz input. Tunable varactor-based low-pass and high-pass phase shifters with 45° phase difference at 62.5 GHz and tunable all-pass phase shifters with 45° phase difference at 125 GHz are designed to generate four 45° phase-shifted differential LO signals at the LO ports of the double-balanced passive mixers. The downconverted signals are then fed to baseband amplifiers/active filters and continuous-time linear equalizers (CTLEs). Due to the high
Figure 2.5: (a) 6-Stage common-emitter-based LNA circuit schematic, (b) simultaneous noise and power matching in the first LNA stage ($Z_{opt}$ and $Z_{in}^*$ curves).

frequency of operation, and therefore sparse multi-path and small delay-spread [77], a simple CTLE suffices to compensate for bandwidth limitations.

An 8PSK test output is taken at the output of each mixer-baseband to adjust the LO and RF-carrier phase references at $22.5^\circ$ offset during measurement for optimum symbol detection. The CTLE outputs are then fed to four sign-check comparators followed by three XOR gates to extract the three bits per symbol based on the theory discussed in Section III.

Figure 2.6: (a) Second single-ended LNA stage layout with MIM bypass capacitors, (b) second differential LNA stage layout with CPW-based matching networks, and (c) LNA S-parameter and NF simulation results.
2.4.1 LNA circuit design

Before opting for the popular LNA cascode topology, the noise contribution of the cascode device needs to be taken into account due to the high frequency of operation. The current noise density of the cascode device appears as $\overline{i_{n,\text{out}}}^2 > 2kTg_m(f/f_T)^2$ at its output. At the center frequency of 125 GHz and a process $f_T$ of 325 GHz, the effect of the cascode device on the LNA NF is thus no longer negligible. Therefore, a common-emitter-based design is adopted for this high frequency LNA.

The 6-stage common-emitter-based LNA is shown in Fig. 2.5a. The LNA’s first stage is designed at a current-density of $J_{DC} = 0.35 \text{ mA/µm}$ as a compromise between $NF_{\text{min}}$ and $G_{\text{max}}$. The first stage is degenerated with a CPW of length $EL$ to provide simultaneous noise and power matching. Variations of $Z_{\text{opt}}$ and $Z_{\text{in}}^*$ with $EL$ are shown in Fig. 2.5b. For $EL = 47.36^\circ$, simultaneous noise and power match is achieved. The remaining stages are optimized for $G_{\text{max}}$ at a current-density of $J_{DC} = 0.7 \text{ mA/µm}$. The frequency responses of the LNA stages are stagger-tuned to maximize the operation bandwidth.

As an example, the layout of the second single-ended stage of the LNA is shown in Fig. 2.6a. MIM capacitors are used as bypass capacitors and are EM-simulated to provide self-resonance-frequency and therefore a small impedance ($\sim 1-2 \Omega$) at the frequency of operation. Likewise, the layout of the second differential stage of the LNA is shown in Fig. 2.6b. Matching networks are composed of low-loss 50 Ω CPW-based T-sections and finger capacitors. Cascaded bends and T-junctions are the main sources of electrical and magnetic field discontinuity in long CPW routings (e.g. matching networks). Bends and T-junctions are cut with a $45^\circ$ angle to improve their reflection coefficient and loss compared to a standard bend and T-junction with $90^\circ$ corner\(^1\). Finger capacitors are realized on (M5-)M6-M7 for values above $30 \text{ fF}$ and on M8 for values below $15 \text{ fF}$, and their port-to-port EM-simulated phase

\(^1\)By applying this technique in a 55-nm SiGe BiCMOS process, a single 50 Ω CPW bend with width of 6.5 µm and lateral ground spacing of 9.4 µm can achieve 8 dB better reflection coefficient at 125 GHz.
delay is taken into account during the design of conjugate matching networks. S-parameter and NF simulation results of the LNA are shown in Fig. 2.6c. LNA achieves a maximum gain of 22.6 dB and a minimum NF of 9.2 dB across the RF bandwidth.

2.4.2 1-to-4 balun-based splitter design

A 1-to-4 splitter is required at the LNA output to feed the downconversion mixers with four in-phase differential RF signals. A balun-based splitter is designed to avoid cross-overs and the high loss of long λ/4 CPWs in conventional differential Wilkinson splitters. The splitter is composed of six baluns, as indicated in Fig. 2.7a [78]. Ground walls on metal stack M1-M8 for the two input baluns and on M1-M6 for the four output baluns are designed to reduce coupling between adjacent baluns. Each balun is double-tuned with input and output finger capacitors to match the primary to 50 Ω and the secondary to 100 Ω differential.

The outputs of the two input baluns are connected to the four output baluns with 100 Ω differential CPWs. Assuming loaded quality factors of $Q_p$ and $Q_s$ for the primary and secondary of each balun, respectively, the overall quality factor is: $Q_{net} = \frac{1}{\sqrt{0.5(1/Q_p^2 + 1/Q_s^2)}}$. The frequency response of a double-tuned balun is maximally flat if the balun’s coupling factor is $K = 1/Q_{net}$ [78]. On the other hand, for $K < 1/Q_{net}$ the response is still flat over a narrower
Figure 2.8: (a) Voltage-mode double-balanced passive mixer followed by 3-stage amplification (last amplification stage is CTLE), (b) simplified circuit of a double-balanced passive mixer, and (c) mixer input matching from 110 GHz to 140 GHz.

bandwidth, and for $K > 1/Q_{net}$ primary and secondary resonance frequencies will be further apart, resulting in wider fractional bandwidth at the cost of higher in-band ripples. \( \sim 45 \) pH primary and secondary inductors of each balun are double-tuned with 29 fF and 62 fF single-ended finger capacitors, respectively. Coupling factor of each balun is designed for \( K \approx 0.48 \) close to \( 1/Q_{net} \) to achieve a maximally flat response. To provide isolation between output ports of the splitter, they are cross-connected with 100 $\Omega$ resistors. The S-parameter simulation results of the 1-to-4 balun-based splitter are shown in Fig. 2.7b. The splitter shows a low loss of 8.3 dB across a wide bandwidth of 40 GHz. The isolation between output ports remains better than 15 dB across the bandwidth. The amplitude and phase errors of the splitter are shown in Fig. 2.7c. The amplitude and phase errors between output ports are less than 0.2 dB and 1.5°, respectively. The differential output ports phase error from 180° is less than 0.5°.

2.4.3 Mixer-baseband circuit design

Each mixer-baseband is composed of a voltage-mode double-balanced passive mixer followed by three amplification stages with the last stage being a CTLE (Fig. 2.8a). Shunt peaking and capacitive neutralization are used in the first two stages to increase bandwidth with-
out sacrificing gain or noise performance. CTLE is an RC-degenerated differential stage with tunable resistors and capacitors. RC-degeneration introduces a zero in the baseband frequency response to compensate for bandwidth limitations imposed by the receiver frontend. CTLE incorporates 3-bit tuning for both degeneration resistors and capacitors. Proper CTLE code is chosen during measurement to optimize the baseband frequency response based on the existing trade-off between bandwidth enhancement and noise integration for wider bandwidths. The long routing between CTLE output and the next-stage sign-check comparator serves as series peaking inductor and helps further increase in the baseband bandwidth.

The simplified circuit of the voltage-mode double-balanced passive mixer is shown in Fig. 2.8b. To calculate RF-referred conversion gain, high-side and low-side RF injections are defined as $V_{RF} = A_{RF}\cos(\omega_H t + \phi)$ and $V_{RF} = A_{RF}\cos(\omega_L t + \phi)$, respectively, where $\omega_H = \omega_{LO} + \omega_B$ and $\omega_L = \omega_{LO} - \omega_B$ and $\phi$ is the phase offset between LO and RF-carrier (i.e. 22.5° based on Section III). Assuming square-wave 50% LO, the high-side and low-side conversion gains of this voltage-mode double-balanced mixer are [79]:

$$G_{\text{conv}}(\omega) = \begin{cases} 
\frac{(2/\pi)Z_{BB}(\omega_H)}{Z_s(\omega_H) + R_{sw} + Z_{BB}(\omega_B)}, & \omega = \omega_H \\
\frac{(2/\pi)Z_{BB}(\omega_L)}{Z_s(\omega_L) + R_{sw} + Z_{BB}(\omega_B)}, & \omega = \omega_L
\end{cases}$$

(2.9)

Therefore, by designing $Z_s(\omega)$ to exhibit resonance at the mixer input (i.e. $Z_s(\omega)$ real), high-side and low-side conversion gains can be equalized and a symmetrical frequency response is achieved at the passband of mixer’s input impedance. The mixer’s input impedance from 110 GHz to 140 GHz is shown in Fig. 2.8c. The low-side and high-side conversion gains (both are almost equal) of mixer-baseband for four different CTLE code settings are shown in Fig. 2.9a. The mixer-baseband achieves a maximum conversion gain of 24.7 dB (mixer conversion gain = -5.7 dB).
Due to reciprocity of passive mixers and 50% LO design, interaction between switches needs to be studied. Assuming a high-side RF injection, the RF voltage at the input of the mixer driven by differential 50% LO $(LO_{m \times 45^\circ}, LO_{180^\circ + m \times 45^\circ})$ is:

$$V_{RF,m}(\omega) = \frac{(4A_{RF}/\pi^2)Z_{BB}(\omega_B)}{Z_s(\omega_H) + R_{sw} + Z_{BB}(\omega_B)} \times \cos(\omega_H t + \phi)$$

$$+ \frac{(4A_{RF}/\pi^2)Z_{BB}^*(\omega_B)}{Z_s(\omega_L) + R_{sw} + Z_{BB}^*(\omega_B)} \times$$

$$\begin{cases}
+\cos(\omega_L t - \phi), & m = 0 \\
-\sin(\omega_L t - \phi), & m = 1 \\
-\cos(\omega_L t - \phi), & m = 2 \\
+\sin(\omega_L t - \phi), & m = 3
\end{cases}$$

(2.10)

where the first term is produced by upconversion of the baseband signal to the main frequency (high-side) and the second term is produced by upconversion of the baseband signal to the image frequency. Similarly, the RF voltages at the input of mixers for low-side injection can be derived by replacing $\omega_B$ with $-\omega_B$ and $\omega_H$ with $\omega_L$ and vice versa. Extra gain stages can be placed at the input of four constituent double-balanced mixers to reduce the cross-talk between them, but such gain stages will degrade the linearity and bandwidth.

Figure 2.9: (a) Conversion gain of mixer-baseband for four different CTLE code settings, and (b) mixer-baseband NF.
and will increase the receiver power dissipation. Therefore, as mentioned in Section IV.B, a passive approach is pursued. Splitter output ports are isolated to minimize the interaction between four mixers.

Each double-balanced passive mixer is followed by differential amplification stages. Assuming high enough gain from each amplification stage to suppress the noise of latter stages, only the first differential stage is considered in NF calculation. For a real $Z_s(\omega) = R_s$ (due to matching) at the mixer input, the input-referred NF is:

$$NF = \frac{\pi^2}{4} \left| \frac{R_s + R_{sw} + Z_{BB}(\omega_B)}{Z_{BB}(\omega_B)} \right|^2 \times \frac{2(R_{sw} + r_b) + \left(\frac{g_m}{\beta}\right) r_b^2 + \frac{g_m + 2/R_L}{g_m}}{2R_s}$$

(2.11)

where $g_m$, $\beta$, $r_b$, and $R_L$ are the trans-conductance, current factor, base and load resistor of the first differential stage. The simulated NF of the mixer-baseband is shown in Fig. 2.9b. The integrated NF across the baseband bandwidth is 9.23 dB.
2.4.4 Comparator

CML-based comparator design is adopted to drive CML-based logic gates and to enable high-speed pipeline operation. Each comparator and the following retiming DFF are combined into one comparator-DFF block (Fig. 2.10a). The comparator is biased with a tail current of 1.6 mA for a Gain x BW product of 105.13 GHz. To compensate for offset across PVT, a calibration differential pair with tunable reference voltage is added in parallel with the comparator input differential pair. The reference voltage is fed from a 5-bit resistive ladder. This design enables calibrating offset voltages from -17 mV to 17 mV with 2 mV accuracy. After offset calibration, each master- and slave-latch can fully regenerate input signals as small as 10 mVp at 12 GHz clock-rate.

To estimate BER degradation due to random noise, the effect of noise in both sampling and regeneration periods is considered. Comparator is a linear periodically time-variant (LPTV) system with cyclo-stationary noise. SNR at the output of the comparator is thus calculated from differential output voltage and integrated noise power spectral density at each time instance within one clock period, i.e., $BER_{B,noise} = Q(\sqrt{SNR_{out}}) = Q\left(\frac{V_{out,comp}}{V_{n,RMS}}\right)$, where $V_{out,comp}$ and $V_{n,RMS}$ are the differential output voltage and RMS noise voltage at
the comparator output. The simulated steady-state output voltage, RMS noise voltage and $SNR_{out}$ are shown in Fig. 2.10b. Signal and noise (accumulated in sampling period) are amplified by exponential gain of the cross-coupled pair in the regeneration period. However, as the comparator output latches to a large-signal digital value, signal and noise will no longer affect BER and output noise will be heavily compressed. Therefore, BER is determined by $SNR_{out}$ at the time instance when the gain of master-latch is maximized but still not compressed. Based on Fig. 2.10b, $SNR_{out}$ at this time instance (i.e. $t_{BER}$) is $\sim 18$ dB resulting in a BER less than 1e-15 for a 12 Gbps input signal.

Another source of BER degradation is metastability due to very small input signals. A timing diagram of the comparator-DFF operation is shown in Fig. 2.11. $V_{in,swing}$, $V_{in,sens}$, and $V_{out,sens}$ are the input full-swing, comparator sensitivity level, and slave-latch sensitivity level, respectively. As long as the input signal is large enough for the master-latch to regenerate it to the slave-latch sensitivity level, the slave-latch will be able to regenerate it to a digital value and correctly detect the bit at half-period clock sampling. The comparator and slave-latch outputs for this case are shown in black in Fig. 2.11 (for very small $V_{in,sens}$, the gray traces will be expected). The regeneration time-constant of the comparator cross-coupled pair terminated with $C_L$ at each output is equal to $\tau_{reg} = R_L(C_\pi + 4C_\mu + C_L)/(g_m R_L - 1)$. Assuming $G_{amp}$ is the unlatched gain of the comparator, the output amplitude at the end of the regeneration period is:

$$V_{out,comp}(t = \frac{T_s}{2}) = V_{out,comp}(0)e^{\frac{T_s}{2\tau_{reg}}}$$

$$= \left( G_{amp} \cdot V_{in,comp}(0) \right) e^{\frac{T_s}{2\tau_{reg}}}$$

(2.12)

where $V_{in,comp}$ and $V_{out,comp}$ are the comparator input and output voltages. The comparator’s bit error probability is equal to the probability of $|V_{in,comp}(0)| < |V_{in,sens}|$, which is:

$$BER_{B,meta} = P(|V_{in,comp}(0)| < |V_{in,sens}|)$$

$$= \frac{V_{out,sens}}{G_{amp} \cdot V_{in,swing}} e^{\frac{-T_s}{2\tau_{reg}}},$$

(2.13)
therefore, smaller $\tau_{\text{reg}} \propto (\text{Gain} \times \text{BW})^{-1}$ results in smaller BER. The designed $\tau_{\text{reg}}$ is 1.45 ps resulting in a BER less than 1e-13 for a 12 Gbps input signal. Therefore, in the proposed direct-demodulation architecture, baseband imposes negligible degradation to BER of the demodulated output bits.

### 2.4.5 LO Generation and Distribution Network

LO network is designed to generate four 45° phase-shifted differential LO signals with close to 0 dBm power at the LO ports of four double-balanced mixers. The bondwired input signal at 20.83 GHz is first buffered and then fed to a tripler to reduce the tripler’s output power sensitivity to bondwire inductance variation. Transistor size and biasing point of the tripler are set to maximize the third harmonic current. A balun tuned at 62.5 GHz provides the optimum load to the tripler at the third harmonic and converts the single-ended output to differential for the rest of the LO chain. The schematic and simulation results of the tripler are shown in Figs. 2.12a and 2.12b, respectively. The saturated output power of the tripler is -1.2 dBm at 62.5 GHz.

The 62.5 GHz tripler output is then split with a differential Wilkinson splitter to two low-
pass and high-pass LO branches. Tunable varactor-based low-pass and high-pass phase shifters with 45° phase difference at 62.5 GHz are designed at each Wilkinson’s output port. The layouts of low-pass and high-pass phase shifters are shown in Figs. 2.13a and 2.13b, respectively. Low-pass/high-pass phase tuning range is 29.9° (Fig. 2.13c) and the loss is less than 1 dB at 62.5 GHz (Fig. 2.13d). These low-pass and high-pass LO branches are then followed by 62.5 GHz capacitively-neutralized buffers and CPW-based doublers to generate quadrature LOs at 125 GHz. The circuit schematic and simulation results of the doubler are shown in Figs. 2.14a and 2.14b, respectively. The saturated output power of the doubler is -1.8 dBm at 125 GHz. The output of the doublers from two quadrature branches are then fed to differential 1-to-2 balun-based splitters (similar to the 1-to-4 balun-based splitter in section IV.B) and followed by 125 GHz capacitively-neutralized buffers to create four LO
branches. Four tunable varactor-based all-pass phase shifters with 45° phase difference at 125 GHz are designed at each of these splitters output ports. The layouts of all-pass phase shifters are shown in Figs. 2.15a and 2.15b. All-pass phase tuning range is 14.1° (Fig. 2.15c) and the loss remains less than 1.3 dB (Fig. 2.15d). These four 45° phase-shifted LOs are then further amplified with 125 GHz capacitively-neutralized buffers and local LO amplifiers at the LO ports of the four differential double-balanced mixers. The saturated output power of the LO network at different output center frequencies and for multiple harmonics of the bondwired 20.83 GHz LO input are shown in Figs. 2.16a and 2.16b, respectively.

2.5 Measurement Results

The die micrograph of the chip fabricated in a 55-nm SiGe BiCMOS process is shown in Fig. 2.17. It occupies 2.5×3.5 mm² of die area including PADs and test circuits (2.5 mm² active area). The direct-demodulation receiver consumes a total DC power of 200.25 mW from 1.5-V/2.0-V supplies (LNA: 16 mW, mixer-basebands: 101 mW, LO network: 49.15 mW, and demodulator: 34.1 mW).

To test the chip a six-layer PCB (3 copper signal layers and 3 copper internal ground planes)

![Figure 2.14: (a) Doubler circuit schematic, and (b) doubler output power.](image)
Figure 2.15: (a) Lower-frequency-tuned all-pass phase shifter, and (b) higher-frequency-tuned all-pass phase shifter, (c) 45° phase-shift tuning range, and (d) S-parameter simulation results.

Figure 2.16: (a) LO network output saturated power for different LO center frequencies, and (b) LO network output power at different harmonics (6th, 4th, 3rd, 2nd, 1st) of the bondwired 20.83 GHz LO input.
was developed. Rogers RO4350B dielectric with a relative permittivity of 3.66 F/m and a loss tangent of 0.0037 at 10 GHz was used to provide low-loss PCB routings for bondwired LO and clock input signals. The measurement setup and lab photo are shown in Figs. 2.18a and 2.18b, respectively. CW or modulated output of a 65-GSa/s, 25-GHz BW AWG (Keysight M8195A) was fed to a WR-8.0 upconversion mixer (SAGE SFB-08-E2) on the transmit side. A WR-8.0 frequency extender (Keysight E8257DV08) provided the LO signal for upconversion mixer.

To measure the frequency response of the receiver, RF output of the WR-8.0 upconversion mixer was directly connected to the RF GSG probe with a WR-8.0 waveguide section, and the 8PSK test output of the receiver was measured with a differential GSGSG probe. For conversion gain measurement, the RF input tone power and the CTLE control code were set to -42 dBm and code 7, respectively, and the LO network was saturated. Conversion gain of
the receiver is shown in Fig. 2.19. A maximum conversion gain of 32 dB was measured across
the RF bandwidth. For NF measurement, the RF input of the receiver was terminated to 50
Ω, CTLE was set to code 7, and the LO network was again saturated. The receiver achieves
a minimum DSB NF value of 10.3 dB and remains less than 14 dB at baseband frequencies
up to 12 GHz (Fig. 2.20a). The measured worst-case input-referred 1-dB compression point
(IP1dB) of the receiver is -31.4 dBm (Fig. 2.20b). Under modulation, the measured power
spectrum of a 36 Gbps modulated signal at the 8PSK test output of the receiver is shown
in Fig. 2.21.

For wireless measurements, WR-8.0 horn antennas with 25 dBi gain were used at WR-

Figure 2.18: (a) Wireless measurement setup schematic, and (b) photo.
Figure 2.19: Measured receiver conversion gain at the 8PSK test output.

Figure 2.20: (a) Measured DSB NF and (b) IP1dB at the 8PSK test output.

Figure 2.21: 36 Gbps power spectrum at the 8PSK test output.

8.0 upconversion mixer output and the receiver input to transmit and receive the 8PSK-modulated wireless data. An 80-GSa/s 33-GHz bandwidth real-time oscilloscope (Keysight DSAV334A) was used to measure the 8PSK eye-diagram at the 8PSK test output after downconversion and before direct-demodulation. External references of the signal generators
for receiver and transmitter LOs were both synchronized with the AWG external reference. A known non-random 8PSK pattern (repetition of $22.5^\circ$, $67.5^\circ$, ..., $337.5^\circ$) was received. Transient waveform at the 8PSK test output after direct-downconversion was observed on the real-time oscilloscope and phase reference of the receiver LO was tuned externally until a 4-level symmetrical eye-diagram with ratios of $+0.92/+0.38/-0.38/-0.92$ was achieved (one-time calibration in a fixed point-to-point test setup). Figs. 2.22a and 2.22b show the wirelessly measured 30 Gbps and 36 Gbps 8PSK eye-diagrams at a maximum distance of 0.3 m (limited by the measurement setup). The 8PSK constellation was reconstructed from two
Figure 2.24: Wirelessly measured eye-diagrams of demodulated 8PSK 3-bit streams for (a) 30 Gbps and (b) 36 Gbps overall data-rates.

8PSK baseband data downconverted by two 90° phase-shifted LOs. Figs. 2.23a and 2.23b show the wirelessly measured 30 Gbps and 36 Gbps 8PSK constellations.

For direct-demodulation, the reference clock of AWG was externally synchronized with the reference clock of the signal generator that feeds the symbol-rate clock to the receiver. Furthermore, AWG timing delay was tuned for optimum sampling point to minimize BER. The measured BPSK eye-diagrams of the 8PSK-demodulated 3-bit streams for 30 Gbps and 36 Gbps total data-rates are shown in Fig. 2.24. A BER of 1e-6 for 36-Gbps PRBS-7 sequence was wirelessly measured at a 0.3 m distance. The measured receiver sensitivity at this BER is -41.28 dBm. Variation of BER for different values of received input power is shown in
Fig. 2.25, which follows 8PSK waterfall profile. The performance of the proposed 8PSK receiver is compared with state-of-the-art direct-demodulation receivers in Table 2.3. This work achieves the highest speed and lowest BER with excellent energy efficiency among all previously reported high data-rate direct-demodulation receivers to date.

### 2.6 Conclusion

In this paper, a novel RF-correlation-based direct-demodulation 8PSK receiver was disclosed. The proposed receiver is the highest speed direct-demodulation receiver with the highest modulation-order to date. The proposed RF-to-bits receiver architecture obviates the need for power-hungry high-speed-resolution ADCs and significantly improves the energy efficiency of the receiver compared to other high-speed receivers in an RF-to-bit scenario.
System analysis in terms of the bit error probability of the multi-phase RF-correlation idea was detailed and high-frequency circuit design techniques and analysis of critical building blocks were presented. Wireless measurement results were conducted and excellent sensitivity and BER were achieved for the reported data-rate verifying the effectiveness of the direct-demodulation idea. This direct-demodulation idea can readily be extended to realize higher-order 16QAM-Star demodulation by only adding one level of envelope detection.
Bibliography


