# On the Design of Highly Linear Efficient Millimeter-Wave Power Amplifiers and Ultra-Wideband Low-Error Phase Shifters for Emerging Wireless Communication Applications 

by<br>Alireza Asoodeh<br>\section*{A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF<br><br>DOCTOR OF PHILOSOPHY<br><br>in}

The Faculty of Graduate and Postdoctoral Studies
(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
July 2021
(C) Alireza Asoodeh 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

On the design of highly linear efficient millimeter-wave power amplifiers and ultra-wideband low-error phase shifters for emerging wireless communication applications
submitted by Alireza Asoodeh in partial fulfillment of the requirements for the degree of Doctor of Philosophy
in Electrical and Computer Engineering

## Examining Committee:

Shahriar Mirabbasi, Electrical and Computer Engineering, UBC

## Supervisor

Thomas Johnson, School of Engineering, UBCO
Supervisory Committee Member

Juri Jatskevich, Electrical and Computer Engineering, UBC
Supervisory Committee Member

Jeff Young, Physics and Astronomy, UBC
University Examiner

Peyman Servati, Electrical and Computer Engineering, UBC
University Examiner

## Additional Supervisory Committee Members:

Alireza Nojeh, Electrical and Computer Engineering, UBC
Supervisory Committee Member

## Abstract

With the unprecedented growth in the number of connected devices and demand for larger amount of data as well as higher data rate, there is a need to improve our communication systems to address such demands. Addressing these demands have resulted in devising a roadmap for transition from current $4^{\text {th }}$ generation of communication systems (4G) to the $5^{\text {th }}$ generation (5G) and beyond 5 G in the near future.

To address the demand for higher data rates, broadening the frequency spectrum of 5 G and beyond 5 G to millimeter wave bands is being actively pursued. To overcome propagation losses, reduce interferers and improve signal-to-noise ratio, phased-array systems have attracted a lot of attention especially for applications operating in mm-wave bands. In this work, we mainly focus on two important building blocks of phased-array systems, namely, power amplifiers (PAs) and phase shifters.

To fulfill the stringent linearity and efficiency requirements of the 5 G and beyond 5 G systems, a linear and efficient PA is required. We present several design techniques for implementing highly linear and efficient CMOS PAs. The proposed techniques include strategic placement of varactors, a multifunction coplanar-waveguide (CPW)-like power combining structure, and a systematic design approach for the passive networks. A proof-of-concept prototype that operates in the 28 GHz band is designed and fabricated in a $65-\mathrm{nm}$ bulk CMOS process. The design achieves a $\mathrm{P}_{\text {sat }}$ of 23.2 dBm , output $\mathrm{P}_{1 d B}$ of 22.7 dBm , and power-added efficiently (PAE) of $35.5 \%$.

Next, a continuous-mode $360^{\circ} \mathrm{mm}$-wave ultra-wideband phase shifter over the frequency range of 10 GHz to 50 GHz is presented. A proof-ofconcept prototype is also designed and fabricated in a $65-\mathrm{nm}$ bulk CMOS process. To implement such an ultra-wideband phase shifter, design approaches for several key building blocks including balun and quadrature allpass filter (QAF) are proposed. These sub-blocks are separately analyzed. To confirm the validity of the proposed techniques, proof-of-concept prototypes have been designed and tested. Particularly, the continuous-mode $360^{\circ}$ phase shifter prototype achieves $\sim 0.2-\mathrm{dB}$ root-mean square (RMS) amplitude and $<1.4^{\circ}$ phase error over the frequency range from 10 to 50 GHz .

## Lay Summary

The research presented in this thesis focuses on the design and implementation of two main building blocks of 5G phased array systems, namely, power amplifiers (PAs) and phase shifters. 5G systems require both highly linear PAs to minimize the overall distortion and efficient PAs to prolong the battery life, which calls for improving linearity and efficiency of the current state-of-the-art designs. In this context, we have proposed several design techniques to enhance both linearity and efficiency of PA structures. We have also proposed and have experimentally validated the operation of a continuous-mode $360^{\circ}$ phase shifter structure. The key advantage of the proposed phase shifter is that while it operates at high frequencies and over a broad bandwidth, its amplitude and phase errors are low which result in a phased array system with improved overall reliability and performance.

## Preface

All the content presented in this dissertation are original, independent work conducted in System-on-chip (SoC) lab by the author, Alireza Asoodeh. The proposed techniques described in Chapter 2, 3, 4 and 5 were designed preliminary by me. Canadian Microelectronic Corporatioon (CMC Microsystems) provided access to the computer-aided design (CAD) tools as well as access to the technology and chip manufacturing. We designed two different chips and had them fabricated in a $65-\mathrm{nm}$ CMOS process. The research was my own work and I performed the research and wrote the associated manuscripts under the supervision of Professor Shahriar Mirabbasi.

A version of Chapter 2 has been published in the following journal paper:
A. Asoodeh, H. M. Lavasani, M. Cai and S. Mirabbasi, "A Highly Linear and Efficient $28-\mathrm{GHz}$ PA With a $\mathrm{P}_{\text {sat }}$ of $23.2 \mathrm{dBm}, P_{1 d B}$ of 22.7 dBm , and PAE of $35.5 \%$ in $65-\mathrm{nm}$ Bulk CMOS," IEEE Journal of Solid-State Circuits, vol. 56, no. 3, pp. 914-927, March 2021.

I was the leading investigator in developing the concept, designing the circuit and testing it. My colleague, M. Cai, who has recently completed his PhD assisted me with the chip measurements. Dr. H. M. Lavasani gave valuable technical advice. Professor S. Mirabbasi was the research supervisor and provided feedback on the work and assisted with preparing the manuscript.

A version of Chapter 3 has been published in the following journal paper:
A. Asoodeh and S. Mirabbasi, "On the Design of $\mathrm{n}^{\text {th }}$-Order Polyphase All-Pass Filters," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 1, pp. 133-146, Jan. 2019.

I was the leading researcher in developing the concept, designing the circuit and performing the experiments. Professor S. Mirabbasi was the research supervisor and provided feedback on the work and assisted with preparing the manuscript.

A version of Chapter 4 has been submitted to a journal by: A. Asoodeh, M. Cai and S. Mirabbasi. The paper is entitled "Analysis and Design of a Wideband mm-Wave Miniaturized Marchand Balun with a Better than 0.4dB Amplitude and $1.5^{\circ}$ Phase Mismatch," and is 14 double-column pages.

I was the leading researcher in developing the concept, designing the circuit and performing the measurements. Dr. M. Cai prepared the measurement set-up. Professor S. Mirabbasi was the research supervisor for this work and helped with composing the manuscript.

A version of Chapter 5 is being prepared to be submitted to a journal by: A. Asoodeh and S. Mirabbasi. The paper is entitled"A Continuous-Mode $360^{\circ} \mathrm{mm}$-Wave Ultra-Wideband Phase Shifter with $\sim 0.2-\mathrm{dB}$ RMS Amplitude and $<1.4^{\circ}$ Phase Error for Next-Generation Wireless Communication Systems," and is 10 double-column pages.

I am the leading investigator in developing the concept, designing the circuit and testing it. Professor S. Mirabbasi is the research supervisor and has provided feedback on the work and assisted with preparing the manuscript.

## Table of Contents

Abstract ..... iii
Lay Summary ..... iv
Preface ..... V
Table of Contents ..... vii
List of Tables ..... xi
List of Figures ..... xii
List of Abbreviations ..... Xviii
Acknowledgements ..... XX
Dedication ..... xxi
1 Introduction and Overview ..... 1
1.1 Mobile Communication History ..... 1
1.2 Background ..... 2
1.2.1 Building Blocks of Phased-Array Systems ..... (4)
1.3 Thesis Outline and Summary of Research Objectives and Contributions ..... 7
2 A Highly Linear and Efficient 28-GHz PA in 65-nm Bulk CMOS ..... 9
2.1 Architecture of the Proposed Power Amplifier ..... 11
2.2 Linearity and Efficiency Enhancements ..... 13
2.2.1 Linearity improvement ..... 13
2.2.2 Efficiency improvement ..... 20
2.3 Implementation of Passive Circuits ..... 26
2.3.1 Band-pass passive matching networks ..... 26
2.4 Reliability and Stability of the PA ..... 33
2.4.1 Voltage distribution across the stacked device of the output stage ..... 33
2.4.2 Stability of the amplifier ..... 34
2.5 Experimental Results ..... 35
2.6 Summary ..... 43
3 On the Design of $n^{t h}$-Order Polyphase All-Pass Filters ..... 44
3.1 The Proposed $n^{t h}$-Order Polyphase All-Pass Filter ..... 46
3.1.1 Introduction to The Structure: ..... 46
3.1.2 Analysis of the Proposed Structure: ..... 48
3.2 Analysis and Design of $2^{n d}, 3^{r d}$, and $4^{t h}$-Order PAFs ..... 52
3.2.1 $\quad 2^{n d}$-order Polyphase All-Pass Filters: ..... 52
3.2.2 $\quad 3^{\text {rd }}$-Order Polyphase All-Pass Filters: ..... 55
3.2.3 $4^{t h}$ - order Polyphase All-Pass filters: ..... 59
3.3 Design Example ..... 63
3.4 Effects of Nonidealities on the Circuit Performance ..... 65
3.4.1 Loading Effects ..... 65
3.4.2 Effects of Component Value Deviation ..... 67
3.4.3 Influence of the Limited Quality Factor of inductors on the QM and Input Impedance of the Proposed Struc- ture ..... 70
3.5 Experimental Results ..... 74
3.6 Summary ..... 75
4 Analysis and Design of a Wideband mm-Wave Miniaturized Marchand Balun with a Better than 0.4-dB Amplitude and $1.5^{\circ}$ Phase Mismatch ..... 79
4.1 Bandwidth Extension of Miniaturized Marchand Baluns ..... 80
4.1.1 Brief Discussion on the Capacitively Loaded Marc- hand Baluns ..... 80
4.1.2 Capability to Improve the Bandwidth of the Capaci- tively Loaded Balun ..... 83
4.2 An Alternative Methodology for Calculating the Components' Values of Capacitively Loaded Baluns ..... 93
4.2.1 The proposed procedure ..... 93
4.2.2 Generalization of the approach to multi-stage balun design with a broader overall bandwidth ..... 94
4.3 Calculating the Component values of the Balun Using the Proposed Approach ..... 95
4.4 Experimental Results ..... 99
4.5 Summary ..... 103
5 A Continuous-Mode $360^{\circ} \mathrm{mm}$-Wave Ultra-Wideband Phase Shifter ..... 104
5.1 On the Selection of the Quadrature All-Pass Filter (QAF) Structure ..... 106
5.1.1 Loading Effect ..... 106
5.1.2 Study of Effect of Resistor Variations on the Output Performance ..... 108
5.1.3 Effect of the Reactive Component Value Deviations ..... 110
5.1.4 Study of the Effect of the Limited Quality Factor of Inductors on Quadrature Outputs ..... 111
5.2 The Complete Structure of the Proposed Vector-Sum Phase Shifter ..... 114
5.2.1 QAF Selection and VGA Structure ..... 114
5.2.2 Baseband Circuit Structure ..... 116
5.2.3 Output Buffer ..... 119
5.3 Design and Layout Considerations ..... 119
5.4 Experimental Results ..... 123
5.5 Summary ..... 128
6 Conclusion and Future Works ..... 129
6.1 Conclusion ..... 129
6.2 Future Works ..... 130
6.2.1 Performance Improvement ..... 131
6.2.2 System-Level Integration ..... 131
Bibliography ..... 132
Appendices
A Solution of $10^{\text {th }}$-Order Polynomial Equation ..... 144
B Calculation of the Notch Frequencies of the Insertion Loss ..... 146
C Evaluation of the Capacitively Loaded Balun ..... 149
D Evaluation of Equation (4.6) ..... 151

E Evaluation of Equation (4.7) . . . . . . . . . . . . . . . . . . . 154

## List of Tables

2.1 Performance comparison of state-of-the-art 5G PAs. ..... 42
3.1 Component values based on the structure of Fig. 3.8 ..... 76
4.1 Performance comparison of state-of-the-art structures ..... 102
5.1 Performance comparison of state-of-the-art phase shifters. . . 127

## List of Figures

1.1 The different generations of mobile communication ..... [2
1.25 G mm -wave phased-array transceiver architecture ..... [3]
1.3 Block diagram of a phased array receiver with equivalent input-referred noise of each path ..... (3)
1.4 Basic block diagram of switched network phase shifter. ..... 5
1.5 Basic concept of vector modulator circuits. (b) Four vectors spaced $90^{\circ}$ apart, that are typically used for the realization of vector modulator circuits. (c) Three vectors ( $120^{\circ}$ apart) that can also be used for the realization of vector modulator circuits. ..... 5
2.1 The complete structure of the proposed PA: (a) the architec- ture, (b) the output stage with an adaptive feedback network enhancing the linearity of the output stage through improv- ing both AM-AM and AM-PM conversions, (c) the proposed filter and CPW-like combiner. ..... 11
2.2 A simplified model of a MOS transistor. ..... 14
2.3 Conceptual schematics of (a) the conventional cascode, and (b) proposed stacked architecture. ..... 14
2.4 Linearity improvement technique: (a) complete schematics (b) AM-PM and, (c) AM-AM reduction mechanisms. ..... 16
2.5 Varactor capacitance variation versus the gate amplitude of the envelop detector for different biases of the envelop detec- tor ( $V_{\text {BE.en }}$ ). ..... 17
2.6 Simulation results comparing the designs with and without linearity improvement technique: (a) gate capacitance vari- ation percentage of the power cell, (b) $M_{c}$ current increase compared to the case that no varactor exists, (c) drain to gate voltage ratio of $M_{g m 1}$ in Fig. 2.4 ..... 18
2.7 The architecture of the output stage of the PA. ..... 19
2.8 Frequency response of the envelope detector, (a) amplitude,(b) phase response.21
2.9 Simplified equivalent circuit of the proposed filter for source inductance cancellation ..... 22
2.10 Simplified equivalent circuit of the network used for the power combining stage. ..... 22
2.11 4-way CPW-like output power combiner and its equivalent circuit. ..... 24
2.12 The normalized $Q_{L}$ of a simple L-constant CPW versus the distance, $S$. ..... 25
2.13 Load impedance matching network using $n$-tuned topology. ..... 26
$2.14 Q_{L} B W_{\text {nor }}$ product versus return loss for different tuned match- ing networks. ..... 27
2.15 Input matching network: (a) detailed structure, (b) final structure including the splitter. ..... 27
2.16 Interstage matching network: (a) detailed structure, (b) sim- plified network and its input reflection coefficient (c) Norton transformation step, (d) final structure. ..... 31
2.17 Equivalent circuit model of the output stages including the power combiner and filter. ..... 32
2.18 Output combiner efficiency with and without shield lines. ..... 32
2.19 Voltage swing of the drain-gate terminal of the stacked FET. ..... 33
2.20 Simulated stability factors at different corners, temperatures and power levels (Full-Power $=23 \mathrm{dBm}$; Small-Power $=-7 \mathrm{dBm}$ ). 3
2.21 Output spectrum for (a) no input signal, (b) $P_{\text {out }} \approx 23 \mathrm{dBm}$. 3 ..... 35
2.22 Die micrograph of the proposed PA. ..... 36
2.23 Measured and simulated $S$-parameters results.
2.24 Measured large-signal behavior of the PA showing $P_{\text {out }}$, gain, and $P A E$ at 28 GHz . ..... 37
2.25 Measured $P_{o-s a t}, P_{o 1 d B}, P A E_{\max }$ and $P A E_{1 d B}$ versus frequency. 38
2.26 Measured $P_{o-s a t}$, and $P A E_{\max }$ for 20 samples at 28 GHz , (a)$P_{\text {o-sat }},(\mathrm{b}) P A E_{\max }$.38
2.27 Linearity results with, without the varactor, and with varac- tor at the bottom FET gate: (a) AM-AM and (b) AM-PM. ..... 39
2.28 Measured AM-PM Results: (a) AM-PM versus frequency, (b) AM-PM for 20 samples at 28 GHz . ..... 40
2.29 Modulation measurement results $\left(f_{\text {carrier }}=28 \mathrm{GHz}\right)$. ..... 40
2.30 Measured EVM, and ACPR for 20 samples for data rate of 2.5 GS/s, (a) EVM, (b) ACPR. ..... 40
2.31 Measured EVM and ACPR versus output power for data rate of $2.5 \mathrm{GS} / \mathrm{s}(\mathrm{PAPR}=8.3 \mathrm{~dB})$. ..... 41
3.1 Key building block of the proposed system. ..... 45
3.2 The proposed all-pass filter structure capable of generating in-phase and quadrature signals. ..... 46
3.3 Examples of $4^{\text {th }}$-order $\mathcal{N}_{i}$ network. (a) Proper type, and (b) improper type ..... 48
$3.4 \quad 2^{\text {nd }}$-order poly-phase all-pass filter (PAF) ..... 52
3.5 Typical phase difference graph of the $2^{\text {nd }}$-order PAF ..... 54
$3.6 \quad 3^{\text {rd }}$-order poly-phase all-pass filter (PAF). ..... 55
3.7 Typical phase difference graph of the $3^{r d}$-order PAF ..... 56
$3.84^{\text {th }}$-order poly-phase all-pass filter (PAF). ..... 59
3.9 Example of the typical phase difference graph of the proposed$4^{\text {th }}$-order PAF.61
3.10 The quadrature phase difference graph of different order PAFs ..... 64
3.11 The quadrature mismatch in terms of parasitic capacitance. ..... 66$3.122^{\text {th }}$-order PAF errors in terms of $\mathrm{x}\left(=\frac{\Delta R}{R}\right)$, which x is a mea-sure showing how much Eq. (3.1) deteriorates. a) Quadra-ture mismatch (dB). b) Input impedance variation (100 $\times$$\left.\frac{Z_{\text {in,eq}}-Z_{\text {in,ideal }}}{Z_{\text {in,ideal }}}\right)$69
$3.133^{\text {rd }}$-order PAF errors in terms of $\mathrm{x}\left(=\frac{\Delta R}{R}\right)$. a) Quadrature mismatch (dB). b) Input impedance variation $\left(100 \times \frac{Z_{\text {in,eq }}-Z_{\text {in,ideal }}}{Z_{i n, \text { ideal }}}\right) 70$
$3.144^{\text {th }}$-order PAF errors in terms of $\mathrm{x}\left(=\frac{\Delta R}{R}\right)$. a) Quadrature mis$\operatorname{match}(\mathrm{dB})$. b) Input impedance variation $\left(100 \times \frac{Z_{\text {in,eq }}-Z_{\text {in,ideal }}}{Z_{\text {in,ideal }}}\right) 7$
$3.152^{\text {th }}$-order PAF errors in terms of Q , which Q is the quality factor of on-chip inductors. a) Quadrature mismatch (dB). b) Input impedance variation $\left(100 \times \frac{Z_{i n, e q}-Z_{i n, \text { ideal }}}{Z_{i n, \text { ideal }}}\right) \ldots$.73
$3.163^{\text {rd }}$-order PAF errors in terms of Q. a) Quadrature mismatch (dB). b) Input impedance variation ( $\left.100 \times \frac{Z_{i n, \text { eq }}-Z_{\text {in, ideal }}}{Z_{\text {in, ideal }}}\right) \cdot$
$3.174^{\text {th }}$-order PAF errors in terms of Q. a) Quadrature mismatch (dB). b) Input impedance variation $\left(100 \times \frac{Z_{\text {in,eq}}-Z_{\text {in,ideal }}}{Z_{\text {in,ideal }}}\right) \cdots 75$
3.18 The photograph of $\mathrm{I} / \mathrm{Q}$ generator network ..... 76
3.19 Input return loss of the I/Q network ..... 77
$3.20 I$ - and $Q$ - magnitudes. ..... 77
3.21 Quadrature error characteristics of the I/Q network. (a) I/Qphase difference. (b) I/Q amplitude mismatch.78
4.1 Balun structure, (a) conventional Marchand balun, (b) capac- itive loading technique to reduce the line length, (c) reduced- size Marchand balun ..... 81
4.2 Normalized notch frequency of the capacitively loaded balun versus the electrical length of the coupled lines (Eq. (4.3)). ..... 82
4.3 The typical frequency response of a band-pass filter: (a) $\left|S_{11}\right|$ (lumped structure), (b) $\ln \left|S_{11}\right|^{-1}$ (lumped one), (c) $\left|S_{11}\right|$ (distributed structure), (d) $\ln \left|S_{11}\right|^{-1}$ (distributed one) . . . ..... 84
4.4 The approach to the calculation of the input admittance of the capacitively loaded balun structure: (a) complete schematic assuming that $C_{1}=C_{2}=0$ and $C_{4}=\infty,(\mathrm{b})$ odd-mode stim- ulation, (c) even-mode stimulation. ..... 86
$4.5 \quad R_{0} C_{3} \omega_{l}$ in terms of $x$ for the different values of $k$. ..... 89
4.6 Normalized values of (4.6) and (4.7) in terms of the differential output impedance. ..... 91
4.7 The ratio of the lower and upper 3 dB frequencies of different baluns for two different coupling factors $(k)$ : (a) $k=0.5$, (b) $k=\frac{\sqrt{2}}{2}$. ..... 92
4.8 The proposed flowchart to find the components values. ..... 94
4.9 The balun structure based on multi-stage topology ..... 95
4.10 The performance comparison of different multi-stag baluns as well as conventional Marchand one with the source and differential load impedances of $50 \Omega$ and $200 \Omega$, respectively. ..... 96
4.11 The coupled line implemented in $65-\mathrm{nm}$ bulk CMOS, (a) its schematic, (b) its coupling factor, $k$. ..... 97
4.12 The capacitively loaded balun, (a) lumped version, (b) dis- tributed variant $\left(\mathrm{R}_{\mathrm{s}}=2 \mathrm{R}_{\mathrm{L}}=50 \Omega\right)$ ..... 97
4.13 Frequency response of the different structure. ..... 98
4.14 The microphotograph of the chip ..... 99
4.15 The simulated and measured insertion loss. ..... 100
4.16 The simulated and measured return loss ..... 100
4.17 The simulated and measured amplitude mismatch. ..... 101
4.18 The simulated and measured phase mismatch ..... 101
5.1 Vector-sum phase shifter building block. ..... 105
5.2 The $4^{t h}$-order quadrature all-pass filters (QAFs): (a) asym- metrical topology, (b) symmetrical one. ..... 107
5.3 The I/Q performance of two topologies versus $C_{L}$ assuming$C_{L_{1}}=C_{L_{2}}$ at three different frequencies: (a) amplitude mis-match, (b) phase mismatch.109
5.4 The I/Q performance of two topologies versus resistance vari- ation at three different frequencies: (a) amplitude mismatch, (b) phase mismatch. ..... 110
5.5 The I/Q performance of two topologies versus capacitance variation at three different frequencies: (a) amplitude mis- match, (b) phase mismatch. ..... 112
5.6 The I/Q performance of two topologies versus Q at three dif- ferent frequencies: (a) amplitude mismatch, (b) phase mis- match. ..... 113
5.7 Simplified block diagram of the proposed vector-sum active phase shifter. ..... 114
5.8 Proposed approach to convert the unbalanced topology of Fig. 5.2(a) to the balanced one. ..... 115
5.9 VGA structure: (a) conceptual architecture, (b) current-mode adder implementation (transistor-level), (c) complete structure. 116
5.10 Bias circuitry of the VGAs. ..... 117
5.11 Relation between the input voltage of the bias circuit, $\mathrm{VB}_{\mathrm{in}}$, and control voltage of the phase shifter, $\mathrm{V}_{\mathrm{CTRL}}$, at different quadrants. ..... 118
5.12 The control circuitry that generates $\mathrm{VB}_{\text {in }}$ and the control signal of switches $S W_{1-4}$ from $\mathrm{V}_{\mathrm{CTRL}}$. ..... 118
5.13 Opamp structure employed in the control circuit. ..... 118
5.14 Output buffer structure. ..... 119
5.15 Amplifier topology: (a) single-ended, (b) differential structure. 1205.16 QAF layout with the VGAs blocks located on both sides ofthe layout: (a) type I, (b) type II.122
5.17 QAF performance of two different layouts shown in Fig. 5.16:
(a) quadrature amplitude ratio, (b) quadrature phase difference. 122
5.18 Die micrograph of the proposed phase shifter. ..... 123
5.19 Output performance of the phase shifter: (a) S parameter results, (b) phase shift at 16 different states. ..... 124
5.20 Quadrature characteristics of the QAF measured at the out- put: (a) I/Q amplitude ratio. (b) I/Q phase difference. ..... 124
5.21 Output performance of the phase sifter versus frequency: (a)RMS amplitude error. (b) RMS phase error.125
5.22 (a) RMS amplitude and (b) phase error for 20 samples mea-sured at 30 GHz .125
5.23 Output phase shift versus the input control $\mathrm{V}_{\mathrm{CTRL}}$ at 30 GHz .126
5.24 Large signal performance at reference state $\left(\mathrm{V}_{\text {CTRL }}=0\right)$ : (a) measured gain versus input power at 30 GHz : (a), (b) $\mathrm{P}_{1 \mathrm{~dB}}$ for 20 samples at 30 GHz .126
B. 1 The approach to the calculation of the input impedance of the capacitively loaded balun structure: (a) complete schematic with the different capacitors connected to the input $\left(C_{3}\right)$, O.C. $\left(C_{4}\right)$ and outputs $\left(C_{5}\right)$, (b) odd-mode stimulation, (c) even-mode stimulation. . . . . . . . . . . . . . . . . . . . . . . 147
D. 1 The selection of contour, (a) The Contour with the branch cut due to the logarithmic function, (b) a simpler contour by moving the branch point into the left-half plane.152
E. 1 The contour considered for the revised quantity. . . . . . . . . 157
xvii

## List of Abbreviations

| 1G | First Generation |
| :---: | :---: |
| 2G | Second Generation |
| 3G | Third Generation |
| 4G | Fourth Generation |
| 5G | Fifth Generation |
| ACPR | Adjacent Channel Power Ratio |
| AM-AM | Amplitude Modulation to Amplitude Modulation |
| AM-PM | Amplitude Modulation to Phase Modulation |
| AMPS | Advanced Mobile Phone System |
| BW | Bandwidth |
| CDMA | Code Devision Multiple Access |
| CMOS | Complementary Metal Oxide Semiconductor |
| CPW | Coplanar Waveguide |
| D-AMPS | Digital AMPS |
| EMBB | Enhanced Mobile Broadband Communication |
| EVM | Error Vector Magnitude |
| GSM | Global System for Mobile communication |
| HSPA | High-Speed Packet Access |
| IC | Integrated Circuit |
| IL | Insertion Loss |

MEMS ........ Micro Electro-Mechanical System
MIMO ......... . Multiple Input Multiple Output
MMTC ........ Massive Machine Type Communication
NF .............. Noise Figure
NMT .......... Nordic Mobile Telephony
$P_{1 d B} \ldots \ldots \ldots \ldots$.......... Compression point
PA .............. Power Amplifier
PAE ............ Power Added Efficiency
PAF ............ Polyphase All-Pass Filter
$P_{a v}$............. Average Power
$P_{D C} \ldots \ldots . . .$. . Personal Digital Cellular
$P_{\text {sat }} \ldots \ldots \ldots .$. ........ Saturated Power
QAF ............ Quadrature All-Pass Filter
QAM .......... Quadrature Amplitude Modulation
RL .............. Return Loss
RMS ............ Root Mean Square
TACS .......... Total Access Communication System
UE ............... User Equipment
URLLC ....... Ultra-Reliable and Low Latency Communication
VGA ............ Variable Gain Amplifier
WCDMA ..... Wideband Code Division Multiple Access

## Acknowledgements

I would like to express my deepest gratitude to Professor Shahriar Mirabbasi for his continuous self-giving support, generous encouragement and professional academic guidance during the course of this research. I am also indebted to him for his support, patience and friendship.

My deep gratitude also goes to my colleagues and friends in System-on-a-Chip research group with whom I have shared an enjoyable experience of learning in a very friendly environment. I would also like to thank Steve Hall and Balvinder Virdee from Keysight Technologies for providing access to measurement equipment, and Roozbeh Mehrabadi and Dr. Roberto Rosales from the Department of Electrical and Computer Engineering at the University of British Columbia for providing CAD tools and technical support.

My wife, Shiva, has always been a source of encouragement and drive behind my achievements. I can hardly find words to utter my appreciation for her love, understanding, support, and devotion. I hope that we share a happy and memorable life in the years to come.

## Dedication

To my wife Shive, son Liam

## Chapter 1

## Introduction and Overview

### 1.1 Mobile Communication History

As shown in Fig. 1.1, four generations of mobile communications have been established over the past 40 years. Emerging around 1980, the first generation (1G) was based on analog data transmission. The systems based on this generation were limited to voice services and made mobile telephony accessible to ordinary people for the first time. In the early 1990s, the second generation (2G) of mobile communication was introduced which was based on digital transmission on the radio link. Although the target service was still voice, the system was able to provide some limited data services to the users. Initially, several different second-generation technologies were developed including GSM (Global System for Mobile communication), D-Advanced Mobile Phone System (D-AMPS), PDC (Personal Digital Cellular), and CodeDivision Multiple Access (CDMA)-based technologies. However, with the passage of time, GSM technology which was initiated in Europe, penetrated from Europe to other parts of the world and became a dominant approach among the second-generation technologies. Due to the success of GSM, 2G systems were adopted by many countries and mobile communication systems became popular for the majority of the world's population. Even today there are many places in the world where GSM is dominant despite the introduction of third and fourth-generation of communication systems (3G and 4G) and their deployment across the globe. In early $2000 \mathrm{~s}, 3 \mathrm{G}$ was introduced. By developing 3G, a main stepping-stone to high-quality mobile broadband, fast wireless Internet access was enabled by the 3G evolution known as HSPA (High-Speed Packet Access) [1]. The 4G was established by deployment of the LTE (Long Term Evolution) technology [2]. LTE improves the performance of HSPA and provides a higher efficiency and a more enhanced mobile-broadband experience in terms of higher end-user data rates. This is provided by means of orthogonal frequency division multiplexing (OFDM)based transmission, which enables wider transmission bandwidths and more advanced multi-antenna technologies.

We are currently in the transition from 4G to fifth generation (5G) sys-
tems of mobile communications. Initial discussions on 5G were initiated around 2012. The term 5 G in many discussions refers to the specific new 5G radio-access technology. However, 5G is also often used in a much wider context, not just referring to a specific radio-access technology but rather to a wide range of new services envisioned to be enabled by emerging mobile communication. The main objective of the 5 G is to enable larger data volumes and further enhanced user experience, for example, by supporting higher data rates with low latency and high reliability. To increase the data rates, broadening the operating frequency spectrum of 5 G to millimeterwave ( mm -wave) bands has become inevitable.


Figure 1.1: The different generations of mobile communication.

### 1.2 Background

Based on the recent research, an optimal technique to counter heavy propagation losses at mm-wave frequencies [3], [4] in the emerging wireless communication systems ( 5 G and beyond 5 G ) is the use of antenna arrays (phasedarray systems). It is expected that such antenna arrays would be integrated into both base stations and user equipment (UE). Moreover, non-line-ofsight (NLOS) 28 GHz coverage is demonstrated in urban cells [10]. Fig. 1.2 shows a typical block diagram of phased-array transceiver. The power received by receiver can be calculated from the Friis' equation [5].

Phased-array systems bring several advantages to the wireless system. Starting from the transmitting side of the phased array system (Fig. 1.2),


Figure 1.2: 5G mm-wave phased-array transceiver architecture.
the signals transmitted by the antennas can constructively add up in certain direction(s) in space through spatial power combining. In comparison to a single-antenna transmitter, transmitting an output power of $P_{0}$, each path of the phased-array transmitter can transmit an output power of $P_{0} / N$ and keep the sum of the output power equal to $P_{0}$. Therefore, the equivalent isotropic radiated power (EIRP) of the phased-array transmitter will be $P_{0} \times N$, which is increased by $10 \times \log N \mathrm{~dB}$ as compared to the single-antenna transmitter. Moreover, the output power of a phased array transmitter can be controlled by simply turning on or off a certain number of transmitters. On the receiver side, the signals received by multiple antennas can be added


Figure 1.3: Block diagram of a phased array receiver with equivalent inputreferred noise of each path.
up constructively. On the other hand, as shown in Fig. 1.3, the noise of
different receiver paths, dominated by the contributions of separate antennas and the low noise amplifiers after the antennas ( $N_{11}, N_{12}, \ldots, N_{1 N}$ ), are generally uncorrelated. As a result, if $N$ antennas are used in the receiver, the output signal-to-noise ratio and thus the sensitivity of the receiver can be improved by $10 \times \log N \mathrm{~dB}$. Furthermore, a phased array system can place nulls in undesired direction(s), which improves channel multipath profile and reduces interference to/from other systems. Consequently, a phased array system facilitates achieving a higher system capacity, larger range and better interference suppression, which are highly beneficial properties for a mm -wave wireless system. In the remainder of this section, we will focus on two main building blocks of a typical phased-array transceiver.

### 1.2.1 Building Blocks of Phased-Array Systems

## Phase shifting blocks

Phase shifters, serving as electronic beam steering elements, are the key building blocks of any phased array systems [6], [7]. A phase shifter can provide a continuously variable phase shift (analog), or a discrete set of phase states (digital). Compared to a continuously variable phase shifter, a discrete-step phase shifter has phase quantization errors. The advantage of using a discrete-step phase shifter is that it can be digitally controlled, which allows for a simple control and better immunity to noise on the control lines. Phase shifters can be either passive or active each of which is briefly described below.

Passive phase shifters: Semiconductor or micro-electro-mechanicalsystem (MEMS) based passive phase shifters can be categorized into two main groups.

- Transmission type: Although there are various types of phase shifters falling into this category, the most popular one is switched network type. A basic block diagram of switched network phase shifter is shown in Fig. 1.4. When the input signal, originally passing through network 1 , is switched to pass through network 2 , we get a differential phase shift ( $\phi_{2}-\phi_{1}$ ), where $\phi_{1}$ and $\phi_{2}$ are the input-output phase difference of network 1 and 2 , respectively. The most commonly used networks in switched network phase shifters are the low-pass and high-pass filter configuration.
- Reflection type: In reflection type phase shifters, the basic design unit is a one-port network, and it is the phase shift of the reflected signal
that is changed by the control line. These basic one-port phase shifters can be converted into two-port components either by using a circulator or a hybrid. The reflection type phase shifter are also able to generate a continuous $360^{\circ}$ phase shift.


Figure 1.4: Basic block diagram of switched network phase shifter.


Figure 1.5: Basic concept of vector modulator circuits. (b) Four vectors spaced $90^{\circ}$ apart, that are typically used for the realization of vector modulator circuits. (c) Three vectors ( $120^{\circ}$ apart) that can also be used for the realization of vector modulator circuits.

Active phase shifters: While passive phase shifters are realized by transistors in the passive mode (switch), use of transistors in the active
mode makes several phase shifter designs possible. Although there are several classes of active phase shifters, the most popular one is phase shifting using vector modulators. A vector modulator is a circuit that is capable of independently varying the amplitude and the phase of an input signal by desired amounts. The concept is shown schematically in Fig. 1.5. In principle, $\phi$ could be anywhere between $0^{\circ}$ and $360^{\circ}$, and the magnitude A could be any feasible positive value. Two main schemes have been proposed for implementing the vector modulator circuit. One of them uses four vectors spaced $90^{\circ}$ apart and pointing in four different directions as shown in Fig. 1.5(b). Amplitudes of the four component vectors $A_{1}, A_{2}, A_{3}$, and $A_{4}$ are controlled independently by four different gain-controlled amplifiers. Only two of these four components are nonzero at any time. Thus, the vector sum could be in any of the four quadrants and the amplitude could be adjusted by controlling the gains of two amplifiers active at the given time. The second scheme, instead of four component vectors uses three vectors spaced $120^{\circ}$ apart, as shown in Fig. 1.5(c). This scheme provides an alternative implementation for vector modulator circuits. In general, the vector type phase shifter has the ability to change the output phase in a continuous manner.

Active versus passive phase shifters: Active phase shifters as compared to their passive counterparts can achieve much finer resolution over a broader bandwidth. Moreover, contrary to passive phase shifters, active ones do not suffer from the high insertion loss, which can adversely affect the noise figure (NF) of the system. When integrated, they also generally occupy less die area as compared to passive ones. On the other hand, active phase shifters have poorer linearity [8], [58], [62] and [66].

## Power Amplifiers (PAs)

Similar to any wireless system, 5G systems also require power amplifiers so that the signal level can reach an acceptable one before being delivered to antenna. PAs are usually categorized based on their efficiency and linearity. While the former measure is indicative of the amount of energy drawn from the power supply for a specific output power, the latter provides an indication of the level of adjacent channel power as a result of spectral regrowth and amplitude compression. Since the PA in the user equipment (UE) devices obtains its energy from the battery, a high power efficiency for low-cost UE devices is an important factor. Moreover, for high data-rate communication, the use of higher-order complex modulation schemes, e.g., quadrature amplitude modulation (QAM), with high peak-to-average power
ratio (PAPR) and large RF bandwidth, e.g., $\geq 100 \mathrm{MHz}$, is being envisioned. These modulation schemes require a highly linear power amplifier. Furthermore, they have a large PAPR and their sensitivity to distortion forces the PA to operate in 8 to 10 dB power back-off, which can drastically lower the efficiency of the PA. From this perspective, the GaAs PAs would be favorable as they do not suffer from problems such as substrate conductivity or low breakdown voltage as compared to CMOS technologies.

On the other hand, due to having many building blocks, low cost and high level of integration are the key points for phased array systems. Therefore, in the context of phased array systems, CMOS PAs are the preferred choice over more power-efficient PAs that are implemented in other technologies such as GaAs, due to the high level of integration, lower cost and higher yield offered by CMOS.

Thus, design and implementation of a highly linear and efficient CMOS PA for 5G UEs have been a popular topic of research in both academia and industry. Since, in a high-volume production setting, cost and complexity preclude the use of calibration, e.g., using digital pre-distortion, due to differing nonlinear behavior among PAs in an integrated array, it is desired that the PAs have inherent circuit-level linearity.

### 1.3 Thesis Outline and Summary of Research Objectives and Contributions

The focus of this research is on the above-mentioned main building blocks of phased-array systems, namely, the PA and the phase shifter, in particular, those that are intended for 5G (or beyond 5G) applications. The objectives and contributions of this work are summarized as follows:

- A highly linear and efficient $28-\mathrm{GHz}$ power amplifier (PA) suitable for 5G phased-array systems is presented in Chapter 2, A proof-ofconcept prototype is designed, fabricated and successfully tested in a $65-\mathrm{nm}$ bulk CMOS process. It achieves the saturated output power $\left(P_{\text {sat }}\right)$ of over 23 dBm at 28 GHz and AM-PM distortion of less than $1.9^{\circ}$ over the frequency band of 26 GHz to 31 GHz with a peak poweradded efficiency (PAE) of $35.5 \%$. Several techniques including strategic placement of capacitors and varactors are used to improve the efficiency and linearity of the design leading to a negligible difference between $P_{1 d B}$ and $P_{s a t}$. The contributions here include several design techniques that facilitate achieving high linearity and high efficiency.
- Chapter 3 is focused on the $n^{\text {th }}$-order quadrature-based poly-phase allpass filters (PAFs). This block is the key building block of the active phase shifters. It generates the in-phase and quadrature (I/Q) signals required for the proper operation of active phase shifters. Since it plays a crucial role in the overall performance, the I/Q generator should produce the quadrature signals with a low error in both quadrature gain and phase over the operation bandwidth. Thus, a network with the ability to operate over a broad bandwidth while having a low error is proposed and analyzed.
- An ultra-wideband miniaturized passive balun (balanced-unbalanced converter) is presented in Chapter 4. As the objective is to eventually design and implement a millimeter-wave phase shifter, the optimal approach is to distribute the sensitive high frequency signals in a balanced manner to reduce the noises coupled to the signal lines, interference effect, higher order, particularly even-order, harmonics. Furthermore, as explained in Chapter 5, the QAF itself intrinsically has a differential structure, it needs to have a balun at the input. The proposed balun converting the single-ended impedance of $50 \Omega$ to the differential one of $50 \Omega$ exhibits the amplitude and phase mismatch of less than 0.4 dB and $1.5^{\circ}$ over the frequency band of 10 GHz to 50 GHz , while it occupies a core area of $180 \mu \mathrm{~m} \times 130 \mu \mathrm{~m}$.
- Finally, in Chapter 5, a fully continuous-mode $360^{\circ}$ ultra-wideband mm-wave active phase shifter that operates over a wide frequency range from 10 GHz to 50 GHz is presented. The quadrature all-pass filter (QAF), which is based on the quadrature PAF described in Chapter 3, and balun proposed in Chapter 4 are used in the proposed phase shifting block. The design techniques presented here facilitate achieving low amplitude and phase errors with a power consumption that compares favorably with that of state-of-the-art designs.


## Chapter 2

## A Highly Linear and Efficient $28-\mathrm{GHz}$ Power Amplifier (PA) in $65-\mathrm{nm}$ Bulk CMOS

To address the ever-growing demand for broadband cellular data communication, emerging wireless systems, including $5^{t h}$-generation (5G) devices, are moving towards using mm-wave frequency bands to facilitate $\mathrm{Gb} / \mathrm{s}$ data rates 9$]-[10]$. Several mm-wave frequency bands, including spectra around $28 \mathrm{GHz}, 37 \mathrm{GHz}$, and 39 GHz have been allocated by the Federal Communications Commission (FCC) for use in 5G radio [11]. Furthermore, to achieve high data rates, the use of higher-order complex-modulation schemes, e.g., quadrature amplitude modulation (QAM), with high peak-to-average power ratio (PAPR) and wide bandwidths are being envisioned. Owing to the high PAPR, 5G power amplifiers (PAs) are required to have a high linearity in both amplitude and phase over a wide power range to ensure a satisfactory quality of service and user experience.

PAs are considered to be one of the most critical building blocks in wireless transceivers whose performance usually dictates the linearity and energy efficiency of the overall wireless transmitter [12]-[13]. In fact, a mmwave PA with high linearity and high efficiency is essential for 5G massive multi-input multi-output (MIMO) and phased-array systems. CMOS-based PAs are the preferred choice over the more power-efficient PAs that are implemented in high performance technologies such as GaAs, due to the higher level of integration, lower cost and higher yield of CMOS circuits. Nevertheless, CMOS PAs typically suffer from low output power ( $P_{\text {out }}$ ), poor linearity, and reduced power added efficiency ( $P A E$ ). Several researchers have reported on improving the intrinsic linearity of PAs by using techniques to reduce their AM-PM distortion [14]- [15]. However, these approaches deliver medium-level output power with limited PAE improvement.

In this chapter, a $28-\mathrm{GHz}$ highly linear and efficient CMOS PA is presented whose complete structure is shown in Fig. 2.1. To enhance both
linearity and efficiency of the system, several techniques are proposed. A varactor-based feedback technique is used to improve the linearity of the PA. It dynamically changes the impedance seen by the $g_{m}$ cell of the output stage to linearize the input capacitance of the $g_{m}$ cell. The feedback network also facilitates driving more current to the output as output power increases. Consequently, unlike conventional techniques where the varactor network helps to reduce the AM-PM distortion [16] or the ones that use separate mechanisms for AM-PM and AM-AM conversion reduction [17], the proposed approach improves the linearity through reducing both AM-PM and AM-AM conversions using the same varactor network. Furthermore, contrary to the conventional stacked structures in which the feedback loop operates at the carrier rate, the proposed topology runs the feedback loop at the envelope rate. Since the envelope frequencies are much lower than the carrier, the design constrains in the proposed technique are less critical. Moreover, compared to the existing PAs using the adaptive biasing scheme which is typically focused on the feedforward compensation to reduce AMAM distortion, the proposed feedback linearization technique enhances the linearity through reduction of both AM-AM and AM-PM distortions. To improve the efficiency several techniques are proposed. The first approach resonates out the inductive portion of the source interconnect parasitics by adding a capacitor (and creating an $L C$ filter). It can also help to partially cancel out the poly gate resistance. The second technique, which is layoutbased, is to twist and combine the layouts of the proposed filter and output power combiner to make a coplanar-waveguide (CPW) like combiner. This is done without using the additional ground planes. In fact, one of the components of the filter, that acts as a DC path for the power stage and also helps to improve the stability of the whole system at lower frequencies, can also serve as the GND plane for the power combiner (Fig. 2.1(c)). Finally, to ensure that all passive matching networks including input, interstage, and output ones are designed with the minimum number of passive components, a systematic approach based on [18] is employed. The aim is to achieve the frequency response with an acceptable flatness in the frequency band of interest while the overall system is optimized in terms of efficiency and area.

The organization of the chapter is as follows. Section 2.1 describes the architecture of the proposed two-stage PA. Linearity and efficiency improvement techniques are described in Section 2.2 . Section 2.3 presents the proposed systematic approach for designing the passive matching networks to achieve the optimum performance over the frequency band of interest. Section 2.4 discusses the reliability and stability of the overall structure. Measurement results and performance comparison with state-of-the-art PAs are
presented in Section 2.5.


Figure 2.1: The complete structure of the proposed PA: (a) the architecture, (b) the output stage with an adaptive feedback network enhancing the linearity of the output stage through improving both AM-AM and AM-PM conversions, (c) the proposed filter and CPW-like combiner.

### 2.1 Architecture of the Proposed Power Amplifier

Fig. 2.1(a) shows the architecture of the proposed PA. The input is first split into two in-phase signals, each of which is fed into its own driver amplifier (DA) whose structure is a single $g_{m}$ cell with inductive degeneration. Then, the output of each DA is further split into two in-phase signals before it is fed to four PA cells. The PA cell outputs are then combined by a 4 -way
microstrip-based power combiner.
An inductively degenerated $g_{m}$ cell $\lfloor 24]$ is preferred over the cascode topology for the DA since it simplifies the design of the interstage matching network. In comparison with a cascode topology, an inductively degenerated structure results in a lower input to output impedance ratio for the interstage network, making the design of a low-loss interstage network over the given bandwidth more convenient. Furthermore, inductive degeneration facilitates impedance matching to an input source with real impedance if desired. Finally, including an inductor in the source of the transistor improves both the linearity and stability of the structure.

The stacked configuration [21] is chosen for the output stage (power cell). To improve the linearity, an adaptive feedback network is proposed to dynamically change the impedance seen by the $g_{m}$ cell of the power stage based on the output power level. The feedback loop is comprised of a single varactor along with an envelope detector circuit. The envelope detector consists of a single-transistor common-drain buffer with a capacitor at its output (Fig. 2.1(b)). Furthermore, the proposed technique reduces the voltage stress on the gate-drain terminal of the stacked transistor compared to the conventional cascode structure where the gate of the cascode device is typically an AC ground. The structure of the output stage is shown in Fig. 2.1(c).

The input and interstage splitters act as both power splitters and impedance matching networks, and help to maximize the power delivered to the next stage. The interstage network converts the input impedance of the output cell to the optimum load required for the driver stage so that the maximum power can be delivered to the power cell in the frequency range of interest. To match the input impedance of the driver cell to $50 \Omega$ (which is used for this standalone PA chip), another matching network is used. To have an optimum matching network with the minimum number of reactive components for input, interstage, and output passive networks, Lopez approach [18]-[20] is employed, which is a revised and improved version of Bode-Fano technique [22]-[23]. The objective is to use a (simple) passive structure that meets both the bandwidth and optimum power transfer requirements. In this context, a systematic approach is chosen to make sure no redundant passive components (that could degrade the overall performance) are used. The output power combiner converts the $50 \Omega$ impedance of the antenna to the required load at the drain of the output stage which results in optimum power delivery and efficiency. This load is obtained at full power to take the large-signal output impedance of the transistors into account. To ensure a unified ground path with the same potential across the chip, a large
ground plane is implemented outside the active area (the shaded area in Fig. 2.1(a)).

### 2.2 Linearity and Efficiency Enhancements

### 2.2.1 Linearity improvement

Two dominant sources of nonlinearity in any linear PA are AM-AM and AMPM conversion. While the former is typically dominated by the nonlinearity of the transconductance, i.e., $g_{m}$, of the active device, the key contributors to the latter are the intrinsic nonlinear capacitor of the device [25]. Conventionally, to improve the overall linearity through AM-PM compensation, if the input transistor of the power stage is an NMOS transistor, a PMOSCAP is placed in parallel with its gate-source capacitance, so that the phase shift caused by input capacitance changes can be kept approximately constant. The bias voltage of this varactor can be fixed [16] or dynamically controlled for the optimum performance [27]. However, this technique can only alleviate the AM-PM distortion and AM-AM distortion still needs to be improved. While AM-PM distortion causes rotation in the signal constellation, AMAM distortion compresses the signal constellation. Such compression effect can be problematic for higher-order modulations, e.g., 256 QAM. Another technique that helps to improve the overall linearity is to put a degeneration inductance in the source of the power transistor [28]. This inductance, by creating a negative feedback, reduces the distortion and thus enhances the linearity. The main drawback of this technique is its negative impact on the output swing and efficiency of the PA.

In this work, a technique is developed that enhances the overall linearity of the PA by reducing both AM-AM and AM-PM distortions with minimal impact on the output power and efficiency. Considering the simplified transistor model shown in (Fig. 2.2), the equivalent Miller capacitance seen at the gate due to gate-drain capacitance is:

$$
\begin{equation*}
C_{m}=\left(1+g_{m} R_{L}\right) \times C_{g d} \tag{2.1}
\end{equation*}
$$

where, $R_{L}$ is the equivalent load resistance.
Now, if a mechanism is introduced that $C_{g s}$ and $C_{m}$ show complementary behaviors as a function of the input voltage, the equivalent input capacitance seen at the gate of the power cell (seen by the driver stage) will be approximately constant. The approach can also enhance the efficiency by alleviating the trade-off between the linearity and efficiency since it would provide an


Figure 2.2: A simplified model of a MOS transistor.
option to use more efficient classes such as class B or C while keeping the linearity at an acceptable level. For instance, biasing the output transistor at a low $V_{G S}$ of around $V_{t h}$ or even lower, significantly reduces the bias current of the power cell, helping to further improve the efficiency as compared to the topologies biased in class AB. The effective gate-source capacitance for the given bias is more voltage-dependent and increases as the gate voltage amplitude increases [24] and [26]. So, if $C_{m}$ decreases when the amplitude of the input signal increases, then the overall equivalent capacitance seen at the input of the power cell can be kept approximately constant. One technique for achieving this is to dynamically control the value of resistor $R_{L}$ (see Eq. (2.1)). For example, to compensate for the increase in $C_{g s}$ for the transistor biased around $V_{t h}$, the value of $R_{L}$ should decrease as the signal amplitude increases.


Figure 2.3: Conceptual schematics of (a) the conventional cascode, and (b) proposed stacked architecture.

## Implementation of a simple variable load

The conventional cascode structure is shown in Fig. 2.3(a). The impedance seen by the $g_{m}$ device $M_{g m}$ is approximately equal to $\left(g_{m-M_{c}}+j \omega C_{g s_{-}}\right)^{-1}$, where $g_{m-M_{c}}$ and $C_{g s_{-M_{c}}}$ are the transconductance and gate-source capacitance of the transistor $M_{c}$, respectively. If the operating frequency is much
smaller than the transit frequency, $f_{T}$, of the device, the impedance can be simplified to $\left(g_{m-M_{c}}\right)^{-1}$. If the impedance presented by $M_{c}$ to $M_{g m}$ can change dynamically with the signal amplitude, the Miller capacitance $C_{m}$ can also vary, compensating for the variations of the gate-source capacitance. Shown in Fig. 2.3(b), is a stacked structure with an impedance $Z$ attached to the gate of the transistor $M_{c}$. The revised impedance seen by $M_{g m}$ is:

$$
\begin{equation*}
\frac{1+Z(j \omega) \times j \omega C_{g s_{-M_{c}}}}{g_{m-M_{c}}+j \omega C_{g s_{-M_{c}}}} \approx \frac{1+Z(j \omega) \times j \omega C_{g s_{-M_{c}}}}{g_{m-M_{c}}} \tag{2.2}
\end{equation*}
$$

If $Z(j \omega)$ is replaced with a varactor (as shown in Fig. 2.3(b)), the impedance seen by the $g_{m}$ cell can be written as:

$$
\begin{equation*}
\frac{1+\frac{C_{g s}-M_{c}}{C_{v}(v)}}{g_{m-M_{c}}} \tag{2.3}
\end{equation*}
$$

where $C_{v}(v)$ is the voltage-dependent capacitance of the varactor. This equation can provide an impedance whose value decreases as the control voltage increases. The control voltage can be implemented using an envelope detector that detects the amplitude of the output signal. Accordingly, the control voltage of the varactor varies with the change in the output amplitude. The proposed structure is shown in Fig. 2.4(a). In this circuit, as the output signal amplitude increases, the control voltage of the varactor increases, which leads to an increase in the capacitance of the varactor. Thus, the impedance seen by the $g_{m}$ device, given by (2.3), decreases. To minimize the nonlinearity effect of varactor, varactors are connected in a back-to-back manner. The detector consists of a simple common-drain structure with a capacitance connected to its output. A thick-oxide device is chosen for the envelop detector to minimize the chance of breakdown.

To further improve the performance of the proposed linearity enhancement technique, an inductor is added at node X to cancel out the capacitance at this node, which mainly originates from the drain-bulk capacitance of the $g_{m}$ device and source-bulk capacitance of the stacked device. Due to adding this inductance, the impedance seen by the $g_{m}$ cell would be mostly resistive over the frequency band of interest. Unlike the structure presented in [28] which uses a series inductor, the power gain and output swing are largely preserved in the proposed design which employs a parallel inductor.

To maximize the impedance modulation expressed in (2.3), the gatesource capacitance of the stacked transistor should be closer to the maximum capacitance of the varactor, $C_{\max }$. Note that due to body effect, the


Figure 2.4: Linearity improvement technique: (a) complete schematics (b) AM-PM and, (c) AM-AM reduction mechanisms.
gate-source capacitance of the stacked transistor, $C_{g s_{-M_{c}}}$, does not usually exhibit the same behavior as that of the $g_{m}$ device, and its variations is less than that of the gate-source capacitance of the $g_{m}$ device. Moreover, the varactor helps with minimizing the variations in $C_{g s_{-} M_{c}}$ as it limits the voltage drop across the capacitance. With these design considerations in mind, as it will be shown later, the overall capacitance seen at the gate of the $g_{m}$ cell can be kept approximately constant despite the variations in the amplitude of the input signal of the $g_{m}$ cell. The technique can also help with reducing the AM-AM distortion as graphically shown in Fig. 2.4(c). Based on (2.3) and as shown in Fig. 2.4b, the effective impedance seen by the $g_{m}$ device decreases at higher output levels due to the impedance modulation directing a larger portion of the current generated by $M_{g m}$ to the output. Consequently, the $P_{1 d B}$ of the PA will be improved leading to lower AM-AM distortion. Fig. 2.5 shows the capacitance variations of the varactor versus the amplitude of the signal at the gate of the envelope detector for different bias voltages, $V_{B E}$. As can be seen from the figure, $V_{B E}$, adjusts the voltage where the varactor capacitance starts to change


Figure 2.5: Varactor capacitance variation versus the gate amplitude of the envelop detector for different biases of the envelop detector ( $\mathrm{V}_{\text {BE-en }}$ ).
(shifts the capacitance versus gate amplitude voltage curve) and plays a key role in the behavior of the capacitance seen at the input of the power cell or by the driver stage. The performance of the circuit can be further enhanced by inserting a variable gain amplifier after the envelope detector (Fig. 2.4(a)). It causes the slope of the capacitance curve in Fig. 2.5 to become variable, which adds more flexibility to the overall structure. In this work, our focus is on improving the performance with minimum added circuitry to the proposed system shown in Fig. 2.4(a) and thus a variable gain amplifier is not used. Fig. 2.6 shows the simulation results, comparing some of the features of the proposed structure with those of the traditional stacked structure. As Fig. 2.6(a) shows, the gate capacitance of the proposed structure stays approximately constant. In contrast, the traditional stacked structure shows more than $30 \%$ variation in the input capacitance versus the gate voltage amplitude of the main transistor $M_{g m 1}$ for the same bias and transistor sizes. In the proposed structure, the input capacitance remains approximately constant for a wide range of the gate voltage amplitude.

The current shunted into the stacked transistor $M_{c 1}$ can increase up to around $40 \%$ using the proposed technique as compared to the case without a varactor (Fig. 2.6(b)). Therefore, the gain compression at higher gate voltages is reduced, and consequently the AM-AM distortion is improved. The drain-to-gate voltage ratio of the transistor $M_{m g 1}$ is shown in Fig. 2.6(c). As can be seen from Fig. 2.6(c), in the proposed technique the voltage gain of


Figure 2.6: Simulation results comparing the designs with and without linearity improvement technique: (a) gate capacitance variation percentage of the power cell, (b) $M_{c}$ current increase compared to the case that no varactor exists, (c) drain to gate voltage ratio of $M_{g m 1}$ in Fig. 2.4.
the main transistor drops faster than that of the conventional stacked topology and thus the effective Miller capacitance drops in a way that the sum of the gate capacitance and the Miller capacitance stays almost constant over the wide range of input voltage amplitude variations.


Figure 2.7: The architecture of the output stage of the PA.

## A brief discussion on feedback loop dynamics

As shown in Fig. 2.7, the input signal of the output stage is fed forward to the output through path $A$ (stacked structure) and the envelope of the output signal is fed back to pin $c$ through path $B$ (envelope detector). It should be noted that due to the use of the envelope detector, the frequency contents of the input and output of the feedback network are different. Therefore, to gain a better qualitative understanding of the behavior of the feedback system the following approach is considered.

Since the varactor capacitance is modulated by the envelope signal, the overall PA, and in particular, its output stage, should ideally remain stable for all capacitance values of the varactor. Thus, the stability of the system is evaluated over different process corners, supply voltages, and temperatures (PVT variations) as well as different input (output) power levels. These results are presented in Section V. To achieve the optimal performance from the proposed technique, the impedance seen by the $g_{m}$ transistor has to be minimum (maximum) when the signal at its input is maximum (minimum). To satisfy this condition, the envelope signal at pin $c$ (Fig. 2.7) should have a minimum phase delay with respect to the envelope of the signal at pin $a$. Furthermore, to avoid unnecessary distortions, the group delay or $\tau_{g}$, namely, the rate of change of the total phase shift with respect to frequency, of the envelope signal at pin $c$ should have little or no variations (ideally
be constant) over its bandwidth. As a result, it is desirable to keep the delay of the signal in the most critical path, from pin $b$ to $c$, at minimum (ideally zero) over the frequency band of the envelope signal. Therefore, in this work, the feedback path is implemented with a simple envelope detector without using any additional amplification stage. This approach would also facilitate a wider bandwidth for the envelope detector. Nonetheless, one can include an amplification stage to potentially improve the performance as pointed out in the previous subsection, at the cost of including another block, namely equalizer, to equalize for the delay and/or amplitude in the feedback path. Fig. 2.8 shows the frequency response of the envelope detector. It confirms that the envelope signal is received by the varactor with a negligible magnitude/phase variation over the frequency band of $\sim 1 \mathrm{GHz}$.

### 2.2.2 Efficiency improvement

This section highlights the proposed efficiency improvement methods. A filtering technique that cancels the source parasitics of the transistors is proposed. The objective is to manipulate the structure to create a local reference at the source of the output transistors since any parasitics at this node adversely affects both the output power and efficiency. Moreover, the layout floor plan is modified to create a CPW-like power combiner without requiring any additional ground trace. In the following, the proposed techniques are described.

## Source-Inductance cancellation network

The parasitic inductance at the source of the $g_{m}$ transistor has a profound impact on the performance of the PA since it limits both the power gain and output swing of the common-source sub-PA. A common approach to neutralize the inductance is to add a series capacitance. However, adding the capacitance to the source blocks the DC component of the device current. To bias the transistor, an inductor can be placed in parallel with the series combination. The added inductor will interact negligibly with the series LC path over the frequency band of interest since the series LC creates a lowimpedance path to GND over the frequency band. The simplified equivalent circuit of this arrangement is shown in Fig. 2.9, where $L_{p}$ creates a DC path for the power cells to GND and $C_{s}$ is added to cancel the source inductance $L_{s}$. At the frequency band of interest, a low impedance path is created by the series resonance, $f_{s}$, of source interconnect and $C_{s}$. Inductance $L_{p}$, in addition to the DC path creation for the $g_{m}$ devices, plays two other roles.


Figure 2.8: Frequency response of the envelope detector, (a) amplitude, (b) phase response.

First, it helps to improve the stability of the structure through its parallel resonance whose value is:

$$
\begin{equation*}
f_{p}=\frac{1}{2 \pi \sqrt{C_{s}\left(L_{s}+L_{p}\right)}} . \tag{2.4}
\end{equation*}
$$

Since $f_{p}$ is always smaller than $f_{s}$, the lower corner of the frequency band can be determined by $f_{p}$. Consequently, any unwanted low-frequency oscillation is dampened. Moreover, it acts as a GND trace for the output power
combiner described in the next subsection to improve the passive efficiency of the combiner. More specifically, the layout of the proposed arrangement together with that of the power combiner can be co-designed to have a smaller size and enhance the overall performance.


Figure 2.9: Simplified equivalent circuit of the proposed filter for source inductance cancellation.


Figure 2.10: Simplified equivalent circuit of the network used for the power combining stage.

## Power combiner structure

Fig. 2.10 shows a simplified lumped circuit model of the structure used for combining the output power of four sub-PAs. $C_{o}$ and $L_{M}$ convert the resistance $R_{L}$ to the optimum impedance seen by each PA. While $C_{o}$ controls the real part of the optimum load, the imaginary part of the optimum load
is controlled by $L_{M} . C_{o}$ and $L_{M}$ are:

$$
\begin{gather*}
C_{o}=\frac{1}{R_{L} \omega_{o}} \sqrt{\frac{4 R_{L}}{R_{o p t}}-1}  \tag{2.5}\\
L_{M}=\frac{X_{o p t}}{4 \omega_{o}}+\frac{C_{o} R_{L}^{2}}{1+C_{o}^{2} R_{L}^{2} \omega_{o}^{2}}, \tag{2.6}
\end{gather*}
$$

where $\omega_{o}, R_{o p t}$ and $X_{o p t}$ are the center frequency, and the real and imaginary parts of the optimum load impedance seen by each sub-PA, respectively. It should be noted that, with the same layout strategy presented here, the structure can accommodate a higher order filter if a higher bandwidth is desired. For example, one can add capacitors at nodes $a, o$, and $a^{\prime}$ (the gray capacitors shown in Fig. 2.10, which are not included in this work). To save area and increase the quality factor of the capacitor $C_{o}$, namely, $Q_{C}$, the capacitor, which is implemented by three top metal layers and optimized using electromagnetic (EM) simulations, is located underneath the output pad. To increase the quality factor of inductor $L_{M}$, namely, $Q_{L}$, a coplanar waveguide (CPW) like structure is used. Compared to microstrip line inductors on silicon substrate, CPWs have a considerably higher $Q_{L}$ due to the close proximity of the ground plane to the signal line [29]. Fig. [2.11 shows the proposed power combiner. The interconnect lines are divided into three segments. The trace carrying the highest current, trace $T_{3}$, is implemented based on CPW (to have the maximum $Q$ over the frequency band of interest). To calculate the lengths of the traces, the following approach is used.

One of the important tasks of the output passive stage design is determining the distance $S$ in Fig. 2.11 since the layouts of the power combiner and the filter are intertwined. EM simulations for a CPW are performed to estimate the inductive $Q, Q_{L}$, of the line as a function of $S$. To ensure that the total impedance seen by each sub-PA does not change for any distance $S$, the length of $T_{3}$ is adjusted with $S$. Given that the length of power combiner line is electrically small (i.e., much smaller than the wavelength) and the bandwidth is not very broad, the power combiner is modeled with a series inductor, and the $Q_{L}$ around this inductance becomes part of the optimum load seen by each sub-PA. Note that the parallel capacitance is typically small compared to the capacitor at the output of the power combiner and can be either ignored or incorporated into the output capacitance (Fig. 2.11). Fig. 2.12 shows the $Q_{L}$ normalized to that of a simple line without any GND traces, i.e., $S=\infty$. Two different initial line lengths of $100 \mu \mathrm{~m}$ and $200 \mu \mathrm{~m}$ are considered. When $S$ is changed the length of the
line is also changed correspondingly so that its inductance stays constant. As can be seen from Fig. [2.12, there is an optimum value for $Q_{L}$ which for both cases occurs at almost the same value of $S$. Once the optimum $S$ is estimated, the lengths of three different segments $T_{1}, T_{2}$ and $T_{3}$ can be found. Denoting $L_{T_{1}}, L_{T_{2}}$ and $L_{T_{3}}$ as the inductance of segments $T_{1}, T_{2}$ and $T_{3}$, respectively, the equivalent inductance of the combiner (Fig. 2.10) can be written as: $L_{M}=L_{T_{3}}+0.5 L_{T_{2}}+0.25 L_{T_{1}}$. Since $L_{M}$ is a weaker function of $L_{T_{1}}$, the length of $T_{1}$ (and consequently, $L_{T_{1}}$ ) is minimized. Therefore, the length of $T_{2}$ covers most of $S$. Finally, the length of $T_{3}$ can be calculated after calculating the equivalent inductance of traces $T_{1}$ and $T_{2}$. Once the sizes are determined, further EM simulations are performed on the power combiner structure that is shielded with the GND traces while the distance $S$ is swept in the vicinity of the optimum value obtained above. The proposed approach provides a good initial estimate for the parameters of the power combiner.


Figure 2.11: 4-way CPW-like output power combiner and its equivalent circuit.


Figure 2.12: The normalized $Q_{L}$ of a simple L-constant CPW versus the distance, $S$.

## Co-designing the layout of the power combiner and source inductance cancellation network

The process begins with placing two lines (slot lines in Fig. 2.11) on each side of the power combiner at the distance of $S$ from the high current segment $T_{L 3}$ obtained from the previous subsection. To create a low-impedance path from the source of power cell (points $C$ and $C^{\prime}$ in Fig. 2.11) to GND, two capacitors are placed in series with the slot lines in the lower half of line $D D^{\prime}$ (Fig. 2.11). Their values can be obtained once the length of the source interconnect is found. The length is constrained by the layout of the overall structure, e.g., the location of RF pads, active cells, etc. The other end of capacitors are connected to GND pads. Since these capacitors cancel out the interconnect inductance, nodes $C$ and $C^{\prime}$ remain almost at the same potential as GND pads over the frequency band of interest. The sources of $g_{m}$ devices are connected to nodes $C$ and $C^{\prime}$. The slot lines located in the upper half of line $D D^{\prime}$ act as the inductance $L_{p}$ (Fig. 2.9), GND lines for high current trace of the output power combiner, $T_{L 3}$, and DC current path for the output stage transistors. $L_{p}$ is calculated using (2.4) and based on its value its length is estimated. This inductance determines the lower corner of the frequency band and dampens the unwanted low frequency oscillations. If the length is close to $T_{L 3}$ that is obtained in the previous subsection, the (near)
optimum performance can be expected. If not, another EM simulation is required to recalculate the circuit parameters, e.g., the length of the source interconnect, $S$, etc. In this design, $C_{s}$ and $L_{s}$ are $\sim 250 \mathrm{fF}$ and $\sim 130 \mathrm{pH}$, respectively. A wide ground plane (Fig. 2.1) consisting of two metal layers is implemented outside the active area to connect input and output GND pads with minimum parasitics. This keeps both input and output GND pads at (approximately) the same potential. Consequently, the trace representing $L_{p}$ can also be considered as the ground shield for the power combiner in the frequency band since one end of this trace is connected to GND and the other end connected to the source, nodes $C$ and $C^{\prime}$, remain near GND. In fact, without adding any extra GND traces for power combiner and only by combining these layouts, the high current trace $T_{3}$ is shielded. To further enhance the efficiency, the effective impedance at the source can be designed to be slightly capacitive, and thus the equivalent impedance seen at the gate will have a negative resistance which partially cancels out the gate resistance. Depending on the choice of the technology which affects the device performance, the device parasitics such as gate resistance can degrade the PA performance metrics, e.g, PAE in mm-wave PAs.


Figure 2.13: Load impedance matching network using $n$-tuned topology.

### 2.3 Implementation of Passive Circuits

### 2.3.1 Band-pass passive matching networks

Fig. 2.13 shows a passive $L C$ structure which converts any complex load to the real one. For the given configuration, it can be shown that the following relationship exists among the quality factor of the load, $Q_{L}$, the normalized bandwidth $\left(B W_{n o r}\right)$ and the maximum permissible reflection magnitude,


Figure 2.14: $Q_{L} B W_{\text {nor }}$ product versus return loss for different tuned matching networks.


Figure 2.15: Input matching network: (a) detailed structure, (b) final structure including the splitter.
$\Gamma_{\max }$, of the band-pass matching network [18].

$$
\begin{equation*}
Q_{L} \times B W_{n o r}=\frac{1}{b_{n} \sinh \left(\frac{1}{a_{n}} \ln \left(\frac{1}{\Gamma_{\max }}\right)\right)+\frac{1-b_{n}}{a_{n}} \ln \left(\frac{1}{\Gamma_{\max }}\right)}, \tag{2.7}
\end{equation*}
$$

where $B W_{\text {nor }}$ is:

$$
\begin{equation*}
B W_{\text {nor }}=\frac{f_{\max }-f_{\min }}{\sqrt{f_{\min } \times f_{\max }}} . \tag{2.8}
\end{equation*}
$$

The matching network in Fig. 2.13 is a lossless $n$-stage network with contiguous branches that alternate between series and parallel combination of $L \mathrm{~s}$ and $C \mathrm{~s} . n$ is the number of series or parallel stages of the matching network. The coefficients $a_{n}$ and $b_{n}$ are constants whose values can be found in [18.

## Input matching network implementation

To design the input matching network with the minimum number of passive components, it should be noted that any degeneration inductor (inductor in the source of the driver transistor) will degrade the quality factor of the impedance seen at the input of the driver stage, namely, $Q_{L}$. As can be seen from Fig. 2.14, for given $B W_{n o r}$ and $\Gamma_{\max }$, Eq. (2.7) can be satisfied with a smaller $n$, i.e., lower number of tuned stages, for lower $Q_{L}$. For example, for bandwidth of 5 GHz (from 26 GHz to 31 GHz ), with $Q_{L}$ of $2, \Gamma_{\max }$ can be higher than 15 dB even if a single-tuned topology is used. In this work, $Q_{L}$ is chosen to be 3 so as to provide some additional passive gain to compensate for the passive network losses. Consequently, $\Gamma_{\max }$ (i.e., the maximum return loss) of better than 12 dB can be attained over the frequency band of interest. Next step is to convert the perceived resistance to $50 \Omega$. In Fig. 2.13 this is achieved by an ideal transformer. Here, a simple L-type matching network is used. Fig. 2.15 (a) shows the resulting structure. The final matching network for driving two DAs is shown in Fig. 2.15(b). $L_{i 1}$ and $L_{i 2}$ are incorporated into the splitter.

## Interstage matching network implementation

Contrary to the input or output matching networks, complex-to-complex impedance transformation is generally performed by an interstage matching network. Therefore, the technique presented in [18] is modified to develop an approach to find the minimum number of required segments for a given $\Gamma$. This is because the technique presented in [18] only works for matching a complex impedance to a real impedance or vice versa. Thus, as shown in Fig. 2.16(a), a combination of two cascaded segments, namely, an ideal transformer and a reactive absorptive network, are used. The left subnetwork converts the complex impedance $Z_{S}$ to real impedance $R_{m 1}$ and the right sub-network converts the complex impedance $Z_{L}$ to the real impedance $R_{m 2}$. The turn ratios of ideal transformers are chosen such that $R_{m 1}=$ $R_{m 2}$. Then, the procedure described in $[18]$ can be followed for each side separately.

The quality factors of the impedances seen at the driver side and the input of sub-PAs, are denoted as $Q_{S}$ and $Q_{L}$, respectively. Both of these quality factors (or one of them depending on their ratio) play a prominent role in determining the bandwidth of the interstage network. This is also shown graphically in Fig. 2.16 (a). In this figure, $Z_{S}$ is the complex conjugate of the driver load impedance that is required for the optimum power transfer
to the succeeding stage and $Z_{L}$ is the impedance seen at the input of the power stage. Using (2.7), the maximum attainable return loss over the frequency band of 26 to 31 GHz for the single- and double-tuned topologies for $Q_{L}$ of 10 are $\sim 5$ and $\sim 10 \mathrm{~dB}$, respectively. Since in this work, $\frac{Q_{L}}{Q_{S}}$ is much higher than 1 , the bandwidth is limited by the right sub-network in Fig. 2.16(a). Consequently, a double-tuned or higher-order interstage matching network topology is needed for the right sub-network shown in Fig. 2.16(a), while a single-tuned network would be sufficient for the left subnetwork. This is due to the fact that over the frequency band of interest, the maximum attainable return loss of a single-tuned topology used for the left sub-network $\left(Q_{S}\right)$ is still much higher than that of a double-tuned network used for the right sub-network $\left(Q_{L}\right)$.

The systematic approach presented thus far can be used for finding the sub-network(s) with the minimum number of components for a given return loss. It can also provide us with some insights into the design of a network whose input and output impedances are complex which is more difficult to achieve than the case where one side of the network is terminated to a $50 \Omega$ (purely real impedance). In what follows, some simplifications including a circuit technique, namely, Norton transformation [30]- [31] are discussed. These simplifications result in a more compact network.

Starting with the inductor $L_{M L 1}$ in the left sub-network and capacitor $C_{M R 1}$ in the right sub-network, they can be transferred to the right and left sub-networks, respectively, with an appropriate turn ratio of the transformers. Transferring the capacitor $C_{M R 1}$ to the left also provides the option to incorporate the output capacitance of the driver stage into it. Then, the two transformers can be merged and modeled with a single transformer with a turn ratio of $1: n_{1} \times n_{2}$, where $1: n_{1}$ and $1: n_{2}$ are the turn ratios of the left and right transformers, respectively. The simplified network, which consists of the reactive absorptive components and a single transformer, is shown in Fig. 2.16(b). As expected and also confirmed by the simulation results shown in Fig. 2.16(b), the bandwidth is limited by the reactive absorptive network of the right sub-network in Fig. 2.16(a) since the maximum reflection coefficient of the final interstage network still remains around -10 dB.

As shown in Fig. 2.16(c), the inductor $L_{M R 2}$ is split into three parts, namely, $L_{1}, L_{2}$ and $L_{s p}$. Then, using a Norton transformation, the righthanded $L$-network consisting of inductors $L_{M}$ and $L_{1}$ is transformed to the left-handed $L$-network (with appropriate scaling). Shown in the red box in Fig. 2.16(c), the resulting circuit can be modeled with two inductors $L_{\text {int1 }}$ and $L_{\text {int2 } 2}$ that are coupled with the coupling factor of $k$ (Fig. 2.16(c)). The
following equations can be used to properly split the inductor $L_{M R 2}$ so that the final $T$-network as well as the ideal transformer can be modeled with two coupled lines.

$$
\begin{gather*}
L_{1}=L_{M}\left(\sqrt{1+\frac{L_{M R 2}-L_{S P}}{L_{M}}}-1\right)  \tag{2.9}\\
L_{2}=L_{M}\left(\frac{L_{M R 2}-L_{S P}}{L_{M}}-\sqrt{1+\frac{L_{M R 2}-L_{S P}}{L_{M}}}+1\right), \tag{2.10}
\end{gather*}
$$

where $L_{M R 2}$ and $L_{s p}$ are the series inductance of the right reactive absorptive network shown in Fig. 2.16(a) and the one that is required for the splitter $S P_{\text {int }}$, respectively.

$$
\begin{gather*}
n=\frac{1}{k}=1+\frac{L_{1}}{L_{M}}  \tag{2.11}\\
L_{i n t 1}=\frac{L_{M}}{N^{2}} \text { and } L_{i n t 2}=\frac{L_{M}}{k^{2}}, \tag{2.12}
\end{gather*}
$$

where $n, N, k, L_{i n t 1}$ and $L_{i n t 2}$ are the Norton transformation ratio, $n_{1} n_{2}$, the coupling factor, the inductance of the primary and secondary terminals of the coupled lines, respectively. The final interstage matching network is shown in Fig. 2.16(d) that is more compact than the structure shown in Fig. 2.16(a). Moreover, the driver and power stages can be biased independently through the coupled lines as the bias voltages can be applied to the inductors $L_{\text {int } 1}$ and $L_{\text {int } 2}$, respectively. As a result, no DC-block capacitor, whose added parasitics (particularly to the substrate) degrade the high frequency performance, is needed in the interstage network.

## Output power combiner implementation

The objective of this block is to convert the load impedance of the antenna to the impedance obtained from the load-pull for the optimum power delivery with acceptable efficiency. While the optimum susceptance is meant to negate the output parasitic capacitance of the power cell, the optimum conductance is approximately given by:

$$
\begin{equation*}
G_{o p t}^{-1} \approx \frac{V_{D D}-V_{o v}}{I_{D C}}, \tag{2.13}
\end{equation*}
$$

where $V_{D D}, V_{o v}$, and $I_{D C}$ are the DC power supply, overdrive voltage, and DC current of the transistor at the power for which load-pull is performed, respectively [41]-[42]. Since $Q_{L}$, that is here the ratio of optimum imaginary


Figure 2.16: Interstage matching network: (a) detailed structure, (b) simplified network and its input reflection coefficient (c) Norton transformation step, (d) final structure.
and real parts, is low ( $\sim 0.75$ ), even with a single-tuned topology, a considerably high return loss ( $>23 \mathrm{~dB}$ ) can be obtained over the frequency band of 26 to 31 GHz . Consequently, the bandwidth of the output power combiner is limited by the transformer, in contrast to that of the interstage matching network that is limited by the reactive absorptive network. Fig. 2.17 shows the complete equivalent circuit model of the output stages including the power combiner and filter. Fig. 2.18 compares the passive efficiency
of the output power combiner with and without shielding lines. Using the CPW-like technique, the combiner efficiency is improved by $5 \%$ and a passive efficiency of over $93 \%$ (equivalent to loss $\sim 0.3 \mathrm{~dB}$ ) near 28 GHz is achieved. Moreover, the combining efficiency remains higher than $90 \%$ over a bandwidth of $\sim 7 \mathrm{GHz}$.


Figure 2.17: Equivalent circuit model of the output stages including the power combiner and filter.


Figure 2.18: Output combiner efficiency with and without shield lines.

### 2.4 Reliability and Stability of the PA

### 2.4.1 Voltage distribution across the stacked device of the output stage

Since the stacked configuration is used for the output stage of the PA, the voltage distributed across the stacked device should be taken into account to make sure that the device is not overly stressed. Among all voltage distributions across different terminals of the stacked device, the more critical ones are the gate-drain and drain-bulk (in bulk CMOS technologies) voltages as they are connected to a high-voltage node [32]. Given that a two-stacked configuration is used in this work, the drain-bulk terminal is not as critical as the gate-drain. In fact, based on [32], the maximum allowed number of stacked devices in bulk CMOS technologies is reported to be three or four (while using three and four times of the nominal supply, respectively) and this number is limited by the junction breakdown voltage. Thus, scaling the supply by a factor of two is tolerable. As for the gate-drain terminal, its maximum allowable swing is $2 \times V_{D D}$ [32]-[33]. Fig. 2.19 shows the voltage swing of the drain-gate terminal of the stacked FET, confirming that it remains under $2 V_{D D}(\sim 2.4 \mathrm{~V}$ here). Also, the varactor, which is connected to the gate of the stacked FET, can help reducing the voltage swing across the gate-drain terminal under large-signal conditions, further improving the reliability.


Figure 2.19: Voltage swing of the drain-gate terminal of the stacked FET.

### 2.4.2 Stability of the amplifier

At higher frequencies, the stability of amplifiers is a major concern as the reverse isolation of active components is low. A practical approach to make a common-source amplifier more unilateral and thus unconditionally stable is to employ the cascode or stacked topology. While the latter is used in the power stage, the stability may be compromised since the gate of the stacked transistor is connected to a varactor instead of directly going to AC GND. To guarantee that the proposed structure, which uses a varactor, remains stable over different gate and drain biases as well as process corners, the approach used in [34] is carried out. The limited quality factor of the varactor along with that of the inductor placed at the source of $M_{c}$ (node $X$ in Fig. 2.4(a)) helps to enhance the stability. The limited $Q$ of the varactor creates a resistance which partially cancels the unwanted negative resistance seen at the gate. The inductance at node X causes the negative impedance seen at the gate of $M_{c}$ to be considerably smaller.

Another concern is the series capacitor placed at the source of $g_{m}$ devices to form an $L C$ filter so that the inductive part of the interconnect parasitics can be canceled out. Since this filter may have a capacitive behavior at low frequencies, it may adversely affect the stability of the overall system. However, as the frequency drops, the parallel resonance frequency, $f_{p}$, becomes dominant, improving the stability by creating a high impedance path at the source of the PA cells and thus reducing the gain of the system. From this property, the lower corner of the frequency band can be set as discussed previously since it creates a parallel self-resonance. Furthermore, the interstage coupled lines, which are designed for the optimum performance in the frequency band of interest, lower the gain at low frequencies and thus help to dampen any unwanted low frequency signal. At the other end of the frequency band, the high-quality matching capacitor placed at the output of the combiner sets the upper corner frequency of the system and attenuates the gain of the overall system.

The even-mode stability ( $\mu$-factor) of the structure is evaluated for each stage (driver and output) separately and also the complete structure at different PVT corners and different bias voltages. Fig. 2.20 shows the simulated stability factor of the whole structure at different corners, temperatures and two different output power levels of $-7 /+23 \mathrm{dBm}$ at 28 GHz . The $\mu$-factor remains greater than one over the wide frequency range indicating that it is unconditionally stable. To ensure that any odd-mode oscillation [35], which leads to gain drop due to internal oscillation, is dampened, the resistors between each two inputs and also each two output nodes of the
power cell with the values of $1.5 k \Omega$ are added (the resistors in Fig. 2.16(d)) [35]. An off-chip resistive biasing network [35] consisting of resistors and capacitors is used to prevent low-frequency oscillations. The measured output spectrum of the PA is also monitored over the frequency range of DC to 60 GHz (with and without an input signal) and no oscillation is observed (Fig. 2.21)


Figure 2.20: Simulated stability factors at different corners, temperatures and power levels (Full-Power $=23 \mathrm{dBm}$; Small-Power=-7 dBm).


Figure 2.21: Output spectrum for (a) no input signal, (b) $P_{\text {out }} \approx 23 \mathrm{dBm}$.

### 2.5 Experimental Results

The PA is designed and fabricated in 1-poly 9 -metal (1P9M) $65-\mathrm{nm}$ bulk CMOS general purpose process. The chip micrograph is shown in Fig. 2.22. The chip size including the pads is $500 \mu \mathrm{~m} \times 800 \mu \mathrm{~m}$. The chip is mounted on a printed-circuit board (PCB) and the PA is measured through on-wafer
probes with all DC pads wire bonded to the PCB. S-parameters are measured with a Keysight N5225B PNA network analyzer. The S-parameter measurements showing the small-signal gain and input matching results are shown in Fig. 2.23. The measured results are in good agreement with simulations. Fig. 2.24 shows the large-signal behavior of the PA including the output power, gain, and PAE versus the input power for a single tone input at 28 GHz . The measurement results demonstrate $P_{\text {sat }}$ of 23.2 dBm , output $P_{1 d B}$ of 22.7 dBm , and peak PAE of $35.5 \%$, at 28 GHz . Fig. 2.25 shows the variation of PAE and $P_{\text {out }}$ versus frequency. The proposed PA has a $P A E_{\max }$ and $P_{\text {sat }}$ of higher than $30 \%$ and 22.1 dBm , respectively, across the frequency range of 26 GHz to 31 GHz . The histograms of $P_{o-s a t}$ and $P A E_{\max }$ for 20 measured samples are shown in Fig. [2.26. The measurement results of 20 samples confirm that the proposed techniques are reliable. To perform these measurements, first, one sample is chosen randomly and the biasing voltages for that sample are adjusted to achieve its optimum performance. Then the same biasing setup is used for the rest of the samples.


Figure 2.22: Die micrograph of the proposed PA.
Fig. 2.27 shows measured and simulated AM-AM and AM-PM distortions confirming high linearity of the proposed structure. AM-AM and AM-PM have the overshoot of less than 0.3 dB and $0.2^{\circ}$, respectively, at 28 GHz . Also, by including the varactor at the gate of the stacked transistor, the output $P_{1 d B}$ is improved by more than 5.5 dB . The proposed technique also shows a favorable performance as compared to the case where the varactor is put at the gate of the bottom transistor (Fig. 2.27). AM-PM


Figure 2.23: Measured and simulated $S$-parameters results.


Figure 2.24: Measured large-signal behavior of the PA showing $P_{\text {out }}$, gain, and $P A E$ at 28 GHz .
distortion at $P_{1 d B}$ is shown in Fig. 2.28. The results of Fig. 2.28(a) show that AM-PM remains less than $1.9^{\circ}$ over the frequency band of 26 GHz to 31 GHz . Moreover, chip to chip variation of AM-PM, which is measured for 20 samples at 28 GHz , is negligible confirming the reliability of the proposed linearity technique (Fig. 2.28(b)). Experimental results for a 64-QAM signal with $1 \mathrm{GS} / \mathrm{s}(6 \mathrm{~Gb} / \mathrm{s})$ and $2.5 \mathrm{GS} / \mathrm{s}(15 \mathrm{~Gb} / \mathrm{s})$ are shown in Fig. 2.29. Due to the linearity and efficiency improvements, the PA shows the average PAEs of $16.1 \%$ and $14.9 \%$ at the high average output power levels of


Figure 2.25: Measured $P_{o-s a t}, P_{o 1 d B}, P A E_{\max }$ and $P A E_{1 d B}$ versus frequency.


Figure 2.26: Measured $P_{o-s a t}$, and $P A E_{\max }$ for 20 samples at 28 GHz , (a) $P_{o-s a t}$, (b) $P A E_{\max }$.
16.5 dBm and 16.1 dBm for the $1 \mathrm{GS} / \mathrm{s}$ and 2.5 GS/s modulated signals, respectively, without using any external pre-distortion circuitry. The histograms of EVM and ACPR for 20 measured samples are shown in Fig. 2.30. The results show typical EVM/ACPR around $-26.4 \mathrm{~dB} /-28.3 \mathrm{dBc}$ with small die-to-die variations (EVM and ACPR variations are confined within $\pm 0.8 \mathrm{~dB}$ and $\pm 1 \mathrm{~dB}$, respectively). Fig. 2.31 shows the EVM and ACPR versus output power for the $2.5 \mathrm{GS} / \mathrm{s}$ modulated signal validating the performance of the proposed technique. Table 5.1 summarizes the performance of the proposed PA and compares it with that of state-of-the-art PAs. The results confirm that the proposed CMOS stack-based topology has a comparable efficiency and linearity with those of the state-of-the-art
common-source CMOS structures, while offering a favorable output power (even compared to PAs implemented in higher performance processes and using higher supply voltages). Thus, while the output swing is improved by using the stacked configuration, its linearity and efficiency are also enhanced by means of the proposed techniques.


Figure 2.27: Linearity results with, without the varactor, and with varactor at the bottom FET gate: (a) AM-AM and (b) AM-PM.


Figure 2.28: Measured AM-PM Results: (a) AM-PM versus frequency, (b) AM-PM for 20 samples at 28 GHz .


Figure 2.29: Modulation measurement results $\left(f_{\text {carrier }}=28 \mathrm{GHz}\right)$.


Figure 2.30: Measured EVM, and ACPR for 20 samples for data rate of 2.5 GS/s, (a) EVM, (b) ACPR.


Figure 2.31: Measured EVM and ACPR versus output power for data rate of $2.5 \mathrm{GS} / \mathrm{s}(\mathrm{PAPR}=8.3 \mathrm{~dB})$.

Table 2.1: Performance comparison of state-of-the-art 5G PAs.

|  | This Work | 14] | 15] | 21] | 371 | 381 | [39] | 401 | 431 | $44]$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Technology | 65 nm CMOS | 65 nm CMOS | 40 nm CMOS | 45 nm SOI | 130 mm SiGe | 130 mm SiGe | 45 nm SOI | 130nm SiGe | 45 nm SOI | 45 nm SOI |
| Freq. (GHz) | 28 | 28 | 27 | 41 | 28.5 | 28 | 27 | 40 | 28 | 30 |
| Gain (dB) | 18 | 15.8 | 22.4 | 8.9 | 20 | 15* | 19.1 | 23.4 | 20.5 | 7 |
| $P_{\text {out }, 1 d B}(\mathrm{dBm})$ | 22.7 | 14 | 13.7 | NA | 15.2 | NA | 22.4 | NA | 19.1 | NA |
| $P_{\text {out,sat }}(\mathrm{dBm})$ | 23.2 | 15.6 | 15.1 | 21.6 | 17 | 23 | 23.3 | 23.7 | 20.4 | 17 |
| $P A E_{1 d B}(\%)$ | 34 | 34.7 | 31.1 | NA | 39.2 | 40.3 | 39.4 | NA | 42.5 | NA |
| $P A E_{\text {max }}(\%)$ | 35.5 | 41 | 33.7 | 25.1 | $44^{* *}$ | 41.4 | 40.1 | 28.5 | 45 | $50.5^{* * *}$ |
| AM-PM ( ${ }^{\circ}$ ) | 1.6 | 0.7 | NA | NA | 3 | NA | <5 | NA | NA | NA |
| @ $P_{1 d B}$ |  |  |  |  |  |  |  |  |  |  |
| Modulated signal | 64QAM | 64QAM | 64QAM |  | 64QAM | 64QAM | 64QAM | 16QAM | $\begin{aligned} & 5 \mathrm{G} 64 \mathrm{QAM} \\ & \text { OFDM } \end{aligned}$ | 64QAM |
|  | $15 \mathrm{~Gb} / \mathrm{s}$ | $2 \mathrm{~Gb} / \mathrm{s}$ | OFDM 8-CC |  | $18 \mathrm{~Gb} / \mathrm{s}$ | OFDM | $6 \mathrm{~Gb} / \mathrm{s}$ | $4 \mathrm{~Gb} / \mathrm{s}$ | $0.8 \mathrm{~Gb} / \mathrm{s}$ | $3 \mathrm{~Gb} / \mathrm{s}$ |
| $P_{o, \text { avg }}(d B m)$ | 16.1 | 9.8 | 6.7 |  | 9.8 | 14.3 | 15.9 | 16.4 | 11.3 | 10.2 |
| EVM (dB) | -26.2 | -26.4 | -25 | NA | -25 | -30.5 | -25.3 | -18.2 | -25.1 | -28 |
| ACPR (dBc) | -28.32 | -30 | NA |  | NA | -33 | -29.6 | $-25^{*}$ | -25.6 | -26.4 |
| PAE(\%) | 14.9 | 18.2 | 11 |  | 18.4 | 25.3 | 29.1 | 19.9*** | 16.6 | 31.2 |
| Active area | 0.2 | 0.24 | 0.225 | $0.3^{* *}$ | 0.291 | 0.56 ** | 0.52 | 0.96 | 0.21 | $1.65^{* *}$ |
| $V_{D D}$ | 2.2 | 1.1 | 1.1 | 5 | 1.9 | 4 | 2 | , | , | 1.2 |
| Architecture | $\begin{aligned} & \text { 4-way } \\ & \text { Combining } \end{aligned}$ | Transformer-Based AM-PM Correction | 1-Way Transformer-Based Combining | $\begin{array}{\|c\|} \hline \text { 4-stacked } \\ \text { configuration } \end{array}$ | Differential Continuous-Mode Harmonically-Tuned | Triaxial Balun Outphasing | 3-Bit Doherty | 4-Bit Asymmetrical | $\begin{gathered} \text { Compensated } \\ \text { Distributed } \\ \text { Balun } \end{gathered}$ | Outphasing |

* Estimated from a plot ${ }^{* *}$ Including pads ${ }^{* * *}$ Drain efficiency


### 2.6 Summary

In this chapter, a $28-\mathrm{GHz}$ CMOS PA suitable for 5 G applications is designed and characterized. The design uses four sub-PAs whose outputs are combined to deliver $P_{\text {sat }}=23.2 \mathrm{dBm}$ at 28 GHz . To improve the efficiency, a proposed low-loss CPW-like power combiner with the passive efficiency of $93 \%$ is used along with an LC filter at the output of the PA, yielding $P A E_{\max }$ in excess of $35 \%$. A single varactor controlled by an envelope detector is also proposed to improve the linearity of each sub-PA resulting in improvement of both AM-AM and AM-PM distortions. The difference between $P_{1 d B}$ and $P_{\text {sat }}$ remains negligible ( $\sim 0.5 \mathrm{~dB}$ ), confirming the high linearity of the circuit. A systematic approach to the design of matching networks is also presented. The PA is capable of delivering an EVM of -26.2 dB when transmitting a $2.5 \mathrm{GS} / \mathrm{s}(15 \mathrm{~Gb} / \mathrm{s}) 64-\mathrm{QAM}$ modulated signal at average Pout of 16.1 dBm . Implemented in a $65-\mathrm{nm}$ bulk CMOS, the performance of the proof-of-concept PA compares favorably with that of the state-of-the-art designs implemented in advanced CMOS or other higher performance processes, in particular, in terms of $P_{\text {out }, 1 d B}, P_{\text {out , sat }}$ and area. The fact that $P_{\text {out, }, 1 d B}$ and $P_{\text {out, sat }}$ of the proposed design are relatively high and close to each other is also an indication of improved linearity of the system. Please also note that for the relatively high output power the proposed design offers the reasonable efficiency.

## Chapter 3

## On the Design of $n^{\text {th }}$-Order Polyphase All-Pass Filters

Polyphase filters (PPFs) that generate outputs with known phase differences are among the key building blocks of modern wireless communication systems [24, [45]- 46]. For example, polyphase filters are often used to generate quadrature local oscillator outputs that are required in direct-conversion and low-intermediate-frequency (low-IF) receivers as well as in vector-sumbased phase shifters [47]-[50] [58]-[61]. LC filters have been traditionally used for quadrature signal generation. However, the networks that are conventionally used are mainly first or second order [60] 61]. Furthermore, they produce the quadrature phase shift for either a single frequency or over a relatively narrow frequency band. The conventional RC polyphase filters, in spite of their widespread use, suffer from two major issues including a poor input matching (to $50 \Omega$ ) and trade-off between the loss and phase difference accuracy [51]. Although a poor input matching (to $50 \Omega$ ) may not be a major concern in transceivers where PPFs function at IF frequencies, this drawback cannot be ignored in radio-frequency (RF) applications including RF vector-sum phase shifters [50], [62]-[67]. For applications where input matching is important, the overall performance of the system will be adversely affected by poor input matching. Due to the trade-off between the amplitude loss and phase accuracy, although polyphase filters are widely used in the LO signal path where the signal amplitude level is large, the loss introduced by such networks is a major issue if they are to be used in the main RF signal path. The situation will be exacerbated when for wideband operation multistage polyphaser filters are used. The proposed $n^{\text {th }}$-order polyphase all-pass filter (PAF) addresses the abovementioned disadvantages.

In this chapter, the analyses are mainly focused on polyphase all-pass filters (PAFs). Without loss of generality and to avoid long and cumbersome equations, we provide the detailed analysis of the system for up to $4^{\text {th }}$ order PAFs. Note that in practical implementation of PAFs, especially in integrated circuits (ICs), the order of PAFs is rarely beyond 4. However,
the equations for higher order PAFs can be derived similar to the approach presented here.

The organization of the chapter is as follows: Section 3.1 introduces the proposed structure and discusses its properties. Section 3.2 presents the mathematical analysis of $2^{\text {nd }}, 3^{\text {rd }}$, and $4^{\text {th }}$-order PAFs. In Section 3.3, the design procedure for the proposed PAF is presented in the context of an example. Section 3.4 is devoted to practical considerations and analysis of some errors caused by non-ideal effects including parasitic capacitances, component tolerance and mismatch, and quality factor (Q) of the inductor. Since in the proposed structure the input impedance is independent of phase accuracy and the overall filter loss, in contrast to the conventional structures [51], the value of the input impedance can be set independently and it does not impact the output performance. The proposed design is verified experimentally through a proof-of-concept prototype and measurement results confirm the operation of the structure over the frequency band of 10 to 100 MHz . Note that this proof-of-concept prototype using discrete components if one would implement the structure monolithically, higher frequency of operation can be achieved. The experimental results are presented in Section 3.5.


Figure 3.1: Key building block of the proposed system.

### 3.1 The Proposed $n^{\text {th }}$-Order Polyphase All-Pass Filter

### 3.1.1 Introduction to The Structure:

The key building block of the proposed filter (as shown in Fig. 3.1) is a passive network consisting of two sub-blocks, $\mathcal{N}_{1}$ and $\mathcal{N}_{2}$ ). $\mathcal{N}_{1}$ and $\mathcal{N}_{2}$ are in turn passive networks that consist of reactive components only and the two networks are dual of each other. Furthermore, we assume that we have the following relationship between the input impedance of the two networks:

$$
\begin{equation*}
Z_{1}(s) \times Z_{2}(s)=R^{2}, \tag{3.1}
\end{equation*}
$$

where, $Z_{1}(s)$ and $Z_{2}(s)$ are the input impedances of networks $\mathcal{N}_{1}$ and $\mathcal{N}_{2}$, respectively, $s$ is the complex frequency, and $R$ is the desired input impedance of the overall circuit. Given the above assumptions, it can be shown that for the overall passive network shown in Fig. 3.1, we have:
i) The input impedance is equal to $R$ and is independent of the frequency of operation.
ii) The magnitude of the overall voltage transfer function is one regardless of the order of $\mathcal{N}_{1}$ and $\mathcal{N}_{2}$.


Figure 3.2: The proposed all-pass filter structure capable of generating inphase and quadrature signals.

It will be shown in the next subsection that by proper design of two such networks and putting them together as shown in Fig. 3.2 one can generate in-phase and quadrature signals. This proposed structure has the potential of generating the quadrature signals with exactly $90^{\circ}$ phase difference at $n$ distinct frequencies, while it still maintains the aforementioned properties
(constant input impedance and magnitude transfer function). This is in contrast with the conventional polyphase filters [49] [51] in which the magnitude of the overall transfer function deteriorates when more stages are added to improve the phase accuracy by increasing the order of the filter. In this discussion, $n$ denotes the total order of the networks $\mathcal{N}_{1}\left(\mathcal{N}_{2}\right)$ and $\mathcal{N}_{3}\left(\mathcal{N}_{4}\right)$. Let us review some conditions that if satisfied the structure produces the quadrature signals at $n$ distinct frequencies.

We will first discuss the requirements for reactive networks $\mathcal{N}_{i}$ where $i \in\{1,2\}$ such that Eq. (3.1) is satisfied. From circuit theory, we know that the product of the impedance of two reactive networks is constant if and only if they are dual of each other [52]. Consequently, in the event $\mathcal{N}_{2}$ is the dual network of $\mathcal{N}_{1}, Z_{1}(s) \times Z_{2}(s)$ will be a constant value. Now, the equation shown in (3.1) will be met if this constant value is a positive real number. To satisfy this condition, we require that the impedance of each component in network $\mathcal{N}_{2}$ is $R^{2}$ divided by the impedance of the corresponding component in the dual network, $\mathcal{N}_{1}$. Reactive networks $\mathcal{N}_{i}$ where $i \in\{3,4\}$ should also meet the same condition.

Furthermore, we assume that the number of poles or zeros of the impedance of each network $\mathcal{N}_{i}, i \in\{1,3\}$ is equal to the total number of the reactive components of the network. That is, the reactive components of the network cannot be combined in such a way to reduce both the number of zeros and poles of the network, e.g.,there is no pole-zero cancellation in the circuit. If this condition is held, the proposed structure ( $\mathrm{Fig} \sqrt{3.2}$ ) would be capable of producing quadrature signals at up to $n=n_{1}+n_{3}$ distinct frequencies, where $n_{1}$ and $n_{3}$ are the total number of the reactive components used in the networks $\mathcal{N}_{1}$ and $\mathcal{N}_{3}$, respectively.

Due to duality, if the above condition is satisfied for networks $\mathcal{N}_{1}$ and $\mathcal{N}_{3}$, it will be also satisfied for the networks $\mathcal{N}_{2}$ and $\mathcal{N}_{4}$ since the poles of $\mathcal{N}_{1}$ $\left(\mathcal{N}_{3}\right)$ are the zeros of $\mathcal{N}_{2}\left(\mathcal{N}_{4}\right)$ and vice versa [52. Let us now further clarify these conditions by an example. For $n_{1}=4$, example proper and improper structures are shown in Fig. 3.3. The structure of Fig. 3.3(b) does not meet the above condition as there is one pole-zero cancellation happening at its impedance function which causing the number of poles and zeros of its impedance function to be smaller than the total number of the reactive components of the network.

In what follows, we will prove that the structure of Fig. 3.2 is capable of generating quadrature signals whose phase difference is exactly $90^{\circ}$ at $n$ distinct frequencies where $n$ is the order of the structure. We will also show that by increasing the order of the structure, one can either improve the phase accuracy of the system over the bandwidth of interest, or achieve the


Figure 3.3: Examples of $4^{\text {th }}$-order $\mathcal{N}_{i}$ network. (a) Proper type, and (b) improper type.
same phase accuracy over a wider bandwidth.

### 3.1.2 Analysis of the Proposed Structure:

In this subsection, we analyze the proposed structure for the case where there are $n_{1}$ and $n_{3}$ reactive components in networks $\mathcal{N}_{1}$ and $\mathcal{N}_{3}$, respectively, and thus the structure can produce quadrature outputs (with $90^{\circ}$ phase difference) at $n=n_{1}+n_{3}$ distinct frequencies.

In this analysis, we assume $n_{1}, n_{3}$ and $\frac{n_{1}+n_{3}}{2}$ are all even numbers. The analysis for each of the other cases (different combinations of even and odd numbers), is similar to that of the case provided here and therefore for the purpose of brevity we omit the analysis. However, we summarize the cases that provide viable solutions.

From basic circuit theory [52], the input impedance of a single-port network containing only reactive components is purely reactive, otherwise Tellegen's theorem is violated. Thus, the input impedance of the network $\mathcal{N}_{1}$, with $n_{1}$ reactive components (while $n_{1}$ is even) can be written as:

$$
\begin{equation*}
Z_{1}(s)=\frac{\alpha s \times\left(1+\sum_{k=1}^{\frac{n_{1}}{2}-1} a_{2 k}^{\prime} s^{2 k}\right)}{1+\sum_{k=1}^{\frac{n_{1}}{2}} a_{2 k} s^{2 k}} \tag{3.2}
\end{equation*}
$$

where $\alpha, a_{2 k}$, and $a_{2 k}^{\prime}$ are real coefficients. $s$ is equal to $j \omega$. Note that depending on the network, the first term in the numerator, i.e., $\alpha s$ may
appear in the denominator instead, in which case the impedance of the dual network will have the similar term in its numerator or vice versa.

Given Eq. (3.2), the input impedance of network $\mathcal{N}_{2}$ (the dual of $\mathcal{N}_{1}$ ) is given by:

$$
\begin{equation*}
Z_{2}(s)=\frac{R^{2}}{Z_{1}(s)} \tag{3.3}
\end{equation*}
$$

Similarly, the input impedance of the network $\mathcal{N}_{3}$ is:

$$
\begin{equation*}
Z_{3}(s)=\frac{\beta s \times\left(1+\sum_{k=1}^{\frac{n_{3}}{2}-1} b_{2 k}^{\prime} s^{2 k}\right)}{1+\sum_{k=1}^{\frac{n_{3}}{2}} b_{2 k} s^{2 k}}, \tag{3.4}
\end{equation*}
$$

where $\beta, b_{2 k}$, and $b_{2 k}^{\prime}$ are real coefficients. $s$ is equal to $j \omega$. Also, the input impedance of $\mathcal{N}_{4}$ (dual of $\mathcal{N}_{3}$ ) will be:

$$
\begin{equation*}
Z_{4}(s)=\frac{R^{2}}{Z_{3}(s)} . \tag{3.5}
\end{equation*}
$$

It can be shown that the output voltages of the structure shown in Fig. 3.2, namely, $V_{I}$ and $V_{Q}$ can be written as:

$$
\begin{equation*}
V_{I}(s)=\frac{R-Z_{1}(s)}{R+Z_{1}(s)} \times V_{i n} \tag{3.6}
\end{equation*}
$$

and

$$
\begin{equation*}
V_{Q}(s)=\frac{R-Z_{3}(s)}{R+Z_{3}(s)} \times V_{i n} . \tag{3.7}
\end{equation*}
$$

Thus, the phases of the output voltages are given by:

$$
\begin{equation*}
\angle V_{I}=-2 \times \arctan \left(\omega \frac{\alpha}{R} A(\omega)\right) \tag{3.8}
\end{equation*}
$$

and

$$
\begin{equation*}
\angle V_{Q}=-2 \times \arctan \left(\omega \frac{\beta}{R} B(\omega)\right), \tag{3.9}
\end{equation*}
$$

where

$$
\begin{equation*}
A(\omega)=\frac{1+\sum_{k=1}^{\frac{n_{1}}{2}-1}(-1)^{k} a_{2 k}^{\prime} \omega^{2 k}}{1+\sum_{k=1}^{\frac{n_{1}}{2}}(-1)^{k} a_{2 k} \omega^{2 k}} \tag{3.10}
\end{equation*}
$$

and

$$
\begin{equation*}
B(\omega)=\frac{1+\sum_{k=1}^{\frac{n_{3}}{2}-1}(-1)^{k} b_{2 k}^{\prime} \omega^{2 k}}{1+\sum_{k=1}^{\frac{n_{3}}{2}}(-1)^{k} b_{2 k} \omega^{2 k}} \tag{3.11}
\end{equation*}
$$

The overall output phase difference can therefore be written as:

$$
\begin{equation*}
\theta=\angle V_{I}-\angle V_{Q}=2 \arctan \frac{\omega \frac{\beta}{R} B(\omega)-\omega \frac{\alpha}{R} A(\omega)}{1+\frac{\omega^{2}}{R^{2}} \alpha \beta A(\omega) B(\omega)} . \tag{3.12}
\end{equation*}
$$

For the output phase difference $\theta$ to be equal to $\frac{\pi}{2}$ (i.e., $90^{\circ}$ ), the following condition should be met.

$$
\begin{equation*}
\omega \frac{\beta}{R} B(\omega)-\omega \frac{\alpha}{R} A(\omega)=1+\frac{\omega^{2}}{R^{2}} \alpha \beta A(\omega) B(\omega) . \tag{3.13}
\end{equation*}
$$

By substituting Eqs. (3.10) and (3.11) into Eq. (3.13), and simplifying the equation, we have:

$$
\begin{equation*}
1+\sum_{k=1}^{n_{1}+n_{3}=n} x_{k} \omega^{k}=0 \tag{3.14}
\end{equation*}
$$

where $x_{k}$ is calculated from Eq. (3.15).

$$
x_{k}=\left\{\begin{array}{l}
(-1)^{\frac{k}{2}}\left(\sum_{i=0}^{\frac{n}{2}} a_{2 i} b_{k-2 i}-\sum_{i=0}^{\frac{n}{2}-1} \frac{\alpha \beta}{R^{2}} a_{2 i}^{\prime} b_{k-2 i-2}^{\prime}\right) \text { when } k \text { is even }  \tag{3.15}\\
(-1)^{\frac{k-1}{2}}\left(\sum_{i=0}^{\frac{n}{2}-1} \frac{\alpha}{R} a_{2 i}^{\prime} b_{k-2 i-1}-\sum_{i=0}^{\frac{n}{2}} \frac{\beta}{R} a_{2 i} b_{k-2 i-1}^{\prime}\right) \text { when } k \text { is odd } \\
\text { and } a_{0}=a_{0}^{\prime}=b_{0}=b_{0}^{\prime}=1
\end{array}\right.
$$

For example, if $n_{1}=n_{3}=4$, from Eq. (3.15) one can calculate $x_{k}$ 's as follows:

$$
\begin{gathered}
x_{8}=a_{4} b_{4}, \\
x_{7}=-\frac{\alpha}{R} a_{2}^{\prime} b_{4}+\frac{\beta}{R} a_{4} b_{2}^{\prime}, x_{6}=-a_{4} b_{2}-a_{2} b_{4}+\frac{\alpha \beta}{R^{2}} a_{2}^{\prime} b_{2}^{\prime}, \\
x_{5}=\frac{\alpha}{R}\left(a_{2}^{\prime} b_{2}+b_{4}\right)-\frac{\beta}{R}\left(a_{4}+a_{2} b_{2}^{\prime}\right),
\end{gathered}
$$

$$
\begin{gathered}
x_{4}=a_{4}+a_{2} b_{2}+b_{4}-\frac{\alpha \beta}{R^{2}}\left(a_{2}^{\prime}+b_{2}^{\prime}\right) \\
x_{3}=-\frac{\alpha}{R}\left(a_{2}^{\prime}+b_{2}\right)+\frac{\beta}{R}\left(a_{2}+b_{2}^{\prime}\right) \\
x_{2}=-a_{2}-b_{2}+\frac{\alpha \beta}{R^{2}} \text { and } x_{1}=\frac{\alpha}{R}-\frac{\beta}{R}
\end{gathered}
$$

The roots of Eq. (3.14) are the frequencies where the phase difference between $V_{I}$ and $V_{Q}$ is $\frac{\pi}{2}$ or $90^{\circ}$. Thus, for an $n^{\text {th }}$-order system, we would like to have $n$ positive real roots for Eq. (3.14). Based on Descartes' rule of signs [53] and [54], for a polynomial of degree $n$ with real coefficients, the necessary condition for having up to $n$ positive real roots is that the sign of the coefficients of the polynomial alternatively changes, that is, if we represent the coefficients of the polynomial by $x_{k}, k \in 0,1,2, \ldots, n$ then $x_{k}=(-1)^{k}\left|x_{k}\right|$ for all $k \in 0,1,2, \ldots, n$, i.e., there are $n$ sign changes in the coefficients. Note that since all the components, inductors and capacitors, have a real value, the coefficients $x_{k}$ are all real. Furthermore, from Eq. (3.1), note that the coefficient of the lowest degree of polynomial, $x_{0}$, is one and thus it is positive. So, based on Descartes' rule of signs, the even-order coefficients of Eq. (3.1) should be positive and its odd-order coefficients should be negative. Eq. (3.15) shows that $x_{k}$ for $k=n$ is always positive because it consists of the multiplication of two positive coefficients ( $a_{n} b_{n}=R^{n} C_{1} \ldots C_{n}$ ). Thus, both $x_{0}$ and $x_{n}$ are positive. Now, for the rest of $n-1$ coefficients, namely, $x_{k}, k \in\{1,2, \ldots, n-1\}$ one can show that we have $n$ unknowns (in the form of $\alpha, \beta, a_{2}, b_{2}, a_{2}^{\prime}, b_{2}^{\prime}, \ldots$ ) in $n-1$ independent equations (inequations, more accurately) for $x_{k}, k \in\{1,2, \ldots, n-1\}$. The reason why these equations are independent is that first it is assumed that there is no pole-zero cancellation, leading some unknowns $\alpha, \beta, a_{2}, b_{2}, a_{2}^{\prime}, b_{2}^{\prime}, \ldots$ to be dependent together or reducing the order of filter. Moreover, carefully examining Eq. (3.15), we note that there is no dependency between two different coefficients $x_{k}$, namely, $x_{i}$ and $x_{j}$. Thus, in general, this system of $n-1$ equations and $n$ unknowns has infinite number of solutions. Since each $x_{k}$ is summation of a positive and negative term, out of infinite solutions, there are ones that are obtained from real components and, in turn, purely positive variables $\alpha, \beta, a_{2}, b_{2}, a_{2}^{\prime}, b_{2}^{\prime}, \ldots$. Therefore, based on Descartes' rule of signs, since there are $n$ sign changes in the coefficients $x_{k}, k \in\{0,1,2, \ldots, n\}$, the polynomial can produce up to $n$ positive real roots.

Recall that the analysis above was based on the assumption that $n_{1}, n_{3}$ and $\frac{n_{1}+n_{3}}{2}$ are all even numbers and we showed that in such case Eq. 3.14
can produce $n$ distinct positive real roots. As we mentioned, other cases with different combinations of even and odd numbers for $n_{1}$ and $n_{3}$ can be analyzed in a similar fashion, however, some of such cases will not be able to provide $n$ distinct positive solutions and some cases will do. For brevity their analyses are omitted, however, the cases that lead to $n$ positive roots are:

- $n_{1}, n_{3}$ and $\frac{n_{1}+n_{3}}{2}$ are all even
- $n_{1}, n_{3}$ and $\frac{n_{1}+n_{3}}{2}$ are all odd
- one of $n_{1}$ or $n_{3}$ is even and the other is odd

In the following section we will present analyses of the $2^{\text {nd }}, 3^{r d}$, and $4^{\text {th }}-$ order polyphase all-pass filters (PAFs) which are representative examples of different combinations of $n_{1}$ and $n_{3}$. Furthermore, in practical implementations PAFs with orders of second, third or fourth are more commonly used as compared to those with orders beyond 4 .

### 3.2 Analysis and Design of $2^{\text {nd }}, 3^{\text {rd }}$, and $4^{\text {th }}$-Order PAFs

In this section we provide the analysis and calculation of component values for each of the structures.


Figure 3.4: $2^{\text {nd }}$-order poly-phase all-pass filter (PAF).

### 3.2.1 $\quad 2^{\text {nd }}$-order Polyphase All-Pass Filters:

The proposed structure for $2^{\text {nd }}$-order PAFs is shown in Fig. 3.4. Note that in this structure, networks $\mathcal{N}_{i}, i \in\{1,2,3,4\}$ each only contains one reactive component and $\mathcal{N}_{2}$ and $\mathcal{N}_{4}$ are the dual networks of $\mathcal{N}_{1}$ and $\mathcal{N}_{3}$, respectively.

Also, note that although the overall structure has 4 reactive components, due to pole-zero cancellation, the overall network is a $2^{\text {nd }}$-order network.

## Component Value Calculation

According to Eq. (3.1) presented in Section II, for the structure to operate properly, the following condition should be met.

$$
\begin{equation*}
\frac{L_{1}}{C_{1}}=\frac{L_{3}}{C_{3}}=R^{2} \tag{3.16}
\end{equation*}
$$

Taking Eq. (3.16) into account, the relationship between input and outputs of the structure can be written as:

$$
\left[\begin{array}{l}
V_{I}  \tag{3.17}\\
V_{Q}
\end{array}\right]=\left[\begin{array}{c}
\frac{1-R C_{1} s}{1+R C_{1} s} \\
\frac{1-R C_{3 s} s}{1+R C_{3} s}
\end{array}\right] \times V_{i n} .
$$

The phase difference between the two outputs of the network, namely, $\theta=\angle V_{I}-\angle V_{Q}$, can be written as:

$$
\begin{equation*}
\theta(\omega)=2 \tan ^{-1} \frac{R C_{3} \omega-R C_{1} \omega}{1+R^{2} \omega^{2} C_{1} C_{3}}, \tag{3.18}
\end{equation*}
$$

where $\omega$ is the angular frequency. The phase difference curve have a maximum at the following frequency:

$$
\begin{equation*}
\omega_{\max }=\frac{1}{R \sqrt{C_{1} C_{3}}} . \tag{3.19}
\end{equation*}
$$

The solution for having $\theta=\frac{\pi}{2}$ results in:

$$
\begin{equation*}
R^{2} C_{1} C_{3} \omega^{2}-R\left(C_{3}-C_{1}\right) \omega+1=0 \tag{3.20}
\end{equation*}
$$

According to Descartes's rule of signs, the necessary condition in order for the above equation to have two positive roots is $C_{3}>C_{1}$. As a consequence, the graph of $\theta$ is similar to Fig. [3.5. Based on Eqs. (3.19) and (3.20), $\omega_{\max }$ is the geometric mean of the frequencies $\omega_{1}$ and $\omega_{2}$ since the product of the roots of the quadratic Eq. (3.20) is equal to the value of $\omega_{\max }$. The desired frequency range of operation should include $\omega_{1}$ and $\omega_{2}$. Let us assume that the frequency range of interest is from $\omega_{l}$ to $\omega_{h}$. One can show that if the following condition is met:

$$
\begin{equation*}
\frac{\pi}{2}-\theta\left(\omega_{l}\right)=\frac{\pi}{2}-\theta\left(\omega_{h}\right) \tag{3.21}
\end{equation*}
$$



Figure 3.5: Typical phase difference graph of the $2^{\text {nd }}$-order PAF.
that is, the deviation of $\theta(\omega)$ from $\frac{\pi}{2}$ at both ends of the frequency band of interest are equal to each other, then, $\omega_{\max }$ is also the geometric mean of the frequencies $\omega_{l}$ and $\omega_{h}$. As shown in Fig. 3.5, $\omega_{1}$ and $\omega_{2}$ are the frequencies at which the phase difference $\theta(\omega)$ is equal to $\frac{\pi}{2}$.

If we further assume that:

$$
\begin{equation*}
\theta\left(\omega_{\max }\right)-\frac{\pi}{2}=\frac{\pi}{2}-\theta\left(w_{h}\right) \tag{3.22}
\end{equation*}
$$

then we will have:

$$
\begin{equation*}
R\left(C_{3}-C_{1}\right)=\sqrt{\frac{2\left(\omega_{l}+\omega_{h}\right)}{\omega_{\max }^{3}}} . \tag{3.23}
\end{equation*}
$$

In this case, the maximum phase error from $\frac{\pi}{2}$ occurs at frequencies $\omega_{\max }$, $\omega_{l}$, and $\omega_{h}$. If we define the ratio of $\omega_{h}$ to $\omega_{l}$ as $\alpha=\frac{\omega_{h}}{\omega_{l}}$, we can write the maximum phase difference from $\frac{\pi}{2}$ as:

$$
\begin{equation*}
\Delta \theta_{\max }=\left|\frac{\pi}{2}-2 \tan ^{-1} \sqrt{\frac{1+\alpha}{2 \sqrt{\alpha}}}\right| \tag{3.24}
\end{equation*}
$$

The above relation is obtained by using the Eqs. (3.18), (3.19), (3.23), and $\omega_{\max }^{2}=\omega_{l} \omega_{h}$.

## Design Procedure

Assuming that the lower and upper corners of frequency band are given, the design procedure of this PAF is summarized as follows:
I. Calculation of $\omega_{\max }$ using $\omega_{\max }^{2}=\omega_{l} \omega_{h}$
II. Calculation of $C_{1}$ and $C_{3}$ using Eqs. (3.19) and (3.23)
III. Calculation of inductor values using Eq. (3.16)


Figure 3.6: $3^{r d}$-order poly-phase all-pass filter (PAF).

### 3.2.2 $\quad 3^{\text {rd }}$-Order Polyphase All-Pass Filters:

For a $3^{r d}$-order PAF, the number of reactive components for the networks $\mathcal{N}_{1}$ and $\mathcal{N}_{3}$ (and their dual networks, $\mathcal{N}_{2}$ and $\mathcal{N}_{4}$ ) should be two and one, respectively, or vice versa. Taking the discussions of Section II into account, Fig. 3.6 shows a generic structure for the proposed $3^{r d}$-order PAF. Within the frequency range of interest, this structure would be capable of producing quadrature signals (with exactly $90^{\circ}$ phase difference) at three distinct frequencies. Note that the structures presented in [50] and [57] can also produce $90^{\circ}$ phase difference at three distinct frequencies, however, the proposed solution here can achieve this with fewer reactive components. In particular, as compared to the structures in [50] and [57], the structure shown in Fig. 3.6 has two fewer inductors and two fewer capacitors. Given that passive components especially inductors occupy a large on-chip area, the proposed structure is more amenable to integration than the previous ones.

## Component Value Calculations

As discussed earlier, for the structure to operate properly, Eq. (3.1) presented in Section II, should be met which results in the following condition for this structure:

$$
\begin{equation*}
\frac{L_{1}}{C_{1}}=\frac{L_{2}}{C_{2}}=\frac{L_{3}}{C_{3}}=R^{2} . \tag{3.25}
\end{equation*}
$$

Taking Eq. (3.25) into account, the relationship between the input and output voltages of the structure of Fig. 3.6 can be written as:

$$
\left[\begin{array}{l}
V_{I}  \tag{3.26}\\
V_{Q}
\end{array}\right]=\left[\begin{array}{c}
\frac{1+C_{1} C_{2} R^{2} s^{2}-R C_{2} s}{1+C_{1} C_{2} R^{2} s^{2}+R C_{2} s} \\
\frac{1-R C_{3} s}{1+R C_{3} s}
\end{array}\right] \times V_{i n} .
$$

From Eq. (3.26) and by replacing $s$ with $j \omega$ one can show that the phase difference between $V_{I}$ and $V_{Q}$, namely, $\theta$, at a given frequency, $\omega$ is :

$$
\begin{equation*}
\theta(\omega)=2 \tan ^{-1} \frac{R\left(C_{2}-C_{3}\right) \omega+R^{3} C_{1} C_{2} C_{3} \omega^{3}}{1+R^{2}\left(C_{2} C_{3}-C_{1} C_{2}\right) \omega^{2}} \tag{3.27}
\end{equation*}
$$

The necessary condition for the phase difference curve to intersect with $\frac{\pi}{2}$ line at three distinct frequencies is given by:

$$
\begin{equation*}
C_{1}<C_{3}<C_{2} \tag{3.28}
\end{equation*}
$$



Figure 3.7: Typical phase difference graph of the $3^{\text {rd }}$-order PAF.
If this condition is met, the phase difference will have one relative minimum and one relative maximum over the frequency range of interest (as
shown in Fig. 3.7). The relation between two frequencies associated with these extremum points is:

$$
\begin{equation*}
\omega_{\min } \times \omega_{\max }=\sqrt{\frac{C_{2}-C_{3}}{C_{1} C_{2} C_{3}\left(C_{2} C_{3}-C_{1} C_{2}\right) R^{4}}} . \tag{3.29}
\end{equation*}
$$

Again, if we assume that the phase difference deviation from $90^{\circ}$ at the end points of the desired frequency range as well as at the minimum and maximum deviation points are all equal (referring to Fig. 3.7), once can write:

$$
\begin{equation*}
\left|\frac{\pi}{2}-\theta\left(\omega_{l}\right)\right|=\left|\frac{\pi}{2}-\theta\left(\omega_{\min }\right)\right|=\left|\frac{\pi}{2}-\theta\left(\omega_{\max }\right)\right|=\left|\frac{\pi}{2}-\theta\left(\omega_{h}\right)\right| . \tag{3.30}
\end{equation*}
$$

It can be shown that the absolute value of the difference between the phase difference and $\frac{\pi}{2}$ at $\omega_{\min }$ and $\omega_{\max }$ will be equal if and only if the following condition are held.

$$
\begin{align*}
& \frac{C_{2} C_{3}-C_{1} C_{2}}{C_{1} C_{2} C_{3}\left(C_{2}-C_{3}\right)}=\frac{\left(C_{2}-C_{3}\right)^{2}}{\left(C_{2} C_{3}-C_{1} C_{2}\right)^{2}}= \\
& \frac{1}{\sqrt[3]{C_{1}^{2} C_{2}^{2} C_{3}^{2}}}=\sqrt{\frac{C_{2}-C_{3}}{C_{1} C_{2} C_{3}\left(C_{2} C_{3}-C_{1} C_{2}\right)}} \tag{3.31}
\end{align*}
$$

It can be shown that this condition is met if and only if $C_{3}$ is the geometric mean of the capacitors $C_{1}$ and $C_{2}$. That is,

$$
\begin{equation*}
C_{3}^{2}=C_{1} C_{2} . \tag{3.32}
\end{equation*}
$$

Considering the above relation, the Eq. (3.29) can be simplified to:

$$
\begin{equation*}
\omega_{\min } \times \omega_{\max }=\frac{1}{R^{2} C_{3}^{2}} \tag{3.33}
\end{equation*}
$$

Moreover, it can be further shown that if $\omega_{l} \times \omega_{h}=\omega_{\text {min }} \times \omega_{\text {max }}$, then, $\left|\frac{\pi}{2}-\theta\left(\omega_{l}\right)\right|=\left|\frac{\pi}{2}-\theta\left(\omega_{h}\right)\right|$. To have $\left|\frac{\pi}{2}-\theta\left(\omega_{l}\right)\right|=\left|\frac{\pi}{2}-\theta\left(\omega_{\text {min }}\right)\right|$, the following relations should hold:

$$
\begin{array}{r}
R^{2} \omega_{l} \omega_{\min }\left(C_{2}-C_{3}\right)^{2}-R^{2} C_{3}\left(\omega_{l}^{2}+\omega_{\min }^{2}\right)\left(C_{2}-C_{3}\right)-  \tag{3.34}\\
1+C_{3}^{4} R^{4} \omega_{l}^{2} \omega_{\min }^{2}+C_{3}^{2} R^{2} \omega_{l} \omega_{\min }=0
\end{array}
$$

Solving the above equation for $C_{2}-C_{3}$ results in:

$$
\begin{align*}
& C_{2}-C_{3}=\frac{C_{3}}{2}\left(\frac{\omega_{l}}{\omega_{\min }}+\frac{\omega_{\min }}{\omega_{l}}+\right.  \tag{3.35}\\
& \left.\sqrt{\left(\frac{\omega_{l}}{\omega_{\min }}+\frac{\omega_{\min }}{\omega_{l}}\right)^{2}+4\left(1+\frac{\omega_{\min }}{\omega_{h}}+\frac{\omega_{h}}{\omega_{\min }}\right)}\right) .
\end{align*}
$$

Since $\omega_{\min }$ is an extremum point, we have $\left.\frac{d \theta}{d \omega}\right|_{\omega=\omega_{\text {min }}}=0$. This equation can be solved in terms of $C_{2}-C_{3}$ which results in the following equation:

$$
\begin{align*}
& C_{2}-C_{3}= \\
& \frac{C_{3}}{2}\left(\frac{\omega_{l} \omega_{h}}{\omega_{\text {min }}^{2}}+\frac{\omega_{\min }^{2}}{\omega_{l} \omega_{h}}+\sqrt{\left(\frac{\omega_{l} \omega_{h}}{\omega_{\text {min }}^{2}}+\frac{\omega_{\min }^{2}}{\omega_{l} \omega_{h}}\right)^{2}+12}\right) . \tag{3.36}
\end{align*}
$$

From Eqs. (3.35) and (3.36), we have the following $10^{\text {th }}$-order polynomial in terms of $\omega_{\text {min }}$ :

$$
\begin{align*}
& -\omega_{\min }^{10}+3 \omega_{h}^{2} \omega_{\min }^{8}+2\left(\omega_{h} \omega_{l}^{2}-\omega_{h}^{3}\right) \omega_{\min }^{7}- \\
& 2 \omega_{l}^{2} \omega_{h}^{2} \omega_{\min }^{6}-2 \omega_{l}^{2} \omega_{h}^{4} \omega_{\min }^{4}+ \\
& 2\left(\omega_{l}^{2} \omega_{h}^{5}-\omega_{l}^{4} \omega_{h}^{3}\right) \omega_{\min }^{3}+3 \omega_{l}^{4} \omega_{h}^{4} \omega_{\min }^{2}-  \tag{3.37}\\
& \omega_{l}^{4} \omega_{h}^{6}=0 .
\end{align*}
$$

This equation is discussed in Appendix A, and it is shown that the equation has a double root at $\omega_{h}$, conjugate roots at $\pm j \sqrt{\omega_{l} \times \omega_{h}}$, one root at $\sqrt{\omega_{l} \omega_{h}}$ and one root at $-\sqrt{\omega_{l} \omega_{h}}$. Factoring out these roots, the remaining term of Eq. (3.37) will be the following $4^{\text {th }}$-order polynomial:

$$
\begin{equation*}
-\omega_{\min }^{4}-2 \omega_{h} \omega_{\min }^{3}+2 \omega_{l}^{2} \omega_{h} \omega_{\min }+\omega_{l}^{2} \omega_{h}^{2}=0 \tag{3.38}
\end{equation*}
$$

The solution to such $4^{\text {th }}$-order polynomial is discussed in literature, e.g., in [55], and one can solve for $\omega_{\text {min }}$. The result is provided in Eq. (3.39). Again, assuming $\alpha=\frac{\omega_{h}}{\omega}$, we can write the maximum phase difference from $\frac{\pi}{2}$, i.e., $\Delta \theta_{\max }$, as provided in Eq. (3.40).

$$
\begin{align*}
& \omega_{\text {min }}=\frac{\omega_{h}}{2} f(\alpha), \text { where } f(\alpha)= \\
& \left(\sqrt{\left.2+\sqrt[3]{\frac{4}{\alpha^{2}}\left(1-\frac{1}{\alpha^{2}}\right)}-\frac{2\left(1-\frac{2}{\alpha^{2}}\right)}{\sqrt{1-\sqrt[3]{\frac{4}{\alpha^{2}}\left(1-\frac{1}{\alpha^{2}}\right)}}}+\sqrt{1-\sqrt[3]{\frac{4}{\alpha^{2}}\left(1-\frac{1}{\alpha^{2}}\right)}}-1\right) \alpha \leq \sqrt{2}}\right.  \tag{3.39}\\
& \left(\sqrt{2+\sqrt[3]{\frac{4}{\alpha^{2}}\left(1-\frac{1}{\alpha^{2}}\right)}}+\frac{2\left(1-\frac{2}{\alpha^{2}}\right)}{\sqrt{1-\sqrt[3]{\frac{4}{\alpha^{2}}\left(1-\frac{1}{\alpha^{2}}\right)}}}-\sqrt{1-\sqrt[3]{\frac{4}{\alpha^{2}}\left(1-\frac{1}{\alpha^{2}}\right)}}-1\right) \alpha>\sqrt{2}
\end{align*}
$$

$$
\begin{align*}
& \Delta \theta_{\max }=\frac{\pi}{2}- \\
& 2 \tan ^{-1} \frac{\frac{\sqrt{\alpha} f(\alpha)}{2}\left(\frac{2}{\alpha f^{2}(\alpha)}+\frac{\alpha f^{2}(\alpha)}{8}+\sqrt{3+\left(\frac{2}{\alpha f^{2}(\alpha)}+\frac{\alpha f^{2}(\alpha)}{8}\right)^{2}}\right)+\left(\frac{\sqrt{\alpha} f(\alpha)}{2}\right)^{3}}{1+\frac{\alpha f^{2}(\alpha)}{4}\left(\frac{2}{\alpha f^{2}(\alpha)}+\frac{\alpha f^{2}(\alpha)}{8}+\sqrt{3+\left(\frac{2}{\alpha f^{2}(\alpha)}+\frac{\alpha f^{2}(\alpha)}{8}\right)^{2}}\right)}, \alpha=\frac{\omega_{h}}{\omega_{l}} \tag{3.40}
\end{align*}
$$

## Design Procedure

For the $3^{r d}$-order PAF, the summary of the design procedure of the proposed structure is given below.
I. Calculation of $\omega_{\min }$ using Eq. (3.39)
II. Calculation of $C_{3}$ using $\omega_{l} \times \omega_{h}=\frac{1}{R^{2} C_{3}^{2}}$
III. calculation of $C_{2}$ using Eq. (3.35) or (3.36)
IV. Calculation of $C_{1}$ using Eq. (3.32)
V. Calculation of inductor values using Eq. (3.25)

### 3.2.3 $4^{\text {th }}$ - order Polyphase All-Pass filters:

For a $4^{\text {th }}$-order PAF, the number of reactive components for both networks $\mathcal{N}_{1}$ and $\mathcal{N}_{3}$ (and their dual networks, $\mathcal{N}_{2}$ and $\mathcal{N}_{4}$ ) should be two. Fig. 3.8 shows the generic form of the structure.


Figure 3.8: $4^{\text {th }}$-order poly-phase all-pass filter (PAF).

## Component Value Calculation

For the structure shown in Fig. 3.8, the following relationship should be met so that Eq. (3.1) is satisfied.

$$
\begin{equation*}
\frac{L_{1}}{C_{1}}=\frac{L_{2}}{C_{2}}=\frac{L_{3}}{C_{3}}=\frac{L_{4}}{C_{4}}=R^{2} \tag{3.41}
\end{equation*}
$$

Taking Eq. (3.41) into account, the relationship between the input and output voltages of the structure of Fig. 3.8 can be written as:

$$
\left[\begin{array}{l}
V_{o 1}  \tag{3.42}\\
V_{o 2}
\end{array}\right]=\left[\begin{array}{l}
\frac{1+C_{1} C_{2} R^{2} s^{2}-R C_{2} s}{1+C_{1} C_{2} R^{2} s^{2}+R C_{2} s} \\
\frac{1+C_{3} C_{4} R^{2} s^{2}-R C_{4} s}{1+C_{3} C_{4} R^{2} s^{2}+R C_{4} s}
\end{array}\right] \times V_{i n}
$$

The phase difference between $V_{I}$ and $V_{Q}$, namely, $\theta$ can be written as:

$$
\begin{align*}
& \theta(\omega)=2 \tan ^{-1} \\
& \frac{R\left(C_{4}-C_{2}\right) \omega+R^{3}\left(C_{2} C_{3} C_{4}-C_{1} C_{2} C_{4}\right) \omega^{3}}{1+R^{2}\left(C_{2} C_{4}-C_{1} C_{2}-C_{3} C_{4}\right) \omega^{2}+R^{4} C_{1} C_{2} C_{3} C_{4} \omega^{4}} \tag{3.43}
\end{align*}
$$

Based on Descartes's rule of signs, the necessary condition for the phase difference curve to cross $\frac{\pi}{2}$ line at up to four points is:

$$
\begin{equation*}
C_{1}<C_{4}, C_{2}<C_{4}, C_{1} C_{2}+C_{3} C_{4}<C_{2} C_{4} \tag{3.44}
\end{equation*}
$$

If the phase difference curve intersects $\frac{\pi}{2}$ line at four points, then the curve has three extrema similar to the typical example shown in Fig. 3.9.

Similar to the discussion in the previous subsections, we would like to calculate the values of components given that the following equalities are met. These equalities lead to having an equal maximum magnitude of the phase difference from $\frac{\pi}{2}$ at the three extrema frequencies (namely, $\omega_{\max 1}, \omega_{\min }$, and $\omega_{\max 2}$ ) and the endpoints of the frequency range of interest, namely, $\omega_{l}$ and $\omega_{h}$.

$$
\begin{array}{r}
\left|\theta\left(\omega_{l}\right)-\frac{\pi}{2}\right|=\left|\theta\left(\omega_{\max 1}\right)-\frac{\pi}{2}\right|=\left|\theta\left(\omega_{\min }\right)-\frac{\pi}{2}\right|= \\
\left|\theta\left(\omega_{\max 2}\right)-\frac{\pi}{2}\right|=\left|\theta\left(\omega_{h}\right)-\frac{\pi}{2}\right| \tag{3.45}
\end{array}
$$

To have $\left|\theta\left(\omega_{\max 1}\right)-\frac{\pi}{2}\right|=\left|\theta\left(\omega_{\max 2}\right)-\frac{\pi}{2}\right|$, the following condition needs to


Figure 3.9: Example of the typical phase difference graph of the proposed $4^{\text {th }}$-order PAF.
be met:

$$
\begin{align*}
& R^{2} \omega_{\max 1} \times \omega_{\max 2}= \\
& \frac{\sqrt[3]{C_{4}-C_{2}}}{\sqrt[3]{C_{1} C_{2} C_{3} C_{4}\left(C_{2} C_{3} C_{4}-C_{1} C_{2} C_{4}\right)}}=  \tag{3.46}\\
& \frac{C_{2} C_{3} C_{4}-C_{1} C_{2} C_{4}}{C_{1} C_{2} C_{3} C_{4}\left(C_{4}-C_{2}\right)}=\frac{C_{4}-C_{2}}{C_{2} C_{3} C_{4}-C_{1} C_{2} C_{4}} .
\end{align*}
$$

The above condition is satisfied if and only if

$$
\begin{equation*}
C_{1} C_{4}=C_{2} C_{3} . \tag{3.47}
\end{equation*}
$$

Substitution of Eq. (3.47) into Eq. (3.46) gives:

$$
\begin{equation*}
\omega_{\max 1} \times \omega_{\max 2}=\frac{1}{R^{2} C_{1} C_{4}} \tag{3.48}
\end{equation*}
$$

$\omega_{\text {min }}$ can obtained from the relations $\frac{d \theta(\omega)}{d \omega}=0$ and $\frac{d^{2} \theta(\omega)}{d \omega^{2}}>0$ is:

$$
\begin{equation*}
\omega_{\min }=\frac{1}{R \sqrt{C_{1} C_{4}}} \tag{3.49}
\end{equation*}
$$

Moreover, to have $\left|\theta\left(\omega_{l}\right)-\frac{\pi}{2}\right|=\left|\theta\left(\omega_{h}\right)-\frac{\pi}{2}\right|$, the following condition should be met.

$$
\begin{equation*}
\omega_{l} \times \omega_{h}=\omega_{\max 1} \times \omega_{\max 2} \tag{3.50}
\end{equation*}
$$

Taking into account:

$$
\begin{equation*}
\left|\theta\left(\omega_{l}\right)-\frac{\pi}{2}\right|=\left|\theta\left(\omega_{\min }\right)-\frac{\pi}{2}\right|, \tag{3.51}
\end{equation*}
$$

we have

$$
\begin{gather*}
R^{2}\left(C_{2} C_{4}-C_{1} C_{2}-C_{3} C_{4}\right)=\frac{1}{\omega_{l}^{2}} \times\left(\frac{2 \sqrt{\alpha}+2 / \sqrt{\alpha}+2}{\alpha}\right)  \tag{3.52}\\
=\frac{1}{\omega_{l}^{2}} f_{1}(\alpha), \text { where } \alpha=\frac{\omega_{h}}{\omega_{l}} .
\end{gather*}
$$

The values of $\omega_{\max 1}$ and $\omega_{\max 2}$, which are calculated from the relation $\frac{d \theta(\omega)}{d \omega}=0$ and $\frac{d^{2} \theta(\omega)}{d \omega^{2}}<0$, are:

$$
\begin{align*}
& \omega_{\max 1, \max 2}= \\
& \omega_{h} \times \sqrt{\frac{f_{1}(\alpha)-\frac{4}{\alpha} \pm \sqrt{\left(f_{1}(\alpha)-\frac{4}{\alpha}\right)^{2}-\left(\frac{2}{\alpha}\right)^{2}}}{2}}  \tag{3.53}\\
& =\omega_{h} \times f_{2}(\alpha),
\end{align*}
$$

where minus and plus signs result in $\omega_{\max 1}$ and $\omega_{\max 2}$, respectively.
Finally, solving for

$$
\begin{equation*}
\left|\theta\left(\omega_{l}\right)-\frac{\pi}{2}\right|=\left|\theta\left(\omega_{\max 1}\right)-\frac{\pi}{2}\right| \tag{3.54}
\end{equation*}
$$

results in the following relationship:

$$
\begin{equation*}
R\left(C_{4}-C_{2}\right)=\frac{g(\alpha)}{\omega_{l}} \tag{3.55}
\end{equation*}
$$

where, $g(\alpha)$ is

$$
\begin{equation*}
\sqrt{\frac{\left(1+\alpha^{2}\left(1+f_{1}(\alpha)\right)\right)\left(f_{2}^{2}(\alpha)\left(f_{1}(\alpha)+f_{2}^{2}(\alpha)\right)+\alpha^{-2}\right)}{f_{2}(\alpha)(1+\alpha)\left(1+\alpha f_{2}^{2}(\alpha)\right)}} . \tag{3.56}
\end{equation*}
$$

The maximum magnitude of the phase difference from $\frac{\pi}{2}$ which occurs at $\omega_{l}, \omega_{\min }, \omega_{\max 1, \max 2}$, and $\omega_{h}$, can be written as:

$$
\begin{equation*}
\Delta \theta_{\max }=\left|\frac{\pi}{2}-2 \tan ^{-1} \frac{g(\alpha)+\alpha^{-1} g(\alpha)}{1+f_{1}(\alpha)+\alpha^{-2}}\right| \tag{3.57}
\end{equation*}
$$

## Design Procedure

Given $\omega_{l}$ and $\omega_{h}$, the capacitor values can be calculated using the following equations, which are obtained from Eqs. (3.47), (3.48), (3.50), (3.52) and (3.55).

$$
\begin{gather*}
C_{2}=\frac{1}{R \omega_{l}} \times\left(-\frac{g(\alpha)}{2}+\right. \\
\left.\sqrt{\frac{f_{1}(\alpha)+\frac{2}{\alpha}+\frac{g^{2}(\alpha)}{2}+\sqrt{\left(f_{1}(\alpha)+\frac{2}{\alpha}\right)^{2}+4 \frac{g^{2}(\alpha)}{\alpha}}}{2}}\right)  \tag{3.58}\\
C_{4}=C_{2}+\frac{g(\alpha)}{R \omega_{l}}  \tag{3.59}\\
C_{3}=\frac{1}{R^{2} C_{2} \omega_{l} \omega_{h}}  \tag{3.60}\\
C_{1}=\frac{C_{2} C_{3}}{C_{4}} \tag{3.61}
\end{gather*}
$$

The inductor values can be calculated using Eq. (3.41) and the values of the capacitors.

### 3.3 Design Example

In this section, we provide a design example for each of the $2^{\text {nd }}-, 3^{r d_{-}}, 4^{t h}{ }_{-}$ order PAFs that we discussed in the previous section. The frequency band of interest for all of these examples is assumed to be $\left[f_{l}, f_{h}\right]=[10 \mathrm{MHz}$, $100 \mathrm{MHz}]$. We begin with the $2^{\text {nd }}$-order PAF. Since $\omega_{\max }=2 \pi f_{\max }$ is the geometric mean of the $\omega_{l}$ and $\omega_{h}$, we can write:

$$
f_{\max }^{2}=f_{l} \times f_{h}=1000 \Longrightarrow f_{\max }=31.62 \mathrm{MHz}
$$

Then, the values of reactive components can be calculated using Eqs. (3.19) and (3.23) assuming that $\mathrm{R}=100 \Omega$.

$$
\begin{aligned}
C_{1} & =16.92 \mathrm{pF}, C_{3}=149.67 \mathrm{pF} \\
L_{1} & =169.2 \mathrm{nH}, L_{3}=1.497 \mu \mathrm{H}
\end{aligned}
$$

The maximum phase shift error, according to Eq. (3.24), is then:

$$
\Delta \theta_{\max } \approx 15.66^{\circ}
$$

For the $3^{\text {rd }}$-order PAF, the design begins with the calculation of $\omega_{\text {min }}$. Its value can be calculated from Eq. (3.39) and the corresponding value of $f_{\text {min }}=\frac{\omega_{\text {min }}}{2 \pi}$ is 18.44 MHz . For $\mathrm{R}=100 \Omega$, the values of $C_{3}, C_{2}$, and $C_{1}$ are $50.33,253$, and 10 pF , respectively. The values of $L_{3}, L_{2}$, and $L_{1}$ are $503.3 \mathrm{nH}, 2.53 \mu \mathrm{H}, 100 \mathrm{nH}$, respectively. The maximum phase shift error, according to $(3.40)$, is approximately $4.13^{\circ}$.

For the $4^{t h}$-order PAF, assuming that $\mathrm{R}=100 \Omega$, the components values are:

$$
\begin{aligned}
& C_{1}=7 \mathrm{pF}, C_{2}=91.9 \mathrm{fF}, C_{3}=27.6 \mathrm{fF}, C_{4}=362.5 \mathrm{fF} \\
& L_{1}=70 \mathrm{nF}, L_{2}=919 \mathrm{nH}, L_{3}=276 \mathrm{nH}, L_{4}=3.63 \mu \mathrm{H}
\end{aligned}
$$

and the maximum phase shift error, according to Eq. (3.57) is approximately $1.08^{\circ}$.

Fig. 3.10 shows the quadrature phase difference graphs for all PAFs based on the components values calculated in this section. The frequency range of operation is assumed to be 10 MHz to 100 MHz . As can be seen from this figure, by increasing the order of the PAF structure, one can either improve the phase accuracy of the structure over the same bandwidth of interest, or achieve the same phase accuracy over a wider bandwidth.


Figure 3.10: The quadrature phase difference graph of different order PAFs

### 3.4 Effects of Nonidealities on the Circuit Performance

In this section, the effects of nonidealities on the performance of the proposed structure are discussed.

### 3.4.1 Loading Effects

This subsection analyzes the effect of load mismatch on the quadrature outputs, as well as the effect of loading on the input impedance. In practice, the loading is caused by the input impedance of the next stage, which in CMOS designs is mainly a capacitive loading.

## Effect of Load Mismatch on the Quadrature Outputs

It can be shown that the input and output voltages of the structure of Fig. 3.1, when the load impedance of $Z_{L}$ is connected to the output, are related to each other as follows:

$$
\begin{equation*}
\frac{V_{o}(s)}{V_{i n}(s)}=\frac{Z_{L}}{Z_{L}+R} \tag{3.62}
\end{equation*}
$$

Note that the output voltage is independent of the reactive component values of the network providing that Eq. (3.1) is satisfied.

For the system with two quadrature outputs, the quadrature amplitude mismatch is defined as below:

$$
\begin{equation*}
\text { Quadrature Mismatch }(\mathrm{QM})=\left|\frac{V_{I}}{V_{Q}}\right| \tag{3.63}
\end{equation*}
$$

So,

$$
\begin{equation*}
\mathrm{QM}=\left|\frac{Z_{L 1}}{Z_{L 2}} \times \frac{Z_{L 2}+R}{Z_{L 1}+R}\right| \tag{3.64}
\end{equation*}
$$

where $Z_{L 1}$ and $Z_{L 2}$ are the load impedances seen by $I$ - and $Q$ - terminals (Fig. 3.2), respectively. Assuming that

$$
\begin{aligned}
Z_{L 1} & =Z_{c m}+\frac{\Delta Z}{2} \\
Z_{L 2} & =Z_{c m}-\frac{\Delta Z}{2}
\end{aligned}
$$

where

$$
Z_{c m}=\frac{Z_{L 1}+Z_{L 2}}{2}
$$

and

$$
\Delta Z=Z_{L 1}-Z_{L 2},
$$

Eq. (3.64) can be approximated as follows providing that $\left|\frac{\Delta Z}{Z_{c m}}\right| \ll 1$.

$$
\begin{equation*}
\mathrm{QM} \approx\left|1+\frac{z_{r}}{1+\frac{Z_{c m}}{R}}\right| \quad \text { where } z_{r}=\frac{\Delta Z}{Z_{c m}} \tag{3.65}
\end{equation*}
$$

The above equation is verified for a capacitive load of $C_{c m}=20 \mathrm{pF}$ shown in Fig. 3.11. The quadrature amplitude mismatch between the output terminals remains around zero assuming that $z_{r}$ stays around zero. The reason is that the subsequent stages connected to the $I$ - and $Q$ - terminals of PAF are ideally identical and thus the deviation between their input capacitances is typically small (assuming a reasonable layout design).


Figure 3.11: The quadrature mismatch in terms of parasitic capacitance.

## Effect of Load on the Input Impedance

Since $z_{r}$ is typically small, that is the input impedance of the stages connected to $I$ and $Q$ are ideally identical, the input impedance of the network of Fig. 3.2 in the presence of the loading can be calculated as follows:

$$
\begin{equation*}
Z_{i n, e q}=\frac{0.5 R\left(Z_{L}+R\right)}{Z_{L}+R\left(\frac{R^{2}+f^{2}(s)}{(R+f(s))^{2}}+\frac{R^{2}+g^{2}(s)}{(R+g(s))^{2}}\right)} . \tag{3.66}
\end{equation*}
$$

It can be shown that the expression into parenthesis in the denominator of Eq. 3.66 remains around one. So, if $\left|\frac{R}{Z_{L}}\right| \ll 1$, an approximation for $Z_{i n, e q}$ can be expressed as:

$$
\begin{align*}
& Z_{i n, e q} \approx \\
& 0.5 R\left(1-\frac{R}{Z_{L}}\left(\frac{\left(R^{2}-f(s) g(s)\right)^{2}+R^{2}(f(s)-g(s))^{2}}{(R+f(s))^{2}(R+g(s))^{2}}\right)\right. \tag{3.67}
\end{align*}
$$

where $f(s)$ and $g(s)$ are the input impedances of the networks $\mathcal{N}_{1}$ and $\mathcal{N}_{3}$, respectively.

### 3.4.2 Effects of Component Value Deviation

In this section we analyze the IQ amplitude mismatch and input impedance error originated from component value deviation. Here, we assume that instead of Eq. [3.1, due to component value deviations, we have:

$$
\begin{equation*}
\frac{L}{C}=(R+\Delta R)^{2}=R^{2}\left(1+\frac{\Delta R}{R}\right)^{2} . \tag{3.68}
\end{equation*}
$$

That is, any capacitor or inductor value variation makes the above ratio deviate from its nominal value $\left(R^{2}\right)$. Assuming that in Fig. 3.2, $x_{i}=\frac{\Delta R_{i}}{R}, i \in$ $\{1,2,3,4\}$ captures the deviation of $\frac{L}{C}$ for each of the networks $\mathcal{N}_{i}, i \in$ $\{1,2,3,4\}$, the approximations for the output and input impedance error of one of the arms of the $4^{\text {th }}$-order PAF can be found from Eqs. (3.69) and (3.70), where $V_{1}$ and $V_{3}$ are equal to $V_{I}$ and $V_{Q}$, respectively. In this analysis, to keep the formulas more tractable and without loss of generality we have assumed that inductances have their nominal values and only capacitances deviate from their nominal values so that $\frac{L}{C}$ deviates from $R^{2}$.

$$
\begin{align*}
& V_{i} \approx \frac{1-\frac{2 x_{i}\left(R^{2} C_{i+1}^{2} s^{2}-1-R^{2} C_{i} C_{i+1} s^{2}\right)-2 x_{i+1}\left(1+R^{2} C_{i} C_{i+1} s^{2}\right)-4 x_{i} x_{i+1}}{\left(1+R^{2} C_{i} C_{i+1} s^{2}\right)^{2}-R^{2} C_{i+1}^{2} s^{2}}}{\left(1+\frac{2 x_{i+1}}{1+R^{2} C_{i} C_{i+1} s^{2}+R C_{i+1} s}\right)\left(1+\frac{2 x_{i}\left(1+R C_{i+1} s\right)}{1+R^{2} C_{i} C_{i+1} s^{2}+R C_{i+1} s}\right)} \times V_{i n} \\
& \text { if } x_{i}=\frac{\Delta R_{i}}{R} \text { and } x_{i+1}=\frac{\Delta R_{i+1}}{R} \ll 2 \tag{3.69}
\end{align*}
$$

$$
\begin{align*}
& \Delta Z_{i} \approx \frac{\frac{R^{2} C_{i+1} s\left(2 R^{2} C_{i} C_{i+1} s^{2} x_{i}+2 x_{i+1}\left(1+2 x_{i}\right)\right)}{\left(1+R C_{i+1} s+R^{2} C_{i} C_{i+1} s^{2}\right)^{2}}}{1+\frac{\left(R^{2} C_{i} C_{i+1} s^{2}+R^{2} C_{i+1}^{2} s^{2}+2 R C_{i+1} s+1\right) 2 x_{i}+\left(R^{2} C_{1} C_{i+1} s^{2}+2 x_{i}+1\right) 2 x_{i+1}}{\left(1+R C_{i+1} s+R^{2} C_{i} C_{i+1} s^{2}\right)^{2}}} \\
& \text { if } x_{i}=\frac{\Delta R_{i}}{R} \text { and } x_{i+1}=\frac{\Delta R_{i+1}}{R} \ll 2 \tag{3.70}
\end{align*}
$$

Based on Eqs. (3.69) and (3.70), the quadrature mismatch of the $2^{\text {nd }}$ order PAF can be calculated as:

$$
\begin{equation*}
\left.Q M \approx\left|\lim _{\substack{C_{2} \rightarrow \infty \\ C_{4} \rightarrow \infty}} \frac{V_{1}}{V_{3}}\right|_{\substack{\Delta R_{2}=0 \\ \Delta R_{4}=0}} \right\rvert\,, \tag{3.71}
\end{equation*}
$$

and its input impedance is:

$$
\begin{equation*}
Z_{i n, e q} \approx \lim _{\substack{C_{2} \rightarrow \infty \\ C_{4} \rightarrow \infty}}\left(R+\Delta Z_{1}\right) \|\left.\left(R+\Delta Z_{3}\right)\right|_{\substack{\Delta R_{2}=0 \\ \Delta R_{4}=0}} \tag{3.72}
\end{equation*}
$$

Fig. 3.12 shows the analytical and simulated curves of the quadrature mismatch for the $2^{\text {nd }}$-order filter over the frequency range of 10 MHz to 100 MHz . The curves are in terms of $\Delta R$ (with the assumption that $\Delta R_{1}=$ $\left.\Delta R_{2}=\Delta R\right)$ at the centre frequency of 55 MHz . The simulations are done using Keysight Advance Design System (ADS).

For the $2^{\text {nd }}$-order PAF, if $\Delta R_{1,2} \ll \frac{R}{2}$, the closed forms of $Q M$ and $Z_{i n, e q}$ can be simplified as:

$$
\begin{equation*}
r C l Q M \approx\left|1-\frac{2 C_{1} s \Delta R_{1}}{\left(1-R^{2} C_{1}^{2} s^{2}\right)}+\frac{2 C_{2} s \Delta R_{2}}{\left(1-R^{2} C_{2}^{2} s^{2}\right)}\right| \tag{3.73}
\end{equation*}
$$

and

$$
\begin{equation*}
r C l Z_{i n, e q} \approx \frac{R}{2}\left(1+\frac{C_{1} s \Delta R_{1}}{\left(1+R C_{1} s\right)^{2}}+\frac{C_{2} s \Delta R_{2}}{\left(1+R C_{2} s\right)^{2}}\right) \tag{3.74}
\end{equation*}
$$

The quadrature mismatch and input impedance of the $3^{r d}$-order PAF can be calculated from Eqs. (3.69) and (3.70) as:

$$
\begin{gather*}
\left.Q M \approx\left|\lim _{C_{4} \rightarrow \infty} \frac{V_{1}}{V_{3}}\right|_{\Delta R_{4}=0} \right\rvert\,  \tag{3.75}\\
Z_{i n, e q} \approx \lim _{C_{4} \rightarrow \infty}\left(R+\Delta Z_{1}\right) \|\left.\left(R+\Delta Z_{3}\right)\right|_{\Delta R_{4}=0} \tag{3.76}
\end{gather*}
$$



Figure 3.12: $2^{\text {th }}$-order PAF errors in terms of $\mathrm{x}\left(=\frac{\Delta R}{R}\right)$, which x is a measure showing how much Eq. (3.1) deteriorates. a) Quadrature mismatch (dB). b) Input impedance variation $\left(100 \times \frac{Z_{i n, e q}-Z_{i n, i d e a l}}{Z_{i n, \text { ideal }}}\right)$

As mentioned earlier, for the $4^{\text {th }}$-order PAF we have:

$$
\begin{gather*}
Q M \approx\left|\frac{V_{1}}{V_{3}}\right|  \tag{3.77}\\
Z_{i n, e q} \approx\left(R+\Delta Z_{1}\right) \|\left(R+\Delta Z_{3}\right) . \tag{3.78}
\end{gather*}
$$

The validity of Eq. (3.75)-(3.78) is verified and the results are shown in Figs. 3.13 and 3.14. As confirmed by Figs. 3.12 to 3.14 , the proposed structure is relatively insensitive to its component value variations. For instance, the quadrature mismatch of $2^{\text {nd }}, 3^{\text {rd }}$ and $4^{\text {th }}$-order PAF remains
below 0.21 dB for $\frac{\Delta R}{R}$ of up to $\pm 20 \%$. It translates into the capacitive deviation of $[-30 \%$ to $+56 \%]$.


Figure 3.13: $3^{r d}$-order PAF errors in terms of $\mathrm{x}\left(=\frac{\Delta R}{R}\right)$. a) Quadrature mismatch (dB). b) Input impedance variation $\left(100 \times \frac{Z_{\text {in,eq }}-Z_{\text {in,ideal }}}{Z_{i n, \text { ideal }}}\right)$

### 3.4.3 Influence of the Limited Quality Factor of inductors on the QM and Input Impedance of the Proposed Structure

Since the quality factor $(\mathrm{Q})$ of the on-chip inductors is typically limited, in this subsection, the effect of inductor Q on PAF performance is investigated.

Similar to the previous subsection, the differential output voltage ( $V_{I}$ or $\left.V_{Q}\right)$ and input impedance error of one of the arms of the $4^{t h}$-order PAF


Figure 3.14: $4^{t h}$-order PAF errors in terms of $\mathrm{x}\left(=\frac{\Delta R}{R}\right)$. a) Quadrature mismatch $(\mathrm{dB})$. b) Input impedance variation $\left(100 \times \frac{Z_{\text {in,eq }}-Z_{\text {in,ideal }}}{Z_{\text {in,ideal }}}\right)$
containing ideal capacitors and inductors with limited Q are expressed in Eqs. (3.79) and (3.80).
$V_{i}=$
$\frac{1-\frac{\frac{R^{2} C_{i+1} \omega s}{Q_{i+1}}\left(C_{i+1}-C_{i}-R^{2} C_{i}^{2} C_{i+1} s^{2}\right)-\frac{R^{4} C_{i}^{2} C_{i+1}^{2} \omega^{2} s^{2}}{Q_{i} Q_{i+1}}-\frac{R^{2} C_{i} C_{i+1} \omega s\left(1+R^{2} C_{i} C_{i+1} s^{2}\right)}{Q_{i}}}{\left(1+R^{2} C_{i} C_{i+1} s^{2}\right)^{2}-R^{2} C_{i+1}^{2} s^{2}}}{\left(1+\frac{R^{2} C_{i} C_{i+1} \omega s}{Q_{i}\left(1+R^{2} C_{i} C_{i+1} s^{2}+R C_{i+1} s\right)}\right)\left(1+\frac{R C_{i+1} \omega\left(1+R C_{i} s\right)}{Q_{i+1}\left(1+R^{2} C_{i} C_{i+1} s^{2}+R C_{i+1} s\right)}\right)} \times V_{i n}$ $(s=j w)$

$$
\begin{align*}
& \Delta Z_{i}= \\
& \frac{\frac{\frac{R^{2} C_{i+1} \omega}{Q_{i+1}}+\frac{R^{2} C_{i} C_{i+1} \omega s}{Q_{i}}\left(\frac{R^{2} C_{i+1} \omega}{Q_{i+1}}+R^{2} C_{i+1} s\right)}{\left(1+R C_{i+1} s+R^{2} C_{i} C_{i+1} s^{2}\right)^{2}}}{1+\frac{\frac{R^{2} C_{i+1}^{2} \omega s}{Q_{i+1}}+R^{2} C_{i} C_{i+1} \omega s\left(\frac{1}{Q_{i}}+\frac{1}{\left.Q_{i+1}+\frac{R^{2} C_{i} C_{i+1} s^{2}}{Q_{i}}+\frac{R^{2} C_{i} C_{i+1} \omega s}{Q_{i} i_{i+1}}+\frac{2 R C_{i+1} s}{Q_{i+1}}+\frac{R^{2} C_{i} C_{i+1} s^{2}}{Q_{i+1}}\right)}\right.}{\left(1+R C_{i+1} s+R^{2} C_{i} C_{i+1} s^{2}\right)^{2}}} \\
& (s=j w)
\end{align*}
$$

The I/Q mismatch and the input impedance of the $2^{\text {nd }}$-order PAF can be derived from Eqs. (3.79) and (3.80) as follows:

$$
\begin{equation*}
Q M=\left|\frac{\lim _{C_{2} \rightarrow \infty} \lim _{Q_{2} \rightarrow \infty} V_{1}}{\lim _{C_{4} \rightarrow \infty} \lim _{Q_{4} \rightarrow \infty} V_{3}}\right|, \tag{3.81}
\end{equation*}
$$

and

$$
\begin{equation*}
Z_{i n, e q}=\lim _{\substack{C_{2} \rightarrow \infty \\ C_{4} \rightarrow \infty \\ Q_{2} \rightarrow \infty \\ Q_{4} \rightarrow \infty}} \lim _{1}\left(R+\Delta Z_{1}\right) \|\left(R+\Delta Z_{3}\right) . \tag{3.82}
\end{equation*}
$$

The accuracy of the above relations are verified and the results are shown in Fig. 3.15.

With the realistic assumption of $Q_{i} \geq 5, i \in\{1,2\}$, the equations for the $2^{\text {nd }}$-order PAF can be simplified as:

$$
\begin{equation*}
Q M \approx 1-\frac{R C_{1} \omega}{Q_{1}\left(1+R^{2} C_{1}^{2} \omega^{2}\right)}+\frac{R C_{2} \omega}{Q_{2}\left(1+R^{2} C_{2}^{2} \omega^{2}\right)}, \tag{3.83}
\end{equation*}
$$

and

$$
\begin{equation*}
Z_{i n, e q} \approx \frac{R}{2}\left(1+\frac{R C_{1} \omega}{2 Q_{1}\left(1+j R C_{1} \omega\right)^{2}}+\frac{R C_{2} \omega}{2 Q_{2}\left(1+j R C_{2} \omega\right)^{2}}\right), \tag{3.84}
\end{equation*}
$$

Eqs. (3.85) to (3.88) express QM and input impedance of the $3^{r d}$ and $4^{\text {th }}$-order PAF, respectively. For the $3^{\text {rd }}$-order PAF, we have:

$$
\begin{equation*}
Q M=\left|\lim _{C_{4} \rightarrow \infty} \lim _{Q_{4} \rightarrow \infty} \frac{V_{1}}{V_{3}}\right|, \tag{3.85}
\end{equation*}
$$

and

$$
\begin{equation*}
Z_{i n, e q}=\lim _{C_{4} \rightarrow \infty} \lim _{Q_{4} \rightarrow \infty}\left(R+\Delta Z_{1}\right) \|\left(R+\Delta Z_{3}\right) \tag{3.86}
\end{equation*}
$$



Figure 3.15: $2^{\text {th }}$-order PAF errors in terms of Q , which Q is the quality factor of on-chip inductors. a) Quadrature mismatch (dB). b) Input impedance variation $\left(100 \times \frac{Z_{i n, e q}-Z_{\text {in,ideal }}}{Z_{\text {in, ideal }}}\right)$

For the $4^{\text {th }}$-order PAF, the relevant equations are:

$$
\begin{equation*}
Q M=\left|\frac{V_{1}}{V_{3}}\right|, \tag{3.87}
\end{equation*}
$$

and its input impedance is:

$$
\begin{equation*}
Z_{\text {input }}=\left(R+\Delta Z_{1}\right) \|\left(R+\Delta Z_{3}\right) . \tag{3.88}
\end{equation*}
$$

Figs. 3.15 to 3.17 confirm that the performance of the proposed structure is relatively insensitive to the quality factor of the inductors.


Figure 3.16: $3^{r d}$-order PAF errors in terms of Q. a) Quadrature mismatch $(\mathrm{dB})$. b) Input impedance variation $\left(100 \times \frac{Z_{i n, e q}-Z_{i n, \text { ideal }}}{Z_{i n, \text { ideal }}}\right)$

### 3.5 Experimental Results

In this section, we present the results of experimental verification of a $4^{t h}{ }_{-}$ order PA (Fig. 3.8). The prototype is implemented on a printed circuit board and using discrete components, and the operating frequency is chosen to be between 10 to 100 MHz , a ratio of 1 to 10 . The prototype is shown in Fig. 3.18. Table 5.1 provides the component values. The performance of the circuit is measured using Keysight E5061B Network Analyzer and an active probe enabling the output terminals to be connected to the Network Analyzer with minimal loading effect. Fig. 3.19 shows the $S_{11}$, which is lower than -25 dB over the band of interest. Figs. 3.20 and 3.21 reflect


Figure 3.17: $4^{\text {th }}$-order PAF errors in terms of Q. a) Quadrature mismatch (dB). b) Input impedance variation $\left(100 \times \frac{Z_{\text {in,eq}}-Z_{\text {in, ideal }}}{Z_{\text {in, ideal }}}\right)$
the I/Q network characteristics. The dashed and solid curves correspond to simulations and measurements, respectively. Based on these figures, there is a good agreement between measurement and simulation results.

### 3.6 Summary

In this chapter, a structure for an $n^{\text {th }}$-order all-pass filter (PAF) is proposed that can generate fairly accurate quadrature signals while alleviating the problems of the conventional structures including poor input matching, and the trade-off between the accuracy of quadrature phase difference and insertion loss. The effects of loading, component value deviations and lim-


Figure 3.18: The photograph of I/Q generator network.
Table 3.1: Component values based on the structure of Fig. 3.8

| Component Values |  |  |  |
| :--- | :--- | :--- | :--- |
| Num (tol.) | $\mathrm{C}($ tol.) $)$ | $\mathrm{R}($ tol. $)$ |  |
| 1 | $68(2 \%) \mathrm{nH}$ | $5.9(2 \%) \mathrm{pF}$ | $108(<5 \%) \Omega$ |
| 2 | $900(5 \%) \mathrm{nH}$ | $100(2 \%) \mathrm{pF}$ | $93(<5 \%) \Omega$ |
| 3 | $270(2 \%) \mathrm{nH}$ | $25(2 \%) \mathrm{pF}$ | $105(<5 \%) \Omega$ |
| 4 | $3.5(2 \%) \mu \mathrm{H}$ | $350(5 \%) \mathrm{pF}$ | $102(<5 \%) \Omega$ |

ited inductor quality factor on the performance of the proposed structure are studied analytically, and the results are confirmed by simulations. The proposed structure is relatively insensitive to component variations. The performance of the proposed structure is confirmed using a proof-of-concept prototype and the measurement and simulation results are in good agreement.


Figure 3.19: Input return loss of the $\mathrm{I} / \mathrm{Q}$ network.


Figure 3.20: $I$ - and $Q$ - magnitudes.


Figure 3.21: Quadrature error characteristics of the I/Q network. (a) I/Q phase difference. (b) I/Q amplitude mismatch.

## Chapter 4

## Analysis and Design of a Wideband mm-Wave Miniaturized Marchand Balun with a Better than $0.4-\mathrm{dB}$ Amplitude and $1.5^{\circ}$ Phase Mismatch

The majority of sensitive analog and radio-frequency (RF) circuits and systems transmit/receive the signal in a balanced manner as differential signalling reduces the effects of the noise and interferences coupled to the signal lines as well as higher order (particularly even-order) harmonics. Such signalling, in turn, enhances the dynamic range of the overall system. A balun is a typical component that converts an unbalanced signal, e.g., the received single-ended antenna signal, to a balanced version. In wireless communication systems, the balun is a key component for realizing building blocks such as double-balanced mixers, balanced amplifiers, frequency multipliers, and balanced antennas [68]-[70].

Among various types of active or passive baluns [71] - [73], the distributedtype passive ones [5], [74] are more desired for the broadband applications since they generally exhibit a better performance in terms of amplitude and phase imbalance over a broader bandwidth. In multi-GHz or mm-wave applications, distributed-type baluns are generally used. Such distributed-type baluns are typically categorized into hybrid $180^{\circ}$ (rat-race) and Marchand [74] baluns. Distributed-type baluns occupy a relatively large area. For example, although the hybrid $180^{\circ}$ baluns have a fairly good frequency response in the microwave frequency band, they require line lengths of $3 \lambda / 4$ and $\lambda / 4$ where $\lambda$ is the wavelength of the signal. On the other hand, the commonly used Marchand balun consists of two identical sections of coupled
lines with the length of $\lambda / 4$. Furthermore, in contrast to the hybrid $180^{\circ}$, both amplitude and phase imbalance of Marchand baluns are ideally zero independent of frequency. Nonetheless, they still occupy a relatively large area due to the use of two quarter-wavelength coupled lines. Moreover, using long-length lines on chip causes performance degradation at higher frequencies, e.g., millimeter wave (mm-wave), due to increased losses such as substrate loss.

Many efforts [75], [76] have been made to reduce the size of the conventional Marchand balun to make them smaller and lower cost with better electrical performance at higher frequencies. However, the main issue that almost all the techniques suffer from is the trade-off between the bandwidth and size. In fact, although such techniques offer smaller size design, the bandwidth of the balun is also reduced as compared to the original Marchand balun.

In this paper, the bandwidth performance of both Marchand balun and its reduced-size counterpart, also referred to as capacitively loaded balun, is analyzed by means of contour integration. It is shown that the conventional approach to the design of the reduced sized Marchand balun is not optimal. Moreover, the analysis shows that the capacitively loaded balun for some specific loads exhibits a broader bandwidth than that of Marchand balun. Finally, a methodology for the optimal design of the reduced-size Marchand baluns is proposed which leads to a balun with both improved bandwidth and reduced size, where the balun has also less sensitivity to load variations.

The organization of the paper is as follows. Section 4.1 describes the bandwidth extension of the reduced-size capacitively loaded Marchand balun (miniaturized Marchand). The proposed methodology to obtain the optimal performance from the structure is described in Section 4.2. The design of the balun using the proposed approach is described in Section 4.3. Section 4.4 presents the experimental results.

### 4.1 Bandwidth Extension of Miniaturized Marchand Baluns

### 4.1.1 Brief Discussion on the Capacitively Loaded Marchand Baluns

Fig. 4.1(a) shows the structure of the conventional Marchand balun. Replacing the quarter-wavelength lines of Fig. 4.1(a) with its capacitively loaded equivalent shown in Fig. 4.1(b) [77] and also further manipulating the con-


Figure 4.1: Balun structure, (a) conventional Marchand balun, (b) capacitive loading technique to reduce the line length, (c) reduced-size Marchand balun.
ventional structure, one can arrive at the structure shown in Fig. 4.1(c). In fact, the length of the segments can be reduced by means of the capacitors connected to both ends of the coupled lines. The component values are:

$$
\begin{gather*}
C_{1}=\frac{\cos \left(\theta^{\prime}\right)}{\omega_{0} Z_{e}}, C_{2}=\frac{\cos \left(\theta^{\prime}\right)}{2 \omega_{0} Z_{o}}-0.5 C_{1}, \\
C_{5}=C_{1}, C_{3}=C_{4}=C_{1}+C_{2},  \tag{4.1}\\
Z_{e}^{\prime}=\frac{Z_{e}}{\sin \left(\theta^{\prime}\right)}, Z_{o}^{\prime}=\frac{Z_{o}}{\sin \left(\theta^{\prime}\right)}, \tag{4.2}
\end{gather*}
$$

where $\omega_{0}=2 \pi f_{0}, \theta^{\prime}, Z_{e}$ and $Z_{o}$ are the center frequency, desired electrical length(within ( $0, \pi / 2]$ ), and even- and odd-mode characteristic impedances of the coupled lines in Fig. 4.1(a), respectively.

However, as the segments length shrinks, the bandwidth is also reduced. To show the trade-off between the bandwidth and size, one can show that the insertion loss $\left(S_{21}\right)$ has at least one notch in the interval $\left[f_{0}, 2 f_{0}\right]$. The reason why the upper bound of frequency is considered to be $2 f_{0}$ is that the first notch in the insertion loss of the conventional Marchand structure (Fig. 4.1(a)) designed for center frequency of $f_{0}$ occurs at $2 f_{0}$. In fact, the notch frequencies are at $k \pi f_{0} / \theta_{0}=2 k f_{0}$ where $k \in Z$ and $\theta_{0}=\pi / 2$
is the electrical length of the segment at center frequency of $f_{0}$. Thus, to prove that the capacitively loaded balun designed at the center frequency of $f_{0}$ has a smaller bandwidth as compared to the conventional Marchand balun, it should be shown that its notch frequency is smaller than $2 f_{0}$. Using Eqs. (B.7), (4.1) and (4.2), the notch frequency of the capacitively loaded structure can be obtained from the following equation:

$$
\begin{equation*}
\frac{f}{f_{0}}=\frac{\cot \left(0.5 \theta^{\prime}\right)}{\cot \left(\theta_{0}^{\prime}\right)} \tag{4.3}
\end{equation*}
$$

where $\theta_{0}^{\prime}$ and $\theta^{\prime}$ are the electrical length of the reduced-size coupled lines at the center frequency $f_{0}$, and $\theta_{0}^{\prime} f / f_{0}$, respectively. To complete the proof, we define the function $g(f)$ as:

$$
\begin{equation*}
g(f)=\frac{f}{f_{0}}-\frac{\cot \left(0.5 \theta_{0}^{\prime} f / f_{0}\right)}{\cot \left(\theta_{0}^{\prime}\right)} . \tag{4.4}
\end{equation*}
$$

Since $g(f)$ is continuous over the interval $\left[f_{0}, 2 f_{0}\right]$ and also $g\left(f_{0}\right) \times g\left(2 f_{0}\right)<$ 0 , there is always at least one root in the interval $\left[f_{0}, 2 f_{0}\right]$. Moreover, because $g^{\prime}(f)>0$ in the interval $\left[f_{0}, 2 f_{0}\right]$, then $g(f)$ has only one root in this interval. It can also be shown that since $g\left(f_{0}\right)$ and $g\left(\sqrt{2} f_{0}\right)$ are both negative quantities and thus $g\left(f_{0}\right) \times g\left(\sqrt{2} f_{0}\right)>0$, the root always remains greater than $\sqrt{2} f_{0}$ for $0<\theta_{0}^{\prime}<\pi / 2$. Fig. 4.2 confirms the aforementioned results and also shows that the bandwidth is inversely proportional to the electrical length of the segment.


Figure 4.2: Normalized notch frequency of the capacitively loaded balun versus the electrical length of the coupled lines (Eq. (4.3)).

In the next subsection, we will show that the direct dependency of the capacitance values $C_{3}, C_{4}$ and $C_{5}$ on $C_{1}$ and $C_{2}$ is the main cause of the bandwidth degradation as the segment length is reduced. In fact, we will show that the structure can even exhibit a boarder bandwidth than that of the conventional Marchand balun under specific load conditions and with only a minor manipulation of the structure.

### 4.1.2 Capability to Improve the Bandwidth of the Capacitively Loaded Balun

## Some relevant definitions

Here, the objective is to achieve the maximally flat bandwidth of the capacitively loaded balun. The balun behaves similar to a band-pass filter with the lower and upper 3 dB frequencies of $f_{l}$ and $f_{h}$, respectively. One possible approach to extend the bandwidth of a network is to increase the area under the curve of the $S_{21}$ (or insertion loss) response. Before discussing the approach for maximizing this area, let us review an interesting property of the lossless networks, that is, at any port of an $N$-port lossless network, the sum of the reflected power at that port and the transmitted power from that port to all other ports of the network is equal to one, as presented below [5].

$$
\begin{equation*}
\sum_{k=1}^{N}\left|S_{k i}\right|^{2}=1, \quad 0 \leq\left|S_{k i}\right| \leq 1 \tag{4.5}
\end{equation*}
$$

where $S_{k i}$ are the scattering parameters of the network, $N$ is the total number of ports, and indices $i$ and $k$ refer to the port number $(1 \leq i, k \leq N)$. Specifically, for the two-port lossless network shown in (Fig. 4.1(c)), assuming that the port 1 is the single-ended input node ( $I N$ ), and port 2 is the differential output $\left(O_{1}-O_{2}\right)$, we have $\left|S_{11}\right|^{2}+\left|S_{21}\right|^{2}=1$. From this property, to maximize $\left|S_{21}\right|$ one could minimize $\left|S_{11}\right|$. Fig. 4.3(a) shows the frequency response for $\left|S_{11}\right|$ of a generic band-pass filter. As shown, for a properly designed filter, $\left|S_{11}\right|$ reaches zero at the center frequency of the filter and gets back to one for the frequencies below and above the lower and upper 3 dB cut-off frequencies. Given that the value of $\left|S_{11}\right|$ over the bandwidth of the filter should be 0 or near 0 , to maximize the bandwidth of the filter one can maximize the area under $1 /\left|S_{11}\right|$ over the frequency range of $f_{l}$ to $f_{h}$. Let us define the following quantity:

$$
\begin{equation*}
\int_{0}^{+\infty} \ln \frac{1}{\left|S_{11}(j \omega)\right|} d \omega \approx \int_{\omega_{l}}^{\omega_{h}} \ln \frac{1}{\left|S_{11}(j \omega)\right|} d \omega, \tag{4.6}
\end{equation*}
$$

where $\ln$ is the natural logarithm. Due to the use of $\left|S_{11}^{-1}\right|$ in (4.6) the integrand is greater than zero. The generic profile of this integrand is shown in Fig. 4.3(b). Given that the desired value of $\left|S_{11}(j \omega)\right|$ outside the bandwidth of the filter is 1 , the value of integrand outside the band of interest is 0 or close to zero. Therefore, the larger the bandwidth of the filter is, the larger the value of the integral (i.e., the area under $\ln \left|S_{11}^{-1}\right|$ ) will be. As


Figure 4.3: The typical frequency response of a band-pass filter: (a) $\left|S_{11}\right|$ (lumped structure), (b) $\ln \left|S_{11}\right|^{-1}$ (lumped one), (c) $\left|S_{11}\right|$ (distributed structure), (d) $\ln \left|S_{11}\right|^{-1}$ (distributed one)
it is shown in Appendix D, for a typical bandpass filter (as well as lowpass filters) the integral is bounded and its value can be obtained using complex integration. Note that the response shown in Fig. 4.3(a) is based on the assumption of the circuit having lumped components. If the filter contains distributed components, the response will be periodic. Therefore, to make sure that the integral in (4.6) is a finite quantity for any group of bandpass or low-pass filters consisting of lumped or distributed components, the integrand is modified as follows,

$$
\begin{equation*}
\int_{0}^{+\infty} \frac{1}{1+\left(\frac{\omega}{\omega_{n o}}\right)^{2 n}} \ln \frac{1}{\left|S_{11}(j \omega)\right|} d \omega \tag{4.7}
\end{equation*}
$$

where $\omega_{n o}$ and $n$ are the frequency of the first notch of $\left|S_{11}\right|$ and an integer value of greater than 0 , respectively. In effect, the original response is lowpass filtered to select the fundamental portion of the response and attenuate the rest. That is, $\left(1+\left(\frac{\omega}{\omega_{n o}}\right)^{2 n}\right)^{-1}$, serves as a low-pass filter of order $n$ that selects the fundamental peak and attenuates the rest so that the overall integral value is finite.

## Bandwidth enhancement of the capacitively loaded balun

Since this work mostly focuses on improving the bandwidth performance of the miniaturized Marchand balun, the main objective of this subsection is the performance analysis of the capacitively loaded balun whose lines have small electrical lengths, e.g. $\leq 10^{\circ}$. Referring to the structure of Fig. 4.1(c), to simplify the analysis and reduce the number of the unknown variables, one can start from a bare minimum structure that has minimum number of capacitors. Thus, we assume $C_{1}$ and therefore $C_{2}$ are zero ( $C_{1}$ and $C_{2}$ are dependent as proved in Appendix (C) and $C_{4}=\infty$. Note that $C_{3}$ and $C_{5}$ represent the input and output capacitances of the balun so in general they have a nonzero finite value. Also, note that the assumption of $C_{4}$ being zero is not a good choice, because that means the right port in Fig. 4.1(c) is open circuit and therefore input impedance matching will be challenging (mainly reactive) given that the line length is small. Therefore, for simplifying assumption we start with a short or $C_{4}=\infty$. It should be noted that with these simplifying assumptions we will show that the structure offers a competitive performance as compared to the case where the components are calculated based on either Eqs. (4.1) and (4.2) or that of the conventional Marchand balun. Thus, one can expect that the performance of the system may be further improved if one has the option of choosing values other than 0 and $\infty$ for $C_{1}, C_{2}$, and $C_{4}$.

The above-mentioned simplifying assumptions are shown in Fig. 4.4(a). From this figure, the input admittance can be calculate by:

$$
\begin{align*}
& Y_{\text {in }}=j \omega C_{3}+0.5 Y_{\text {odd }}^{\prime}+0.5 Y_{\text {even }}^{\prime}= \\
& j \omega C_{3}+0.5\left(\frac{Y_{11} Y_{L}^{\prime}+Y_{11}^{2}-Y_{41}^{2}}{Y_{L}^{\prime}+Y_{11}}\right)+0.5\left(Y_{11}-\frac{Y_{21}^{2}}{Y_{11}}\right) \tag{4.8}
\end{align*}
$$



Figure 4.4: The approach to the calculation of the input admittance of the capacitively loaded balun structure: (a) complete schematic assuming that $C_{1}=C_{2}=0$ and $C_{4}=\infty$, (b) odd-mode stimulation, (c) even-mode stimulation.
where

$$
\begin{align*}
Y_{11} & =-j 0.5 \cot \theta^{\prime}\left(Z_{e}^{\prime-1}+Z_{o}^{\prime-1}\right) \\
Y_{21} & =\frac{j 0.5}{\sin \theta^{\prime}}\left(Z_{e}^{\prime-1}+Z_{o}^{\prime-1}\right)  \tag{4.9}\\
Y_{41} & =\frac{j 0.5}{\sin \theta^{\prime}}\left(Z_{e}^{\prime-1}-Z_{o}^{\prime-1}\right), Y_{L}^{\prime}=R_{L}^{-1}+j \omega C_{5}
\end{align*}
$$

As pointed out earlier, assuming that the electrical length $\theta^{\prime}$ is small, e.g., $\theta^{\prime} \leq 10^{\circ}, \cot \theta^{\prime}$ and $\sin \theta^{\prime}$ can be approximated by $\theta^{\prime-1}$ and $\theta^{\prime}$, respectively. It can be shown that by using complex frequency $s$, the coupled line Yparameters of Eqs. 4.9 can be approximated:

$$
\begin{equation*}
Y_{11} \approx \frac{1}{s L\left(1-k^{2}\right)}, Y_{21} \approx \frac{-1}{s L\left(1-k^{2}\right)}, Y_{41} \approx \frac{k}{s L\left(1-k^{2}\right)} \tag{4.10}
\end{equation*}
$$

where $L$ and $k$ are:

$$
\begin{equation*}
\frac{l}{2 c}\left(Z_{e}^{\prime}+Z_{o}^{\prime}\right), \frac{Z_{e}^{\prime}-Z_{o}^{\prime}}{Z_{e}^{\prime}+Z_{o}^{\prime}} \tag{4.11}
\end{equation*}
$$

and, $l$ and $c$ are the length of the coupled line and the speed of light, respectively. By substituting Eq. (4.10) into Eq. (4.9), and rearranging it and also using $\left(Y_{i n}^{-1}+R_{0}\right) /\left(Y_{i n}^{-1}-R_{0}\right)$ for $\Gamma^{-1}$, the inverse of the input return loss for the capacitively loaded balun is given by:

$$
\begin{equation*}
\Gamma^{-1}=-\frac{N(s)}{D(s)} \tag{4.12}
\end{equation*}
$$

where $N(s)$ and $D(s)$ are as follows:

$$
\begin{align*}
& N(s)=R_{0} C_{3} C_{5} L^{2}\left(1-k^{2}\right) s^{4}+ \\
& L^{2}\left(1-k^{2}\right)\left(Y_{L} R_{0} C_{3}+C_{5}\right) s^{3}+ \\
& \left(R_{0}\left(C_{3}+0.5 C_{5}\right)+Y_{L}\left(1-k^{2}\right) L\right) L s^{2}+  \tag{4.13}\\
& \left(0.5 Y_{L} R_{0}+1\right) L s+0.5 R_{0}
\end{align*}
$$

and

$$
\begin{align*}
& D(s)=R_{0} C_{3} C_{5} L^{2}\left(1-k^{2}\right) s^{4}+ \\
& L^{2}\left(1-k^{2}\right)\left(Y_{L} R_{0} C_{3}-C_{5}\right) s^{3}+ \\
& \left(R_{0}\left(C_{3}+0.5 C_{5}\right)-Y_{L}\left(1-k^{2}\right) L\right) L s^{2}+  \tag{4.14}\\
& \left(0.5 Y_{L} R_{0}-1\right) L s+0.5 R_{0} .
\end{align*}
$$

$Y_{L}$ is $R_{L}^{-1}$ and $R_{0}$ is the reference impedance. From Appendix D , for the capacitively loaded balun, the integral defined in (4.6) can be written as:

$$
\begin{equation*}
\frac{\pi}{R_{0} C_{3}}-\pi \sum_{j} a_{j}, \tag{4.15}
\end{equation*}
$$

where $a_{j}$ 's are the roots of $D(s)$ in the complex $s$-plane that are located on the right side of the $j \omega$ axis. Note that the objective is to maximize the expression given in (4.15) so that the maximum flat bandwidth is achieved with the capacitively loaded balun structure. Given that the second term in (4.15), namely $\pi \sum_{j} a_{j}$, is the sum of the roots of $D(s)$ that are on the right half of the complex $s$-plane, this term is a non-negative real number and since it is being subtracted from the first term of (4.15), intuitively speaking, one can maximize (4.15) by moving all the right-half plane roots of $D(s)$ onto the $j \omega$ axis. Note that, for impedance matching to reference impedance $R_{0}$, that is having $S_{11}=0$, some roots of $D(s)$ have to be on the $j \omega$ axis. Thus, in the best case scenario, all the roots of $D(s)$ should reside on the $j \omega$ axis. The necessary condition for a polynomial (with real coefficients), in this case $D(s)$, to have all its roots on the $j \omega$ axis is that it only consists of the even-order terms, i.e., $s^{0}, s^{2}$ and $s^{4}$. Thus, the coefficients of the odd powers of $s$ in Eq. (4.14) should be zero. Two conditions under which these coefficients of $D(s)$ will be zero are:

$$
\begin{equation*}
Y_{L} R_{0}=2 \text { and } C_{5}=2 C_{3} . \tag{4.16}
\end{equation*}
$$

By substituting these two conditions in Eq. (4.14), then:

$$
\begin{align*}
& D(s)=2 R_{0} C_{3}^{2} L^{2}\left(1-k^{2}\right) s^{4}+  \tag{4.17}\\
& \left(2 R_{0} C_{3}-2 R_{0}^{-1}\left(1-k^{2}\right) L\right) L s^{2}+0.5 R_{0}
\end{align*}
$$

For all four roots of the polynomial in Eq. (4.17) to be purely imaginary (i.e., to be on the $j \omega$ axis), we need:

$$
\begin{equation*}
\frac{L}{C_{3}}=x R_{0}^{2}\left(\frac{1-\sqrt{1-k^{2}}}{1-k^{2}}\right) \tag{4.18}
\end{equation*}
$$

where $0<x \leq 1$. Assuming that one of the frequency corners is known so that both the capacitively loaded and conventional structures can be designed and compared, we obtain the relationship between the components values and one of the frequency corners, namely, the lower frequency corner, $-\omega_{l}$. Note that the goal of extending the bandwidth of the structure will be achieved by realizing a structure that has a larger value for (4.6) between $\omega_{l}$ and $\omega_{h}$. As mentioned earlier, since $|\Gamma|=\sqrt{2} / 2$ at the $3-\mathrm{dB}$ frequencies of a lossless two-port network, then from Eq. (4.12) we can reach to the following relation between $C_{3}$ and $\omega_{l}$.

$$
\begin{equation*}
a_{8} C_{3}^{8}+a_{6} C_{3}^{6}+a_{4} C_{3}^{4}+a_{2} C_{3}^{2}+1=0 \tag{4.19}
\end{equation*}
$$

where the polynomial coefficients are as follows:

$$
\begin{align*}
& a_{2}=8 R_{0}^{2} \omega_{l}^{2} x\left(\frac{1}{\sqrt{1-k^{2}}}-1\right) \times \\
& \left(\frac{x\left(3 k^{2}-1\right)}{1-k^{2}}+\frac{x-1+(1-3 x) k^{2}}{\sqrt{\left(1-k^{2}\right)^{3}}}\right) \\
& a_{4}=8 R_{0}^{4} \omega_{l}^{4} x^{2}\left(\frac{1}{\sqrt{1-k^{2}}}-1\right)^{2} \times \\
& \left(\frac{3-\left(1+2 x^{2}\right) k^{2}+4 x(x-1)\left(1-\sqrt{1-k^{2}}\right)}{1-k^{2}}\right)  \tag{4.20}\\
& a_{6}=32 R_{0}^{6} \omega_{l}^{6} x^{3}\left(\frac{1}{\sqrt{1-k^{2}}}-1\right)^{3}\left(\frac{x-1}{\sqrt{1-k^{2}}}-x\right) \\
& a_{8}=16 R_{0}^{8} \omega_{l}^{8} x^{4}\left(\frac{1}{\sqrt{1-k^{2}}}-1\right)^{4}
\end{align*}
$$

Next step is to find the value of $x$ for which the expression (4.15) becomes maximum. Fig. 4.5 shows $R_{0} C_{3} \omega_{l}$ in terms of the value $x$ for the different
values of the coupling factor, $k$. It is extracted from Eq. (4.19). As observed, it is a strictly decreasing function over the given range. Thus, for each value of the coupling factor, $k$, the minimum value of $C_{3}$ (maximum value of expression (4.15)) for the given values of $R_{0}$ and $\omega_{l}$ occurs at $x=1$. In fact, for $x=1$, the polynomial (4.17) gives two double roots. To further enhance the bandwidth, but at the expense of some in-band ripple in the frequency response of the insertion loss, the double roots can be split. it occurs for $x<1$. Consequently, the maximally flat bandwidth is obtained for the maximum value of the quantity defined in (4.6).


Figure 4.5: $R_{0} C_{3} \omega_{l}$ in terms of $x$ for the different values of $k$.
As the value of the reference impedance is typically $50 \Omega$, based on (4.16), there is only one case for which all the roots of Eq. (4.17) will be located on the $j \omega$ axis. This case occurs for $R_{L}=Y_{L}^{-1}=25 \Omega$.

However, in a scenario in which $Y_{L} R_{0} \neq 2$, all four roots of Eq. (4.14) cannot be on the $j \omega$ axis as one of the conditions in (4.16) is violated. Since the roots on the $j \omega$ axis appear in the complex conjugate form, we assume that two out of four roots of Eq. (4.14) can be placed on the $j \omega$ axis. For a $4^{\text {th }}$-order polynomial of the form $a_{4} s^{4}+a_{3} s^{3}+a_{2} s^{2}+a_{1} s+a_{0}=0$, to have only two imaginary roots (i.e., two roots on the $j \omega$ axis), it can be shown that the necessary and sufficient conditions are given by:

$$
\begin{equation*}
\frac{a_{0} a_{3}}{a_{1} a_{4}}=\frac{a_{2}}{a_{4}}-\frac{a_{1}}{a_{3}} \text { and } a_{1} a_{3}>0 . \tag{4.21}
\end{equation*}
$$

Using the first condition in (4.21), the $4^{\text {th }}$-order polynomial can be factorized
as:

$$
\begin{equation*}
\left(s^{2}+\frac{a_{1}}{a_{3}}\right)\left(a_{4} s^{2}+a_{3} s+\frac{a_{0} a_{3}}{a_{1}}\right) \tag{4.22}
\end{equation*}
$$

If $D(s)$ is of the form shown in (4.22), then it has two roots on the $j \omega$ axis and two other roots on either half planes of the complex frequency plane depending on the sign of $a_{1}$. One can show that if $a_{1}<0\left(a_{1}>0\right)$, then two roots will be on the right (left) half plane. By replacing the coefficients of the polynomial (4.14) in the first condition of (4.21), the following relation can be obtained.

$$
\begin{align*}
& L=\frac{R_{0}\left(C_{3}+0.5 C_{5}\right)}{Y_{L}\left(1-k^{2}\right)}+\frac{0.5 R_{0}\left(C_{5}-R_{0} Y_{L} C_{3}\right)}{Y_{L}\left(0.5 R_{0} Y_{L}-1\right)}-  \tag{4.23}\\
& \frac{R_{0} C_{3} C_{5}\left(0.5 R_{0} Y_{L}-1\right)}{\left(1-k^{2}\right)\left(R_{0} Y_{L} C_{3}-C_{5}\right)}
\end{align*}
$$

Next, we find a relation between components values and the lower frequency corner $\omega_{l}$ for this scenario. Given that $|\Gamma|=\sqrt{2} / 2$, from Eq. (4.12), the following relation can be obtained.

$$
\begin{align*}
& g\left(C_{3}, C_{5}, L\right)=4 R_{0}^{2} C_{3}^{2} C_{5}^{2} L^{4}\left(1-k^{2}\right)^{2} \omega_{l}^{8}+ \\
& \left(4 L^{4}\left(1-k^{2}\right)^{2}\left(Y_{L}^{2} R_{0}^{2} C_{3}^{2}+C_{5}^{2}\right)-\right. \\
& \left.4 R_{0}^{2} C_{3} C_{5} L^{3}\left(1-k^{2}\right)\left(C_{5}+2 C_{3}\right)\right) \omega_{l}^{6}+ \\
& \left(4 R_{0}^{2} C_{3} C_{5} L^{2}\left(1-k^{2}\right)+\left(C_{5}+2 C_{3}\right)^{2} L^{2} R_{0}^{2}+\right.  \tag{4.24}\\
& \left.4 L^{4}\left(1-k^{2}\right)^{2} Y_{L}^{2}-4 L^{3}\left(1-k^{2}\right)\left(2 C_{5}+Y_{L}^{2} R_{0}^{2} C_{3}\right)\right) \omega_{l}^{4} \\
& +\left(2 R_{0}\left(6 Y_{L} L^{2}\left(1-k^{2}\right)-\left(C_{5}+2 C_{3}\right) L R_{0}\right)-\right. \\
& \left.\left(2+Y_{L} R_{0}\right)^{2} L^{2}+2 L^{2}\left(Y_{L} R_{0}-2\right)^{2}\right) \omega_{l}^{2}+R_{0}^{2}=0
\end{align*}
$$

The last step is to maximize the term in (4.15). Given that in this scenario, $Y_{L} R_{0}$ can be smaller or greater than two, the quantity defined in (4.15) can have two different values as given below.

$$
\int_{0}^{+\infty} \ln \frac{1}{\left|S_{11}(j \omega)\right|} d \omega= \begin{cases}\frac{\pi}{R_{0} C_{3}} & \text { if } Y_{L} R_{0}>2  \tag{4.25}\\ \frac{\pi Y_{L}}{C_{5}} & \text { if } Y_{L} R_{0}<2\end{cases}
$$

Given that in practice, the integral quantity should be a continuous curve as $Y_{L} R_{0}$ changes, therefore, $\frac{\pi}{R_{0} C_{3}}=\frac{\pi Y_{L}}{C_{5}}$ which is a general form of conditions presented in (4.16) for the case of $Y_{L} R_{0}=2$.

To maximize the value of the integral in (4.25), we should use the minimum value of $C_{3}$ when $Y_{L} R_{0}>2$ or the minimum value of $C_{5}$ when $Y_{L} R_{0}<2$. Thus, depending on whether the value of $Y_{L} R_{0}$ is greater or smaller than 2, the corresponding relation below should be used to find the minimum value for $C_{3}$ or $C_{5}$. If $Y_{L} R_{0}=2$ either one of the relations can be used. Note that for the purpose of brevity, in arriving at the rightmost relation, some intermediate steps are omitted.

$$
\left\{\begin{array}{l}
\text { if } Y_{L} R_{0}>2 \Rightarrow \frac{d C_{3}}{d C_{5}}=0 \Rightarrow \frac{\partial g}{\partial C_{5}}+\frac{\partial g}{\partial L} \frac{\partial L}{\partial C_{5}}=0  \tag{4.26}\\
\text { if } Y_{L} R_{0}<2 \Rightarrow \frac{d C_{5}}{d C_{3}}=0 \Rightarrow \frac{\partial g}{\partial C_{3}}+\frac{\partial g}{\partial L} \frac{\partial L}{\partial C_{3}}=0
\end{array}\right.
$$



Figure 4.6: Normalized values of $(4.6)$ and $(4.7)$ in terms of the differential output impedance.

Fig. 4.6 shows the plots of (4.6) and (4.7) as a function of the differential load $\left(2 R_{L}\right)$ for the capacitively loaded and Marchand balun, respectively, normalized to the lower frequency corner $\left(\omega_{l}\right)$. The components values for Marchand balun are calculated based on the relations presented in [79]. Appendix E is used to calculate (4.7). Depending on the value of $Y_{L} R_{0}$, the relations in (4.16) - (4.26) are used to obtain the components values for the


Figure 4.7: The ratio of the lower and upper 3 dB frequencies of different baluns for two different coupling factors $(k)$ : (a) $k=0.5$, (b) $k=\frac{\sqrt{2}}{2}$.
capacitively loaded structure. As can be seen from the figure, for a range of values of the differential load, the normalized quantity of the integral in (4.6) for the capacitively loaded balun is higher than that of (4.7) for the Marchand balun. That is, the capacitively loaded balun can potentially have a wider bandwidth than that of the Marchand balun. Fig. 4.7 compares the output performance of the capacitively loaded baluns and the standard Marchand balun as a function of the differential load for two different values of the coupling factor. Note that for calculating the components of the capacitively loaded baluns three different approaches are used, namely, the components are calculated based on the conventional approach based on Eqs. (4.1) and (4.2), the proposed approach using Eqs. (4.16) - (4.26), and
another proposed approach which will be discussed in the next section. In these plots all the baluns are designed to have the same $\omega_{l}$. Also, while the electrical length of the coupled line for Marchand balun is $90^{\circ}$, for all the capacitively loaded baluns it is $10^{\circ}\left(\ll 90^{\circ}\right)$. Based on the results shown in Fig 4.7 (a) and (b), the following observations can be made: 1) the bandwidth of the structure is proportional to the coupling factor $k ; 2$ ) the conventional approach for calculating the components values of the capacitively loaded balun is not optimal. For instance, even with the lower number of lumped components in our approach (with assumption of $C_{1}=C_{2}=0$ and $C_{4}=\infty$ ) and using Eqs. (4.16) - (4.26), the output performance of the proposed capacitively loaded balun is better than that of the conventional capacitively loaded balun over a wide range of the differential output impedance values; 3) the performance of the proposed balun can be further enhanced using the approach presented in the next section that assumes finite values for $C_{1}, C_{2}$ and $\left.C_{4} ; 4\right)$ the capacitively loaded baluns whose components are obtained using the approaches presented in this work offer a better performance as compared to the conventional Marchand balun for a range of the output impedance values; 5) the overall performance of the capacitively loaded baluns whose components are calculated based on the approaches presented in this work are less sensitive to the load variations.

### 4.2 An Alternative Methodology for Calculating the Components' Values of Capacitively Loaded Baluns

### 4.2.1 The proposed procedure

The proposed approach to finding the value of components of a capacitively loaded balun presented in the previous section is based on the simplifying assumption of $C_{1}=C_{2}=0$ and $C_{4}=\infty$. In this section, we present an alternative approach in which $C_{1}, C_{2}$ and $C_{4}$ can have finite values. The flowchart of this alternative approach is shown in Fig. 4.8. In this approach, the independent parameters that one needs to find are $C_{1}, C_{3}, C_{4}, C_{5}$ and $Z_{o}^{\prime}$. Note that $C_{2}$ is not considered an independent parameter, since it is dependent on $C_{1}$ as shown in Appendix C. $Z_{e}^{\prime}$ is also dependent on $Z_{o}^{\prime}$ through the relation $(1+k) /(1-k) \times Z_{o}^{\prime}$. In this alternative methodology, the main objective is to find the values of the components that maximize the integral value of (4.6) normalized to the lower frequency corner $\omega_{l}$. The values of the parameters are found using the iterations shown in the left hand side of

Fig. 4.8 (Path A). To find the solution set(s) that do not have an in-band notch, a user-defined parameter $\Delta \omega$ is included in the flowchart. In fact, the approach is to pick the set(s) whose notch and upper 3 dB frequencies satisfy $\omega_{\text {notch }}-\omega_{h} \geq \Delta \omega$. If no solution set(s) that can satisfy both of the following conditions, namely, (a) maximizing the integral value (4.6) normalized to its $\omega_{l}$ and (b) the set(s) would not result in any in-band notch, then the right-hand path of the flowchart (Path B) is used to find another solution set(s). In this path, by including another user-defined parameter, namely, $\Delta P$, an effort is made to obtain a suboptimal solution set(s) that result in the integral value of $(P-i \times \Delta P)$ (which is $P$ minus $\Delta P$ or minus an integer multiple of $\Delta P$ ). Furthermore, the procedure in this Path selects the solution set(s) for which there is either no in-band notch frequency or the notch frequency of the output response is out of band and $\omega_{\text {notch }}-\omega_{h} \geq \Delta \omega$. If more than one solution set is found, the one that offers the maximum ratio $\omega_{h} / \omega_{l}$ is selected.


Figure 4.8: The proposed flowchart to find the components values.

### 4.2.2 Generalization of the approach to multi-stage balun design with a broader overall bandwidth

The proposed balun structure has a single-stage topology. The balun structure can be generalized to a multi-stage topology with a broader overall bandwidth. Fig. 4.9 shows one such generalized structure. Based on what is explained in Appendix C, one can show that the amplitude and phase mismatch of the multi-stage structure will ideally remain independent of the frequency if the following conditions are met.

$$
\begin{equation*}
r=\frac{1+k}{1-k}=\frac{Z_{e i}}{Z_{o i}} \quad 1 \leq i \leq n \tag{4.27}
\end{equation*}
$$



Figure 4.9: The balun structure based on multi-stage topology.

$$
\begin{equation*}
\frac{C_{1, i}}{C_{2, i}}=\frac{2}{r-1}, \tag{4.28}
\end{equation*}
$$

where $k$ is the coupling factor between the two lines of each stage and is assumed to be the same for all the stages. Fig. 4.10 compares the frequency performance of one, two, and three stage baluns as well as the conventional Marchand balun for the source and differential load impedances of $50 \Omega$ and $200 \Omega$, respectively. For the designs whose frequency response is shown in Fig. 4.10, the sum of the electrical line lengths of each arm of the multi-stage design is selected to be $15^{\circ}$ at 30 GHz . That is, $\sum_{i=1}^{n} \theta_{n}=15^{\circ}$, where $n$ for the single, two, and three-stage baluns are 1,2 and 3 , respectively. The component values of the multi-stage balun are obtained from the proposed flowchart. The conventional Marchand balun is designed based on [79]. In this design, to have a fair comparison on the performance improvement of the multi-stage balun as the number of stages increases, the lower 3 dB frequency corner of all the baluns is set to be equal to that of the conventional Marchand balun, which is 8.1 GHz . As can be seen from the figure, the frequency performance of the three-stage balun with the total electrical line length of $15^{\circ}$ for each arm is comparable with that of the conventional Marchand balun with the electrical line length of $90^{\circ}$. Thus, while the size can be much more compact than that of the Marchand balun, a similar frequency performance can be achieved.

### 4.3 Calculating the Component values of the Balun Using the Proposed Approach

This section is mainly focused on the design of the capacitively loaded balun using the proposed approach. As a proof-of-concept a designed balun, which converts the unbalanced signal with the source impedance of $50 \Omega$ to the balanced output with the differential impedance of $50 \Omega$, is implemented in a $65-\mathrm{nm}$ bulk CMOS process. The design procedure begins with finding the


Figure 4.10: The performance comparison of different multi-stag baluns as well as conventional Marchand one with the source and differential load impedances of $50 \Omega$ and $200 \Omega$, respectively.
optimal coupling factor, $k$, in the given technology. The process technology used in this work offers 9 metal layers plus one additonal layer, where the top two metal layers and the additional layer are thicker than the rest. To obtain the optimal $k$, the primary (main) transmission line is surrounded by two secondary (coupled) transmission lines. Fig. 4.11(a) shows the coupled line structure. Fig. 4.11(b) shows the coupling factor obtained using (electromagnetic) EM simulations. Its value is around 0.75 over the frequency band of 1 GHz to 100 GHz .

Since the main objective here is to design a wideband miniaturized balun, two assumptions are made. First, the coupled lines in Fig. 4.1(c) are approximated by two coupled inductors (Fig. 4.12 (a)). In fact, given the small size of the balun, each of its small transmission lines can be approximated with a single inductor. Second, for the purpose of comparison, the lower 3 dB frequency, $\omega_{l}$, of the designed balun is chosen to be equal to that of the conventional Marchand balun. With these assumptions, the component values are obtained using the equations derived earlier for a small $\theta^{\prime}$. Once the design is completed, the coupled inductors are replaced with their corresponding transmission lines (Fig. 4.12(b)). Then using the proposed flowchart that is presented in the previous section, if needed, the component values are adjusted to achieve the best possible bandwidth.

Considering that (as mentioned in the next section) our desired band-


Figure 4.11: The coupled line implemented in $65-\mathrm{nm}$ bulk CMOS, (a) its schematic, (b) its coupling factor, $k$.


Figure 4.12: The capacitively loaded balun, (a) lumped version, (b) distributed variant $\left(\mathrm{R}_{\mathrm{s}}=2 \mathrm{R}_{\mathrm{L}}=50 \Omega\right)$.
width of operation is from 10 to 50 GHz , therefore, the center frequency is 30 GHz . From Fig. 4.12 (b) the coupling factor of the coupled lines at 30 GHz is 0.75 . Therefore, using $k=0.75$ and the equations presented in [34], we design the Marchand balun. Its $\omega_{l}$ is 10.8 GHz . Using (4.19) with known $k$ and $\omega_{l}$, two acceptable values for $C_{3}$ are obtained. The values are 99.2 fF and 503.9 fF . The lower one is acceptable since the objective is to maximize Eq. (4.25). Then, the values of $C_{5}$ and $L$ can be found using (4.16) and (4.18), respectively. Their values are 198.4 fF and 191.92 pH . Replacing the inductors in Fig. 4.12(a) with the transmission lines and using the relation below (which gives a relationship among the inductance $L$, the odd mode impedance $Z_{o}^{\prime}$ and electrical length $\theta^{\prime}$ at the center frequency, $f_{0},[78]$ ), we can estimate the initial value of $Z_{o}^{\prime}$.

$$
\begin{equation*}
L=\frac{\sin \theta^{\prime}}{2 \pi f_{0}(1-k)} \times Z_{o}^{\prime} \tag{4.29}
\end{equation*}
$$

From EM simulation, the line based on the structure shown in 4.11 (a) with the electrical length of $\sim 17^{\circ}$ exhibits the inductance of $\sim 192 \mathrm{pH}$. Con-
sidering $f_{0}$ to be 30 GHz , the initial value of $Z_{o}^{\prime}$ from Eq. (4.29) is $31.7 \Omega$. Now, with the initial values obtained for $C_{1}, C_{3}, C_{4}, C_{5}$ and $Z_{o}^{\prime}$, which are $0,99.2 \mathrm{fF}, \infty, 198.4 \mathrm{fF}$ and $31.7 \Omega$, respectively, we can use the proposed flowchart to obtain the optimum component values to achieve the widest bandwidth. Given that the electrical length of the lines is relatively small $\left(\sim 17^{\circ}\right)$, for $C_{3}, C_{5}$ and $Z_{o}^{\prime}$ the range of their variation for tweaking them is considered within $0.5 \times$ to $1.5 \times$ of their initial values. The value of $C_{1}$ changes from 0 to 50 fF . $C_{4}$ varies from 500 fF to 2 pF assuming that the capacitance of 2 pF is large enough. With these assumptions, the final values of $C_{1}, C_{3}, C_{4}, C_{5}$ and $Z_{o}^{\prime}$ are $0,76.6 \mathrm{fF}, \infty, 182 \mathrm{fF}$ and $31.4 \Omega$, respectively. Fig. 4.13 shows the frequency response of the following baluns which are designed based on: 1) the proposed approach for designing the capacitively loaded structure (Fig. 4.12(b)), 2) the structure with lumped components (Fig. 4.12 (a)) whose component values are calculated based on the proposed approach, 3) the conventional Marchand structure whose component values are obtained based on [79], 4) the conventional reduced-size capacitively loaded Marchand balun whose component values are calculated using Equations (Eqs. (4.1) and (4.2)), and 5) the structure presented in [76]. As can be seen from the figure, over the frequency band of 10 GHz to 50 GHz , the structure designed with the proposed approach offers the best performance. It has the highest bandwidth among the five balun structures.


Figure 4.13: Frequency response of the different structure.

### 4.4 Experimental Results

The proposed approach is used to design a capacitively loaded balun for operation over the frequency range of 10 GHz to 50 GHz . A proof-of-concept prototype is designed and implemented in a $65-\mathrm{nm}$ bulk CMOS technology. The balun occupies the total area of $130 \mu \mathrm{~m} \times 180 \mu \mathrm{~m}$ excluding the pads ( $250 \mu \mathrm{~m} \times 200 \mu \mathrm{~m}$ including the pads). It converts an unbalanced signal with a source impedance of $50 \Omega$ to a balanced output with a differential load impedance of $50 \Omega$. The capacitance $C_{1}$ for the given input and output impedances is very small and therefore is approximated to be equal to zero. Moreover, the capacitance $C_{4}$ is approximated to be infinity (short) since it is large in the frequency band of interest. To make the coupled lines, two thick top metals with the thickness of $3.4 \mu \mathrm{~m}$ and $1.65 \mu \mathrm{~m}$ are used. The electrical length and coupling factor of the lines are $\sim 17^{\circ}$ at 30 GHz and 0.72 , respectively. The input and output capacitors values are 77 fF and 180 fF , respectively. Fig. 4.14 shows the micrograph of the chip. The input insertion loss and return loss are shown in Figs. 4.15 and 4.16. The amplitude and phase mismatches are shown in Figs. 4.17 and 4.18, respectively. As observed, the amplitude and phase mismatches remain below 0.4 dB and $1.5^{\circ}$, respectively, over the frequency band of 10 to 50 GHz . As can be seen from the figures, the simulated and measured results are in good agreement further confirming the validity of all the proposed analysis and methodology. Table 4.1 compares the measurement results of this balun with those of the state-of-the-art.


Figure 4.14: The microphotograph of the chip.


Figure 4.15: The simulated and measured insertion loss.


Figure 4.16: The simulated and measured return loss.


Figure 4.17: The simulated and measured amplitude mismatch.


Figure 4.18: The simulated and measured phase mismatch.

Table 4.1: Performance comparison of state-of-the-art structures.

|  | This Work | $80]$ | $81]$ | $82]$ | $[83]$ | $84]$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Technology | $65-\mathrm{nm}$ CMOS | $0.13-\mu \mathrm{m} \mathrm{SiGe}$ | InP | $0.18-\mu \mathrm{m} \mathrm{CMOS}$ | $0.13-\mu \mathrm{m} \mathrm{SiGe}$ | $0.18-\mu \mathrm{m} \mathrm{SiGe}$ |
| 3dB Bandwidth (BW), GHz | $10->50$ | $21.5-95$ | $70-110$ | $\sim 15-\sim 60^{* *}$ | $6.5-28.5$ | $34-110$ |
| 3dB Bandwidth Ratio, GHz | $>133.3$ | 126.2 | 44.4 | 120 | 125.7 | 105.6 |
| Insertion loss (IL), dB | $1.2^{*}$ | $1.8^{*}$ | $<4.5$ | $8.5^{*}$ | $<3$ | $1.7 @ 54 \mathrm{GHz}$ |
| Amplitude Mismatch $\left(A_{e}\right), \mathrm{dB}$ | $<0.4$ | $<0.4$ | 2 | $<1$ | $< \pm 1$ | 1.5 |
| Phase mismatch $\left(\theta_{e}\right),^{\circ}$ | $<1.5$ | 2.5 | 2.5 | $< \pm 5$ | $< \pm 1.65$ | 7 |
| Active area $\left(\mathrm{mm}^{2}\right)^{* * *}$ | 0.023 | 0.085 | 0.06 | 0.231 | 0.054 | 0.11 |

*Midband IL, ${ }^{* *}$ Estimated from the graph, ${ }^{* * *}$ This is the physical area and the effects of frequency scaling on the area are not taken into account.

### 4.5 Summary

In this chapter, through mathematical analysis of two types of baluns, namely, the Marchand balun and the capacitively loaded variant of it, it is shown that the conventional approach for designing the capacitively loaded balun may not necessarily lead to the optimum performance. In fact, it is shown that the capacitively loaded baluns can have wider bandwidth even with a fewer number of capacitive components. Furthermore, the reducedsize structure can exhibit a better performance than that of the conventional Marchand balun for some loads. A proof-of-concept prototype implemented in a $65-\mathrm{nm}$ bulk CMOS technology, confirms that the proposed compact size balun has a measured performance that compares favorably with that of the state-of-the-art in terms of insertion loss, and amplitude and phase mismatches, and the measurements are in good agreement with simulation results.

## Chapter 5

## A Continuous-Mode $360^{\circ}$ mm-Wave Ultra-Wideband Phase Shifter with $\sim 0.2-\mathrm{dB}$ RMS Amplitude and $<1.4^{\circ}$ Phase Error for Next-Generation Wireless Communication Systems

Phased array systems play a crucial role in the emerging and next generation of wireless communication systems [86]- 91]. Although beam-forming operation can be carried out in the radio-frequency ( RF ) path [92, 93], local-oscillator (LO) path [94] or intermediate-frequency (IF)/baseband path [95, 96], the RF-based approaches have the following advantages. First, all the building blocks after the phase shifter can be shared and second, the reactive components required in the RF path are smaller as compared to those in the baseband frequencies. While the first advantage leads to saving a considerable amount of area and power as compared to the LO-based or IF/baseband approaches, the second advantage enables a more compact area.

Depending on the structure and components used in the phase shifters they can be categorized as passive or active phase shifters [59], [97]-[100]. While passive phase shifters are typically lossy and occupy larger chip area, they offer a higher linearity. In contrast, active phase shifters exhibit a higher gain, a smaller area, simpler structure and a higher accuracy at cost of higher power consumption and lower linearity.

Yet, in another classification, depending on the operation of the phase shifters, they can be categorized as analog phase shifters which provide


Figure 5.1: Vector-sum phase shifter building block.
continuous changes in the phase shift [100] or digital phase shifters where the changes in the phase shift are discrete [59]. In general, analog phase shifters, which are typically varactor-based, have a simpler structure and lower loss. However, as compared to their digital counterparts, they exhibit a lower bandwidth, a more limited phase shift range and a higher dependency of the insertion loss of the structure on the amount of the phase shift. The dependency of the insertion loss on the amount of phase shift is mainly due to the use of the varactors that are employed in the phase shifting block. Furthermore, the bandwidth of analog phase shifters is limited by the quadrature signal splitting block, such as hybrid $90^{\circ}$ that they typically use. To alleviate the trade-off between the insertion loss and phase shift and also increase the phase shift range up to $360^{\circ}$, the vector-sum phase shifters have been considered. As shown in Fig. 5.1, in such vector-sum phase shifters, the phase shift is generated by changing the magnitude of two quadrature components of the input signals and adding them. However, the overall input-output gain of the vector-sum phase shifters is usually kept constant [100]. Since the performance of vector-sum phase shifters is mainly limited by that of their in-phase/quadrature (I/Q) generator, their accuracy over the bandwidth of interest is generally dominated by their I/Q generator.

In this chapter, we first extend the work in [101] to propose a $4^{\text {th }}$-order quadrature all-pass filter (QAF) structure. Then, we employ it in an active phase shifter, and present a low-error continuous-mode ultra-wideband $360^{\circ}$ vector-sum phase shifter.

The organization of the chapter is as follows. Section 5.1 presents the QAF structure that is used in this work and discusses its properties. Then, based on the presented QAF block, the complete structure of the vectorsum modulator is described in Section 5.2. Section 5.3 presents some design and layout considerations for the implementation of the phase shifter. Measurement results and performance comparison with other phase shifters are
presented in Section 5.4 .

### 5.1 On the Selection of the Quadrature All-Pass Filter (QAF) Structure

In [101], the design of a general $n^{\text {th }}$-order polyphase all-pass filter (PAF) has been presented that improves the trade-off between output phase accuracy and filter loss. As discussed in [101], in practice, especially in integrated circuits (ICs), the order of PAFs is rarely beyond 4. A $4^{\text {th }}$-order PAF offers a good compromise between the quadrature phase accuracy over its bandwidth and the filter's implementation complexity. Furthermore, as compared to the $3^{\text {rd }}$-order PAFs the structure of the $4^{\text {th }}$-order PAF is more amenable to symmetric implementation. Therefore, in this work we build on and modify the $4^{t h}$-order PAF to present a $4^{\text {th }}$-order QAF for the wideband phase-shifting application of this work.

Based on the design approach presented in Chapter 3 [101], Fig. 5.2(a) shows two different variants of the $4^{\text {th }}$-order PAF topology. The topology of Fig. 5.2(a) is the one presented in [101. The building blocks $N_{1}$ and $N_{3}$ comprise of the series LC networks. $N_{2}$ and $N_{4}$ are the dual networks of $N_{1}$ and $N_{3}$, respectively. One can show that the structure of Fig. 5.2(b) exhibits the same performance as that of Fig. 5.2(a) with the same component values obtained for the structure of 5.2 (a). In contrast to the structure of Fig. 55.2(a) that is intrinsically asymmetrical, the structure of 5.2 (b) has a symmetric topology. This symmetry could facilitate the design of the circuit layout at high frequencies and for wideband applications.

The rest of this section further studies the effect of various non-idealities on the output performance of these two topologies.

### 5.1.1 Loading Effect

In this subsection, the loading effect (mainly capacitive loading of next stage) on the quadrature outputs (magnitude and phase) of two topologies are analyzed and compared with each other. The reason why the load is considered to be capacitive is that in practice, since the filter is typically loaded by an active variable gain amplifier (VGA), the load will be mainly the input capacitance of the CMOS VGA. After some analysis, the the ratio of the output voltages in terms of output load capacitances for the topology of


Figure 5.2: The $4^{\text {th }}$-order quadrature all-pass filters (QAFs): (a) asymmetrical topology, (b) symmetrical one.

Fig. 5.2(a) will be as follows:

$$
\begin{equation*}
\frac{V_{o 1}}{V_{o 2}}=A(s) \times \frac{1+R C_{L_{2}} s}{1+R C_{L_{1}} s} \tag{5.1}
\end{equation*}
$$

where the load capacitors for $V_{o 1}$ and $V_{o 2}$ are denoted as $C_{L_{1}}$ and $C_{L_{2}}$, respectively, and $A(s)$ is the output ratio $\left(V_{o 1} / V_{o 2}\right)$ in the absence of the load capacitances and is defined as:

$$
\frac{Z_{1}(s)-R}{Z_{1}(s)+R} \times \frac{Z_{2}(s)+R}{Z_{3}(s)-R}
$$

$Z_{1}(s)$ and $Z_{3}(s)$ are the effective impedances of $N_{1}$ and $N_{3}$, respectively. It should be noted that the impedance of the dual networks $N_{2}$ and $N_{4}$, that is, $Z_{2}(s)$ and $Z_{4}(s)$, respectively, and that of the networks $N_{1}$ and $N_{3}$, namely, $Z_{1}(s)$ and $Z_{3}(s)$, are related to each other by the following relation [101].

$$
\begin{equation*}
Z_{2,4}(s)=\frac{R^{2}}{Z_{1,3}(s)} \tag{5.2}
\end{equation*}
$$

To meet the above relation, the inductive (capacitive) component of the networks $N_{1,3}$ and capacitive (inductive) component of their dual networks $N_{2,4}$ are required to satisfy the following relation [101].

$$
\begin{equation*}
R^{2}=\frac{L}{C} \tag{5.3}
\end{equation*}
$$

If $\left|R C_{L_{i}} \omega\right| \ll 1$, the amplitude and phase of the topology of Fig. 5.2(a) can be approximated as below.

Amplitude mismatch:

$$
\begin{equation*}
\left|\frac{V_{o 1}}{V_{o 2}}\right| \approx 1+\left(R C_{L} \omega\right)^{2} x \tag{5.4}
\end{equation*}
$$

Phase mismatch:

$$
\begin{equation*}
\angle\left(\frac{V_{o 1}}{V_{o 2}}\right)-\angle A(j \omega) \approx R C_{L} \omega x \tag{5.5}
\end{equation*}
$$

where $C_{L}$ and $x$ are the arithmetic mean of $C_{L_{1}}$ and $C_{L_{2}}$ and $\left(C_{L_{2}}-C_{L_{1}}\right) / C_{L}$, respectively. One can show that the outputs ratio for the topology shown in Fig. 5.2(b) will be:

$$
\begin{equation*}
\frac{V_{o 1}}{V_{o 2}}=A(s) \times \frac{1+\frac{2 R^{2} C_{L_{2}} s Z_{3}(s)}{\left(Z_{3}(s)+R\right)^{2}}}{1+\frac{2 R^{2} C_{1} s Z_{1}(s)}{\left(Z_{1}(s)+R\right)^{2}}}, \tag{5.6}
\end{equation*}
$$

Comparing Eqs. (5.1) and (5.6), it can be seen that the topology of Fig. 5.2(a) is independent of the load capacitance assuming that the output loads are equal. However, the topology of Fig. 5.2 (b) does not show the same property. Fig. 5.3 confirms this observations through simulations. It shows the simulation results at three different frequencies, i.e. $10 \mathrm{GHz}, 22.36 \mathrm{GHz}$ (geometric mean of the lower and upper frequency corners), and 50 GHz . The desired frequency band of the design is from 10 GHz to 50 GHz . The resistance is $50 \Omega$. The design procedure is based on what is presented in [101] for the $4^{\text {th }}$-order polyphase all-pass filter (PAF). Assuming $C_{L_{1}}=C_{L_{2}}=C_{L}$, the sensitivity of the topology of Fig. 5.2(b) to the load is nonzero, in contrast to that of the structure of Fig. 5.2 (a) where the design is insensitive to (i.e., independent of) the load.

### 5.1.2 Study of Effect of Resistor Variations on the Output Performance

Another factor that can affect the performance is the deviation of the resistor values from their typical value. Note that compared to $L, R$ and $C$ have much more variations over the process corners. The relations below present the ratio of the voltages $V_{o 1}$ and $V_{o 2}$ as a function of resistor variation $\Delta R$ for both topologies. Here, it is assumed that the resistance values are all the same, however, they vary by $\Delta R$ from the nominal value of


Figure 5.3: The $\mathrm{I} / \mathrm{Q}$ performance of two topologies versus $C_{L}$ assuming $C_{L_{1}}=C_{L_{2}}$ at three different frequencies: (a) amplitude mismatch, (b) phase mismatch.
$R$. This assumption can be justified due to the fact that in an integrated implementation, the resistors are in close proximity of each other. Thus, all changes due to process variation are the same.

For the topology of Fig. 5.2 (a), in the presence of deviations in $R$, we have:

$$
\begin{equation*}
\frac{V_{o 1}}{V_{o 2}}=A(s) \times \frac{1+\frac{\Delta R}{R}+\frac{\Delta^{2} R}{R\left(\sqrt{Z_{3}(s)}+\frac{R}{\sqrt{Z_{3}(s)}}\right)^{2}}}{1+\frac{\Delta R}{R}+\frac{\Delta^{2} R}{R\left(\sqrt{Z_{1}(s)}+\frac{R}{\sqrt{Z_{1}(s)}}\right)^{2}}} \tag{5.7}
\end{equation*}
$$

For $\Delta R / R \ll 1$, the second term in the above relation can be approximated with 1. In fact, for a small deviation, the topology of Fig. 5.2(a) is insensitive to $\Delta R$.

However, for the topology of Fig. 5.2 (b), we have:

$$
\begin{equation*}
\frac{V_{o 1}}{V_{o 2}}=A(s) \times \frac{1+\frac{\Delta R\left(Z_{3}(s)+\frac{R^{2}}{Z_{3}(s)}\right)}{R\left(\sqrt{Z_{3}(s)}+\frac{R}{\sqrt{Z_{3}(s)}}\right)^{2}}}{1+\frac{\Delta R\left(Z_{1}(s)+\frac{R^{2}}{Z_{1}(s)}\right)}{R\left(\sqrt{Z_{1}(s)}+\frac{R}{\sqrt{Z_{1}(s)}}\right)^{2}}} \tag{5.8}
\end{equation*}
$$

Fig. 5.4 shows the I/Q performance of the two structures shown in Fig. 5.2 in terms of resistance variation. Similar to the previous subsection, the sensi-
tivity of the topology of Fig. 5.2(a) is much less than that of Fig. 5.2 (b). For instance, the quadrature phase mismatch of the structure of Fig. 5.2 (a) remains below $4^{\circ}$ for total resistance variation $(\Delta R / R)$ of $\pm 30 \%$, whereas, for the same resistance variations, the topology of Fig. 5.2 (b) shows a maximum phase mismatch iof $20^{\circ}$.


Figure 5.4: The I/Q performance of two topologies versus resistance variation at three different frequencies: (a) amplitude mismatch, (b) phase mismatch.

### 5.1.3 Effect of the Reactive Component Value Deviations

In this subsection, the quadrature error originated from the reactive component value deviation is analyzed. In fact, any component value variation would cause deviations in either side of the equality in (5.3) and thus could affect the overall performance. In the following, to keep the formulas more tractable, we assume that only the capacitance values deviate from their nominal values. This is a reasonable assumption since, in practice, in ICs, the inductance variations are typically much less than that of the capacitance variations. Furthermore, again due to proximity of the components, we assume that the capacitances deviation from their nominal values due to process variation are equal. The outputs ratio versus the capacitance
deviation $\Delta C$ for both topologies are presented below.

$$
\begin{align*}
& \frac{V_{o 1}}{V_{o 2}}=A(s) \times \frac{1+\frac{\frac{Z_{2}^{2}(s) C_{1} s x}{1+2_{2}(s) C_{15 x}}-\frac{x}{(1+x) C_{2} s}}{Z_{1}(s)-Z_{2}(s)}}{1-\frac{\left(Z_{2}(s)+\alpha R\right) Z_{2}(s) C_{1} s x+\frac{\left(1+\frac{\alpha Z_{2}(s)}{R}+Z_{2}(s) C_{1} s x\right) x}{1(1+x) C_{2} s}}{\left(1+Z_{2}(s) C_{1} s x\right)\left(\sqrt{Z_{1}(s)}+\sqrt{Z_{2}(s)}\right)^{2}}} \times  \tag{5.9}\\
& \frac{1-\frac{\left(g_{4}(s)+\alpha R\right) Z_{4}(s) C_{3} s x+\frac{\left(1+\frac{\alpha Z_{4}(s)}{R}+Z_{4}(s) C_{3} s s\right) x}{\left.(1+x)^{2}\right)}}{\left(1+Z_{4}(s) C_{3} s x\right)\left(\sqrt{\left.Z_{3}(s)+\sqrt{Z_{4}(s)}\right)^{2}}\right.}}{1+\frac{Z_{4}^{2}(s) C_{3} s x}{1+Z_{4}\left(s C_{3} s x\right.}-\frac{x}{1+x) C_{4} s}} Z_{3}(s)-Z_{4}(s)
\end{align*},
$$

where $x=\Delta C_{1} / C_{1}=\Delta C_{2} / C_{2}=\Delta C_{3} / C_{3}=\Delta C_{4} / C_{4} . \alpha$ is a constant parameter. It is 1 or 2 for the topology of Fig. 5.2 (a) or (b), respectively. For $x \ll \min \left\{1,\left|1-\frac{1}{R^{2} C_{1(3)} C_{2(4)} \omega^{2}}\right|\right\}$, the above relation can be approximated by:

$$
\begin{align*}
& \frac{V_{o 1}}{V_{o 2}} \approx A(s) \times \\
& \frac{1+\frac{Z_{2}^{2}(s) C_{1} s-\frac{1}{C_{2} s}}{Z_{1}(s)-Z_{2}(s)} x}{1-\frac{\left(Z_{2}(s)+\alpha R\right) Z_{2}(s) R C_{1} C_{2} s^{2}+R+\alpha Z_{2}(s)}{R C_{2} s\left(\sqrt{Z_{1}(s)}+\sqrt{Z_{2}(s)}\right)^{2}} x} \times  \tag{5.10}\\
& \frac{1-\frac{\left(Z_{4}(s)+\alpha R\right) Z_{4}(s) R C_{3} C_{4} s^{2}+R+\alpha Z_{4}(s)}{R C_{4} s\left(\sqrt{Z_{3}(s)}+\sqrt{Z_{4}(s)}\right)^{2}} x}{1+\frac{Z_{4}^{2}(s) C_{3} s-\frac{1}{C_{4} s}}{Z_{3}(s)-Z_{4}(s)} x} .
\end{align*}
$$

The simulation results are shown in Fig. 5.5. Overall, as can be seen from the figure, the structure of Fig. 5.2 (a) shows much less sensitivity to capacitance variations as compared to that of Fig. 5.2(b).

### 5.1.4 Study of the Effect of the Limited Quality Factor of Inductors on Quadrature Outputs

Given that the quality factor $(\mathrm{Q})$ of on-chip inductors is typically limited, this subsection is mainly focused on evaluating the effects of the inductor $Q$ on the output performance. The relation below approximates the ratio of


Figure 5.5: The I/Q performance of two topologies versus capacitance variation at three different frequencies: (a) amplitude mismatch, (b) phase mismatch.
the outputs in terms of the inductor Q . In deriving this equation, we have assumed that $Q^{2} \gg 1$ or $Q>3.16$ (that is, $Q^{2}>10$ ) so we can approximate $1+Q^{2}$ with $Q^{2}$ and $Q^{2} /\left(1+Q^{2}\right)$ with 1.

$$
\begin{align*}
& \frac{V_{o 1}}{V_{o 2}} \approx A(s) \times \\
& \frac{1+\frac{\frac{R^{2} C_{1} \omega}{Q}+\frac{Z_{2}^{2}(s)}{Z_{2}(s)+R^{2} C_{2} \omega Q}}{1+\frac{\frac{R^{2} C_{1} \omega}{Q}\left(Z_{2}(s)+R^{2} C_{2} \omega Q\right)+\alpha R^{3} Z_{2}(s) C_{1} C_{2} \omega^{2}-Z_{2}^{2}(s)-\alpha R Z_{2}(s)}{\left(Z_{2}(s)+R^{2} C_{2} \omega Q\right)\left(\sqrt{Z_{1}(s)}+\sqrt{Z_{2}(s)}\right)^{2}}} \times}{1+\frac{\frac{R^{2} C_{3} \omega}{Q}\left(Z_{4}(s)+R^{2} C_{4} \omega Q\right)+\alpha R^{3} Z_{4}(s) C_{3} C_{4} \omega^{2}-Z_{4}^{2}(s)-\alpha R Z_{4}(s)}{\left(Z_{4}(s)+R^{2} C_{4} \omega Q\right)\left(\sqrt{Z_{3}(s)}+\sqrt{Z_{4}(s)}\right)^{2}}}  \tag{5.11}\\
& 1+\frac{\frac{R^{2} C_{3} \omega}{Q}+\frac{Z_{4}^{2}(s)}{Z_{4}(s)+R^{2} C_{4} \omega Q}}{Z_{3}(s)-Z_{4}(s)}
\end{align*}
$$

Same as the previous subsection, $\alpha$ is 1 for the topology of Fig. 5.2(a), and it is 2 for that of Fig. 5.2(b). The above relation can be further simplified if
$Q \gg\left|R^{2} C_{1(3)} C_{2(4)} \omega^{2}-1\right|^{-1}$ as shown below.

$$
\begin{align*}
& \frac{V_{o 1}}{V_{o 2}} \approx A(s) \times \\
& \frac{1+\frac{R^{2} C_{1} \omega+\frac{Z_{2}^{2}(s)}{R^{2} c_{2} \omega}}{Z_{1}(s)-Z_{2}(s)} Q^{-1}}{1+\frac{R^{2} C_{1} \omega+\alpha R Z_{2}(s) C_{1} \omega-\frac{Z_{2}^{2}(s)}{R^{2} C_{2} \omega}-\frac{\alpha Z_{2}(s)}{R C_{2} \omega}}{\left(\sqrt{Z_{1}(s)}+\sqrt{Z_{2}(s)}\right)^{2}} Q^{-1}} \times  \tag{5.12}\\
& \frac{1+\frac{R^{2} C_{3} \omega+\alpha R Z_{4}(s) C_{3} \omega-\frac{Z_{4}^{2}(s)}{R^{2} C_{4} \omega}-\frac{\alpha Z_{4}(s)}{R C_{4} \omega}}{\left(\sqrt{Z_{3}(s)}+\sqrt{Z_{4}(s)}\right)^{2}} Q^{-1}}{1+\frac{R^{2} C_{3} \omega+\frac{Z_{4}^{2}(s)}{R^{2} 4_{4} \omega}}{Z_{3}(s)-Z_{4}(s)} Q^{-1}} .
\end{align*}
$$

Simulating the $\mathrm{I} / \mathrm{Q}$ quadrature error of the two structures in terms of Q , we can see the topology of Fig. 5.2 (a) shows a better performance as compared to that of Fig. 5.2(b).


Figure 5.6: The I/Q performance of two topologies versus Q at three different frequencies: (a) amplitude mismatch, (b) phase mismatch.

Based on the above-mentioned studies, the topology of Fig. 5.2(a) is much more robust and insensitive to the process variations as compared to the topology of Fig. 5.2(b). However, the topology in Fig. 5.2(a) is inherently asymmetric, Given that for high-frequency and wideband designs the symmetry of the circuit further facilitates the implementation and layout
of the circuit, in the following section we propose an alternative topology based on that of Fig. 5.2(a), which also offers circuit symmetry.


Figure 5.7: Simplified block diagram of the proposed vector-sum active phase shifter.

### 5.2 The Complete Structure of the Proposed Vector-Sum Phase Shifter

### 5.2.1 QAF Selection and VGA Structure

The simplified structure of the proposed active phase shifter is shown in Fig. 5.7. In this design, a passive balun is used to convert the single-ended input signal to a differential (balanced) signal that is applied to the QAF. As mentioned earlier, the QAF in this architecture plays a crucial role and its amplitude and phase mismatches would adversely affect the overall performance. In the previous section, the output performance of two different QAFs were compared and it was shown that the asymmetric topology of Fig. 5.2 (a) offers a better performance. However, in practice, especially at high frequencies, one would prefer a symmetric (i.e., balanced) structure, since a balanced distribution of the on-chip traces is less prone to electromagnetic interference. Furthermore, the differential variable-gain amplifiers that follow the QAF, typically exhibit a lower than nominal common-mode rejection ratio (CMRR) at high frequencies, and thus any asymmetry that causes a common-mode can adversely affect the overall performance. Therefore, if the topology of Fig. 5.2(a) is meant to be used in the phase shifting block, it first requires some manipulation to be transformed to a balanced structure. In the following, we present an approach to address this issue.

To convert the unbalanced structure of Fig. 5.2(a) to a balanced one, the approach shown in Fig. 5.8 is proposed. As shown in this figure, for the input


Figure 5.8: Proposed approach to convert the unbalanced topology of Fig. 5.2(a) to the balanced one.
signal to see similar impedances as that of Fig. 5.2(a), the QAF structure is duplicated and the value of impedances of each replica circuit is halved (as compared to that of the original circuit). Thus, for a differential input signal, node G would be a virtual ground. Since the QAF structure is modified and it now has 8 output nodes ( 4 differential outputs, where as shown in Fig. 5.8, $o_{1(3)},-o_{1(3)}, o_{2(4)},-o_{2(4)}$ are the outputs associated with the inphase (quadrature) outputs of the QAF, the conventional differential VGA [59, 100] is no longer suitable for this work. Therefore, it should be modified to accommodate for the 4 differential outputs of the proposed symmetric QAF. Fig. 5.9 presents an alternative VGA structure for the in-phase (and quadrature) channels of the revised QAF network. As conceptually shown in Fig. 5.9(a), each VGA consists of three adders to combine the signals coming from the I (or Q) channel of the QAF. Fig. 5.9(b) shows the transistor-level implementation of each VGA in which the adding operations are performed in the current domain rather than the voltage domain. In fact, adding in the current domain is much simpler than adding in the voltage domain. Moreover, the proposed VGA, in contrast to the simple differential VGA, shows a more robust behaviour with respect to the input common-mode signal as the common-mode is attenuated by both the current-mode adder and also the differential nature of the structure. The complete structure of the I-channel and Q-channel VGAs is shown in Fig. 5.9(c). At different quadrants, only 2 out of 4 current sources $I_{n}, I_{n}^{\prime}, I_{m}$, and $I_{m}^{\prime}$ should be on. The current sources that are off in each quadrant are shown in Fig. 5.9(c).


Figure 5.9: VGA structure: (a) conceptual architecture, (b) current-mode adder implementation (transistor-level), (c) complete structure.

### 5.2.2 Baseband Circuit Structure

This subsection mainly focuses on the implementation of the baseband structures that include ADC, full scaler and bias circuits.

Fig. 5.10 shows the bias circuitry of the VGAs. It should be noted that nodes $\mathrm{X}, \mathrm{X}^{\prime}, \mathrm{X}_{\mathrm{B}}, \mathrm{X}_{\mathrm{B}}^{\prime}, \mathrm{Y}, \mathrm{Y}^{\prime}, \mathrm{Y}_{\mathrm{B}}$ and $\mathrm{Y}_{\mathrm{B}}^{\prime}$ in this figure are connected to the corresponding nodes shown in Fig. 5.9. The current shared between the two transistors $M B_{1}$ and $M B_{2}$ is controlled by the input voltage, $\mathrm{VB}_{\text {in }}$. The input inverters create balanced (differential) signals for the differential pair $M B_{1}$ and $M B_{2}$. Here, transistors $M B_{7}$ and $M B_{8}$ reproduce a voltage equal to the source voltage of $M B_{1}$ and $M B_{2}$ at the drain of $M B_{6}$, ensuring that the source-drain voltage of $M B_{5}$ and $M B_{6}$ are approximately equal. This reduces the current deviation due to channel-length modulation [26]. It also helps the current source $I_{B}$ to be more stable due to the negative feedback loop consisting of $M B_{6}$ and $M B_{7,8}$. The control switches $S W_{1-4}$ turn on/off the tail current sources of VGAs, based on the desired quadrant


Figure 5.10: Bias circuitry of the VGAs.
of operation. Here, as shown in Fig. 5.9, we would like to implement a fully continuous $360^{\circ}$ phase shifter with only one control signal, i.e., $\mathrm{V}_{\mathrm{CTRL}}$. Thus, the control inputs of the 4 switches $S W_{1-4}$ need to be derived from $\mathrm{V}_{\text {CTRL }}$. In other words, we need to design a circuit whose input is $\mathrm{V}_{\text {CTRL }}$ and it produces the control signals for the the 4 switches, $S W_{1-4}$, and the analog input voltage of the bias circuit, $\mathrm{VB}_{\text {in }}$. The relation between $\mathrm{V}_{\text {CTRL }}$ and $V B_{\text {in }}$ is shown in Fig. 5.11. As observed, the control voltage range from 0 to $\mathrm{V}_{\mathrm{DD}}$ is divided into four sub-regions each of which represents the specific quadrant. Note that the input voltage of the bias circuit is scaled from 0 to $V_{D D}$ in each sub-region. The relation between $\mathrm{VB}_{\text {in }}$ and $\mathrm{V}_{\text {CTRL }}$ is a triangular waveform. As the control voltage increases from zero, the input voltage of the bias circuit should rise with the slope of four. At $\mathrm{V}_{\mathrm{CTRL}}=0.25 \mathrm{~V}_{\mathrm{DD}}$, the input voltage should return to zero with the slope of -4 . At $\mathrm{V}_{\mathrm{CTRL}}=0.5 \mathrm{~V}_{\mathrm{DD}}$, the output phase change is expected to be $180^{\circ}$. For the third and fourth quadrants, the same approach is repeated as illustrated in Fig. 5.11.

Fig. 5.12 shows the structure of the control circuit. It includes two subblocks, namely, an 2-bit flash analog-to-digital converter (ADC) and a full scaler. The 2-bit flash ADC controls the switches $S W_{1-4}$ in the bias circuit for the selection of the proper quadrant while the control voltage of the phase shifter, namely, $V_{\text {CTRL }}$ is changing from 0 to $V_{D D}$. The single-pole


Figure 5.11: Relation between the input voltage of the bias circuit, $\mathrm{VB}_{\text {in }}$, and control voltage of the phase shifter, $\mathrm{V}_{\mathrm{CTRL}}$, at different quadrants.
double throw (SPDT) switches used in the full scaler are also contorted by the ADC. The full scaler makes the triangular waveform that is required to be formed. The operational amplifier (opamp) architecture employed in the structure of Fig. 5.12 is the constant- $g_{m}$ rail-to-rail input/output opamp proposed in [104, 105]. The amplifier topology is shown in Fig. 5.13.


Figure 5.12: The control circuitry that generates $\mathrm{VB}_{\text {in }}$ and the control signal of switches $S W_{1-4}$ from $V_{\text {CTRL }}$.


Figure 5.13: Opamp structure employed in the control circuit.

### 5.2.3 Output Buffer

To facilitate measurements, the output buffer shown in Fig. 5.14 is used. To enhance the bandwidth, the series/shunt peaking technique ( $\mathrm{L}_{\mathrm{s}}-\mathrm{L}_{\mathrm{p}}$ inductors) is employed [102, 103]. The inductor $\mathrm{L}_{\mathrm{o}}$ and capacitor $\mathrm{C}_{\mathrm{o}}$ are added at the output of the buffer to improve the output matching.


Figure 5.14: Output buffer structure.

### 5.3 Design and Layout Considerations

Given that the proposed circuit is a high-frequency broadband phase shifter, special attention should be paid to the routing, parasitics, and layout of the design to ensure that design has the desired bandwidth and does not exhibit any instability over its bandwidth of interest. Fig 5.15(a) shows a simple common-source amplifier with inductive load. The total inductance $L$ represents both the nominal value of the desired inductor $\left(\mathrm{L}_{a}\right)$ as well as that of the parasitic inductance $\left(\mathrm{L}_{t}\right)$. The parasitic inductance is mainly due to the interconnect traces of the load inductor and here we have assumed that the parasitic inductance dominates and the parasitic capacitances can be ignored. Eq. (5.13) shows the input admittance of the amplifier and as can be seen from the equation, the structure shows a negative input resistance at frequencies lower than $1 / \sqrt{L C_{g d}}$.

$$
\begin{equation*}
Y_{i n}(j \omega)=\frac{g_{m}}{1-\left(L C_{g d} \omega^{2}\right)^{-1}}+j \omega\left(C_{g s}+\frac{C_{g d}}{1-L C_{g d} \omega^{2}}\right) \tag{5.13}
\end{equation*}
$$



Figure 5.15: Amplifier topology: (a) single-ended, (b) differential structure.

To reduce (dampen) the undesired effects of this negative resistance, possible solutions include making the inductance smaller, putting a resistor in parallel with the input, and lowering the quality factor of the load. Shown in Fig. 5.15 (b) , is the differential counterpart of Fig. 5.15(a). Considering that the VGA structure used in this design is also a differential structure, the stability of the differential structure should be studied. Two scenarios, namely odd- and even-mode stability of the design should be considered. Here, the odd (even)-mode instability is evaluated when the amplifier is stimulated in the differential (common) mode manner. For this design, the odd (differential)-mode stability is less critical for the following three reasons. First, the equivalent parasitic inductance of the traces $L_{t}$ can be minimized by routing the differential interconnects in the close proximity of each other, so that the differential currents on each side (which are in opposite direction of each other) result in a lower effective magnetic field and associated magnetic flux, and thus lower parasitic trace inductance. Second, since nodes G in Fig. 5.8 are virtual ground, the resistors of the QAF stage appear as parallel resistors at the input of the amplifier. Thirdly, considering that the designed phase shifter has a broad bandwidth, the resistive term of the load can be increased so that the VGA and therefore the phase shifter show a flat response in the bandwidth of interest. In other words, the inductive loads will have effectively a lower quality factor. Due to the aforementioned reasons, the design is more stable in the differential-mode. As for the even (common)-mode instability, the situation is a bit more challenging. In this scenario, since the output currents passing through the traces on each side of the amplifier are in phase, the equivalent parasitic output inductance of the traces is higher than the differential-mode case. On the other hand, the equivalent transconductance of the amplifier for the common-mode operation is lower than that of the differential-mode operation, which in turn
results in a parallel resistance with a larger magnitude and thus less effects on the rest of the circuit. Furthermore, to ensure that the undesired effects are mitigated, the virtual ground nodes in Fig. 5.8 is connected to an actual GND. Thus, the input of the amplifier would be terminated to ground with a real resistance in both differential- and common-mode operations.

The layout of VGAs/current-mode adder integrated with that of QAF can also affect the overall performance. Fig. 5.16 shows two different layouts both of which includes the QAF and VGAs (more specifically, the differential traces connecting the differential outputs of both VGAs together to implement current-mode adding). As shown, the VGAs in these layouts are located at either end of the design. QAF layouts in both designs are the same, however, two different layout approaches are taken to combine the outputs of the two VGAs (current-mode adder). In Fig. 5.16(a), the two traces in green, which are combined at the center, serve as the adder. The main problem with this approach is that the adder layout is not symmetrical as the differential lines before combining do not see the same structures. Consequently, the coupling on adder lines may not appear in the common-mode manner. This asymmetry would affect the performance (through different magnetic coupling on the green traces on left and right side of the center). Furthermore, the assumetry issue accompanied by the closely spaced adder lines and QAF inductors at some ponits before adding is carried out can even casuse instability through loop creation between the adder lines and QAF inductors. One possible approach to ameliorate this issue is to shield the differential traces. However, shielding necessitates accessing a good quality GND with minimum parasitics across the die which is not trivial. The approach taken here (Fig. 5.16(b)) is that the differential input traces fed into the QAF are used to serve as the shielding lines as well for the differential output lines of the VGAs. As illustrated in Fig. 5.16(b), since both lines are differential and are laid out symmetrically, any signal from each line coupled to the other lines can be neutralized by its out-of-phase (differential) signal on the other line. So, it is expected that the performance in this situation would be better than that of Fig. 5.16(a). Fig. 5.17 shows the simulated performance of both layouts, confirming the improved performance of the design in Fig. 5.16(b). As can be seen, the quadrature performance is degraded at around 28 GHz for the layout of Fig. 5.16(a), while very little degradation can be observed for the structure of Fig. 5.16(b).


Figure 5.16: QAF layout with the VGAs blocks located on both sides of the layout: (a) type I, (b) type II.


Figure 5.17: QAF performance of two different layouts shown in Fig. 5.16. (a) quadrature amplitude ratio, (b) quadrature phase difference.


Figure 5.18: Die micrograph of the proposed phase shifter.

### 5.4 Experimental Results

The vector-based phase shifter is designed and implemented in a 1-poly 9 -metal (1P9M) $65-\mathrm{nm}$ bulk CMOS process. Fig. 5.18 shows the chip micrograph. The chip size, including the pads, is $700 \mu \mathrm{~m} \times 930 \mu \mathrm{~m}$. The chip is mounted on a printed-circuit board (PCB) and measured through on-wafer probes with all dc pads wire bonded to the PCB. S-parameters are measured with a Keysight N5225B performance network analyzer (PNA). The S-parameter measurements showing the small-signal gain at 16 different states and input matching results at the reference state are presented in Fig. 5.19 (a). As can be seen from the figure, the phase shifter exhibits a reasonably flat response over the bandwidth of interest. Fig. 5.19(b) shows the output phase shift at 16 different states with the phase step of $22.5^{\circ}$. A fairly flat phase response is also another feature of the proposed phase shifter.

To obtain the QAF performance including the in-phase to quadrature amplitude ratio and in-phase and quadrature phase difference, the measurement of both gain and phase of the phase shifter at two different cases of reference ( $\mathrm{I}_{\mathrm{m}} \neq 0$ and $\mathrm{I}_{\mathrm{n}}=0$ ) and $90^{\circ}$ phase shift with respect to reference ( $\mathrm{I}_{\mathrm{m}}=0$ and $\mathrm{I}_{\mathrm{n}} \neq 0$ ) are carried out. These two cases represent I and Q vectors of the QAF, respectively. The measured amplitude ratio and phase difference with respect to frequency are shown in Fig. 55.20. As can be seen from the figure, the in-phase to quadrature amplitude ratio and in-phase to quadrature phase difference show an error of less than 0.15 dB and $0.6^{\circ}$, respectively, over the frequency range of 10 GHz to 50 GHz .

To evaluate the performance of the overall phase shifter over the fre-


Figure 5.19: Output performance of the phase shifter : (a) S parameter results, (b) phase shift at 16 different states.


Figure 5.20: Quadrature characteristics of the QAF measured at the output: (a) I/Q amplitude ratio. (b) I/Q phase difference.
quency range of interest, RMS amplitude error and RMS phase error, as defined in [62] are used. Fig. 5.21 shows these two performance metrics. The phase shifter exhibits an RMS amplitude error that varies between 0.18 to 0.205 dB and an RMS phase error of $<1.4^{\circ}$ from 10 GHz to 50 GHz , respectively. Note that for the measurements, the output amplitude and phases are measured for phase steps of $5.625^{\circ}$, which is equivalent to that of a 6 -bit phase shifter. The measurements are performed for 20 different chips and the RMS amplitude and phase error histograms for the 20 samples measured at 30 GHz are shown in Fig. 5.22 . The results show a typical RMS amplitude and phase error of around 0.2 dB and $1.2^{\circ}$ with die-to-die variations of from $\sim 0.05$ to 0.45 dB in amplitude error and from $\sim 0.9^{\circ}$ to $\sim 1.4^{\circ}$ in phase error.

Fig. 5.23 shows the output phase shift versus the input control $\mathrm{V}_{\mathrm{CTRL}}$ at 30 GHz . The output phase shift transfer characteristic can be further linearized by improving the linearity of the differential pair stage $\mathrm{MB}_{1}$ and $\mathrm{MB}_{2}$ in Fig. 5.10. The gain at reference state $\left(\mathrm{V}_{\mathrm{CTRL}}=0\right)$ versus input power


Figure 5.21: Output performance of the phase sifter versus frequency: (a) RMS amplitude error. (b) RMS phase error.


Figure 5.22: (a) RMS amplitude and (b) phase error for 20 samples measured at 30 GHz .
and the histogram of $\mathrm{P}_{1 \mathrm{~dB}}$ for the 20 samples that are measured, are shown in Fig. 5.24. Table 5.1 summarizes the performance of the proposed phase shifter and compares it with that of the state-of-the-art designs. For the purpose of performance comparison, the following figure-of-merit (FOM) is used

$$
\begin{equation*}
F O M=\frac{f_{\max }+f_{\min }}{2\left(f_{\max }-f_{\min }\right)} \times \theta_{e, \max } \times A_{e, \max } \tag{5.14}
\end{equation*}
$$

where $\theta_{e, \max }$ is the maximum RMS phase error in degrees and $A_{e, \max }$ is the linear value of the maximum RMS amplitude error.


Figure 5.23: Output phase shift versus the input control $\mathrm{V}_{\mathrm{CTRL}}$ at 30 GHz .


Figure 5.24: Large signal performance at reference state $\left(\mathrm{V}_{\mathrm{CTRL}}=0\right)$ : (a) measured gain versus input power at 30 GHz : (a), (b) $\mathrm{P}_{1 \mathrm{~dB}}$ for 20 samples at 30 GHz .

Table 5.1: Performance comparison of state-of-the-art phase shifters.

|  | This Work | 106 | 107 | 108 | 109 | 110 | 111 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Technology | 65 nm CMOS | $\begin{gathered} 28 \mathrm{~nm} \text { FDSOI } \\ \text { CMOS } \end{gathered}$ | 28nm CMOS | 65 nm CMOS | $\begin{gathered} 250 \mathrm{~nm} \\ \text { BiCMOS } \end{gathered}$ | 40nm CMOS | 65 nm CMOS |
| Topology | Vector-Sum | Vector-Sum | Vector-Sum | II VGA | APN | I-DAC, VS | Vector-Sum |
| Freq. (GHz) | 10-50 | 78.8-92.8 | 22-44 | 51-66.3 | 14-50 | 3-7 | 8-12 |
| Gain (dB) | 0 | $\begin{gathered} 2.3 @ \\ 87.4 \mathrm{GHz} \end{gathered}$ | -5.81~-0.36 | -3.8(peak) | 5-16** | $-1.1 \pm 1.5^{* * *}$ | $\begin{gathered} 1.5 @ \\ 10 \mathrm{GHz}^{* * * *} \end{gathered}$ |
| RMS Gain Error, $\mathrm{A}_{\mathrm{e}}$ (dB) | $\sim 0.2$ | $<2$ | $\leq 0.59$ | $\leq 0.72$ | $<0.94$ | $\leq 0.89$ | <1.08 |
| RMS Phase Error, $\theta_{\mathrm{e}}$ (Deg) | $<1.4$ | $<11.9$ | $\leq 1.02$ | $\leq 7$ | $<9.7$ | $\leq 1.67$ | $<3.1$ |
| Phase Range (Deg) | 360 | 360 | 360 | 360 | 360 | 360 | 360 |
| $\mathrm{P}_{1 \mathrm{~dB}, \text { in }}(\mathrm{dBm})$ | -0.5 | -7 | -2.89~-0.48 | -0.23 | - | - | - |
| Resolution (bits) | Continuous | 4 | 7 | 5 | 3 | 8 | 6 |
| Voltage Supply (V) | 1 | 1.2 | 0.9 | 1 | - | 1.1 | 1.2 |
| Power Consumption (mW) | 24.8 | 21.6 | 35 | 5 | 0 | 16.2 | 14.8 |
| Area (mm ${ }^{2}$ ) | 0.65 | 0.12* | 0.74 | $0.3^{*}$ | 0.48* | 0.19* | - |
| FOM | 1.05 | 91.8 | 1.64 | 29.15 | 9.61 | 2.31 | 8.78 |

*Active Area ${ }^{* *}$ Insertion Loss ${ }^{* * *}$ Conversion Gain ${ }^{* * * *}$ Graphically estimated average gain

### 5.5 Summary

In this chapter, a low-error ultra-wideband active phase shifter suitable for 5G and beyond 5G wireless applications is presented. The QAF, the key building block of the phase shifter, is based on the PAF presented in Chapter 3. The input balun placed at the input of the phase shifter to convert the unbalanced input signal to a balanced signal is based on the balun presented in Chapter 4. To improve the stability and performance of the proposed phase shifter over a wide bandwidth of interest, several design and layout techniques are presented. A proof-of-concept of the continuous $360^{\circ}$ phase shifter is designed and fabricated in a $65-\mathrm{nm}$ bulk CMOS. Measured results show a typical RMS amplitude error of around 0.2 dB with a reasonably flat gain response and an RMS phase error of less than $1.4^{\circ}$ over the broad bandwidth of 40 GHz (from 10 GHz to 50 GHz ). The measurement results of 20 different samples further confirm the reliability and reproducibility of the proposed low-error phase shifter structure.

## Chapter 6

## Conclusion and Future Works

### 6.1 Conclusion

One of the main goals in the development of cellular networks from the first to fourth generation ( 1 G to 4 G ) has been improving the system capacity. Most of the systems in the market have been able to achieve (near-)optimal performance in terms of capacity. However, with regards to other system parameters such as latency, energy efficiency, connection density, etc there are still rooms for improvements and the previous generation systems (1G to 4G) do not offer efficient solutions. In this context, 5G (and beyond 5G) systems have the ambitious objectives and are expected to cover a wide variety of application areas including eHealth, factory automation, automated vehicles, and critical communication. To enhance over-the-air (OTA) efficiency 5G systems are relying on using multiple-in-multiple-out (MIMO), beamsteering antennas, and phased-array technologies. Furthermore, to satisfy the need for high data rates of multi $\mathrm{Gb} / \mathrm{s}$, the use of mm -wave frequency bands are envisioned due to the availability of wider frequency spectrum in those bands. At these higher frequencies, beam-steering antennas are required to direct radiated energy from the base station antenna array to the end user and vice versa mainly to overcome the higher path losses occurring at these frequencies. The main focus of this research is on the design and implementation of two main building blocks of the phased-array systems, namely, power amplifiers and phase shifters.

Since 5G systems use more complex modulation schemes to provide higher data rates, they require a highly linear PA. Moreover, to increase the battery longevity, such PAs need to be as efficient as possible. In this context, we present a highly linear and efficient $28-\mathrm{GHz}$ power amplifier. A proof-of-concept prototype is designed and fabricated in a $65-\mathrm{nm}$ bulk CMOS process and is successfully tested to validate the proposed design techniques. To increase the overall output power, the outputs of four sub-

PAs are combined. The saturated output power $P_{\text {sat }}$ of the proposed design is 23.2 dBm at 28 GHz . To enhance efficiency design techniques including using a low-loss CPW-like power combiner with the passive efficiency of $93 \%$ along with an LC filter at the output of the PA are proposed. By doing so, the PA achieves $P A E_{\max }$ in excess of $35 \%$. The linearity improvement technique is performed by means of a single varactor controlled by an envelope detector. It results in reducing both AM-AM and AM-PM distortions. The difference between $P_{1 d B}$ and $P_{\text {sat }}$ remains negligible ( -0.5 dB ), confirming the high linearity of the circuit. The PA is able to deliver an EVM of -26.2 dB when transmitting a $2.5 \mathrm{GS} / \mathrm{s}(15 \mathrm{~Gb} / \mathrm{s}) 64$-QAM modulated signal at average $P_{\text {out }}$ of 16.1 dBm . The performance of the proof-of-concept PA compares favorably with that of state-of-the-art designs implemented in more advanced CMOS or other higher performance processes. This favorable performance shows that the proposed techniques for improving the linearity and efficiency are promising.

In addition to the PA, a low-error ultra-wideband active phase shifter suitable for 5 G and beyond 5 G wireless applications is also designed and implemented. To make sure that the proposed block operates properly over the wide bandwidth of interest, several design and layout techniques are presented. A proof-of-concept of the continuous $360^{\circ}$ phase shifter is designed and fabricated in a $65-\mathrm{nm}$ bulk CMOS. Measured results show a typical RMS amplitude error of around 0.2 dB with a reasonably flat gain response and an RMS phase error of less than $1.4^{\circ}$ over a wide bandwidth of 40 GHz (from 10 GHz to 50 GHz ). The histograms of the errors of 20 different implemented samples measured at the center frequency of 30 GHz also confirm the reliability of the system.

### 6.2 Future Works

Although the design techniques proposed in this work are presented in the context of mm-wave PAs and phase shifters and are validated in a $65-\mathrm{nm}$ CMOS technology, many of the techniques are general and can be applied to other mm-wave building blocks and in different technologies. Applying these techniques to other building blocks and in different technologies can be subject of the future work. Furthermore, specific to the contributions presented in this thesis, the following areas of relevant research can be exploited in the future.

### 6.2.1 Performance Improvement

In the context of PA design, the focus of this work was on improving the linearity and efficiency. The validity of the proposed techniques are shown through implementing and testing proof-of-concept prototypes in a mature technology node, i.e., in $65-\mathrm{nm}$ CMOS. Nevertheless, some other design aspects can be further improved. These design parameters include:

- The operation bandwidth of the PA can be increased to cover a frequency band of more than 5 GHz .
- Another performance metric is the back-off efficiency. In fact, PAs do not usually operate at peak efficiency most of the time. The situation would be usually worse for the more spectrally efficient modulation schemes such as OFDM which are essential for achieving high data rates. Thus, back-off efficiency would become important factor to consider.
- Techniques to further enhance the linearity of the PA system, including using pre-distortion can be explored.


### 6.2.2 System-Level Integration

Another important task for the future is the integration of both PA and phase shifter and potentially other relevant building blocks on a single chip. Furthermore, due to the shorter wavelength of mm-wave signals the antennas are smaller and thus they are more amenable to integration. Consequently, a low-cost single-chip phased-array transmitter system that includes phase shifter, PA and antennas can be explored.

## Bibliography

[1] T. Chapman, E. Larsson, P. von Wrycza, E. Dahlman, S. Parkvall, J. Skold, HSPA Evolution: The Fundamentals for Mobile Broadband, Academic Press, 2014.
[2] E. Dahlman, S. Parkvall, J. Skold, 4G LTE-Advanced Pro and the Road to 5G, Elsevier, 2016.
[3] F. Aryanfar, J. Pi, H. Zhou, T. Henige, G. Xu, S. Abu-Surra, D. Psychoudakis, and F. Khan, "Millimeter-wave base station for mobile broadband communication," in 2015 IEEE MTT-S International Microwave Symposium, IMS 2015, 2015.
[4] W. Roh, J. Y. Seol, J. H. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-Wave Beamforming as an Enabling Technology for 5G Cellular Communications: Theoretical Feasibility and Prototype Results," IEEE Commun. Mag., 2014.
[5] D. M. Pozar, Microwave Engineering, 4th ed. Hoboken, NJ, USA: Wiley, 2011.
[6] K. M. Simon, M. J. Schindler, V. A. Mieczkowski, P. F. Newman, M. E. Goldfarb, E. Reese, and B. A. Small, "A Production-Ready, $6-18-\mathrm{GHz}, 5$-b Phase Shifter with Integrated CMOS-Compatible Digital Interface Circuitry," IEEE J. Solid-State Circuits, 1992.
[7] P. S. Wu, H. Y. Chang, M. D. Tsai, T. W. Huang, and H. Wang, "New Miniature $15-20-\mathrm{GHz}$ Continuous-Phase/Amplitude Control MMICs Using 0.18- $\mu \mathrm{m}$ CMOS Technology," IEEE Transactions on Microwave Theory and Techniques, 2006.
[8] Y. Yu, K. Kang, C. Zhao, Q. Zheng, H. Liu, S. He, Y. Ban, L. L. Sun, and W. Hong, "A $60-\mathrm{GHz} 19.8-\mathrm{mW}$ Current-Reuse Active Phase Shifter with Tunable Current-Splitting Technique in $90-\mathrm{nm}$ CMOS," IEEE Trans. Microw. Theory Tech., 2016.
[9] T. S. Rappaport et al., "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," IEEE Access, vol. 1, pp. 335-349, 2013.
[10] S. Rangan, T. S. Rappaport and E. Erkip, "Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges," Proceedings of the IEEE, vol. 102, no. 3, pp. 366-385, March 2014.
[11] https://apps.fcc.gov/edocs_public/attachmatch/FCC-1689A1.pdf.
[12] S. Cripps, RF Power Amplifiers for Wireless Communications, $2^{\text {nd }}$ ed. Boston, MA, USA: Artech House, 2006.
[13] F. H. Raab et al., "Power amplifiers and transmitters for RF and microwave," IEEE Transactions on Microwave Theory and Techniques, vol. 50, no. 3, pp. 814-826, March 2002.
[14] S. N. Ali, P. Agarwal, J. Baylon, S. Gopal, L. Renaud and D. Heo, "A $28 \mathrm{GHz} 41 \%$-PAE linear CMOS power amplifier using a transformer-based AM-PM distortion-correction technique for 5G phased arrays," 2018 IEEE International Solid - State Circuits Conference - (ISSCC), San Francisco, CA, 2018, pp. 406-408.
[15] S. Shakib, M. Elkholy, J. Dunworth, V. Aparin and K. Entesari, "A wideband 28 GHz power amplifier supporting $8 \times 100 \mathrm{MHz}$ carrier aggregation for 5G in 40nm CMOS," 2017 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, 2017, pp. 44-45.
[16] M. Vigilante and P. Reynaert, "A 29-to-57GHz AM-PM compensated class-AB power amplifier for 5 G phased arrays in 0.9 V 28 nm bulk CMOS," 2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC), Honolulu, HI, 2017, pp. 116-119.
[17] S. Kousai, K. Onizuka, T. Yamaguchi, Y. Kuriyama and M. Nagaoka, "A 28.3 mW PA-closed loop for linearity and efficiency improvement integrated in a +27.1 dBm WCDMA CMOS power amplifier," 2012 IEEE International Solid-State Circuits Conference, San Francisco, CA, 2012, pp. 84-86.
[18] A. R. Lopez, "Review of narrowband impedance-matching limitations," IEEE Antennas and Propagation Magazine, vol. 46, no. 4, pp. 88-90, Aug. 2004.
[19] A. R. Lopez, "Rebuttal to Fano limits on matching bandwidth," IEEE Antennas and Propagation Magazine, vol. 47, no. 5, pp. 128-129, Oct. 2005.
[20] A. R. Lopez, "Wheeler and Fano Impedance Matching [Antenna designer's notebook]," IEEE Antennas and Propagation Magazine, vol. 49, no. 4, pp. 116-119, Aug. 2007.
[21] H. Dabag, B. Hanafi, F. Golcuk, A. Agah, J. F. Buckwalter and P. M. Asbeck, "Analysis and design of stacked-FET millimeter-wave power amplifiers," IEEE Transactions on Microwave Theory and Techniques, vol. 61, no. 4, pp. 1543-1556, April 2013.
[22] H. W. Bode, Network Analysis and Feedback Amplifier Design, Van Nostrand, New York, 1945.
[23] R. M. Fano, "Theoretical limitations on the broad-band matching of arbitrary impedances," J. Franklin Inst., vol. 249, pp. 57-83, Jan. 1950 and pp. 139-154, Feb. 1950.
[24] B. Razavi, RF Microelectronics- Second Edition, New York: Prentice Hall, Dec 2012.
[25] G. Hueber, A. M. Niknejad, Millimeter-Wave Circuits for 5G and Radar, Cambridge University Press, 2019.
[26] B. Razavi, Design of Analog CMOS Integrated Circuits, New York: McGraw-Hill, Jan 2016.
[27] Y. Palaskas et al., "A 5-GHz 20-dBm Power Amplifier With Digitally Assisted AM-PM Correction in a $90-\mathrm{nm}$ CMOS Process," IEEE Journal of Solid-State Circuits, vol. 41, no. 8, pp. 1757-1763, Aug. 2006.
[28] S. Shakib, H. Park, J. Dunworth, V. Aparin and K. Entesari, "A Highly Efficient and Linear Power Amplifier for $28-\mathrm{GHz} 5 \mathrm{G}$ Phased Array Radios in 28-nm CMOS," IEEE Journal of Solid-State Circuits, vol. 51, no. 12, pp. 3020-3036, Dec. 2016.
[29] A. M. Niknejad, H. Hashemi, mm-Wave Silicon Technology 60 GHz and Beyond, New York: Springer, 2008.
[30] Medley, M. W., and Jr., Microwave and RF Circuits: Analysis, Synthesis, and Design, Norwood, MA: Artech House, 1993.
[31] L. Besser, R. Gilmore, Practical RF Circuit Design for Modern Wireless Systems Volume I- Passive Circuits and Systems, Norwood, MA: Artech House, 2003.
[32] A. Chakrabarti and H. Krishnaswamy, "High-power high-efficiency class-E-like stacked mm-Wave PAs in SOI and bulk CMOS: theory and implementation," IEEE Transactions on Microwave Theory and Techniques, vol. 62, no. 8, pp. 1686-1704, Aug. 2014.
[33] A. Mazzanti, L. Larcher, R. Brama and F. Svelto, "Analysis of reliability and power efficiency in cascode class-E PAs," IEEE Journal of Solid-State Circuits, vol. 41, no. 5, pp. 1222-1229, May 2006.
[34] A. Komijani and A. Hajimiri, "A $24 \mathrm{GHz},+14.5 \mathrm{dBm}$ fullyintegrated power amplifier in $0.18 \mu \mathrm{~m}$ CMOS," Proceedings of the IEEE 2004 Custom Integrated Circuits Conference, Orlando, FL, USA, 2004, pp. 561-564.
[35] I. J. Bahl, Fundamentals of RF and Microwave Transistor Amplifiers, Hoboken, New Jersey: Wiley, 2009.
[36] S. Shakib, H. Park, J. Dunworth, V. Aparin and K. Entesari, "A 28 GHz efficient linear power amplifier for 5 G phased arrays in 28 nm bulk CMOS," 2016 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, 2016, pp. 352-353.
[37] T. Li, M. Huang and H. Wang, "A continuous-mode harmonically tuned $19-$ to- 29.5 GHz ultra-linear PA supporting $18 \mathrm{~Gb} / \mathrm{s}$ at $18.4 \%$ modulation PAE and $43.5 \%$ peak PAE," 2018 IEEE International Solid - State Circuits Conference - (ISSCC), San Francisco, CA, 2018, pp. 410-412.
[38] B. Rabet and J. Buckwalter, "A high-efficiency 28 GHz outphasing PA with 23 dBm output power using a triaxial balun combiner," 2018 IEEE International Solid - State Circuits Conference - (ISSCC), San Francisco, CA, 2018, pp. 174-176.
[39] F. Wang, T. Li and H. Wang, "A Highly Linear Super-Resolution Mixed-Signal Doherty Power Amplifier for High-Efficiency mmWave 5G Multi-Gb/s Communications," 2019 IEEE International Solid- State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2019, pp. 88-90.
[40] C. R. Chappidi, X. Wu and K. Sengupta, "Simultaneously Broadband and Back-Off Efficient mm-Wave PAs: A Multi-Port Network Synthesis Approach," IEEE Journal of Solid-State Circuits, vol. 53, no. 9, pp. 2543-2559, Sept. 2018.
[41] S. C. Cripps, RF Power Amplifiers for Wireless Communications, MA: Artech House, 1999.
[42] E. L. Griffin, "Application of loadline simulation to microwave high power Amplifiers," IEEE Microwave Mag., vol. 1, pp. 58-66, June 2000.
[43] F. Wang and H. Wang, "An Instantaneously Broadband UltraCompact Highly Linear PA with Compensated DistributedBalun Output Network Achieving $>17.8 \mathrm{dBm} P_{1 d B}$ and $>36.6 \%$ $P A E_{P 1 d B}$ over 24 to 40 GHz and Continuously Supporting 64-/256QAM 5G NR Signals over 24 to 42 GHz ," 2020 IEEE International Solid - State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2020, pp. 372-374.
[44] K. Ning, Y. Fang, N. Hosseinzadeh and J. F. Buckwalter, "A 30GHz CMOS SOI Outphasing Power Amplifier with Current Mode Combining for High Back-off Efficiency and Constant Envelope Operation," IEEE Journal of Solid-State Circuits, vol. 55, no. 5, pp. 1411-1421, May 2020.
[45] F. Behbahani, Y. Kishigami, J. Leete, and A. A. Abidi, "CMOS mixers and polyphase filters for large image rejection," IEEE Journal of Solid-State Circuits, vol. 36, no. 6, pp. 873-887, Jun. 2001.
[46] J. Savoj, "LR polyphase filter," U.S. Patent 2011 0,092,169 A1, Apr. 21, 2011.
[47] A. A. Abidi, "Direct-conversion radio transceivers for digital communications," IEEE Journal of Solid-State Circuits, vol. 30, no. 12, pp. 1399-1410, Dec. 1995.
[48] J. Crols and M. S. J. Steyaert, "A single-chip 900 MHz CMOS receiver front-end with a high performance low-IF topology," IEEE Journal of Solid-State Circuits, vol. 30, no. 12, pp. 1483-1492, Dec. 1995.
[49] Y.-Y. Huang, H. Jeon, Y. Yoon, W. Woo, C.-H. Lee, and J. S. Kenney, "An ultra-compact, linearly-controlled variable phase Shifter designed with a novel RC Poly-Phase filter," IEEE Transactions on Microwave Theory and Techniques, vol. 60, no. 2, pp. 301-310, Feb. 2012.
[50] A. Asoodeh and M. Atarodi, "A full $360^{\circ}$ vector-sum phase Shifter with very low RMS phase error over a wide bandwidth," IEEE Transactions on Microwave Theory and Techniques, vol. 60, no. 6, pp. 1626-1634, Jun. 2012.
[51] J. Kaukovuori and K. Stadius, J. Ryynanen, and K. Halonen, "Analysis and design of passive Polyphase filters," IEEE Transactions on Circuits and Systems I, vol. 55, no. 10, pp. 3023-3037, Nov. 2008.
[52] C. A. Desoer and E. S. Kuh, Basic Circuit Theory, New York: McGraw-Hill, 1969.
[53] L. Mejlbro, Methods for Finding (Real and Complex) Zeros in Polynomials, Bookboon, 2013.
[54] E. J. Barbeau, Polynomials: a Problem book in Mathematics, New York: Springer-Verlag, 1989.
[55] M. Abramowitz, and I. A. Stegun, "Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables," New York: Dover, pp. 17-18, 1972.
[56] J. Spanier, K. B. Oldham, "An Atlas of Functions," Washington, DC: Hemisphere, pp. 131-147, 1987
[57] A. Asoodeh and M. Atarodi, "A 6-bit active digital phase shifter," IEICE Electronics Express, vol. 8, no. 3, pp. 121-128, 2011.
[58] S. Y. Kim, D.-W. Kang, K.-J. Koh, and G. M. Rebeiz, "An improved Wideband all-pass I/Q network for millimeter-wave phase shifters," IEEE Transactions on Microwave Theory and Techniques, vol. 60, no. 11, pp. 3431-3439, Nov. 2012.
[59] K. J. Koh, J. W. May, and G. B. Rebeiz, "A millimeter-wave $(40-45 \mathrm{GHz}) 16$-element phased-array transmitter in $0.18-\mu \mathrm{m}$ SiGe BiCMOS technology," IEEE Journal of Solid-State Circuits, vol. 44, no. 5, pp. 1498-1509, May 2009.
[60] K. T. Christensen, "LC quadrature generation in integrated circuits," IEEE International Symposium on Circuits and Systems, Sydney, Australia, 2001, pp. 41-44.
[61] K. Christensen, "Polyphase Filters in Silicon Integrated Circuit Technology," US Patent 6,696,885, Feb. 2004.
[62] K.-J. Koh and G. M. Rebeiz, "0.13-um CMOS phase shifters for X-, Ku, and k-band phased arrays," IEEE Journal of Solid-State Circuits, vol. 42, no. 11, pp. 2535-2546, Nov. 2007.
[63] Y. Zheng and C. E. Saavedra, "Full $360^{\circ}$ Vector-Sum Phase-Shifter for Microwave System Applications," IEEE Transactions on Circuits and Systems I, vol. 57, no. 4, pp. 752-758, 2010.
[64] M. Mohsenpour and C. E. Saavedra, "Variable $360^{\circ}$ Vector-Sum Phase Shifter With Coarse and Fine Vector Scaling," IEEE Transactions on Microwave Theory and Techniques, vol. 64, no. 7, pp. 2113-2120, 2016.
[65] F. Akbar and A. Mortazawi, "A Frequency Tunable $360^{\circ}$ Analog CMOS Phase Shifter With an Adjustable Amplitude," IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 64, no. 12, pp. 1427-1431, 2017.
[66] D. Pepe and D. Zito, "Two mm-Wave Vector Modulator Active Phase Shifters With Novel IQ Generator in 28 nm FDSOI CMOS," IEEE Journal of Solid-State Circuits, vol. 52, no. 2, pp. 344-356, 2017.
[67] F. Hameau, C. Jany, B. Martineau, A. Larie, and E. Mercier, "A highly linear bidirectional phase shifter based on vector modulator for 60 GHz applications," IEEE MTT-S International Microwave Symposium (IMS), pp. 1707-1710, 2017.
[68] I. T. E. Elfergani et al., "Balanced antenna structure with slotted ground plane for LTE dual-band," 2016 Loughborough Antennas \& Propagation Conference (LAPC), Loughborough, 2016, pp. 1-5.
[69] H. T. Nguyen, and H. Wang, "A coupler-based differential mmwave Doherty power amplifier with impedance inverting and scaling baluns," IEEE Journal of Solid-State Circuits, vol. 55, no. 5, pp. 1212-1223, May 2020.
[70] C. Lien, C. Wang, C. Lin, P. Wu, K. Lin and H. Wang, "Analysis and design of reduced-size Marchand rat-race hybrid for millimeterwave compact balanced mixers in $130-\mathrm{nm}$ CMOS process," IEEE Transactions on Microwave Theory and Techniques, vol. 57, no. 8, pp. 1966-1977, Aug. 2009.
[71] H. Ma, S.J. Fang and F. lin, "Accurate and tunable active differential phase splitters in RFIC applications," United States Patent, no. $6,121,809$, Sept. 2000.
[72] Y. J. Yoon et al., "Design and characterization of multilayer spiral transmission-line baluns," IEEE Transactions on Microwave Theory and Techniques, vol. 47, no. 9, pp. 1841-1847, Sept. 1999.
[73] C. Cho and K. C. Gupta, "A new design procedure for single-layer and two-layer three-line baluns," IEEE Transactions on Microwave Theory and Techniques, vol. 46, no. 12, pp. 2514-2519, Dec. 1998.
[74] N. Marchand, "Transmission line conversion transformers," Electronics, vol. 17, no. 12, pp. 142-145, Dec. 1944.
[75] L. Xu, Z. Wang, Q. Li and J. Xia, "Modelling and design of a wideband Marchand balun," 2010 Asia-Pacific International Symposium on Electromagnetic Compatibility, Beijing, 2010, pp. 1374-1377.
[76] T. Johansen and V. Krozer, "Analysis and design of lumped element Marchand baluns," $17^{\text {th }}$ International Conference on Microwaves, Radar and Wireless Communications, Wroclaw, 2008, pp. 1-4.
[77] M. C. Scardelletti, G. E. Ponchak and T. M. Weller, "Miniaturized Wilkinson power dividers utilizing capacitive loading," IEEE Microwave and Wireless Components Letters, vol. 12, no. 1, pp. 6-8, Jan. 2002.
[78] I. J. Bahl, Lumped Elements for RF and Microwave Circuits, Artech House, 2003.
[79] H. Ahn and S. Nam, "New Design Formulas for ImpedanceTransforming 3-dB Marchand Baluns," IEEE Transactions on Microwave Theory and Techniques, vol. 59, no. 11, pp. 2816-2823, Nov. 2011.
[80] S. Chakraborty, L. E. Milner, X. Zhu, O. Sevimli, A. E. Parker and M. C. Heimlich, "An Edge-Coupled Marchand Balun With Partial Ground for Excellent Balance in $0.13 \mu \mathrm{~m}$ SiGe Technology," IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 1, pp. 226-230, Jan. 2021.
[81] M. Hossain, T. K. Johansen, M. Hrobak, W. Heinrich and V. Krozer, "A Compact Broadband Marchand Balun for Millimeterwave and Sub-THz Applications," 2020 German Microwave Conference (GeMiC), Cottbus, Germany, 2020, pp. 60-63.
[82] Y. Hsiao, C. Meng and Y. Peng, "Broadband CMOS SchottkyDiode Star Mixer Using Coupled-CPW Marchand Dual-Baluns," IEEE Microwave and Wireless Components Letters, vol. 27, no. 5, pp. 500-502, May 2017.
[83] H. J. Qian and X. Luo, "Compact 6.5-28.5 GHz On-Chip Balun With Enhanced Inband Balance Responses," IEEE Microwave and Wireless Components Letters, vol. 26, no. 12, pp. 993-995, Dec. 2016.
[84] I. Song, R. L. Schmid, D. C. Howard, Seungwoo Jung and J. D. Cressler, "A 34-110 GHz wideband, asymmetric, broadsidecoupled Marchand balun in 180 nm SiGe BiCMOS technology," 2014 IEEE MTT-S International Microwave Symposium (IMS2014), Tampa, FL, 2014, pp. 1-4.
[85] D. S. Mitrinovic, J.D. Keckic, "The Cauchy method of residues: Theory and applications," D. Reidel Publishing Company, 1984.
[86] Y. Yu, P. G. M. Baltus, A. de Graauw, E. van der Heijden, C. S. Vaucher and A. H. M. van Roermund, "A 60 GHz Phase Shifter Integrated With LNA and PA in 65 nm CMOS for Phased Array Systems," IEEE Journal of Solid-State Circuits, vol. 45, no. 9, pp. 1697-1709, Sept. 2010.
[87] F. Caster, L. Gilreath, S. Pan, Z. Wang, F. Capolino and P. Heydari, "A 93-to-113GHz BiCMOS 9-element imaging array receiver utilizing spatial-overlapping pixels with wideband phase and amplitude control," 2013 IEEE International Solid-State Circuits Conference, San Francisco, CA, 2013, pp. 144-145.
[88] J. Paramesh, R. Bishop, K. Soumyanath and D. J. Allstot, "A four-antenna receiver in $90-\mathrm{nm}$ CMOS for beamforming and spatial diversity," IEEE Journal of Solid-State Circuits, vol. 40, no. 12, pp. 2515-2524, Dec. 2005.
[89] T. Yu and G. M. Rebeiz, "A 22-24 GHz 4-Element CMOS Phased Array With On-Chip Coupling Characterization," IEEE Journal of Solid-State Circuits, vol. 43, no. 9, pp. 2134-2143, Sept. 2008.
[90] S. Shahramian, Y. Baeyens, N. Kaneda and Y. Chen, "A 70-100 GHz Direct-Conversion Transmitter and Receiver Phased Array Chipset Demonstrating $10 \mathrm{~Gb} / \mathrm{s}$ Wireless Link," IEEE Journal of Solid-State Circuits, vol. 48, no. 5, pp. 1113-1125, May 2013.
[91] S. Kundu and J. Paramesh, "A Compact, Supply-Voltage Scalable $45-66 \mathrm{GHz}$ Baseband-Combining CMOS Phased-Array Receiver," IEEE Journal of Solid-State Circuits, vol. 50, no. 2, pp. 527-542, Feb. 2015.
[92] S. Lin, K. B. Ng, H. Wong, K. M. Luk, S. S. Wong and A. S. Y. Poon, "A 60 GHz digitally controlled RF beamforming array in 65 nm CMOS with off-chip antennas," 2011 IEEE Radio Frequency Integrated Circuits Symposium, Baltimore, MD, 2011, pp. 1-4.
[93] E. Cohen, C. Jakobson, S. Ravid and D. Ritter, "A Bidirectional TX/RX Four-Element Phased Array at 60 GHz With RF-IF Conversion Block in $90-\mathrm{nm}$ CMOS Process," IEEE Transactions on Microwave Theory and Techniques, vol. 58, no. 5, pp. 1438-1446, May 2010.
[94] A. Natarajan, A. Komijani, X. Guan, A. Babakhani and A. Hajimiri, "A 77-GHz Phased-Array Transceiver With On-Chip Antennas in Silicon: Transmitter and Local LO-Path Phase Shifting," IEEE Journal of Solid-State Circuits, vol. 41, no. 12, pp. 2807-2819, Dec. 2006.
[95] K. Raczkowski, W. De Raedt, B. Nauwelaers and P. Wambacq, "A wideband beamformer for a phased-array 60 GHz receiver in 40 nm digital CMOS," 2010 IEEE International Solid-State Circuits Conference, San Francisco, CA, 2010, pp. 40-41.
[96] M. Tabesh et al., "A 65 nm CMOS 4-Element Sub-34 mW/Element 60 GHz Phased-Array Transceiver," IEEE Journal of Solid-State Circuits, vol. 46, no. 12, pp. 3018-3032, Dec. 2011.
[97] B. Min and G. M. Rebeiz, "Single-Ended and Differential Ka-Band BiCMOS Phased Array Front-Ends," IEEE Journal of Solid-State Circuits, vol. 43, no. 10, pp. 2239-2250, Oct. 2008.
[98] A. Valdes-Garcia et al., "A Fully Integrated 16-Element PhasedArray Transmitter in SiGe BiCMOS for $60-\mathrm{GHz}$ Communications," IEEE Journal of Solid-State Circuits, vol. 45, no. 12, pp. 2757-2773, Dec. 2010.
[99] J. Staudinger, "Delay line phase shifter with selectable phase shift," U.S. Patent 2013 0,194,017 A1, Aug. 1, 2013.
[100] A. Asoodeh and M. Atarodi, "A Full $360^{\circ}$ Vector-Sum Phase Shifter With Very Low RMS Phase Error Over a Wide Bandwidth," IEEE Transactions on Microwave Theory and Techniques, vol. 60, no. 6, pp. 1626-1634, June 2012.
[101] A. Asoodeh and S. Mirabbasi, "On the Design of $n^{\text {th }}$-Order Polyphase All-Pass Filters," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 1, pp. 133-146, Jan. 2019.
[102] K. Kanda et al., " $40 \mathrm{~Gb} / \mathrm{s} 4: 1$ MUX/1:4 DEMUX in 90nm standard CMOS," ISSCC 2005 IEEE International Digest of Technical Papers. Solid-State Circuits Conference, San Francisco, CA, USA, 2005, pp. 152-590.
[103] S. Galal and B. Razavi, "40-Gb/s amplifier and ESD protection circuit in $0.18-/ \mathrm{spl} \mathrm{mu} / \mathrm{m}$ CMOS technology," IEEE Journal of Solid-State Circuits, vol. 39, no. 12, pp. 2389-2396, Dec. 2004.
[104] R. Hogervorst et al.. "CMOS low-voltage operational amplifiers with constant-gm rail-to-rail input stage," Analog Integrated Circ. Signal Proc., vol. 5, no. 2, pp. 135-146, Mar. 1994.
[105] R. Hogervorst, J. P. Tero, R. G. H. Eschauzier and J. H. Huijsing, "A compact power-efficient 3 V CMOS rail-to-rail input/output operational amplifier for VLSI cell libraries," IEEE Journal of SolidState Circuits, vol. 29, no. 12, pp. 1505-1513, Dec. 1994.
[106] D. Pepe and D. Zito, "Two mm-Wave Vector Modulator Active Phase Shifters With Novel IQ Generator in 28 nm FDSOI CMOS," IEEE Journal of Solid-State Circuits, vol. 52, no. 2, pp. 344-356, Feb. 2017.
[107] J. Zhou, H. J. Qian and X. Luo, "High-Resolution Wideband Vector-Sum Digital Phase Shifter With On-Chip Phase Linearity Enhancement Technology," IEEE Transactions on Circuits and Systems I: Regular Papers.
[108] G. H. Park, C. W. Byeon and C. S. Park, "A 60-GHz Low-Power Active Phase Shifter with Impedance-Invariant Vector Modulation in 65-nm CMOS," IEEE Transactions on Microwave Theory and Techniques, vol. 68, no. 12, pp. 5395-5407, Dec. 2020.
[109] E. V. P. Anjos, D. M. M. -. Schreurs, G. A. E. Vandenbosch and M. Geurts, "A $14-50-\mathrm{GHz}$ Phase Shifter with All-Pass Networks for 5G Mobile Applications," IEEE Transactions on Microwave Theory and Techniques, vol. 68, no. 2, pp. 762-774, Feb. 2020.
[110] H. J. Qian, B. Zhang and X. Luo, "High-Resolution Wideband Phase Shifter With Current Limited Vector-Sum," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 2, pp. 820-833, Feb. 2019.
[111] K. Tang et al., "A 4TX/4RX Pulsed Chirping Phased-Array Radar Transceiver in $65-\mathrm{nm}$ CMOS for X-Band Synthetic Aperture Radar Application," IEEE Journal of Solid-State Circuits, vol. 55, no. 11, pp. 2970-2983, Nov. 2020.

## Appendix A

## Solution of $10^{\text {th }}$-Order Polynomial Equation

In this section the solution of Eq. (3.37) and also the relation for $\omega_{\min }$ at which the $3^{r d}$-order PAF phase difference has a relative minimum are provided.

$$
\begin{align*}
& -\omega_{\min }^{10}+3 \omega_{h}^{2} \omega_{\text {min }}^{8}+2\left(\omega_{h} \omega_{l}^{2}-\omega_{h}^{3}\right) \omega_{\text {min }}^{7}- \\
& 2 \omega_{l}^{2} \omega_{h}^{2} \omega_{\text {min }}^{6}-2 \omega_{l}^{2} \omega_{h}^{4} \omega_{\text {min }}^{4}+ \\
& 2\left(\omega_{l}^{2} \omega_{h}^{5}-\omega_{l}^{4} \omega_{h}^{3}\right) \omega_{\text {min }}^{3}+3 \omega_{l}^{4} \omega_{h}^{4} \omega_{\text {min }}^{2}- \\
& \omega_{l}^{4} \omega_{h}^{6} \\
& =  \tag{A.1}\\
& \\
& -\omega_{\text {min }}^{7}\left(\omega_{\text {min }}^{3}-3 \omega_{h}^{2} \omega_{\text {min }}+2 \omega_{h}^{3}\right)+ \\
& 2 \omega_{h} \omega_{l}^{2} \omega_{\text {min }}^{3}\left(\omega_{\text {min }}^{4}-\omega_{h} \omega_{\text {min }}^{3}-\omega_{h}^{3} \omega+\omega_{h}^{4}\right)- \\
& \omega_{l}^{4} \omega_{h}^{3}\left(2 \omega_{\text {min }}^{3}-3 \omega_{h} \omega_{\text {min }}^{2}+\omega_{h}^{3}\right)
\end{align*}
$$

The factorization of the terms in parentheses generates

$$
\begin{align*}
& -\omega_{\min }^{7}\left(\omega_{\min }+2 \omega_{h}\right)\left(\omega_{\text {min }}-\omega_{4} h\right)^{2}+ \\
& 2 \omega_{h} \omega_{l}^{2} \omega_{\text {min }}^{3}\left(\omega_{\min }^{2}+\omega_{h} \omega_{\text {min }}+\omega_{h}^{2}\right)\left(\omega_{\text {min }}-\omega_{h}\right)^{2} \\
& -\omega_{l}^{4} \omega_{h}^{3}\left(2 \omega_{\text {min }}+\omega_{h}\right)\left(\omega_{\text {min }}-\omega_{h}\right)^{2} \\
= & \\
& \left(\omega_{\text {min }}-\omega_{h}\right)^{2}\left(-\omega_{\text {min }}^{8}-2 \omega_{h} \omega_{\text {min }}^{7}+\right.  \tag{A.2}\\
& 2 \omega_{h} \omega_{l}^{2} \omega_{\text {min }}^{5}+2 \omega_{h}^{2} \omega_{l}^{2} \omega_{\text {min }}^{4}+2 \omega_{h}^{3} \omega_{l}^{2} \omega_{\text {min }}^{3}- \\
& \left.2 \omega_{l}^{4} \omega_{h}^{3} \omega_{\text {min }}-\omega_{l}^{4} \omega_{h}^{4}\right)
\end{align*}
$$

Rearrangement of the terms in the second parenthesis produces

$$
\begin{align*}
&\left(\omega_{\min }-\omega_{4}\right)^{2}\left(\left(-\omega_{\min }^{8}+2 \omega_{4}^{2} \omega_{0}^{2} \omega_{\min }^{4}-\omega_{0}^{4} \omega_{4}^{4}\right)\right. \\
&-2 \omega_{4} \omega_{\min }^{3}\left(\omega_{\min }^{4}-\omega_{4}^{2} \omega_{0}^{2}\right)+ \\
&\left.2 \omega_{4} \omega_{0}^{2} \omega_{\min }\left(\omega_{\min }^{4}-\omega_{0}^{2} \omega_{4}^{2}\right)\right) \\
&=  \tag{A.3}\\
&\left(\omega_{\min }-\omega_{4}\right)^{2}\left(-\left(\omega_{\min }^{4}-\omega_{0}^{2} \omega_{4}^{2}\right)^{2}-\right. \\
& 2 \omega_{4} \omega_{\min }^{3}\left(\omega_{\min }^{4}-\omega_{0}^{2} \omega_{4}^{2}\right)+ \\
&\left.2 \omega_{4} \omega_{0}^{2} \omega_{\min }\left(\omega_{\min }^{4}-\omega_{0}^{2} \omega_{4}^{2}\right)\right) \\
&= \\
&\left(\omega_{\min }-\omega_{4}\right)^{2}\left(\omega_{\min }^{4}-\omega_{0}^{2} \omega_{4}^{2}\right) \times \\
&\left(-\omega_{\min }^{4}-2 \omega_{4} \omega_{\min }^{3}+2 \omega_{4} \omega_{0}^{2} \omega_{\min }+\omega_{0}^{2} \omega_{4}^{2}\right)
\end{align*}
$$

As shown above, the $10^{t h}$-order polynomial has a double root at $\omega_{h}$, conjugate roots at $\pm j \sqrt{\omega_{l} \omega_{h}}$, one zero at $\sqrt{\omega_{l} \omega_{h}}$ and a root at $-\sqrt{\omega_{l} \omega_{h}}$. Since one can show that $\omega_{2}$ in Fig. 3.7 is equal to $\sqrt{\omega_{l} \omega_{h}}$, none of these roots is a solution to $\omega_{\min }$. Thus, to obtain an expression of $\omega_{\min }$, the $4^{\text {th }}$-order polynomial should be solved. Given that the solution to such $4^{\text {th }}$-order polynomial can be found in the literature [55]- [56], for the purpose of brevity the solutions are not included here.

## Appendix B

## Calculation of the Notch Frequencies of the Insertion Loss of the Capacitively Loaded Balun

Since the network is assumed to be lossless, to calculate the notch frequencies one can find the frequencies at which the input impedance of the structure becomes purely reactive. In fact, because the insertion loss is the ratio of the available power to the real power delivered to the load, and when the input impedance is purely reactive no real power can be delivered to the load, therefore, the frequencies at which the input impedance is purely reactive coincide with the notch frequencies for the insertion loss of the structure. Furthermore, it can be shown that for lossless network, at the frequency of the notch of the insertion loss the input impedance is purely reactive. To find the input impedance of the capacitively loaded structure, the following circuit approach is taken.

Shown in Fig. B.1(a), the left- and right-hand structures are equivalent, however, to find the input impedance more easily, the input capacitor $C_{3}$ is decomposed into two capacitors $C_{3}-C_{4}$ and $C_{4}$. By doing so, excluding the capacitance $C_{3}-C_{4}$, the resulting network is symmetrical around the dashed line $C C^{\prime}$. Thus, to find the input impedance, one can find $Z_{i n}^{\prime}$ and $Z_{\text {in }}$ is the parallel combination of $Z_{i n}^{\prime}$ and $C_{3}-C_{4}$. To calculate the impedance $Z_{i n}^{\prime}$, the odd- and even-mode half circuit models are used as shown in Fig. B.1(b) and (c). Assuming that the input impedances obtained from the oddand even-mode stimulations are denoted as $Z_{\text {odd }}^{\prime}$ and $Z_{\text {even }}^{\prime}$, respectively, the impedance $Z_{\text {in }}^{\prime}$ is equal to $0.5 Z_{\text {odd }}^{\prime}+0.5 Z_{\text {even }}^{\prime} . Z_{\text {odd }}^{\prime}$ and $Z_{\text {even }}^{\prime}$, can be obtained as follows:

$$
\begin{align*}
& Z_{o d d}^{\prime}= \\
& \frac{Y_{L}+Y_{11}+j \omega\left(C_{2}+C_{5}\right)}{\left(Y_{11}+j \omega C_{4}\right)\left(Y_{L}+Y_{11}+j \omega\left(C_{2}+C_{5}\right)\right)-Y_{41}^{2}} \tag{B.1}
\end{align*}
$$



Figure B.1: The approach to the calculation of the input impedance of the capacitively loaded balun structure: (a) complete schematic with the different capacitors connected to the input $\left(C_{3}\right)$, O.C. $\left(C_{4}\right)$ and outputs $\left(C_{5}\right)$, (b) odd-mode stimulation, (c) even-mode stimulation.

$$
\begin{equation*}
Z_{\text {even }}^{\prime}=\frac{Y_{11}+j \omega\left(C_{1}+C_{2}\right)}{\left(Y_{11}+j \omega C_{4}\right)\left(Y_{11}+j \omega\left(C_{1}+C_{2}\right)\right)-Y_{21}^{2}}, \tag{B.2}
\end{equation*}
$$

where $Y_{L}$ and $Y_{11,21,41}$ are $1 / R_{L}$ and Y-parameters of the coupled lines, respectively. These parameters are presented below.

$$
\begin{gather*}
Y_{11}=-j 0.5 \cot \left(\theta^{\prime}\right)\left(\frac{1}{Z_{e}^{\prime}}+\frac{1}{Z_{o}^{\prime}}\right)  \tag{B.3}\\
Y_{21}=\frac{j 0.5}{\sin \left(\theta^{\prime}\right)}\left(\frac{1}{Z_{e}^{\prime}}+\frac{1}{Z_{o}^{\prime}}\right)  \tag{B.4}\\
Y_{41}=\frac{j 0.5}{\sin \left(\theta^{\prime}\right)}\left(\frac{1}{Z_{e}^{\prime}}-\frac{1}{Z_{o}^{\prime}}\right) \tag{B.5}
\end{gather*}
$$

The node numbers of the coupled lines are shown in Fig. B.1(b) and (c).
Finally, the input admittance $Y_{i n}$ is:

$$
\begin{equation*}
Y_{\text {in }}=\frac{2}{Z_{\text {odd }}^{\prime}+Z_{\text {even }}^{\prime}}+j \omega\left(C_{3}-C_{4}\right) \tag{B.6}
\end{equation*}
$$

Putting Eqs. ( $\overline{\mathrm{B} .1}$ ) and ( $\overline{\mathrm{B} .2}$ ) into Eq. ( $\overline{\mathrm{B} .6) \text {, one can show that the }}$ input impedance becomes purely reactive, if:

$$
\begin{equation*}
\left(Y_{11}+j \omega C_{4}\right)\left(Y_{11}+j \omega\left(C_{1}+C_{2}\right)\right)=Y_{21}^{2} \tag{B.7}
\end{equation*}
$$

Substituting Eqs. (B.3), (B.4) and (C.4) into Eq. (B.7), we can further simplify the relation as follows:

$$
\begin{array}{r}
4 \omega^{2} C_{2} C_{4}-2 \omega\left(C_{2}\left(\frac{1}{Z_{o}^{\prime}}+\frac{1}{Z_{e}^{\prime}}\right)+C_{4}\left(\frac{1}{Z_{o}^{\prime}}-\frac{1}{Z_{e}^{\prime}}\right)\right) \cot \left(\theta^{\prime}\right)= \\
\frac{1}{Z_{o}^{\prime 2}}-\frac{1}{Z_{e}^{\prime 2}} \tag{B.8}
\end{array}
$$

The roots of Eq. (B.8) are the notch frequencies of the insertion loss of the structure.

## Appendix C

## Evaluation of the <br> Capacitively Loaded Balun in Terms of Amplitude and Phase imbalance over the bandwidth

The main objective of this appendix is to find out if there would be any condition for the capacitively loaded balun (Fig. B.1(a)) making both its amplitude and phase imbalance (relations below) zero independent of the frequency.

$$
\begin{equation*}
\left|\frac{O_{2}}{O_{1}}\right|=0 \mathrm{~dB} \text { and } \angle O_{2}-\angle O_{1}-\pi=0 \mathrm{rad} \tag{C.1}
\end{equation*}
$$

To realize whether such the condition would exist, one can use the odd-/even-mode half circuit models as shown in Fig. B.1(b) and (c). It can be shown that both amplitude and phase imbalance can be kept at zero if and only if the even-mode output voltage is zero. As explained in Appendix B, the symmetry line can be defined for the capacitively loaded balun after a minor manipulation of the network. Consequently, using the half circuit model shown in Fig. B.1(c), the even-mode output voltage is:

$$
\begin{equation*}
V_{o, \text { even }}=\frac{Y_{21}\left(Y_{42}-j \omega C_{2}\right)-Y_{41}\left(Y_{11}+j \omega\left(C_{1}+C_{2}\right)\right)}{2 \times D E N} I \tag{C.2}
\end{equation*}
$$

$$
\begin{aligned}
& D E N= \\
& Y_{41} N U M+\left(Y_{11}+j \omega C_{4}\right)\left(\left(Y_{11}+j \omega\left(C_{1}+C_{2}\right)\right) \times\right. \\
& \left.\left(Y_{11}+j \omega\left(C_{2}+C_{5}\right)\right)-\left(Y_{42}-j \omega C_{2}\right)^{2}\right) \\
& +Y_{21}\left(\left(Y_{42}-j \omega C_{2}\right) Y_{41}-Y_{21}\left(Y_{11}+j \omega\left(C_{2}+C_{5}\right)\right)\right),
\end{aligned}
$$

where $N U M$ is the numerator of $V_{o, \text { even }} . Y_{11,21,41}$ can be found from Appendix B, and $Y_{42}$ is:

$$
\begin{equation*}
Y_{42}=-j 0.5 \cot \left(\theta^{\prime}\right)\left(\frac{1}{Z_{e}^{\prime}}-\frac{1}{Z_{o}^{\prime}}\right) . \tag{C.3}
\end{equation*}
$$

Evaluating the numerator of $V_{o, \text { even }}$, we can see that the only condition for the numerator to be zero independent of the frequency is as follows:

$$
\begin{equation*}
C_{1}=\frac{2 C_{2}}{\frac{Z_{o}^{\prime}}{Z_{o}^{\prime}}-1} \tag{C.4}
\end{equation*}
$$

As a result, if the above relation between $C_{1}$ and $C_{2}$ is met, the capacitively loaded balun exhibits both zero amplitude and phase imbalance.

## Appendix D

## Evaluation of Equation (4.6)

The quantity defined in (4.6), can be calculated by means of complex integration. From Cauchy's integral theorem, if the function $f(s)$ is analytic everywhere within a simply-connected region, then

$$
\begin{equation*}
\oint_{C} f(s) d s=0 \tag{D.1}
\end{equation*}
$$

for every simple closed path $C$ within the region. In this work, the function $f(s)$ is defined as $\ln \frac{1}{S_{11}(s)}$, where $S_{11}(s)$ is:

$$
\begin{equation*}
\frac{Z_{i n}-R_{0}}{Z_{i n}+R_{0}}, \tag{D.2}
\end{equation*}
$$

and $Z_{i n}$ and $R_{0}$ are the input impedance and the reference resistance of the network, respectively. Since $f(s)$ is a logarithmic function, the selection of the contour $C$ needs a special attention due to the existence of branch points. The associated branch cuts may fall in the region of intereset. Fig. D.1(a) shows a typical contour considered for the function $f$ assuming that there are some branch points on the right-half plane. In fact, since there is the term $Z_{\text {in }}-R_{0}$ in $S_{11}$, this term can have zeros and/or poles on the right-half plane, and therefore, one needs to add branch cuts to the contour (Fig. D.1(a)). Thus, the contour for evaluating Cauchy's residue theorem should exclude the branch cuts. It should be noted that we are assuming that passive components are used, and thus the poles and zeros of the term $Z_{i n}+R_{0}$ cannot be in the right half plane. To move the possible branch points from the right-half plane into the left-half plane (as shown in Fig. D.1(b)), the following approach is used. By doing so, the function $f$ will be analytic in the right-half plane and thus the Cauchy's theorem (Eq. (D.1)) can be applied.

Let us denote the branch points in the right-half plane by $a_{1}, a_{2} \ldots$ $a_{n}$. Each $a_{j}$ will be replaced by the corresponding branch points in the left-half plane when $1 / S_{11}$ is multiplied by the all-pass term of the form $\left(s-a_{j}\right) /\left(s+a_{j}\right)$. Since the function is now analytic in the right-half plane,


Figure D.1: The selection of contour, (a) The Contour with the branch cut due to the logarithmic function, (b) a simpler contour by moving the branch point into the left-half plane.
using Eq. (D.1), we can write:

$$
\begin{align*}
& \oint_{C} g(s) d s= \\
& \oint_{C} \ln \left[\frac{Z_{i n}(s)+R_{0}}{Z_{i n}(s)-R_{0}} \times \frac{\left(s-a_{1}\right) \ldots\left(s-a_{n}\right)}{\left(s+a_{1}\right) \ldots\left(s+a_{n}\right)}\right] d s=0 \tag{D.3}
\end{align*}
$$

The contour considered for this integral is shown in Fig. D.1(b). The semicircular part of the path is extremely large (i.e., has a radius of infinity), while the small indentations on the $j \omega$ axis are included to avoid any singularities in the integrand which may exist on the imaginary axis. The integral can be broken into an integration around the semicircle, around the small indentations and along the the imaginary axis. It can be shown that the integral around the small indentations is zero, and thus Eq. (D.3) can be written as:

$$
\begin{equation*}
\int_{-\infty}^{+\infty} g(j \omega) j d \omega+\int_{\Gamma} g(s) d s=0 . \tag{D.4}
\end{equation*}
$$

The limits of integration across the imaginary axis are from $-\infty$ to $+\infty$. The path $\Gamma$ in the second term of the above equation represents the semicircular portion of the path $C$. Because $S_{11}(j \omega)=\left|S_{11}(j \omega)\right| \exp \left(j \angle S_{11}(j \omega)\right)$, the first term of the equation can be represented in polar format using its magnitude and the phase as below.

$$
\begin{equation*}
\int_{-\infty}^{+\infty} \ln \frac{1}{\left|S_{11}(j \omega)\right|} j d \omega+\int_{-\infty}^{+\infty}\left[\angle S_{11}(j \omega)+\angle k(j \omega)\right] d \omega, \tag{D.5}
\end{equation*}
$$

where $\angle k(j \omega)$ is the phase of $\left[\left(j \omega-a_{1}\right) \ldots\left(j \omega-a_{n}\right)\right] /\left[\left(j \omega+a_{1}\right) \ldots\left(j \omega+a_{n}\right)\right]$. The second term in the above equation is zero since the phase of a real function is always an odd function of the frequency. To solve for the second term in Eq. (D.4), as the integration is performed around the semicircle with the extremely large radius, $g(s)$ is approximated by its Laurent series. Assuming that the function $\ln \left[1 / S_{11}(s)\right]$ is analytic at infinite frequency, using the Laurent series, it can be expanded as follows:

$$
\begin{equation*}
j \beta+\frac{\sum_{i} \lambda_{z i}-\sum_{i} \lambda_{p i}}{s}+\frac{A_{2}^{\infty}}{s^{2}}+\ldots \tag{D.6}
\end{equation*}
$$

where $\beta$ is 0 or $\pi$ depending on the sign of $S_{11}$, and $\lambda_{z i}$ and $\lambda_{p i}$ are the zeros and poles of $S_{11}$, respectively. The function $\ln \left[\left(s-a_{1}\right) \ldots\left(s-a_{n}\right)\right] /[(s+$ $\left.\left.a_{1}\right) \ldots\left(s+a_{n}\right)\right]$ is also analytic at infinity. Thus, by Laurent series, it can be expanded as:

$$
\begin{equation*}
-\frac{2 \sum a_{j}}{s}-\frac{A_{3}}{s^{3}}-\ldots \tag{D.7}
\end{equation*}
$$

Integrating both Eqs. (D.6) and (D.7) around the semicircle, we have:

$$
\begin{equation*}
\int_{\Gamma} g(s) d s=-j \pi\left(\sum_{i} \lambda_{z i}-\sum_{i} \lambda_{p i}\right)+j 2 \pi \sum_{j} a_{j} \tag{D.8}
\end{equation*}
$$

By substituting Eqs. (D.5) and (D.8) into (D.4), we have:

$$
\begin{equation*}
\int_{0}^{+\infty} \ln \frac{1}{\left|S_{11}(j \omega)\right|} d \omega=\frac{\pi}{2}\left(\sum_{i} \lambda_{z i}-\sum_{i} \lambda_{p i}\right)-\pi \sum_{j} a_{j} \tag{D.9}
\end{equation*}
$$

## Appendix E

## Evaluation of Equation (4.7)

Due to only using the transmission lines of the electrical length of $\pi / 2$ in the structure of Marchand balun, its frequency response is periodic with the period $2 f_{0}$. Consequently, the quantity defined in (4.6) is unbounded when integated from 0 to $+\infty$. Manipulating the quantity defined in (4.6) as what is presented in (4.7), we can now examine the frequency response of the structure within a single period. In fact, the term $1 /\left[1+\left(\frac{\omega}{\omega_{n o}}\right)^{2 n}\right]$ in (4.7) attenuates the repetitive parts of the response. The attenuation will be higher as $n$ increases.

The $\Gamma^{-1}(j \omega)$ of the structure shown in Fig. 4.1(a) is:

$$
\begin{equation*}
\Gamma^{-1}(j \omega)=S_{11}^{-1}(j \omega)=\frac{Z_{\text {in }}(j \omega)+R_{0}}{Z_{\text {in }}(j \omega)-R_{0}}=\frac{N(j \omega)}{D(j \omega)}, \tag{E.1}
\end{equation*}
$$

where $N(j \omega)$ and $D(j \omega)$ are as follows:

$$
\begin{align*}
& N(j \omega)= \\
& j \frac{\cot ^{3} \theta}{Z_{e} Z_{o}}-\cot ^{2} \theta\left(Z_{e}^{-1}+Z_{o}^{-1}\right)\left(0.5 Y_{L}+\frac{R_{0}}{Z_{e} Z_{o}}\right)  \tag{E.2}\\
& -j 0.5 \cot \theta\left(\frac{Y_{L}}{R_{0}}+\left(Z_{e}^{-1}+Z_{o}^{-1}\right)^{2}\left(0.5+R_{0} Y_{L}\right)\right) \\
& +Y_{L}\left(Z_{e}^{-1}+Z_{o}^{-1}\right)
\end{align*}
$$

and

$$
\begin{align*}
& D(j \omega)= \\
& j \frac{\cot ^{3} \theta}{Z_{e} Z_{o}}-\cot ^{2} \theta\left(Z_{e}^{-1}+Z_{o}^{-1}\right)\left(0.5 Y_{L}-\frac{R_{0}}{Z_{e} Z_{o}}\right)  \tag{E.3}\\
& -j 0.5 \cot \theta\left(\frac{Y_{L}}{R_{0}}+\left(Z_{e}^{-1}+Z_{o}^{-1}\right)^{2}\left(0.5-R_{0} Y_{L}\right)\right)
\end{align*}
$$

$Z_{e}, Z_{o}, R_{0}, Y_{L}$ and $\theta$ are the even- and odd-mode characteristic impedance, the reference impedance, $1 / R_{L}$ and $\omega l / c$, respectively. $\omega, l$ and $c$ are the
angular frequency $(2 \pi f)$, segment length and speed of light, respectively. In the above relations, it is assumed that $\sqrt{2 Y_{L} / R_{0}}=\left(1 / Z_{o}-1 / Z_{e}\right)[79]$.

To analyze the performance of Marchand balun using the revised quantity defined in (4.7), the complex function $f(s)$ is defined as:

$$
\begin{equation*}
f(s)=\frac{1}{1+\left(\frac{s}{j \omega_{n o}}\right)^{2 n}} \ln \frac{1}{\Gamma(s)}, \tag{E.4}
\end{equation*}
$$

where

$$
\begin{equation*}
\Gamma(j \omega)=\left.\Gamma(s)\right|_{s=j \omega} . \tag{E.5}
\end{equation*}
$$

The numerator $(N(s))$ and denominator $(D(s))$ of $\Gamma^{-1}(s)$ are given by:

$$
\begin{align*}
& N(s)= \\
& \frac{\operatorname{coth}^{3} \theta^{\prime}}{Z_{e} Z_{o}}+\operatorname{coth}^{2} \theta^{\prime}\left(Z_{e}^{-1}+Z_{o}^{-1}\right)\left(0.5 Y_{L}+\frac{R_{0}}{Z_{e} Z_{o}}\right)  \tag{E.6}\\
& +0.5 \operatorname{coth} \theta^{\prime}\left(\frac{Y_{L}}{R_{0}}+\left(Z_{e}^{-1}+Z_{o}^{-1}\right)^{2}\left(0.5+R_{0} Y_{L}\right)\right) \\
& +Y_{L}\left(Z_{e}^{-1}+Z_{o}^{-1}\right)
\end{align*}
$$

and

$$
\begin{align*}
& D(s)=\operatorname{coth} \theta^{\prime} \times \\
& {\left[\frac{\operatorname{coth}^{2} \theta^{\prime}}{Z_{e} Z_{o}}+\operatorname{coth} \theta^{\prime}\left(Z_{e}^{-1}+Z_{o}^{-1}\right)\left(0.5 Y_{L}-\frac{R_{0}}{Z_{e} Z_{o}}\right)\right.}  \tag{E.7}\\
& \left.+0.5 \times\left(\frac{Y_{L}}{R_{0}}+\left(Z_{e}^{-1}+Z_{o}^{-1}\right)^{2}\left(0.5-R_{0} Y_{L}\right)\right)\right],
\end{align*}
$$

where $\theta^{\prime}$ and coth are $s l / c$ and the hyperbolic cotangent, respectively. As mentioned earlier, the numerator $N(s)$ does not have any roots in the righthalf plane. However, the denominator $D(s)$ can have roots in the right-half plane. Thus, to move any potential branch points and therefore branch cuts from the right-half plane into the left-half one, the approach adopted in Appendix D is used here again. The zeros/poles of the function $D(s)$ that are located on the $j \omega$ axis are of the form:

$$
\begin{equation*}
s_{j \omega}=j \frac{c k \pi}{2 l}, k \in Z \tag{E.8}
\end{equation*}
$$

In $D(s)$, the expression in the brackets, which has a quadratic form in terms of coth $\theta^{\prime}$, can have some right-half plane roots. The roots of the quadratic
equation in terms of $\operatorname{coth} \theta^{\prime}$, are of the general form of $a_{1}+j b_{1}$ and $a_{2}+j b_{2}$, where $a_{1,2}$, and $b_{1,2}$ are the real and imaginary parts of the roots, and cot $\theta^{\prime}$ is equal to these roots. Assuming that $\theta^{\prime}=x+j y, x$ and $y$ in terms $a_{i}+j b_{i}$ where $i=\{1,2\}$ can be found from the relations below.

$$
\begin{align*}
& x_{1}=\tanh ^{-1}\left(\frac{a_{i}^{2}+b_{i}^{2}+1}{2 a_{i}}+\sqrt{\frac{b_{i}^{2}}{a_{i}^{2}}+\frac{\left(a_{i}^{2}+b_{i}^{2}-1\right)^{2}}{4 a_{i}^{2}}}\right) \\
& y_{1}=\tan ^{-1}\left(\frac{a_{i}^{2}+b_{i}^{2}-1}{2 b_{i}}+\sqrt{\left.1+\frac{\left(a_{i}^{2}+b_{i}^{2}-1\right)^{2}}{4 b_{i}^{2}}\right)}\right.  \tag{E.9}\\
& x_{2}=\tanh ^{-1}\left(\frac{a_{i}^{2}+b_{i}^{2}+1}{2 a_{i}}-\sqrt{\frac{b_{i}^{2}}{a_{i}^{2}}+\frac{\left(a_{i}^{2}+b_{i}^{2}-1\right)^{2}}{4 a_{i}^{2}}}\right) \\
& y_{2}=\tan ^{-1}\left(\frac{a_{i}^{2}+b_{i}^{2}-1}{2 b_{i}}-\sqrt{1+\frac{\left(a_{i}^{2}+b_{i}^{2}-1\right)^{2}}{4 b_{i}^{2}}}\right)
\end{align*}
$$

Finally,

$$
\begin{equation*}
s_{k}=\frac{c}{l} \theta^{\prime}=\frac{c}{l}\left(x_{i}+j\left(y_{i}+k \pi\right)\right), \quad i=\{1,2\}, k \in Z \tag{E.10}
\end{equation*}
$$

It should be noted that since $-1<\tanh (x)<+1$, the argument of $\tanh ^{-1}$ has to remain between -1 and +1 . To move the possible branch points located on the right-half plane onto the left-half one, the function $f(s)$ can be manipulated as follows.

$$
\begin{equation*}
g(s)=\frac{1}{1+\left(\frac{s}{j \omega_{n o}}\right)^{2 n}} \ln \frac{D(s)}{S_{11}(s)}, \tag{E.11}
\end{equation*}
$$

where $D(s)$ is:

$$
\begin{cases}1 & \operatorname{Re}\left\{s_{k}\right\}<0  \tag{E.12}\\ \frac{\left(s-s_{0}\right)}{\left(s+s_{0}\right)} \prod_{k=1}^{+\infty} \frac{\left(s-s_{k}\right)\left(s-s_{-k}\right)}{\left(s+s_{k}\right)\left(s+s_{-k}\right)} & \operatorname{Re}\left\{s_{k}\right\}>0\end{cases}
$$

$s_{k}$ 's are the branch points of Eq. (E.10). Now, the function $g(s)$ is analytic everywhere on the right-half plane except at some points, namely, removable singularities [85]. Using Cauchy's residue theorem, the revised quantity is
given by:

$$
\begin{align*}
& \int_{0}^{+\infty} \frac{1}{1+\left(\frac{\omega}{\omega_{n o}}\right)^{2 n}} \ln \frac{1}{\left|S_{11}(j \omega)\right|} d \omega= \\
& \pi\left|\operatorname{Re}\left\{\sum_{m=0}^{n-1} \lim _{s \rightarrow a_{m}}\left(s-a_{m}\right) g(s)\right\}\right| \tag{E.13}
\end{align*}
$$

The singularities $a_{m}$ are:

$$
\begin{equation*}
a_{m}=-j w_{n o} e^{j \frac{2 m+1}{2 n} \pi} \tag{E.14}
\end{equation*}
$$

The contour considered for the integration is shown in Fig. E.1. The small indentations on the $j \omega$ axis are considered to include the singularities whose values can be found from Eq. (E.8). From Eq. (E.10), since the real part of $s_{k}$ is independent of $k, k$ in Eq. (E.12) does not need to change to an infinitely large value. In fact, the terms $\left(s-s_{ \pm k}\right)$ and $\left(s+s_{ \pm k}\right)$ can be neglected for those values of $k$ that meet both of the following conditions.

$$
\begin{align*}
& \left|\operatorname{Re}\left(s \pm s_{ \pm k}\right)\right|_{s=a_{m}} \ll\left|\operatorname{Im}\left(s \pm s_{ \pm k}\right)\right|_{s=a_{m}} \\
& |k| \gg \frac{l}{\pi c}\left|\operatorname{Im}\left(s \pm\left(s_{k}-\frac{c}{l} k \pi\right)\right)\right|_{s=a_{m}} \tag{E.15}
\end{align*}
$$



Figure E.1: The contour considered for the revised quantity.

