HIGH PERFORMANCE RADIO-FREQUENCY AND
MILLIMETER-WAVE FRONT-END INTEGRATED CIRCUITS
DESIGN IN SILICON-BASED TECHNOLOGIES

A Dissertation
Presented to
The Academic Faculty

by

Jihwan Kim

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
August 2011

Copyright © Jihwan Kim 2011
HIGH PERFORMANCE RADIO-FREQUENCY AND MILLIMETER-WAVE FRONT-END INTEGRATED CIRCUITS DESIGN IN SILICON-BASED TECHNOLOGIES

Approved by:

Dr. Kevin T. Kornegay, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Geoffrey Ye Li
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Emmanouil M. Tentzeris
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Paul A. Kohl
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Chang-Ho Lee
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: April 18, 2011
To my wife and two sons
ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation to my advisor, Prof. Kevin T. Kornegay, for all the opportunities, guidance, and encouragement that he has provided. I am also grateful to Prof. Emmanouil M. Tentzeris, Prof. Geoffrey Ye Li, Dr. Chang-Ho Lee, and Prof. Paul A. Kohl for their time in reviewing my dissertation and serving as my defense committee members. I especially thank to Dr. Chang-Ho Lee for his great support and guidance given to me throughout this study. I am also indebted to Dr. Kyutae Lim for his effort to provide the best research environment, and Dr. Joy Laskar for giving me research opportunities.

Much appreciation is extended to my colleagues and friends at Georgia Institute of Technology, especially Mr. Hyungwook Kim, Mr. Youngchang Yoon, Mr. Hamhee Jeon, Mr. Hyun-Woong Kim, Mr. Eungjung Kim, Mr. Yan-Yu Huang, Mr. Kun Seok Lee, Mr. Stephen Kim, Mr. Kwanyeob Chae, Mr. Seungho Baek, Dr. Kilhoon Lee, Mr. Michael Oakely, Mr. Seungil Yoon, Dr. Kyu Hwan An, Dr. Ockgoo Lee, Dr. Joonhoi Hur, Dr. Taejoong Song, Dr. Sanghyun Woo, Dr. Jaehyouk Choi, Dr. Jong Min Park, Dr. Kwan-Woo Kim, Dr. Sang Min Lee, Dr. Hyungsoo Kim, and Dr. Dong Ho Lee, for their assistance and cooperation. I am especially grateful to Dr. Kyu Hwan An for his helpful guidance through this research. I would also like to acknowledge Dr. Tonmoy Mukherjee, Mr. Mohammad Omer, Mr. Duane Howard, Mr. Michael Kroger, and Dr. Javier Alvarado, Jr. for sharing ideas and discussions.

I would like to thank Samsung Design Center engineers, Dr. Woonyun Kim, Dr. Jae Joon Chang, Dr. Wangmyong Woo, Dr. Changhyuk Cho, Dr. Yunseo Park, Dr.
Minsik Ahn, Dr. Ki Seok Yang, Dr. Jeongwon Cha, and Dr. Jeonghu Han for their support and assistance.

Special thanks to the Cambridge Culture Foundation, Korea for the financial support of a part of my graduate study.

I am especially grateful to my parents, my grandmother who passed away a year ago, and my sister for their continued support of my education. Last but not least, I would like to express my deepest love and gratitude to my wife, Eun Jung Kwon, for her invaluable support through my half of life. She has been a person with whom I can find peace whenever I struggled with difficulties that I faced pursuing my degree. This dissertation would not have been completed without the help and support of all of these individuals.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv  
LIST OF TABLES ix  
LIST OF FIGURES x  
LIST OF SYMBOLS AND ABBREVIATIONS xvii  
SUMMARY xxi  

CHAPTER  
1. Introduction 1  
1.1. Wireless Communications 1  
1.2. RF and MMW Transceivers 5  
1.3. Silicon-Based Technologies for RF and MMW ICs 6  
1.4. Motivation 9  
1.5. Organization 11  

2. Challenges and Issues of Transceiver Front-End IC Designs in Silicon-Based Technologies 13  
2.1. Introduction 13  
2.2. Issues of Silicon Substrate 14  
2.3. Limitations of Passive Components 16  
2.3.1. Inductors 17  
2.3.1.1. Eddy current 17  
2.3.1.2. Proximity effect and skin effect 19  
2.3.1.3. Techniques for inductor Q enhancement 21  
2.3.1.4. Inductor model 23
### 2.3.2. Capacitors
- **2.3.2.1. Metal-insulator-metal (MIM) capacitor**
- **2.3.2.2. Parasitic extraction of MIM capacitor**

### 2.3.3. Transmission lines
- **2.3.3.1. Hybrid-type transmission line**
- **2.3.3.2. Characteristic impedance**

### 2.4. Limitations of Active Components
- **2.4.1. MOSFETs**
  - **2.4.1.1. Transconductance**
  - **2.4.1.2. Device reliability**
  - **2.4.1.3. Linearity**
- **2.4.2. SiGe HBTs**
  - **2.4.2.1. Speed consideration**
  - **2.4.2.2. Noise-gain trade-off**
  - **2.4.2.3. Device breakdown**

### 2.5. Conclusion

### 3. Design of Multi-Mode CMOS Power Amplifiers for High Data-Rate Wireless Communications

#### 3.1. Introduction

#### 3.2. Linear Multi-Mode CMOS Power Amplifier
- **3.2.1. Power combining transformer**
  - **3.2.1.1. Transformers as power combiners**
  - **3.2.1.2. Efficiency of power combining transformers**
- **3.2.2. Analysis on efficiency degradation of transformers**
- **3.2.3. Design of a linear CMOS PA**
  - **3.2.3.1. Concurrent resizing and combining technique**
5.4. Design of a W-band Low-Noise Amplifier 145
   5.4.1. Noise characteristics of a SiGe HBT 145
   5.4.2. Circuit design 146
5.5. Design of a W-band Down-Conversion Mixer 151
   5.5.1. Double-balanced mixer topology 153
   5.5.2. Circuit design 155
   5.5.3. Measurement results 158
5.6. Design of an Integrated W-Band Receiver 161
   5.6.1. Receiver structure 161
   5.6.2. Measurement results 162
5.7. Conclusion 167

6. Conclusions and Future Works 169
   6.1. Contribution to the field 169
   6.2. Future Research Directions 170

PUBLICATIONS 172
REFERENCES 174
VITA 183
LIST OF TABLES

Table 1. Commercial PA products for WiMAX/LTE applications. 7
Table 2. Conductivity, resistivity, and skin depth values of commonly-used metals at 10 GHz. 20
Table 3. PAPRs of various modulation schemes. 44
Table 4. Summary of device sizes. 70
Table 5. Performance comparison of watt-level linear CMOS PAs. 90
Table 6. Calculated input impedances and efficiencies of the combining transformers. 99
Table 7. EM characteristic of the half-section of the PSCT (at 2.4 GHz). 113
Table 8. Performance comparison of high-power linear CMOS PAs. 125
Table 9. Performance comparison of W-band down-conversion mixers in SiGe technologies. 159
Table 10. Performance comparison of W-band receivers in SiGe technologies. 167
LIST OF FIGURES

Figure 1. Revenue for the telecommunication market. (a) Worldwide. (b) In the United States. 2
Figure 2. Wireless services revenue by category in the United States. 3
Figure 3. Frequency spectrum and various applications. 4
Figure 4. Block diagram of a typical direct-conversion transceiver system. 5
Figure 5. Evolution of CMOS and SiGe technology nodes. 8
Figure 6. Research outline. 10
Figure 7. Parasitic-aware design procedure of RF and MMW ICs. 14
Figure 8. Signal coupling and noise injection in the silicon substrate. 15
Figure 9. Substrate eddy current generation in the silicon substrate. 18
Figure 10. Uneven current redistribution of AC current due to eddy current in a spiral inductor. 19
Figure 11. Spiral inductor employing patterned ground shield (PSG). 21
Figure 12. Spiral inductor with varying widths of metal traces. 22
Figure 13. Double-stacked inductor. 23
Figure 14. Commonly used equivalent model of a monolithic spiral inductor. 24
Figure 15. MIM capacitor. (a) Cross section. (b) Equivalent model. 25
Figure 16. Comparison of MIM capacitor models (100 fF) (a) Capacitance values. (b) $Q$ of capacitors. 26
Figure 17. Hybrid-type transmission line. 27
Figure 18. Characteristic impedances of conventional CPW transmission lines ($H=9.25 \mu m$). 29
Figure 19. Cross-section of the hybrid-type transmission line in IBM BiCMOS 8HP process. 30
Figure 20. Characteristic impedances of the hybrid-type transmission lines obtained by embedded impedance calculator in the IBM 8HP design-kit ($H=9.25 \mu m$).

Figure 21. Calculation of the characteristic impedance. (a) A transmission line with length of $L$. (b) A transmission line in the even-mode. (c) A transmission line in the odd-mode.

Figure 22. 3-D structure of the hybrid-type transmission line in IBM 8HP design-kit.

Figure 23. Characteristic impedances of the hybrid-type transmission lines ($L=300 \mu m$, $W=5 \mu m$, $H=9.25 \mu m$).

Figure 24. $S_{11}$ of the quarter-wavelength transmission line with ground termination.

Figure 25. Resistance and inductance of the quarter-wavelength hybrid-type transmission line. (a) Resistance. (b) Inductance.

Figure 26. Equivalent NMOS transistor model.

Figure 27. Equivalent SiGe HBT model.

Figure 28. Block diagram of a parallel-combining PA with discrete power-control.

Figure 29. Operation of a multi-power mode PA.

Figure 30. Two types of power combining transformers. (a) PCT. (b) SCT.

Figure 31. Equivalent transformer models. (a) Single 1:n transformer. (b) Equivalent model of a single 1:n transformer. (c) Equivalent model of an $M$-primary PCT. (d) Equivalent model of an $M$-primary SCT.

Figure 32. Diagram of a parallel-combining PA using a PCT in case when one of PAs is disabled.

Figure 33. Equivalent circuit models of $M$-primary PCT and SCT with $N$-idle inputs. (a) PCT. (b) SCT. (c) Three different cases of idle primary input terminations.

Figure 34. Analyzed efficiencies of 4-primary 1:2 transformers. (a) PCT with $N$-idle primary inputs that are open-circuited. (b) SCT with $N$-idle primary inputs that are open-circuited.

Figure 35. Analyzed efficiencies of 4-primary 1:2 transformers. (a) PCT with $N$-idle primary inputs that are shorted to ground. (b) SCT with $N$-idle primary inputs that are shorted to ground.
Figure 36. Analyzed efficiencies of 4-primary 1:2 transformers. (a) PCT with \(N\)-idle primary inputs that are loaded by a capacitor. (b) SCT with \(N\)-idle primary inputs that are loaded by a capacitor.

Figure 37. 3D structures of two-primary 1:2 transformers. (a) PCT. (b) SCT.

Figure 38. Simulated maximum available efficiencies of two-primary 1:2 transformers with EM-simulated models. (a) PCT. (b) SCT.

Figure 39. Structures of parallel-combining PAs with discrete power-control using a PCT. (a) Conventional combining structure. (b) Concurrent combining structure.

Figure 40. Implementation of a concurrent combining PA. (a) Diagram of a PA with an integrated two-primary 1:2 PCT. (b) Generalized schematic of each driver and power stage. (c) Operation conditions of the PA.

Figure 41. Simplified schematic of designed PA.

Figure 42. EM simulation setup of the two-primary PCT using HFSS.

Figure 43. Lead inductance de-embedding setup. (a) For lead1. (b) For lead2.

Figure 44. EM simulation results of two-primary PCT. (a) Maximum available gain. (b) \(Q\).

Figure 45. A Chip micro-photo of the concurrent combining multi-mode PA.

Figure 46. Measurement results of multi-mode linear CMOS PA: gain and PAE (at 2.1 GHz).

Figure 47. Measurement results of multi-mode linear CMOS PA: EVM (at 2.1 GHz).

Figure 48. Measurement results of multi-mode linear CMOS PA: S-parameters. (a) \(S_{11}\) and \(S_{22}\). (b) \(S_{21}\) and \(S_{12}\).

Figure 49. A varactor-based tunable capacitive matching circuit for load impedance modulation.

Figure 50. Simulated voltage swings at the drain of each power stage when RF input power is -3 dBm. (a) For different VSWRs with same angle (0°). (b) For VSWR of 10:1 and different angles.

Figure 51. Simulation results of the varactor-based matching circuit. (a) \(Q\). (b) Capacitance.

Figure 52. Simulation results of the PA with the varactor-based matching circuit. (a) Gain. (b) PAE.
Figure 53. Entire schematic of designed linear multi-mode CMOS PA with load impedance modulation technique.

Figure 54. Micro-photograph of designed multi-mode linear CMOS PA with load impedance modulation technique (1.1 × 1.8 mm²).

Figure 55. Measurement results of multi-mode PA with load impedance modulation in three operation modes: gain and PAE (at 2.5 GHz).

Figure 56. Measurement results of multi-mode linear PA with load impedance modulation in three operation modes: DC current (at 2.5 GHz).

Figure 57. Measurement results of multi-mode linear PA with load impedance modulation: S-parameters. (a) S₁₁ and S₂₂. (b) S₂₁ and S₁₂.

Figure 58. Measurement results of multi-mode linear PA with load impedance modulation in three operation modes: EVM (at 2.5 GHz). (a) For WLAN OFDM 64-QAM signal. (b) For WiMAX OFDM 64-QAM signal.

Figure 59. Measurement results of multi-mode linear PA with load impedance modulation in the high power mode: constellation and spectrum (at 2.5 GHz) (a) For WLAN OFDM 64-QAM modulated signal (P_out=20.5 dBm). (b) For WiMAX OFDM 64-QAM modulated signal (P_out=16 dBm).

Figure 60. Output power combining technique for high-power PA applications.

Figure 61. Conventional power combining transformers. (a) PCT. (b) SCT.

Figure 62. Calculated efficiencies of the power combining transformers with varying turn ratio (n) and the number of combining inputs (M). (a) For the PCT. (b) For the SCT.

Figure 63. Calculated input impedances of the power combining transformers with varying turn ratio (n) and the number of combining inputs (M). (a) For the PCT. (b) For the SCT.

Figure 64. Schematic diagram of proposed PSCT.

Figure 65. Calculated efficiencies and input impedances of three power combining transformers: PCT, SCT, and PSCT with four combining inputs (M=4). (a) Efficiencies. (b) Input impedances.

Figure 66. Schematic diagram of proposed PSCT with T-equivalent transformer model to account for non-ideal coupling effect.
Figure 67. Implementation of proposed 4-primary 1:2 PSCT (a) Conceptual diagram. (b) Actual layout.

Figure 68. Simulation setup for characterization of half-section PSCT. (a) For the primary winding. (b) For the secondary winding. (c) For the coupling effect and the mutual inductance.

Figure 69. Size comparison between the proposed PSCT and SCT. (a) proposed PSCT ($M=4$, $n=2$). (b) SCT ($M=4$, $n=1$).

Figure 70. Simulated performances of the proposed PSCT ($M=4$, $n=2$) and SCT ($M=4$, $n=2$). (a) Insertion loss. (b) Input impedance.

Figure 71. Diagram of proposed PSCT-based high-power linear PA.

Figure 72. Schematic diagram of designed CMOS PA using the PSCT as a power combiner.

Figure 73. Micro-photograph of fabricated PA (2.1 x 1.64 mm²).

Figure 74. Measurement results of designed PA. (a) Gain, DE, and PAE. (b) Gain versus $P_{out}$ at different frequencies.

Figure 75. Measurement results of the PSCT-based high-power linear PA: S-parameters. (a) $S_{11}$ and $S_{22}$. (b) $S_{21}$ and $S_{12}$.

Figure 76. Measurement results of designed PA: EVM. (a) For WLAN 54 Mbps 64-QAM signal. (b) For WiMAX 54 Mbps 64-QAM signal.

Figure 77. Measurement results of the PSCT-based high-power linear PA: constellation and spectrum (at 2.4 GHz). (a) For WLAN OFDM 64-QAM signal ($P_{out}=23$ dBm). (b) For WiMAX OFDM 64-QAM signal ($P_{out}=19$ dBm).

Figure 78. Diagram of a direct down-conversion receiver system.

Figure 79. Diagram of the 60 GHz receiver in SiGe demonstrated in [30].

Figure 80. Diagrams of reported W-Band SiGe receivers. (a) In [35]. (b) In [36].

Figure 81. Diagram of proposed W-band (94 GHz) receiver front-end system.

Figure 82. Cross section of SiGe HBT.

Figure 83. Schematic of a planar-type Marchand balun.

Figure 84. Coupled quarter-wavelength transmission line.

Figure 85. Total capacitance in a coupled line.
Figure 86. Cross-section of the hybrid-type coupled transmission line. 136

Figure 87. Simulated coupling coefficient of the hybrid-type coupled transmission line ($W=5 \, \mu m$, $S=20 \, \mu m$). 136

Figure 88. 3-D structure of the hybrid-type coupled quarter-wavelength transmission line ($W=5 \, \mu m$, $S=20 \, \mu m$, $D=5 \, \mu m$). 137

Figure 89. Simulated characteristics of the hybrid-type coupled quarter-wavelength transmission line. 138

Figure 90. Configuration of a coupled quarter-wavelength transmission line as a balun. 139

Figure 91. Simulated amplitude balance of a coupled quarter-wavelength transmission line. 139

Figure 92. Coupled transmission line with unequal ground-plane widths. 140

Figure 93. Coupling coefficient of coupled transmission line with unequal ground-plane widths. 141

Figure 94. 3-D structure of designed W-band planar-type Marchand balun. 142

Figure 95. Performance of designed W-band planar-type Marchand balun. (a) Micro-photo of the fabricated balun. (b) Amplitude balance. (c) Phase balance. 143

Figure 96. Primary noise sources in an HBT. 145

Figure 97. Cascode structure for the W-band LNA. 147

Figure 98. Simulated characteristics of the cascode structure as a function of $J_C$. (a) $f_T$ and $N_{F_{\text{min}}}$. (b) Associate gain. (c) $R_{S,\text{opt}}$. (d) $\Gamma_{\text{opt}}$. 148

Figure 99. Schematic of designed W-band LNA. 149

Figure 100. Simulation results of designed W-band LNA. 149

Figure 101. Noise figure in the mixer. (a) Single-side band (SSB) noise figure. (b) Double-side band noise figure. 151

Figure 102. Double-balanced mixer topology. 152

Figure 103. Signal waveforms. (a) RF signal. (b) LO signal. (c) IF signal. 153

Figure 104. Schematic of designed W-band mixer. 155

Figure 105. Design flow for the W-band mixer. 156
Figure 106. Micro-photograph of designed W-band mixer.

Figure 107. Measurement results of designed W-band mixer. (a) Conversion gain, noise figure, and port-to-port isolations. (b) $P_{1dB\text{ in}}$ with LO power of 0 dBm.

Figure 108. Designed W-band receiver. (a) Diagram of the receiver. (b) Chip micro-photograph of fabricated W-band receiver (the size is 1.82 mm$^2$).

Figure 109. Measurement setup for the W-band receiver. (a) For a conversion gain. (b) For a noise figure.

Figure 110. Measurement results of designed W-band receiver. (a) Conversion gain vs RF input frequency. (b) Noise figure vs RF input frequency.

Figure 111. Measurement results of designed W-band receiver. (a) Conversion gain and noise figure vs LO input power. (b) IF output power vs RF input power. (c) RF and LO input matching. (d) Port-to-port isolations.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGC</td>
<td>automatic gain control</td>
<td></td>
</tr>
<tr>
<td>AM-AM</td>
<td>amplitude-to-amplitude</td>
<td></td>
</tr>
<tr>
<td>AM-PM</td>
<td>amplitude-to-phase</td>
<td></td>
</tr>
<tr>
<td>BEOL</td>
<td>back-end-of-line</td>
<td></td>
</tr>
<tr>
<td>BiCMOS</td>
<td>bipolar and CMOS</td>
<td></td>
</tr>
<tr>
<td>BV</td>
<td>breakdown voltage</td>
<td></td>
</tr>
<tr>
<td>BV_{CBO}</td>
<td>breakdown voltage when the emitter is open-circuited</td>
<td></td>
</tr>
<tr>
<td>BV_{CEO}</td>
<td>breakdown voltage when the base is open-circuited</td>
<td></td>
</tr>
<tr>
<td>CDMA</td>
<td>code division multiple access</td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>common gate</td>
<td></td>
</tr>
<tr>
<td>CMOS</td>
<td>complementary metal oxide semiconductor</td>
<td></td>
</tr>
<tr>
<td>CPW</td>
<td>coplanar waveguide</td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>common source</td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>driver amplifier</td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>drain efficiency</td>
<td></td>
</tr>
<tr>
<td>DSB</td>
<td>double-side band</td>
<td></td>
</tr>
<tr>
<td>EBR</td>
<td>efficiency back-off ratio</td>
<td></td>
</tr>
<tr>
<td>EBR_{achieved}</td>
<td>achieved EBR</td>
<td></td>
</tr>
<tr>
<td>EER</td>
<td>efficiency enhancement ratio</td>
<td></td>
</tr>
<tr>
<td>EM</td>
<td>electro-magnetic</td>
<td></td>
</tr>
<tr>
<td>EVM</td>
<td>error vector magnitude</td>
<td></td>
</tr>
<tr>
<td>FCC</td>
<td>Federal Communications Commission</td>
<td></td>
</tr>
<tr>
<td>FGCPW</td>
<td>finite ground-plane coplanar waveguide</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>$f_{\max}$</td>
<td>maximum unity power-gain frequency</td>
<td></td>
</tr>
<tr>
<td>FMCW</td>
<td>frequency modulated continuous wave</td>
<td></td>
</tr>
<tr>
<td>$f_T$</td>
<td>maximum transit frequency</td>
<td></td>
</tr>
<tr>
<td>GaAs</td>
<td>gallium arsenide</td>
<td></td>
</tr>
<tr>
<td>GSM</td>
<td>global system for mobile</td>
<td></td>
</tr>
<tr>
<td>HBT</td>
<td>hetero-junction bipolar transistor</td>
<td></td>
</tr>
<tr>
<td>HEMT</td>
<td>high electron mobility transistor</td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>high-power</td>
<td></td>
</tr>
<tr>
<td>IC</td>
<td>integrated circuit</td>
<td></td>
</tr>
<tr>
<td>IF</td>
<td>intermediate frequency</td>
<td></td>
</tr>
<tr>
<td>IIP$_2$</td>
<td>input-referred second-order intercept point</td>
<td></td>
</tr>
<tr>
<td>IIP$_3$</td>
<td>input-referred third-order intercept point</td>
<td></td>
</tr>
<tr>
<td>IL</td>
<td>insertion loss</td>
<td></td>
</tr>
<tr>
<td>IMD$_3$</td>
<td>third order inter-modulation distortion</td>
<td></td>
</tr>
<tr>
<td>InGaP</td>
<td>indium gallium phosphide</td>
<td></td>
</tr>
<tr>
<td>InP</td>
<td>indium phosphide</td>
<td></td>
</tr>
<tr>
<td>ISM</td>
<td>industrial, scientific, and medical</td>
<td></td>
</tr>
<tr>
<td>LNA</td>
<td>low-noise amplifier</td>
<td></td>
</tr>
<tr>
<td>LO</td>
<td>local oscillator</td>
<td></td>
</tr>
<tr>
<td>LP</td>
<td>low-power</td>
<td></td>
</tr>
<tr>
<td>LPF</td>
<td>low pass filter</td>
<td></td>
</tr>
<tr>
<td>LTE</td>
<td>long term evolution</td>
<td></td>
</tr>
<tr>
<td>MIM</td>
<td>metal-insulator-metal</td>
<td></td>
</tr>
<tr>
<td>MMW</td>
<td>millimeter-wave</td>
<td></td>
</tr>
<tr>
<td>MOSFET</td>
<td>metal oxide semiconductor field effect transistor</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------</td>
<td></td>
</tr>
<tr>
<td>MP</td>
<td>medium-power</td>
<td></td>
</tr>
<tr>
<td>NF&lt;sub&gt;min&lt;/sub&gt;</td>
<td>minimum noise figure</td>
<td></td>
</tr>
<tr>
<td>NF</td>
<td>noise figure</td>
<td></td>
</tr>
<tr>
<td>OFDM</td>
<td>orthogonal frequency division multiplexing</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;1dB&lt;/sub&gt;</td>
<td>output referred 1-dB compression point</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;1dB_in&lt;/sub&gt;</td>
<td>input referred 1-dB compression point</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>power amplifier</td>
<td></td>
</tr>
<tr>
<td>PAE</td>
<td>power added efficiency</td>
<td></td>
</tr>
<tr>
<td>PAPR</td>
<td>peak-to-average power ratio</td>
<td></td>
</tr>
<tr>
<td>PCB</td>
<td>printed circuit board</td>
<td></td>
</tr>
<tr>
<td>PCT</td>
<td>parallel-combining transformer</td>
<td></td>
</tr>
<tr>
<td>PGS</td>
<td>patterned ground shield</td>
<td></td>
</tr>
<tr>
<td>P&lt;sub&gt;sat&lt;/sub&gt;</td>
<td>saturated (peak) output power level</td>
<td></td>
</tr>
<tr>
<td>PSCT</td>
<td>parallel-series combining transformer</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>quality-factor</td>
<td></td>
</tr>
<tr>
<td>QAM</td>
<td>quadrature amplitude modulation</td>
<td></td>
</tr>
<tr>
<td>RF</td>
<td>radio frequency</td>
<td></td>
</tr>
<tr>
<td>SCT</td>
<td>series-combining transformer</td>
<td></td>
</tr>
<tr>
<td>SiGe</td>
<td>silicon germanium</td>
<td></td>
</tr>
<tr>
<td>SNR</td>
<td>signal-to-noise ratio</td>
<td></td>
</tr>
<tr>
<td>SRF</td>
<td>self-resonant frequency</td>
<td></td>
</tr>
<tr>
<td>SSB</td>
<td>single-side band</td>
<td></td>
</tr>
<tr>
<td>VCO</td>
<td>voltage-controlled oscillator</td>
<td></td>
</tr>
<tr>
<td>VNA</td>
<td>vector network analyzer</td>
<td></td>
</tr>
<tr>
<td>WCDMA</td>
<td>wideband CDMA</td>
<td></td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>WiMAX</td>
<td>worldwide interoperability for microwave access</td>
<td></td>
</tr>
<tr>
<td>WLAN</td>
<td>wireless local area network</td>
<td></td>
</tr>
<tr>
<td>WPAN</td>
<td>wireless personal area network</td>
<td></td>
</tr>
</tbody>
</table>
SUMMARY

The wireless communication market has been explosively growing while being fueled by innovative advancements in digital communication/processing technologies along with revolutionary progression in integrated circuit (IC) technologies, manifested both in the speed of the transistors and the complexity of IC systems. The silicon-based process technologies such as complementary metal oxide semiconductor (CMOS) and silicon-germanium (SiGe) are the key driver in the wireless communication IC industry that can enable implementations of fully integrated single-chip transceivers and hence true one-chip systems of mobile terminal devices at low cost. However, developments of high-performance IC blocks and systems overcoming intrinsic vulnerabilities of the silicon-based technologies have been difficult and laborious tasks.

The objective of this research is to explore limitations and challenges of radio frequency (RF) and millimeter-wave (MMW) front-end IC designs using the silicon-based technologies, and to develop effective circuit topologies and design techniques to improve the target performance of the ICs. CMOS power amplifiers (PAs) for high data-rate mobile communications (e.g., WiMAX) and SiGe receiver front-end circuits and systems for 90 to 94 GHz (W-band) long-range wireless accesses are aimed for exploration in this research.

Detailed contents presented in this dissertation are categorized as following:

1. Fundamental limitations of active and passive devices of the silicon-based technologies (silicon CMOS and SiGe) are discussed from the perspective of transceiver IC designs.
2. The design methodology for a multi-mode linear PA fabricated in a 0.18 µm CMOS technology with enhanced low-power efficiency is presented along with thorough analysis on the novel power combining structure. The design technique and procedure to implement a multi-mode (low-power, medium-power, and high-power modes) multi-standard (WLAN and WiMAX) class-AB CMOS PA is discussed. The PA performance at various operation modes is optimized by employing a novel load impedance modulation technique using varactor-based tunable matching networks. These research works will be published in IEEE Journal of Solid-State Circuits (JSSC) in May 2011 [Publication 1]. A part of the research works has been presented in IEEE Radio Frequency Integrated Circuits (RFIC) Symposium in 2010 [Publication 2].

3. A new type of power combining transformer, which is appropriate for high-power (i.e., greater than 33-dBm output power) PA applications, is introduced. The proposed transformer performs the parallel (current) combining and the series (voltage) combining simultaneously in a single structure, supplementing drawbacks of conventional types of power combining transformers. The class-AB PA with the proposed combining transformer is implemented in a 0.18 µm CMOS technology, achieving a $P_{1dB}$ of 31.5 dBm and a $P_{sat}$ of 34 dB with a peak PAE of 34.9%. Measurement results show the effectiveness of the proposed PA structure for high-data rate wireless communication standards such as WiMAX. The result of this research work is submitted to IEEE Journal of Solid-State Circuits (JSSC) and is under review [Publication 3].
4. The design methodologies for W-band (90 to 94 GHz) receiver building blocks such as a low-noise amplifier (LNA), a balun (balanced-to-unbalanced), a mixer using a SiGe technology is presented focusing on performance optimization techniques. Detailed design procedures for individual circuits are discussed with supportive simulation and measurement results. An integrated receiver system is also implemented in a 200-GHz-$f_T$ SiGe technology, achieving a maximum conversion gain of 36.3 dB and a minimum noise figure of 10 dB at 91 GHz. The designed receiver demonstrated the highest conversion gain among recently reported receivers built in silicon-based technologies operating beyond 90 GHz. These research works have been presented in IEEE Radio Frequency Integrated Circuits (RFIC) Symposium in 2008 [Publication 4], and published in Electronics Letters in 2009 [Publication 5].
1.1. Wireless Communications

The global telecommunication market has been rapidly growing thanks to advancements in communication theories, digital signal processing, and integrated circuits (ICs) technologies. The revenue for the telecommunication industry is expected to reach $4.94 trillion worldwide and $278.4 billion in the United States by 2013 (Figure 1) [1]. The trend of this remarkable expansion of the wireless communication market is also observed from the fact that the number of wireless subscribers in the United States has become over 90% of the population in 2010 and will grow further to over 95% of the population by 2013 [1]. Undoubtedly, the wireless market is the mainspring of the growing telecommunication industry.

Meanwhile, ever-increasing demands from consumers for broadband wireless communications has accelerated developments of smart mobile devices that adopt various forms of wireless communication standards at radio frequency (RF) bands. Evolved from early generations of wireless communication standards such as global system for mobile (GSM) or code division multiple access (CDMA), more advanced mobile communication standards such as IEEE 802.11a/b/g wireless local area network (WLAN), wideband CDMA (WCDMA), worldwide interoperability for microwave access (WiMAX), and/or long term evolution (LTE) have become available to terminal users for accommodation of high speed data communications. As shown in the graph of Figure 2, data communications will account for about 44% ($93 billion) of overall services spending in
Figure 1. Revenue for the telecommunication market. (a) Worldwide. (b) In the United States.
Inevitably, developments of faster and more reliable mobile communication technologies are desired for manufacturers more than ever to meet the demands of consumers.

As frequency resources are being exhausted by competing communication standards at the RF bands, millimeter-wave (MMW) bands (greater than 20 GHz) are getting increasing interests in the telecommunication industry as a breakthrough. Much wider channel bandwidths can potentially realize Gb/s data throughput with a relaxed requirement on spectral efficiency. Promising frequency bands for the broadband next-generation wireless accesses are 59 to 64 GHz in the V-band (50 to 75 GHz), 77 to 79 GHz and 94 GHz in the W-band (75 to 110 GHz). The Federal Communications Commission (FCC) has assigned a spectrum of 59 to 64 GHz, which is called industrial,
scientific, and medical (ISM) band, for general unlicensed purposes [2]. The IEEE 802.15.3.c broadband wireless personal area network (WPAN) [3] and the WirelessHD™ [4] are good examples of emerging applications in the 59 to 64 GHz band. However, the path loss is particularly large at this frequency band due to the absorption lines of the oxygen (e.g., the attenuation of the path is about 20 dB/km at 60 GHz). Therefore, a wireless communication range is limited within a few tens of meters. In contrast, the propagation attenuation is not problematic in the W-band (75 to 110 GHz). Thus, more reliable long-range wireless accesses can be realized. Automotive radars for collision avoidance (77 to 79 GHz), passive imaging for astronomy, defense, or security purposes (94 GHz), and extremely wideband communications (110 to 120 GHz) are potential applications in this MMW frequency range. Various applications at different frequency bands with available process technologies are graphically summarized in Figure 3 [1].

![Figure 3. Frequency spectrum and various applications.](image-url)
1.2. RF and MMW Transceivers

In a mobile communication device, a transceiver (transmitter and receiver) front-end system is the most essential component which is interfacing with physical channels. Although there are diverse digital modulation schemes for various applications invented in recent years, the transceiver architecture has not been fundamentally changed for decades. In a generic homodyne transceiver system, as illustrated in Figure 4, a received signal from an antenna is amplified by a low-noise amplifier (LNA) preserving the signal-to-noise ratio (SNR) as much as possible. Down-conversion is then performed by a mixer to shift down the frequency of the signal to the baseband using a frequency synthesizer as a local oscillator (LO) source. The signal is amplified again by automatic gain control (AGC) blocks and filtered by a low pass filter (LPF) before reaching a

![Figure 4. Block diagram of a typical direct-conversion transceiver system.](image-url)
digital processing unit. The transmitter side consists of almost same functional blocks except for a power amplifier (PA) that transmits the up-converted signal to the antenna at required power level. This structure, so called a direct-conversion transceiver, has been widely adopted to implement RF and MMW wireless communication systems.

1.3. Silicon-Based Technologies for RF and MMW ICs

Full integration of digital, analog, and RF functional blocks has been the ultimate goal of the wireless communication industry for minimization of the manufacturing cost. A true single-chip transceiver implemented in silicon-based process technologies can enable a total integration of all the sub-systems within a mobile terminal device. Complementary metal oxide semiconductor (CMOS) or silicon germanium (SiGe) BiCMOS (bipolar and CMOS) technologies may be great solutions for the full integration of the radio terminal thanks to their low-cost material and good versatility [5]. The silicon wafer costs less than one-fiftieth of that of III-V compound semiconductor technologies such as gallium arsenide (GaAs), indium phosphide (InP), or indium gallium phosphide (InGaP).

Vigorous research and industrial activities to realize the fully integrated CMOS RF front-end systems for various applications have been made in recent years [6]-[8]. Nonetheless, the PA is still left as the biggest bottleneck toward the realization of the single-chip transceiver in a CMOS technology since intrinsic weakness of metal oxide semiconductor field effect transistors (MOSFETs) in terms of linearity and reliability makes the design of CMOS PAs for high-power and high-data-rate applications very challenging [9]-[14]. Therefore, despite of great exertions on improving performances of
the CMOS PAs, III-V compound hetero-junction bipolar transistors (HBTs) or high electron mobility transistors (HEMTs) have been dominating technologies of choice in the PA market as observed from the Table 1 which lists recent PA products for the fourth generation (4G) wireless communication standards such as WiMAX or LTE.

### Table 1. Commercial PA products for WiMAX/LTE applications.

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Model</th>
<th>Freq (GHz)</th>
<th>Gain (dB)</th>
<th>P_{1dB} (dBm)</th>
<th>EVM**</th>
<th>Power Supply (V)</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analog Devices [15]</td>
<td>ADL5571</td>
<td>2.5-2.7</td>
<td>29</td>
<td>29</td>
<td></td>
<td>3.3</td>
<td>GaAs HBT</td>
</tr>
<tr>
<td>Anadigics [16]</td>
<td>AWM6422</td>
<td>2.3-2.4</td>
<td>30</td>
<td>30</td>
<td>2.5%@22dBm</td>
<td>4%@23.5dBm</td>
<td>3.3 InGaP HBT</td>
</tr>
<tr>
<td>SiGe Semicon [17]</td>
<td>SE7262L</td>
<td>2.5-2.7</td>
<td>34</td>
<td>-</td>
<td></td>
<td>3.3</td>
<td>SiGe HBT</td>
</tr>
<tr>
<td>RFMD [18]</td>
<td>SZM-3066Z</td>
<td>3.3-3.8</td>
<td>33</td>
<td>33.5</td>
<td>2.5%@26dBm</td>
<td></td>
<td>InGaP HBT</td>
</tr>
<tr>
<td>RFMD [18]</td>
<td>SZA-3044Z</td>
<td>2.7-3.8</td>
<td>-</td>
<td>31</td>
<td>2.5%@24dBm</td>
<td></td>
<td>InGaP HBT</td>
</tr>
<tr>
<td>TriQuint [19]</td>
<td>TGA2702 SM</td>
<td>2.3-2.8</td>
<td>-</td>
<td>29.8</td>
<td>2.5%@23dBm</td>
<td></td>
<td>InGaP HBT</td>
</tr>
<tr>
<td>Skyworks Solutions [20]</td>
<td>SKY77441</td>
<td>2.3-2.5</td>
<td>-</td>
<td>-</td>
<td>3.5%@27dBm</td>
<td></td>
<td>InGaP HBT</td>
</tr>
</tbody>
</table>

* P_{1dB}: output referred 1-dB compression point
** EVM: error vector magnitude

The MMW transceiver front-end circuits and systems have been also developed typically using the III-V compound semiconductor technologies [21]-[26]. With comparative advantages, the III-V technologies have demonstrated their appropriateness in implementations of MMW ICs. In terms of device performances such as a maximum transit frequency \( f_T \), a maximum unity power-gain frequency \( f_{max} \), a breakdown
voltage, and the availability of high quality-factor (Q) passives, GaAs/InP/InGaP HBTs/HEMTs are far superior to traditional MOSFETs and/or early generations of SiGe HBTs. However, due to the considerable cost effectiveness and a higher level of integration, the silicon-based technologies have also started attract the MMW IC industry. Along with continuous down-scaling of the technology node, bringing advancements in the speed of transistors, a deep sub-micrometer CMOS or advanced generations of SiGe technologies are becoming plausible options to implement the MMW front-end circuit components as well as the integrated transceivers (Figure 5). With incorporation of innovative design and layout techniques to compensate for the inherent drawbacks and limitations of the silicon substrate and active/passive devices, performances of CMOS/SiGe MMW transceivers and composing circuits are becoming comparable to III-V counterparts [27]-[41].

Figure 5. Evolution of CMOS and SiGe technology nodes.
1.4. Motivation

Obviously, the silicon-based technologies are advantageous for the reduced production cost and the higher level of integration of wireless communication systems as mentioned in the previous section. However, many difficulties in designing RF and MMW ICs using CMOS or SiGe technologies, which will be discussed in Chapter 2, still makes III-V compound technologies more suitable for the implementation of high performance front-end circuits. For example, a significantly low breakdown voltage of the MOSFETs and SiGe HBTs compared to GaAs/InP/InGaP HBTs/HEMTs is the most critical obstacle in designing PAs targeting for generation of watt-level output power. Furthermore, a bias-dependent input/output capacitance of the MOSFET is the fundamental drawback which severely limits linearity performance of the circuits. Lossy silicon substrate is also a big design hurdle that constrains implementation of high-quality passive components such as inductors, capacitors, or transmission lines. In the MMW operating frequencies (e.g., in the W-band), these issues become more challenging, and must be effectively handled.

The objective of the research in this dissertation is to explore limitations and challenges of RF and MMW front-end IC designs using the silicon-based technologies, and to investigate effective design techniques to improve the circuit performances. As illustrated in the research outline of Figure 6, the RF band of 2.4 to 2.5 GHz and the MMW W-band of 90 to 94 GHz are aimed for the research.

In the RF band, the PA, the most intricate building block in a transceiver system, is particularly focused to betterment its performance for next-generation wireless communications standards. To accommodate multiple standards and various power
modes, the PA must support flexible operation conditions with optimized performance. Specifically, a low-efficiency characteristic of the PA at lower output power levels (backed-off power levels from the peak output power level) must be resolved by innovative approaches with a structural viewpoint. Thus, fully integrated CMOS PAs with a novel combining structure that enables the efficiency optimization in various operation modes are implemented. Thorough investigations and analyses on monolithic transformers are performed as a primary part of the performance enhancement for the PA. Furthermore, a novel power combining transformer, which is suitable for the generation of output power of greater than 33 dBm, is proposed and designed. To demonstrate the effectiveness of the proposed power combining transformer, a fully-integrated high-power (31.5-dBm P_{1dB} and 34-dBm P_{sat}) linear CMOS PA is implemented. This research...
attempts to elevate potential appropriateness of a CMOS technology for high-end PA designs (i.e., CMOS PAs for the WiMAX application).

The receiver building blocks; an LNA, a mixer, and a passive balun (balanced-to-unbalanced) for the W-band applications are designed using 200-GHz-$f_T$ SiGe technology. Critical passive structures are electro-magnetically (EM) characterized for an accurate modeling. The integrated W-band receiver chain achieves a very high conversion gain (greater than 36 dB) due to the effective design and layout techniques, which successfully demonstrate the suitability of SiGe technology for the high-performance MMW front-end ICs. Various design hurdles originated from the inherent drawbacks of the silicon-based technology are handled by innovative design techniques with well-optimized design procedures.

1.5. **Organization**

Based on the motivation described in the previous section, this dissertation is organized for the best presentation of the research achievements as following.

Chapter 2 discusses various issues and challenges of the silicon-based technologies for RF and MMW IC designs. Substrate-related constraints and limitations of passive devices such as inductors, capacitors, and transmission lines are explained. Moreover, intrinsic characteristics of MOSFETs and SiGe HBTs are reviewed from the perspective of front-end circuit designs.

Chapter 3 presents a new power combining and a discrete resizing technique using an on-chip transformer to implement a multi-mode PA in a 0.18 μm CMOS technology. A novel combining structure that maximizes the efficiency of the PA at
lower power modes are introduced with theoretical analysis and experimental data. A load impedance modulation technique to optimize the PA performance at different power modes is also explored.

Chapter 4 introduces a novel power combining transformer structure: a parallel-series combining transformer (PSCT). Using the proposed power combining transformer proposed, a high-power linear PA is implemented in a 0.18 μm CMOS technology. Characteristics and advantages of the proposed transformer comparing to conventional types of monolithic power combining transformers are discussed.

Chapter 5 presents the design methodology of W-band front-end building blocks such as an LNA, a balun, and a mixer using 200-GHz-\(f_T\) SiGe technology. Detailed design procedures of MMW circuits to maximize the performance of individual circuit components and an integrated receiver chain are described followed by supporting simulation and measurement results.

Finally, the summary of the dissertation and conclusions is followed in Chapter 6. Future works and directions are also proposed.
CHAPTER 2

CHALLENGES AND ISSUES OF TRANSCEIVER FRONT-END IC DESIGNS IN SILICON-BASED TECHNOLOGIES

2.1. Introduction

The design of RF and MMW transceiver front-end ICs using silicon-based technologies accompanies various challenges and design issues: (1) The conductive substrate raises isolation problems between closely placed active and passive devices, (2) a high quality-factor ($Q$) for monolithic inductors and transmission lines is not achievable, (3) accurate device models are required as wavelengths of AC signals become comparable to physical dimensions of the devices, (4) parasitic components of active and passive devices must be well-extracted and properly included in actual circuit designs, and (5) constraints on device reliabilities must be carefully considered. Therefore, thorough awareness of limitations of active and passive devices and understanding of parasitic effects on circuit performances at high operating frequencies are strongly required in RF and MMW IC designs using silicon-based technologies. The design procedure becomes an amalgamation of utilizations of various design tools such as schematic simulators for transient or harmonic balance analyses and parasitic extractors or EM simulators for a modeling of secondary parasitic effects on active/passive devices as illustrated in Figure 7.

This chapter introduces typical limitations of silicon-based process technologies from the perspective of RF and MMW IC designs. Isolation issues due to the conductive silicon substrate are discussed in Section 2.2, and constraints of passive components for
high-frequency applications are explained in Section 2.3. Section 2.4 deals with characteristics and intrinsic limitations of active devices (MOSFETs and SiGe HBTs). The conclusion follows in Section 2.5.

2.2. Issues of Silicon Substrate

A major superiority of III-V compound process technologies such as GaAs or InP to silicon-based technologies such as CMOS or SiGe is their highly resistive low-loss substrate. Due to small dimensions of modern process technologies and a high level of integration, active and passive devices that perform completely different functions are placed very close to each other in state-of-the-art integrated systems. Inevitably, cross-
talks between adjacent functional blocks become critical issues in RF and MMW system designs in which individual building blocks must be well isolated from one another. As presented in the simplified diagram of the silicon substrate in Figure 8, a finite resistance (about 10 Ω-cm) through the $p$-type silicon substrate associated with oxide and diffusion capacitors introduces non-negligible conductive paths between adjacent active or passive devices. Because of these unwanted signal paths, a serious signal coupling or a noise injection may occur in a highly integrated system that has thousands of transistors acting as critical noise sources. Moreover, large signal swings coming from transceiver front-end blocks such as PAs can be coupled to other critical blocks (e.g., frequency synthesizers or LNAs) that must be absolutely isolated because of their high sensitivity to the noise. In addition, a low resistivity of the silicon substrate provides the RF signals with conductive paths to ground attenuating strength and quality of the signal.

Figure 8. Signal coupling and noise injection in the silicon substrate.
Various techniques can be employed to minimize the unwanted signal coupling and the noise injection through the conductive silicon substrate. For example, guard-rings and deep $N$-wells in a triple well CMOS process are widely utilized to suppress the cross-talk and improve the signal integrity [42], [43]. Similarly, deep trench (vertical wall) techniques in a SiGe process can effectively enhance the inter-device isolations. In addition, extensive use of substrate contacts that tie up the silicon substrate with ground enables bypassing of the coupled noise to ground with a little penalty of the signal attenuation [44]. However, despite of these attempts, the perfect isolation and the zero signal/noise coupling between adjacent blocks are impossible to achieve. To consider various effects caused by conductive substrate, accurate schematic models of a substrate resistance and a capacitance must be properly included in simulations.

### 2.3. Limitations of Passive Components

At high frequencies, performances of circuits and systems are strongly affected by limitations of passive components that are functioning critical roles. The most frequently used passive components in RF and MMW designs are spiral inductors, capacitors, and transmission lines. Because of the conductive silicon substrate, complicated loss mechanisms inevitably degrade qualities of the passive components. The quality-factor, $Q$, of the passive components is defined as the ratio of the amount of energy stored in the component to the amount of energy dissipated per a cycle. Therefore, an expression of the $Q$ can be given as [45]

$$Q = \frac{E_{\text{stored}}}{E_{\text{dissipated}}} = \frac{2\pi}{\omega} \left( \frac{W_m + W_e}{P_e \times T} \right) = \omega \left( \frac{W_m + W_e}{P_e} \right),$$ (1)
where $W_m$ is the magnetically stored energy, $W_e$ is the electrically stored energy, $P_e$ is the electrically dissipated energy, and $T$ is the time period of the excited sinusoidal input. As indicated in the (1), the $Q$ of the passive structures is reduced as the magnetic and electrical energy stored within the structure is decreased. In silicon-based process technologies, considerable amount of the EM energy leaks through the lossy substrate, which degrades the $Q$ of the passive components.

2.3.1. Inductors

2.3.1.1. Eddy current

For monolithic spiral inductors, the substrate loss is a critical source of performance degradation [46]. The substrate loss is mainly categorized as a capacitive loss and a magnetic loss. The capacitive loss is due to a voltage difference between the conductor and the grounded substrate that gives rise to a capacitive coupling between them. The magnetic loss is generated by induced currents from penetrating magnetic fields of the conductor into the substrate that cause an extra resistive loss. The substrate eddy current is a dominant source of the substrate loss. As illustrated in Figure 9 [46], a secondary AC current, so called the eddy current, is induced in the substrate in an opposite direction to the conductor AC current. The eddy current, which tends to have a higher current density as the distance between the conductor and the substrate is decreased, generates a negative mutual inductance and reduces the magnetic field of the conductor. Thus, an overall inductance (reactance) of the conductor is reduced, which consequently reduces the $Q$ the inductor. If the substrate has a high conductivity (or a
small resistivity) as in the silicon substrate, a substantial amount of the eddy current is produced and degradation of the inductor’s $Q$ becomes a serious issue.

The eddy current also introduces an uneven redistribution of the AC current density in metal conductors for monolithic spiral inductors. Figure 10 represents the mechanism of the current density redistribution in a spiral inductor caused by the eddy currents [46]. The time-varying magnetic field produced by outer metal traces ($B_{\text{coil}}$) partially impedes inner metal traces and generates a secondary magnetic field inside the conductor ($B_{\text{eddy}}$) in such a way that both magnetic fields cancel out each other. This secondary magnetic field becomes a source of the circular eddy current in the metal conductor. The inner side of the conductor will have the eddy current that has a same direction as the conductor’s AC current, and the current density will be increased. On the
other hand, the opposite side of the conductor will have the eddy current with an opposite
direction to the conductor’s AC current, which reduces the effective current density.
Therefore, the current density within a single conductor will be redistributed unevenly
enlarging the AC resistance (or RF resistance) of the conductor at high frequencies. This
uneven current redistribution puts limitations on implementing on-chip spiral inductors
with a large number of turns. Thus, a large effective inductance with a high $Q$ value is not
realizable.

2.3.1.2. Proximity effect and skin effect

The proximity effect and the skin effect [47] also introduce AC current density
redistributions to the metal conductors. The proximity effect occurs when conductors are
placed very close to each other. Neighboring magnetic fields from adjacent conductors
interact with the self-magnetic field of the conductor canceling the mutual inductance.

Figure 10. Uneven current redistribution of AC current due to eddy current in a spiral inductor.
Therefore, a uniform current distribution is disturbed and the effective AC resistance of the conductor is increased. The skin effect is caused by time-varying magnetic fields of the conductor that raises the electric field that impedes the current flowing on the inner side of the conductor. Thus, the surface of the conductor will have a higher AC current density than that of the core of the conductor, and the AC resistance of the conductor is enlarged. The depth below the surface at which the AC current density decays to \(1/e\) (about 0.37) of the current density of the infinitely thick plane conductor is called “skin depth”, and is expressed as

\[
\delta = \sqrt{\frac{2\rho}{\omega \mu}},
\]

where \(\delta\) (m) is the skin depth, \(\rho\) (C/m\(^3\)) is charge density, and \(\mu\) (H/m) is the permeability.

To reduce the skin effect, the thickness, the width, and the conductivity of the metal traces must be controlled to minimize the AC resistance. The skin depths of typical metal conductors at the frequency of 10 GHz are summarized in Table 2.

<table>
<thead>
<tr>
<th>Metals</th>
<th>Conductivity ((10^5 \text{ s/cm}))</th>
<th>Resistivity ((10^{-6} \Omega/\text{cm}))</th>
<th>Skin Depth ((\mu\text{m}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminum</td>
<td>3.8</td>
<td>2.6</td>
<td>0.817</td>
</tr>
<tr>
<td>Copper</td>
<td>5.8</td>
<td>1.7</td>
<td>0.660</td>
</tr>
<tr>
<td>Gold</td>
<td>4.1</td>
<td>2.4</td>
<td>0.786</td>
</tr>
<tr>
<td>Silver</td>
<td>6.2</td>
<td>1.6</td>
<td>0.640</td>
</tr>
</tbody>
</table>
2.3.1.3. Techniques for inductor $Q$ enhancement

Considering diverse EM effects on the monolithic spiral inductors, the AC resistance (RF resistance) of inductors can be expressed as

$$R_{RF} = \frac{l}{w \rho \delta (1 - e^{-t/\delta})}, \quad (3)$$

where $l$, $w$, and $t$ are the length, the width, and the thickness of the metal trace, $\rho$ is the conductivity of the metal, and $\delta$ is the skin depth as in the (2). In a CMOS technology, various techniques have been utilized to enhance the inductor’s $Q$ by minimizing the AC resistance. To reduce the effect of the substrate eddy current, a patterned ground shield (PGS) underneath the metal conductors can be employed as illustrated in Figure 11 [48].

The metal slots in a 90°-shifted to the direction of AC current on the metal traces reduce the eddy current generation in the ground plane and effectively increase the $Q$ of the

![Figure 11. Spiral inductor employing patterned ground shield (PSG).](image_url)
inductor. However, this technique is not appropriate for high frequency applications because a substantial parasitic capacitance is generated between the conductor and the ground plane, which lowers a self-resonant frequency (SRF) of the inductor.

The uneven redistribution of the AC current density caused by time-varying magnetic fields within a spiral inductor can be mitigated if the inductor has varying widths for metal traces [49]. A strong coupling of the magnetic field from the outer trace to the inner trace will be reduced by relatively narrower widths of the inner metal traces (Figure 12). Multi-stacked spiral inductor is another way to implement high-$Q$ on-chip spiral inductors with a large inductance values as shown in Figure 13 [50]. For example, two spiral inductors on different metal layers can be connected in series such that the AC current can flow in a same direction in two inductors. Thus, the total inductance of the connected inductor becomes

![Figure 12. Spiral inductor with varying widths of metal traces.](image-url)
\[ L_T = L_1 + L_2 + 2M, \quad (4) \]

where \( L_1 \) and \( L_2 \) are self inductance of two spiral inductors, respectively, and \( M \) is the mutual inductance between two inductors. The coupling coefficient is given as

\[ k = \frac{M}{\sqrt{L_1 L_2}}, \quad (5) \]

and the total inductance will be \( 4L \) if we assume that \( k=1 \) and \( L_1=L_2=L \). In general, a \( n \)-stacked inductor can have an inductance of \( n^2L \) in an ideal case. However, most of standard CMOS processes do not support multiple thick-metal layers, which limits the usage of this approach.

![Double-stacked inductor](image)

\textbf{Figure 13. Double-stacked inductor.}

2.3.1.4. Inductor model

To reflect various EM effects on inductor performance into the circuit simulation, an accurate equivalent model of the inductor can be used to reduce simulation time and
cost. A commonly used inductor model for the silicon-based process technologies is presented in Figure 14. The model exhibits ohmic resistances and parasitic capacitances. $L_S$, $R_S$, and $C_S$ are the self inductance, the series resistance and the capacitance, respectively. $C_S$ accounts for the metal-to-metal capacitance between two input terminals of the inductor that is caused by fringing field between separate metal layers. $C_{ox}$ represents the parasitic oxide capacitance between the conductor metal and the substrate. $R_{si}$ and $C_{si}$ are the series resistance inside the silicon substrate due to finite resistivity. $R_{se}$ accounts for an additional ohmic loss due to the eddy current caused by the magnetic coupling between the spiral inductor metal trace and the low-resistivity silicon substrate. To use the equivalent model of the inductors adequately, experimental verifications are required to determine the appropriate values of each components of the model.

![Figure 14. Commonly used equivalent model of a monolithic spiral inductor.](image-url)
2.3.2. Capacitors

2.3.2.1. Metal-insulator-metal (MIM) capacitor

Monolithic capacitors are mainly classified into three kinds: microstrip capacitors, interdigital capacitors, and metal-insulator-metal (MIM) capacitors. Open-stub microstrip transmission lines can be used to realize a capacitance at MMW frequencies. The interdigital capacitors have applications where moderate capacitance values are needed. The MIM capacitors are the most widely used for RF and MMW IC designs, but they require multi-level process because the effective capacitance is generated between parallel metal planes (or electrodes). All capacitors exhibit a parasitic inductance due to their finite size and series resistance due to contact/electrode resistance. The cross section and the equivalent model of the MIM capacitor in multi-level CMOS/SiGe processes are presented in Figure 15. The parasitic inductance ($L_s$) associated with the series

![Figure 15. MIM capacitor. (a) Cross section. (b) Equivalent model.](image)
capacitance \( (C_s) \), the parasitic parallel capacitance \( (C_p) \), and the parasitic capacitance to ground \( (C_{si}) \) may form the SRF of the capacitor below the operating frequency, so the size of the MIM capacitor must be carefully determined for high frequency applications.

2.3.2.2. Parasitic extraction of MIM capacitor

Accurate models of the MIM capacitors can be obtained by parasitic extractions or EM simulations. For MMW applications, small capacitance values (a few hundreds of \( fF \)) are often required. Thus, a parasitic resistance due to metal contacts and vias may contribute to the insertion loss (IL), substantially degrading the \( Q \) of the capacitor. This effect becomes critical if the operating frequency of the application is beyond 60 GHz, where the \( Q \) of the capacitor is comparable to that of the inductors. Figure 16 shows difference between a schematic model and a parasitic-extracted model of 100-\( fF \) MIM capacitor (IBM 8HP BiCMOS) up to frequency of 110 GHz. For large MIM capacitors,

![Figure 16. Comparison of MIM capacitor models (100 fF) (a) Capacitance values. (b) \( Q \) of capacitors.](image-url)
long inter-connections also act as inductive transmission lines, and must be EM extracted to be accurately accounted in simulations.

2.3.3. Transmission lines

2.3.3.1. Hybrid-type transmission line

Conventional microstrip transmission lines are widely used for MMW applications. However, absence of ground shields at the top layer complicates an analysis and an accurate modeling of EM characteristics. To take advantages of side ground shields, we can implement a hybrid-type transmission line that is similar to finite ground-plane coplanar waveguide (FGCPW) transmission lines except for the existence of via connections between the side ground shields and the bottom ground plane. The cross section of the hybrid-type transmission line is presented in Figure 17, where $W$ is the width of the top conductor, $S$ is the gap between the conductor and the side ground shield, and $H$ is the distance between the top conductor and the bottom ground plane. Due to the via connections, same potential around all ground shields is ensured. Comparing to the

Figure 17. Hybrid-type transmission line.
conventional microstrip transmission lines, the hybrid-type lines support well-confined electric and magnetic fields, so more accurate analyses and computer-aided simulations are possible. These transmission lines can be used for interconnections between active circuits including matching networks, enhancing signal isolations by providing well-defined ground return paths and minimizing an unwanted EM coupling to adjacent structures. Furthermore, they can be easily integrated with active circuits without entailing any parasitic discontinuities in the ground plane.

2.3.3.2. Characteristic impedance

Given geometric parameters, the characteristic impedance of conventional coplanar waveguide (CPW) lines can be calculated letting \( a = W \) and \( b = W + 2S \) as [51]

\[
Z_o = \frac{60\pi}{\sqrt{\varepsilon_{\text{eff}}}} \cdot \frac{1}{\frac{K(k)}{K(k')} + \frac{K(kl)}{K(kl')}}
\]  \hspace{1cm} (6)

where

\[
k = \frac{a}{b}
\]  \hspace{1cm} (7)

\[
k' = \sqrt{1 - k^2}
\]  \hspace{1cm} (8)

\[
kl' = \sqrt{1 - kl^2}
\]  \hspace{1cm} (9)

\[
k = \frac{\tanh(a\pi / 4H)}{\tanh(b\pi / 4H)}
\]  \hspace{1cm} (10)

\[
\varepsilon_{\text{eff}} = \frac{1 + \varepsilon_r}{1 + \frac{K(k')}K(kl)}
\]  \hspace{1cm} (11)
and $K(k)$ is elliptic integrals of the first kind. Because of the foundry process, $H$ will be fixed parameter (9.25 $\mu$m in IBM 8HP BiCMOS technology), whereas other parameters can be chosen arbitrarily. A few examples of characteristic impedances of the CPW lines calculated with varying $W$ and $S$ are presented in Figure 18. Although this calculation is very well approximating the trend of the characteristic impedance with various dimensions, actual characteristic impedance of the hybrid-type transmission line will be different since it has additional vias between the top ground shields and the bottom plane.

The IBM 8HP BiCMOS process technology provides the hybrid-type transmission line models up to 200 GHz. This process features seven back-end-of-line (BEOL) metal layers with 4-$\mu$m-thickness top aluminum metallization. The cross section of the hybrid-type transmission line built on the IBM 8HP BiCMOS process is presented in Figure 19. The design-kit also provides a characteristic impedance calculator.
Characteristic impedances of the transmission lines with various dimensions are obtained using the embedded impedance calculator and are plotted in Figure 20. As observed from the Figure 18 and Figure 20, there is significant difference in calculated characteristic impedances of the hybrid-type transmission lines obtained by embedded impedance calculator in the IBM 8HP design kit ($H=9.25 \, \mu m$).

Figure 19. Cross-section of the hybrid-type transmission line in IBM BiCMOS 8HP process.

Figure 20. Characteristic impedances of the hybrid-type transmission lines obtained by embedded impedance calculator in the IBM 8HP design kit ($H=9.25 \, \mu m$).
impedances even though we consider that Figure 18 is for the conventional CPW transmission lines which do not employ side vias. Therefore, EM simulations are necessary to verify the transmission line models.

The ADS Momentum was used to simulate EM characteristics of the quarter-wavelength hybrid-type transmission line. The line features $W=5 \, \mu m$ and $L=300 \, \mu m$. The values of $S$ were varied from $4 \, \mu m$ to $20 \, \mu m$ with a step of $4 \, \mu m$. With EM simulated data, characteristic impedance of the transmission line was obtained from S-parameters. As shown in Figure 21, the transmission line with a length of $L$ applied by two ports at each side has two different modes of excitation that are considered. First, in the even-mode case, total signal leaving the port 1 is given by

$$S_e = S_{11} + S_{12} = \frac{z \cdot \coth(\gamma l) - 1}{z \cdot \coth(\gamma l) + 1},$$

where $l=L/2$, $\gamma=\alpha+j\beta$ ($\alpha$ is attenuation constant and $\beta$ is a propagation constant), and $z=Z/Z_0$. Therefore, we can have

$$z \cdot \coth(\gamma) = \frac{1+S_e}{1-S_e}. \quad (13)$$

Similarly, in the odd-mode case, total signal leaving the port 1 can be expressed as

$$S_o = S_{11} - S_{12} = \frac{z \cdot \tanh(\gamma l) - 1}{z \cdot \tanh(\gamma l) + 1}, \quad (14)$$

and we can obtain

$$z \cdot \tanh(\gamma l) = \frac{1+S_o}{1-S_o} \quad (15)$$

Thus, by the (13) and the (15), the complex propagation constant and the characteristic impedance can be calculated as
Based on these equations, the characteristic impedances of the hybrid-type transmission line models with a length of 300 µm (as shown in Figure 22) are calculated and plotted in Figure 23. Because of the side ground shields, the characteristic impedances are calculated as follows:

\[
\tanh^2 (\gamma l) = \frac{(1 + S_o)/(1 - S_e)}{(1 + S_e)/(1 - S_o)}
\]  
\[
z^2 = \left( \frac{1 + S_e}{1 - S_e} \right) \left( \frac{1 + S_o}{1 - S_o} \right).
\]  

Figure 21. Calculation of the characteristic impedance. (a) A transmission line with length of L. (b) A transmission line in the even-mode. (c) A transmission line in the odd-mode.
impedances for all various dimension parameters have smaller values compared to ones for the conventional CPW transmission lines. The EM model has smaller characteristic impedances for all different values of the distance between the top conductor and the side ground shields. However, this data for the EM model is still different from the data for the model provided by the design-kit. This is because the EM simulator accounts for a fringing capacitance between the signal line and the bottom ground plane to more extent than the impedance calculator in the design-kit does.

Figure 22. 3-D structure of the hybrid-type transmission line in IBM 8HP design-kit.
The scattering parameters (S-parameters) are also simulated to compare two models: the EM simulated model and a model from IBM 8HP design-kit. The quarter-wavelength transmission line is supposed to transform the zero impedance (shorted to ground) to the infinite impedance (open) at the design frequency. As shown in Figure 24, the EM simulated model features a smaller quarter-wavelength frequency compared to the design-kit model. Beyond the quarter-wavelength frequency, the function of the transmission line is changed from the inductive one to the capacitive one as illustrated in Figure 25. We can observe that resistance and inductance values of two models change their polarities at 89 GHz (EM model) and 94 GHz (design-kit model), respectively. It must be considered that the design-kit model shows a more optimistic characteristic in terms of the resistive loss.

Figure 23. Characteristic impedances of the hybrid-type transmission lines ($L=300$ μm, $W=5$ μm, $H=9.25$ μm).
2.4. Limitations of Active Components

Active devices, or transistors, are the most critical components in circuit blocks. They consume a DC current from external power supplies and perform necessary functions. Important figure-of-merits of active devices includes a transconductance ($g_m$), a noise factor, a maximum transit frequency ($f_T$), and a maximum unity power-gain.
frequency \((f_{\text{max}})\). Reliability and linearity of the active devices are also important characteristics to be carefully considered for optimal circuit designs. In this section, MOSFETs are reviewed from the perspective of the transconductance, reliability, and linearity characteristics, and SiGe HBTs are discussed in terms of a noise, a speed, and reliability characteristics. Circuit designers must thoroughly understand these characteristics and limitations of different active devices to appropriately determine right design environment for specific target applications.

## 2.4.1. MOSFETs

### 2.4.1.1. Transconductance

The fundamental function of active devices is amplifying the signal with a finite gain. The transconductance, or \(g_m\), which is the ratio of the input voltage to the output current, for MOSFET device is given as

\[
g_m = \frac{2I_{\text{DC}}}{V_{GS} - V_t},
\]

where \(V_{GS}\) is the gate-source voltage, \(V_t\) is the threshold voltage of the transistor, and \(I_{\text{DC}}\) is the DC current flowing through the transistor. The \(g_m\) of BJT device has relation of

\[
g_m = \frac{qI_{\text{DC}}}{kT}.
\]

Typically, the \(g_m\) of MOSFET is lower than that of BJT at a room temperature due to lower value of \(q/kT\) as compared to \(V_{GS} - V_t\) in common process technologies. The higher \(g_m\) is always desired in analog and RF circuit designs to achieve a high gain. Moreover, higher \(g_m\) is also necessary to minimize the input-referred noise contribution from each noise source of transistors in order to enhance the performance of the LNAs and down-
conversion mixers. In RF PA designs, a low $g_m$ necessitates the use of large device cells to accommodate required power capability, which results in a large power consumption. A low power-gain due to a low $g_m$ also requires high driving input power for the PA, consequently reducing the power added efficiency (PAE). In short, a low $g_m$ of MOSFETs is a big challenge in RF PA designs in terms of the load impedance matching and power efficiency [52].

2.4.1.2. Device reliability

A low breakdown voltage of MOSFET devices is one of key challenges in RF and MMW IC designs. Generally, PAs in transmitters deal with large swings of AC signals, and the device reliability is the most important factor for robust performance of the circuit. When the transistor is under the condition of high voltage stress, excessive electric fields cause the breakdown of the gate-oxide dielectric. This gate-oxide breakdown, especially in between the drain and gate terminals of the MOSFET device, is the biggest problem because it may destruct the device. The drain-to-substrate junction breakdown is also critical problem even though it is not as devastating to the transistor as the drain-to-gate oxide breakdown. The junction breakdown occurs when the high reverse-biased voltage is applied across the $pn$-junction. Excessive charge due to the high reverse-biased voltage causes the avalanche breakdown in the $pn$-junctions. The channel breakdown (between the drain and the source) due to the impact ionization or the gate leakage current, along with the hot-carriers-degradation effect, is another mechanism that is also destructive to the device. These breakdown effects on the MOSFET devices limit
the maximum voltage swing that the device can handle, bringing serious design hurdles to the large-signal circuit designs.

2.4.1.3. Linearity

Every circuit blocks in the transceiver system needs to operate as linearly as possible regardless of the signal strength they deal with. The linearity of active devices is obviously the key factor that determines overall linearity of circuits and systems. Inherently, the MOSFET device has various intrinsic sources of non-linearity. These are the gate-source junction capacitance, the gate-source transconductance, the gate-drain transconductance, and the drain-body junction diodes. These parameters with other parasitic components of an \( n \)-type MOSFET transistor are depicted in Figure 26. The gate-source junction capacitance (\( C_{gs} \)) may be the most dominant factor in non-linearity of a device, which seriously causes the phase distortion in the circuits. For example, a sensitive dependence of \( C_{gs} \) on the gate bias voltage introduces time-varying input reactance of a common source CMOS amplifier, originating the memory effect and amplitude-to-phase (AM-PM) distortion. In addition, non-vanishing third order term of the gate-source transconductance of which magnitude is also dependent on the gate bias voltage generates the third order inter-modulation distortion (IMD\(_3\)) and harmonic distortion, severely limiting the linearity performance of the CMOS amplifiers.

There have been efforts to analyze the effects of these dominant factors among other components [53]. However, most of researches are focused on the small signal analysis although critical linearity issues are raised from large signal operations. The RF
and MMW IC designers must adopt proper design techniques that compensates for non-linearity of active devices to achieve best performance of circuits and systems.

2.4.2. SiGe HBTs

2.4.2.1. Speed consideration

One of key parameters of active device for RF and MMW applications are the $f_T$ and the $f_{\text{max}}$. These parameters are strongly associated with gain/noise characteristics of the transistors. The equivalent circuit model of a general SiGe HBT is presented in Figure 27 [54]. From the transistor model, the $f_T$ can be derived as

\[
\frac{1}{2\pi f_T} = \tau_B + \tau_C + \frac{kT}{q}(C_{je} + C_{cb}) + (R_{ex} + R_c)C_{cb},
\]

(20)
where $\tau_B$ and $\tau_C$ are the base and collector transit time, respectively, $C_{je}$ is the emitter-base junction capacitance, $C_{cb}$ is the collector-base junction capacitance, and $R_{ex}$ and $R_c$ are the parasitic emitter and collector resistances, respectively. As seen in the (21), the $f_T$ is dominated by the base and collector transit time with other secondary effects originated from the parasitic components. The $f_{max}$ can be obtained from the $f_T$ with relation of

$$f_{max} = \sqrt[8]{\frac{f_T}{\pi R_b C_b}},$$

where $R_b$ and $C_b$ are the base resistance and the base capacitance, respectively.

As a rule of thumb in MMW IC designs, the $f_T$ and the $f_{max}$ are required to be greater than twice of the operating frequency of circuits. Therefore, the implementation of the transceiver circuit blocks operating beyond tens of GHz requires highly advanced process technology with very high $f_T$ and $f_{max}$. 

Figure 27. Equivalent SiGe HBT model.
2.4.2.2. Noise-gain trade-off

The noise performances of active devices are critical in LNAs and down-conversion mixers. Considering intrinsic parasitic parameters of active devices, minimum noise factor ($F_{min}$) of SiGe HBT can be simplified as [55]

$$F_{min} = 1 + \frac{1}{\beta} + \sqrt{2 g_m R_b} \sqrt{\frac{1}{\beta} + \left(\frac{f}{f_T}\right)^2}, \tag{22}$$

where $g_m R_b$ is assumed to be greater than 0.5. The associated gain, which is defined as the highest power gain achievable when the transistor source is noise matched for the minimum noise figure, is given by

$$G_a = \frac{1}{\omega^2 C_i} \left[ g_m R_b + \frac{1}{2} g_m^2 \frac{\beta}{\beta} + \frac{\left(\omega C_i\right)^2}{2 g_m R_b} \right], \tag{23}$$

where $C_i = C_{be} + C_{bc}$. As seen in the (21), $F_{min}$ is minimized with a high DC current gain ($\beta$), a high $f_T$, and a low $R_b$. Due to the advanced band-gap engineering in the SiGe technology, a high $\beta$, a high $f_T$, and a low $R_b$ can be simultaneously achieved for active devices, which makes SiGe devices as the optimal solution for low noise circuit designs. However, the associated gain, as expressed in the (22), is decreased as the $\beta$ becomes higher, implying that the trade-off between the low noise characteristic and the high gain performance exists even in a single transistor. The low gain of active devices for low noise performance limits the overall gain and the dynamic range of the receiver system.

2.4.2.3. Device breakdown

The other key issue of active devices is the breakdown of the transistor, which is a big obstacle in implementation of large signal circuits in the transmitter system such as
PAs. For the SiGe HBTs, the breakdown issue is determined by effects of high electric fields that lead the device to fail. The avalanche multiplication in the collector-base region is the fundamental source of the breakdown of bipolar devices, and collector-emitter breakdown is also problematic [54]. The low breakdown voltage (BV) of the transistors makes design of circuits that handle large voltage swings very intricate. Furthermore, the breakdown voltage has trade-off relation to the $f_T$ of the transistor as known as Johnson limit [56]. The Johnson limit describes that the product of the breakdown voltage and the $f_T$ is always material-related constant. Therefore, extending $BV \cdot f_T$ product for enhancing the performance of RF and MMW circuits has been key challenges for many years.

The HBT breakdown voltages are characterized by the collector-emitter breakdown voltage when the emitter is open-circuited ($BV_{CBO}$) or when the base is open-circuited ($BV_{CEO}$). The $BV_{CBO}$ is normally greater than $BV_{CEO}$ because of the current gain in the emitter-base region. The relation of the two breakdown voltages can be given as [54]

$$BV_{CEO} \approx \frac{BV_{CBO}}{\beta^{1/n}}$$  \hspace{1cm} (24)

where the $\beta$ normally varies between 2 and 5.

2.5. Conclusion

Various challenges and issues of passive and active components for RF and MMW IC designs using silicon-based process technologies were discussed. Monolithic inductors, capacitors, and transmission lines were reviewed with particular focus on EM characteristics that determines performances of the components. Limitations and issues of
MOSFETs and SiGe HBTs for RF and MMW circuit designs were also discussed. These characteristics of passive and active devices must be carefully considered for accurate and optimal designs of each circuit blocks in the transceiver system.
CHAPTER 3

DESIGN OF MULTI-MODE CMOS POWER AMPLIFIERS FOR
HIGH DATA-RATE WIRELESS COMMUNICATIONS

3.1. Introduction

With explosively growing demands on developments of true single-chip radio transceivers for mobile wireless communication devices at low cost, advances in research have led to successful implementations of CMOS PAs for various applications [57]-[61]. However, significant challenges in implementing CMOS PAs, such as low break-down voltages of transistors and lossy passive components on the highly conductive substrates, have necessitated the use of sophisticated design techniques to compensate for PA performance limitations.

Dealing with RF signals modulated with high complexity, PAs must simultaneously achieve high output power, high peak-power efficiency, and good linearity. In addition, advanced modulation schemes for high data-rate mobile wireless communications require PAs to operate at backed-off power levels to support high peak-to-average power ratio (PAPR). The PAPRs of common modulation schemes that are widely used in these days are summarized in Table 3. Because the PAs exhibits high power efficiency only near its peak output power level (P_{sat}), where transmitted signals are highly distorted, efficiency enhancement techniques at backed-off power levels (linear operation region) is strongly desired. Furthermore, enhanced efficiency at the lower output power levels extends the battery lifetime of mobile devices because the average efficiency of the PA is dominated by a probability distribution function in which
most of the transmission power is concentrated at much lower power levels than the peak output power [62].

<table>
<thead>
<tr>
<th>Table 3. PAPRs of various modulation schemes.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>QPSK</td>
</tr>
<tr>
<td>64-QAM</td>
</tr>
<tr>
<td>WCDMA</td>
</tr>
<tr>
<td>OFDM</td>
</tr>
</tbody>
</table>

In this chapter, design techniques to implement a fully-integrated watt-level linear CMOS PAs for advanced mobile communications are discussed. Section 3.2 introduces a novel power combining and discrete resizing technique to implement a multi-power mode operation with enhanced low-power efficiencies. Measured results verify the effectiveness of the proposed structure for WiMAX and WLAN multi-standards. Section 3.3 introduces a new figure-of-merit for multi-mode PAs that takes into account efficiency enhancement at power back-offs. The conclusion is followed in Section 3.4.

### 3.2. Linear Multi-Mode CMOS Power Amplifier

To design CMOS PAs with high output power and good efficiency/linearity, parallel-combining (or, parallel-amplification) techniques, as illustrated in Figure 28, can
be an attractive solution. By assembling multiple sub-PAs in parallel using output combining networks, the entire PA system can possibly transmit high output power while relaxing a linearity burden on individual sub-PAs. Moreover, the parallel combining structure can easily adopt a discrete power-control scheme to realize the necessary power back-offs by adequately resizing PA cells (i.e., disabling parts of parallel amplifier paths), thereby achieving enhanced low-power efficiency by reducing power consumption.

![Block diagram of a parallel-combining PA with discrete power-control.](image)

The transformer-based parallel combining technique is a practical way to realize fully-integrated parallel-combining multi-mode CMOS PAs. They are superior to transmission-line based parallel combining structures [63] in terms of integration ability and design simplicity. They also have wider operating bandwidth and less optimization issues than the LC-based network exploited in [64]. Two different ways of using the transformer-type power combiners associated with linear CMOS PAs have been
demonstrated thus far [65]-[68]. In one approach, the transformer is used as a current combiner [65]. The other approach uses the transformer to add up voltages at the output [66]-[68]. Integrated within PA systems, both structures were used to combine output power of individual PAs and support the control of the transmitted power levels in a discrete manner (Figure 29). However, by simply turning off each amplifier path to save power, these approaches neglected degradation effects in the transformer efficiency due to inactive primary inductors that create a leakage path back from the secondary inductor. Hence, there is room to further improve the efficiency characteristics in the lower-power modes of the PAs if the leakage is minimized.

In this section, we analyze the efficiency degradation effects of the multi-input-single-output transformers of which some of the inputs are disabled. Furthermore, a
design technique using varactor-based tunable matching networks to further improve the PA performance is introduced.

3.2.1. **Power combining transformer**

3.2.1.1. Transformers as power combiners

Power combining transformers can be mainly categorized as parallel-combining transformers (PCTs) that combine currents or series-combining transformers (SCTs) that aggregate voltages. Both combining transformers together with parallel PAs are illustrated in Figure 30. The number of turns of primary windings and secondary windings are denoted as $N_1$ and $N_2$, respectively. For the PCT, the secondary winding is shared by multiple primary windings and sums all AC currents induced by each primary winding. For the SCT, multiple secondary windings are connected in series, each of which is correspondingly coupled to individual primary windings. All AC voltages across

![Figure 30. Two types of power combining transformers. (a) PCT. (b) SCT.](image)
each secondary winding are summed at the top of the secondary node. With the multi-
input nature of these transformers, it is also possible to enable/disable some of the inputs
in a discrete manner, which is very attractive feature when they are integrated within a
PA system.

3.2.1.2. Efficiency of power combining transformers

The efficiencies of both PCTs and SCTs have been analyzed in [69] ignoring
coupling losses and assuming ideal mutual inductance. The transformer efficiency, which
is the ratio of the transferred power at the load to the total input power, for both
combining transformers can be expressed as

\[
\eta_{\text{transformer}} = \frac{P_{\text{Load}}}{P_{\text{In total}}} 
\]

(25)

\[
\eta_{\text{PCT}} = \frac{R_L}{(R_Ln^2)/M + R_2 + R_L} 
\]

(26)

\[
\eta_{\text{SCT}} = \frac{R_L}{MR_Ln^2 + MR_2 + R_L}, 
\]

(27)

where \( M \) is the number of combining paths, \( n \) is the turn ratio between the primary and
the secondary windings (i.e., \( n = N_2/N_1 \)), \( R_L \) and \( R_2 \) are respective series resistances of the
primary and the secondary windings, and \( R_L \) is the load impedance.

More elaboration to formulate the efficiencies of both transformers can be made
including the finite coupling effect between the primary and the secondary windings.
Replacing each single transformer with the T-equivalent circuit as shown in Figure 31(a)
and 31(b) [70], the PCT and the SCT can be also equivalently modeled as presented in
Figure 31(c) and 31(d), respectively (single-ended configurations are considered for
Figure 31. Equivalent transformer models. (a) Single 1:n transformer. (b) Equivalent model of a single 1:n transformer. (c) Equivalent model of an M-primary PCT. (d) Equivalent model of an M-primary SCT.
Due to the finite coupling coefficient, $k$ ($< 1$), between the primary winding ($L_1$) and the secondary winding ($L_2$), leakage inductance will exist on both sides. For the PCT, the secondary leakage inductance will be formed in a parallel manner, and the effective inductance seen by total induced current at the secondary side becomes $(1-k)L_2/M$. In contrast, the effective secondary leakage inductance of the SCT is $M(1-k)L_2$ because all the secondary windings are connected in series.

In Figure 4(c), assuming that the current at the secondary winding is equally attributed from the primary windings, relations describing the currents and the voltages at each node are

$$M(I_1 - I_{sl}) = nI_2$$  \hspace{1cm} (28)

$$nV'_1 = V'_2$$  \hspace{1cm} (29)

$$I_{sl} = \frac{X}{j\omega L_1 n}I_2$$  \hspace{1cm} (30)

$$I_1 = I_2 \left( \frac{n}{M} + \frac{X}{j\omega L_1 n} \right)$$  \hspace{1cm} (31)

$$V_i = I_2 \left( R_i + j\omega (1-k)L_1 \right) \left( \frac{n}{M} + \frac{X}{j\omega L_1 n} \right) + \frac{X}{n},$$  \hspace{1cm} (32)

where

$$X = j\omega \left( \frac{1-k}{M} \right) L_2 + R_s + R_L.$$  \hspace{1cm} (33)

Therefore, the efficiency of the PCT can be derived as

$$\eta_{PCT} = \frac{P_{\text{load}}}{P_{\text{in, total}}} = \frac{\text{Re}(I_2V'_2)}{M \times \text{Re}(I_2V'_2)} = \frac{R_i}{\frac{M \times \text{Re} \left( A \left( R_i + j\omega (1-k)L_1 A \right) + \frac{X}{n} \right)}{-}}$$  \hspace{1cm} (34)
where
\[ A = \frac{n}{M} + \frac{X}{j\omega L_M n} \]  \hspace{1cm} (35)

Assuming ideal coupling \((k=1)\) and ignoring mutual inductance \((I_M=0)\), the (34) becomes equal to the (26) as derived in [69].

For the SCT, a similar flow can be taken to formulate the efficiency. In Figure 31(d), the current and the voltage relations are derived as

\[ I_1 - I_M = nI_2 \]  \hspace{1cm} (36)

\[ MnV'_1 = V'_2 \]  \hspace{1cm} (37)

\[ I_M = \frac{Y}{j\omega L_M Mn} I_2 \]  \hspace{1cm} (38)

\[ I_1 = I_2 \left( n + \frac{Y}{j\omega L_M Mn} \right) \]  \hspace{1cm} (39)

\[ V_1 = I_2 \left( (R_1 + j\omega (1-k)L_1) \left( n + \frac{Y}{j\omega L_M Mn} \right) + \frac{Y}{Mn} \right) \]  \hspace{1cm} (40)

where

\[ Y = j\omega M (1-k)L_2 + MR_2 + R_L \]  \hspace{1cm} (41)

Hence, the efficiency of the SCT can be represented as

\[ \eta_{SCT} = \frac{R_L}{M \times \text{Re} \left\{ B \left( (R_1 + j\omega (1-k)L_1) B + \frac{Y}{Mn} \right) \right\}^*}, \]  \hspace{1cm} (42)

where

\[ B = n + \frac{Y}{j\omega L_M Mn} \]  \hspace{1cm} (43)
Note that the (42) also becomes equal to the (27) when ideal coupling \((k=1)\) is assumed and mutual inductance is ignored \((I_M=0)\).

3.2.2. Analysis on efficiency degradation of transformers

To determine the efficiencies of combining transformers described in the previous section, we only considered the case when all primary inputs participate in power combining. However, some of the primary inputs must be de-activated to realize discrete power back-offs. Although the overall power efficiency of the entire PA system appears to be increased at backed-off power levels at a first glance, substantial enhancement on the efficiency is difficult to achieve due to reverse-induced current from the secondary windings back to the idle primary windings that will inevitably dissipate power. This occurrence is illustrated in Figure 32. There is a de-activated PA among multiple PAs assembled in parallel using a combining transformer. Each active primary winding of the transformer induces AC currents at the secondary winding, and the secondary current induces another AC current at the idle primary windings with the same coupling coefficient. This reverse-induced current causes non-negligible power leakage through the idle primary sides of the transformer, degrading overall transformer efficiency. Depending on the input terminations of the idle primary windings (e.g., open-circuited, shorted to ground, and loaded by a capacitor), various frequency responses in the transformer efficiency characteristic may be obtained.

Using the approach taken in the previous section, the efficiency of the multi-primary-single-secondary transformers with some of the inputs disabled can be also derived. In Figure 33, the transformers have \(N\) idle primary inputs among \(M\) total primary
inputs. Therefore, the number of active primary inputs becomes \( M-N \). Each single transformer with an 1:n turn ratio can be replaced by the equivalent T-model as shown in Figure 31(b). Three different idle input terminations are considered: open-circuited, shorted to ground, and loaded by a capacitor. First, we define the impedance looking from the ideal transformer back to the idle primary input (excluding the mutual inductance) as \( Z_R \). Therefore, \( Z_R \) is infinity for the open-circuited case, \( R_1 + j\omega(1-k)L_1 \) for the case when the idle primary side is shorted to ground, and \( R_1 + j\omega(1-k)L_1 + 1/(j\omega C_R) \) for the case when the idle primary side is terminated by a series capacitor, \( C_R \). The reverse-induced current and voltage from the secondary current \( (I_2) \) are denoted as \( I_R \) and \( V_R \), respectively. Since the active primary windings serve as current sources to the secondary winding in the PCT (Figure 33(a)), we recognize that \( M-N \) of active primary currents \( (I_1-\)

Figure 32. Diagram of a parallel-combining PA using a PCT in case when one of PAs is disabled.
Figure 33. Equivalent circuit models of $M$-primary PCT and SCT with $N$-idle inputs. (a) PCT. (b) SCT. (c) Three different cases of idle primary input terminations.
\( I_M \) induce the secondary current, \((I_1-I_M)/n\), while some of the secondary current leak through the idle primary windings via reverse-induced current. Therefore, relations of the currents and the voltages at each node can be obtained as

\[
(M - N)(I_1 - I_M) = nI_2 + N(I_R + I_M)
\]

\[
I_M = \frac{X}{j\omega L_n} I_2
\]

\[
I_R = \frac{X}{nZ_R} I_2
\]

\[
I_1 = \frac{I_2}{M - N} \left( n + \frac{NX}{nZ_R} + \frac{MX}{j\omega L_n} \right)
\]

\[
V_1 = I_2 \left( \frac{R_i + j\omega(1-k)L_i}{M - N} + \frac{X}{n} \right)
\]

where \( X \) is the same as the (33) and

\[
C = n + \frac{NX}{nZ_R} + \frac{MX}{j\omega L_n}.
\]

Therefore, the efficiency of the \( M \)-primary PCT with \( N \)-idle primary windings becomes

\[
\eta_{PCT,N-idle} = \frac{Re(I_2 V_1^*)}{(M - N) \times Re(I_2 V_1^*)} = \frac{R_i}{Re \left( C \left( \frac{R_i + j\omega(1-k)L_i}{M - N} + \frac{X}{n} \right) \right)}.
\]

On the other hand, for the SCT in Figure 33(b), the active primary winding should be regarded as a voltage source. The AC voltages across each secondary winding that are induced by the corresponding active primary windings are aggregated at the very top of the secondary node. The secondary current, \((I_1-I_M)/n\), induced by the primary current, \( I_1-I_M \), flows through all the serially-connected secondary windings, subsequently inducing the AC current, \( nI_2 \), at the idle primary windings and generating induced voltage, \( V_R \). As
a result, the reverse-induced current, $I_R$, is produced. The currents and the voltages relations can be given as

$$I_I - I_M = n I_2$$  \hspace{1cm} (51)

$$V'_i = \frac{V'_2 - N n V_R}{(M - N)n}$$  \hspace{1cm} (52)

$$V_R = -n I_2 \frac{Z_R j \omega k L_I}{Z_R + j \omega k L_I}$$  \hspace{1cm} (53)

$$V_2 = \frac{R_2}{Y} V'_2$$  \hspace{1cm} (54)

$$I_I = V_2 \left( \frac{D}{j \omega k L_I n (M - N)} + \frac{n}{R_L} \right)$$  \hspace{1cm} (55)

$$V_i = \frac{V_2}{(M - N)n} D + I_I (R_I + j \omega (1 - k) L_I),$$  \hspace{1cm} (56)

where $Y$ is the same as the (41) and

$$D = \frac{Y}{R_L} + \frac{N^n Z_R j \omega k L_I}{R_L (Z_R + j \omega k L_I)}.$$  \hspace{1cm} (57)

As a result, the efficiency of the M-primary SCT with $N$-idle primary sides can be derived as

$$\eta_{SCT, M-n-idle} = \frac{\text{Re}(I, V'_i)}{(M - N) \times \text{Re}(I, V'_i)} = \frac{R_L}{\text{Re} \left( \frac{D}{n} + E \left( R_I + j \omega (1 - k) L_I \right) (M - N) \right)},$$  \hspace{1cm} (58)

where

$$E = \frac{R_I D}{j \omega k L_I n (M - N)} + n.$$  \hspace{1cm} (59)

As observed from the (50) and the (58), the efficiencies of the $M$-primary PCT and SCT...
is a function of not only the number of combining inputs \( M \), but also the number of idle primary sides \( N \) and the impedance looking toward the idle primary input \( Z_R \). Note that if \( Z_R \) is very close to zero at specific frequencies, which is the case for LC resonance when the idle primary side is loaded by a capacitor, the denominators of the (50) and the (58) are radically increased and the efficiencies accordingly approach zero.

The analyzed efficiencies of both \( M \)-primary PCT and SCT having \( N \)-idle primary inputs with three different idle primary terminations are plotted in Figure 34, 35, and 36. A turn ratio of \( n=2 \) is considered. The self-inductance values of the primary and the secondary windings are set as \( L_1=1 \) nH and \( L_2=n^2\cdot L_1 \) nH. The series resistance values, \( R_1 \) and \( R_2 \), are 0.8 Ω and \( n^2\cdot R_1 \) Ω, respectively. The quality-factor \( Q \) of spiral inductors and the coupling coefficient \( k \) are assumed to be 15 and 0.7 at 2.5 GHz, respectively. For both the PCT and the SCT, zero \( N \) represents that all primary inputs are enabled, and the efficiencies in this case are the maximum limit of the transformers. It is clearly observed that the overall efficiency is noticeably degraded as \( N \) increases. When the idle primary sides are perfectly open-circuited, the amount of efficiency degradation is minimal among all different cases. For the case when the idle primary sides are shorted to ground, the efficiency is reduced substantially across the entire frequency range. Furthermore, the resonant characteristic in the efficiency versus frequency plot is clearly seen in the case when the idle primary inputs are terminated by a capacitor (e.g., 4 pF in this analysis) for both transformers.

The transformer efficiency based on the calculation of maximum available gain [71] can be obtained in simulations. Example implementations of a two-primary PCT (Figure 37(a)) and an SCT (Figure 37(b)) with a turn ratio of 1:2 were electro-
Figure 34. Analyzed efficiencies of 4-primary 1:2 transformers. (a) PCT with $N$-idle primary inputs that are open-circuited. (b) SCT with $N$-idle primary inputs that are open-circuited.
Figure 35. Analyzed efficiencies of 4-primary 1:2 transformers. (a) PCT with \( N \)-idle primary inputs that are shorted to ground. (b) SCT with \( N \)-idle primary inputs that are shorted to ground.
Figure 36. Analyzed efficiencies of 4-primary 1:2 transformers. (a) PCT with \( N \)-idle primary inputs that are loaded by a capacitor. (b) SCT with \( N \)-idle primary inputs that are loaded by a capacitor.
magnetically modeled using Ansoft-HFSS and embedded into the simulations. As presented in Figure 38(b) and 38(c), maximum efficiencies across the entire frequency range were obtained only when all the primary inputs are transferring power, which coincides with derived results in the analysis above. When one of the primary inputs was disconnected from the power source with an open termination, about 10% efficiency drop

Figure 37. 3-D structures of two-primary 1:2 transformers. (a) PCT. (b) SCT.
Figure 38. Simulated maximum available efficiencies of two-primary 1:2 transformers with EM-simulated models. (a) PCT. (b) SCT.
was observed. There was more significant efficiency degradation when one of the input ports was shorted to ground. Moreover, severe efficiency reduction occurred at resonances when one of the primary input ports had an additional capacitive load (2-pF and 4.5-pF capacitors were used in the simulation).

3.2.3. Design of a linear CMOS PA

3.2.3.1. Concurrent resizing and combining technique

As examined in above sections, all the primary inputs of the combining transformer must be active in order to utilize the transformer efficiency maximally. In a parallel-combining PA structure, instead of turning off each power-cell one at a time as shown in the structure in Figure 39(a), we can split each power-cell into multiple sub-cells and discretely enable/disable some of them in a concurrent manner. The idea is illustrated in Figure 39(b). There are several PA stages in parallel to be combined by the transformer at the output. Each PA stage consists of multiple sub-cells (i.e., PA_1 to PA_K) that are selectively combined in a concurrent way, meaning that only PA_1s in each PA path are combined when the lowest output power is needed, and all the sub-cells in each PA path are enabled and combined when the maximum output power is required. The number of enabled sub-power-cells may vary gradually depending on changes of required output power levels. Hence, all the primary inputs of the transformer can participate in power combining in every operation modes, preventing power losses due to idle primary inputs of the transformer. Although the idea is exemplified using a PCT in Figure 39, the SCT can be also used for the implementation.
Additionally, the concurrent combining scheme has another advantage over the conventional technique particularly when the PCT is used. The combining structure in
Figure 39(a) has a tendency to decrease input load impedance at the output node of each PA when some PAs are turned off [65], thereby resulting in a reduction in the voltage gain and power efficiency. The optimal load impedance of a smaller power cell should be greater than that of a larger power cell because the load-line changes in accordance with changes in the current level that the power cell can sustain [52]. Although the variation of the load impedance can be compensated with additional matching circuitry, excessive variation of the load impedance must be avoided for ease of design. In Fig. 39, the load impedance, which is a ratio of the output voltage ($V_1$) to the output current ($I_1$), becomes

$$Z_{\text{Load}} = \frac{V_1}{I_1} = \frac{1}{n} \frac{V_2}{I_2} = \frac{(M-N)V_2}{n^2 I_2} = \frac{M-N}{n^2} R_L,$$

(62)

where $M$ is the total number of parallel PA paths and $N$ is the number of disabled PA paths. Thus, $Z_{\text{Load}}$ in the conventional structure is decreased as $N$ is increased to realize power back-offs, whereas it remains constant in the concurrent combining structure as long as equally-sized sub-power-cells in each path are enabled and combined (i.e., $N$ is always zero in this structure). Without severely disturbing the optimal load impedance, improved efficiency characteristics in the lower-power modes can be achieved as compared to the conventional combining scheme using the PCT. However, additional tuning of the load impedance in such a way that each PA has increased load impedance when its size is reduced may be necessary to further improve the PA efficiency. This idea will be discussed later in this section.

3.2.3.2. Design and implementation
A fully-integrated multi-mode PA has been fabricated in IBM 1P6M 0.18-μm CMOS technology. The diagram of designed two-stage PA is shown in Figure 40(a). The PA consists of an input balun with matching circuits, a driver stage, an inter-stage matching network, two parallel power stages, and a two-primary 1:2 PCT \((M=2\) and \(N=0\)) as a power combiner. The driver and each of power stages are divided into three sub-cells \((K=3)\) for discrete resizing and combining. The PCT was chosen for this implementation because of its superiority in size as compared to the SCT. Due to the serially-connected secondary windings, the outer dimension of the SCT is approximately twice of that of the PCT as illustrated in Figure 37. Generalized schematic of each driver/power stage is shown in Figure 40(b). Although \(K=3\) was chosen for this design, higher number of \(K\) (number of parallel driver/power stages) is also realizable to enable more continuous selections of the power mode. Operation conditions for designed PA are graphically presented in Figure 40(c). As required output power is increased, more parallel driver/power stages can be enabled with different bias conditions. At the maximum output power level that the PA can achieve, higher gate bias voltage is applied to operate the PA in clean class-A condition. For smaller power generation, relatively lower gate bias voltages are applied to enhance the overall efficiencies.

A simplified schematic of the entire PA is shown in Figure 41. Individual driver/power stages have pseudo-differential configuration to reduce the source degeneration effect by bonding wires. In addition, the differential topology minimizes self-mixing of even-order harmonics with the fundamental signal, thus reducing IMD3. Each sub-cell has a cascode configuration with thick gate-oxide transistors for the common gate (CG) devices, which not only relaxes voltage stress on the individual
Figure 40. Implementation of a concurrent combining PA. (a) Diagram of a PA with an integrated two-primary 1:2 PCT. (b) Generalized schematic of each driver and power stage. (c) Operation conditions of the PA.
transistors, but also eases the implementation for discrete power control. That is, by applying adequately low bias voltage to the gate of the CG devices, the signal paths through the devices are switched off and do not participate in the power combining. The common source (CS) devices in each path are always biased to operate in the saturation mode, thus overall input capacitance of driver/power stages remains constant. In contrast, there is finite change in the total output capacitance value depending on the number of enabled sub-cells because the CG devices in each stage are transitioned from the cut-off region (turned off) to the saturation region (turned on). This variation in the output capacitance should also be considered for the optimal matching, and will be effectively absorbed into the variable output matching circuitry.

Figure 41. Simplified schematic of designed PA.
Both the driver and the power stages are biased at class-AB compromising efficiency and linearity. The driver stage is biased a little close to class-B so that it has the gain expansion to compensate for the compression of the power stage gain. An RC feedback is also employed for the individual cells to achieve better stability as well as a flatter gain response, or amplitude-to-amplitude (AM-AM) response. Three different operation modes are available with respect to required output power levels although more diverse combinations of individual sub-cells are also implementable. For maximum output power generation, all sub-cells in the driver stage (DA_{1} + DA_{2} + DA_{3}) and the power stages (PA_{1} + PA_{2} + PA_{3}) are enabled, and the PA operates in the high-power (HP) mode. By disabling the largest cells in each stage (i.e., enabling DA_{1} + DA_{2} and PA_{1} + PA_{2}) the medium-power (MP) mode operation is realized with adequate power back-off. For the low-power (LP) mode operation, only the smallest sub-cells in each stage (DA_{1} and PA_{1}) are operational for the power generation. Device sizes and configurations of each transistor in both stages are summarized in Table 4.

Design of the combining transformer is also critical because its power loss may dominate the overall power efficiency of the PA system even though the PA itself is well optimized. The outer dimension (size) of the transformer and widths of metal traces determine the self-inductance values and the quality factors of the primary and the secondary winding inductors. The spacing between two adjacent metal traces affects the effective coupling between the primary and the secondary windings as well as SRF of the transformer. Therefore, careful considerations should be taken to choose appropriate design parameters for the transformer while iteratively verifying with EM simulations. A turn ratio, \( n=2 \), was chosen because around 15-\( \Omega \) load impedance is required for
individual power stages to generate maximum power (this value was obtained by load-pull simulations using ADS). The 30-μm width and 10-μm spacing for the windings were chosen in this design. The PCT with a size of 780 × 800 μm² has the primary and the secondary self-inductance of 1.5 nH and 4.2 nH at 2.5 GHz, respectively. The quality factor of each winding is greater than 12 (peak of 16) below 3 GHz, and the effective coupling coefficient is 0.72. The maximum available gain of designed PCT is -0.63 dB at 2.5 GHz. The transformer was implemented using a stack of two thick-metals connected through whole via trace as presented in Figure 40(a), which helps to achieve relatively high quality factor and coupling coefficient.

The characteristics of the PCT were obtained from EM simulations using Ansoft-HFSS. Figure 42 shows the simulation setup for the PCT. There are perfect-conductor leads for port excitations since the distance between the ground wall of the boundary box
and excitation ports is quite significant. This lead inductance must be obtained from independent simulations and effectively de-embedded from the overall inductance of the transformer. The separate EM simulation was performed for the conductor leads: Lead1 and Lead2. Then, the inductance of individual lead \( L_{\text{Lead}} \) is calculated from the setup presented in Figure 43 using equations

\[
L_{\text{Lead1}} = \frac{\text{Im}(Z_{\text{in}})}{2\pi f} \tag{61}
\]

\[
L_{\text{Lead2}} = \frac{1}{2} \frac{\text{Im}(Z_{\text{in}})}{2\pi f} \tag{62}
\]
Negative values of these lead inductances must be added to the overall transformer inductance to obtain actual self inductance. Simulation results of the PCT are presented in Figure 44.

![Diagram](image)

**Figure 43.** Lead inductance de-embedding setup. (a) For lead1. (b) For lead2.

![Graph](image)

**Figure 44.** EM simulation results of two-primary PCT. (a) Maximum available gain. (b) \( Q \).
3.2.3.3. Measurement results

A micro-photograph of designed PA is presented in Figure 45. A printed circuit board (PCB) was used for measurements, and the loss of signal lines in the PCB was de-embedded. Due to bond-wire effects in the matching networks, optimal operating frequency was a little shifted down from the desired frequency. One-tone power sweep with continuous-wave signal at the RF frequency of 2.1 GHz was performed, and the measurement results are presented in Figure 46, 47, and 48. V_{DD} was 3.3 V and, the gate bias voltages for the low power (LP), the medium power (MP), and the high power (HP) modes were set at 0.5 V, 0.55 V and 0.6 V, respectively. Measured P_{sat}/P_{1dB} are 30.7/27.5 dBm with PAE of 35.8/20% for the HP mode, 25/21.8 dBm with PAE of 19.8/13% for the MP mode (6-dB back-off) and 19/16.5 dBm with PAE of 10.5/9.8% for the LP mode (12-dB back-off). Although we observed that 7% of efficiency enhancement was achieved at 17-dBm P_{out}, this is not as significant enhancement as anticipated in the

Figure 45. A Chip micro-photo of the concurrent combining multi-mode PA.
Figure 46. Measurement results of multi-mode linear CMOS PA: gain and PAE (at 2.1 GHz).

Figure 47. Measurement results of multi-mode linear CMOS PA: EVM (at 2.1 GHz).
Figure 48. Measurement results of multi-mode linear CMOS PA: S-parameters. (a) $S_{11}$ and $S_{22}$. (b) $S_{21}$ and $S_{12}$. 
simulation. We believe that the optimal performance was limited at some point by the fixed matching network. We can improve the performance by adapting tunable capacitors in the output matching. For linearity test, 802.11g WLAN modulated signal was used to evaluate EVM. With 54 Mbps 64 QAM OFDM signal, -25-dB EVM specification was satisfied at output powers of 21 dBm (HP), 14.8 dBm (MP), and 10 dBm (LP), consuming DC quiescent current of 560 mA (HP), 200 mA (MP), and 100 mA (LP). We could verify that the linearity of designed PA is comparable to the conventional PA with enhanced low-power efficiency.

3.2.4. Load impedance modulation

To further optimize the PA performance in the lower-power modes, the load impedance seen by individual power stages must be tuned by variable matching circuitry. As the PA changes its operation mode (e.g., from HP mode to LP mode) the effective size of the power cell changes accordingly, requiring different values for the load impedance (Figure 49). Although the transformer has a fixed impedance transformation

Figure 49. A varactor-based tunable capacitive matching circuit for load impedance modulation.
ratio due to its physical turn ratio, the transformed impedance value can still be adjusted (or, modulated) by adding a variable shunt capacitor to the output node of the power stages. Moreover, as fore-mentioned, the variation in the effective drain capacitance in accordance with the change in the cell size should also be taken into account. Absorbing this variation into the output capacitance, required capacitance values for the HP mode and the LP mode power stages are found to be 2.8 pF and 1.2 pF, respectively including device parasitic capacitances. The required capacitance for the MP mode power stage is not greatly different from the one for the LP mode, thus 1.2-pF capacitance was chosen for the necessary value covering the LP and MP modes.

This tunable capacitor was implemented using accumulation-mode (A-mode) MOS varactors. As presented in Figure 49, two varactors are connected symmetrically and embraced by fixed MIM capacitors on top and bottom. The A-mode varactor is an NMOS transistor on N-well of which the drain (D) and the source (S) are tied together. As the varactor changes its state from the depletion to the accumulation by adequate DC biasing which is applied through large resistors, different capacitance values can be achieved. When the biasing voltage difference between the gate and the source/drain is negative (e.g., $V_G=-1.5$ V and $V_D/S=0$ V), the varactor is in depletion mode having smaller effective capacitance (the oxide capacitance in series with the depletion capacitance). On the other hand, the varactor exhibits larger capacitance (the oxide capacitance only) when the biasing voltage difference between the gate and the source/drain is positive (e.g., $V_G=1.5$ V and $V_D/S=0$ V) as the varactor is in the accumulation mode.

Because the reliability is the critical concern when using the varactors, a stacked
structure was exploited to relax the voltage stress on the gate-source/drain oxides by dividing applied voltage across the entire tunable matching circuit. The amplitude of the voltage swing across the gate-source/drain should be maintained within a few volts not to cause gate-oxide break-down. This was verified with simulated waveforms. As presented in Figure 50, when the PA is transmitting at almost saturated output power (with RF input power of -3 dBm), the voltage swing at the drain of each power stage remains under 11 V even in various load mismatch conditions (9 V in the optimal load condition). In this case, a maximum 4.2-Vpp (3.7-Vpp in the optimal load condition) swing is applied across the gate-source/drain of individual varactors ($V_1-V_2$ and $V_3-V_2$ in Figure 49). On the other hand, less than 1-Vpp swing is applied across the varactor when the PA is operating in the LP mode. Therefore, with the stacking approach, the varactors operate under safe conditions.

It should be also noted that a large voltage swing across the varactors when the
PA is transmitting near saturated output power levels may modulate the instantaneous capacitance value during the signal period. However, the average capacitance value is still determined by applied gate voltage of the varactors, and the linearity performance of the PA in terms of the EVM and harmonics are negligibly affected by the varactors because the PA operates at backed-off power levels in most of the times. The varactor-based tunable matching circuit was used to implement $C_{OUT1}$ and $C_I$ in Figure 41.

The tunable matching circuit has $Q$ of greater than 35 in the HP mode and 25 in the LP mode (Figure 51). Compared to much lower $Q$ of the on-chip transformer, the effect of the varactor-based matching circuit on the overall $Q$ of the output matching network is negligible. In the simulation, only 0.5-dB difference in the gain performance, and 1-% difference in the efficiency performance were seen (Figure 52). The entire schematic of the linear multi-mode CMOS PA with load impedance modulation technique is presented in Figure 53.

![Figure 51. Simulation results of the varactor-based matching circuit. (a) Q. (b) Capacitance.](image)
Figure 52. Simulation results of the PA with the varactor-based matching circuit. (a) Gain. (b) PAE.

Figure 53. Entire schematic of designed linear multi-mode CMOS PA with load impedance modulation technique.
3.2.5. Measurement results of a multi-mode CMOS PA with load impedance modulation

The PA chip was mounted on the FR-4 printed circuit board (PCB) for measurements. The micro-photograph of the PA is presented in Figure 54. The PCB line loss was de-embedded, but the loss due to bonding-wires is included in the experimental results. A single-tone continuous-wave signal with the frequency of 2.5 GHz was applied to the PA to measure the gain, $P_{1\text{dB}}$, $P_{\text{sat}}$, and PAE in three different operation modes. The measured results are plotted in Figure 55. Power back-offs are realized by concurrently resizing the driver and the power-cells. When all sub-cells in the driver and the power stages were enabled for the HP mode operation, gain of 31.3 dB, $P_{\text{sat}}$ of 31 dBm, $P_{1\text{dB}}$ of 28 dBm, and peak PAE of 34.8% were achieved. For the MP and the LP mode operations, two different cases were tested. First, only discrete resizing and concurrent power combining were exploited to implement power back-offs. Next, the inter-stage and the

![Figure 54. Micro-photograph of designed multi-mode linear CMOS PA with load impedance modulation technique (1.1 × 1.8 mm²).](image)
output matching circuits were additionally tuned by adjusting the varactor biasing voltage (bias voltages were applied externally). By only disabling the largest cells in the driver and the power stages, a power back-off of about 5 dB was realized (MP mode), achieving 26.5 dB, 26 dBm, 22.5 dBm, and 20.2% for the gain, $P_{\text{sat}}$, $P_{1\text{dB}}$, and peak PAE, respectively. With additional tuning of matching networks, gain expansion was obtained with boosted peak PAE. The small-signal gain did not have noticeable change, but the $P_{1\text{dB}}$ and the peak PAE were increased to 23.8 dBm and to 22.5%, respectively. For the LP mode operation, only the smallest cells in the driver and the power stages were enabled. The $P_{\text{sat}}$ was backed-off by about 3 dB more from the one for the MP mode. An achieved gain, $P_{\text{sat}}$, $P_{1\text{dB}}$, and peak PAE with additional tuning of the matching networks were 22.4 dB, 22.3 dBm, 20.5 dBm, and 15%, respectively. The peak PAE in the LP

Figure 55. Measurement results of multi-mode PA with load impedance modulation in three operation modes: gain and PAE (at 2.5 GHz).
mode without tuning of the matching networks was only 12%. As clearly indicated in Figure 55, by employing the proposed multi-mode scheme, 9-% point increase in the PAE at the output power of 20 dBm (from 5% to 14% at 11-dB back-off from P_{sat}) and 5-% point increase at the output power of 25 dBm (from 14% to 19% at 6-dB back-off from P_{sat}) were achieved as compared to the case when the PA is operating solely in the HP mode. The measured DC currents and S-parameters of designed PA in three operation modes is plotted in Figure 56 and 57. In the HP, MP, and LP modes, designed PA consumes idle DC current of 590 mA, 225 mA, and 105 mA, respectively at a 3.3 V power supply. The gate biasing voltages for the HP, MP, and LP modes were 0.58 V, 0.54 V, and 0.5 V, respectively for the power stages and 0.56 V, 0.52 V, and 0.5 V, respectively for the driver stage. Relatively high gate biasing voltages were applied in the

![Figure 56](image-url)

*Figure 56. Measurement results of multi-mode linear PA with load impedance modulation in three operation modes: DC current (at 2.5 GHz).*
Figure 57. Measurement results of multi-mode linear PA with load impedance modulation: S-parameters. (a) $S_{11}$ and $S_{22}$. (b) $S_{21}$ and $S_{12}$. 
HP mode operation concerning linearity as the output power level approaches $P_{1\text{dB}}$. The gate of CG devices of the driver stage and the power stages were biased at 3.3 V and 3 V, respectively for all measurements. The PA achieves small signal gain of 31-dB. $S_{11}$ and $S_{22}$ are less than -12 dB and -5 dB at 2.5 GHz, respectively.

The EVM was measured to determine the linearity of designed PA with the IEEE 802.11g WLAN OFDM 64-QAM and 802.16e WiMAX OFDM 64-QAM modulated signals at 2.5 GHz. Measured EVM for all three operation modes with two different modulated signals is presented in Figure 58(a) and (b). With WLAN 54-Mbps 64-QAM modulated signal applied, the PA meets the linearity requirement (EVM of -25 dB or 5.6%) at the output power of 21 dBm, 18 dBm, and 14.5 dBm for the HP, MP, and LP modes, respectively. For WiMAX 54-Mbps 64-QAM modulated signal (10 MHz channel band-width), the measured EVM exceeds -31-dB (2.8%) limit at the output power of 17 dBm (HP), 14.5 dBm (MP), and 12 dBm (LP). Note that the PAE of the LP mode PA at the linear output power for the WLAN 64-QAM modulated signal ($P_{\text{Linear}}$ of 15 dBm) is about 8% with the proposed multi-mode operation scheme, while the PAE in the HP mode at the same output power is less than 4%. Therefore, the designed PA shows potential operability for WLAN/WiMAX multi-standards multi-mode applications that can support multiple modulation schemes with varying output power levels. The incorporation of proposed combining structure with discrete resizing technique using the 2-primary PCT effectively achieved optimization of the PA performance at lower power levels. Although demonstrated PA employed the PCT, other combining transformers such as SCT also can be utilized in the proposed structure. Measured constellation and output spectrum of the HP mode PA with two modulated signals are presented in Figure 59.
Figure 58. Measurement results of multi-mode linear PA with load impedance modulation in three operation modes: EVM (at 2.5 GHz). (a) For WLAN OFDM 64-QAM signal. (b) For WiMAX OFDM 64-QAM signal.
3.3. Efficiency Figure-of-Merit

An indication number as a measure of efficiency enhancement at specific power back-offs can be used. Assuming that the optimal load impedance condition is preserved, we can theoretically calculate an efficiency back-off ratio (EBR) by a ratio of the efficiency at the backed-off power level to the peak efficiency as [52], [72].

Figure 59. Measurement results of multi-mode linear PA with load impedance modulation in the high power mode: constellation and spectrum (at 2.5 GHz) (a) For WLAN OFDM 64-QAM modulated signal (P_{out}=20.5 \text{ dBm}). (b) For WiMAX OFDM 64-QAM modulated signal (P_{out}=16 \text{ dBm}).
\[
EBR = 2\sqrt{10^{BO/10}} \left( \frac{1}{2} \cdot \frac{\alpha_{BO} - \sin(\alpha_{BO})}{\sin(\alpha_{BO}/2) - \alpha_{BO} \cos(\alpha_{BO}/2)} \right) = \frac{1}{\sqrt{10^{BO/10}}},
\]

where \(BO\) is the amount of power back-off in dB, \(\alpha_{BO}\) and \(\alpha_{peak}\) are the current conduction angles at the backed-off power and the peak output power, respectively, and the bias condition is assumed to be unchanged as the output power is reduced. This EBR represents how much low-power efficiency can be obtained with respect to the peak efficiency in a single mode PA when the amount of power back-off is specified. We can also define and obtain an achieved EBR (\(EBR_{achieved}\)) at the same power back-offs as

\[
EBR_{achieved} = \eta_{BO} / \eta_{peak} \times \eta_{BO,EET} / \eta_{peak} = \eta_{BO,EET} / \eta_{peak},
\]

where \(\eta_{peak}\) is achieved (or, measured) peak efficiency, \(\eta_{BO}\) is achieved back-off efficiency without employing efficiency enhancement techniques, and \(\eta_{BO,EET}\) is achieved back-off efficiency with efficiency enhancement techniques employed. Thus, we can define the efficiency enhancement ratio at power back-offs (\(EER_{BO}\)) as

\[
EER_{BO} = \frac{EBR_{achieved}}{EBR} = \sqrt{10^{BO/10}} \left( \frac{\eta_{BO,EET}}{\eta_{peak}} \right).
\]

By using the (65) rather than using a ratio of \(\eta_{BO,EET}/\eta_{BO}\) as \(EER_{BO}\), it is possible to reflect appropriate weighting for the amount of power back-offs and avoid exaggeration of the amount of efficiency enhancement in a case when \(\eta_{BO}\) is excessively lower than \(\eta_{BO,EET}\).

3.4. Conclusion
The efficiency degradation effects of the power combining transformers due to idle primary inputs (in cases of open-circuited, shorted to ground, and terminated by a capacitor) are fully analyzed. To avoid efficiency degradation in lower-power modes, a concurrent resizing and combining technique has been exploited in combination with tunable matching networks to implement a fully-integrated multi-mode class-AB CMOS PA. Load impedance modulation technique using the tunable matching networks improves the performance of designed PA when operating modes of the PA changes. The measured results validate that the proposed technique can be effectively employed as a solution for multi-mode multi-standard linear CMOS PA applications. The performance comparison of state-of-the-art watt-level linear CMOS PAs is shown in Table 5.
Table 5. Performance comparison of watt-level linear CMOS PAs.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>[65]</th>
<th>[68]</th>
<th>[73]</th>
<th>[74]</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Freq [GHz]</strong></td>
<td>2.4</td>
<td>2.4</td>
<td>2.45</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td><strong>Gain [dB]</strong></td>
<td>37</td>
<td>28</td>
<td>32</td>
<td>18</td>
<td>31.3</td>
</tr>
<tr>
<td>$P_{1dB}/P_{sat}$ [dBm/dBm]</td>
<td>27/31</td>
<td>27.7/30.1</td>
<td>27.5/31.5</td>
<td>29.5/32</td>
<td>28/31</td>
</tr>
<tr>
<td>$P_{out@-25\text{-dB EVM}}$ [dBm]</td>
<td>22</td>
<td>22.7</td>
<td>25.5</td>
<td>25</td>
<td>21</td>
</tr>
<tr>
<td>Peak PAE [%]</td>
<td>27</td>
<td>33</td>
<td>25</td>
<td>48</td>
<td>34.8</td>
</tr>
<tr>
<td>Linearization Technique</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Multi Power Mode w/ EET</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Power Back-off [dB]</td>
<td>6</td>
<td>11</td>
<td>12</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td><strong>PAE@Back-off [%]</strong></td>
<td><strong>w/o EET</strong></td>
<td>12</td>
<td>4</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td><strong>w/ EET</strong></td>
<td>15</td>
<td>6</td>
<td>10.4</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>EER$_{90}^\circ$</td>
<td>1.11</td>
<td>0.79</td>
<td>1.25</td>
<td>0.63</td>
<td>0.96</td>
</tr>
<tr>
<td>Supply [V]</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>Size [mm$^2$]</td>
<td>2</td>
<td>4.32</td>
<td>2.7</td>
<td>2.25</td>
<td>1.98</td>
</tr>
<tr>
<td>Technology</td>
<td>180 nm CMOS</td>
<td>90 nm CMOS</td>
<td>65 nm CMOS</td>
<td>90 nm CMOS</td>
<td>180 nm CMOS</td>
</tr>
</tbody>
</table>
CHAPTER 4

DESIGN OF HIGH-POWER LINEAR CMOS PA WITH NOVEL
POWER COMBINING TRANSFORMER

4.1. Introduction

CMOS PAs have been gaining increasing interests from the wireless communication IC industry due to their low development cost and high integration levels as mentioned in Chapter 1. From early generations of cellular applications (e.g., GSM/GPRS) to high data-rate applications with moderate transmission power levels (e.g., IEEE 802.11 a/g/n WLAN), the territory of applications for CMOS PAs has been rapidly expanded with help of novel design techniques overcoming diverse issues and challenges of silicon CMOS technologies [61], [65]-[69]. Recently, CMOS PAs have even demonstrated their suitability for more advanced wireless communication standards such as WiMAX with good achievements in output power, efficiency, and linearity performances [73]-[75]. However, it is still challenging for CMOS PAs to compete with their counterparts implemented in III-V compound semiconductor technologies (e.g., GaAs/InP/InGaP), thus more sophisticated design techniques are required to further improve performances of CMOS PAs for adaptation to next generation wireless communications.

Although linearity is one of the important performance metric of PAs, generation of high output power with good power efficiency is a more fundamental requirement for mobile applications. Due to high PAPR of OFDM-based modulation schemes that are commonly used for high data-rate communications, average output power levels for PAs
must be backed off from the peak output power level by a significant amount. With considerations of non-ideality of CMOS PAs in terms of amplitude and phase distortion, the required amount of power back-off becomes at least 8 to 9 dB for tolerable data transmission [68]. Therefore, a flat gain-response for a PA up to 30 to 31 dBm (i.e., output referred $P_{1\text{dB}}$ of greater than 30 dBm) is required to meet a linearity specification at the average output power levels of around 22 dBm in a long-range mobile communication environment such as WiMAX. If a margin for additional loss due to a front-end switch with an antenna is taken into account, the required output power level is further increased (i.e., $P_{1\text{dB}}$ of greater than 31 dBm).

To implement such high-power PAs in a CMOS technology with good power efficiency, transistor stacking and/or output power combining techniques have been frequently employed to overcome design issues caused by low breakdown voltage of

![Diagram of output power combining technique for high-power PA applications.](image)

**Figure 60.** Output power combining technique for high-power PA applications.
MOSFET devices (Figure 60). By parallel combining of multiple PAs using an output power combining network, the requirement on the voltage swing for individual PAs can be alleviated, which results in improvements in overall linearity and efficiency. Moreover, the excessively small load impedance value that is needed to be seen by a single large PA cell can be avoided by assembling smaller sub-PA cells in parallel, so the amount of loss at the output matching networks (or the power combiner) is reduced because of lower impedance transformation ratio. The key issue is then how we can implement the low-loss monolithic power combiner in a small form-factor that can be easily integrated with the PAs in CMOS technologies.

In recent years, transformer-based power combiners have attained particular attention for implementations of single-chip high-power CMOS PAs generating watt-level output power [65]-[69]. Two different types of monolithic power combining transformers: parallel-combining transformers (PCTs) and series-combining transformers (SCTs) have been successfully used thus far (These transformers are also explained in Chapter 3). Although reported high-power CMOS PAs using these transformers achieved good performance, there are still design issues and limitations in terms of feasibility of implementation and intrinsic characteristics of the transformers (e.g., efficiency and impedance transformation ratio) especially when they are incorporated to combine more than three parallel PAs to generate higher output power.

In this Chapter, we investigate characteristics of the PCT and the SCT focusing on the efficiency and transformed impedance levels at each input terminal with varying numbers of combining inputs. To overcome limitations of the PCT and the SCT, we propose a novel monolithic power combining transformer that is appropriate for
generation of output power level of greater than 33 dBm (2 W). The proposed transformer performs parallel and series combining simultaneously in a single structure to combine four sub-PAs, exhibiting an insertion loss of 0.73 dB at 2.4 GHz. Using the proposed transformer, a two-stage class-AB CMOS high-power PA for WiMAX applications is designed and demonstrated. This Chapter is organized as follows. In Section 4.2, conventional power combining transformers are characterized and compared from the perspective of high-power PA designs. Then, a new combining transformer, parallel-series combining transformer (PSCT), is introduced in Section 4.3 along with analysis and comparison to other types of combining transformers. Section 4.4 presents design example of PSCT-based CMOS PA in detail, while the measurements results and conclusions are presented in Section 4.5 and 4.6, respectively.

4.2. Conventional Power Combining Transformers: PCTs and SCTs

Schematic diagrams of conventional power combining transformers: the PCT and the SCT are illustrated in Figure 61. The transformers are composed of sub-transformers with a primary winding \( L_1 \) and a secondary winding \( L_2 \) that are coupled to each other. These sub-transformers are connected in parallel (PCT) or in series (SCT) to form entire combining transformers. The schematic diagrams in Figure 61 include series parasitic resistances of each primary winding \( R_1 \) and the secondary winding \( R_2 \). Since the SCT has secondary windings connected in series according to the number of combining inputs \( M \), the total series resistance at the secondary side becomes \( MR_2 \). In contrast, the parasitic resistance at the secondary side of the PCT comes from a single secondary winding that is shared by all primary windings. \( I_1 \) and \( I_2 \) respectively represent the
primary and the secondary AC currents. \( V_1 \) is the AC voltage applied at each input terminal of the primary winding, and \( V_1' \) is the actual voltage across the primary winding, which is smaller than \( V_1 \) due to the voltage drop through the series parasitic resistance, \( R_1 \). On the secondary side, \( V_2' \) represents the induced AC voltage across the actual secondary winding, and \( V_2 \) represents the voltage at the load, \( R_L \). These transformers can

Figure 61. Conventional power combining transformers. (a) PCT. (b) SCT.
be used as a power combiner to implement monolithic high-power PAs. From exited input power at each input terminal at the primary side of the transformers, the output power is generated by combining the secondary AC currents (in the PCT) or the AC voltages (in the SCT). Multiple PAs can be assembled in parallel such that their outputs are combined by the transformers to effectively increase the output power level delivered to the load. These transformers also serve as an impedance transforming network with a physical turn ratio \( n \) between the primary and the secondary winding to step down the load impedance (50 Ω) to required impedance level seen by individual PAs at the input terminals.

To characterize the power combining transformers, the efficiency and the input impedance of the transformers must be quantified. The transformer efficiency represents a ratio of the amount of delivered power at the load to the total amount of input power excited to the transformer as described in the (26) in Chapter 3. The input impedance is defined as the ratio of the input voltage to the input current at each input terminal (i.e., \( R_{in}=V_1/I_1 \)). While the transformer efficiency simply indicates how much power is dissipated or lost through the transformer out of transferred power to the load, the effect of the input impedance of the transformer must be considered in conjunction with sub-PA cells under specific matching conditions. Based on required output power level (accordingly, the size of the sub-PA cells and current capability), optimal load impedance value \( R_{opt} \) for individual sub-PA cells is determined. If the input impedance of the combining transformer is largely varied from the required value for optimal load-line matching of the sub-PA cell, the entire PA performance such as output power and efficiency will be significantly degraded.
Ignoring the finite coupling effect and the mutual inductance, the efficiency and the input impedance, $R_{IN}$, of both the PCT and the SCT in Figure 61 can be expressed as

$$\eta_{PCT} = \frac{1}{1 + \left(\frac{n^2 R_1 + R_2}{M}\right)/R_L}$$

(66)

$$\eta_{SCT} = \frac{1}{1 + \left(Mn^2 R_1 + MR_2\right)/R_L}$$

(67)

$$R_{IN\_PCT} = R_1 + \frac{M}{n^2} (R_2 + R_L)$$

(68)

$$R_{IN\_SCT} = R_1 + \frac{1}{Mn^2} (MR_2 + R_L)$$

(69)

These quantities can be calculated using realistic values for $R_1$ and $R_2$ (1 $\Omega$) assuming that the inductance of the primary winding is 1 nH with a quality-factor ($Q$) of 15, which is reasonable assumption based on simulated EM characteristics of the transformer. With varying $n$ and $M$, calculated input impedance and efficiency for both transformers are summarized in Table 6 and graphically presented in Figure 62 and 63.

As indicated in equations above and observed from plotted graph in Figure 62 and 63, the input impedance and the efficiency of the PCT is directly proportional to the number of combining inputs, $M$, whereas those of the SCT is inversely proportional to $M$ (Note that these values do not account for power leakage due to non-ideal electromagnetic (EM) coupling between the primary and the secondary winding inductors). These relations imply that the PCT may be superior to the SCT in terms of the efficiency especially when a large $M$ is required. This is because the SCT has a larger series resistance at the secondary side, which is a critical factor for the efficiency degradation.
The amount of the secondary series resistance of the SCT is proportionally increased as the number of combining inputs, $M$, is increased, significantly degrading the transformer efficiency. In contrast, the secondary winding of the PCT is equally shared by all primary windings, and the number of the secondary windings does not need to be increased for larger $M$ (i.e., there is only one secondary winding with fixed physical size). This superiority of the PCT becomes more apparent when higher turn ratios are used since multi-turn windings have longer physical lengths and larger series resistance.

However, it is not practical to choose the PCT to combine more than three PAs for two reasons. The first reason is the difficulty in physical implementation of the PCT with a large $M$. For the parallel combining using planar inductors, multiple primary

![Table 6. Calculated input impedances and efficiencies of the combining transformers.](image)

<table>
<thead>
<tr>
<th>Turn ratio, n</th>
<th>No. of inputs, M</th>
<th>Efficiency (%)</th>
<th>$R_{IN}$(Ω)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCT</td>
<td>SCT</td>
<td>PCT</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>96.2</td>
<td>96.2</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>97.1</td>
<td>92.6</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>97.4</td>
<td>89.3</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>97.6</td>
<td>86.2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>90.9</td>
<td>90.9</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>94.3</td>
<td>83.3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>95.5</td>
<td>76.9</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>96.2</td>
<td>71.4</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>83.3</td>
<td>83.3</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>90.1</td>
<td>71.4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>92.6</td>
<td>62.5</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>93.9</td>
<td>55.5</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>74.6</td>
<td>74.6</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>84.7</td>
<td>59.5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>88.8</td>
<td>49.5</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>90.9</td>
<td>42.4</td>
</tr>
</tbody>
</table>
Figure 62. Calculated efficiencies of the power combining transformers with varying turn ratio \( n \) and the number of combining inputs \( M \). (a) For the PCT. (b) For the SCT.
Figure 63. Calculated input impedances of the power combining transformers with varying turn ratio ($n$) and the number of combining inputs ($M$). (a) For the PCT. (b) For the SCT.
windings must be evenly interwoven with the secondary winding to achieve equal coupling coefficient preserving equal self-inductances of the primary windings. An unequal coupling coefficient and winding inductances will cause mismatches between distributed RF signals in each combining path, consequently generating distortion at the output signal. Even distribution and placement of more than three primary windings coupled to one secondary winding in a planar structure is difficult to implement. If higher turn ratios are required, implementation of the transformer becomes more challenging. Secondly, the PCT with \( M>3 \) has an excessively large input impedance transformed from the load impedance (50 \( \Omega \)), which is undesirable for PA designs. For example, the PCT with \( M=4 \) and \( n=1 \) exhibits very large input impedance (205 \( \Omega \)), thus additional matching network must be employed to down-transform the impedance to an acceptable level. However, overall loss at the output matching network will be substantially large due to such a high impedance transformation ratio. Although the excessively large input impedance value of the PCT with large \( M \) can be lowered by choosing larger value of the turn ratio \((n>3)\), the PCT with larger turn ratio is not realizable as fore-mentioned.

Therefore, the SCT seems to be a more appropriate option to implement a high-power PA combining more than three sub-PAs. For example, the SCT with \( M=4 \) and \( n=1 \) has the input impedance close to the optimal load impedance value for the individual sub-PAs (i.e., the sub-PA cell for 28-dBm output power under 3.3-V supply requires around 15-\( \Omega \) \( R_{opt} \), and the SCT with \( M=4 \) and \( n=1 \) exhibits \( R_{IN} \) of 14.5 \( \Omega \)). However, the large series resistance of the secondary winding, which is proportionally increased with increasing number of combining inputs, severely degrades the transformer efficiency. The large physical size of the SCT with large \( M \) is another design hurdle toward the
implementation of compact single-chip PAs (e.g., the SCT with \( M=4 \) requires four primary windings placed side by side and four serially connected secondary windings coupled to all primary windings [66]).

In short, both the PCT and the SCT are not practical solutions to implement a single-chip high-power CMOS PA combining more than three sub-PAs. Although the SCT with \( n=1 \) is a plausible choice to implement a high-power PA by combining four PAs, its limited efficiency and large size must be further improved.

4.3. Parallel-Series Combining Transformer (PSCT)

To resolve the contradiction in the characteristics between the PCT and the SCT, a novel approach to implement compact and efficient power combining transformer is proposed. The proposed transformer is a combination of the PCT and the SCT, and performs parallel (current) combining and series (voltage) combining at the same time. This parallel-series combining transformer, or PSCT, can generate adequate input impedance levels independent of the number of combining inputs when the numbers of parallel-combining and series-combining are preserved to be same, which is advantageous as compared to the conventional combining transformers that have varying input impedance values with a varying \( M \). In addition, the PSCT effectively halves the number of secondary windings as compared to the SCT for the same number of total combining inputs, which leads to improved efficiency and reduced overall transformer size.

4.3.1. Characteristics of the PSCT
The structure of proposed PSCT is presented in Figure 64. There are \(N_P\)-primary windings that are coupled to one secondary winding (i.e., \(N_P\)-parallel-combining) and \(N_S\)-sets of \(N_P\)-parallel combining transformers connected in series via multiple secondary windings (i.e., \(N_S\)-series-combining). Thus, proposed transformer performs \(N_P\)-parallel-combining and \(N_S\)-series-combining simultaneously in a single structure, and effective number of combining inputs becomes \(M=N_PN_S\). Ignoring the non-ideal coupling effect, relations for voltages and currents at each node in Figure 64 can be obtained as

\[nN_3V_1' = V_2'\]  
(70)

\[V_1' = V_1 - I_1R_i\]  
(71)

\[V_2 = V_2' - I_2N_SR_2\]  
(72)

\[nI_2 = N_PI_1\]  
(73)
and the efficiency of the transformer and the transformed input impedance at each input terminal can be derived as

$$
\eta_{PSCT} = \frac{1}{N_pN_s} \frac{I_2V_2}{I_1V_1} = \frac{1}{1 + \left( \frac{N_s}{N_p} n^2 R_1 + N_s R_2 \right) / R_L}
$$

(75)

$$
R_{IN\_PSCT} = \frac{V_i}{I_i} = R_1 + \frac{N_p}{N_s} \frac{1}{n^2} (R_L + N_s R_2).
$$

(76)

It is observed that the efficiency of the PSCT is less sensitive to the series resistance of the primary winding \((R_1)\) when \(N_P\) equals \(N_S\). In contrast to the PCT or the SCT, the effect of \(R_1\) on the efficiency of the PSCT becomes independent of the number of combining inputs in this case (i.e., \(N_P=N_S\)). However, the secondary series resistance \((R_2)\) still dominantly affects the efficiency in accordance with the number of series combining inputs \((N_S)\) as in the SCT, but its effect is mitigated due to the reduced number of series combining inputs for the same number of total combining inputs (i.e., \(M=2N_S\) if \(N_P=N_S\)).

We can expect that the efficiency of the PSCT will be lower than that of the PCT because of the series combining parts, but higher than that of the SCT if the inductances of primary and secondary windings are assumed to be same for both transformers.

Unlike for the conventional combining transformers, the input impedance of the PSCT can also be independent of the number of combining inputs when \(N_p\) becomes equal to \(N_S\). In other words, for any number of combining inputs with equal \(N_P\) and \(N_S\), the PSCT exhibits almost constant input impedance that is only dominantly dependent on the turn ratio, \(n\). In an ideal case, the input impedance is simply expressed as
Thus, we can implement the combining transformer that realizes required input impedance by only determining the turn ratio of parallel combining part. Avoiding the effect of the number of combining inputs on the input impedance, the proposed transformer can have more degrees of freedom in designing PAs. The efficiency and input impedance of three power combining transformers: PCT, SCT, and PSCT with four combining inputs \((M=4)\) are calculated and compared in Figure 65. For the case of \(n=2\), the PSCT has an \(R_{IN}\) of 14 \(\Omega\) with an efficiency of 89.3 \%, while the SCT with \(M=4\) and \(n=2\) shows an \(R_{IN}\) of 4.4 \(\Omega\) with an efficiency of 71.4 \%. It is more reasonable to compare the PSCT with \(M=4\) and \(n=2\) to the SCT with \(M=4\) with \(n=1\) considering their comparable input impedance values (i.e., \(R_{IN}\) of 14 \(\Omega\) for the PSCT and 14.5 \(\Omega\) for the SCT). Even in this comparison, the PSCT has a higher efficiency than that of the SCT (86.2 \%).

A more realistic equation for the efficiency of the PSCT with \(N_P\)-parallel-combining and \(N_S\)-series-combining can be derived including non-ideal coupling effect between the primary and the secondary windings (i.e. considering finite coupling coefficient and mutual inductance between \(L_1\) and \(L_2\)). Each transformer can be replaced with a T-equivalent model [70] as presented in Figure 66. Due to finite coupling coefficient \((k<1)\), there exists leakage inductance at each primary and the secondary side. At the secondary side, the leakage inductance of a single secondary winding is magnified by series-combining part and distributed by parallel-combining part to form an effective leakage inductance of \(N_S(1-k)L_2/N_P\). First, we define the overall impedance at the secondary side excluding the ideal transformer as

\[
R_{IN_{PSCT}} = \frac{R_L}{n^2}.
\]

(77)
Figure 65. Calculated efficiencies and input impedances of three power combining transformers: PCT, SCT, and PSCT with four combining inputs ($M=4$). (a) Efficiencies. (b) Input impedances.
From voltages and currents at each node, we can obtain following relations:

\[ N_S nV'_1 = V'_2 \]  
\[ I_M = \frac{V'_2}{j\omega kL N_S} = \frac{I_2 Z_2}{j\omega kL N_S} \]  
\[ I_1 = \frac{nI_2}{N_p} + I_M = \frac{nI_2}{N_p} + \frac{I_2 Z_2}{j\omega kL N_S} = I_2 \left( \frac{n}{N_p} + \frac{Z_2}{j\omega kL N_S} \right) \]  

We can define a current transfer factor, \( F \), as

\[ \frac{1}{F} = \frac{n}{N_p} + \frac{Z_2}{j\omega kL N_S} \]  

thus the secondary current, \( I_2 \), is expressed as a function of the primary current, \( I_1 \), as
\[ I_2 = F \cdot I_1, \]  
\[ (83) \]

The input voltage, \( V_1 \), is then derived as
\[ V_1 = I_1 \left( R_1 + j \omega (1-k)L_1 \right) + \frac{I_2 Z_2}{nN_S} = I_2 \left( \frac{1}{F} \left( R_1 + j \omega (1-k)L_1 \right) + \frac{Z_2}{nN_S} \right), \]
\[ (84) \]

and the efficiency of the PSCT can be obtained as
\[ \eta_{\text{PSCT}} = \frac{1}{N_S N_p} \left( \frac{\text{Re}(I_2 V_1^*)}{\text{Re}(I_1 V_1^*)} \right) = \frac{1}{N_S N_p} \cdot \frac{R_L}{\text{Re} \left( \frac{1}{F} \left( R_1 + j \omega (1-k)L_1 + \frac{Z_2}{nN_S} \right) \right)} \]
\[ (85) \]

**4.3.2. Implementation of the PSCT**

As observed from the previous section, the PSCT can supplement drawbacks of the conventional on-chip power combining transformers for high-power PA designs. We can legitimize the utilization of the PSCT for the following reasons:

1. To implement a high-power PA with greater than 33-dBm output power, it is practical to have four sub-PAs with 28-dBm individual output power (\( R_{opt} = 14 \Omega \)), assuming about 1-dB insertion loss at the power combiner.
2. The PCT with four primary inputs is not a plausible option due to difficulty in physical implementation.
3. The 4-primary 1:1 SCT (i.e., \( M=4 \) and \( n=1 \)) is suitable to use in terms of transformed input impedance (\( R_{IN} = 14.5 \Omega \)), but its size is excessively large due to serially connected secondary windings.
4. The PSCT with two parallel combining inputs (\( N_p = 2 \)) and two series combining inputs (\( N_s = 2 \)) having \( n=2 \) for parallel combining parts exhibits appropriate input impedance (\( R_{IN} = 14 \Omega \)). In addition, its size is approximately
half of the 4-primary 1:1 SCT because the number of series-combining is halved, which results in higher or comparable efficiency (depending on the actual dimension of implemented transformer).

Therefore, we propose a 4-primary 1:2 PSCT (i.e., $M=4$ and $n=2$) as illustrated in Figure 67. There are four primary inputs that are supposed to be connected to the outputs of

![Figure 67. Implementation of proposed 4-primary 1:2 PSCT (a) Conceptual diagram. (b) Actual layout.](image)
individual PAs. Two PCTs are composed of two primary windings that are interwoven with each other to achieve equal coupling effects to the secondary winding with a turn ratio of $n=2$, and are connected in series. Thus, the transformer performs a series-combining of two parallel-combining parts (i.e., $N_P=N_S=2$). The turn ratio of $n=2$ was chosen to realize the impedance transformation ratio of 1:4 based on the (77). The size of winding inductors must be carefully decided based on iterative EM simulations because it affects the optimal operating frequency and the bandwidth of the transformer in terms of the efficiency. The primary inductance of 1 nH was selected for this design. Since the transformer was implemented using two thick metals (4-$\mu$m thick aluminum and 3-$\mu$m thick copper) that are connected through whole via trace as presented in Figure 67(b), high-$Q$ winding inductors can be realized. The width of windings and the spacing between adjacent windings are 30 $\mu$m and 10 $\mu$m, respectively. EM simulations using Ansoft HFSS were performed to characterize the PSCT. The simulation setup is shown in Figure 68. Because the PSCT is the series connection of two PCTs, only half-section of the PSCT, which is the 1:2 PCT with two primary inputs, can be simulated to characterize the self-inductance of the primary and the secondary windings ($L_1$ and $L_2$), the mutual inductance ($M$) and the coupling coefficient ($k$) between the primary and the secondary winding, the quality factors of the primary and secondary winding ($Q_1$ and $Q_2$), and the effective turn ratio ($n$). These parameters can be obtained with following equations:

$$L_{1,2} = \frac{Im(Z_{11} + Z_{22} - 2Z_{12})}{\omega}$$

(86)

$$Q_{1,2} = \frac{Im(1/Y_{11})}{Re(1/Y_{11})}$$

(87)
Figure 68. Simulation setup for characterization of half-section PSCT. (a) For the primary winding. (b) For the secondary winding. (c) For the coupling effect and the mutual inductance.
The transformer efficiency based on the calculation of maximum available gain is given by

$$\eta = 10^{\frac{\text{MaxGain(dB)}}{10}} \times 100$$

Table 7 summarizes the EM characteristics of the half-section PSCT. The efficiency of the entire PSCT structure was also obtained from the EM simulation. The PSCT achieves a maximum available efficiency of 84% at 2.4 GHz (i.e., insertion loss of 0.73 dB) when it has differential configuration and 73.5% when it has single-ended configuration.

Table 7. EM characteristic of the half-section of the PSCT (at 2.4 GHz).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_1$</td>
<td>0.925 nH</td>
</tr>
<tr>
<td>$L_2$</td>
<td>3.17 nH</td>
</tr>
<tr>
<td>$R_1$</td>
<td>1.125 Ω</td>
</tr>
<tr>
<td>$R_2$</td>
<td>2.634 Ω</td>
</tr>
<tr>
<td>$Q_1$</td>
<td>13.2</td>
</tr>
<tr>
<td>$Q_2$</td>
<td>17.9</td>
</tr>
<tr>
<td>$M$</td>
<td>1.045 nH</td>
</tr>
<tr>
<td>$k$</td>
<td>0.65</td>
</tr>
<tr>
<td>$n$</td>
<td>1.85</td>
</tr>
<tr>
<td>Efficiency</td>
<td>88%</td>
</tr>
</tbody>
</table>
The conventional 4-primary 1:1 SCT was also implemented and simulated to compare its performance to the proposed structure. Layout of the PSCT and the SCT with 4-primary inputs are illustrated in Figure 69. As mentioned in the previous section, the size of the SCT (2.72 x 0.55 mm$^2$) is almost twice of the size of the PSCT (1.43 x 0.58 mm$^2$). The simulated insertion losses and transformed input impedance of these two transformers are compared in Figure 70. While the SCT exhibits an insertion loss of 0.85 dB at 2.4 GHz, the PSCT achieves a lower insertion loss of 0.73 dB due to the reduced number of series combining parts. The input impedance of the PSCT is 16.5 Ω at 2.4 GHz, which is very close to the required load impedance for individual PAs ($R_{opt}$) to generate targeted output power. The SCT has relatively lower input impedance levels.
Figure 70. Simulated performances of the proposed PSCT (M=4, n=2) and SCT (M=4, n=2). (a) Insertion loss. (b) Input impedance.
because the secondary winding is surrounding the primary windings on the outer side, which results in larger secondary self-inductance of each coupled part leading to a higher effective turn ratio. Additional inductance at the cross-section area of the adjacent primary windings also increases the effective turn ratio, lowering transformed input impedance.

4.4. Design of a PSCT-based High-Power CMOS PA

4.4.1. Structure of the PSCT-based CMOS PA

Using the proposed PSCT, a fully-integrated PA is implemented in CMOS technology. The diagram of the proposed PA structure is presented in Figure 71. There are four PA branches that are combined by the PSCT at the output. Adjacent two PA branches are combined in parallel (i.e., output AC currents are combined), and two sets of
parallel combinations are connected in series (i.e., output AC voltages are combined). The PSCT performs not only the power combining but also the impedance down-transformation that is realized by the turn ratio between the primary and the secondary windings. Thus, additional output matching networks are unnecessary (except for shunt capacitors to tune out the reactive component at the drain node of sub-PAs). The entire PA system consists of a driver stage and a combined power stage that are interconnected by LC matching network. An on-chip transformer as an input balun is used along with capacitive matching network to convert a single-ended external signal to a differential internal signal. The schematic diagram of proposed PA is presented in Figure 72.

![Schematic diagram of designed CMOS PA using the PSCT as a power combiner.]

**Figure 72.** Schematic diagram of designed CMOS PA using the PSCT as a power combiner.

### 4.4.2. Circuit Design

The driver and power stage amplifiers are pseudo differential amplifiers with cascode configuration. The thick gate-oxide transistors ($L=0.4 \ \mu m$) are used for the
common gate (CG) devices \((M_2)\) in the driver and power stage amplifiers to avoid the
device breakdown issues. The width of common source (CS) devices \((M_1)\) in the power
stage amplifiers is \(3072 \mu m\) (384 fingers with the unit finger width of \(8 \mu m\)). This device
size was determined through the load-pull simulation in the Agilent ADS in such a way
that each power-cell can generate the maximum output power of 28 dBm with reasonable
efficiency. The size of the CG devices in the power stage amplifiers is \(3936 \mu m\) (492
fingers with the unit finger width of \(8 \mu m\)). Because four power-cells are combined in
parallel, effective cell size of the entire power stage becomes 12.288 mm \((3072 \mu m \times 4)\).
The cell size of the driver stage is determined in such a way that the output power of the
driver stage amplifier is not compressed until the output power of the entire power stage
begins to compress. Considering the gain of the combined power stage, a minimum \(P_{1\text{dB}}\)
level of the driver stage amplifier can be obtained, and the size of one-fifth of the entire
power stage \((2560 \mu m, 321 \text{ fingers with the unit finger width of } 8 \mu m)\) was chosen for
this design. Multiple bonding-wires are used to ground the source of the driver and power
stages amplifiers to minimize the effect of wire inductance and resistance. To reduce the
second-order distortions, large bypass capacitors \(\left(C_{\text{bypass}}\right)\) are used to connect the center-
tab of the primary windings of the PSCT to source of each power stage amplifiers. While
the stability of the entire PA system may be degraded with excessively large \(C_{\text{bypass}},\) a
small capacitance may reduce the gain of the PA system. Therefore, the size of the
bypass capacitor must be carefully chosen compromising the gain and the stability of the
PA. \(12\text{-pF for } C_{\text{bypass}}\) was chosen for this design. Each cascode differential amplifier has
a resistive feedback network from the drain of the CG transistors to the gate of the CS
transistors to improve AM-AM characteristic and stability. DC supply currents for the
driver and power stage are fed through the center-tab of the differential inductor or the transformer. Since no bonding-wires are included in the matching network, accurate frequency matching can be realized. The CS transistors are biased by external voltage source applied to the gate through large resistors. The power stage amplifiers are biased for class-AB operation for good efficiency and linearity. The driver stage is biased to act more like a class-B (i.e., deep class-AB) amplifier to maximize the efficiency. The deep class-AB operation also expands the gain of the driver stage amplifier and compensates for the gain compression of the power stage amplifiers. Expanded gain (or better AM-AM characteristic) is beneficial to achieve better EVM performance for the entire PA system.

4.4.3. Measurement Results

The high-power linear CMOS PA has been fabricated in 0.18-μm CMOS (IBM 1P6M) technology. A micro-photograph of fabricated PA is presented in Figure 73. The PA has the size of 2.1 x 1.64 mm². The PA chip was mounted on the FR-4 PCB for measurements. The PCB line loss (0.3 dB) was de-embedded to obtain actual PA performances. However, the effect of bonding-wires is included in the measurement results. The gain, $P_{1dB}$, $P_{sat}$, and PAE were measured with single-tone continuous wave signal. The gate bias voltages for the driver and power stages amplifiers are 0.43 V and 0.46 V, respectively, at which the overall PA has the highest $P_{1dB}$ with small gain variation. The gate bias voltages for the CG devices in the driver and power stage amplifiers are chosen to be 3.3 V and 3.2 V, respectively. Relatively higher bias voltage for CG devices improves AM-AM performance of the PA with small penalty in the
efficiency. The measured gain and efficiency of designed PA at the operating frequency of 2.4 GHz is presented in Figure 74(a). The PA achieves a gain of 22 dB, a $P_{1dB}$ of 31.5 dBm, a $P_{sat}$ of 34 dBm, a peak PAE of 34.9%, and a peak drain efficiency (DE) of 41%. The gain variation is less than 0.5 dB. The measured gain/$P_{1dB}$ at 2.3 GHz and 2.5 GHz are 21.5 dB/32.5 dBm and 22 dB/29 dBm, respectively, as shown in Figure 74(b).

The measured S-parameters are plotted in Figure 75. The input and output return losses ($S_{11}$ and $S_{22}$) are better than 15 dB at 2.4 GHz. The small signal gain ($S_{21}$) of the PA is greater than 19 dB from 1.8 to 3.2 GHz (i.e., 3-dB bandwidth is 1.4 GHz). The EVM was also measured to observe the linearity of designed PA using the IEEE 802.11g WLAN OFDM 64-QAM (channel bandwidth is 20 MHz) and 802.16e WiMAX OFDM 64-QAM (channel bandwidth is 24 MHz) modulated signals at 2.4 GHz. The PA satisfies WLAN EVM specification (-25 dB or 5.6%) up to 23.5-dBm output power, and WiMAX

Figure 73. Micro-photograph of fabricated PA (2.1 x 1.64 mm$^2$).
Figure 74. Measured results of designed PA. (a) Gain, DE, and PAE. (b) Gain versus $P_{\text{out}}$ at different frequencies.
Figure 75. Measurement results of the PSCT-based high-power linear PA: S-parameters. (a) $S_{11}$ and $S_{22}$. (b) $S_{21}$ and $S_{12}$.
Figure 76. Measurement results of designed PA: EVM. (a) For WLAN 54 Mbps 64-QAM signal. (b) For WiMAX 54 Mbps 64-QAM signal.
EVM specification (-31 dB or 2.8%) up to 19.5-dBm output power (Figure 76). Measured spectrums and constellations for WLAN and WiMAX signals are presented in Figure 77. The PA consumes an idle DC current of 270 mA (30 mA for the driver stage amplifier and 60 mA for each of power stage amplifier) from 3.3 V supply. The performance of designed PA is compared to other state-of-the-art linear CMOS PAs in Table 8.

Figure 77. Measurement results of the PSCT-based high-power linear PA: constellation and spectrum (at 2.4 GHz). (a) For WLAN OFDM 64-QAM signal ($P_{\text{out}}$=23 dBm). (b) For WiMAX OFDM 64-QAM signal ($P_{\text{out}}$=19 dBm).
4.5. Conclusion

A novel power combining transformer, parallel-series combining transformer (PSCT), was proposed to implement a single-chip high-power linear CMOS PA. The PA has been fabricated in a standard CMOS technology and achieved a high $P_{1\text{dB}}$ (31.5 dBm), a high $P_{\text{sat}}$ (34 dBm), and a high peak PAE (34.9%) with a gain of 22 dB at the frequency of 2.4 GHz. While meeting the EVM specifications for WLAN and WiMAX OFDM 64-QAM signals at 23.5 dBm and 19.5 dBm, respectively, the designed PA demonstrated the highest $P_{1\text{dB}}$ and $P_{\text{sat}}$ with good efficiency among reported CMOS linear PAs. The effectiveness of proposed transformer, PSCT, as a power combiner for implementations
of high-power linear CMOS PAs has been successfully validated. However, the linearity performance of the PA can be further improved by employing other linearization techniques.
5.1. Introduction

The most auspicious frequency bands in the MMW range for the next-generation broadband wireless communications are the V-band (50 to 75 GHz) and the W-band (75 to 110 GHz). The V-band has been vigorously exploited to develop applications of short-range high data-rate (> 1 Gb/s) wireless accesses such as WPAN or WirelessHD™. The W-band provides more available frequency resources without severe propagation attenuation problem. With abundant spectral bandwidth available, the W-band range is being actively utilized for many commercial (e.g., automotive cruise control systems), and military (e.g., aerospace/satellite smart radar systems and missile control systems) applications [32]-[36].

The structure of the receiver systems for MMW applications also adopts conventional topologies that are widely used for RF applications. The direct down-conversion receiver illustrated in Figure 78 is perhaps the most commonly used architecture. There is no need for an image rejection filter and the second down-conversion step as in the heterodyne receiver system because the signal in the carrier frequency is directly down-converted to the baseband. Therefore, more compact implementation of the system with reduced power consumption is realizable, which is the biggest advantage for many MMW applications.
The most successful implementation of the V-Band direct down-conversion receiver system using SiGe technology has been demonstrated in [30]. The block diagram of this V-Band receiver is presented in Figure 79, where shaded region indicates actually integrated receiver. This receiver consists of an LNA with single-ended-input and differential-output (working as an active balun), two down-conversion mixers that are based on conventional Gilbert-cell multipliers, a frequency tripler to generate the LO input, a differential branch-line directional coupler, and buffers. The novelty in this structure is the integration of the directional coupler to generate 90°-phase-shifted signals, which was implemented using quarter-wavelength transmission lines. The measured performance of critical circuit blocks is also summarized in Figure 79, where the down-converter includes the mixer and the buffer. Although the measurement data for the entire receiver chain was not reported, mathematical calculation of cascaded system shows 5.8-dB noise figure (NF), 31-dB gain, -22-dBm input-referred intercept point (IIP₃), and -32-dBm input referred P₁dB (P₁dB_in).

Figure 78. Diagram of a direct down-conversion receiver system.
In the W-Band, integrated receiver systems using SiGe technology have been implemented employing relatively simple architectures [35], [36]. The diagrams of these W-Band receivers are presented in Figure 80. Applications in the W-band frequencies (e.g., automotive control systems, passive imagers, etc) do not exploit advanced modulation schemes, thus separation of I/Q signal paths within the receiver chain is not necessary. For the automotive radar application in 77 GHz, the demonstrated receiver in [35] achieved 30-dB conversion gain, 11.5-dB NF, -26-dBm $P_{1\text{dB},\text{in}}$, and -21.6-dBm $IIP_3$, consuming 440 mW from 5.5 V supply. This W-Band receiver features a single-stage cascode LNA, an active balun, and a Gilbert-cell-based double balanced mixer with an LO balun. The LO balun was implemented using LC passive networks of which inductive components were realized by transmission lines. Another W-Band SiGe receiver that integrated a two-stage differential cascode LNA with a double-balanced mixer and a voltage-controlled oscillator (VCO) has been reported in [36]. This receiver achieved 40-dB conversion gain, 8-dB NF, -38-dBm $P_{1\text{dB},\text{in}}$ at 77 GHz, consuming 195 mW.

<table>
<thead>
<tr>
<th></th>
<th>LNA</th>
<th>Downconverter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq.</td>
<td>61.5 GHz</td>
<td>61.5 GHz</td>
</tr>
<tr>
<td>Gain</td>
<td>14.7 dB</td>
<td>16 dB</td>
</tr>
<tr>
<td>NF</td>
<td>4.5 dB</td>
<td>14.8 dB</td>
</tr>
<tr>
<td>$P_{1\text{dB},\text{in}}$</td>
<td>-20 dBm</td>
<td>-17 dBm</td>
</tr>
<tr>
<td>$IIP_3$</td>
<td>-8.5 dBm</td>
<td>-7 dBm</td>
</tr>
<tr>
<td>$S_{11}$</td>
<td>-6 dB</td>
<td>-12 dB</td>
</tr>
<tr>
<td>DC Power</td>
<td>6 mA @ 1.8 V</td>
<td>112 mA @ 2.7 V</td>
</tr>
</tbody>
</table>

Figure 79. Diagram of the 60 GHz receiver in SiGe demonstrated in [30].
As a next step toward higher operating frequencies, we are motivated to explore potential usability of advanced SiGe technology to build a high-dynamic-range receiver front-end system operating at 90 to 94 GHz. This frequency band is suitable for long-range high-resolution frequency modulated continuous wave (FMCW) radar applications [76]-[78]. Proposed structure of a W-band (94 GHz) receiver is illustrated in Figure 81. The receiver adopts the direct down-conversion topology. The balun transforms single-ended signals that are amplified by an LNA to differential signals to be down-converted by a mixer. The balun also performs conversion of single-ended reference clock signal from the external oscillator to differential LO signal for the mixer input. Double-balanced mixer can improve linearity and dynamic range of the receiver system.

This chapter presents the design procedures and implementation issues of circuit blocks essential to a W-band receiver front-end system: a balun, an LNA, and a mixer. Considering diverse design challenges for active and passive circuits operating at such a high frequency (94 GHz), careful characterizations and optimizations of individual functional blocks are performed. Section 4.2 briefly describes SiGe HBT BiCMOS

Figure 80. Diagrams of reported W-Band SiGe receivers. (a) In [35]. (b) In [36].
technology that was used to fabricate the circuits. The design of the planar-type Marchand balun is introduced in Section 4.3. Section 4.4 and 4.5 present design issues and methodologies of the LNA and the down-conversion mixer along with simulation and measurement results, respectively. Implementation of an integrated W-band receiver is presented in Section 4.6. The conclusion follows in Section 4.7.

5.2. **SiGe Technology**

Although III-V technologies (GaAs and/or InP) provide significant performance advantages, silicon-based technologies may be preferred to be used to reduce the development cost and ease the integration with baseband digital circuitries. An advanced generation (third or fourth generation as in Figure 5) SiGe BiCMOS technology can be a good solution to implement high performance MMW circuit components since it

![Diagram of proposed W-band (94 GHz) receiver front-end system.](image)
provides improved performance of active/passive components with advantages of the low-cost silicon substrate. A 0.12-µm-emitter-width IBM 8HP SiGe BiCMOS technology was chosen to fabricate the W-band receiver front-end circuits. The technology supports SiGe HBTs with peak $f_T/f_{max}$ of 200/280 GHz with 0.13-µm 1.2-V Si CMOS transistors and passive elements including MIM capacitors, thin-film resistors, on-chip inductors, and transmission lines. Seven levels of metallization are available: two top aluminum layers and five copper bottom layers. The top aluminum metal features 4-µm thickness that enables implementations of high-$Q$ inductors and transmission lines.

The advancements in SiGe HBT performance is mainly achieved by smaller base-collector capacitance ($C_{BC}$) which is realized by eliminating out-diffusion of the extrinsic base [54]. Employment of laterally scaled emitter width minimizes base resistance ($R_b$) and improves frequency/noise characteristics. Thus, high speed, low noise, and high gain devices are realizable in this technology, which is a great advantage for MMW circuit designs. The cross section of the SiGe HBT is presented in Figure 82 [54].

![Figure 82. Cross section of SiGe HBT.](image-url)
5.3. Design of a W-band balun

Differential functionality of circuits operating at high frequencies is strongly desired because it reduces common mode noise and improves linearity and dynamic range. Although single-ended (or unbalanced) signals are inevitably generated by antennas or other preceding functional blocks, a conversion to differential (or balanced) signals is still realizable using an additional circuit block.

A balanced-to-unbalanced (balun) signal conversion can be realized using active or passive circuits. Active baluns normally employ differential amplifiers of which only one side of input or output terminals is taken for applying or extracting the signals. The active baluns are advantageous in terms of generating extra gain in addition to converting signals, but have critical drawbacks. The biggest problem is amplitude/phase mismatch issue that is raised by unbalanced differential functionality of the amplifier especially when the operating frequency of the circuit is very high (as in MMW frequency range) [79]. Additional nonlinearity caused by the active balun also degrades overall linearity performance of the receiver system. Moreover, the active baluns also consume power which is commonly comparable amount to power consumed by other critical functional blocks. For these reasons, passive baluns with various configurations have been widely used for applications in MMW integrated circuits [80]-[81]. Among them, the planar-type of the Marchand balun [84] is perhaps one of the most attractive solutions due to its wideband performance and ease of design. The planar Marchand balun consists of two coupled quarter-wavelength (λ/4) transmission lines, which can be realized using microstrip lines, CPW lines, or spiral coils. The schematic representation of the planar Marchand balun is shown in Figure 83. One of each coupled lines has ground termination
at the input side, and opposite coupled lines are shorted together at the output side. An unbalanced input port and balanced output ports of the balun are terminated by impedance $Z_S$ and $Z_L$, respectively.

![Diagram of a planar-type Marchand balun.](image)

**Figure 83. Schematic of a planar-type Marchand balun.**

### 5.3.1. Quarter-wavelength coupled line

To understand how the balun works, it is necessary to review characteristics of a coupled quarter-wavelength transmission line which is illustrated in Fig. 84. The coupled line is composed of two closely placed straight transmission lines, and the coupling coefficient between two lines is denoted as $k$. When terminal impedances of the individual lines are chosen to be equal to the characteristic impedance of the lines (i.e., $Z_S=Z_L=Z_0$), the $S$-parameters of the four port network can be represented as [85]

$$
[S]_{coupler} = \begin{bmatrix}
0 & -j\sqrt{1-k^2} & k & 0 \\
-j\sqrt{1-k^2} & 0 & 0 & k \\
k & 0 & 0 & -j\sqrt{1-k^2} \\
0 & k & -j\sqrt{1-k^2} & 0
\end{bmatrix}
$$

(92)
The coupling coefficient can be calculated from effective capacitance of unit length of coupled lines using the equation given by

\[
k = \frac{C_{C_{\text{total}}}}{\sqrt{(C_{S1_{\text{total}}} + C_{C_{\text{total}}})(C_{S2_{\text{total}}} + C_{C_{\text{total}}})}},
\]  

(93)

where \( C_{C_{\text{total}}} \) denotes the total equivalent coupling capacitance between two lines and \( C_{S1_{\text{total}}} \) and \( C_{S2_{\text{total}}} \) are the total equivalent capacitances between the conductors and ground plane as presented in Figure 85. As can be anticipated, the coupling coefficient is increased if \( C_{C_{\text{total}}} \) is large as in a case when two lines are placed very closely. The coupling coefficient is also inversely proportional to \( C_{S1_{\text{total}}} \) and \( C_{S2_{\text{total}}} \) implying that large distance between the conductors and ground-plane is desired for higher coupling.
The coupled line supports two types of excitations: the even mode, where the currents in the conductors are equal in amplitude and in the same direction, and the odd mode, where the currents in the strip conductors are equal in amplitude but in opposite directions. The characteristic impedances in each excitation modes ($Z_{O,e}$ and $Z_{O,o}$) can be derived as [84]

$$Z_{O,e} = \sqrt{Z_S Z_L} \cdot \sqrt{\frac{1+k}{1-k}}$$

(94)

$$Z_{O,o} = \sqrt{Z_S Z_L} \cdot \sqrt{\frac{1-k}{1+k}}$$

(95)

Therefore, we can get relations between the characteristic impedance and the coupling coefficient, $k$, as

$$\frac{Z_{O,e}}{Z_{O,o}} = \frac{1+k}{1-k}$$

(96)

$$k = \frac{Z_{O,e} - Z_{O,o}}{Z_{O,e} + Z_{O,o}}$$

(97)

For better performance in terms of the bandwidth, the ratio of $Z_{O,e}/Z_{O,o}$ of the coupled line must be maximized [82], so higher coupling between two lines is desired.

The hybrid-type coupled transmission line has been simulated using IBM 8HP design-kit model. The length of each conductor line is 300 $\mu$m which is a quarter-wavelength at 94 GHz. The width of signal conductor is 5 $\mu$m to feature the characteristic impedance of 65 $\Omega$ (same structure as one explained in Chapter 2). This value was chosen since the balun will be loaded by output impedance of the preceding LNA and input impedance of following mixer that have higher impedances than 50 $\Omega$. The cross section of the hybrid-type quarter-wavelength transmission line is shown in Figure 86.
The simulated coupling coefficient as a function of the distance between two conductor lines ($D$) is plotted in Figure 87.

Figure 86. Cross-section of the hybrid-type coupled transmission line.

Figure 87. Simulated coupling coefficient of the hybrid-type coupled transmission line ($W=5$ $\mu$m, $S=20$ $\mu$m).
The characteristics of the hybrid-type coupled quarter-wavelength transmission line ($L=300 \ \mu m$, $W=5 \ \mu m$, $S=20 \ \mu m$, and $D=5 \ \mu m$) are compared between the design-kit model and the EM simulated model. The structure of the coupled transmission line for EM simulation is illustrated in Figure 88. The four-port S-parameters are obtained from the simulation to observe characteristics of the coupled line (through, coupling, and isolation). The comparison is presented in Figure 89. The EM model has slightly higher insertion loss (2.1 dB at 94 GHz) than the design-kit model (1.2 dB at 94 GHz). The coupling of the EM model is better than that of the design-kit model ($k$ of the EM model and the design-kit model is 0.67 and 0.625 at 94 GHz, respectively). The isolations for both models are greater than 18 dB at 94 GHz. The design-kit model features relatively optimistic characteristics in terms of the line loss (resistance and inductance as shown in Figure 25 in Chapter 2), but the coupling and isolation characteristics for both models are relatively close.
The quarter-wavelength coupled transmission line itself can be used to implement the balun function. One of two terminals at one side is loaded by the input port that has impedance of $Z_s$, while the other terminal is open circuited for full reflection (it may be shorted to ground as well). Two terminals at the opposite side are loaded by output ports that have impedance of $Z_L$ (this configuration is presented in Figure 90). Two output ports will have signals with the phase difference of $180^\circ$ and the amplitude imbalance of

$$\frac{S_{31}}{S_{21}} = \frac{Z_{o,e} - Z_{o,o}}{Z_{o,e} + Z_{o,o}} = k,$$

where $Z_s$ and $Z_L$ are assumed to be matched to the characteristic impedance of the line.

Calculated amplitude imbalance of the quarter-wavelength coupled transmission line with $Z_{O,o}$ of 35 Ω as a function of $Z_{O,e}$ is plotted in Figure 91. As observed, very high $Z_{O,e}$ is

Figure 89. Simulated characteristics of the hybrid-type coupled quarter-wavelength transmission line.
required to achieve reasonable amplitude balance, which is the reason why the single section of a quarter-wavelength coupled transmission line is not widely used as a balun.
When the coupled transmission line has side ground plane (ground shields) with unequal widths, the ratio of the even-mode characteristic impedance to the odd-mode characteristic impedance \( \frac{Z_{O,e}}{Z_{O,o}} \) is not preserved as it is for the coupled line with equal ground plane widths [87]. The ratio of \( Z_{O,e}/Z_{O,o} \) is maximized when the widths of side ground planes are equal (i.e., symmetric structure has the maximum \( Z_{O,e}/Z_{O,o} \)). The asymmetric structure of the coupled transmission line is illustrated in Figure 92. If the ratio of \( B_2/B_1 \) is enlarged, \( Z_{O,e}/Z_{O,o} \) is decreased, which results in smaller coupling coefficient between two conductors and narrower operating bandwidth. However, this effect is mitigated if the distance between the conductor and the side ground-plane is much larger than the width of the conductor line because the transmission line acts more like a microstrip line rather than a CPW line (i.e., most of the magnetic field from the AC current in the conductor is diminished before it reaches the side ground shield). With different values of \( B_2/B_1 \), the coupling coefficients of the structure are obtained from EM

![Coupled transmission line with unequal ground-plane widths.](image)

**Figure 92.** Coupled transmission line with unequal ground-plane widths.
simulations and presented in Figure 93. There is no significant variation in the values of the coupling coefficient as the difference between widths of two side ground shields is increased. That is, the coupling coefficient is changed from 0.625 for the case of $B_2/B_1=1$ to 0.616 for the case of $B_2/B_1=16$. It was verified with simulations that this amount of variation in the coupling coefficient does not critically deviate the characteristic of the coupled transmission line (e.g., the characteristic impedance).

![Figure 93. Coupling coefficient of coupled transmission line with unequal ground-plane widths.](image)

5.3.2. W-band Marchand balun

By cascading S-parameters of two coupled lines configured as shown in Figure 94 and assuming the source and load impedance are matched, overall S-parameters of the three-port balun can be obtained as
Figure 94. 3-D structure of designed W-band planar-type Marchand balun.

\[
[S]_{balun} = \begin{bmatrix}
\frac{1-3k^2}{1+k^2} & j\frac{2k\sqrt{1-k^2}}{1+k^2} & -j\frac{2k\sqrt{1-k^2}}{1+k^2} \\
-j\frac{2k\sqrt{1-k^2}}{1+k^2} & \frac{1-k^2}{1+k^2} & \frac{2k^2}{1+k^2} \\
-j\frac{2k\sqrt{1-k^2}}{1+k^2} & \frac{2k^2}{1+k^2} & \frac{1-k^2}{1+k^2}
\end{bmatrix}.
\] (99)

where \( k \) is same parameter as for the single coupled lines. If we engineer \( k \) so that it minimizes input reflection of the balun (i.e., \( k=0.567 \) and \( S_{11}=0 \)), we can have

\[
[S]_{balun} = \begin{bmatrix}
0 & \frac{j}{\sqrt{2}} & -\frac{j}{\sqrt{2}} \\
\frac{j}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\
-\frac{j}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2}
\end{bmatrix}.
\] (100)
It is noticed that the coupling coefficient, $k = 1/\sqrt{3}$, for the best of input return loss causes relatively poor output matching and narrow operating bandwidth (note that higher $k$ will help the balun to have flatter amplitude balance over a wide bandwidth). Thus, some compromise is necessary to select appropriate value for $k$ to have balanced performance of the balun. According to simulated data in Figure 87, $k$ of 0.625 can be achieved if $D$ of the coupled line is selected to be 5 $\mu$m. Instead of having $k$ of 0.567,

![Figure 95. Performance of designed W-band planar-type Marchand balun. (a) Micro-photo of the fabricated balun. (b) Amplitude balance. (c) Phase balance.](image_url)
higher $k$ was chosen for this design to have wider operating bandwidth with a little penalty of return loss. The EM simulation was performed using ADS Momentum. Measured performance of the balun with EM simulation results is presented in Figure 95. The balun achieved about -8-dB $S_{21}$ and -9.5-dB $S_{31}$ at 94 GHz. The phase imbalance between two output ports is about 20° at 94 GHz.

5.4. Design of a W-band Low-Noise Amplifier

An LNA is the most critical building block among other components in a receiver chain. Because its noise performance determines the overall noise figure of the receiver, considerably low noise characteristic is required for the LNA. In addition, maximized gain and linearity performances are desired for a high dynamic range of the system, particularly when the receiver exploits a direct down-conversion topology. Good input and output return losses are important as well to be effectively integrated with preceding/following components, and low power dissipation is demanded due to limited power budget.

5.4.1. Noise characteristics of a SiGe HBT

Primary noise sources in a HBT are: 1) base thermal noise, 2) base shot noise, and 3) collector shot noise. These noise sources compose equivalent noise voltage ($v_n^2$) and noise current ($i_n^2$) with cross-correlation factor ($\gamma$) in an HBT as shown in Figure 96 [88]. A noise resistance ($R_n$), an optimal source admittance for minimum noise ($Y_{S,opt}$), and a minimum noise factor ($F_{min}$) for the HBT can be expressed as

$$ R_n = \frac{v_n^2}{4kT\Delta f} $$

(101)
If we consider the HBT as a two-port network, we can express the noise parameters of the HBT as [89], [90]

\[
R_n = r_b + \frac{1}{2g_m} \tag{104}
\]

\[
G_{S,\text{opt}} = \sqrt{\frac{g_m}{2R_n\beta}} + \frac{(\omega C_i)^2}{2g_m R_n} \left(1 - \frac{1}{2g_m R_n}\right) \tag{105}
\]

where \(r_b\) is the base resistance, \(g_m\) is the transconductance, \(\beta\) is the DC current gain, \(C_i\) is total input capacitance (i.e., \(C_i = C_{be} + C_{bc}\)). Then, the noise factor \((F)\) is determined as

\[
F = F_{\text{min}} + \frac{G_n}{R_S} \left|Z - Z_{S,\text{opt}}\right|^2, \tag{106}
\]
where \( G_n \) is the noise conductance and \( R_S \) is the real part of the source resistance. The optimal source impedance, \( Z_{S,\text{opt}} \), reflects specific matching point of the input at which the device has a minimum noise factor (i.e., \( F = F_{\text{min}} \)). The real part of \( Z_{S,\text{opt}} \) is obtained as

\[
R_{S,\text{opt}} = 50 \cdot \text{Re}\left(\frac{1 - \Gamma_{\text{opt}}}{1 + \Gamma_{\text{opt}}}\right),
\]

where \( \Gamma_{\text{opt}} \) is the optimal input reflection coefficient of the device for minimum noise normalized to 50 Ω. Given the operating frequency (\( f \)), dimension of the device, and collector bias current (\( I_C \)), we can obtain \( R_{S,\text{opt}} \) as

\[
R_{S,\text{opt}} = \frac{f_L}{f} \sqrt{\frac{2r_b}{kT I_C^2 q}},
\]

which is inversely proportional to the physical size (emitter length) of the device.

### 5.4.2. Circuit design

The LNA design starts with investigating the optimal bias point of the device in terms of low \( F_{\text{min}} \) and high gain. Within a power budget, the device can be biased at specific collector current level (\( I_C \)) in such a way that the device has reasonable gain and \( F_{\text{min}} \). Once desired bias current density level (\( J_C \)) is found, physical size of the device can be determined to have \( R_{S,\text{opt}} \) very close to 50 Ω (assuming that impedance of preceding component such as an antenna or a switch is 50 Ω). The imaginary part (reactance component) of \( Z_{S,\text{opt}} \) can be compensated by an input matching network. The conjugate output matching can be implemented to have higher gain. The available gain of the amplifier with noise-matched input and conjugate-matched output is called an associate gain (\( G_a \)).
Figure 97. Cascode structure for the W-band LNA.

Figure 97 presents a unit cascode structure for the W-band LNA. The cascode configuration provides better immunity from the Miller effect, hence better isolation between input and output of the LNA can be achieved [90]. Moreover, improved isolation brings better stability over wide operating bandwidth. We selected emitter length ($L_E$) of transistors $Q_1$ and $Q_2$ as 4 $\mu$m and 2.5 $\mu$m, respectively. Multiplicity of two for the transistors was chosen. Therefore, effective emitter area ($A_E$) becomes 0.96 $\mu$m$^2$ for $Q_1$ and 0.6 $\mu$m$^2$ for $Q_2$. Simulated noise and gain characteristics of the cascode unit structure are shown in Figure 98. Depending on values of $J_C$, different maximum cut-off frequency ($f_T$) and $NF_{min}$ can be achieved with opposite tendencies. To have high $f_T$ and low $NF_{min}$, $J_C$ of 3.8 mA/$\mu$m$^2$ was selected as a bias point (Figure 98(a)). This value of $J_C$ provides about 15-dB associate gain at 94 GHz (Figure 98(b)). The
multiplicity of two for the transistors was chosen to have $R_{S,\text{opt}}$ of about 50 Ω (Figure 98(c)). Selected size of the transistors and bias current level enable well-compromised noise and gain characteristics of the input stage of the LNA. Figure 98(d) presents $I_{\text{opt}}$ of the cascode structure for different values of $J_C$ at 94 GHz.

Figure 98. Simulated characteristics of the cascode structure as a function of $J_C$. (a) $f_T$ and $NF_{\text{min}}$. (b) Associate gain. (c) $R_{S,\text{opt}}$. (d) $I_{\text{opt}}$. 
The schematic of designed LNA and simulation results are presented in Figure 99 and 100. The W-band LNA exploits two stages to achieve higher gain enough to compensate the loss at the following stage (the passive balun). The second stage has identical transistor sizes as the first stage with the same bias condition. Base currents for

![Figure 99. Schematic of designed W-band LNA.](image)

![Figure 100. Simulation results of designed W-band LNA.](image)
both stages are fed through quarter-wavelength transmission lines which are part of matching networks. The inter-stage matching network transforms the input impedance of the second stage to the conjugate value of the output impedance of the first stage. The output matching network also performs conjugate matching to the input impedance of the balun. Simulated gain and noise figure of the LNA are 26-dB and 8.4 dB around 96 GHz, respectively. The 3-dB bandwidth is 4 GHz. Input and output return losses are better than 13 dB at the operating bandwidth.

5.5. Design of a W-band Down-Conversion Mixer

For the implementation of transceiver systems in MMW frequency bands, direct up/down-conversion techniques have been widely utilized due to their design simplicity and low power consumption. However, these homodyne structures, especially the receiver side, suffer from critical drawbacks such as vulnerability to even-order distortions and LO self-mixing [92]. In order to mitigate these problems, a fully differential operation is desired for a mixer within a receiver chain. In addition, high noise-floor level at the LO output in the W-band requires better common-mode noise rejection at the mixer input, which also necessitates the use of a double-balanced topology in the W-band mixers. Critical performance metrics of the mixer are voltage conversion gain, noise figure, linearity, and power consumption. The voltage conversion gain is determined as the ratio of $r_{ms}$ voltage amplitude of the down-converted signal to that of the RF input signal. The noise figure of the mixer indicates how much SNR of the down-converted signal is degraded as the signal passes through the mixer. Depending on the type of LO injection, two different definitions of the noise figure can be considered: 1)
single-side band (SSB) noise figure and 2) double-side band (DSB) noise figure (Figure 101). If the LO frequency is chosen lower or higher than the RF frequency, DSB noise figure is appropriate to measure the noise performance of the mixer. The linearity of the mixer is mainly measured by $P_{\text{1dB,in}}$ and IIP$_3$. The linearity characteristic of the mixer dominates the linearity performance of entire receiver front-end system. Moreover, because the mixer is followed by baseband circuitries, second-order linearity characteristic is also important. If the mixer generates significant second-order distortions, which is normally evaluated by input-referred second-order intercept point (IIP$_2$), considerable DC offset is produced, which will desensitize the baseband amplifiers. If the receiver features the direct down-conversion architecture, the DC offset problem becomes tremendously critical since the total gain (or dynamic range) of the baseband blocks is large [92].

![Diagram of noise figure in the mixer](image)

Figure 101. Noise figure in the mixer. (a) Single-side band (SSB) noise figure. (b) Double-side band noise figure.
5.5.1. Double-balanced mixer topology

The mixer down-converts (in a receiver) or up-converts (in a transmitter) the input signals to necessary frequency bands. In a direct down-conversion receiver, the mixer is required to provide a high gain to maximize the dynamic range of the system while the noise performance is less important (i.e., noise figure of the entire receiver is dominated by the LNA). Furthermore, fully differential operation is required to minimize the second-order distortion. Hence, a double-balanced topology based on a Gilbert-cell multiplier, presented in Figure 102, is commonly adopted for mixers for MMW applications [93]-[98]. The RF input signal in a voltage domain is converted to a current domain by a transconductance stage (Q1 and Q2). A switching pairs (Q3 through Q6) that

![Figure 102. Double-balanced mixer topology.](image-url)
are excited by a high LO swing perform a signal multiplication between the RF signal and the LO signal. This multiplication of two signals results in a frequency translation of the RF signal to an intermediate frequency (IF) band. If the frequency of the LO signal is chosen to be same (or sufficiently close) as the frequency of the RF signal, the IF signal will be at DC or a very low frequency band. The IF signal is finally converted to a voltage domain by load resistors. The mixer operation is graphically illustrated in Figure 103. The RF and LO signals can be expressed as

\[
V_{RF}(t) \quad V_{RF}(\omega) \\
V_{LO}(t) \quad V_{LO}(\omega) \\
V_{IF}(t) \quad V_{IF}(\omega)
\]

Figure 103. Signal waveforms. (a) RF signal. (b) LO signal. (c) IF signal.
\[ V_{RF}(t) = V_1 \cos(\omega_{RF} t) \]  

\[ V_{LO}(t) = V_2 \sum_{n=1,\text{odd}}^{\infty} A_n \cos(n\omega_{LO} t), \]  

where \( A_n \) represents coefficients for voltage amplitudes at odd-harmonics of the LO frequency. Thus, converted IF signal which is multiplication of the RF and the LO signals is expressed as

\[ V_{IF}(t) = k \sum_{n=1}^{\infty} A_n V_{RF} \cos(\omega_{RF} t) \cos(\omega_{LO} t) \]

\[ = k \sum_{n=1}^{\infty} A_n V_{RF} \frac{1}{2} \left( \cos\left( (n\omega_{LO} + \omega_{RF}) t \right) + \cos\left( (n\omega_{LO} - \omega_{RF}) t \right) \right), \]  

and the voltage conversion gain is given by

\[ G_c = \frac{kA_m}{V_i} = \frac{2}{\pi} g_m R, \]  

where \( g_m \) is the transconductance of a gain stage (Q_1 and Q_2) and \( R \) is the load resistance. This is the upper limit of the voltage conversion gain when the LO swing is high enough to fully switch Q_3 through Q_6 without causing current leakage by simultaneously turned-on the switching pair (i.e., if two transistors in the switching pair are both turned on for a short period of time, converted RF current is wasted by common-mode amplification by a switching differential pair). Therefore, high LO signal swing is desired for a maximum voltage conversion gain as well as minimum noise contribution from the switching transistors.

### 5.5.2. Circuit design

As shown in the circuit schematic in Figure 104, the Gilbert-cell-based double-balanced mixer mainly consists of RF transconductance transistors (Q_1 and Q_2) and LO
switching transistors (Q3 through Q6). Q1 and Q2 dominantly determine overall circuit performances. That is, the linearity of the mixer heavily depends on these transistors (i.e., it is improved as the size of Q1 and Q2 increases with constant collector current density). However, the shot noise of these transistors, which is the main source of noise in the mixer, is also increased as their size and resulting DC current density are increased. Therefore, the device size and the bias current level of Q1 and Q2 should be carefully optimized so that they have maximum transconductance as well as a reasonable noise factor and linearity at the operation point. We chose a 12-µm emitter length (6 µm × 2) and 5.7-mA bias current for this design. This is equivalent to $J_C$ of 3.96 mA/µm$^2$ and the transconductance stage has maximum available gain ($G_{max}$) of 16 dB and $NF_{min}$ of 5.7 dB at 94 GHz. After selecting the bias current level, sizing of the Q3 to Q6 and load resistors ($R_I$) is followed. Since immediate and complete switching maximizes the conversion gain and minimizes the noise figure, the switching transistors must be properly sized and biased at a peak $f_T$ current density level. While small transistors having large input

Figure 104. Schematic of designed W-band mixer.
impedance are desirable for a large LO swing, a large base resistance of excessively small transistors would contribute more noise to the output. By compromising the conversion gain and the noise figure, we chose 2-µm emitter lengths for Q3 to Q6, which results in \( J_C \) of 9.6 mA/µm\(^2\) for the switching transistors. The transistors biased at selected \( J_C \) achieves maximum \( f_T \). The value of \( R_1 \) was selected to maximize the voltage swing at the output node maintaining the bandwidth higher than IF of 500 MHz. Total extracted capacitance at the collector node of switching transistors was 170 fF, and the load resistor values was chosen to be 75 Ω. The buffer, consisting of an emitter follower and a differential amplifier was also included. The emitter follower acts as a level shifter and has small size (1.2 µm) to minimize the loading effect. The differential amplifier was designed to have unity voltage gain with 50-Ω loads and not to degrade linearity of the mixer core. The emitter length of Q9 and Q10 is 8 µm and the value of \( R_2 \) is 300 Ω. The two-stage buffer draws 16.6 mA from a 3.3 V supply. Matching networks for the RF and

Figure 105. Design flow for the W-band mixer.
the LO inputs were realized with the hybrid-type transmission lines (W=5 \, \mu m and \, S= 20 \, \mu m) in combination with designed Marchand baluns. Long metal inter-connections (about 120 \, \mu m) from the mixer core to input baluns were also modeled using the transmission lines (M_1 and M_2). A design flow of the W-band mixer is summarized in Figure 105.

5.5.3. Measurement results

The micro-photo of the mixer with input baluns is shown in Figure 106. All measurements were done on-wafer using GSG and GSSG probes. To measure the conversion gain, we used two W-band source modules (Agilent 83558A and OML S10MS) for RF and LO signal generations. An IF output port delivers down-converted signal to a spectrum analyzer through a DC blocking capacitor. The single-ended output

![Figure 106. Micro-photograph of designed W-band mixer.](image)
Figure 107. Measurement results of designed W-band mixer. (a) Conversion gain, noise figure, and port-to-port isolations. (b) $P_{1dB_{in}}$ with LO power of 0 dBm.
was measured and adjusted by 3 dB to obtain the differential output power level. Figure 107 presents measurement results of designed W-band mixer. The measured conversion gain varies from 14.5 dB to 17.2 dB depending on operating frequencies. The low-side LO injection was used for the down-conversion, and the IF was fixed at 500 MHz for all measurements. Measured conversion gain, SSB noise, port-to-port isolations, and input-referred 1-dB gain compression points (P_{1dB, in}) at three different RF frequencies are also plotted in Figure 107. The measured P_{1dB, in} at three different input frequencies are -15, -16, and -18 dBm at 77, 94, and 110 GHz, respectively. At the LO power level of 3 dBm, the conversion gains start to saturate at all three input frequencies. For the noise figure measurement, we used a combination of a W-band noise source (QNS-FB12LW), a wave-guide mixer, and external amplifiers for down-conversion of the noise. The measured minimum and maximum SSB noise figures are 17.4 dB and 19.5 dB, respectively. The port-to-port isolations were measured using a 110 GHz vector network.

### Table 9. Performance comparison of W-band down-conversion mixers in SiGe technologies.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>[93]</th>
<th>[94]</th>
<th>[95]</th>
<th>[96]</th>
<th>[97]</th>
<th>[98]</th>
<th>[98]</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq (GHz)</td>
<td>65-90.8</td>
<td>75-78</td>
<td>75-79</td>
<td>75-85</td>
<td>75-79</td>
<td>70-80</td>
<td>89-97</td>
<td>75-110</td>
</tr>
<tr>
<td>CG (dB)</td>
<td>&gt; 24</td>
<td>&gt; 10</td>
<td>&gt; 11</td>
<td>&gt; 15</td>
<td>&gt; 13.5</td>
<td>20</td>
<td>15</td>
<td>&gt; 12</td>
</tr>
<tr>
<td>NF (dB)</td>
<td>&lt; 18</td>
<td>18.4</td>
<td>&lt; 16.5</td>
<td>&lt; 16.5</td>
<td>&lt; 12.5</td>
<td>&lt; 13.4</td>
<td>&lt; 16.6</td>
<td>&lt; 19.5</td>
</tr>
<tr>
<td>P_{1dB, in} (dBm)</td>
<td>-30</td>
<td>-12</td>
<td>-0.3</td>
<td>-3</td>
<td>2.5</td>
<td>-14.7</td>
<td>-10.8</td>
<td>-15</td>
</tr>
<tr>
<td>DC power</td>
<td>60mA, -5V</td>
<td>39mA, 4.5V</td>
<td>75mA, 5.5V</td>
<td>34mA, 5.5V</td>
<td>61mA, 5.5V</td>
<td>120mA, 3V</td>
<td>120mA, 3V</td>
<td>28mA, 3.3V</td>
</tr>
<tr>
<td>Size (mm²)</td>
<td>0.247</td>
<td>0.275</td>
<td>0.63</td>
<td>0.533</td>
<td>0.679</td>
<td>3.63</td>
<td>3.63</td>
<td>0.81</td>
</tr>
</tbody>
</table>
analyzer (VNA). The LO-to-RF isolation, which is critical in homodyne receiver architectures, is better than 30 dB across the entire W-band. The mixer consumes 28 mA from a 3.3 V supply. The size of fabricated chip is 0.81 mm². The performance of designed W-band mixer is compared with other SiGe W-band mixers in Table 9.

5.6. Design of an Integrated W-Band Receiver

5.6.1. Receiver structure

Individual building blocks (the LNA, the Marchand balun as RF and LO inputs, and the mixer) were integrated to implement a W-band receiver chain. The diagram of the receiver and micro-photograph of the fabricated chip are presented in Figure 108. The LNA features two-stage cascode topologies with single-ended input and output. The balun integrated with input matching network of the mixer converts the single-ended input signal to differential signal. Another balun at the LO input also converts the single-ended LO signal from the external signal generator to the differential signal to perform down-conversion in the mixer. In order to minimize unwanted feed-through problems (or maximize the isolation), substrate contacts were densely placed surrounding critical active core blocks and interconnecting transmission lines to tie the substrate to ground and effectively reduce signal couplings. The output matching network of the LNA and the input matching network of the mixer/balun combination were modified to consider actual input/output impedances of preceding/following blocks. (i.e., a stand-alone LNA and a mixer had been matched to 50 Ω).
5.6.2. Measurement results

The measurement for the W-Band circuits is also intricate work. The limitation in the highest available frequency of signal sources necessitates the incorporation of other external measurement devices such as up-conversion harmonic mixers, wave-guide...
Figure 109. Measurement setup for the W-band receiver. (a) For a conversion gain. (b) For a noise figure.
mixers, external amplifiers, etc. The measurement setups for a conversion gain and a noise figure are presented in Figure 109. On-wafer measurements using GSG probes for RF and LO input were performed. A GSSG probe was used for the IF output. To generate 85 to 100 GHz external signals, two MMW source modules, the OML S10MS and the Agilent 83558A, were used in combination with Hewlett Packard 20 GHz signal generators (83622B). Using the Agilent W-band power sensor (W8486A) and the power meter (E4419B) a reference power level table was made through the frequency range of interest to de-embed all possible losses due to RF cables and wave-guide adapters. The minimum power level that can be generated from the signal source was around -60 dBm. The RF power level of -50 dBm was chosen to observe the variation of the conversion gain while sweeping RF input frequency or applied LO power level. In all cases, the IF output frequency was fixed at 500 MHz. The single-side band (SSB) noise figure was measured using the Quinstar Technology W-band noise source (QNS-FB12LW) and the Agilent noise figure analyzer (N8972A). Careful calibration was performed with the W-band noise source in combination with a wave-guide mixer (QMB-FBFBWS), an LNA (QLW-75B05015-I1) and a PA (QPW-75B01315-I1) before measuring the actual noise figure.

The measured performances of the W-band receiver are presented in Figure 110 and 111 with comparison to simulation results. A maximum conversion gain of 36.3 dB was achieved with 2-dBm LO power level at 91 GHz. In the frequency range of 87 to 93 GHz, the conversion gain is greater than 32 dB. The measured minimum SSB noise figure is 10 dB at 92 GHz. From 85 to 96 GHz, the SSB noise figure is less than 15 dB. The peak-gain frequency was shifted by 3 GHz compared to the simulation result in
Figure 110. Measurement results of designed W-band receiver. (a) Conversion gain vs RF input frequency. (b) Noise figure vs RF input frequency.
which the transmission line models in the design-kit (transmission lines with smaller characteristic impedance) are used. With input/output matching networks in which EM model of transmission lines are used, the conversion gain versus frequency plot is shifted down by 3 dB, which is coincident with measurement result. The measured $P_{1dB_{in}}$ is -36 dBm. The RF frequency and LO power were fixed at 91 GHz and 2 dBm, respectively. The S-parameters were measured from 20 to 110 GHz using Agilent 8510C VNA. The return losses for RF and LO ports are better than 10 dB in the range of 87 to 94 GHz.

Figure 111. Measurement results of designed W-band receiver. (a) Conversion gain and noise figure vs LO input power. (b) IF output power vs RF input power. (c) RF and LO input matching. (d) Port-to-port isolations.
The RF to LO and the LO to RF isolation are better than 28 dB. The integrated W-band receiver performance is summarized in Table 10 with comparison to other SiGe W-band receivers.

5.7. Conclusion

The issues and design methodologies for W-band receiver building blocks: a balun, an LNA, a down-conversion mixer were presented in detail. Individual circuit components implemented using SiGe BiCMOS technology with $f_T$ of 200 GHz. The planar-type Marchand balun was designed using hybrid-type coupled transmission lines achieving -8 dB and -9.5 dB for $S_{21}$ and $S_{31}$, respectively. The two-stage LNA with cascode structure was designed showing a gain of 26 dB and a noise figure of 8.4 dB at 96 GHz. The double-balanced down-conversion mixer was also implemented with integration of the Marchand balun for RF and LO inputs. The fabricated stand-alone mixer achieved a conversion gain of greater than 12 dB and a SSB noise figure of less than 19.5 dB across the entire W-band (75 to 110 GHz) demonstrating very wideband performance. Finally, integrated W-band receiver chain was implemented. Achieved maximum conversion gain is 36.3 dB (at 91 GHz) and minimum SSB noise figure is 10 dB (at 92 GHz).
Table 10. Performance comparison of W-band receivers in SiGe technologies.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>[35]</th>
<th>[36]</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq (GHz)</td>
<td>77</td>
<td>77</td>
<td>91</td>
</tr>
<tr>
<td>CG (dB)</td>
<td>30</td>
<td>40</td>
<td>36.3</td>
</tr>
<tr>
<td>NF (dB)</td>
<td>11.5</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>$P_{1dB_{in}}$ (dBm)</td>
<td>-26</td>
<td>-38</td>
<td>-36</td>
</tr>
<tr>
<td>$IIP_3$ (dBm)</td>
<td>-21.6</td>
<td>-26</td>
<td>-26</td>
</tr>
<tr>
<td>DC power (mW)</td>
<td>440</td>
<td>195</td>
<td>109.7</td>
</tr>
</tbody>
</table>
CHAPTER 7
CONCLUSIONS AND FUTURE WORKS

7.1. Contributions to the Field

In this dissertation, various approaches with novel design techniques were employed to implement high-performance transceiver front-end ICs operating at RF (2.4 GHz) and MMW (90 to 94 GHz) bands using silicon-based technologies. In Chapter 2, fundamental limitations of silicon-based technology for high-frequency IC designs were reviewed. Implantation issues for RF PAs using CMOS technology were explored in Chapter 3 and 4. Design methodologies and techniques to develop MMW receiver front-end ICs using SiGe technology were discussed in Chapter 5.

To meet increasing demand from modern wireless transceiver industry, adaptation of multi-standards and multi-modes operation schemes to PAs is strongly desired. Inherent difficulties of implementing multi-standards multi-mode PAs for high-data rate applications using CMOS technologies includes low power-efficiency, poor linearity, and limited output power capability. Effective techniques to improve the performance of linear CMOS PA in terms of efficiency optimization for multi-mode operation and maximization of the output power capability were proposed and demonstrated. Employed techniques are:

1. A novel discrete resizing and concurrent power combining structure for minimization of reverse-induced current leakage through disabled combining transformer inputs.
2. A varactor-based tunable matching network to implement the load impedance optimization.
3. A novel parallel-series combining transformer (PSCT) as an output power combiner which overcomes limitations of conventional combining transformers.
such as the parallel-combining transformer (PCT) or the series-combining transformer (SCT).

Theoretical analysis and EM simulations were performed to investigate proposed passive structures verifying the effectiveness of them in applications of high-end CMOS PA designs. With these achievements, potential appropriateness of CMOS technologies to develop RF PAs for advanced wireless communications standards such as WiMAX and LTE have been demonstrated, which is the biggest contribution to the related field.

W-band receiver building blocks (an LNA, a balun, and a mixer) and integrated receiver were implemented using SiGe BiCMOS technology with $f_T$ of 200 GHz. Optimization of the circuit performance using well-defined design methodologies were thoroughly performed to demonstrate the usability of SiGe technology for W-band applications (90 to 94 GHz). The operating frequency of designed circuits is almost half of $f_T$ of the process technology. The implementation of receiver building blocks in such a high frequency is an amalgamation of careful EM simulations, parasitic extractions, circuit simulations, and layout optimizations. Given limited design resources including accurate models for active and passive devices, first-time success of the circuit demonstration is not easy. However, implemented circuit blocks and integrated receiver achieved good performances which well coincides with simulation results, verifying effectiveness of proposed structure and design methodology. The performance of the integrated W-band receiver showed highest conversion gain with reasonable SSB noise figure among reported W-band receivers operating beyond 90 GHz.

### 7.2. Future Research Directions

Although a part of this research focused on efficiency and power improvement techniques for RF CMOS PAs, linearity performance of CMOS PAs is also important for advanced wireless communications. To implement CMOS PAs that completely satisfy
specifications of high data-rate wireless communication standards, sophisticated design techniques with a device/structure view point must be incorporated to improve the linearity performance as well. In addition, the load-sensitivity problem of CMOS PAs is another issue to be overcome as a future research.

For the part of the research on the W-band SiGe IC designs, more accurate modeling of passive devices such as transmission lines must be performed. In addition, various topologies and techniques that have been utilized for applications in RF bands can be employed to enhance the linearity and noise performance of the receiver.
PUBLICATIONS


quadrature coupler,” submitted to Solid-State Circuits, IEEE Journal of (under review)

REFERENCES


VITA
JIHWAN KIM

Jihwan Kim was born in Suwon, Korea in 1978. He received his B.S. degree in electrical and computer engineering from Hanyang University, Seoul, Korea in 2005. In 2007, he received his M.S. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, where he is currently pursuing his Ph.D. degree. His major research interests are designing RF and mm-wave front-end integrated circuits including power amplifiers (PAs), mixers, low-noise amplifiers (LNAs), and voltage-controlled oscillators (VCOs) using CMOS/SiGe technologies.

He has worked as an intern at Samsung Design Center, Atlanta, GA, in 2008 designing CMOS LNAs for GSM/GPRS applications and at RF Micro Devices (RFMD), Torrance, CA, in 2009 designing high performance GaAs pHEMT LNAs and PAs for millimeter-wave applications.