CMOS Radio-Frequency Power Amplifiers

For Multi-Standard Wireless Communications

A Dissertation
Presented to
The Academic Faculty

by

Hyungwook 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 © Hyungwook Kim 2011
CMOS Radio-Frequency Power Amplifiers

For Multi-Standard Wireless Communications

Approved by:

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

Dr. Saibal Mukhopadhyay
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Jongman Kim
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Jun-Hee Heu
School of Electrical Engineering and Computer Science
Seoul National University

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

Date Approved: May 18, 2011
To my wife and parents
ACKNOWLEDGEMENTS

First of all, I would like to express my sincere appreciation to my advisor, Professor Emmanouil M. Tentzeris, for monumental support and enthusiastic supervision. Without his guidance and inspiration, I would not have achieved this research goal. I am also grateful to all the committee members, Professor Jongman Kim, Professor Chang-Ho Lee, Professor Saibal Mukhopadhyay, and Dr. Jun-Hee Heu for their time and effort in reviewing my dissertation and serving as my defense committee members. I especially thank to Dr. Chang-Ho Lee for all the opportunities, guidance, and encouragement that he has provided, and I would like to specially express my gratitude to Dr. Joy Laskar for his great support and guidance throughout this research.

I would also like to thank Dr. Kyutae Lim, Dr. Jaejoon Chang, Dr. Woonyun Kim, Dr. Changhyuk Cho, Dr. Wangmyong Woo, Dr. Yunseo Park, Dr. Ki Seok Yang, Dr. Minsik Ahn, Dr. Jeonghu Han, Dr. Dong-Ho Lee, Dr. Jeongwon Cha, and Dr. Jaejoon Kim and for their technical guidance and assistance.

Much appreciation is extended to my colleagues and friends at Georgia Institute of Technology. I am indebted my colleague members in Microwave Application Group for their invaluable support. I would like to express my appreciation to Dr. Franklin Bien, Dr. Jinsung Park, Dr. Taejoong Song, Dr. Kyu Hwan An, Dr. Ockgoo Lee, Dr. Jongmin Park, Dr. Kwan-Woo Kim, Dr. Hyoungsoo Kim, Dr. Jaehyouk Choi, Dr. Sang Min Lee, Dr. Sanghyun Woo, Dr. Joonhoi Hur, and Dr. Kilhoon Lee for their assistance and cooperation. I also would like to express my thanks to Mr. Youngchang Yoon, Mr. Jihwan Kim, Mr. Kun Seok Lee, Mr. Eungjung Kim, Mr. Kwanyeob Chae, Mr. Hamhee
Jeon, Mr. Hyun-Woong Kim, Mr. Yan-Yu Huang, Mr. Michael Oakely, Mr. Seungil Yoon, Mr. Seungho Baek, and Mr. Stephen Kim for their assistance and cooperation. I am especially grateful to Dr. Kyu Hwan An, Dr. Ockgoo Lee, Mr. Youngchang Yoon, and Mr. Jihwan Kim for their insightful discussion and assistance.

And most of all, I am especially grateful to my parents, Jong-Deuk Kim and Keum-Ran Park, and my parents-in-law, Jong-Chul Lim and Kyung-Ja Ryu, for their unconditional love. I can’t express my love and gratitude enough to my wife, Namhee Lim, for her endless support throughout my life. Without their encouragement and support, I would not be able to complete this work.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ............................................................................................... iv

LIST OF TABLES .............................................................................................................. x

LIST OF FIGURES .......................................................................................................... xi

LIST OF ABBREVIATIONS ............................................................................................ xvii

SUMMARY ......................................................................................................................... xix

CHAPTER 1 INTRODUCTION ......................................................................................... 1

1.1. Backgrounds and Motivations .............................................................................. 1

1.2. Challenges .............................................................................................................. 6

1.2.1. Reliability issues of CMOS technology due to the low breakdown voltage . . 6

1.2.2. Lossy silicon substrate in CMOS technology ................................................. 8

1.3. Organization of the Thesis .................................................................................. 13

CHAPTER 2 MULTI-STANDARD CMOS RADIO-FREQUENCY PAs ........... 15

2.1. Introduction ......................................................................................................... 15

2.2. CMOS PA with Polar Modulation ...................................................................... 16
2.3. CMOS PA with the Doherty Configuration................................. 18

CHAPTER 3 DESIGN OF CMOS POWER CELLS WITH CASCODE TOPOLOGY

................................................................. 22

3.1 Introduction................................................................. 22

3.2. Power Cells with CMOS Cascode Topology.................... 23

3.2.1. DC characteristics according to the cascode topology .... 27

3.2.2. RF power and efficiency characteristics according to the cascode topology ................................................................. 31

3.2.3. Linearity characteristics according to the cascode topology ... 37

3.3. Optimization of Device Sizing and Biasing Conditions .......... 41

3.4. Conclusion ................................................................ 49

CHAPTER 4 DESIGN OF GSM/EDGE DUAL-MODE POWER AMPLIFIERS.. 51

4.1. Introduction................................................................. 51

4.2. Dual-Mode Operation and Matching Network Transition........ 51

4.3. Design of Sub-Blocks for a Watt-Level PA ...................... 54

4.3.1. Differential Cascode topology PA core with cross-coupled capacitors...... 55

4.3.2. Optimum load impedances for GSM and the EDGE applications ...... 58

4.3.3. CMOS on-chip transformer ............................................. 61

4.4. Conclusion ................................................................. 69
CHAPTER 5 IMPLEMENTATION AND MEASUREMENT OF DUAL-MODE PA 71

5.1. Implementation of Single Stage GSM/EDGE Dual Mode PA ......................... 71

5.2. Measurement Results of Single Stage GSM/EDGE Dual Mode PA ............. 73
  5.2.1. Measurement setup ...................................................................................... 73
  5.2.2. PA performance in the GSM application ..................................................... 73
  5.2.3. PA performance in the EDGE application ................................................... 77

5.3. Design Methodology of Two-stage Dual-Mode PA ........................................... 82

5.4. Implementation and Measurement of Two-Stage GSM/EDGE PA ............... 89

5.5. Linearity Improvement with Pre-distortion ...................................................... 94

5.6. Conclusion ....................................................................................................... 97

CHAPTER 6 A TRIPLE-MODE POWER AMPLIFIER FOR

GSM/EDGE/WCDMA APPLICATIONS ................................................. 99

6.1 Introduction ..................................................................................................... 99

6.2. Design of Multi-mode PA with Body-switched PA Core ............................. 100
  6.2.1. Body-switched PA core ............................................................................. 101
  6.2.2. Comparison of PA cores with direct connection and body switches ....... 104

6.3. Implementation and Measurements of GSM/EDGE/WCDMA PA .............. 108
  6.3.1. GSM performance of body-switched PA ..................................................... 108
  6.3.2. Verification of the body-switch PA core .................................................... 110
6.3.3. Body-switched PA in the EDGE application............................................. 115

6.3.4. Body-switched PA in the WCDMA application........................................ 115

6.4. Conclusion ........................................................................................................ 117

CHAPTER 7 CONCLUSION AND FUTURE WORK........................................ 119

7.1. Technical Contributions and Impacts of the Dissertation......................... 119

7.2. Scope of Future Research ............................................................................. 122

REFERENCES .................................................................................................. 124

VITA .............................................................................................................. 130
LIST OF TABLES

Table 1. Breakdown mechanisms of CMOS technology.................................................. 8

Table 2. Loss mechanism of CMOS technology.......................................................... 10

Table 3. Comparison of Pout and efficiency according to the cascode topology...... 34

Table 4. Electrical properties of the designed transformer........................................ 66

Table 5. Bias condition and designed values for the tunable capacitor. ................. 69

Table 6. Summary of 1-tone measurements in linear mode operation. .................... 80

Table 7. Designed values for the tunable capacitor for the inter-stage matching..... 89

Table 8. Bias conditions and values of tunable capacitors of two-stage dual-mode PA. ......................................................................................... 90
LIST OF FIGURES

Figure 1. Evolution of wireless communication standards ................................................................. 2
Figure 2. Prediction of global handset shipments by technology .............................................................. 3
Figure 3. Loss mechanisms in the CMOS technology ............................................................................ 12
Figure 4. Block diagram of polar modulated PA .................................................................................. 18
Figure 5. A PA with the Doherty configuration (a) Block diagram of the PA ........................................ 19
Figure 6. Two CMOS cascode types according to body connection of CG transistor. (a) BG-cascode topology. (b) BS-cascode topology ................................................................................. 25
Figure 7. Cross-sectional views of CMOS cascode with DNWs. (a) BG-cascode topology. (b) BS-cascode topology .................................................................................................................. 26
Figure 8. Measured $I_{DS}$-$V_{DS}$ characteristics for BG-cascode and BS-cascode as a function of gate voltage of CS device, $V_{G,CS}$. (a) $V_{G,CG} = 3.0$ V. (b) $V_{G,CG} = 2.4$ V. ........................................ 28
Figure 9. Measured $I_{DS}$-$V_{G,GS}$ characteristics and transconductance of BS-cascode according to the $V_{G,CG}$ at fixed $V_{DS} = 3.0$ V ........................................................................................................ 29
Figure 10. Comparison of $I_{DS}$-$V_{G,GS}$ and $g_m$ characteristics for BG-cascode and BS-cascode at $V_{DS} = 3.0$ V. (a) $V_{G,CG} = 3.0$ V. (b) $V_{G,CG} = 2.4$ V .................................................................................. 30
Figure 11. Conventional source-pull and load-pull measurement setup .............................................. 32
Figure 12. Output power and efficiency contours on Smith chart as the results of a load-pull simulation ....................................................................................................................................... 33
Figure 13. Comparison of power performances according to the input power sweep at two different $V_{G,CG}$ values. (a) $V_{G,CG} = 3.0$ V. (b) $V_{G,CG} = 2.4$ V .................................................................................. 35
Figure 14. Power performance versus gate voltage of CS device, $V_{G,CS}$, with $V_{G,CG} = 2.4$ V and input power of 0 dBm at 1.9 GHz. ................................................................. 36

Figure 15. Optimum load impedance variation as a function of $V_{G,CG}$ at fixed $V_{G,CS} = 0.6$ V ............................................................................................................................. 38

Figure 16. Comparison of linearity according to the cascode types. (a) IMD$_3$ characteristics of BG-cascode as a function of $V_{G,CG}$. (b) IMD$_3$ characteristics of BS-cascode as a function of $V_{G,CG}$. (c) output power and efficiency comparison at the C/I$_3$ of 25 dBC. .............................................................................................................. 40

Figure 17. Microphotograph of the fabricated CMOS cascode power cells .................... 42

Figure 18. Dependency on the device size and $V_{G,CG}$ : (a) Saturated $P_{out}$. (b) Saturated power added efficiency. ......................................................................................................................... 43

Figure 19. Bias dependency at optimum device size. (a) Saturated $P_{out}$. (b) Saturated power added efficiency. ......................................................................................................................... 45

Figure 20. Dependency on the device size and $V_{G,CG}$ in linear operation. (a) Maximum linear output power. (b) Maximum linear PAE. ................................................................. 46

Figure 21. C/I$_3$ versus $P_{out}$ as a function of $V_{G,CS}$. at optimum device ratio and $V_{G,CG}$..... 48

Figure 22. Performance requirements of power amplifier according to the application... 52

Figure 23. Output matching networks and output impedances of power amplifiers. (a) PA with fixed output matching network and its output impedances. (b) PA with tunable output matching network and its output impedances.................................................. 53

Figure 24. Schematic of designed power amplifier core.................................................... 56

Figure 25. AM-AM and AM-PM characteristics according to the cross-coupled capacitors. (a) Gain variation. (b) Phase variation................................................................. 57
Figure 26. Tunable matching network and optimum impedances. (a) Schematic of tunable matching network with 2×1:2 parallel combining transformer. (b) Optimum impedances for the GSM and EDGE applications at 1.8 GHz. ................................. 59

Figure 27. Cross-sectional view of the RF CMOS process. ........................................... 62

Figure 28. Parallel combining transformer(PCT). (a) Schematic of 2×1:2 parallel combining transformer. (b) Physical layout of 2×1:2 step-up transformer............. 63

Figure 29. Designed transformer. (a) Physical layout of 2×1:2 step-up transformer with vertical metal structure. (b) Simulated transformer efficiency. ......................... 65

Figure 30. High-power tunable capacitor. (a) Schematic of tunable capacitor with 4 MIM capacitors and a switch transistor. (b) Characteristics of “ON” state. (c) Characteristics of “OFF” state. ................................................................................. 67

Figure 31. Schematic of single stage dual-mode power amplifier with a tunable matching network................................................................................................................. 72

Figure 32. Microphotograph of the fully integrated single stage dual-mode power amplifier with a tunable matching network. .................................................. 72

Figure 33. Measurement test bench. ................................................................. 74

Figure 34. Single-tone measurement results for the GSM application. (a) Pout and drain efficiency as a function of frequency. (b) Pout and drain efficiency as a function of supply voltage. ...................................................................................... 75

Figure 35. PA performances with single tone as a function of input power for the GSM application.......................................................................................................... 76

Figure 36. Measured output spectrum of GMSK modulated signal and GSM spectrum emission mask at 1.76 GHz of operation frequency. ................................. 77
Figure 37. PA performances with single tone as a function of input power for the EDGE application.

Figure 38. Measured performances of EDGE-mode PAs with tunable matching network according to the output impedances.

Figure 39. Measured EVMs and efficiencies according to the output impedances with EDGE.

Figure 40. Measured output spectrum with EDGE modulated signal at 27.5 dBm of Pout and 1.76 GHz of operation frequency.

Figure 41. Schematic of fully integrated two-stage GSM/EDGE dual-mode power amplifier with tunable matching networks.

Figure 42. Simulated gain and phase variations according to the gate bias of CS device in the power stage. (a) Gain variation. (b) Phase variation.

Figure 43. Linearization of two-stage PA with pre-distortion of driver stage. (a) Gain variation. (b) Phase variation.

Figure 44. Microphotograph of the fully integrated two-stage GSM/EDGE dual-mode PA with tunable matching networks.

Figure 45. Measured saturated output power and PAE of two-stage GSM/EDGE dual-mode PA according to the frequency for the GSM application.

Figure 46. Measured gain, EVM, and PAE of two-stage GSM/EDGE dual-mode PA according to the output power for the EDGE application.

Figure 47. Output spectra of two-stage GSM/EDGE dual-mode PA. (a) Output spectrum of GSM modulation signal. (b) Output spectrum of EDGE modulation signal.
Figure 48. Comparison of EVM characteristics according to the matching networks of single and two-stage PAs. .........................................................................................................................95

Figure 49. Comparison of output spectra with EDGE modulated signal at 27.5 dBm of Pout and 1.76 GHz of operation frequency according to the matching networks of single and two-stage PAs. .........................................................................................................................96

Figure 50. Schematic of designed body-switched power amplifier core. .........................101

Figure 51. Simulated resistances and capacitances of switch as a function of total gate width........................................................................................................................................103

Figure 52. Simulated PA characteristics with conventional direct connected cascodes and body-switched cascode. (a) Comparison between conventional BS-cascode and BS-cascode with body switches. (b) Comparison between conventional BG-cascode and BG-cascode with body switches. ........................................................................................................105

Figure 53. Power characteristics of body-switch cascode PA for BS and BG connection with similar gain variation and PAE 1.8 GHz........................................................................................................106

Figure 54. Linearity comparison of body-switched cascode PA for BS- and BG-types with similar gain variation and PAE. ...........................................................................................................107

Figure 55. Microphotograph of fully integrated GSM/EDGE/WCDMA PA chip on evaluation board. ........................................................................................................................................109

Figure 56. Measured output power and PAE versus frequency of body-switched PA for the GSM application. ........................................................................................................................................110

Figure 57. Measured gain variation and PAE versus output power of body switched cascode PA with similar values. ...........................................................................................................111
Figure 58. Measured EVM of body-switched cascode PA according to the body connection.................................................................112

Figure 59. Measured IMD3 of body-switched cascode PA according to the body connection.................................................................112

Figure 60. Measured small-signal S-parameters for linear PA application. (a) $S_{11}$ and $S_{22}$. (b) $S_{21}$ and $S_{12}$.................................................................113

Figure 61. Measured gain, EVM, and PAE versus output power of BG-cascode PA with body switches for the EDGE application.................................................................114

Figure 62. Measured ACLR of body-switched cascode PA according to body connection with WCMDA modulated signal.................................116

Figure 63. Output spectra of body-switched-cascode PA according to body connection with WCDMA modulated signal.................................117
LIST OF ABBREVIATIONS

AC .......................................................... alternating current
AM-AM .................................................. amplitude-to-amplitude
AM-PM .................................................. amplitude-to-phase
BG ......................................................... body tied to ground
BJT ...................................................... bipolar junction transistor
BS ........................................................ body tied to source
CDMA .................................................... code division multiple access
CG ........................................................ common gate
CMOS ................................................... complementary metal oxide semiconductor
CS ......................................................... common source
dAT ...................................................... distributed active transformer
DC ......................................................... direct current
DE ......................................................... drain efficiency
DNW .................................................... deep n-well
DUT ..................................................... device under test
EDGE ................................................... enhanced data rates for GSM evolution
EVM ...................................................... error vector magnitude
FET ....................................................... field effect transistor
GaAs .................................................... gallium arsenide
GSM ..................................................... global system for mobile
HBT ...................................................... hetero-junction bipolar transistor
IC................................. integrated circuit
IMD................................. inter-modulation distortion
IMD3................................. third order inter-modulation distortion
LDMOS............................... laterally diffused MOS
LNA................................. low-noise amplifier
LTE................................. long term evolution
MIM................................. metal-insulator-metal
MOSFET............................. metal oxide semiconductor field effect transistor
P1dB................................. output referred 1-dB compression point
PA................................. power amplifier
PAE................................. power added efficiency
PAPR............................... peak-to-average power ratio
PCB................................. printed circuit board
PCT................................. parallel-combining transformer
Psat................................. saturated (peak) output power level
RF................................. radio frequency
SCT................................. series-combining transformer
SOG................................. silicon on glass
SOI................................. silicon on insulator
VCO................................. voltage-controlled oscillator
WCDMA............................. wideband CDMA
WiMAX............................. worldwide interoperability for microwave access
WLAN.............................. wireless local area network
SUMMARY

The development of multi-standard wireless communication systems with low cost and high integration is continuously requested and accompanied by the explosive growth of the wireless communication market. Although CMOS technology can provide most building blocks in RF transceivers, the implementation of CMOS RF power amplifiers is still a challenging task. The objective of this research is to develop design techniques to implement fully-integrated multi-mode power amplifiers using CMOS technology.

In this dissertation, a load modulation technique with tunable matching networks and a pre-distortion technique in a multi-stage PA are proposed to support multi-communication standards with a single PA. A fully-integrated dual-mode GSM/EDGE PA was designed and implemented in a 0.18 \( \mu \text{m} \) CMOS technology to achieve high output power for the GSM application and high linearity for the EDGE application. With the suggested power amplifier design techniques, fully-integrated PAs have been successfully demonstrated in GSM and EDGE applications.

In addition to the proposed techniques, a body-switched cascode PA core is also proposed to utilize a single PA in multi-mode applications without hurting the performance. With the proposed techniques, a fully-integrated multi-mode PA has been implemented in a 0.18 \( \mu \text{m} \) CMOS technology, and the power amplifier has been demonstrated successfully for GSM/EDGE/WCDMA applications.

In conclusion, the research in this dissertation provides CMOS RF power amplifier solutions for multiple standards in mobile wireless communications with low cost and high integration.
CHAPTER 1
INTRODUCTION

1.1. Backgrounds and Motivations

In modern society, mobile devices for personal communication have become one of the daily necessities. After exposure to these modern conveniences, it would not be easy to live without their benefits. However, even two decades ago, wireless communication was exclusive property for specific purposes such as military, voyage and aviation. It was unpopular and rare for the public until the mid or late 1990s. The growth of the wireless communication market has been tremendously explosive with advancements in semiconductor technology and the wireless devices industry.

In the early days of public personal wireless communication, mobile devices carried out a relatively simple “voice communication” function, which was enough to satisfy the users at that time. However, as the wireless industry grows, the consumers of mobile terminals are continuously demanding cheaper hand-held wireless devices with more advanced features including voice and video communications, multimedia functions, and even Internet access. The various demands of consumers have steered continuously wireless applications to the multiple standards which include global systems for mobile communications (GSM), code division multiple access (CDMA), enhanced data rates for GSM evolution (EDGE), wideband CDMA (WCDMA), wireless local area network (WLAN), worldwide interoperability for microwave access (WiMax), long term evolution (LTE), and so on. As shown in Figure 1, the wireless communication standards
have been evolved continuously in the direction of high data rate and high mobility to support the needs of consumers. Thus, recent wireless terminals contain two or more transceivers to support several communication standards with a single device.

Along with the evolution of wireless communication standards, the wireless communication market has grown explosively. The total worldwide subscriptions of cellular wireless communications have risen from approximately 500 million at the end of 1999 to approximately 5.2 billion at the end of 2010, and it is expected that the aggregate number of global subscriptions could exceed 6.8 billion in 2014 from the forecast reported by iSuppli Corporation. From the wireless market forecasting, the worldwide handset sales are expected to grow approximately 10% annually from 2011 through 2014 as shown in Figure 2, which shows the global handset shipments. This
The figure also shows the prediction of the market share by the air interface technology node. According to the forecasting, the wireless market will shift to the advanced 3G and 4G technologies, and the shipments of 3G/4G devices are predicted to increase to more than 50% of the market in 2013. The growth of the advanced technology does not mean the fadeaway of 2G and 2.5G technologies. It means that 3G/4G handsets have to support more communication standards including not only 3G/4G technologies but also 2G/2.5G technologies in a single device.

In addition to the multi-standard application, the manufacturers of wireless devices have been pressured to find a solution for cheaper wireless devices because of the strong demands of low-cost mobile terminals from the market. Therefore, the integration of different functional blocks, such as digital, analog, and even radio frequency (RF)
building blocks, has become an important task to realize low-cost wireless devices. Among the various process technologies, complementary metal oxide semiconductor (CMOS) technology, which has the capability of a high-level of integration and is already a low-cost, matured process technology, is the most adequate solution for RF transceivers in wireless applications.

Although CMOS technology has several advantages such as good thermal characteristics, low cost and an ability of integration superior to other technologies, it has intrinsic drawbacks for the application to large signal components such as power amplifiers (PAs) and RF switches due to poor reliability related to its low breakdown characteristics and low quality (Q) factor related to its lossy substrate. Thus, RF PAs and RF switches have been the bottleneck for fully integrated CMOS transceivers, and the implementation of those components with CMOS technology is still a very challenging task. Due to these reasons, the majority of commercial PA products adopt compound semiconductor processes such as Gallium Arsenide (GaAs) based technology, and RF switches adopt GaAs, Semiconductor-on-Glass (SOG), and Silicon-on-Insulator (SOI) technology.

In the history of the semiconductor industry, it is well known that the concept of the FET was proposed earlier than the bipolar junction transistor (BJT), however, the invention of BJT preceded that of FET [1]. As a high-quality insulator for gate oxide appeared, the FET took over leadership in the market with high integration. Due to the explosive expansion of integrated circuits (ICs) and the memory industries, there has been a tremendous and steady progress in the development of CMOS technology.
Therefore, CMOS technology has dominated in most of the digital and analog IC industries by its appealing low-cost solution.

However, it has been a different story in the RFIC industry because of high operating frequencies. For early transceivers, not only PAs but also other RF building blocks, such as low noise amplifiers (LNA), voltage controlled oscillators (VCO), and mixers had been implemented using compound semiconductor technologies such as GaAs technology. But, continuous demands of cost reduction and miniaturization from the market have made manufacturers and circuit engineers choose low-cost CMOS technology and accelerate the development of CMOS transceivers. As a result, all of the digital and analog blocks have been integrated in CMOS technology. Even some RF building blocks also have been integrated together into one chip in modern transceivers of hand-held wireless devices.

Although the trends of CMOS transceivers have been aiming for a fully integrated RF solution, also called “a single radio,” there are still many difficulties in the integration of some RF components, such as PAs and RF switches, and even in the implementation of them using CMOS technology. The limitations come from intrinsic drawbacks of CMOS technology, such as low breakdown voltage of MOSFETs [2] and low Q factor of passive devices due to the lossy silicon substrate [3].

Compared to the relatively small-signal components, which have been already implemented using CMOS and integrated with other components, PAs in hand-held devices deliver high output power at the end of a transmitter to satisfy modern communication standards with high efficiency and linearity. Watt-level PAs with low impedances force CMOS technology to handle such large voltage and current signals that
make the drawbacks of CMOS become more problematic from an RF perspective. According to these reasons, in spite of higher cost, GaAs Hetero-junction Bipolar Transistors (HBTs) have been mainstreamed in the PA market with their inherent superiorities to MOSFETs, such as higher breakdown voltage, higher current density, and better linearity.

Nevertheless, the full integration of digital, analog, and even radio frequency (RF) functions has been an inevitable task in the modern wireless communication industry. There have been great efforts to design PAs using CMOS technology to realize a true single chip. In the following section, the challenges of current CMOS technologies and some approaches of designing CMOS PAs will be discussed in detail.

1.2. Challenges

1.2.1. Reliability issues of CMOS technology due to the low breakdown voltage

The scaling of CMOS technology can give benefits such as faster switching speed, lower energy per switching, and higher integration for digital and analog integrated circuitry. As CMOS technology scales down to sub-μm gate length, the typical supply voltage is also scaled down. However, for PA products for mobile terminals, watt-level output power is required with low supply voltage. In order to achieve watt-level output power with low supply voltage, the required small load impedance can make the amplitudes of voltage and current swings easily exceed normal operating range of CMOS transistors.

High voltage stress on the devices can cause reliability problems including device breakdown of transistor. According to the causes and physical phenomena, the
breakdown mechanisms of FETs can be divided into several categories, such as junction breakdown, channel breakdown, hot carrier degradation, and oxide breakdown [4, 5]. Among the breakdown mechanisms, some are nondestructive phenomena, and the device will operate normally without excess voltages. However, some are destructive processes that degrade device characteristics severely.

Junction breakdown occurs at the p-n junctions between the substrate and the drain or source regions. If the reverse bias applied to them exceeds certain values, the junctions conduct large currents. However, junction breakdown is a nondestructive phenomenon, and the device will function properly if the large voltage is removed. While the device is on, the electron-hole pairs with high kinetic energy cause channel breakdown and hot carrier degradation near the drain region where the electric field is high. Due to impact ionization of carriers with high energy, excessive currents flow through the channel and cause channel breakdown, which is nondestructive. However, these electrons and holes can cause hot-carrier degradation, which is a destructive process. The velocities of carriers with high energy, i.e., hot carriers, can be sufficient to penetrate carriers into the oxide under a high electric field. The accumulated carriers in the oxide change the amount of oxide charges and cause variation of the threshold voltage of the device. An excessive voltage applied at the gate insulator makes the electric field exceed the dielectric strength of the gate oxide, the so-called gate-oxide breakdown, which is a destructive process. After oxide breakdown, the gate and channel have a resistive connection, and the FET will lose its original characteristics. The breakdown mechanisms of an FET are summarized in Table 1.
These breakdown processes are caused by large voltages between the terminals of the FET. The gate-oxide breakdown is mainly related to the voltage between gate and drain ($V_{DG}$), and the channel breakdown and hot carrier degradation are related to the voltage between drain and source ($V_{DS}$). The excessive voltage between drain and substrate ($V_{DB}$) causes junction breakdown. In order to ensure safe operation, the voltage swings between terminals should not exceed twice the nominal supply voltage [2, 6].

<table>
<thead>
<tr>
<th>Voltage</th>
<th>Breakdown</th>
<th>Cause and effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>drain-gate ($V_{DG}$)</td>
<td>Oxide Breakdown</td>
<td>▪ Excessive vertical electric field on the gate oxide</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▪ Gate leakage current due to gate channel linkage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▪ Destructive process</td>
</tr>
<tr>
<td>drain-source ($V_{DS}$)</td>
<td>Channel Breakdown</td>
<td>▪ Impact ionization in the channel due to excessive lateral electric field</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▪ Gate leakage current due to gate channel shortage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▪ Not inherently destructive process</td>
</tr>
<tr>
<td></td>
<td>Hot Carriers degradation</td>
<td>▪ Carriers injection into the gate oxide</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▪ Variation of $V_{th}$ due to the oxide charge</td>
</tr>
<tr>
<td></td>
<td>Junction Breakdown</td>
<td>▪ Avalanche breakdown in pn junction due to the excessive reverse bias in pn diode</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▪ pn diode breakdown due to the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▪ Not inherently destructive process</td>
</tr>
</tbody>
</table>

1.2.2. Lossy silicon substrate in CMOS technology

Due to the lossy Si substrate, CMOS technology has two major problems. One is coupling through the conductive substrate, and the other is the low Q-factor of passive components. Substrate coupling degrades the isolation characteristics of the devices so that it becomes one of the biggest obstacles to implementing a single radio (including the
PA) with CMOS technology. The lower Q-factor of passive devices means larger loss due to the effective resistive term in the passive components. Because the output matching of CMOS PAs usually exploits transformers and/or inductors, the losses of those components are directly related to the output power and efficiency of PAs. Thus, high Q-factors in passive components are necessary for a PA with high performance.

In CMOS technology, active devices such as MOSFETs have a lateral structure and share the Si substrate. Due to the sharing of the substrate, a parasitic circuit that is composed of parasitic NPN and PNP bipolar transistors in the substrate exists. This parasitic circuit can make a positive feedback loop and create a low impedance path between the power supply rails (e.g. Vdd and GND). This phenomenon is called latch-up, and it can lead to the destruction of CMOS ICs due to overcurrent. Latch-up of CMOS devices is very closely related with the resistivity of the substrate. High substrate resistance, due to the high resistivity of the substrate, triggers the parasitic substrate circuits to disrupt proper functioning of ICs [7]. According to this reason, low-cost conventional CMOS technologies use bulk Si substrates with resistivity of around 10 Ohm-cm, instead of highly resistive substrates.

By exploiting the low-resistive Si substrate in CMOS technology, a low impedance path can be created, and it couples signals from one node to another. The signal transferring through the conductive substrate is called substrate noise coupling, or substrate coupling. The signal isolations between functional blocks become one of the challenging tasks for the full integration of an RF transceiver. When a watt-level CMOS PA is integrated in a CMOS transceiver, the large signal generated by the PA can saturate functional blocks in the receiver such as the LNA due to substrate coupling. Although
some methods, such as guard rings [7] and triple-well processes with a deep n-well [8, 9], are used in RF CMOS processes to suppress cross-talk, it is still not easy to eliminate the substrate coupling perfectly.

The Q factors of integrated passive components such as capacitors, inductors, and transformers are strongly affected by the material properties of metallization and the substrate used to build the ICs. Table 2 shows the different loss mechanisms in CMOS technology with a conductive substrate. The loss mechanisms can be categorized as metal losses and substrate losses, according to the source of loss.

<table>
<thead>
<tr>
<th>Loss mechanism</th>
<th>Causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ohmic loss in metal</td>
<td>▪ DC resistance due to finite metal conductivity</td>
</tr>
<tr>
<td></td>
<td>▪ AC resistance due to skin effect &amp; proximity effect</td>
</tr>
<tr>
<td>Substrate loss</td>
<td>▪ Displacement current due to electric coupling</td>
</tr>
<tr>
<td></td>
<td>▪ Magnetic induced current due to magnetic coupling</td>
</tr>
</tbody>
</table>

At low frequency, the ohmic loss due to the DC resistance in the metal is the dominant loss. Since the metal layers do not have infinite conductivity, energy is lost to heat. This loss can be represented as a resistor connected a lossless passive component in series. Thus, the Q factor of passive components is reduced as the resistance is increased due to the inverse relationship between the Q factor and the series resistance as shown in equation (1).

\[ Quality \ Factor = \frac{X_s}{R_s} \]  

(1)
where $X_S$ is the value of the series reactance and $R_S$ is the series resistance. In order to reduce the ohmic loss, thick top metal with higher conductivity such as Cu metallization, instead of Al metallization, is adopted in advanced RF processes.

At higher frequency, the current distribution in metal layers is not uniform anymore due to the Eddy currents because of the self- and neighboring magnetic fields, which is known as skin effect and proximity effect, respectively. When AC current is flowing in the conductors, time-varying magnetic fields are induced. The induced magnetic fields penetrate the conductors. The induced magnetic fields generate the electric fields that impede the current flowing on the inner side of conductors and reinforce the current flowing on the outer side of conductors due to the circulating Eddy currents. The current density near the surface of conductor is greater than that near the core of conductor. Therefore, the effective area for current flowing is reduced. The effective resistance of the conductor increases and the Q factor is degraded due to the skin effect. Hence more energy is dissipated into heat. As the frequency increases, the skin effects are intensified in the conductor.

The proximity effects due to the neighbor-magnetic field from adjacent conductors also significantly affect the AC resistance. The AC resistance due to the proximity effects can be increased or decreased according to the direction of the induced neighboring magnetic field. In parallel conductors, when their currents flow in the same direction, as in a spiral inductor, the AC resistance due to the proximity effects is increased. When the currents flow in opposing directions, as in a transformer, the AC resistance due to the proximity effects is decreased [10].
In addition to the losses due to the conductors, the losses due the conductive Si substrate also reduces the Q factor of passive elements in CMOS technology. At high frequency, the conductive Si substrate is one of the greater sources of loss, and the loss mechanisms are different according to the sources of the substrate conducting [11].

The loss mechanisms due to the conductive substrate are also shown in Figure 3. For example, electric coupling exists between metal layers and the conductive substrate, and currents are induced in the substrate due to the magnetic fields penetrated into the conductive substrate. The electric energy is transferred to the substrate by electrically induced conduction. This is called electric or capacitive coupling. The electric coupling induces a displacement current in the conductive substrate by coupling electric energy of

Figure 3. Loss mechanisms in the CMOS technology.
the conductor to the substrate through the dielectric material between them. The other mechanism of loss in the substrate is similar to the magnetic coupling between the conductors. An AC current flowing in the conductors generates time-varying magnetic fields that penetrate into the substrate. Because the substrate is conductive, the time varying magnetic fields induce Eddy currents inside the substrate by Lentz’s Law. The induced current flows in a direction that can reduce the penetrated magnetic fields. The induced currents flow through the resistive network of the substrate, which is called a lossy substrate, and reduce the Q factor of passive components. These losses are more problematic when PAs are implemented with CMOS. Both ohmic losses in the metal and substrate losses are increased due to large currents.

However, the value of resistivity in semi-insulating GaAs substrates is very high compared to that of Si substrates in CMOS technology. Typically, semi-insulating GaAs substrates with \(10^5 \sim 10^9\) Ohm-cm of resistivity are used in GaAs HBT processes [12]. Therefore, the losses due to electromagnetic coupling in the substrate are very low in GaAs HBTs by suppressed leakage current and Eddy current in the substrate, while Si substrates have lossy substrate networks due to the conductive substrate.

### 1.3. Organization of the Thesis

Based on the aforementioned technological background and motivation, the purpose of this work is to use CMOS technologies for developing RF PAs for current and future wireless communications. In this work, the research on multi-mode RF PA with design techniques is proposed for a fully-integrated high-power-mode PA to achieve both high output power and good efficiency for constant envelope communication and for a high-
linearity-mode PA to obtain high linearity for non-constant envelope and high-data-rate communications with a single RF PA.

Chapter 1 contains an introduction to the wireless market and current trends, the requirements of RF PAs, and the motivation of this work including CMOS technology and its bottlenecks from the perspective of RF PA design is presented. To provide some background, Chapter 2 presents prior works in multi-standard CMOS PAs to achieve high efficiency and high linearity. In Chapter 3, the performance of a power cell with a CMOS cascode is analyzed. Based on this background, optimization of the power cell follows. Chapter 4 introduces the design of the building blocks for multi-mode PAs, such as a power cell with a differential cascode topology and tunable matching network. Chapter 5 presents the fully-integrated dual-mode CMOS PAs for GSM/EDGE applications. Based on the design techniques of dual-mode PAs, a fully integrated triple-mode PA for GSM/EDGE/WCDMA is introduced in Chapter 6. Finally, Chapter 7 summarizes and makes conclusions about the entirety of this dissertation, and projects research trends for the future.
CHAPTER 2
MULTI-STANDARD CMOS RADIO-FREQUENCY PAS

2.1. Introduction

Due to the inherent inferiority to GaAs technology, the development of PAs with CMOS technology has had more difficulties catching up to the performance of PAs with GaAs HBTs. Most of the efforts in the early days for the development of CMOS PAs were focused on generating enough output power to overcome the low power capability of CMOS devices. With off-chip components such as an RF choke and the off-chip inductors in matching networks, CMOS PAs have become comparable to commercial PAs from the perspective of output power [13-15]. Although watt-level CMOS PAs are available, it is not easy to implement a fully-integrated PA.

With the help of new power combining methods using transformers, such as distributed active-transformers (DATs) [16] and parallel combining transformers (PCTs) [19], CMOS PAs can be fully integrated in a single chip. At the beginning of CMOS PA development, high output power and high efficiency were mainly targeted to realize watt-level CMOS PAs. Therefore, in order to get a high efficiency, nonlinear switching PA schemes, like Class E and Class F, were adopted for applications using a constant envelope signal, especially for the GSM application.

Most of the initial efforts to realize CMOS PAs were focused on cellular applications with high efficiency and over 30 dBm of output power [17-20] and WLAN applications with sufficient linearity but moderate power levels of around 24 dBm of saturated output.
Due to the inherent nonlinearity of CMOS technology [9, 24], CMOS PAs still have poor linearity in watt-level applications with non-constant envelope signals such as EDGE, WCDMA, and WiMax. The development of CMOS PAs with high linearity and high output power is a very challenging task.

In conventional watt-level CMOS PA design, it is not easy to design a linear and efficient PA at the same time due to the direct trade-off between efficiency and linearity. This trade-off is particularly severe in CMOS technology due to the drawbacks of CMOS technology. To overcome this trade-off, some performance enhancement techniques were adopted in the CMOS PA design, such as efficiency enhancement techniques and linearization techniques. Linearization techniques attempt to improve the linearity of an efficient nonlinear PA, whereas efficiency enhancement techniques attempt to improve the efficiency of an inefficient linear PA.

For the multiple standards of wireless communication applications, there were several efforts to achieve linear PAs for high output power. One way is by applying a linearization technique, known as polar modulation, on a nonlinear switching PA [25]. Another way is by applying an efficiency enhancement technique, the Doherty amplifier technique, on a linear PA for GSM and the EDGE applications [26].

2.2. CMOS PA with Polar Modulation

Among the system level linearization techniques, a polar modulation system was successfully implemented in a CMOS process [27]. Figure 4 shows the block diagram of the polar modulated power amplifier. In this figure, the RF voltage at the digital baseband system can be expressed in cartesian coordinates by
where \( I(t) \) is the in-phase voltage, and \( Q(t) \) is the out-of-phase voltage. This RF output voltage in the cartesian representation can be rewritten in polar coordinates as followes.

\[
v_{in}(t) = I(t) \cdot \cos(\omega t) + Q(t) \cdot \sin(\omega t)
\]  

(2)

where \( A(t) \) is amplitude modulation, and \( P(t) \) is phase modulation. The constant envelope phase-modulated signal, \( \cos(\omega t + P(t)) \), is amplified by the RF PA through a constant envelope path that consists of other transmitter components like an up-converter and a PLL. The amplitude modulation signal \( A(t) \) is amplified by the low frequency amplifier through the envelope path.

The constant-envelope phase-modulated signal is amplified by a nonlinear switching PA such as a Class E amplifier for high efficiency. Although the amplitude information of the RF input signal can be lost due to the nonlinear PA in the constant envelope path, by modulating the supply voltage using the amplitude modulating signal, the information can be recovered at the RF output through the envelope path. Thus, the amplitude modulator should have sufficient linearity. However, since the envelope signal varies slowly, it is much easier for amplitude modulator to achieve enough linearity. With the mixing operation by the up-converter, the output of the PA can be represented as

\[
v_{out}(t) = M \cdot A(t) \cdot \cos(\omega_c t + P(t))
\]  

(4)

where \( M \) is voltage gain and \( \omega_c \) is the center frequency.

The major advantages of polar modulation is that the burden of linearity is shifted to the low-frequency linear amplifier, but nonlinear operation of the PA in the RF path
ensures high efficiency of the nonlinear RF amplifier. For full performance, a high level of integration is necessary, and any mismatch between signal paths should be avoided. The advantages and drawbacks are summarized as followes.

**Advantages –**

- Linearity burden on the low frequency linear amplifier.
- High efficient nonlinear PA at the RF path.

**Drawbacks –**

- Bulky system with high level of integration.
- Sensitivity to the mismatch of amplitude and phase.

### 2.3. CMOS PA with the Doherty Configuration

Although the original concept of the Doherty configuration was first proposed in 1936 [28], Doherty amplifiers have been adopted in high-power base station transmitters
for efficiency enhancement in different device technologies such as GaAs and laterally diffused MOS (LDMOS) processes. Recently, due to the stringent linearity requirement for high-data-rate wireless communications such as EDGE, WCDMA, and WiMax, most linear PAs are operated at a back-off power from their peak power. The efficiency at the backed-off output power is dropped significantly, especially for CMOS PAs, and thus the

Figure 5. A PA with the Doherty configuration (a) Block diagram of the PA. (b) Efficiency characteristics.
dropped efficiency reduces the battery lifetime of wireless handsets. In order to improve amplifier efficiency at a lower output power level instead of the peak output power, some CMOS PAs have successfully adopted the Doherty configuration for WLAN applications, which uses a relatively moderate output power level [29, 30]. However, for watt-level PAs for wireless applications, only a few fully-integrated CMOS Doherty PAs have been reported [26]. The main goal of the Doherty technique is to maintain a high efficiency over a wide range of power.

Figure 5 shows the basic block diagram of the Doherty configuration and its efficiency as a function of power. The Doherty amplifier consists of two sub-amplifiers, a main amplifier and an auxiliary amplifier, which are also called the carrier amplifier and the peaking amplifier, respectively. With different PA operating bias points, usually Class AB for the main PA and Class C for the auxiliary amplifier, only the main amplifier operates during low-power operation, and both amplifiers operate during high-power operation.

The main amplifier works alone until its output is limited by its maximum output voltage swing, achieving the first high-efficiency peak. The auxiliary amplifier starts turn on at a further increase of input power. At this time, the impedance looking from the main amplifier decreases as the current of the auxiliary amplifier increases according to the increasing input power. More current and power can be provided from main the amplifier. With the satisfaction of the proper conditions for maintaining the output voltage swing of the main amplifier, the efficiency of the main amplifier can be kept at its peak value. Once the output swing at the load reaches at its maximum, both amplifiers operate with their peak efficiency. Hence, the second peak can be obtained at the
maximum output power of the Doherty PA. High efficiency over a wide range of power levels can be achieved by using this technique, as shown in Figure 5 (b).

The overall linearity of the Doherty PA is only determined by the main PA under ideal conditions such as a lossless impedance inverter network and an infinite output impedance of the main amplifier. However, in the real case, the nonlinearity of the auxiliary PA affects the overall linearity at higher output power due to the finite Q of the impedance inverter network and the finite output impedance of the main amplifier. Therefore, at the higher output power region where the efficiency is enhanced, additional linearity improvement techniques are required to get a high linearity and enhanced efficiency at the same time. The advantages and drawbacks of the Doherty PA are summarized as following.

*Advantages* –

- High efficiency at backed-off power levels over the wide output power range.

*Drawbacks* –

- Bulky structure due to double RF paths with 90° phase shifter and impedance inverter.
- Sensitivity of efficiency enhancement function to RF signal phase mismatch.
CHAPTER 3
DESIGN OF CMOS POWER CELLS WITH CASCODE TOPOLOGY

3.1 Introduction

As mentioned in the previous chapter, a low breakdown voltage is one of the disadvantages of deep sub-micron CMOS technology towards achieving a watt-level CMOS power amplifier. In order to generate high power, large voltage swings are needed at the active devices. However, CMOS PAs may have reliability issues if too much excessive voltage is applied on the transistors. In order to alleviate these issues, two different approaches to maintain the voltage swing below the transistor breakdown limits have been proposed so far. One is reducing the voltage swing itself at the transistors, and the other is the stacking of transistors, by which the large voltage swing can be shared over the stacked transistors.

Usually, watt-level CMOS PAs have a power combining structure that combines powers from several single PAs to guarantee the output power specifications. If the required output power is larger than the power that a single PA can provide, power combining structures are necessary. Among the combining methods, series combining transformers (SCTs), such as the DAT [16, 31] or the figure-8 transformer [32, 33], at the output network of a PA can reduce the voltage swing at the transistors of single PAs. Because the SCTs have voltage boosting characteristics, the voltage swing of a load can be increased by adding the voltage swings of each of the single PAs, which are controlled
below their limits. However, for watt-level PA applications, SCTs have narrow usable ranges of turn ratios and the number of primary windings in the viewpoint of transformer efficiency [19].

Another approach to reducing the reliability issues is to spread a large voltage swing into smaller swings on stacked transistors. By stacking transistors, the voltages differences between the terminals of the transistors can be maintained below the devices’ breakdown limits. The simplest structure of stacking transistors is the cascode topology, which consists of a common gate (CG) transistor and a common source (CS) transistor. By using the cascode configuration, the large voltage swing at the output is divided into the voltages between the drains and sources, $V_{DS}$, of the two transistors. By exploiting a higher gate voltage at the CG transistor, the stress voltage between the drain and gate, $V_{DG}$, on the oxide of the CG transistor is reduced. This is the reason that the cascode configuration can alleviate the reliability issues in CMOS PAs. In order to increase the limit of the output voltage swing, a thick-gate-oxide transistor can be exploited as the CG transistor due to its higher breakdown voltage. A thin-gate-oxide transistor is used as the CS transistor for high transconductance. However, the cascode configuration has more complex DC and RF characteristics than a single transistor configuration such as a common source or a common base configuration.

### 3.2. Power Cells with CMOS Cascode Topology

Although the mechanisms of current generation are different, the terminals of a bipolar transistor and a field effect transistor are analogous from the viewpoint of their functionalities. One difference between bipolar and CMOS is that CMOS FETs have an
additional terminal, called the body terminal. According to the applied voltage between the source and body terminals, $V_{SB}$, the DC characteristics of the transistor are varied, known as the body effect or the backgate effect. Due to the body effect, the threshold voltage of the MOSFET, $V_{th}$, can be describes as

$$ V_{th} = V_{th0} + \gamma \left( \sqrt{2\phi_F + V_{SB}} - \sqrt{2\phi_F} \right) $$ (5)

where $V_{th0}$ is the threshold voltage at zero-$V_{SB}$, $\gamma$ is body effect coefficient, and $\phi_F$ is Fermi potential that is defined by the substrate doping concentration. When the voltages of the source and body terminal are not at the same potentials, the body effect exists. For example, in equation (5), the transistor has a larger value of the threshold voltage than $V_{th0}$ when $V_{SB}$ is positive. In other words, the transistor requires a larger gate voltage to turn on.

Most modern RF CMOS processes adopt deep n-wells for RF active devices to reduce the cross-talking between devices and substrate coupling [8, 9]. By using the DNWs in the implementation of FETs, the body terminals of the devices can be separated from the devices to devices. Thus, there are two options in the implementation of the cascode topology using RF CMOS process according to the connection of the body terminal of the CG transistor. For the CS transistor, typically, the body and source terminals have the same potential. Thus, the CS transistor does not have body effect. However, the body terminal of CG transistor can be connected to the ground or its source terminal (the drain terminal of the CS transistor). When it is connected to the ground, the cascode transistor has body effect due to the non-zero $V_{SB}$. On the contrary, the cascode transistor does not have body effect when the source and body of the CG transistor are tied together. According to the two different the CG transistor’s body connections, two
different cascode topologies can exist although the sizing of devices and their biasing conditions are exactly the same. For the sake of convenience in expression, the two cascode topologies can be named as “BG-cascode” and “BS-cascode”. The body of the CG transistor is tied to the ground in the BG-cascode topology, and it is tied to the source of CG transistor in the BS-cascode topology, as shown in Figure 6 (a) and (b), respectively. The cross-sectional views of the BG- and BS- casdodes are shown in Figure 7.

As mentioned above, the different connection types of the body of the CG transistor cause different threshold voltages. In the viewpoint of device technology, the change in the threshold voltage of the transistor affects almost all of the device parameters. Not only the DC characteristics, but also the small-signal characteristics of the CMOS cascode are varied due to the different body connections. These characteristics are
Figure 7. Cross-sectional views of CMOS cascode with DNWs. (a) BG-cascode topology. (b) BS-cascode topology.
especially crucial in PA design. Therefore, when the cascode configuration is applied to the CMOS power cell for the PA application, it is essential to carefully investigate PA performances such as power and linearity according to the different topologies, BG-cascode and BS-cascode. To examine the CMOS power cell with the cascode configuration, a comparison of the BG-cascode and the BS-cascode has been conducted from the viewpoint of DC characteristics and power performances, such as output power, efficiency, and linearity. With the comparison results, the optimization of the CMOS power cell in the sizing of cascode and biasing has been conducted.

3.2.1. DC characteristics according to the cascode topology

The CMOS cascode power cells have been implemented using 1P6M 0.18 μm CMOS process. The CS device used a thin-gate-oxide transistor with a gate length of 0.18 μm and the CG device used a thick-gate-oxide transistor with a gate length of 0.35 μm. In the designed power cell, the ratio between the size of the CG and CS devices is set as 4:3. The total gate widths of the CG and CS devices are 512 μm (4 x 8 μm x 16 fingers) and 384 μm (3 x 8 μm x 16 fingers), respectively.

Figure 8 shows the measured $I_{DS}$-$V_{DS}$ characteristics as a function of the gate voltage of the CS device, $V_{G,CS}$. $I_{DS}$ and $V_{DS}$ denote the current and voltage from the drain of the CG device to the source of the CS device, as shown in Figure 6. Due to the body effect of CG transistor, the BS-cascode can flow a larger current than the BG-cascode. At a low current level, the current difference between BG- and BS-cascodes is very small because the channel of the CG device is fully opened with this biasing condition. However, at a high current level, the difference becomes enlarged as the $V_{G,CS}$ increases.
Figure 8. Measured $I_{DS}$-$V_{DS}$ characteristics for BG-cascode and BS-cascode as a function of gate voltage of CS device, $V_{G,CS}$. (a) $V_{G,CG} = 3.0\,\text{V}$. (b) $V_{G,CG} = 2.4\,\text{V}$. 
At a high current level, the transconductance of the cascode decreases because the CG device cannot support the current that the CS device provided. In other words, at this biasing condition, the current is limited and determined by the CG device. Although the CS device can flow more current with high $V_{\text{G,CS}}$, the CG device is unable to flow a current beyond the limit of the CG device. Thus, $I_{DS}$ is set as the maximum current that the CG device can handle. The transconductance of the cascode, $g_m$, is degraded with a large gate bias of CG device due to this reason. The onset of decreasing of $g_m$ occurs at a higher $V_{\text{G,CS}}$ as the $V_{\text{G,CG}}$ is raised.

These phenomena can be seen more clearly with the $I_{DS}$-$V_{\text{GS}}$ characteristics, as shown in Figure 9. It shows the $I_{DS}$-$V_{\text{GS}}$ characteristics and the $g_m$ characteristics of the BS-
Figure 10. Comparison of $I_{DS}$-$V_{G,GS}$ and $g_m$ characteristics for BG-cascode and BS-cascode at $V_{DS} = 3.0$ V. (a) $V_{G,CG} = 3.0$ V. (b) $V_{G,CG} = 2.4$ V.
cascode according to the gate bias of the CG device at a fixed drain-source voltage, $V_{DS} = 3.0 \text{ V}$. At low $V_{G,CS}$ values, the current and $g_m$ are almost the same for the two cases because the devices have the same DC characteristics, including the threshold voltage, and the CG device is fully opened with the given $V_{G,CG}$. As $V_{G,CS}$ increases, the currents start to saturate at different onset voltages according to $V_{G,CG}$. The smaller the value of $V_{G,CG}$, the earlier the current starts to saturate. As a result, the $g_m$ collapse occurs at different onset voltages according to the gate voltages of the CG device. For higher $V_{G,CG}$, a larger maximum $g_m$ value and a wide flat-$g_m$ region are achieved.

The comparison of the $I_{DS}$-$V_{G,GS}$ and $g_m$ characteristics for the BG- and BS-cascodes is shown in Figure 10. For both high and low $V_{G,CG}$, at fixed $V_{DS} = 3.0 \text{ V}$, the BS-cascode shows higher current and larger $g_m$. Higher onset voltages of $I_{DS}$ saturation are achieved in the BS-cascode. As mention before, although the device size and biasing conditions are the same, the cascode shows different $g_m$ characteristics.

### 3.2.2. RF power and efficiency characteristics according to the cascode topology

The RF characteristics of a power amplifier are strongly affected by the DC characteristics of the active power cell. Previously, the DC performances of the power cell were different according to the cascode topology. In order to scrutinize the effect on the PA characteristics with regard to the cascode topology, several PA measurements have been conducted. The optimum input and output impedances of the BG- and BS-cascodes are different according to the biasing conditions due to different small-signal parameters. To operate the devices under test (DUTs), the BG- and BS-cascodes in this case, at their optimum impedances, source-pull and load-pull tests are necessary.
Figure 11 shows the diagram of a conventional source-pull and load-pull test setup. With source and load tuners, the input and output impedances can have variable values.

The RF characteristics, such as output power and efficiency, can be achieved at different input and output impedances by using these tuners. The results of the source- and load-pull tests are typically shown on a Smith chart as the contours of specific characteristics such as power and efficiency.

Figure 12 shows the simulation result of the load-pull test. In this figure, the two parameters of interest, the output power and efficiency, are shown in different contours. The contour connects the impedances which have the same value of output power or efficiency. Thus, the impedances for maximum output power and efficiency can be achieved, respectively. It is well known that the impedance points for maximum output power and efficiency are not coincident as the power increases [34]. At low power, the load line in the DC $I_{DS}$-$V_{DS}$ plane is the same for maximum power and maximum efficiency because both voltage and current swing without any clamping. However, as the
power goes up, the voltage and current swings also increase and start to be clamped at a certain power level. By increasing the current swing more, i.e. a steeper slope of the load line in the $I_{DS}$-$V_{DS}$ plane, a higher output power can be achieved, but efficiency is not its maximum due to the higher current consumption. On the other hand, maximum efficiency can be achieved by increasing the voltage swing, i.e. a gentler slope of load line in the $I_{DS}$-$V_{DS}$ plane. In Figure 12, the impedance for maximum output power is smaller than that for maximum efficiency. Therefore, this is a trade-off between maximum power point and maximum efficiency point in choosing the optimum impedance.

Figure 12. Output power and efficiency contours on Smith chart as the results of a load-pull simulation.
In order to compare the power capability and efficiency of the BG-cascode and BS-cascode, source-pull and load-pull simulations and measurements were conducted at different bias conditions. With fixed biases of $V_{DS} = 3.0\,\text{V}$, $V_{G,CS} = 0.75\,\text{V}$, and a RF input power of 6dBm at 1.9 GHz, source-pull and load-pull simulations have been conducted at the different $V_{G,CG}$. Because the impedances looking into the cascode device are varied according to the bias conditions, different optimum impedances have been achieved from the load-pull simulation. A comparison of RF power performances between the BG- and BS-cascodes is summarized in Table 3.

<table>
<thead>
<tr>
<th>Type</th>
<th>$V_{G,CG}$</th>
<th>$P_{\text{max}}$</th>
<th>$\eta_{\text{max}}$</th>
<th>$P_{\text{opt, } \eta_{\text{opt}}}$</th>
<th>$P_{\text{opt, } \eta_{\text{opt}}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG-cascode</td>
<td>3.0 V</td>
<td>20.6 dBm</td>
<td>61.7%</td>
<td>20.1 dBm, 60.1%</td>
<td>19.7 dBm, 60.2%</td>
</tr>
<tr>
<td></td>
<td>2.7 V</td>
<td>20.2 dBm</td>
<td>61.7%</td>
<td>19.9 dBm, 60.4%</td>
<td>19.3 dBm, 58.8%</td>
</tr>
<tr>
<td></td>
<td>2.4 V</td>
<td>19.7 dBm</td>
<td>61.6%</td>
<td>19.5 dBm, 60.0%</td>
<td>18.3 dBm, 57.1%</td>
</tr>
<tr>
<td>BS-cascode</td>
<td>3.0 V</td>
<td>20.8 dBm</td>
<td>66.5%</td>
<td>20.6 dBm, 62.3%</td>
<td>20.3 dBm, 61.1%</td>
</tr>
<tr>
<td></td>
<td>2.7 V</td>
<td>20.6 dBm</td>
<td>66.6%</td>
<td>20.4 dBm, 61.4%</td>
<td>19.9 dBm, 59.3%</td>
</tr>
<tr>
<td></td>
<td>2.4 V</td>
<td>20.1 dBm</td>
<td>66.7%</td>
<td>20.0 dBm, 61.0%</td>
<td>19.2 dBm, 58.7%</td>
</tr>
</tbody>
</table>

As the $V_{G,CG}$ decreases, the maximum output power decreases for both cascode topologies. The reduced current swing causes a decrease of the output power because the $g_m$ collapses earlier with lower $V_{G,CG}$. However, the maximum efficiency shows almost the same values for the different bias conditions. The BS-cascode shows better performance in the maximum output power and the maximum efficiency. At the optimum impedance, the BS-cascode also achieved larger output power and efficiency for all
Figure 13. Comparison of power performances according to the input power sweep at two different $V_{G,CG}$ values. (a) $V_{G,CG} = 3.0$ V. (b) $V_{G,CG} = 2.4$ V.
different values of $V_{G,CG}$. This comes from better current capability of the BS-cascode. The BS-cascode can supply larger current compared to the BG-cascode which has a larger threshold voltage of its CG device due to the body effect.

Single-tone measurements have been conducted at the optimum impedances for the BG- and BS-cascodes. As the input power sweeps, PA characteristics such as output power, power added efficiency (PAE), and gain have been measured and compared between the two cascode topologies at $V_{G,CG}$ of 3.0 V and 2.4 V as shown in Figure 13. At a $V_{G,CG}$ of 3.0 V, BS-cascode shows slightly higher small-signal gain and saturation power than the BG-cascode. The small-signal gain difference becomes larger and thus the efficiency of the BS-cascode is larger at $V_{G,CG}$ of 2.4 V.

Figure 14. Power performance versus gate voltage of CS device, $V_{G,CS}$, with $V_{G,CG} = 2.4$ V and input power of 0 dBm at 1.9 GHz.
The effect of $V_{G,CS}$ on the power performance is shown in Figure 14. With a input power of 0 dBm and $V_{G,CG}$ of 2.4 V, the gate bias of the CS device was swept from 0.3 V to 0.9 V. The output powers are saturated at an input power of 0 dBm. As $V_{G,CS}$ increases, the output power and DC current also increase for both cascode types. However, the BS-cascode has flatter and higher efficiency than BG-cascode over the swept gate voltages of CS device.

3.2.3. Linearity characteristics according to the cascode topology

In order to compare the linearity characteristics, two-tone signal tests have been conducted to measure inter-modulation for both of the BG- and BS-cascode previously described. Before the two-tone measurements, the optimum impedances have been extracted through the two-tone source-pull and load-pull tests with two MHz-spacing signals at a center frequency of 1.9 GHz. Different from the single-tone measurement setup shown in Figure 11, two signal generators were used to apply the two input signals with the same power level and different frequencies. Through the measurements with the two-tone signal, inter-modulation distortions (IMDs) have been measured in addition to the parameters from the single-tone measurement such as output power and efficiency. Among the IMDs, the third order IMD, IMD$_3$, is used to compare the linearity of the power cells according to the cascode types. The IMD characteristics are widely used in linearity tests and can give intuition about the out-of-band linearity such as adjacent channel power ratio (ACPR) or adjacent channel leakage ratio (ACLR).

Not only linearity but also output power and efficiency are important in linear power amplifier design. However, both power and efficiency are not good enough at the optimum impedance for linearity only. To reduce the measurement uncertainty during the
trade-off between interest parameters, the optimum impedances have been selected in the same manner as was done previously for the single-tone measurement.

Figure 15 shows the measured optimum load impedances according to the $V_{G,CG}$ variation at a fixed gate bias for the CS device at 0.6 V. As the $V_{G,CG}$ changes from 3.0 V to 2.1 V with step of 0.3 V, the optimum impedance moves to higher impedance for both cascode types. At the different $V_{G,CG}$ values, the traces of optimum impedance show an identical trend. As the gate bias $V_{G,CG}$ decreases, the current at the higher $V_{G,CS}$ decreases as shown in Figure 8, and the slope of the load line decreases to guarantee higher output power and efficiency. Therefore, the optimum impedance increases when $V_{G,CS}$ decreases.

The IMD3 can be expressed as the ratio of carrier power to 3rd order inter-modulation power, $C/I_3$. A larger value of $C/I_3$ means larger difference between the carrier power and
inter-modulation power, and thus implies better performance in terms of out-of-band linearity. In order to compare linearity characteristics according to the cascode types, the two-tone power measurements were conducted with a variation of $V_{G,CG}$ at a fixed $V_{G,CS}$ of 0.6 V.

In Figure 16 (a) and (b), the measured IMD$_3$s according to the gate bias of the CG device are shown for the BG-cascode and BS-cascode, respectively. At low output power, the C/I$_3$ curves are not affected by the variation of $V_{G,CG}$ and decrease with the slope of around -2 dBc / dBm. These curves reflect very good linearity at this region. As power increases more, the slopes start to increase and reach a flat region for the lower gate bias. For the higher gate bias cases, the C/I$_3$ curves have local maxima with different maximum values. These phenomena are called as “IMD sweet spots” where high linearity shows. These can be explained by the $g_{m3}$ non-linearity cancellation during large signal operation [35]. The sweet spots are mainly affected by the bias condition and appear in various forms and at different positions according to the operation class. Although the gate bias conditions of the CS devices are the same, the IMD sweet spots appear in different shapes according to the gate bias conditions of the CG device because of the different $g_{m}$ characteristics.

The trend of carrier-to-inter-modulation, C/I$_3$, is similar at the same biasing condition for both cascode types. However, as mentioned previously, the output power and the efficiency are important when linearity is comparable. In other words, a PA can be recognized as the more linear when it has higher maximum linear efficiency and maximum linear output power while satisfying the linearity specification. From the data of the two-tone measurements, the efficiency and output power are extracted when the
Figure 16. Comparison of linearity according to the cascode types. (a) IMD$_3$ characteristics of BG-cascode as a function of $V_{G,CG}$. (b) IMD$_3$ characteristics of BS-cascode as a function of $V_{G,CG}$. (c) output power and efficiency comparison at the C/I$_3$ of 25 dBc.
C/I\textsubscript{3} reaches the value of 25 dBc. In Figure 16 (c), the extracted efficiencies and output powers as a function of V\textsubscript{G,CG} are shown and compared according to the types. The BG-cascode has slightly higher maximum linear output power than the BS-cascode over the variation of V\textsubscript{G,CG} at this bias level of the CS device. Because the BG-cascode has a peak efficiency at the smaller output power compared to the BS-cascode with the same bias conditions and device sizes, the BG-cascode usually has higher efficiency than the BS-cascode at the same output power although it has smaller peak efficiency. Therefore, the maximum linear efficiencies of the BG-cascode are higher than those of the BS-cascode.

### 3.3. Optimization of Device Sizing and Biasing Conditions

In the previous section, the BG- and BS-cascodes were compared from the viewpoint of power and linearity performances. The BS-cascode showed higher power capability and efficiency, and the BG-cascode showed better linearity at the same bias condition and the fixed ratio between the size of the CG device and that of the CS device. In order to utilize the previous results in PA design, finding the optimum device sizing and biasing condition is necessary.

For a watt-level PA application, the size of the active device is much larger than the power cell. Multiple power cells are combined to generate the targeted power. However, it is not easy to analyze the characteristics of a transistor with large size because a large transistor is apt to oscillate due to the high gain. The oscillation usually occurs in the DC measurement as well as in RF measurement. To remove the concern of instability during the measurement, power cells with small sizes were fabricated and fully analyzed for the optimization of sizing and biasing conditions. Although the effects of parasitic
components increase when the devices become larger, it is meaningful to analyze and understand the characteristics of a small power cell because the inherent device characteristics are not changed.

The BG- and BS-cascodes have been fabricated using 0.18 μm RF CMOS process with different ratios of between the CG and CS devices as shown in Figure 17. The ratios of CG to CS device sizes cover from 2:3 to 6:3 with a total gate width of 384 μm for the CS device. At various biasing conditions, on-wafer two-tone measurements have been conducted at the optimum input and output impedances that were achieved from the source-pull and load-pull measurements with 2 MHz spacing two-tone signal at a center frequency of 1.9 GHz.

Figure 18 shows the dependency of the saturated output power and efficiency of the BS-cascode on the ratio of CG to CS device sizes and the gate bias of the CG device,
Figure 18. Dependency on the device size and $V_{G,CG}$: (a) Saturated $P_{out}$. (b) Saturated power added efficiency.
The gate voltage of the CS device was set as 0.6 V. At the same device size, both output power and efficiency increase as $V_{G,CG}$ increases, which is the same as the previous results. The efficiency decreases monotonically as the size of device increases at the same $V_{G,CG}$ because of the increasing current consumption with the increase of device size. On the contrary, the output power has a local maximum point around the ratio of 3:3 ~ 3:5 and $V_{G,CG}$ of 3.0 V. Considering both $P_{out}$ and efficiency, the optimum ratio of device size is around 3:3 ~ 4:3, and the optimum $V_{G,CG}$ is the highest one.

With the achieved optimum device size of 4:3, the gate bias of the CS device has been swept from 0.5 V to 0.7 V with 0.1 V step to find the optimum bias condition of the CS device and the dependency of $P_{out}$ and PAE on the bias condition. In Figure 19, the bias dependencies of the saturated $P_{out}$ and PAE are shown. As $V_{G,CS}$ and $V_{G,CG}$ increase, the maximum output power increases. On the contrary, the maximum efficiency is not affected much by the swept $V_{G,CS}$ at higher $V_{G,CG}$. From the viewpoint of PAE only, the efficiency has an optimum region at the $V_{G,CG}$ range of 2.7 V ~ 3.0 V and the $V_{G,CS}$ of 0.5 V ~ 0.7 V. However, because too much current will flow, the efficiency will drop when $V_{G,CS}$ increases further. Therefore, the optimum gate bias of the CS device is around 0.7V.

For the application without a linearity burden, higher $V_{G,CG}$ and $V_{G,CS}$ can guarantee a higher maximum $P_{out}$ and PAE.

In order to investigate optimum bias conditions and device sizes for the linear PA application, the IMD3 characteristics as well as power characteristics have been analyzed through measuring carrier-to-3rd order inter-modulation, $C/I_3$. Different from the measuring the saturated $P_{out}$ and PAE, the maximum linear $P_{out}$ and maximum linear PAE have been extracted from the $C/I_3$ measurement results of the BG-cascode with variations.
Figure 19. Bias dependency at optimum device size. (a) Saturated $P_{\text{out}}$. (b) Saturated power added efficiency.
Figure 20. Dependency on the device size and $V_{G,CG}$ in linear operation. (a) Maximum linear output power. (b) Maximum linear PAE.
of the bias condition and ratios between the size of the CG device and that of the CS device. In these measurements, maximum linear $P_{\text{out}}$ and maximum linear PAE are defined as the values of $P_{\text{out}}$ and PAE when the C/I$_3$ reaches 25 dBc, (IMD$_3$ = -25 dBc). The IMD$_3$ of -25 dBc is very closely related to the linearity specification in the WCDMA application, ACPR of -33 dBc at 5 MHz offset from center frequency with 3.83 MHz channel bandwidth.

Figure 20 shows the dependencies of the maximum linear output power and maximum linear PAE on the device size variation and the gate voltage of the CG device. According to the $V_{G,\text{CG}}$ variation, efficiency and power show the same trends. As the $V_{G,\text{CG}}$ increases, both $P_{\text{out}}$ and PAE increase. The saturated $P_{\text{out}}$ and PAE show smooth contour lines according to the variation of the ratio of the size of the CG device to that of the CS device. However, contrary to the $P_{\text{out}}$ and PAE dependencies in the saturation, the maximum linear $P_{\text{out}}$ and maximum linear PAE show undulant contour lines according to the size variation. This is related to the variation in the “depth” of the sweet spot according to the device sizing, and more profound studies are needed to understand this phenomenon in detail. From the optimum regions of $P_{\text{out}}$ and PAE, the common area shows the ratio of 4:3 and $V_{G,\text{CG}}$ of 3.0 V for optimum values.

At a glance, the achieved optimum values of the device ratio and the $V_{G,\text{CG}}$ for linear operation are almost similar to those for the operation in saturation. However, the optimum gate bias of the CS device is quite different. In Figure 21, the dependency of C/I$_3$ on $V_{G,\text{CS}}$ is presented. At the optimum device ratio of 4:3 and $V_{G,\text{CG}}$ of 3.0 V, the IMD$_3$s have been measured at different values of $V_{G,\text{CS}}$ by sweeping the input power. The measured C/I$_3$ curves show very different patterns according to the gate biasing condition.
of the CS device. At $V_{G,CS}$ of 0.7 V, the linear maximum output power is smallest, and the linearity is worst compared to the others. The maximum output powers for $V_{G,CS}$ of 0.6 V and 0.55 V are almost the same values. The gap in maximum output power is about 4 dB, which is quite big. The biggest difference in C/I3 curves according to the $V_{G,CS}$ is the sweet spot. As mentioned before, the existence of the sweet spot can be explained with the $g_{m3}$ nonlinearity cancellation according to the operation Class, which is very closely related the gate bias in the common source configuration [35]. The numbers and positions of sweet spots can indicate the PA operation Class. At a gate bias of 0.55 V, the quiescent point is located close to the threshold point, so the PA cell operates in “deep” Class AB, which is Class AB close to the Class B operation. In deep Class AB operation, two sweet spots exist. As $V_{G,CS}$ increases, the sweet spot at low output power disappears, and a single sweet spot exists as shown in the curve of the 0.6 V case. As the $V_{G,CS}$
increases further, no benefit of a sweet spot exists as shown in the curve of the 0.7 V case when the operation is close to Class A. Therefore, the optimum gate bias of the CS device is around 0.55 ~0.6 V for linear operation.

3.4. Conclusion

Two types of CMOS cascode power cells, BG-cascode and BS-cascode, power cell have been fully analyzed and compared from a PA application perspective. Although the device size and biasing conditions are exactly the same for both cascode configurations, their characteristics show differences in both DC and RF performances. According to the connection of the body terminal, the BG-cascode has a body effect, whereas the BS-cascode does not. Therefore, the BS-cascode has higher current capability, including larger $g_m$ and wider flat-$g_m$ region, than the BG-cascode. The BS-cascode also shows better power performance from the viewpoint saturated $P_{out}$ and saturated efficiency. On the contrary, the BG-cascode shows better linearity performance in terms of IMD3 characteristics. It shows higher maximum linear $P_{out}$ and maximum linear PAE. From these results, BS-cascode configuration is more suitable for the application in saturation, and the BG-cascode configuration is beneficial for the linear application.

Through the optimization of the ratio between the size of the CG device and the gate biasing conditions for both CG and CS devices, optimum conditions are achieved according to the applications. For both applications, in saturated operation and in linear operation, the values of the optimum device ratio and the gate biasing for the CG device are analogous. However, the optimum bias conditions for the gate of the CS device are different from each other. The BS-type cascode power cell can be a good candidate for
the saturated PA application with high gate biasing for both CS and CG devices, whereas
the BG-type cascode power cell with high gate biasing for CG device and low gate
biasing for CS device shows better performances for the linear PA application.
CHAPTER 4
DESIGN OF GSM/EDGE DUAL-MODE POWER AMPLIFIERS

4.1. Introduction

Recently, there are a lot of research activities on developing CMOS PAs ongoing in the industries and academia because of the capability of integration and low cost of CMOS processes. These two advantages make a CMOS PA a rising candidate in the PA markets. As personal communication markets grow, multiple communication standards are being embedded into one single transceiver in order to ultimately reduce the overall costs while supporting higher data-rate communication. Cell-phone makers need to support these multiple standards in single cell phone. The PA is the biggest challenge toward a multi-standard solution. The conventional methodologies are parallel combinations of PAs for each standard, which increase cost. This is why multi-mode PAs should be utilized in multi-standard applications. In this research, a single CMOS PA with tunable matching network is introduced, which is able to support GSM and the EDGE applications. This approach can be one of the solutions in the implementation of a multi-mode power amplifier.

4.2. Dual-Mode Operation and Matching Network Transition

Figure 22 shows the different PA performance requirements for GSM and the EDGE applications. For the GSM application, high power and high efficiency are required. However, it does not need high linearity because of constant envelope modulation of the
GSM signal. In a CMOS implementation, for higher efficiency, a GSM PA usually adopts switching operation. On the other hand, an EDGE-mode PA needs higher linearity to support a time-varying envelope signal. Although there are several approaches to support an EDGE signal with a switching mode PA, such as a polar modulated PA and an out-phasing structure, however, both approaches are somewhat complex and need additional control block for linear operation.

To satisfy the linearity, a typical PA operates in the linear output power range which is backed-off by certain amount of power from the saturated output power. In the implementation of dual-mode PA, both design targets need to be considered. Figure 23 shows a conceptual diagram that illustrates the summary of performance requirements of GSM and EDGE PAs.
Because of the different performance requirements of GSM and EDGE PAs, they have different optimum output impedances according to the operation mode, which means that the GSM mode needs to have the impedance more optimized for higher output power and efficiency while the EDGE mode needs to have the impedance more optimized for linearity.

As shown in Figure 23 (a) the optimum impedances for two different modes of operation are not the same. With the optimum impedance corresponding to each
operation mode, the PA performance can be maximized at each operation mode, respectively.

However, if we have a fixed output matching network as is typically used, the PA just sees only a single impedance although the optimum impedances for two different modes of operation are not coincided. In this situation, the designer has no other option but to choose a single impedance point for the output matching of the PA enduring some degradation of performance. When the impedance is chosen close to one mode, the performance at the other mode will be sacrificed. Therefore, the compromised impedance must be chosen somewhere between the two impedances.

As shown in Figure 23 (b), if a tunable output matching network is adopted in the dual-mode PA design, variable output impedances can be obtained, and the output impedances can be placed much closer to the optimal points corresponding to each operation mode. It can give the designer an additional degree of freedom in choosing the output impedance of a PA. By the use of this tunable matching component, high performance for two different modes of operation with a single PA can be achieved.

4.3. Design of Sub-Blocks for a Watt-Level PA

The objective of this research is to develop design techniques to implement a fully-integrated dual-mode power amplifier using CMOS technology, primarily focusing on output power and efficiency of switching mode operation under constant envelope signal and linearity for linear operation mode under non-constant envelope signal with a single PA.
4.3.1. Differential Cascode topology PA core with cross-coupled capacitors

Figure 24 shows the detailed schematic of the designed power amplifier core. To reduce voltage stress on the active device, a cascode topology is adopted. The total gate widths for both CG and CS devices are 4 mm, and a thick-gate-oxide device is used for the CG transistor to support a larger voltage swing with less stress at the output ports. Also a differential topology is adopted to reduce unwanted bond-wire effects. Compared to the single-ended topology, a differential amplifier has several advantages in power amplifier applications:

1. It is immune to bond wire effects due to the virtual AC ground.
2. It has smaller even harmonics at the output.
3. It has better linearity due to a smaller input signal for the same output.

For a higher gain of the power amplifier, cross-coupled capacitors (C_{CC}) are adopted, which are connected between the gate of the CS device at one branch of differential pair and the drain of the CS device at the other branch shown in Figure 24. In addition, another important advantage of using the C_{CC} is neutralizing gate-drain capacitance having better linearity [26].

From [26], the nonlinearity of input capacitance is problematic due to the variation of Miller gain (A_{VM}, defined as the gain from the gate of CS device to the drain of CS device) when the CG device enters the triode region while the CS device is in saturation. Therefore, at a large voltage swing, the nonlinearity of the input capacitance due to the Miller effect deteriorates the linearity of the power amplifier. To remove this problem, cross-coupled capacitors can be used in the differential topology, and the input capacitance can be expressed as
where $A_{VM}$ is the Miller gain, which is varied according to the output power. By choosing the value of $C_{CC}$ to eliminate the last term in equation (6), $C_{in}$ can be independent of Miller gain.

By using $C_{CC}$, the overall gain of the amplifier can be boosted and the nonlinearity due to the variation of input capacitance can be reduced. In addition, $C_{CC}$ can give some degree of freedom in designing a linear PA by manipulating AM-AM and AM-PM characteristics. In Figure 25, the simulated gain and phase of the PA according to the output power are shown as a function of $C_{CC}$. A larger value of $C_{CC}$ can give a larger gain but causes larger variations of gain and phase over output power. For a smaller value of...
Figure 25. AM-AM and AM-PM characteristics according to the cross-coupled capacitors. (a) Gain variation. (b) Phase variation.
CCC, the gain degrades at the low output power, thus, the PA has a smaller 1dB compression point of output power. The gain and phase variations are reduced but are not small enough.

As shown in Figure 25, there exists an optimal value of $C_{\text{CC}}$ to minimize AM-AM and AM-PM variations which are strongly correlated with the error vector magnitude (EVM) of a linear PA. The linearity of the PA can be manipulated by exploiting $C_{\text{CC}}$ in the differential topology.

However, the cross-coupled capacitors make positive feedback loops that may cause some stability issue in the circuit. Therefore, the value of $C_{\text{CC}}$ needs to be chosen very carefully not to make PA unstable but to achieve advantages, such as higher gain and better linearity. For more guaranteeing stability of the PA, the RC feedback network between the drain of the CG device and the gate of the CS device was exploited at both branches. The RC feedbacks form a negative feedback from the output port to the input port and make the circuit stable. The impedance of the RC feedback has an optimum value for the purposes of stability and high gain. Too small of an impedance of RC feedback gives too much feedback from output to input and causes gain degradation, whereas too large impedance of feedback would not stabilize the circuit. Considering both performance and stability for the linear PA, 2.5 pF of $C_{\text{CC}}$ and 500 $\Omega$ of $R_{\text{fb}}$ and 0.5 pF of $C_{\text{fb}}$ were selected for the design of the power amplifier core.

4.3.2. Optimum load impedances for GSM and the EDGE applications

The tunable output matching network used in the CMOS PA for dual-mode operation is shown in Figure 26 (a). It is composed of a transformer and two tunable input capacitors, $C_{\text{IN}}$, and an output capacitor, $C_{\text{OUT}}$. Two pairs of a differential PA core were
Figure 26. Tunable matching network and optimum impedances. (a) Schematic of tunable matching network with 2×1:2 parallel combining transformer. (b) Optimum impedances for the GSM and EDGE applications at 1.8 GHz.
used to achieve Watt-level output power. The magnetic coupled transformer can provide power combining and impedance transformation simultaneously so that it is suitable for Watt-level CMOS PA design [9]. An on-chip 2×1:2 parallel-combining transformer (PCT) is used for efficient power combining and impedance transformation [6]. It also serves as an output balun, which converts the differential signal to a single-ended signal to be delivered to the load. The role of $C_{\text{OUT}}$ is to minimize the loss due to the inductances of the transformer by resonating inductance of a secondary winding with $C_{\text{OUT}}$ [10].

The variable $C_{\text{IN}}$ described previously is adopted to vary the load impedance according to the operation mode, and its structure is shown in Figure 26 (a). The capacitance values of the variable capacitor were determined by the optimum load impedances according to the applications. For more convenient ways of determining the capacitance, a single input port was used by tying together two nodes with the same polarity in the two differential input ports of the transformer. The impedance looking into the output matching network from the PA cores is defined as $Z_{\text{out}}$, as shown in Figure 26 (a).

For switching mode operation, the optimum output load impedance can be obtained from the output power requirements [11]. From this initial solution, optimization is followed using harmonic balance simulations to get the optimum impedance for the GSM application. To find the optimum impedance for linear amplifier operation, load-pull simulations are conducted with changing bias points. For higher efficiency and linearity, the bias point is set around class AB operation for the EDGE application.
From the simulations and optimizations, 12.6 + j5.9 of $Z_{\text{out}}$ was achieved for the GSM application and 8.9 + j7.8 of $Z_{\text{out}}$ was achieved for the EDGE application at the frequency of 1.8 GHz. These optimum values of $Z_{\text{out}}$ were shown in Figure 26 (b).

4.3.3. CMOS on-chip transformer

Although PAs with CMOS technology have some drawbacks compared to their competitors, the activity of CMOS PA development is remarkable because of the advantages of CMOS PAs such as low cost and the ability of integration. In watt-level CMOS PA design, in order to satisfy the output power requirement, the power combining technique is usually adopted. The smaller output powers from single PAs are added up to achieve the required power. To further increase the output power, more unit PAs should be combined effectively. By using the power combining technique, it is possible to reduce the burden of the necessary impedance transformation without increasing the impedance transformation ratio too much. Output matching networks with large impedance transformation ratios increase power loss and degrade efficiency of a power amplifier. In addition, by using a transformer as a power combiner, the transformer works as not only a power combiner but also a impedance transformer, allowing a higher load impedance to be transformed to a lower impedance at the output of the PA.

When the output power reaches watt level, the PA efficiency can be degraded considerably due to the resistance of the output matching network. Because the current flowing in the output matching network can have a level of amperes, even a small resistance at the output can cause significant loss. Hence, it is essential to reduce the loss of the matching network in order to obtain high PA efficiency. As mentioned before, the conventional CMOS substrate is lossy so that the quality factor of inductive components
is quite low. However, the reduction of DC resistance itself can lower the loss at the output matching network by using thick metals as signal lines. Usually in RF CMOS processes, thick top metals are adopted to reduce resistances of passive components. Figure 27 shows the cross-sectional view of the RF CMOS process that is used for this on-chip transformer design. The top three layers (3 μm of copper layer E1, 3 μm of copper layer E2, and 4 μm of aluminum layer MA) can be used for the transformer design. Those three layers can be stacked for the signal lines, and a transformer with a lower resistance with reasonable Q-factor can be achieved for Giga-hertz application.

In this power amplifier design, two pairs of differential unit PAs are used to generate the required output power, especially for GSM and the EDGE applications. In order to implement a fully integrated CMOS PA, the transformer-based output matching network...
is used to combine power from each unit PA and convert the differential output to a single-ended load at the same time. Two primary transformer windings are needed because of the two unit-PA pairs. For higher transformer efficiency and smaller chip size, a parallel combing structure is adopted in this design [19]. The 50-Ohm load impedance is transformed to a lower impedance at each unit PA output by using a step-up transformer. According to the required impedance transformation ratio, the turn ratio between primary turns and secondary turns is set as 1:2. Thus, the designed transformer can be notated as a 2×1:2 PCT; that is, a parallel-combining transformer with two 1-turn primary windings and one 2-turn secondary winding.

The characteristics of a transformer such as Q-factor, self inductances of primary and secondary windings, and coupling factor, are truly layout-dependent. In the viewpoint of a PA application, the transformer should have a high Q-factor and a high coupling factor.
to transfer power to the load with a small loss. These electrical properties of transformers are closely related to the physical parameters of transformers such as outer dimension, the width of metal traces, and spacing between primary and secondary windings. Hence, careful optimization using EM simulation is needed to get a highly efficient transformer considering the size and passive efficiency.

Figure 28 shows the simplified schematic and the layout of the designed 2×1:2 parallel combining transformer. It has two input ports (P1, P2) for two primary windings and one output port (S1) for secondary winding. The input ports will be connected to the outputs of the two unit PAs, and the output port will be connected the load. The width and the spacing of the metal traces are 30 μm and 10 μm, respectively. The three stacked top layers as shown in Figure 27 were used to reduce resistances of metal traces. By considering the frequency band of interest and the chip size, the size of the transformer is 0.78 mm x 0.72 mm as shown in Figure 29 (a).

From the 2-port impedance network of a transformer and voltage and current relationships, the electrical characteristics of the transformer can be extracted [16]. The self inductance of the primary winding (L1) and that of secondary winding (L2) are extracted from the following equations

\[
L_1 = \text{Re}(Z_{11})/2\pi f \quad \text{and} \quad L_2 = \text{Re}(Z_{22})/2\pi f
\]

(7)

where \( f \) is operation frequency. From the extracted self inductances, the coupling coefficient, \( k \), mutual inductance, \( M \), and the actual turn ratio, \( N \), also can be given as follows.
Figure 29. Designed transformer. (a) Physical layout of 2x1:2 step-up transformer with vertical metal structure. (b) Simulated transformer efficiency.
\[
    k = \frac{\text{Im}(Z_{12}) \times \text{Im}(Z_{21})}{\text{Im}(Z_{11}) \times \text{Im}(Z_{22})}
\]

(8)

\[
    M = k\sqrt{L_1 \times L_2}
\]

(9)

\[
    N = \sqrt{L_2 / L_1}
\]

(10)

The electrical characteristics of the designed transformer are summarized in Table 4. At the frequency of interest, 1.81 GHz, the primary winding and secondary winding have Q factors of 16.6 and 21.1, respectively. Because of the low resistive metal traces, the transformer has high Q factors in spite of the conductive Si substrate. From maximum available gain when the two input ports are tied, losses are calculated. At 1.81 GHz, the insertion loss of the transformer is 0.51 dB with a differential load, and is 0.8 dB with a single-ended load, which are 89% and 83% of the transformer efficiencies, respectively, as show in Figure 29 (b).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Primary Winding</th>
<th>Secondary Winding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self Inductance ((L_1, L_2))</td>
<td>1.45 nH</td>
<td>3.82 nH</td>
</tr>
<tr>
<td>Resistance ((R_1, R_2))</td>
<td>0.99 Ω</td>
<td>2.07 Ω</td>
</tr>
<tr>
<td>Quality Factor</td>
<td>16.6</td>
<td>21.1</td>
</tr>
<tr>
<td>Coupling Coefficient ((k))</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Mutual Inductance ((M))</td>
<td>1.55 nH</td>
<td></td>
</tr>
<tr>
<td>Actual Turn Ratio ((N))</td>
<td>1.62</td>
<td></td>
</tr>
</tbody>
</table>

4.3.4. Variable capacitor

To implement different impedances looking into the output matching according to the applications, a high-power and highly linear switchable capacitor is used. It consists of 4
Figure 30. High-power tunable capacitor. (a) Schematic of tunable capacitor with 4 MIM capacitors and a switch transistor. (b) Characteristics of “ON” state. (c) Characteristics of “OFF” state.
metal-insulator-metal capacitors and one switch transistor as shown in Figure 30 (a). By turning on and off the switch transistor, the capacitance between the two nodes can be varied. To sustain higher voltage stresses between the terminals of the switch transistor, a thick-gate-oxide transistor is used for the switch transistor.

Because of the large signal at the variable capacitor for a Watt-level PA application, it is very important for the variable capacitor to sustain its capacitance value without degrading its linearity. As the power increases, the increasing voltage swings between the terminals of the switch transistor can turn on parasitic junction diodes and/or the switch transistor itself. Unlike conventional variable capacitors, DC voltages can be applied at all terminals of the switch transistor through the high-resistance resistors in our structure. This can prevent the unwanted turn-on phenomenon and make our variable capacitor sustain high voltage with high linearity. In this design, the peak voltage between variable capacitor terminals can reach up to 9 V at peak power of 32 dBm. The high-power tunable capacitor used in this design does not deteriorate its characteristics up to 10 V.

Figure 30 (b) and (c) show the simulated capacitances and Q-factors of the designed tunable capacitor according to the switch conditions. The capacitances can be expressed as following according to the states of the switch transistor

\[ C_{T,ON} \approx C_4 + C_1 \parallel C_3 \]  \hspace{1cm} (11)

\[ C_{T,OFF} \approx C_4 + C_1 \parallel C_2 \parallel C_3 \]  \hspace{1cm} (12)

where \( \parallel \) denotes series connection of capacitors and the parasitic capacitances of the switch are ignored. The Q-factor of the on-state drops because of the on-resistance of the switch transistor but still has a value that is high enough compared to the other passive components, such as inductor and transformer. The size of the transistor needs to be
chosen carefully by considering the trade-off between the Q factor and the parasitic capacitances of the switch transistor, and 1.2 mm of total gate width was used for the switch transistor in this design. In Table 5, the design values and biasing conditions used in this design were shown to implement the tunable capacitor.

Table 5. Bias condition and designed values for the tunable capacitor.

<table>
<thead>
<tr>
<th>Bias / capacitance / design variables</th>
<th>ON state</th>
<th>OFF state</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{G,SW}$</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>$V_{D,SW}, V_{S,SW}$</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>$V_{B,SW}$</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Capacitance / Q-factor (@1.8 GHz)</td>
<td>4.6 pF / 45.5</td>
<td>3.5 pF / 141</td>
</tr>
<tr>
<td>$C_1, C_3$</td>
<td>4.9 pF</td>
<td></td>
</tr>
<tr>
<td>$C_2$</td>
<td>1.75 pF</td>
<td></td>
</tr>
<tr>
<td>$C_4$</td>
<td>2 pF</td>
<td></td>
</tr>
</tbody>
</table>

4.4. Conclusion

The necessity of the tunable matching is addressed in this chapter. Due to the different requirements of power and linearity according to the communication standards, it is not easy to support multiple communication standards with a single PA. To solve this problem and utilize a single PA for multiple standards, the tunable matching network is introduced and the effectiveness is discussed. For the design of a fully integrated multi-mode PA, the required sub-blocks such as a PA core, transformer, and tunable capacitor are introduced and designed.
The designed PA core utilizes a differential cascode topology. Thick-gate-oxide transistors are used for the common gate device to reduce the reliability issues due to the high voltage swing at the output. To get a higher gain and better linearity, cross-coupled capacitors are used in the differential topology, and the design issues are addressed including stability. Tunable matching with an on-chip transformer and variable capacitors is introduced for the GSM /EDGE dual-mode application. For the dual-mode application, two different optimum impedances are obtained according to the performance requirement of PA according to the application. A highly efficient on-chip transformer and a high-power variable capacitor are designed and discussed to implement the tunable matching network at the PA output.
5.1. Implementation of Single Stage GSM/EDGE Dual Mode PA

To avoid the complexity of a multi-stage PA and verify the effect of the transition of load impedances according to the operation modes, a PA circuit is designed for the power stage only. A schematic of a dual-mode CMOS RF PA is shown in Figure 31. It is composed of three parts, (1) an input matching network including an input balun to convert single-ended signal to differential signal, (2) a single stage PA composed of two pairs of CMOS differential cascodes, and (3) a tunable output matching network with variable capacitors and a on-chip transformer to change the load impedance according to the operation modes.

The input matching network consists of symmetric inductors with a center-tap for DC biasing and an on-chip input balun with input and output capacitors, \( C_{B1} \) and \( C_{B2} \). By using the input balun, the single-ended RF signal is converted into the differential signal. To reduce the loss due to the inductances of the balun, \( C_{B1} \) and \( C_{B2} \) tunes input and output inductances of the balun, respectively. The PA core uses two pairs of differential cascodes and a single stage of PA that is described in the simplified schematic of Figure 31. The sizes of cascode devices are 4 mm for the CS device and 4 mm for the CG device. The CG device uses thick gate oxide. The detailed schematic of the PA core is shown in Figure 24. For the tunable output matching network, 2×1:2 step-up on-chip transformer and tunable capacitors were used.
The dual-mode GSM/EDGE PA is implemented with 0.18 μm RF CMOS process. Figure 32 shows a microphotograph of the chip with a size of 1.7 x 1.1 mm² including the on-chip output transformer and all pads. The size of implemented the tunable capacitor is
0.16 x 0.1 mm$^2$ which is less than 1% of the total chip area. Even more, the tunable capacitor is small enough so that it can be implemented without any increase in the total chip size as shown in Figure 32. The chip is assembled on a 2-layer FR-4 evaluation board for the measurement. The loss from the printed circuit board (PCB) is compensated while that of wire-bonding is included.

5.2. Measurement Results of Single Stage GSM/EDGE Dual Mode PA

5.2.1. Measurement setup

Figure 33 shows the measurement test bench used to characterize PAs on evaluation boards. The test bench consists of a vector signal generator, a vector signal analyzer, a power meter, and power suppliers. The vector signal generator provides input signals, such as single tone and modulated signals to the PAs, and the vector signal analyzer or spectrum analyzer analyzes the output signal of the PAs according to the measurements. The power meter measures the input and output powers through two separate power sensors as shown in Figure 33. Before power measurements, calibration of the power meter should be performed to remove any offsets from the power sensors. The power calibration is very important for precise measurement, and the detailed procedure of power calibration is presented in other literature. All measured voltages and currents come from power supplies during measurements.

5.2.2. PA performance in the GSM application

For GSM operation, the idle current is set at 300 mA with 2.8 V of the CG gate voltage, and the DC supply is 3.4 V. To set the optimum load impedance for GSM
operation, the capacitance of the tunable capacitor is set at 4.6 pF. The switch transistor is biased in its “ON” state by applying 3.4 V at the gate and 0 V at the source and drain terminals of switch transistor, respectively. Figure 34 shows the measured results with a single tone at the GSM application conditions, such as load impedance and bias conditions. In Figure 34 (a), the power amplifier delivers fairly constant output power of more than 31.7 dBm with less than 0.3 dB of variation from 1.6 GHz to 1.9 GHz. However, the efficiency decreases slightly as the frequency increases. The output power has its maximum value of 32 dBm at around 1.75 GHz of operating frequency, and the drain efficiency has its maximum value of 48% at a lower frequency than that for maximum output power by 100 MHz. This is due to the additional parasitic passive
Figure 34. Single-tone measurement results for the GSM application. (a) Pout and drain efficiency as a function of frequency. (b) Pout and drain efficiency as a function of supply voltage.

Pin = 19 dBm
Freq. = 1.76 GHz
components that appeared during the layout but were not considered in the simulation. The $P_{out}$ has a linear relationship with the square of the supply voltage as shown in Figure 34 (b).

Figure 35 shows the $P_{out}$, drain efficiency, and gain characteristics versus input power. At 1.76 GHz of operation frequency, the saturated output power is 32 dBm with 45% drain efficiency as shown in Figure 35. Due to the gain boosting effect of the cross-coupled capacitor, the small-signal power gain is 22.2 dB, which is considerably high for the PA with a single stage.

To verify the potential of the fabricated dual-mode PA in the GSM application, it was tested by applying a Gaussian minimum shift keying (GMSK) modulated signal with bandwidth-bit-time of 0.3 (BT=0.3). Figure 36 shows the measured spectrum at the
maximum output power with the GSM mask. The output spectrum of the PA was well confined in the GSM spectral emission mask up to 32 dBm output power at 1.76 GHz.

5.2.3. PA performance in the EDGE application

As opposed to GSM mode operation, better linearity is required for the EDGE mode operation. So, the bias point of PA is set at deep Class AB by setting the idle current at 140 mA with 2.5 V of the CG gate voltage and the same DC supply voltage. By changing the bias of the switch transistor as “OFF” state with 0 V at the gate and 3.4 V at both source and drain, the tunable capacitor is set at 3.5 pF, which makes the load impedance optimum for EDGE operation. In a conventional fixed output matching structure, the output impedance cannot be adapted to the different performance requirements, and the only thing that can be done is to apply different biasing conditions according to the applications.
Figure 37 shows the results of single-tone measurements with the optimum bias point and the optimum output impedance for the EDGE application. Compared to Figure 35, the characteristics of the GSM-mode PA (PA operating in the biasing condition and output impedance for the GSM application) versus input power, the EDGE-mode PA (PA operating in the biasing condition and output impedance for the EDGE application) has 19.2 dB of the small-signal gain (linear gain) which is 3 dB smaller than that of the GSM-mode PA. The reason of reduced linear gain is majorly due to the change of bias point. As expected, the gain of the EDGE-mode PA is flatter than that of the GSM-mode PA, which means better linearity of the EDGE-mode PA. For the saturation output power ($P_{sat}$), the EDGE-mode PA has 31.6 dBm of $P_{sat}$, which is smaller than that of GSM PA.

Figure 37. PA performances with single tone as a function of input power for the EDGE application.
by 0.4 dB. However, the peak efficiency of the EDGE-mode PA is similar to that of the GSM-mode PA.

To verify the effectiveness of the tunable matching network, several measurements were conducted by comparing the performances of the PAs with two different output impedances, i.e. the output impedances for GSM and EDGE, especially for PA linearity. Figure 38 shows the measurement results of the linear PA with a single tone according to the optimum output impedances for GSM and EDGE. For both cases, the bias conditions are the same. Usually, from the single-tone measurement results, linearity can be expressed with $P_{1\text{dB}}$ which is defined as the output power at the 1 dB compression point compared to the small-signal gain. Typically, a higher $P_{1\text{dB}}$ gives the better linearity. To compare the linearity of both cases, a linear $P_{\text{out}}$ and linear efficiency are defined as the...
P_{out} and efficiency at the P_{1dB} point. By changing the output impedance to the optimum value for the EDGE application, the small-signal gain is slightly reduced, but P_{1dB} increases by about 1 dB. Therefore, linear P_{out} and linear efficiency increase by 1 dB and about 7%, respectively. This implies that the output impedance for EDGE is closer to the optimum impedance point for linear operation, and the load impedance is moved to a better load impedance point in the viewpoint of linear PA. The results of single-tone measurements are summarized in Table 6.

Table 6. Summary of 1-tone measurements in linear mode operation.

<table>
<thead>
<tr>
<th>Measured parameter</th>
<th>Output imp. for GSM</th>
<th>Output imp. for EDGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small signal gain</td>
<td>19.8 dB</td>
<td>19.2 dB</td>
</tr>
<tr>
<td>P_{1dB}</td>
<td>29.5 dBM</td>
<td>30.5 dBM</td>
</tr>
<tr>
<td>P_{sat}</td>
<td>31.5 dBM</td>
<td>31.6 dBM</td>
</tr>
<tr>
<td>Efficiency at P_{1dB}</td>
<td>33 %</td>
<td>40 %</td>
</tr>
</tbody>
</table>

Applying the EDGE modulated signal at 1.76 GHz, the error vector magnitude (EVM) and the output spectrum are measured as shown in Figure 39 and Figure 40, respectively. With adaptation of the output impedance to EDGE mode by changing the capacitance value of the tunable matching network, the linearity of the PA is improved significantly. The output power and efficiency at 5% of EVM value are measured. With the output impedance of EDGE, the RMS value of linear output power has been increased from 26.5 dBm to 27.5 dBm and the modulated efficiency has been increased 24.9% to 28.5% at 5% of EVM.
Figure 39. Measured EVMs and efficiencies according to the output impedances with EDGE modulated signal.

Figure 40. Measured output spectrum with EDGE modulated signal at 27.5 dBm of Pout and 1.76 GHz of operation frequency.
Figure 40 shows the effectiveness of the tunable matching network in the EDGE application. In this figure, the output spectra of linear PAs for different output impedances with an EDGE modulated signal at 1.76 GHz of operating frequency and 27.5 dBm of output power. The output spectrum of the PA moves down at the side band and is confined well within the spectrum mask of the EDGE application by applying the tunable matching network due to the improvement of the PA linearity.

5.3. Design Methodology of Two-stage Dual-Mode PA

Only focused on a single stage, the design and measurement of the PA with a tunable matching network at the output are presented. However, for a sufficient gain and output power, RF power amplifiers in wireless communications usually use several stages, such as one or two driver stages and a power stage. There exist inter-stage matching networks between stages to transform the input impedance of the next stage to the output impedance of the previous stage and vice versa. The same as the power stage, the output matching networks of driver stages also have optimum impedances according to the primarily required performances - output power and linearity. Although the impact of inter-stage matching on the PA performance is smaller than that of output matching, the tunability of inter-stage matching can also give additional improvement on the PA performance.

Therefore, a tunable matching network can also be adopted in the inter-stage matching in multi-stage PAs. Inter-stage tunable capacitors have less voltage stresses due to the smaller voltage swings across the capacitor nodes different to the tunable capacitors at the output matching. According to the PA performances, inter-stage
Figure 41. Schematic of fully integrated two-stage GSM/EDGE dual-mode power amplifier with tunable matching networks.
matching networks of power amplifiers also affect the power and linearity of the whole PA.

In this design, a two-stage PA is designed for high-band GSM/EDGE applications. Figure 41 shows the designed two-stage PA with tunable matching networks for both inter-stage and output. A single pair of driver stage amplifies the input signals, and the amplified signals are transferred to two pairs of power stages as input signals. For the drive stage, a differential cascode with 1 mm / 1 mm of CS / CG device is used, and the power stage has the same structure as the previous single stage PA. By adopting the tunable capacitor at the inter-stage matching network, the driver stage can see different optimum output impedances according to the application modes. The structure of the tunable capacitor is the same as that in the output matching, which is shown in Figure 30 (a).

By using a driver stage to drive power to a power stage, an additional advantage can be obtained in improving linearity. As the applied power increases, the gain and phase of the power stage are distorted. The high voltage swings between the terminals of active device make a PA more nonlinear as power increases. In the viewpoint of linearity, a smaller variation in gain and phase according to the output power shows better linearity. The gain and phase variations according to the power can be stretched to higher output power with flatter values by shaping gain and phase variations of the drive stage. This technique has the same concept of the pre-distortion technique for linearity enhancement. The AM-AM and AM-PM characteristics of the whole 2-stage PA are sums of those of the drive and the power stages. In other words, the variation in AM-AM and AM-PM of the power stage can be compensated by the shaping of AM-AM and AM-PM of the
Figure 42. Simulated gain and phase variations according to the gate bias of CS device in the power stage. (a) Gain variation. (b) Phase variation.
driver stage. With setting AM-AM and AM-PM of the driver stage in opposite direction to those of the power stage, higher linearity can be obtained for the two-stage PA.

Without the help of the driver stage by shaping its gain and phase variations, the AM-AM and AM-PM characteristics of entire PA are not good enough for the EDGE application. Figure 42 shows the AM-AM and AM-PM characteristics of PA with gate bias change of the CS device of the power stage. The gain and phase variations are very sensitive to the biasing condition. At low gate voltages close to the threshold voltage, so called deep Class AB condition, gain expansion occurs. At high gate voltage of the CS device close to Class A condition, gain compression occurs. When gain compression starts to occurs, it starts at the lower output power as the gate voltage of the CS device increases. The phase variation becomes smaller monotonically as the gate bias of the CS device increases. Both compression and excessive expansion in the gain worsen the linearity in the viewpoint of AM-AM distortion. On the other hand, the higher gate biasing condition is better from the AM-PM distortion perspective. To get a higher linearity, both AM-AM and AM-PM distortions need to be minimized simultaneously.

In the designed two-stage PA, the input node, inter-stage node and output node are denoted as A, B, and C, as shown in Figure 41. The differences of gain and phase variations between node A and node B represent the AM-AM and AM-PM characteristics of the driver stage, and those between B and C show the gain and phase variations of the power stage, respectively. In this two-stage PA design, to get a higher linearity for the EDGE application, the gain and phase variations of the driver stage are shaped in an opposite way to those of the power stage as shown in Figure 43. Usually to get linearity and efficiency at the same time, the PA core of the power stage is biased at Class AB for
Figure 43. Linearization of two-stage PA with pre-distortion of driver stage. (a) Gain variation. (b) Phase variation.
linear applications. Therefore, the power stage shows some overshoots in both gain and phase as shown in Figure 43. To compensate this gain expansion and phase variation, the output impedance of the driver stage and the bias condition need to be optimized. Linearization with the pre-distortion is conducted to make the gain and phase variations of the driver stage have an opposite shape to those of the power stage as shown in Figure 43. After optimization of the impedance and biasing conditions of the driver stage, the gain and phase variations of the entire two-stage PA become flatter, and the PA can be used for linear applications by with higher linearity.

To satisfy less than 5% of the EVM up to the maximum linear output power of 27.5 dBm for the EDGE application at the high band, both AM-AM and AM-PM characteristics are carefully investigated up to the output power of 31 dBm due to the peak-to-average-ratio (PAPR) of the EDGE application. From the experimental relationship between the EVM in the EDGE modulation and the single tone results, the gain variation and phase variation are controlled less than 0.2 dB and 5 degrees, respectively.

However, by using the optimum output impedance of the driver stage for a linear application, the output power of the PA is reduced due to the loss in the inter-stage matching. Thus, a conjugated matching is performed in the inter-stage matching for the GSM application. By exploiting the tunable capacitor in the inter-stage matching, the two-stage PA can have two different output impedances of the driver stage according to the allocation. To obtain dual mode operation, the switch transistor in the tunable capacitor has 0.6 mm of total gate width, and the designed capacitance values are summarized in Table 7. By turning on the switch, 1.7 pF of capacitance is used for the
GSM application. For the EDGE application, 1.2 pF of capacitance is used for the optimum impedance.

<table>
<thead>
<tr>
<th>Capacitance / design variables</th>
<th>ON state</th>
<th>OFF state</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacitance (@1.8 GHz)</td>
<td>1.7 pF</td>
<td>1.2 pF</td>
</tr>
<tr>
<td>C1, C3</td>
<td></td>
<td>1.75 pF</td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td>0.5 pF</td>
</tr>
<tr>
<td>C4</td>
<td></td>
<td>0.8 pF</td>
</tr>
</tbody>
</table>

5.4. Implementation and Measurement of Two-Stage GSM/EDGE PA

The fully integrated dual-mode GSM/EDGE two-stage PA is implemented with 1P6M 0.18 μm RF CMOS process. A microphotograph of the fabricated chip is shown in Figure 44. The total chip size is $2.1 \times 1.1 \text{ mm}^2$ including the on-chip output transformer and all pads. The implemented two-stage PA includes an input balun to transform a single-ended signal to a differential signal, drive stage, tunable inter-stage matching network, two pairs of power stages, and a tunable output matching network including a $2 \times 1:2$ transformer. The size of the tunable capacitors in the inter-stage and output matching networks is small enough not to increase the total chip area. All measurements are done using chip-on-board (COB) tests. Two-layer FR-4 PCBs are used as the evaluation board, and the losses from board and cables from power supplies are compensated, while that of wire-bonding is included.

The Measurements of the two-stage PA are performed with the same test bench as in Figure 33. The idle current of the power stage for the GSM application is increased to 400 mA for the higher output power, and that for the EDGE application is the same.
current as the single-stage case. For the drive stage bias, different idle currents are used for the higher output power of the GSM application and the higher linearity of the EDGE application. The biasing conditions and capacitance values of the tunable capacitors for the inter-stage matching ($C_{SH}$) and output matching network ($C_{IN}$) are summarized in Table 8 according to the applications.

**Table 8. Bias conditions and values of tunable capacitors of two-stage dual-mode PA.**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>GSM</th>
<th>EDGE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Drive stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{gs,DS}$</td>
<td>0.55 V</td>
<td>0.50 V</td>
</tr>
<tr>
<td>$V_{fg,DS}$</td>
<td>2.3 V</td>
<td>2.3 V</td>
</tr>
<tr>
<td>$I_{Q,DS}$</td>
<td>40 mA</td>
<td>20 mA</td>
</tr>
<tr>
<td><strong>Power stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{gs,PS}$</td>
<td>0.54 V</td>
<td>0.47 V</td>
</tr>
<tr>
<td>$V_{fg,PS}$</td>
<td>2.8 V</td>
<td>2.5 V</td>
</tr>
<tr>
<td>$I_{Q,PS}$</td>
<td>400 mA</td>
<td>140 mA</td>
</tr>
<tr>
<td>$C_{SH}$</td>
<td>1.7 pF</td>
<td>1.2 pF</td>
</tr>
<tr>
<td>$C_{IN}$</td>
<td>4.6 pF</td>
<td>3.5 pF</td>
</tr>
</tbody>
</table>
Figure 45 shows the single-tone measurement results which show the saturated output power and saturated power added efficiency (PAE) with the inter-stage and output matching networks for GSM and the EDGE applications. The GSM biasing conditions are used for both cases. As expected from the single stage PA results, higher saturated $P_{out}$ and higher saturated PAE are obtained with the matching networks optimized for the GSM application. The small-signal gains are almost the same for both cases having a value of 40 dB. Over the frequency range of 1.6 to 2.0 GHz, the saturated output power and the saturated PAE with matching networks for GSM are higher than those for EDGE by 0.2 ~ 0.4 dB and 1.5 ~ 3.5%, respectively. Although the frequencies of maximum output power and efficiency are shifted to lower frequencies than the simulation results, the output power variation is about 0.6 dB over the frequency range from 1.6 to 2 GHz.
and 0.3 dB over the high band frequency (1710 ~ 1910 MHz). For the GSM application, the two-stage power amplifier achieves 32.3 dBm of saturated output power and 41.7% saturated PAE with 10 dBm of input power at 1.76 GHz.

The measurement results for the EDGE applications are shown in Figure 46. The designed two-stage PA has 30 dB of small-signal gain and 4.7% of EVM and 27% of PAE at an output power of 27.5 dBm. The consumed DC currents are 20 mA and 575 mA for the drive and power stages, respectively. Because of the pre-distortion of the drive stage, the EVM decreases by 0.3% at 27.5 dBm of output power. More improvement is achieved at higher output power, for example 1.5% EVM improvement at 28.3 dBm of output power.

Figure 46. Measured gain, EVM, and PAE of two-stage GSM/EDGE dual-mode PA according to the output power for the EDGE application.
Figure 47. Output spectra of two-stage GSM/EDGE dual-mode PA. (a) Output spectrum of GSM modulation signal. (b) Output spectrum of EDGE modulation signal.
The output spectra of the two-stage dual-mode PA are shown in Figure 47. The amplified GMSK modulated signal at 32.3 dBm of the maximum output power and GSM spectrum emission mask are shown in Figure 47 (a), and the output spectrum is well confined in the GSM spectral emission mask over the entire output power range. Figure 47 (b) shows the amplified EDGE signal at 27.5 dBm of output power and the EDGE spectrum emission mask. The output spectrum is also well confined in the spectrum mask and shows good linearity performance for the EDGE application.

5.5. Linearity Improvement with Pre-distortion

Tunable matching networks at the inter-stage and output enable benefits of having high output power and efficiency of the GSM application and high linearity of the EDGE application. In addition to the advantages of the tunable matching network, the pre-distorted drive stage gives additional improvement in PA linearity for linear applications. To verify the effectiveness of the tunable matching network and the distortion of the drive stage at the two-stage PA, a linearity comparison is performed for the three cases: (1) a single stage PA with an optimum output matching network for the GSM application, (2) a single stage PA with an optimum output matching network for the EDGE application, and (3) a two-stage PA with optimum inter-stage and output matching networks and pre-distortion of the drive stage for the EDGE application. The first case is a conventional one-stage PA with fixed output matching for GSM, and the second case is a one-stage PA with output matching tuned for the EDGE application. The third case is two-stage PA fully optimized for the EDGE application with a pre-distorted drive stage and tunable inter-stage and output matching networks.
Figure 48. Comparison of EVM characteristics according to the matching networks of single and two-stage PAs.

Figure 48 shows the EVM vs. output power curves of the three cases with an EDGE modulated input signal. As the output power goes up, the EVMs of PAs increase sharply due to the nonlinearities in AM-AM and AM-PM. In a linear PA application, the maximum linear output power is usually defined as the maximum output power that satisfies the linearity specs, such as IMD, EVM, ACPR, ACLR, etc. In this measurement, the maximum linear output power can be defined as the output power when the EVM reaches the value of 5% for the EDGE application. In other words, for the high band the EDGE application, the PA is required to have less than 5% of EVM value at an output power of 27.5 dBm. Case (1) shows the lowest maximum linear output power of 26.5 dBm, and the maximum linear output power increases to 27.5 dBm and 27.8 dBm for cases (2) and (3), respectively. At a fixed output power of 27.5 dBm, the EVM values are
7.3%, 5%, 4.7% for cases (a), (b), and (c), respectively. As the output power goes up more, the improvement increases as shown in Figure 48. From the EVM comparison of the three cases, the in-band linearity improves significantly.

Figure 49 shows other evidences for the improvement in linearity. The output spectra of the amplified EDGE signal are compared for the three cases. From the comparison of the output spectrum of case (1) and case (2), the side-band output signal of case (2) is reduced due to the optimum impedance for the EDGE application, as expected. By using a pre-distorted drive stage of the two-stage PA, the output spectrum shows more margin from the emission mask. This comparison reveals the advantage of the improvement in out-of-band linearity.
The results of the linearity comparison in both in-band and out-of-band show the effectiveness of tunable matching and pre-distortion of the drive stage on the PA linearity without hurting the GSM performances.

5.6. Conclusion

In this chapter, the implementation and measurement results of the fully integrated GSM/EDGE dual-mode PA were described. To verify the effectiveness of the tunable matching network in the output matching, a single stage PA is designed and implemented in 0.18 µm RF CMOS process. The tunable output matching network is implemented with an on-chip 2×1:2 PCT and a high-power tunable capacitor that can give two different capacitance values. The implemented fully integrated single stage PA shows 32 dBm of maximum P_{out} and 45% of drain efficiency with the output matching network for the GSM application. For the EDGE application, by changing the output matching optimized for the EDGE application, the P_{1dB} and corresponding drain efficiency increase from 29.5 dBm to 30.5 dBm and 33% to 40%, respectively. With applying an EDGE modulated signal, 27.5 dBm of maximum linear P_{out} and 28.1% of linear drain efficiency are achieved at 5% of EVM. The improvement is 1 dB and 3.6% of maximum linear P_{out} and linear efficiency, respectively. These improvements show the effectiveness of the tunable output matching network, which can give more optimized output impedances according to application modes.

Tunable matching is also applied in the inter-stage matching network in the two-stage dual-mode PA. The designed PA also adopts pre-distortion in the drive stage to have additional linearity. The maximum P_{out} and PAE are 32.3 dBm and 41.7%, respectively,
for the GSM application. For the EDGE application, 4.7% of EVM and 27% of PAE are achieved at an output power of 27.5 dBm. The comparison of the EVM and the output spectrum with applying an EDGE signal also shows the improvement of in-band and out-of-band linearity. The implemented two-stage PA shows additional linearity improvement for the EDGE application without hurting GSM performances. These results show the usefulness of the tunable matching and the effectiveness of pre-distorted drive stage for the multi-standard wireless communications.
CHAPTER 6
A TRIPLE-MODE POWER AMPLIFIER FOR
GSM/EDGE/WCDMA APPLICATIONS

6.1 Introduction

As described in the previous chapters, the requirements of output power and linearity are different according to the applications in a multi-mode PA. For example, the GSM application requires a PA with high output power but less linearity while the EDGE application needs a PA with higher linearity and intermediate output power. For those applications, the dual-mode PA is introduced and described with impedance modulation by virtue of tunable matching and a pre-distorted drive stage in a multi-stage PA to get high output power and high linearity according to the applications.

However, as modern communications grow and support more services including voice, video communication, multimedia functions and internet access, the needs of high data-rate communications are continuously demanded. To support those higher data-rate communications, power amplifiers are also required to have more and more linearity. To achieve more linearity for linear applications and sustain power capability for the application with constant envelope signals, the design of a multi-mode PA is introduced in this chapter.

To get a higher PA linearity, the structure of the PA core is revisited. The PA cores utilize triple-well devices in the RF CMOS process. The triple-well devices, having deep n-wells (DNWs), have some advantages in RF applications. They can give an option in
the connection of the body terminal of the CG device when the cascode topology is adopted for the PA core. One option is that the body of the CG device can be connected with its source terminal (the drain terminal of the CS device), so called BS-type cascode. The other option is that BG-type cascode whose body terminal of the CG device is connected to the source terminal of the CS device, which is usually connected to the ground. The DNW devices also show better device isolation and device linearity compared to those without DNW [9].

From the previous results, the PA core shows some different characteristics in terms of power capability and linearity according to the connection of the body terminal of the CG device in cascode. Because the cascode topology with triple-well devices is adopted, the PA core can have some advantages if it is implemented both types of connection with a single PA core.

In this chapter, in addition to the techniques described previously, a body-switching technique is introduced to achieve both two BS and BG types of PA core with a single PA. The effectiveness of implementing body-switching is discussed by comparing the individual BS and BG types of PA core. This PA core with body-switching technique is used in the GSM/EDGE/WCDMA triple-mode PA, and the implemented triple-mode PA is measured and discussed in the viewpoint of linearity.

6.2. Design of Multi-mode PA with Body-switched PA Core

In Chapter 3, unit cells with the BS-cascode type and the BG-cascode type are implemented and evaluated through the power measurement and load-pull measurements. For the power capability, the BS-type cascode shows better performance while the BG
type cascode shows better linearity performance. To achieve high output power of a dual-mode PA for the GSM application, the BS-type cascode is utilized in the previous chapters.

6.2.1. Body-switched PA core

For the GSM/EDGE/WCDMA triple-mode PA, the PA core is redesigned by adopting the body-switching technique. The designed PA core can use the BS-type cascode for the GSM application to guarantee high output power and utilize the BG-type cascode for linear applications. Figure 50 shows the simplified schematic of the designed PA core with the body-switching technique. By adding switches among the body terminal of the CG device, the source terminal of the CG device and the source terminal of the CS device, the PA core can switch the body connection. The switch between the body terminal and the source terminal of the CG device, SW_{BS}, is turned on, and the switch...
between the body terminal of CG device and the source terminal of CS device, SW\textsubscript{BG}, is turned off to implement BS-type cascode. For the realization of the BG-type cascode, the states of the switches are toggled from those of the BS-cascode type. The switches are implemented with thick-gate-oxide devices.

However, conventional BS- and BG-cascodes use direct connections between the appropriate terminals with low-impedance connection, which are usually metal connections. The designed body-switched cascode should have finite on- and off-resistances and capacitances due to the switches. To take advantage of both cascode types fully in PA operation, the impacts of finite resistances and capacitances should be considered.

The switch can be modeled as a parallel connection of a resistor and a capacitor. When the switch is on, the resistor has a value of on-resistance, which is a channel resistance of the transistor in the triode region. The on-resistance is inversely proportional to the total gate width of the switch. When the switch is off, the resistor has a very high resistance because there is no current path in the ideal off-condition. Because the resistor has high resistance at the gate of the switch, the capacitance value of the switch can be express as

\[ C_{SW} \approx \left( C_{gs} \parallel C_{gd} \right) + C_{para} \tag{13} \]

where \( C_{gs} \) and \( C_{gd} \) are the gate-source capacitance and gate-drain capacitance, and \( C_{para} \) is the parasitic capacitance between the drain and the source thorough the body of the switch. When the switch is on, \( C_{gs} \) and \( C_{gd} \) are composed of oxide capacitances and overlap capacitances while they are composed of overlap capacitances when the switch is off. Both on- and off-capacitances are proportional to the total gate width of the switch.
Figure 51 shows the simulation results of on- and off-resistances and on- and off-capacitances of the switch at an operation frequency of 1.8 GHz. From the S-parameter simulation of the switch with several hundred μm of total gate width at the high band frequency, the on-resistance of the switch is a few Ohms, and the off-resistance is over 10kΩ. The on-capacitance and off-capacitance are on the order of pico-Farads.

The states of the body switches are toggled to realize the BS- and BG-cascodes. When the switch is on, the on-resistance is important. Because the on-resistance is parallel with the on-capacitance, the on-capacitance has much higher impedance, and the on-resistance gives dominant effect on the characteristics of the connection. However, as opposed to the conventional BG- and BS-type cascodes, the cascode with the body switches has a common node at the body terminal of the CG device. Thus, the connection
through the other switch, which is off, must also be considered. As in the ON state, the resistor and capacitor are parallel when the switch is off. As shown in Figure 51, because the off-resistance is high enough compared to the impedance of the off-capacitance, the off-capacitance has an important role in the off connection. As the on-resistance and off-capacitance both become smaller, the characteristics of the body-switched cascode are closer to those of conventional direct connected cascodes. However, the on-resistance becomes smaller with a large size of the switch, but the off-capacitance becomes larger. Therefore, the size of the switches needs to be optimized so that the cascode with the body switches can have similar characteristics to the BS- and BG-cascode at the corresponding states.

6.2.2. Comparison of PA cores with direct connection and body switches

To verify the effectiveness of the body-switched cascode PA core, the conventional two-stage PAs with the direct-connected BS- and BG-cascode PA cores and the two-stage PA with the body-switched cascode PA core are compared. Both switches for the BS-cascode ($SW_{BS}$) and the BG-cascode ($SW_{BG}$) use 600 $\mu$m of total gate width from the optimization results. The designed body switch network is inserted in the PA design, and the PA with body switches is evaluated through single tone simulation.

Figure 52 shows the simulated gain variation and PAE results. Comparisons are conducted in both BS and BG cases. The gain variation and PAE curves of the BS-type cascode are shown in Figure 52 (a), and those of the BG-type cascode are shown in Figure 52 (b), respectively. The corresponding curves are almost identical for both PAs with the BS-type cascode and the BG-type cascode at different bias conditions. These comparisons are good evidence that the body-switched cascode can work well as both a
Figure 52. Simulated PA characteristics with conventional direct connected cascodes and body-switched cascode. (a) Comparison between conventional BS-cascode and BS-cascode with body switches. (b) Comparison between conventional BG-cascode and BG-cascode with body switches.
BS- and BG-cascode. Even though there exist some finite resistances and capacitances in the body connections, the body switches do not affect the performance of the PAs.

The voltages between gate and source/drain terminals, source and drain terminals of the switch are also checked. If the voltages exceed some limits for proper operation of the switch, the switches cause an additional nonlinearity of the PA by unwanted turn-on during some time in the period. Over 33 dBm of output power, it is confirmed that the switches work properly without any additional distortion. In addition to the switch operation check, two-tone simulations are also conducted to compare the linearity of the body-switched cascode PA with direct-connected cascode PA. These simulation results also give a confirmation of no additional nonlinearity due to the switches in the body.

Figure 53. Power characteristics of body-switch cascode PA for BS and BG connection with similar gain variation and PAE 1.8 GHz.
With the body-switched cascode PA core, two PAs with BS- and BG cascode PA cores with single PA core can be implemented. To compare the linearity of the two PAs with different types of PA cores, BS-type and BG-type, two-tone simulations are conducted to get the IMD3. For a fair comparison of linearity, BS- and BG-type PAs are set to have identical gain variation and PAE. As shown in Figure 53, the AM-AMs and PAEs of the BS-type cascode and BG-type cascode PA are almost identical. The two PA modes show almost the same amount of gain overshoot and $P_{1\text{dB}}$ with similar PAE.

Figure 54 shows the simulation results of linearity comparison between BS- and BG-types with the body-switched cascode PA at similar power characteristics as shown in Figure 53. The BG-type PA shows better linearity than the BS-type PA in terms of IMD3,
although the power characteristics are almost identical to each other. These results coincide with the results from the linearity comparison of unit power cells in Chapter 3.

6.3. Implementation and Measurements of GSM/EDGE/WCDMA PA

From the results above, the single PA can work as BS- and BG-types without performance variation by adopting body switches in the PA core. A high-band PA with a body-switched PA core is designed for GSM/EDGE/WCDMA applications. Higher power capability of the BS-cascode PA is used for the GSM application and higher linearity of the BG cascode PA is used for EDGE and WCDMA PAs. Most of the sub-blocks in the 2-stage dual-mode PA are reused in this design, but the PA core is modified for the implementation of the body switch networks. The body-switched PA core is utilized in the power stage only to avoid some complexity in the analysis. The pre-distorted drive stage and impedance modulations in both inter-stage and output matching are also applied in this PA.

A fully integrated triple-mode GSM/EDGE/WCDMA two-stage PA with a PA core utilizing body switches is implemented in 0.18 \( \mu \)m RF CMOS process to verify the usefulness of the body-switched cascode PA core. Figure 55 shows the microphotograph of the fabricated fully integrated GSM/EDGE/WCDMA PA chip on the evaluation board. The total chip size is 2.1 x 1.1 mm\(^2\).

6.3.1. GSM performance of body-switched PA

Measurements of the triple-mode PA are done with the same test-bench previously shown in Figure 33. The bias condition for the GSM application is the same as that for
the GSM application in the 2-stage PA measurements. The body switches are set as the
BS-type cascode PA. Figure 56 shows the results of single-tone measurements for the
GSM application. The frequencies of the maximum output power and maximum
efficiency are shifted to a higher frequency compared to the results of the previous
chapter. The output power is greater than 32 dBm over the high frequency band (1710 ~
1910 MHz) with less than 0.1 dB of variation. The maximum output power is 32.1 dBm
with 39.6% of PAE with 7dBm of input power at the frequency of 1.8 GHz. These results
are similar to the GSM results with the two-stage GSM/EDGE PA in chapter 6. These
results show the BS-type cascode PA with body switches works well and shows similar
results to direct-connected BS-type cascode PA for the GSM application.

Figure 55. Microphotograph of fully integrated GSM/EDGE/WCDMA PA chip on evaluation board.
6.3.2. Verification of the body-switch PA core

To verify the simulation results of the linearity comparison according to the body connection, the body-switched cascode PA is biased to get identical gain variation according to the output. Because the current level of the power stage is different according to the body connection, different bias conditions are applied for the BS- and BG-type cascode PAs. All other biases are the same as the bias condition of the EDGE application as shown in Table 8, but the gate bias of the CS in the power stage is controlled to reduce the effect of the bias change. For the BS connection, the bias condition is the same as that of the two-stage dual-mode PA for the EDGE application as shown in Table 8. The gate of the CS device in the power stage is biased at a voltage 12
mV high for the BG-type cascode PA to get identical gain variation to the BS-type cascode PA. Figure 57 shows the measured gain variation and PAE versus output power according to the body connection of the PA core in the power stage. By changing the gate bias for the BG connection, the gain variation curves are very close up to 28 dBm of output power.

Figure 58 shows the measured EVM results according to the body connection with an EDGE modulated signal at the conditions of similar gain variation. The EVM values are close enough to each other below 27.5 dBm of output power, but the BG-type cascode PA shows better EVM at over 27.5 dBm because of better AM-AM characteristics of the BG-type cascode above 28 dBm of output power.
Figure 58. Measured EVM of body-switched cascode PA according to the body connection.

Figure 59. Measured IMD3 of body-switched cascode PA according to the body connection.
Figure 60. Measured small-signal S-parameters for linear PA application. (a) $S_{11}$ and $S_{22}$. (b) $S_{21}$ and $S_{12}$.
From the two-tone test, the measured IMD3 curves versus output power according to the body connection are shown in Figure 59. The center frequency is 1.8 GHz with 5 MHz of tone spacing. Contrary to the EVM measurements, the measured IMD3 of the BG-type cascode PA shows better linearity through the output power except around sweet spot region. The sweet spot of BG-type cascode PA occurs at a slightly higher output power than the BS-type cascode PA. Although the gain curves of the BS- and the BG-type cascode PAs are very similar to each other, the BG cascode PA shows better linearity in terms of inter-modulation distortion. This is well matched with the simulation results described previously. The BG cascode PA shows better IMD characteristics.

Figure 60 (a) and (b) show the measured S-parameters of the BG-type cascode PA at the bias condition of the EDGE application. The measured $S_{11}$ and $S_{22}$ show good input
and output matching, respectively. At the frequency of 1.8 GHz, -10 dB of $S_{11}$ and over -20 dB of $S_{22}$ are achieved. Figure 60 (b) shows the small-signal gain and a reverse isolation. At the frequency range of interest, the measured $S_{21}$ shows around 30 dB of small-signal gain. The measured $S_{12}$ shows over -40 dB over the frequency range of 1 to 4 GHz indicating good reverse isolation. From the measured S-parameters versus frequency, the PA shows good matching at inter-stage as well as input and output.

6.3.3. Body-switched PA in the EDGE application

With an EDGE modulated signal at the frequency of 1.8 GHz, the measured gain, PAE, and EVM of the BG-type cascode PA are shown in Figure 61. The BG-type cascode PA gives 31 dB of the small-signal gain. At the output power of 27.5 dBm, 4.2% of EVM and 26% of PAE are achieved. By adopting the switched-body PA, the linear output power is slightly increased with the BG-type cascode PA compared to that with the BS-type cascode PA for the EDGE application.

6.3.4. Body-switched PA in the WCDMA application

In linear PA applications, the out-of-band linearity with an applied modulated signal is usually expressed as a ratio between the main channel power and the adjacent channel power, such as adjacent channel power ratio (ACPR) and adjacent channel leakage ratio (ACLR). In the WCDMA application, the out-of-band linearity for the system is required to have less than -33 dBc of ACLR at the first next channel from the main channel (ACLR1) and less than -43 dBc of ACLR at the second next channel from the main channel (ACLR2). According to the relative position of the adjacent channel to the main channel, ACLR has two values, high (upper) and low (lower).
Figure 62 shows the measured ACLR versus output power of the body-switched PA with a WCDMA modulated signal according to the body connection. As with the results of IMD3, the BG cascode type PA shows better ACLR values over the output power except around the sweet spots of the BS-type cascode PA. The location of the sweet spots of the BG-type cascode PA is a little higher than that of the BS-type cascode PA. The BS-type cascode PA has a maximum linear output power of 26.8 dBm, meeting the ACLR1 specification, while the ACLR1 of the BG-type cascode PA is less than -33dBc up to 27.5 dBm of output power. Between the two values of ACLR, lower ACLR and upper ACLR, the worst one is picked up as a representative ACLR value at that output power. The linear PAEs are 23.3% and 26% for BS- and BG-type cascode PA, respectively.
The usefulness of the designed body-switched PA for the WCDMA application is also shown in Figure 63. This figure shows the WCDMA spectrum emission mask and the output spectra of the PA with body switches according to the body connection with a WCDMA modulated signal at 1.8 GHz of operating frequency and 27.5 dBm of output power. The spectral regrowth at the sideband is reduced by adopting the BG-type cascode PA with body switch, and the output spectrum can be confined within the spectrum emission mask up to 27.5 dBm of output power.

**6.4. Conclusion**

In this chapter, the design, implementation, and measurement results of the fully integrated GSM/EDGE/WCDMA triple-mode PA are described. By adopting a body-
switched PA, which uses two switches at the body of the CG device, it is possible to implement both the BS-type cascode PA and the BG-type cascode PA with a single PA.

The effectiveness of the body-switched PA is shown through the linearity improvement for both in-band and out-of-band linearity without hurting the output power capability. The BS-type cascode PA is used for the GSM application for higher output power and the BG-type cascode is used for linear PA applications, EDGE and WCDMA.
CHAPTER 7
CONCLUSION AND FUTURE WORK

7.1. Technical Contributions and Impacts of the Dissertation

The growth wireless communication markets has been tremendously explosive with advancements in semiconductor technology and the wireless device industry over the last few decades. With the development of wireless communications, costumers demand higher data-rate communications with more functionality. This increasing demand brings multiple wireless communication standards into being. Thus, today’s wireless communication devices should be able to support not only several wireless standards such as GSM, EDGE, WCDMA, LTE, and so on but also some non-cellular standards such as WLAN and WiMax etc. However, these standards cannot be support by a single power amplifier, but instead would require several power amplifiers because of the different output power and linearity requirements according to the communication standards. Thus, communication device makers have tried to reduce the number of PAs in order to reduce cost.

In addition to the multiple standards issue, the power amplifier market has been dominated by GaAs-based electronic devices even though Si-based CMOS dominates the transceiver market with a single chip. Thus, usually power amplifiers are sold as front-end modules that include a PA, an RF switch, and control chips. However, if CMOS PAs are available with sufficient performance for modern wireless communication standards, high integration is possible not only in the PA chip but also in the RF communication
chip. Therefore, because of low-cost and the capability of high integration of CMOS technology, there exist lots of activities and advancement of CMOS RF PAs not only in academia but also in the wireless market.

For the purpose of achieving CMOS RF PAs capable of supporting multiple wireless standards, this research shows both theoretical contributions and successful implementations of fully-integrated CMOS RF PAs. The achievements can be summarized as follows.

- The Performance of power cell with a CMOS $\chi_\sigma\chi_\delta\chi_\delta\epsilon$ has been analyzed comprehensively and experimentally according to the body connection of the common-gate device in terms of power capability, efficiency and linearity for PA applications. According to the design parameters, such as body connection type, the ratio and the bias conditions of CS and CG devices, optimum conditions have been achieved for both high-power applications and high-linearity applications through optimization. This analysis can give a guideline to designing of CMOS PA with cascode type power cell.

- A high-power tunable matching has been developed for watt-level PA applications. With the tunable matching network at the output, a load modulation technique has been introduced for dual-mode PA applications with a single PA. A tunable output matching network with the transformer of the PA has been achieved by using a high-power and highly linear tunable capacitor for GSM/EDGE dual-mode applications.

- A fully-integrated single stage GSM/EDGE dual-mode PA has been successfully implemented in 0.18 $\mu$m RF CMOS technology with a tunable matching output
matching network, achieving 32 dBm of output power and 45% of drain efficiency for the GSM application and 27.5 dBm of linear output power and 28.5% of linear drain efficiency for the EDGE application at 1.76 GHz, respectively. With the aid of tunable matching, $P_{1\text{dB}}$ has been increased from 29.5 dBm to 30.5 dBm, and improvements of 1 dB in linear output power and 3.6% in linear efficiency have been achieved.

- A fully-integrated two-stage GSM/EDGE dual-mode PA with a pre-distorted drive stage and tunable matching networks at both inter-stage matching and output matching has been demonstrated using 0.18 $\mu$m RF CMOS technology. At 1.76 GHz, a maximum output power of 32.3 dBm and PAE of 41.7% for the GSM application, and 4.7% of EVM and 27% of PAE at 27.5 dBm of output power for the EDGE application have been achieved, respectively. The proposed pre-distortion drive stage has given an additional linearity improvement.

- A fully-integrated GSM/EDGE/WCDMA triple-mode PA with a body-switched PA core has been successfully implemented in 0.18 $\mu$m RF CMOS technology. The body-switched PA can work as the BS-cascode type PA for the GSM application and the BG-type cascode PA for EDGE and the WCDMA applications. With the body switching technique, 27.5 dBm of output power and 26% of PAE have been obtained at -33 dBc of ACLR1 for the WCDMA application. Without degradation of performance in the GSM application, the linear output power and the linear PAE have been increased by 0.7 dB and 2.7%, respectively.
7.2. Scope of Future Research

In this dissertation, the research focuses on a single CMOS PA to support multiple wireless communication standards. Even with achievements in PA design from the proposed techniques, challenges still remain to implement a fully integrated RF front-end.

First of all, additional research and development of PAs for multi-band applications is required to reduce the number of PAs in the communication devices. For multi-band approaches, tunable matching networks in terms of the operation frequency are required. Thus, future research should be focused on devising a smart way to improve the frequency tunability with some tunable passive components including tunable inductors and tunable transformers without sacrificing the efficiency and linearity performance.

Second, additional efforts are needed to overcome inherent inferiorities in CMOS RF PAs to counterparts of compound semiconductors for more improvement in output power, efficiency, and linearity. Future communication technology needs a more and more linear power amplifier for higher data-rate communications. It is very difficult for a CMOS PA to meet efficiency and linearity requirements simultaneously due to the limitations of the CMOS technology. Therefore, future research should be focused on developing a smart way to improve the linearity without sacrificing the efficiency performance.

Finally, for a single chip transceiver with low cost and high integration, the PA should be effectively isolated from other circuitries, not only sub-blocks in the transmitter chain but also in the receiver chain. The large voltage swing of a PA can be coupled to adjacent circuitries and cause saturation and/or malfunction of the adjacent blocks. This effect is relatively minimized in the SOI technology because of the high-resistive
substrate. However, in bulk CMOS technology, it should be carefully considered to keep the coupling from affecting other circuits.
REFERENCES


[33] P. Haldi, D. Chowdhury, L. Gang, and A. M. Niknejad, "A 5.8 GHz Linear Power Amplifier in a Standard 90nm CMOS Process using a 1V Power Supply," in


VITA

Hyungwook Kim was born in Daegu, Korea, in 1970. He received the B.S. degree in physics education and M.S. degree in electronics engineering from Seoul National University, Seoul, Korea, in 1994 and 1996, respectively. He is currently working toward the Ph.D. degree in electrical and computer engineering at the Georgia Institute of Technology, Atlanta.

From 1996 to 2003, he was with LG Electronics, Seoul, Korea, where he was involved in the design and fabrication of GaAs-based electronic devices and the development of power amplifier for CDMA application. In 2010 he has held internship position at Qualcomm Incorporated, San Diego, CA. His research interests include RF CMOS device modeling, CMOS RF transceivers and power amplifier design for wireless communication.