KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)

CMOS Operational and RF Power Amplifiers for Mobile Communications

Promotor:
Prof. Dr. ir. M. Steyaert

Thesis submitted to obtain the degree of doctor in applied sciences

by

João Ramos

March 2005

KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)

CMOS Operational and RF Power Amplifiers for Mobile Communications

Jury:
Prof. Dr. ir. L. Froyen, Chairman
Prof. Dr. ir. M. Steyaert, Promotor
Prof. Dr. ir. W. Sansen
Prof. Dr. ir. B. Nauwelaers
Prof. Dr. ir. W. Dehaene
Prof. Dr. Eng. Mª. Rosário (IST, Portugal)
Ir. L. Cloetens (ST Microelectronics)

Thesis submitted to obtain the degree of doctor in applied sciences

by

João Ramos

UDC: 621.3.049.77

March 2005

D/2005/7515/23

ISBN 90-5682-588-7

Acknowledgments

On the 17^{th} August 2000 I arrived in Belgium to start what would become a unique experience in my life. In a lot of aspects this Ph.D. was the beginning of a new era – ranging from the enriching experience of living in the multi-cultural city that Leuven is, to the possibility of doing research in a renowned international center for analog circuit design. In particular, I would like to thank:

- Prof. M. Steyaert, my promotor, for his support during my stay in Leuven, by providing the conditions that made this work possible and always having a motivating attitude even during difficult times. I'm particular thankful for the freedom of research that I have enjoyed under his supervision, a freedom that I never had before, nor will ever have again.
- Prof. W. Sansen, head of MICAS. His hard work and excellent qualities in passing on his knowledge have contributed to position the group on an international level.
- W. Sansen and B. Nauwelaers, members of the reading committee for their suggestions, feedback and corrections that helped improve the quality of the text.
- Prof. W. Dehaene, Prof. Mª. Rosário and Ir. L. Cloetens members of the jury and Prof. L. Froyen, chairman of the jury, for their valuable time spent on the evaluation of this research work.
- My colleagues with whom I have shared office 91.21. Kenneth Francken, Xiaohong Peng, Marian Verhelst and Yves Vanderperren. I'm thankful for you having made that period a pleasant time, most of the times! ;-)
- Kenneth Francken and Libin Yao for their help during the defense preparation and for organizing all the paper work that allows me to defend this thesis.
- Yann Zinzius, Libin Yao, Jurgen Deveugele, Philippe Coppejans, Tim Piessens and Bram De Muer for all the discussions concerning analog circuit design, measurements and manuscript preparation. A special thanks to Erik Lauwers for his guidance.
- Those from different technical backgrounds, with whom during this research work I have shared ideas: Kenneth Francken from ESAT-MICAS, Guilherme Lujan and Augusto Redolfi from IMEC, Samuel Xavier-de-Souza from ESAT-SISTA and Manuel Yarleque from ESAT-TELEMIC. Their experience, comments and fruitful discussions have undoubtedly contributed to the success of this doctorate.

- The skilled technical team at MICAS: Elvira Wouters, Ben Geeraerts, Danielle Vermetten, Chris Mertens, Chris Van Grieken, Frederik Daenen, Albert Boon, Noëlla Gaethofs, Michel De Cooman, Viviane Muls, Tony Van Nuland, Piet Vanderwegen. Their dedication and cooperation has allowed me to concentrate only on doing research.
- The ESAT computer administrator team: Frank Schoeters, Piet Hendriks, Marc Mangelschots, Stef Verbruggen, Rik Theys and Eric Olemans, deserve credit for ESAT always having a running computer infrastructure, even during weekends.
- An De Ruyck, Luciana Baracchini and Stijn Speelman for being as they are.
- The Brazilian community, Karin Fischer and Nacho Lopezmalo together with persons from all over the world, has made this experience memorable. Their company at Alma or Pangaea, in going to the cinema or in their houses is not forgotten.
- Paulo Bernardino, Paulo Santos and Augusto Marques for their friendship and for being there whenever needed.
- The *Fundação para a Ciência e a Tecnologia*, Portugal, for the financial support that made this research work possible.
- Most of all, I thank you that have given me unconditional love that to me is the definition of family. God, to whom I have prayed and the One that was always present.

The technical results are summarized in this book that you are about to read. The human experience will always be in my mind as the result of this singular experience. I thank you all!

João Ramos

Leuven, March 2005

Abstract

In this era where the desire to always be within reach of a communication device is larger than ever, there is a growing need for the development of circuits that make the dream of this Global Village a reality. Mobility, while remaining reachable at all times, is possible with the use of wireless communications that have the common characteristic of having a limited power source in the battery. This asks for circuit techniques that increase the portable electronic device autonomy. A higher autonomy makes a product more usable, lighter (because a smaller battery is necessary), and as such more appealing to the consumer. Systems with a higher functionality, higher capacity and a slick design are only possible if power consumption is reduced to the minimum. Furthermore, new products are developed on a regular basis. As a consequence, system integrators and everyone working in such a competitive environment see their time-to-market as a small window of opportunity. All design steps until the product reaches the shop shelves have to be optimized. Usually, the first steps in the engineering development require a large percentage of this time, which has to be minimized to keep the prices low.

The presented work covers the modeling, analysis, design, optimization and IC characterization of two different types of amplifiers in CMOS suitable to be integrated on a transceiver for mobile communications.

The design of a low-power, low-voltage and high efficiency three stage operational amplifier is covered in the first part. A new frequency compensation topology is presented and analyzed in detail in terms of some of the most common characteristics. Furthermore, comparison with other frequency compensation topologies, and the impact of different design characteristics (compensation and load capacitor, optimization emphasis, etc.) on the amplifier performance is presented. This comparison makes it possible to see the performance of each compensation topology without the influence of external design factors by comparing all topologies under the same conditions. Subsequently, the effect of the positioning of the poles and its impact on the power consumption is addressed. Finally, the proposed frequency compensation topology is optimized using a computer design automation approach, simultaneously at the block and transistor level, resulting in considerable power savings.

The second part details the design of a power amplifier for the GSM-850 standard. First, the class E power amplifier design methodology is covered. The simplified state-space model and the automated sizing allow to obtain a simple architectural representation and a highly accurate description of the RF behavior. As a result, they are used to gain insight into the circuit performances where precise equations including circuit parasitics are lacking. Secondly, the possibility to increase the supply voltage beyond

the stated maximum supply voltage of standard CMOS low-voltage technologies is explored, with the objective of increasing the efficiency. And finally, a 850-MHz, 30 dBm class E power amplifier designed in standard CMOS technology is presented. Careful analysis and inclusion of all circuit and test-board parasitics in combination with the automated sizing methodology result in a high efficiency amplifier sized in less than one hour of CPU time. Also, simulation results are in very good agreement with measurements. As a consequence of this work, for a similar output power level and frequency of operation, the designed PA shows the highest efficiency for a circuit implemented in CMOS. Furthermore, this high efficiency is already achieved from low output power levels which is especially important as the power amplifier is thus not always working at the maximum power level.

The two different types of amplifiers described in detail in the following pages, were studied, optimized, designed, manufactured, and measured while simultaneously aiming at power optimization, reduced design cycle and cost saving. The two high-performance amplifiers designed in commercial 0.35 µm CMOS technology have the following summarized characteristics:

- Operational amplifier: with a power consumption of 275 µW and a load of 130 pF//24 kΩ, it achieves a Unity Gain Frequency (UGF) of 2.7 MHz with a Phase Margin (PM) of 52°. An average slew rate of 1.0 V/µs is measured. The positive and negative settling time to 1 % final settling time error is 1.4 µs and 1.0 µs, respectively.
- Class E power amplifier: a maximum Power Added Efficiency (PAE) of 66 % is measured. At a supply voltage of 2.26 V, a maximum output power of 955 mW is obtained for an input signal with a frequency equal to 855 MHz. It has a high efficiency over a broad range of output power levels and the PAE is always greater than 60 % for values above 158 mW.

Contents

Acknowledgments i

Abstract iii

Contents v

List of Figures xi

List of Tables xvii

Symbols and Abbreviations xix

1 Introduction 1

1.1 Motivation and Applications 1

1.1.1 Cost Reduction 2

1.2 Research Work 3

1.3 Outline of the Work 4

2 Positive Feedback Frequency Compensation for Three Stage Amplifier 7

2.1 Introduction 7

2.2 Operational Amplifiers 8

2.2.1 Single Stage Amplifiers 8

2.2.2 Two Stage Amplifiers 8

2.2.3 Three Stage Amplifiers 9

2.3 The NMC Structure 11

2.4 Proposed Compensation 14

2.4.1 Working Principle 14

2.4.2 Small-Signal Transfer Function 15

2.4.3 Unity Gain Frequency and Phase Margin 18

2.4.4 Closed Loop Transfer Function 19

2.4.5 Stability Analysis 20
2.4.6 Design Equations 20
2.4.7 Power Optimization 21
2.5 Design of the Operational Amplifier 22
2.5.1 Circuit Implementation and Layout 22
2.5.2 Test Setup 23
2.5.3 Measured Results 23
2.5.4 Discussion of the Results 24
2.6 Conclusion 24

3 Topology Study of Three Stage Amplifier Frequency Compensation 29
3.1 Introduction 29
3.2 Comparison to other Topologies 30
3.2.1 Topology Comparison and Power Estimation 30
3.2.2 Unity Gain Frequency 31
3.2.2.1 NMC 31
3.2.2.2 DFCFC 31
3.2.2.3 AFFC 32
3.2.2.4 TCFC 33
3.2.2.5 NMCF 34
3.2.2.6 PFC 35
3.2.2.7 Simulation Results 35
3.2.3 Settling Time 36
3.2.4 Final Comments 40
3.3 Optimization 40
3.3.1 Block Level: Comparison to the NMC 40
3.3.1.1 Optimization Method 41
3.3.1.2 Frequency Response 41
3.3.1.3 Slew Rate 42
3.3.1.4 Settling Time 42
3.3.2 Transistor Level 47
3.4 Conclusion 52

4 RF Power Amplifier Classification, Theory and Optimization 53
4.1 Introduction . . . 53
4.2 Definition of Performance Metrics . . . 54
4.3 RF Power Amplifier Classification . . . 55
4.3.1 Linear Power Amplifiers . . . 55
4.3.1.1 Class A . . . 55
4.3.1.2 Class B . . . 55
4.3.1.3 Class C . . . 55
4.3.2 Switching Power Amplifiers . . . 56
4.3.2.1 Class D . . . 56
4.3.2.2 Class E . . . 56
4.3.2.3 Class F . . . 56
4.3.3 Comparison . . . 56
4.4 Class E Design Methodology . . . 57
4.4.1 State-Space Model of the Power Amplifier . . . 59
4.4.1.1 State-Space Description . . . 61
4.4.2 Power Amplifier Automated Sizing . . . 63
4.4.2.1 Introduction . . . 65
4.4.2.2 Tool Methodology . . . 66
4.4.2.3 Automated Analog Circuit Sizing . . . 66
Overview . . . 66
Sizing Methodology . . . 67
4.4.2.4 Device Profilers . . . 69
Inductor Parasitics . . . 69
Capacitor Parasitics . . . 70
Pad and Interconnection Parasitics . . . 71
Transistor Parasitics . . . 72
Other Parasitics . . . 73
4.5 Optimization Experiments . . . 73
4.5.1 Influence of the Supply Voltage and Inductor Quality Factor . . . 74
4.5.2 Influence of Different Output Power Values . . . 77
4.5.3 Influence of Different Transistor Sizes . . . 77
4.5.4 Influence of Different Frequencies of Operation . . . 79
4.5.5 Two Stage Amplifier . . . 82

4.5.6 Comparison with Previous Approaches 85
4.5.7 Final Remarks 85
4.6 Conclusion 86

5 High-Voltage Operation 87
5.1 Introduction 87
5.2 High-Voltage Solutions in CMOS 88
5.2.1 Circuit Breakthroughs 88
5.2.1.1 Cascode 88
5.2.1.2 Self-Biased Cascode 88
5.2.1.3 Non Zero Drain Voltage at Turn-On 89
5.2.2 Customized Silicon Technologies 90
5.2.2.1 Lightly Doped Drain: LDD 90
5.2.2.2 Thick Oxide 91
5.2.2.3 Reduced Surface Field: RESURF 91
5.2.2.4 Fully Customized Technology: EZ-HV™ 92
5.2.2.5 Lateral Double-Diffused MOSFET: LDMOS 92
5.2.3 Reliability Concerns 93
5.2.3.1 Time Dependent Dielectric Breakdown: TDDB . . 93
5.2.3.2 Hot-Carrier Injection: HCI 93
5.2.3.3 Electromigration 94
5.2.3.4 Junction Breakdown 94
5.2.4 Short Channel Effects 94
5.2.4.1 Drain Induced Barrier Lowering: DIBL 94
5.2.4.2 Punchthrough 95
5.2.4.3 Subthreshold and Leakage Current 95
5.3 STI/LOCOS Compatible LDMOS Structure in Standard CMOS . . . 96
5.3.1 Device Isolation Techniques 96
5.3.1.1 Local Oxidation of Silicon: LOCOS 97
5.3.1.2 Shallow Trench Isolation: STI 97
5.3.2 High-Voltage Device Structure 98
5.3.2.1 Improving Breakdown Voltage 98
5.4 Experimental Results 98
5.4.1 DC Experimental Results 99

5.4.1.1 Measurement Setup . . . 101
5.4.1.2 Foundry Transistors . . . 102
Comparison with the SPICE Model from the Foundry 103
5.4.1.3 Multi-Side Polygon Junctions and Transistors. . . . 106
5.4.1.4 LDMOS M_{2_02} . . . 108
5.4.1.5 LDMOS M_{5_03} . . . 108
5.4.1.6 LDMOS M_{A_01} . . . 108
5.4.1.7 LDMOS2 . . . 108
5.4.2 AC Experimental Results . . . 112
5.4.2.1 Measurement Setup . . . 112
5.4.2.2 LDMOS2: Measurements and Characterization of the Device . . . 113
5.4.2.3 Estimation of the Initial Values of the Components . 116
5.4.2.4 Determination of the cut-off Frequency and the Maximum Oscillation Frequency . . . 116
5.4.2.5 Other Measurements . . . 117
5.4.2.6 RF Transistor Model for Class E Power Amplifier . 118
5.4.3 Further Improvements . . . 121
5.5 Design for High-Voltage: an Advantage? . . . 122
5.5.1 Influence of the on-resistance in the DE and PAE . . . 122
5.5.2 Design on a More Advanced Technology . . . 123
5.5.3 Design with an LDMOS Transistor in the same Technology . 125
5.6 Conclusion . . . 125

6 A 850-MHz, 30 dBm Class E Power Amplifier 127
6.1 Introduction . . . 127
6.2 Global System for Mobile Communications (GSM) . . . 127
6.2.1 Specifications . . . 128
6.3 Design for Maximum Drain Efficiency . . . 129
6.3.1 Basic Architecture . . . 129
6.3.2 Two Stage Class E Power Amplifier Optimization . . . 131
6.3.3 Fully Parasitic-Aware Power Amplifier Optimization . . . 133
6.4 Implementation of the Class E Power Amplifier . . . 134
6.4.1 Output Stage . . . 135
6.4.2 Driver Stage . . . 135

6.4.3 Interstage Interconnection 135
6.4.4 Chip Layout 135
6.5 Experimental Evaluation 136
6.5.1 Measurement Setup 137
6.5.2 Measurement Results 137
6.5.2.1 PA-LVMOS2 137
6.5.2.2 Effect in Efficiency of the Number of Bonding Wires to the Ground 140
6.5.3 Performance Comparison 141
6.6 Conclusion 143

7 Conclusions **145**
7.1 Research Overview and Contributions 145
7.2 Possibilities for Future Work 146

List of Publications **149**

Bibliography **151**

List of Figures

2.1 Single stage amplifier. 10

2.2 Two stage amplifier with Miller compensation. 10

2.3 Some of two stage amplifier characteristics for a two stage amplifier with a DC gain of 1000 and a dominant pole at 1 kHz for different values of the error tolerance. Each curve represents a different final error tolerance. The discontinuities are a result of the system natural frequency which occur when response peaks or dips enter the bounded region for increasing phase margin [Yan90]. 10

2.4 Poles of the third order Butterworth filter. 11

2.5 Block diagram of a typical NMC amplifier. 12

2.6 Schematic diagram of the NMC amplifier. 12

2.7 NMC pole zero diagram. 13

2.8 The PFC amplifier. 15

2.9 PFC pole zero diagram. 18

2.10 Influence of the compensation capacitors on the total power consumption of the PFC amplifier. 21

2.11 Schematic diagram of the PFC amplifier. 22

2.12 Microphotograph of the designed operational amplifier. 23

2.13 Measured frequency response of the PFC amplifier. The bump in the frequency response is the result of the zero in LHP. 25

2.14 Measured transient response of the amplifier configured as a unity gain follower (130 pF//24 kΩ). 25

3.1 The DFCFC amplifier topology. 32

3.2 Structure of a AFFC amplifier. 33

3.3 Block diagram of the TCFC amplifier. 34

3.4 Three stage NMCF amplifier topology. 35

3.5 Normalized power as a function of the load capacitor, for C_{m1}=18 pF, C_{m2}=3 pF. Reference power consumption is the one transistor amplifier with $g_m = C_L UGF$. 37

3.6 Normalized power as a function of the load capacitor, for C_{m1}=8pF, C_{m2}=3pF. Reference power consumption is the one transistor amplifier with $g_m = C_L UGF$. 38

3.7 Theoretical transient response of each topology to a step impulse. . . 39

3.8 Normalized settling time as a function of power dissipation and UGF. The light point is the result as given by equations from Chapter 2 and the dark point is the new optimized point. 45

3.9 Different performances of the PFC amplifier. 46

3.10 Pole-zero diagram for the sized amplifiers Table 3.7. 48

3.11 Magnitude and phase margin plots for the original circuit (Section 2.4) and for the last experiment in Table 3.7. 51

3.12 Minimum cost evolution for the experiments in Table 3.7. 52

4.1 The transmitter part of a transceiver. 54

4.2 Ideal waveforms for the drain current and drain voltage for each operation class tuned for maximum power efficiency. Class B and D are usually implemented in push-pull configuration and have each waveform indicated by 1 and 2. 58

4.3 Ideal class E power amplifier. It consists of switch S, the finite dc-feed inductance L_1, the shunt capacitor C_1 the series-tuned (L_0-C_0) in series with a reactance jX and the load resistance R_L. Elements S and C_1 are a representation of the on-resistance and drain-bulk capacitance of a CMOS transistor. 59

4.4 Notation for the nth switching period. 60

4.5 Ideal class E power amplifier as in Fig. 4.3 with the indication of the state-space variables. The switch S is now more conveniently represented by resistance R_k. 61

4.6 Simulated waveforms of the circuit in Fig. 4.5 using two different approaches: by first finding the initial conditions of the state-space description and through solving the differential equations set. 64

4.7 Flow diagram of the PAMPER tool. 67

4.8 Example schematics of possible device models for an inductor serving as input to the DANCE tool. Any customized model representing more complex inductor layouts (octagonal, hexagonal, etc.) or a more accurate representation over frequency can be used. 70

4.9 Possible capacitor elements including parasitics for accurate representation over frequency as used by the DANCE tool. 71

4.10 Lumped circuit models of a shielded and unshielded pad. The DANCE tool can use these and other equivalent models. 72

4.11 Possible equivalent circuit for a transistor in a class E power amplifier. The model includes the off- and on-resistance and the capacitances between the gate and the other three terminals. It gives satisfactory results when applied to a class E power amplifier. 73

4.12 Class E amplifier schematic with the switching device replaced by a NMOS transistor. In this way, the effect of the different capacitances between the gate and the other terminal are considered. Capacitor C_1 in Fig. 4.3 is now the drain-bulk capacitance of the MOS transistor and reactance jX is merged with the inductor L_0. 75

4.13 Drain Efficiency and Power Added Efficiency as a function of supply voltage for different inductor quality factors Q_L. 76

4.14 Efficiency for different output power values. 78

4.15 DE and PAE for an amplifier with different transistor sizes. 80

4.16 Variation of the DE and PAE with varying operation frequency. 81

4.17 Circuit schematic of the differential two stage class E power amplifier. Appropriate parasitics are included for better representation of the losses. Measured efficiency should thus be similar to the optimized value obtained here. 82

4.18 Simulated waveforms of the optimized class E power amplifier: drain current and drain voltage of the output transistor obtained with a sinewave with a 1 V peak voltage (+10 dBm input power). Due to the switching nature of the amplifier, sharper edges on the current waveform are obtained for a higher peak voltage which when on-chip can be obtained without the need to be matched to a 50 Ω resistor. And thus saving in the RF input power. Nevertheless, the transistor/switch current will never be similar to Fig. 4.6(b) due to the reason that the output stage is not driven by a perfect square wave. 84

4.19 Drain efficiency and power added efficiency of the optimized circuit as a function of the input power. 84

4.20 Minimum cost evolution during optimization. The optimization is forced to continue after meeting one of the stop criteria. 85

5.1 Various cascode configurations: (a) Conventional, (b) self-biased and (c) bootstrapped. 89

5.2 Voltage and current waveforms at the switch. (a) As commonly designed [Sok75, Raa78b] and (b) with non zero voltage at turn-on [Por93]. 90

5.3 Device with lightly doped drain implant (LDD, n) surrounded by a high-doped region ($n+$) to implement the source and drain region. . . 91

5.4 In dual gate oxide technology, for applications where a high breakdown voltage is desired, transistors featuring a thick oxide are used, whereas the other devices are used for low-voltage operation. 92

5.5 Cross section of the high-voltage transistor with the poly acting as field plate [Par87]. 92

5.6 Electrons with high kinetic energy upon reaching the drain lattice can generate substrate current or destroy the thin oxide. They are the cause of severe reliability problems in semiconductor devices. The inclusion of LDD implants can reduce electrons speed and minimize their effect. 93

5.7 Effect of DIBL on threshold voltage. 95

5.8 Punchthrough occurs when the depletion region from the drain reach the source side and reduces the barrier for electron injection. It can either occur at the surface (a) or within the bulk (b). 95

5.9 LOCOS vs. STI comparison. 97

5.10 Cross sectional representation of the STI/LOCOS compatible high-voltage structure in standard CMOS, LDMOS M_{5_03}. 99

5.11 Starting with the structure from Fig. 5.10, an improvement in the breakdown voltage and a reduction in the on-resistance is obtained by extending the *n*-well towards the source junction. This creates the above M_{A_01} transistors. 99

5.12 Schematic diagram of the experimental arrangement used for DC characterization. 102

5.13 Designed test chip from run723 including nine large groups subdivided in 25 individual elements. Each of the 225 individual devices is numbered between M_{1_01}–M_{9_25}. 103

5.14 Test chip designed in run815. It includes three families equal to run723 included for run to run comparison: M_{1_xx}, M_{5_xx} and M_{3_xx}. Also included are four new groups including 100 unique transistors numbered between M_{A_01}–M_{D_25}. 104

5.15 Microphotograph from run876. Transistor LDMOS2 is composed of 180 M_{A_01} parallel transistor. Its large size (5.6 mm) minimizes the effect of parasitics during AC characterization. 104

5.16 Set of DC measurements for transistor M_{1_01}. 105

5.17 Comparison between measured and simulated waveforms for transistor M_{1_01}. 107

5.18 Set of DC measurements for transistor M_{2_02} 109

5.19 Set of DC measurements for transistor M_{5_03}. 110

5.20 Set of DC measurements for transistor M_{A_01}. 111

5.21 LDMOS2 output characteristics with gate voltage from 0 V to 1.6 V. This device is made of 180 M_{A_01} unity transistors giving a total length of 5.6 mm. As a result of the high power dissipation in the device, self-heating phenomenon is observed in the saturation region. 112

5.22 Photograph of the dies on a ceramic substrate used for device characterization. 113

5.23 Setup used in the AC measurements of the transistor LDMOS2. . . . 114

5.24 Smith chart representation of measured (○, □, and ×) and modeled (–) s parameters of the LDMOS device at V_{DS}=0 V and V_{GS}=0 V. Values measured over a range of frequencies from 0.1 to 5 GHz. 115

5.25 Simplified transistor model used to estimate initial values for the indicated components. 117

5.26 Transistor equivalent model including both the intrinsic and extrinsic elements. Substrate parasitics are included to increase accuracy of the fitting parameters to the measurements. 118

5.27 Device cut-off frequency and f_{max} at V_{DS}=3 V and V_{GS}=1 V. 119

5.28 f_T (solid line) and f_{max} (dashed line) versus drain current for three different V_{DS}. A maximum value of 3.5 GHz is measured for the f_T. The maximum f_{max} is measured to be 6.3 GHz. 119

5.29 C_{gs} and C_{gd} per millimeter length at 1 GHz with V_{GS}=0. 120

5.30 C_{gs} and C_{gd} per millimeter length at 1 GHz with V_{DS}=0. 120

5.31 C_{ds} per millimeter length at 1 GHz with V_{GS}=0. 120

5.32 Comparison between the SPICE BSIM3 model and the model from Fig. 4.11. Values for the capacitances and on-resistance are taken from Table 5.2 and Table 5.5. Because the substrate is short-circuited with the source, the measured C_{gs} value includes the C_{gs} and C_{gb}. 121

5.33 Cross section of improved LDMOS. 122

5.34 Drain Efficiency and Power Added Efficiency as a function of supply voltage for different inductor quality factor Q_L. Transistor on-resistance for the LDMOS M_{A_01} (○) is 2.2 times greater than for the LVMOS M_{1_01}(+). 124

6.1 Mask limits for the GSM burst as specified by the standard. 130

6.2 Mobile station GSM spectral mask due to GSMK modulation. A 30 kHz measurement bandwidth is used for frequencies below 1800 kHz and 100 kHz otherwise. A higher attenuation is required for higher transmitted power (Table 6.1), resulting in the different horizontal lines. . . 130

6.3 Source/Drain transistor junctions layout. 131

6.4 Complete two stage fully differential class E PA circuit schematic including all parasitics. The performance-limiting parasitics are taken into account during optimization: pad, inductor, discrete SMD capacitor and transistor. Bonding wires connecting to the ground are considered to be enough in order not to limit the performance. To better describe the testing environment, coupling between the inductors is included. Furthermore, the transmission line delay (TD) between L_M and C_M is fine-tuned to maximize efficiency. The die is indicated by a dashed line. 134

6.5 The RF signal enters the chip on the left side through the hexagonal pad that connects directly to the driver stage. The amplified signal is further amplified in the output stage before leaving the chip via the RFOUT pad. The final fully differential amplifier is constructed by a mirror and copy operation. All the remaining pads are used for connecting to the ground. 136

6.6 Setup used in the measurements of the PA-LVMOS2 power amplifier. 138

6.7 Output spectrum for a sinewave input signal measured with a RS FSIQ 26 signal analyzer with a resolution bandwidth of 30 kHz. 140

6.8 Measured output power, DE and PAE over supply voltage. 141

6.9 Output power, DE and PAE *versus* frequency measured with V_{DD}=2 V. 141

6.10 GSM spectrum with the output at the maximum output power level at an input signal of 855 MHz and a filter parameter BT=0.3. The resolution bandwidth of the signal analyzer (RS FSIQ 26) is 30 kHz. . 142

6.11 Frequency domain output of the differential amplifier from 800 MHz up to 2800 MHz with a 855 MHz input signal frequency. 142

List of Tables

2.1 Component values used in the PFC amplifier. 23

2.2 Measured Performance. 25

2.3 Performance Comparison of Different Multistage Amplifiers 26

3.1 Optimization Parameters 31

3.2 Results for C_{m1}=18 pF, C_{m2}=3 pF and C_L=100 pF 37

3.3 Results for C_{m1}=8 pF, C_{m2}=3 pF and C_L=100 pF 38

3.4 Sizing results for sub-section 3.3.1 with C_{m1}=18 pF, C_{m2}=3 pF and C_L=100 pF and comparison with equations from Chapter 2. 44

3.5 Performance Improvement 44

3.6 Simulated performances for the original amplifier design (Section 2.4) and three automatically sized experiments using DANCE (including some optimization statistics). 47

3.7 Device sizes, component values and biasing for the original design (Section 2.4) and several automatically sized experiments using DANCE. . 49

3.8 Dominant poles and zeroes for the sized amplifiers in Table 3.7. Unlike the reference which has explicitly set $g_{m3} = g_{mf}$, the previous transconductances can have different values in the proposed designs according to DANCE. The result is that the complex conjugate pole pair disappears and z_1 compensates for the other real nondominant pole p_3. 50

4.1 Optimized circuit sizes for the circuit in Fig. 4.5. 63

4.2 Inputs for the first experiment: influence of the supply voltage and inductor quality factor. 76

4.3 Effect of the supply voltage and inductor quality factor on the circuit sizes. 76

4.4 Inputs for the second experiment: influence of different output power values. 78

4.5 Circuit sizes for different output power values. 78

4.6 Inputs for the third experiment: influence of different transistor sizes. . 80

4.7 Optimized circuit sizes considering a fixed transistor size. 80

4.8 Inputs for the fourth experiment: influence of different frequencies of operation. 81
4.9 Resulting device values optimized for different frequencies of operation. 81
4.10 Constraint template. 83
4.11 Optimized circuit sizes for the circuit in Fig. 4.17 for the situation where the constraints are applied to both stages or only to the output stage. 84

5.1 Comparison between measured and simulated on-resistance for transistors with W=15 µm operating with V_{GS}=3 V and V_{DS}=0.1 V. . . . 106
5.2 Measured Electrical Parameters of Various Devices 112
5.3 Aluminum oxide (Al_2O_3) ceramic substrate properties. 115
5.4 Measured capacitances per millimeter length. 119
5.5 Simulated capacitances per millimeter length. 119
5.6 Optimized circuit sizes for same conditions as in Table 4.3 with the exception that the switch-on resistance is 2.2 times greater. 123

6.1 GSM Mobile Station Types. 130
6.2 Efficiency for different transistor drain-bulk capacitance per unit width for the circuit in Fig. 4.17 and 1 W output power. 132
6.3 Optimized circuit sizes for the circuit in Fig. 6.4 considering all circuit and ceramic substrate parasitics. 133
6.4 Summary of the class E power amplifier performance. 140
6.5 Overview of CMOS power amplifiers. 144

Symbols and Abbreviations

Symbols

A_k, B_k, C_k, D_k	Constant real matrices of a state-space representation
A_o, A	Open and closed loop gain
C_{gg}	Total input gate capacitance
η	Efficiency
f_{dc}	DC feedback of a closed loop amplifier
f_{max}	Maximum oscillation frequency
f_T	Unity current gain frequency (cut-off frequency)
I	Identity matrix
L_{DRIFT}	Gate lenght over drift region
L_{GATE}	Gate length over channel region
n, n^+, n^-	n-type implant
ω_c	Cut-off frequency
ω_{do}, ω_d	Open and closed loop dominant pole
ω_{no}, ω_n	Open and closed loop undamped natural frequency
Q_L	Inductor quality factor
R_k	Switch resistance on different portion of a switching period
R_{ON}, R_{OFF}	On- and off-resistance of a MOS transistor
$\sigma_{n,k}$, $\sigma_{n,1-3}$	Switching instants
T	Switching period
τ_k, τ_{1-2}	Portion of a switching period
T_{OX}	Thin oxide thickness
$u(t)$	Input vector of a state-space representation
V_{BR}	Transistor or junction breakdown voltage
V_{TH}	Threshold voltage of a MOS transistor

W	Transistor width
x_1, x_2, x_3, x_4	State-space variables
$x_{n,k}(t)$	State vectors of a state-space representation
$y_{n,k}(t)$	Output vector of a state-space representation
ζ_o, ζ	Open and closed loop damping ratio parameter

Abbreviations

AFFC	Active Feedback Frequency Compensation
ACPR	Adjacent Channel Power Ratio
BJT	Bipolar Junction Transistor
CAD	Computer-Aided Design
CMOS	Complementary Metal Oxide Semiconductor
CMP	Chemical Mechanical Planarization
DE	Drain Efficiency
DFCFC	Damping Factor Control Frequency Compensation
DIBL	Drain Induced Barrier Lowering
DMOS	Double - Diffused MOSFET
EEPROM	Electrically Erasable Programmable Read Only Memory
FDMA	Frequency Division Multiple Access
GaAs	Gallium Arsenide
GMSK	Gaussian Minimum Shift Keying
GSM	Global System for Mobile Communications
HCI	Hot Carrier Injection
I/O	Input/Output
LDD	Lightly Doped Drain
LDMOS	Lateral DMOS
LOCOS	Local Oxidation of Silicon
LVMOS	Low Voltage MOS
MAG	Maximum Available Gain
MOSFET	MOS Field Effect Transistor

MOS	Metal Oxide Semiconductor
MSG	Maximum Stable Gain
NMCF	Nested Miller Compensation with Feed-forward
NMC	Nested Miller Compensation
NMOS	n-channel MOSFET
PAE	Power Added Efficiency
PA	Power Amplifier
PFC	Positive Feedback Compensation
PMOS	p-channel MOSFET
PM	Phase Margin
RESURF	REduced SURface Field
RF	Radio Frequency
SiGe	Silicon Germanium
SPICE	Simulation Program with Integrated Circuit Emphasis
STI	Shallow Trench Isolation
TCFC	Transconductance with Capacitances Feedback Compensation
TDDB	Time Dependent Dielectric Breakdown
TDMA	Time Division Multiple Access
UGF	Unity Gain Frequency
WPAN	Wireless Personal Area Network

Chapter 1

Introduction

1.1 Motivation and Applications

Wireless communications have seen a remarkable growth in the last years. Mobile communications are part of our lives, so much so that it is common to ask for somebody's mobile phone number. Together with microprocessors and computers, the wireless market is a driving force behind semiconductor development, with foundries improving their technologies to include extensions that increase the performance of Radio Frequency (RF) circuits. Smaller, lighter, with extra functionality and with higher autonomy are key characteristics to create a more attractive product. Manufacturers thus sell a myriad of such devices: mobile phones, Personal Digital Assistants (PDAs), multimedia players, headsets, GPSs, etc. All these devices are so universal in the developed world that with the exception of the author of this thesis, everybody has at least one type of battery operated wireless device.

Interest in wireless Internet has skyrocketed in recent years. Furthermore, there is a growing interest in sensors that can communicate with one another. The need for an ever increasing data traffic or simpler and cheaper devices have as result that standards are both being developed or enhanced.

For this wireless world to exist, the block that ultimately transmits the signal must necessarily be present. This circuit is commonly called Power Amplifier (PA). In a wireless transmitter the task of a power amplifier is to increase the input signal power to a higher power before delivering it to the antenna. The PA requires special attention due to its characteristics. A considerable percentage of the total power consumption of a mobile device is in this block alone. If the efficiency is low two things happen: more power has to be drained from the batteries and the extra power generated due to its inefficiency is transformed into heat that must be adequately removed from the chip. The heat in itself is a problem because it degrades circuit life-time. Another consequence is that a package that can dissipate heat more efficiently is necessary with the drawback that this is usually more expensive. More power consumption also implies that the user has to charge the batteries more frequently, making this less appealing. By no means unimportant, is that this inefficiency also requires more resources by wasting more electricity with the consequent increase in the cost of operation of the product.

Specifically in the case of mobile phones, the maximum output power is usually only transmitted in the beginning of the handshake with the base station or when on the road, moving to another base station. This happens, as the mobile phone can be instructed,

by the base station, to reduce its power level. The issue here is that commonly the highest efficiency of the PA is only obtained when transmitting its peak output power. It is thus desirable to achieve a high efficiency over a broad range of output power levels to maximize battery autonomy.

With the ever decreasing maximum supply voltage, circuit functionality must be executed without performance loss. Because the supply voltage is linked to the output power by $P = VI$, increasing the circuit current is the only way to keep the output power constant. With each advance in technology, device geometries and metal interconnections become smaller, which makes it difficult to design PAs. Although in a transceiver, it is the PA that frequently has the highest percentage of the total power budget, other circuits play an equally important role. Analog filters, voltage controlled oscillators, A/D, D/A and so on, must also perform in this situation.

Other types of amplifiers are also required in the analog part of a transceiver. Various functions in it depend on how good and how efficient these amplifiers function. One of the most universal type of amplifiers is the operational amplifier (OPAMP). Two of the most common characteristics are its high DC gain and the frequency at which this gain reaches the value of one. More advanced technologies make increasing the unity gain frequency an easier task. However, because the supply voltage is decreased, circuit techniques, different from those used before, must be investigated so the DC gain continues to be equally high. The operational amplifier commonly works in a closed loop with negative feedback which desensitizes circuit variations in the closed loop, if the open loop gain is high.

All wireless protocols such as the Global System for Mobile Communications (GSM) for mobile phones, the IEEE 802.x for wireless local area networks and short range Wireless Personal Area Network (WPAN) which include for example the Bluetooth® and IEEE 802.15.x need in their implementation a different set of amplifiers. However, only the previous two will be dealt with in this work.

1.1.1 Cost Reduction

Evidently, the best circuit in the world is worthless if it cannot be sold because it is too expensive. Its results can be published but the company bottom line will not profit from it. To that end, careful selection of the technology is important. This is where Complementary Metal Oxide Semiconductor (CMOS) plays a role. In general, CMOS performs poorly at high frequency but the potential of large scale integration with the high integration it provides for digital circuits and other supporting analog functions, makes is attractive for single chip solutions. The academia thinks that this is still an unexplored field which can originate new ideas and interesting situations. The reason why CMOS is not used more in the industry is because designers there have mostly used, and feel more comfortable using other technologies. With more students with understanding of CMOS, it is expected that this situation is likely to change. These days, for applications up to a couple of GHz, a full system is a reality while small blocks are already available for frequencies above 50 GHz.

Needless to say, the engineering cost can be reduced, or at least optimized. By carefully

gathering important information used during sizing, a faster design can be obtained the next time the circuit is sized. More important, this second sizing can also occur during or before the chip is first sent for fabrication. If more similar steps are executed in an automatic manner while a somehow intelligent algorithm decides the best way to go, two advantages arise. Firstly, all the knowledge is retained. Secondly, it is possible that a new design might even be proposed, which might turn out to be more efficient. Another advantage includes the reduction in design cycles. The designer can now concentrate on really interesting design aspects instead of having to guarantee that the circuits can indeed work on all process corners.

Many variables influence the total cost: process technology, die area, package type and circuit test to name a few. Cost savings can in this way be obtained by integrating. Two advantages arise from this fact: only one package is necessary and testing is simplified. Furthermore, because the signal can always stay inside the chip, the circuit can be made more power efficient, hence consuming less power so that smaller batteries are required for the same performance. In the long run, more functionality can be crammed into silicon while having similar or increased autonomy. For the final integrator, having to deal with less chips has the advantage of saving real estate and complexity. At the end, cost is the prevailing factor and the cheapest solution often wins.

1.2 Research Work

The presented work deals with the analysis, simulation, design and measurement of two types of amplifiers: the operational amplifier for high-gain and low-power design and the class E power amplifier for RF applications in the current mobile phone standards. The work addressed in this thesis is not extensive for the obvious reason that this field is very extensive. It is nevertheless interesting to join these two types of amplifiers for two different blocks in the same final product. The remaining of this section gives an overview of the carried out research work.

Focusing only on the CMOS technology, two different applications are to be explored. While analog filters or A/D converters require amplifiers having simultaneously low-power and low-voltage operation with maximized efficiency, the amplifier in charge of delivering a strong signal to the antenna for transmission is necessarily a high-power circuit. In common they have the functionality while the frequency of operation, the power level and the constraints in their design sets them apart.

- Multistage operational amplifiers. In the work presented, various aspects of the design of three stage operational amplifiers are presented. In particular, the following characteristics are discussed.
  - A novel frequency compensation for three stage operational amplifiers is presented. A detailed analysis of the different design parameters is given covering the small signal transfer function, Unity Gain Frequency (UGF), Phase Margin (PM), stability and design equations.
  - A method for amplifier comparison where the influence of the compensation and load capacitors is minimized. Optimization both at the block and

transistor level can be used to further minimize the power consumption.

- Radio frequency amplifiers based on non-linear class E power amplifier. The following design aspects are covered in the presented work.
  - A methodology to optimally design and optimize radio frequency amplifiers while accounting for the parasitic effects is presented. This means two things: first, the study of the amplifier topology where the inclusion of simple parasitics makes analytical analysis cumbersome, is possible; secondly, the optimization for maximum efficiency of complex circuits including all device and board parasitics within a couple of hours.
  - The exploration of the possibility to increase the supply voltage beyond the maximum stated by the technology provider in a commercial CMOS technology and later use it for a class E power amplifier design. In particular, this increase in the transistor breakdown voltage is obtained without any change in the already existing process flow.

As a result of this work, two amplifiers have been fabricated of which more details are given in this text. A brief summary of their results is presented below.

- Three stage operational amplifier: It is implemented in 0.35 µm CMOS technology. A Unity Gain Frequency (UGF) of 2.7 MHz with a PM of 52° while loaded by 130 pF//24 kΩ is achieved. The measured average slew rate is 1.0 V/µs. The amplifier power consumption is 275 µW.
- Class E power amplifier: It is implemented in 0.35 µm CMOS technology and reaches a maximum Power Added Efficiency (PAE) of 66 %. At a supply voltage of 2.26 V a maximum output power of 955 mW is obtained for an input signal with a frequency equal to 855 MHz. It has a high efficiency over a broad range of output power levels and the PAE is always greater than 60 % for values above 158 mW.

1.3 Outline of the Work

The outline of the presented work is as follows. The first two chapters discuss the operational amplifier design while the following three chapters cover the implementation of the radio frequency amplifier. A more detailed description of each one of the chapters is now given.

Chapter 2 presents a detailed analysis of the novel three stage frequency compensation. The chapter starts with a brief introduction of the need to use multistage amplifiers in the near future as a result of the supply voltage reduction which decreases the maximum attainable gain per stage. After this, the Nested Miller Compensation (NMC) topology which is usually used as the reference for multistage amplifiers is briefly introduced. Then, the proposed frequency compensation scheme and some of its more common characteristics are presented. Finally, to validate the design procedure, the measurement results of the fabricated operational amplifier are given.

Chapter 3 presents the method to compare various amplifier topologies where the influence of design criteria, such as the value of the compensation and load capacitors or the transconductance, are minimized. Using the design equations given in each of the papers where the topologies are first presented, the results are then given for each of the compensation topologies, for a range of load and compensation capacitors. Using the results already obtained, the small signal settling time is presented for each one of the amplifiers. The second part of this chapter discusses the possibilities to further reduce the power consumption by properly selecting the operating point, either at the block or at the transistor level. Optimizations are used in this case to explore new design possibilities.

Chapter 4 starts by briefly presenting some of the most common classes of amplifiers that are used in RF applications. A simplified state-space model of the class E power amplifier is given alongside with the advantages and disadvantages. A new methodology for optimally designing power amplifiers for maximum efficiency is then given. A computer tool is subsequently built and used to perform simple studies on the basic class E power amplifier with respect to various design criteria. To show the flexibility, a more complex differential two stage class E power amplifier including circuit parasitics is later optimized.

Chapter 5 first discusses existing possibilities to achieve a high breakdown voltage in CMOS. The solutions include both circuit design and technological advances. Then, follow reliability concerns and short channel effects which are enhanced for an increased supply voltage operation. A novel high-voltage structure compatible with the most common isolation techniques used in current technologies is later given. Measurements from manufactured test chips are given where the various low-voltage and high-voltage structures are compared in terms of some of their characteristics. The chapter ends by studying the advantages and disadvantages of a high-voltage transistor in terms of its on-resistance and total input capacitance when applied to the design of a class E power amplifier. For better comparison, the study is done in the same conditions as the one done in Chapter 4

Chapter 6 presents the design and experimental results of an efficient class E power amplifier implemented in CMOS technology. Designed to be employed in the GSM-850 standard, the architecture is first briefly described. Then, the circuit and board parasitics are included in an optimization loop where the tool methodology described in Chapter 4 is used to maximize the circuit efficiency. The chapter ends with the presentation of the measurement results of the circuit, which as a result of proper modeling of all parasitics are in excellent agreement with the simulations.

Chapter 7 presents the conclusions and possible topics for future research.

Chapter 2

Positive Feedback Frequency Compensation for Three Stage Amplifier

2.1 Introduction

Research in analog circuit design is focused on low-voltage low-power battery operated equipment to be used for example in portable equipment, wireless communication products, hearing aids and consumer electronics. A reduced supply voltage is necessary to decrease power consumption to ensure a reasonable battery lifetime in portable electronics. For the same reason, low-power circuits are also expected to reduce thermal dissipation, of increasing importance with the general trend in miniaturization.

Advancements required by International Technology Roadmap for Semiconductors (ITRS), means that with current CMOS standard fabrication processes, circuits must work at supply voltages as low as 1.5 V [ITRS]. Industry and academia are researching new circuit techniques that will make them operate at this voltage. To this end the industry and the academia are doing research in new circuit techniques to enable them to operate at this voltage. Working at lower voltages poses new constraints, especially important for the specific case of analog design. However, a reduction in performance is not desirable. Nevertheless, going to lower voltages and higher efficiencies require innovative circuits to solve current design needs of faster circuits and better performance.

Overall cost is important as well and integration is seen as a possible solution. Another is the use of simpler fabrication processes. CMOS technology is today *de facto* consumer technology, mainly due to its use for microprocessors and memories. This causes the most advanced technologies to be tailored to digital circuits. New analog circuits and solutions that can extract the most from a digital custom-built technology are thus needed.

Specifically in the case of analog design, operational amplifiers continue to be frequently used as building blocks for analog processing. This constitutes a strong motive to do research in circuits able to operate at these low voltages, but without decrease in performance. Particularly at low voltages, cascoding is no longer an advisable technique to achieve high gain, as stacked transistors limit the available voltage swing at the output. As a result the multistage amplifier is being researched as a technique to overcome this limitation. As such, recent designs suggest some solutions for operation at lower supply voltages [Fon91, Ng99, Leu00], higher gain [Bul90, Ng99, Leu00], bandwidth and efficiency.

The goal of this work is to develop a three stage amplifier with a better bandwidth to power efficiency and suitable for driving capacitive loads for circuits that require a high gain such as high accuracy $\Delta\Sigma$ modulators, pipeline A/D converters, linear regulators, etc. Due to its characteristics, it is also appropriate to drive large external capacitive loads in applications where the voltage is intrinsically low. For the CMOS operational amplifier presented in this chapter, a three stage (good compromise between a high voltage gain, complexity of the circuit and power dissipation) amplifier is developed that offers simultaneous operation at low voltages (1.5 V), rail-to-rail output swing, high gain ($>$100 dB) and good bandwidth to power dissipation efficiency.

The organization of this chapter is as follows. In Section 2.2 the single stage and multistage operational amplifiers are briefly explained. In Section 2.3 a brief review of the three stage Nested Miller Compensation (NMC) amplifier topology is given. Then, the Positive Feedback Compensation (PFC) amplifier and its frequency response characteristic are introduced in Section 2.4. Next, the PFC amplifier and its frequency response characteristic are introduced. Consequently, in Section 2.5 follow the measured results from a test chip fabricated and measured to show the effectiveness of the compensation scheme. Finally, the conclusion of the work shown in this chapter is presented in Section 2.6.

2.2 Operational Amplifiers

The term operational amplifier refers to an electronic circuit with two inputs (one inverting (-) and one non-inverting (+)) and one output. More specifically, the term operational comes from the fact that they can be used in an analog computer used to perform mathematical operations such as: sum, subtraction, multiplication, integration and so on. Although the study in this section directly focuses on a first stage with two inputs, it is also valid for non-differential input stages.

2.2.1 Single Stage Amplifiers

Single stage amplifiers have the advantage of intrinsic simplicity (Fig. 2.1). For this reason they can be extensively analyzed and manual optimization is relatively simple. They also feature a good ratio of bandwidth to power dissipation which make them good candidates for high speed circuits. The main reason why they are not used more often is that they show reduced gain for the case of the single transistor or reduced output swing for the case of the telescopic amplifier. For this reason they are commonly replaced by a two stage amplifier for applications where this low gain or output swing is inadequate.

2.2.2 Two Stage Amplifiers

For low voltage applications where high gain and/or high voltage swing is necessary the designer usually selects a two stage amplifier. The tradeoff is, however, stability. While the single stage amplifier is intrinsically stable, two stage amplifiers can have stability problems. For this reason, frequency compensation is required. The simplest of these is shown in Fig. 2.2. The Miller capacitor applies negative feedback around the output

stage, which splits the two real poles so that they remain real even when feedback is applied to the amplifier. This introduces a dominant pole at a frequency low enough to ensure a roll-off of 20 dB/dec in the open loop response, down to the UGF. The correct positioning of the poles affects not only the stability (Fig. 2.3(a)) but also the settling-time (Fig. 2.3(b)). Miller compensation and other methods of frequency compensation are further discussed in [Lak94].

Again, if the supply voltage is low, two stage amplifiers might still not provide the necessary gain. This is especially true if gain boosting cannot be used or the output swing is limited by the available supply voltage. In this case the designer must necessarily use a three stage amplifier. These type of amplifiers are discussed next.

2.2.3 Three Stage Amplifiers

The basic principle in the stabilization of three stage amplifiers is to assume that it has third order Butterworth frequency response with unity gain frequency [Esc92]. Only the position of the poles is taken into account under the assumption that the zeros influence is minimal. The poles are evenly distributed on a circle of radius ω_c (cut-off frequency) in the s plane. For a third order system their relative position is given in Fig. 2.4. The closed loop transfer function $G(s)$ can be obtained from the amplifier open loop transfer function $G_o(s)$ with a DC feedback f_{dc}. It can be modeled as

$$G(s) = \frac{G_o(s)}{1 + f_{dc}G_o(s)};\ f_{dc} = 1 \tag{2.1}$$

$$= \frac{1}{1 + \left(\frac{2}{\omega_c}\right)s + \left(\frac{2}{\omega_c^2}\right)s^2 + \left(\frac{1}{\omega_c^3}\right)s^3} \tag{2.2}$$

$$= \frac{1}{\left(1 + \frac{1}{\omega_c}s\right)\left(1 + \frac{1}{\omega_c}s + \frac{1}{\omega_c^2}s^2\right)} \tag{2.3}$$

and, with poles at

$$-\omega_c;\ \left(-\frac{1}{2} \pm \frac{\sqrt{3}}{2}\right)\omega_c \tag{2.4}$$

The result from (2.3) is obtained when the open loop transfer function $G_o(s)$ is in the form of

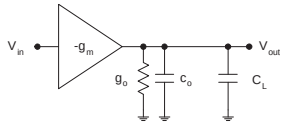


Figure 2.1: Single stage amplifier.

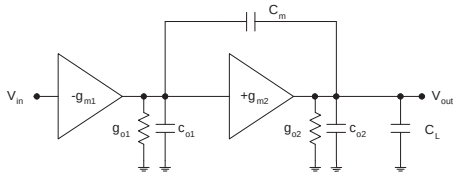
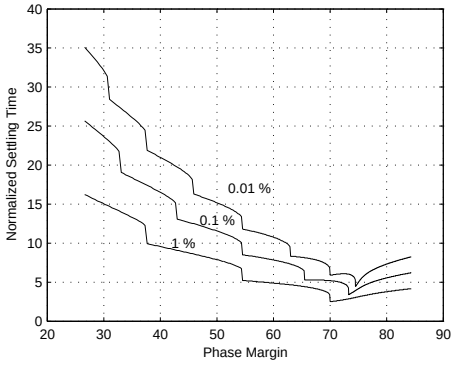
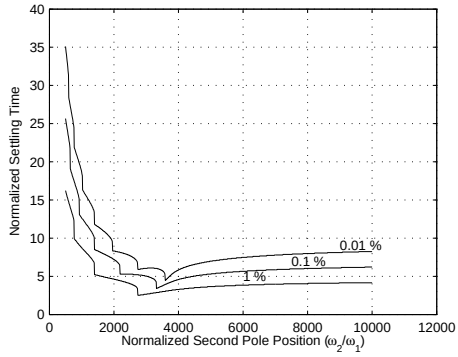


Figure 2.2: Two stage amplifier with Miller compensation.



(a) Normalized small-signal settling time vs. phase margin.



(b) Normalized small-signal settling time vs. normalized second pole position.

Figure 2.3: Some of two stage amplifier characteristics for a two stage amplifier with a DC gain of 1000 and a dominant pole at 1 kHz for different values of the error tolerance. Each curve represents a different final error tolerance. The discontinuities are a result of the system natural frequency which occur when response peaks or dips enter the bounded region for increasing phase margin [Yan90].

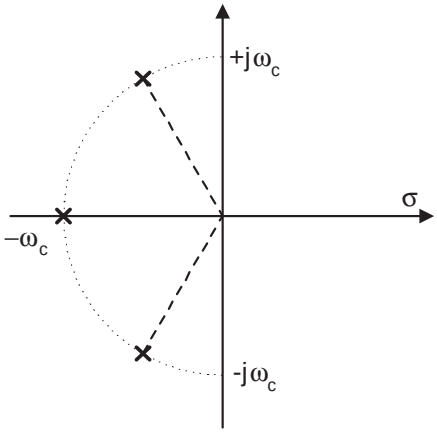


Figure 2.4: Poles of the third order Butterworth filter.

$$G_o(s) = \frac{A_o}{\left(1+\frac{s}{\omega_{do}}\right)\left(1+2\zeta_o\left(\frac{1}{\omega_{no}}\right)s+\left(\frac{1}{\omega_{no}^2}\right)s^2\right)} \tag{2.5}$$

$$\approx \frac{1}{\left(\frac{2}{\omega_c}s\right)\left(1+\left(\frac{1}{\omega_c}\right)s+\left(\frac{1}{2\omega_c^2}\right)s^2\right)} \tag{2.6}$$

By equating the above two second order systems, equation (2.5) and (2.6), the open loop damping ratio parameter (ζ_o) is $1/\sqrt{2}$ and the open loop undamped natural frequency (ω_{no}) is $\sqrt{2}\omega_c$. With these values it is guaranteed that the amplifier open loop response has no peaking in the magnitude response. Nevertheless, a peaking around 10 % of the final value will be found in the closed loop transient response due to the phase margin value.

2.3 The NMC Structure

The three stage amplifier with the NMC has been extensively studied [Fon91, Esc92, Per93, Ng99, Leu00] and only a brief overview will be given. Its basic structure is shown in Fig. 2.5. The small-signal elements $g_{m(1-3)}$, $g_{o(1-3)}$ and $c_{o(1-3)}$ represent, respectively, transconductance, output conductance and output capacitance. It has two compensation capacitors: C_{m1} and C_{m2}. The load capacitor is represented by C_L. The output conductance g_{o3} includes the output resistive load.

The open loop transfer function of Fig. 2.5 is rather complex. A few assumptions are required in order to simplify it, while maintaining accuracy: first, C_{m1}, C_{m2} and $C_L \gg c_{o(1-3)}$; second, the effect of a factor containing the output conductance is negligible

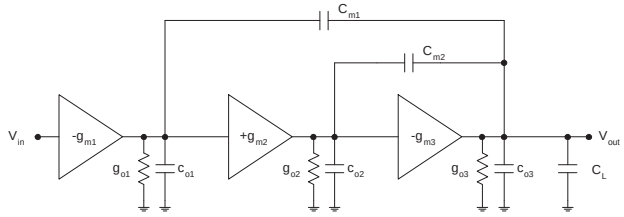


Figure 2.5: Block diagram of a typical NMC amplifier.

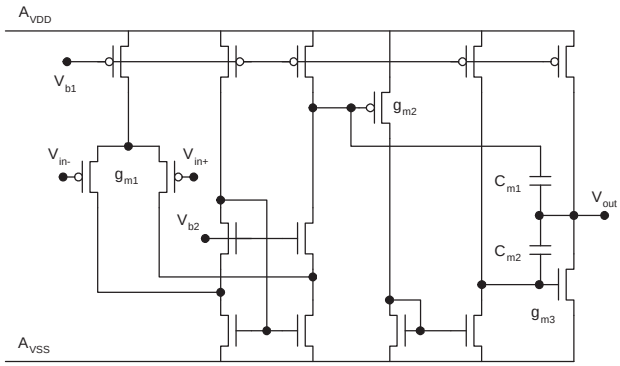


Figure 2.6: Schematic diagram of the NMC amplifier.

when compared to one containing a transconductance. The transfer function, after simplification is

$$G_o(s) = \frac{A_o\left(1 - \frac{C_{m2}}{g_{m3}}s - \frac{C_{m1}(C_{m2}+c_{o2})}{g_{m2}g_{m3}}s^2\right)}{\left(1+\frac{s}{\omega_{do}}\right)\left(1+\frac{C_{m2}(g_{m3}-g_{m2})}{g_{m2}g_{m3}}s+\frac{C_L C_{m2}}{g_{m2}g_{m3}}s^2\right)} \tag{2.7}$$

where the DC gain is given by

$$A_o = \frac{g_{m1}g_{m2}g_{m3}}{g_{o1}g_{o2}g_{o3}} \tag{2.8}$$

with a dominant pole at

$$\omega_{do} = \frac{g_{o1}g_{o2}g_{o3}}{C_{m1}g_{m2}g_{m3}} \tag{2.9}$$

The transfer function has two zeros: one in the left half plane (LHP), and one in the right half plane (RHP). The latter in module is at a lower frequency which degrades the

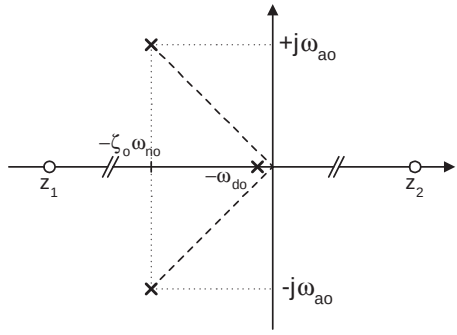


Figure 2.7: NMC pole zero diagram.

phase margin. The zeros are given by

$$z_{1,2} = -\frac{g_{m2}}{2C_{m1}} \mp \sqrt{\left(\frac{g_{m2}}{2C_{m1}}\right)^2 + \frac{g_{m2}g_{m3}}{C_{m1}C_{m2}}} \tag{2.10}$$

The denominator of (2.7) can be simplified, as g_{m3} is usually much larger than g_{m2}. The determination of the compensation capacitors (C_{m1}, C_{m2}) comes from the fact that the NMC amplifier has a third order Butterworth frequency response with unity gain feedback [Esc92]:

$$C_{m1} = 4\left(\frac{g_{m1}}{g_{m3}}\right)C_L \tag{2.11}$$

$$C_{m2} = 2\left(\frac{g_{m2}}{g_{m3}}\right)C_L \tag{2.12}$$

Substituting (2.11) and (2.12) into the denominator of (2.7) results in having a pair of complex conjugated poles at

$$p_{2,3} = -\frac{g_{m3}}{2C_L} \pm j\frac{g_{m3}}{2C_L} \tag{2.13}$$

with the open loop damping ratio ζ_o equal to $1/\sqrt{2}$ resulting in the pole zero diagram given in Fig. 2.7. The UGF and PM can now be determined

$$UGF_{(NMC)} = \frac{1}{4}\left(\frac{g_{m3}}{C_L}\right) \tag{2.14}$$

$$PM_{(NMC)} \approx 60^\circ \tag{2.15}$$

The clear disadvantage is that the assumption that g_{m3} is much larger than g_{m2} might not be necessarily true in a low-voltage design [Leu00]. Also, the only possibility to increase the UGF is to increase the transconductance of the last stage, further increasing the power consumption. This is particularly relevant in the case of large capacitive loads.

The NMC approach has withstood the test of time and its wide acceptance in industry is a prove of its longevity. A new scheme with better bandwidth and slew rate to power ratio is presented in the next section. This new compensation method tries to overcome some of the preceding limitations.

2.4 Proposed Compensation

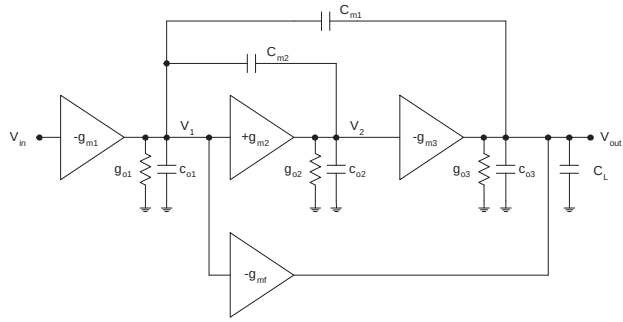
In this section, detailed analysis of the Positive Feedback Compensation (PFC) amplifier is presented. First, the working principle is presented in Section 2.4.1. This is followed by a detailed analysis of the transfer function in Section 2.4.2. The unity gain frequency and the phase margin are discussed in Section 2.4.3. Next, the closed loop transfer function in unity gain feedback is presented in Section 2.4.4 which is later used in Section 2.4.5 to derive the criteria for unconditional stability of the amplifier. This is followed by the design equations of the proposed frequency compensation in Section 2.4.6. The influence of the compensation capacitors value on the power consumption is presented in Section 2.4.7.

2.4.1 Working Principle

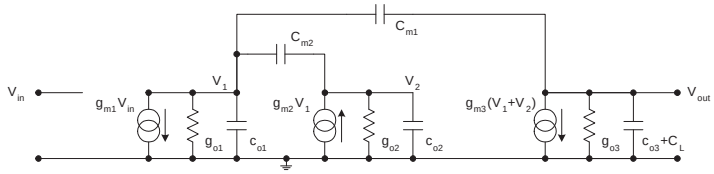
With the NMC amplifier (Fig. 2.5), a new Miller capacitor is added for compensation purposes for every new gain stage added. However, this new capacitor reduces the bandwidth by a factor of two [Esc92]. Moreover, both capacitors load the output, and thus, increase the power requirements for the last stage in order to fulfill the bandwidth and slew requirements. If capacitor C_{m2} is not present, a higher bandwidth is possible, but not without having an unstable amplifier, as now a pair of complex poles with a small damping ratio appear close to the unity gain frequency. If a circuit could be conceived that allows to: control the damping ratio, does not load the output and does not increase the circuit consumption, then the peaking would be eliminated without reducing the bandwidth. The proposed solution for frequency compensation is depicted in Fig. 2.8(a).

- The feed-forward transconductance g_{mf} bypasses all but the first stage at high frequencies to provide a direct path to the output. Consequently, this boosts the bandwidth of the PFC amplifier. This block is implemented using a single MOS transistor (g_{mf} in Fig. 2.11), driven by the output of the first stage and connected to the output node. By using this approach to implement g_{mf}, it is ensured that there is no increase in power consumption and silicon area when comparing it to the NMC [Esc92] counterpart.

- A positive feedback around g_{m2} allows effective control of the damping ratio of the complex poles, and capacitor C_{m2} (2.27) fulfills this condition. This capacitor



(a) Block diagram of the proposed PFC amplifier.



(b) Equivalent small-signal circuit of the PFC amplifier.

Figure 2.8: The PFC amplifier.

is significantly smaller than C_{m1}, and thus, does not limit the slew rate of the first stage. As an extra benefit, a LHP zero is created which helps in improving the phase margin.

The g_m feed-forward has been widely used to maximize the amplifier bandwidth [Esc92, Per93, You97] and will be used here as well, but to the authors knowledge this is the first amplifier to exploit the positive feedback to control the damping ratio in a three stage amplifier.

2.4.2 Small-Signal Transfer Function

Referring to Fig. 2.8(a), the small-signal elements $g_{m(1-3)}$, $g_{o(1-3)}$ and $c_{o(1-3)}$ represent transconductance, output conductance and output capacitance respectively. The two compensation capacitors are represented by C_{m1} and C_{m2}. C_L is the load capacitor. The output conductance g_{o3} includes the output resistive load.

With the purpose of having symmetrical current capability in the output stage of Fig. 2.11 and noting that g_{m3} and g_{mf} are in the same branch, both transconductances in Fig. 2.8(a) are set to have equal values: $g_{mf} = g_{m3}$. The small-signal circuit is thus simplified to

the one presented in Fig. 2.8(b).

The nodal equations that describe the amplifier are given by

$$\begin{cases} g_{m1}V_{in} + V_1 g_{o1} + V_1 s c_{o1} + (V_1 - V_{out})sC_{m1} + (V_1 - V_2)sC_{m2} = 0 \\ -g_{m2}V_1 + V_2 g_{o2} + V_2 s c_{o2} + (V_2 - V_1)sC_{m2} = 0 \\ g_{m3}V_2 + V_{out} g_{o3} + V_{out} s C_L + g_{m3}V_1 + (V_{out} - V_1)sC_{m1} = 0 \end{cases} \tag{2.16}$$

After grouping the terms, the previous system of equations is now given by

$$\begin{cases} (g_{o1} + sC_{nV_1})V_1 & -sC_{m2}V_2 & -sC_{m1}V_{out} & = & -g_{m1}V_{in} \\ (g_{m2} + sC_{m2})V_1 & -(g_{o2} + sC_{nV_2})V_2 & & = & 0 \\ (g_{m3} - sC_{m1})V_1 & +g_{m3}V_2 & +(g_{o3} + sC_{nV_{out}})V_{out} & = & 0 \end{cases} \tag{2.17}$$

where for simplicity the following parameters have been defined

$$C_{nV_1} = c_{o1} + C_{m1} + C_{m2} \tag{2.18}$$

$$C_{nV_2} = c_{o2} + C_{m2} \tag{2.19}$$

$$C_{nV_{out}} = C_L + C_{m1} \tag{2.20}$$

The transfer function can now be calculated for the arrangement of Fig. 2.8. The transfer function of the open loop gain is quite long but has the general form of

$$G_o(s) = \frac{A_o\left(1 - \frac{s}{z_1}\right)\left(1 - \frac{s}{z_2}\right)}{\left(1 + \frac{s}{\omega_{do}}\right)\left(1 + 2\zeta_o\left(\frac{1}{\omega_{no}}\right)s + \left(\frac{1}{\omega_{no}^2}\right)s^2\right)} \tag{2.21}$$

After simplification of (2.17) without loosing accuracy, it can be expressed by

$$G_o(s) = \frac{A_o\left(1 + \frac{2C_{m2}}{g_{m2}}s - \frac{C_{m1}C_{m2}}{g_{m2}g_{m3}}s^2\right)}{\left(1 + \frac{s}{\omega_{do}}\right)\left(1 + \frac{C_{m2}(2g_{m3}C_{m1} - g_{m2}C_L)}{g_{m2}g_{m3}C_{m1}}s + \frac{C_L(C_{m2} - c_{o2})}{g_{m2}g_{m3}}s^2\right)} \tag{2.22}$$

During the derivation of the above equation, it is assumed that: 1) C_{m1}, C_{m2} and $C_L \gg c_{o(1-3)}$; 2) in a polynomial, the effect of a factor containing the output conductance is negligible when compared to one containing a transconductance, and 3) $C_L > C_{m1}$, C_{m2}.

After simplification, the DC gain is given by

$$A_o = \frac{g_{m1}g_{m2}g_{m3}}{g_{o1}g_{o2}g_{o3}} \tag{2.23}$$

with a dominant pole at

$$\omega_{do} = \frac{g_{o1}g_{o2}g_{o3}}{C_{m1}g_{m2}g_{m3}} \tag{2.24}$$

The overall transfer function includes two zeros: one LHP zero and one RHP zero. The latter will have a value one order of magnitude higher than the UGF, minimizing the degradation in the phase margin; a value greater but close to $-\zeta_o\omega_{no}$ for the zero in the LHP enhances the overall phase margin. The convenient location of the zeros

$$z_{1,2} = \left(\frac{g_{m3}}{C_{m1}}\right)\left(1 \mp \sqrt{1 + \frac{C_{m1}}{C_{m2}}\frac{g_{m2}}{g_{m3}}}\right) \tag{2.25}$$

significantly improves the overall stability of the amplifier.

The determination of the values for the second-order system in the denominator of (2.22) follows the assumption that the amplifier has a third order Butterworth frequency response in a unity gain feedback configuration [Esc92]. Although the previous method does not take the zeros' effect into account, due to the existence of z_1 close to the UGF, its effect is included. As a result, after positioning the complex conjugated poles, the zeros are thereafter considered. By comparing the denominator of (2.5) and (2.6) with (2.22) the following set of equations can be written

$$\begin{cases} 2\zeta_o\dfrac{1}{\omega_{no}} = \dfrac{1}{\omega_c} = \dfrac{C_{m2}(2g_{m3}C_{m1} - g_{m2}C_L)}{g_{m2}g_{m3}C_{m1}} \\ \dfrac{1}{\omega_{no}^2} = \dfrac{1}{2\omega_c^2} = \dfrac{C_L(C_{m2} - c_{o2})}{g_{m2}g_{m3}} \end{cases} \tag{2.26}$$

from which the following open loop parameters are obtained

$$\zeta_o = \frac{g_{m3}C_{m1}(2C_{m2} + c_{o2}) - g_{m2}C_{m2}(C_L + C_{m1})}{2C_{m1}\sqrt{g_{m2}g_{m3}C_L(C_{m2} + c_{o2})}} \tag{2.27}$$

$$\omega_{no} = \sqrt{\frac{g_{m2}g_{m3}}{C_L(C_{m2} + c_{o2})}} \tag{2.28}$$

and the pole zero diagram drawn as shown in Fig. 2.9.

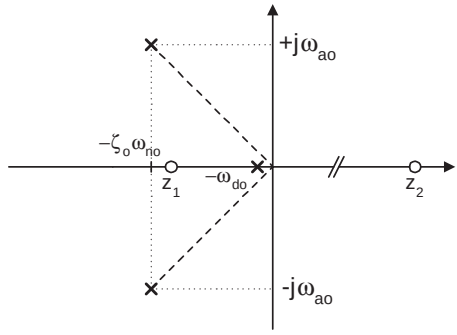


Figure 2.9: PFC pole zero diagram.

2.4.3 Unity Gain Frequency and Phase Margin

During the derivation of the UGF, the influence of the zero z_1 in Fig. 2.9 cannot be overlooked as it is less than this value. This causes an increase in the PM and an underestimation of the unity gain frequency if the simple equation $UGF = g_{m1}/C_{m1}$ is used. Consequently, not only the DC gain and the dominant pole but also z_1 and ω_{no} in (2.21) must be used to obtain an accurate equation. From (2.22) to (2.28), the unity gain frequency of the PFC amplifier can be found to be

$$\begin{cases} UGF = a\dfrac{g_{m1}}{C_{m1}} \\ UGF = b\dfrac{1}{2\sqrt{2}}\omega_{no} \end{cases} \quad \begin{cases} UGF = 1.46 \times \left(\dfrac{g_{m1}}{C_{m1}}\right) \\ UGF = 0.875 \times \omega_{no} \end{cases} \tag{2.29}$$

where a and b are corrective values to the common equations in (2.29), necessary to take the zero z_1 influence into account.

Solving the previous system with the result from (2.28) in order to obtain UGF as a function of the circuit's small-signal parameters

$$UGF_{(PFC)} = \left(\frac{g_{m1}}{C_{m1}}\right) \times 1.46 \tag{2.30}$$

$$= \frac{7}{8}\sqrt{\frac{g_{m3}g_{m2}}{C_L(C_{m2}+c_{o2})}} \tag{2.31}$$

$$= \frac{1}{4}\left(\frac{g_{m3}}{C_L}\right) \times \beta \tag{2.32}$$

$$PM_{(PFC)} \approx 60^\circ \tag{2.33}$$

where

$$\beta = \frac{7}{2}\sqrt{\left(\frac{g_{m2}}{g_{m3}}\right)\left(\frac{C_L}{C_{m2}+c_{o2}}\right)} \tag{2.34}$$

If (2.14) with (2.32), equations describing the UGF of both the NMC and PFC are compared, it is seen that a β factor appears. Reference [Leu00] uses this relation to justify that this criteria can be used to measure the bandwidth improvement over the NMC. On the contrary, β only indicates a relation between the transconductances of the output stage for both topologies. The UGF of an amplifier depends on the relative position of all poles and zeros, and not only on the relation between one transconductance. Nevertheless, if the prevailing source of power dissipation in an amplifier is the output stage, as indicated by (2.34), maximizing g_{m2} with respect to g_{m3} and minimizing parasitic capacitors at the output of the second stage makes this approach more efficient. In applications where one needs to drive large capacitive loads, the value of β can be large since it is proportional to $\sqrt{C_L}$, making the PFC more efficient than the NMC.

2.4.4 Closed Loop Transfer Function

The analysis can be further extended and the derivation of the closed loop transfer function obtained which, assuming a DC feedback f_{dc}, is

$$G(s) = \frac{G_o(s)}{1 + f_{dc}G_o(s)} \tag{2.35}$$

The above result can also be written with the explicit indication of the closed loop parameters: dominant pole (ω_d), undamped natural frequency (ω_n) and damping ratio parameter (ζ)

$$G(s) = \frac{A\left(1+\frac{s}{z_1}\right)\left(1+\frac{s}{z_2}\right)}{\left(1+\frac{s}{\omega_d}\right)\left(1+2\zeta\left(\frac{1}{\omega_n}\right)s+\left(\frac{1}{\omega_n^2}\right)s^2\right)} \tag{2.36}$$

After manipulation and simplification an approximated solution for the small-signal open loop transfer function in (2.35) is calculated as

$$G(s) = \frac{A\left(1+\frac{s}{z_1}\right)\left(1+\frac{s}{z_2}\right)}{1+\frac{C_{m1}}{g_{m1}}s+\frac{C_{m2}(2g_{m3}C_{m1}-g_{m2}C_L)}{g_{m1}g_{m2}g_{m3}}s^2+\frac{C_{m1}C_L(C_{m2}-c_{o2})}{g_{m1}g_{m2}g_{m3}}s^3} \tag{2.37}$$

One of the uses of the above equation is the determination of the system stability which is done in the next sub-section.

2.4.5 Stability Analysis

The linear system given by (2.22) can have its stability determined without the explicit solution of the closed loop characteristic equation using the Routh-Hurwitz stability criterion [Dor04]. The characteristic equation, representing a polynomial in s of the denominator of (2.37) can be written as

$$D(s) = a_0 s^3 + a_1 s^2 + a_2 s^1 + a_3 \tag{2.38}$$

The system is unstable if any of the coefficients are zero or if positive and negative terms exist simultaneously. Otherwise, nothing can be concluded and the coefficients should be arranged in rows and columns in the following matrix form [Dor04]:

$$\begin{array}{c|cc} s^3 & a_0 & a_2 \\ s^2 & a_1 & a_3 \\ s^1 & b_1 & \\ s^0 & c_1 & \end{array} \tag{2.39}$$

After substitution by the coefficient in the denominator of the closed loop transfer function, the above elements in the matrix are written as

$$\begin{array}{c|lcl lcl} s^3 & a_0 & = & \dfrac{C_{m1}C_L(C_{m2}-c_{o2})}{g_{m1}g_{m2}g_{m3}} & a_2 & = & \dfrac{C_{m1}}{g_{m1}} \\ s^2 & a_1 & = & \dfrac{C_{m2}(2g_{m3}C_{m1}-g_{m2}C_L)}{g_{m1}g_{m2}g_{m3}} & a_3 & = & 1 \\ s^1 & b_1 & = & (a_1a_2 - a_0a_3)/a_1 & & & \\ s^0 & c_1 & = & (b_1a_3)/b_1 & & & \end{array} \tag{2.40}$$

For stability, all the coefficients in the first column of the Routh-Hurwitz table must be positive, i.e., it is required

$$\begin{cases} C_{m2} > c_{o2} \\ 2g_{m3}C_{m1} > g_{m2}C_L \end{cases} \tag{2.41}$$

With a_3 equal to one, $c_1 = b_1$ in (2.40). The solution from the remaining coefficient has to be approximated as it does not have a simple representation.

$$(2g_{m3} - g_{m2})C_{m1} > g_{m1}C_L \tag{2.42}$$

2.4.6 Design Equations

Finally, all design equations can be obtained in a simple and straightforward form. Using (2.29) the first stage transconductance is obtained. The remaining variables are derived from (2.27), (2.28) and (2.29). After simplification they are

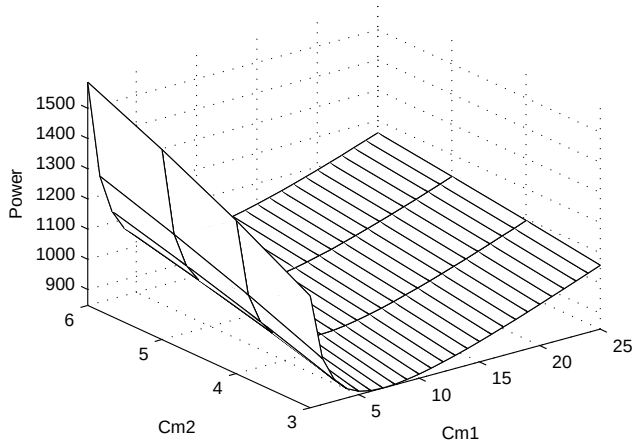


Figure 2.10: Influence of the compensation capacitors on the total power consumption of the PFC amplifier.

$$g_{m1} = \frac{C_{m1} \cdot UGF}{1.46} \tag{2.43}$$

$$g_{m2} = \sqrt{2} \cdot UGF \cdot \frac{4}{7} k \tag{2.44}$$

$$g_{m3} = \sqrt{2} \cdot UGF \cdot \frac{8}{7} C_L C_{m2} \frac{1}{k} \tag{2.45}$$

$$k = \frac{C_{m1} C_L}{C_{m1} + C_L} \left(\sqrt{1 + 4 C_{m2} \frac{C_{m1} + C_L}{C_{m1} C_L}} - 1 \right) \tag{2.46}$$

Using the design equations above, the stability conditions presented in sub-section 2.4.5 are always verified. Electrical simulators are also useful in this situation of evaluating possible sensitivity to element variation. The Monte Carlo analysis (for all devices) of the amplifier around each corner of the bias current ($\pm$ 20 %) and compensation capacitors ($\pm$ 20 %) have shown no unexpected behavior.

2.4.7 Power Optimization

With the design equations available a simple study of the power consumption as a function of the different compensation capacitors value is possible. In Fig. 2.10 is depicted the result of such a study for a fixed load capacitor. As is clear, increasing the load capacitor necessarily requires more power. A more in-depth analysis of the power consumption and optimization is given in Chapter 3.

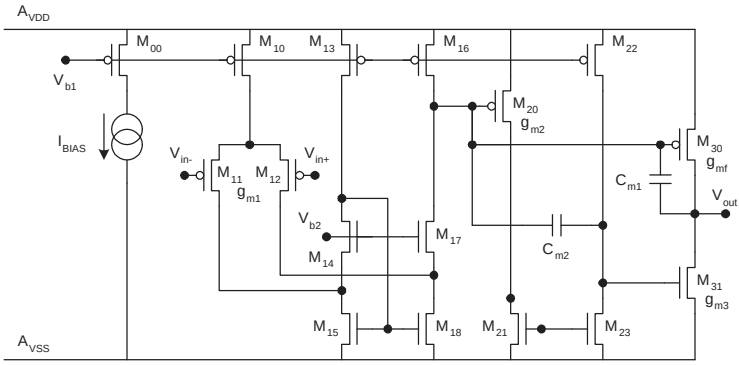


Figure 2.11: Schematic diagram of the PFC amplifier.

2.5 Design of the Operational Amplifier

This section describes the implementation, layout and measurement of the PFC amplifier produced in a 0.35 µm 5M2P CMOS technology. The typical threshold voltages for the NMOS and PMOS transistors are 0.59 V and -0.57 V, respectively.

The experimental results are now presented. Firstly, the component sizes and layout are presented in sub-section 2.5.1. This is followed by the test setup used to characterize the amplifier performance. The measurement results are then presented in sub-section 2.5.3. In brief, the amplifier achieves a UGF of 2.7 MHz with a PM of 52°. An average slew rate of 1.0 V/µs is obtained while loaded by 130 pF//24 kΩ. With a total power consumption of 275 µW, this proves to be the amplifier with the highest figure of merit at the date of publication. Finally, the performance of the implemented PFC amplifier is compared to that of other three stage amplifiers.

2.5.1 Circuit Implementation and Layout

The circuit of Fig. 2.11 is intended to demonstrate the proposed compensation scheme of Fig. 2.8(a). The final component sizes used are presented in Table 2.1.

The bandwidth of the amplifier is chosen to be within the range of the testset for automatic characterization of opamps [Gri93]. The low frequency value of the UGF does not pose significant problems to the layout. Care is taken to minimize parasitic capacitances in the most sensitive nodes. In general, the layout is made compact without impairing the resistive losses due to metal resistance and inclusion of sufficient substrate contacts for proper definition of the local ground.

A microphotograph of the PFC amplifier is shown in Fig. 2.12. As can been seen, most of the area (0.03 mm^2) in the circuit is occupied by the compensation capacitors.

Table 2.1: Component values used in the PFC amplifier.

Parameter	Unit	Value
C_{m1}	pF	15
C_{m2}	pF	3
M^*_{00-10}	µm/µm	10/0.7
M_{11-12}	µm/µm	40/0.7
M^*_{13-16}	µm/µm	10/0.7
M_{14-17}	µm/µm	6/0.7
M_{15-18}	µm/µm	11/0.7
M_{20}	µm/µm	15/0.7
M_{22}	µm/µm	10/0.7
M_{21-23}	µm/µm	2/0.7
M^*_{30}	µm/µm	15/0.7
M^*_{31}	µm/µm	15/0.5
*: multiplication factors of transistors M_{10}, M_{13}, M_{16}, M_{30} and M_{31} are 6, 3, 3, 22 and 5 respectively.		

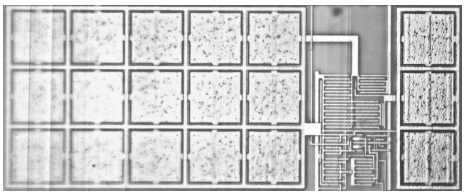


Figure 2.12: Microphotograph of the designed operational amplifier.

2.5.2 Test Setup

The amplifier is to be tested both in the frequency domain and in DC. For this purpose the fabricated die is mounted on a DIL package that fits in the testset [Gri93]. An external capacitor is added in such a way that the total capacitance plus the parasitic capacitance in the testset add up to a total value of 130 pF. A resistive load of 24 kΩ is also included in parallel.

The transient response is measured using the same amplifier connected in voltage follower configuration. A square wave signal caused the amplifier to enter in slew rate, thus making it possible to measure this characteristic. In order not to load the amplifier, an active probe is used for a non-intrusive measure. A digital oscilloscope is used to record the waveform. With subsequent analysis, the settling time is obtained.

2.5.3 Measured Results

The results presented next are obtained using a supply voltage of 1.5 V and a load consisting of a 130 pF capacitor and 24 kΩ resistor. The DC biasing current used is

5 µA. Although an operational amplifier can be considered a system on its own, where different measurements covering various aspects of the design are possible, only the most common characteristics are measured.

The frequency response is shown in the Bode diagram in Fig. 2.13. A UGF of 2.7 MHz with a PM of 52 ° is measured with an HP3577A network analyzer.

The measurement of the transient response in unity gain feedback is shown in Fig. 2.14 and is done with a Tektronix TDS680B oscilloscope. Both the positive and negative slew rate is measured to be 1.0 V/µs. From the figure, it is possible to see that the amplifier output settles to the final value after a small transient. In this case, the positive and negative settling time to 1 % final settling time error is 1.4 µs and 1.0 µs, respectively.

A total quiescent power consumption of 275 µW is measured. This and the other measurements are summarized in Table 2.2.

2.5.4 Discussion of the Results

Considering that each amplifier has its own characteristics, comparing them requires a figure of merit (FOM) that weighs the tradeoff between the bandwidth, load capacitance, slew rate and power consumption. Two of them will be used: one for small-signal [Ng99] and one for large signal performance [Leu00].

$$FOM_S = \frac{UGF_{[MHz]} \cdot C_{L[pF]}}{Power_{[mW]}} \tag{2.47}$$

$$FOM_L = \frac{SR_{[V/\mu s]} \cdot C_{L[pF]}}{Power_{[mW]}} \tag{2.48}$$

The units are shown between brackets and the average value of the SR is used for the calculations. From (2.47) and (2.48) it can be concluded that the higher the FOM, the better the amplifier.

Table 2.3 presents an overview of multistage amplifiers available in the open literature. Based on measured data, both figures of merit are also given. As can be seen, the figure of merit varies by more than two orders of magnitude. Different topologies and how much emphasis has been placed on the optimization can justify this effect. Another reason is the different values of the compensation capacitors, as this has a large influence [Ram03a] on the slew rate and total power consumption.

From the results presented in Table 2.3, it can be concluded that the PFC amplifier outperforms the NMC. Also, the proposed amplifier compares well with other topologies. As indicated by equations (2.29) and (2.32), a higher load capacitor and smaller compensation capacitors (C_{m1}, C_{m2}) will allow a higher FOM_S.

2.6 Conclusion

This chapter has presented the design and measurement results of a novel operational amplifier frequency compensation, where the damping factor is controlled by the posi-

Table 2.2: Measured Performance.

Parameter	Value
Low Frequency Gain	>100 dB
Unity Gain Frequency	2.7 MHz
Phase Margin	52°
Positive Slew Rate	1.0 V/µs
Negative Slew Rate	1.0 V/µs
Positive Settling Time (to 1 %)	1.4 µs
Negative Settling Time (to 1 %)	1.0 µs
Power Consumption	275 µW
Power Supply	±0.75 V
Load Condition	130 pF//24 kΩ
Area	0.03 mm^2

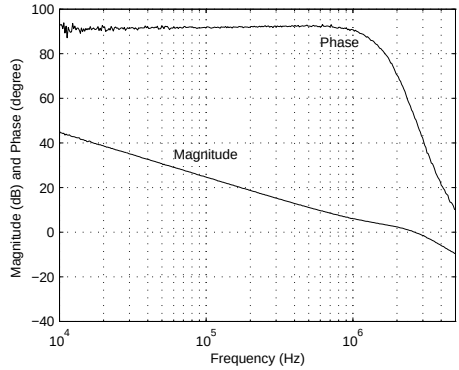


Figure 2.13: Measured frequency response of the PFC amplifier. The bump in the frequency response is the result of the zero in LHP.

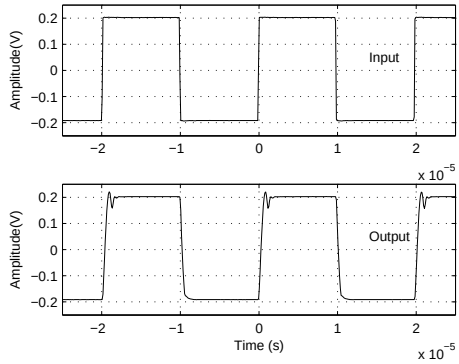


Figure 2.14: Measured transient response of the amplifier configured as a unity gain follower (130 pF//24 kΩ).

Table 2.3: Performance Comparison of Different Multistage Amplifiers

	Gain (dB)	UGF (MHz)	$Power$ (mW@Vdd)	SR (V/µs)	C_L (pF)	FOM_S $(\frac{MHz.pF}{mW})$	FOM_L $(\frac{V/\mu s.pF}{mW})$	C_{m1} (pF)	C_{m2} (pF)	Technology
MPD [Fon91]	117	3.4	0.7@1	1.1	100	485	157	6	6	3/1 GHz f_T BiCMOS
NMC [Esc92]	100	60	76@8	20	100	79	26	20	6	3 GHz f_T BJT
MNMC [Esc92]	100	100	76@8	35	100	131	46	11	8	3 GHz f_T BJT
NNMC [Per93]	—	2	10@5	1.5	250	50	37	17.5	5+12	1.5 µm CMOS
NGCC [You97]	100	1	1.4@2	5	20	14	71	—	—	2 µm CMOS
ETC [Ng99]	102	47	6.9@3	69	40	272	400	—	—	0.6 µm CMOS
DFCFC [Leu00]	>100	2.6	0.42@2	1.32	100	619	314	18	3	0.8 µm CMOS
NGRNMC [Pen02]	100	80	7.8@2.6	56	16	164	115	0.43	0.6	0.35 µm CMOS
AFFC [Lee03a]	>120	4.5	0.4@2	1.49	120	1350	447	7	3	0.8 µm CMOS
DLPC [Lee03b]	>120	7	0.33@1.5	3.3	120	2545	1200	4.8	2.5	0.6 µm CMOS
TCFC [Pen04]	>100	2.85	0.045@1.5	1.04	150	9500	3450	1.1	0.92	0.35 µm CMOS
This work	>100	2.7	0.275@1.5	1.0	130	1276	473	15	3	0.35 µm CMOS

tive feedback capacitor C_{m2} around g_{m2}. A direct path to the output through g_{mf} boosts the bandwidth without increasing the power consumption or increasing the silicon area.

The implemented amplifier has been fabricated in a commercial 0.35 µm CMOS technology. The measurement results show that it achieves a UGF of 2.7 MHz with a PM of 52° while driving a 130 pF load. The total power consumption of the amplifier is 275 µW and it occupies a total area of 0.03 mm^2.

Theoretical analysis has been presented to support the amplifier's stability despite the positive feedback around the second stage transconductance. Measured results have shown good agreement with theory and no oscillations or instability have been seen. The presented design equations allow to quickly design an amplifier following the PFC configuration. Measurements done at low-voltage show its suitability for circuits simultaneously requiring high DC gain and high power efficiency in standard CMOS technologies. During measurements the amplifier has proved to be working up to a supply voltage of 1.25 V. This voltage is limited by the correct operation of the input differential pair.

Comparing different implementations, being each one of them done by a different designer is always a complex issue. Usually, a figure of merit is used which must weight different tradeoffs. The main difficulty is that each architecture put emphasis on a different topic. These can up to a certain point justify the huge difference in the values of the figure of merit seen in Table 2.3. More than two orders of magnitude is seen. It is thus unthinkable that these results bare all point in a operational amplifier design. Matters like the consequence of the multiple feedback and feed-forward loops are commonly overlook. Noise, PSRR, CMRR, and settling performance are not considered. Although these aspects need to be answered when including the block on an application, their analysis is lengthy and outside of the scope of this thesis. Nevertheless, two common performances, the UGF and settling time are studied.

In the next chapter, theoretical analyses of the PFC, the NMC and four other frequency compensation topologies are sized for the same goals under equal design conditions. In this manner it is possible to avoid the influence of different compensation and load capacitors, as well as different optimization emphasis and technology.

Chapter 3

Topology Study of Three Stage Amplifier Frequency Compensation

3.1 Introduction

In the three stage amplifier field, the NMC amplifier is usually taken as the reference. The proponents in [Esc92] present reasons for the four times reduction in the bandwidth in comparison to the single stage amplifier. Each time a new stage is added, an extra capacitor is added from the output to close the external feedback loop. Since then, different topologies have been presented that tried to overcome this limitation.

Another important factor in integration is the size. In an multistage operational amplifier, compensation capacitors generally take a considerable amount of the total circuit. It is important to minimize their value to the bare minimum required by design criteria such as sensitivity to parameters variation, technology guidelines or stability.

Over the years numerous new topologies have been presented in the open literature. They have addressed issues like bandwidth extension, phase margin increase, better suitability for low-power design or size reduction. The common characteristic of most of them is the existence of some kind of feed-forward to compensate the bandwidth reduction. Also, the two compensation loops existing in the NMC are usually avoided or connected in a different way.

The proliferation of all these new solutions necessitated the proposal of a figure of merit. This number allowed a comparison between different works from different designers. Some interesting results have nevertheless appeared. Even for the same topology, where results are expected to be similar, results surpassing 80 % have been shown [Esc92, Leu00]. Compensation capacitors as large as the load capacitor have also been given [Leu00] as well as compensation capacitors which have a considerable unequal value [Esc92, Lee03a]. Equally important is that the measurements are performed on different loads while not accounting for its influence on the amplifier's efficiency.

With topologies claiming improvement over the NMC in the order of 10 to 20 in the bandwidth, while the NMC is known to be four times less efficient than the single stage amplifier, this boils down to the fact that these topologies will consume between two tenths and four tenths ([0.2, 0.4]) of the one stage amplifier for the same Unity Gain Frequency (UGF). Furthermore, each new design is shown to be considerably smaller, and as such, suitable for high integration in CMOS. Missing however is the effect these variations have on the amplifier's specifications. These effects are visible in the data presented in Table 2.3.

Without a theoretical analysis of the amplifiers under the same conditions, it is not possible to say which one is undoubtedly the most efficient, or to know how each design criteria influences each performance.

This chapter intends to give a few guidelines to the designer in the task of better choosing the amplifier and throw some light on how the results in Table 2.3 are possible.

In this chapter, each one of the other four compensation schemes is first presented in Section 3.2, with some of its characteristics and the sizing equations. They are later compared with the NMC structure and the Positive Feedback Compensation (PFC) topology proposed in the previous chapter.

In Section 3.3 a different view of the sizing procedure than the one presented in the previous chapter, where equation are used, is given. Firstly the NMC and PFC are theoretically compared on some common characteristics and optimized on the block level. Later, the PFC amplifier is fully optimized in the transistor level.

The chapter ends with conclusions in Section 3.4

3.2 Comparison to other Topologies

This section introduces the method for topology comparison between three stage amplifier frequency compensation schemes. In sub-section 3.2.1 the rules with which each amplifier is sized are given along with the method for power estimation. As described next, the influence of the compensation capacitors and load capacitor in the total power dissipation is minimized by sizing each one of them for the same conditions. The two compensation topologies described in the previous chapter and four new schemes are then compared.

In sub-section 3.2.2 each compensation topology is first briefly presented and sized for the maximum UGF. Afterwards, in sub-section 3.2.3, the transient step response is given for the resulting amplifier. Finally, some comments are given in sub-section 3.2.4 concerning the analysis presented here.

3.2.1 Topology Comparison and Power Estimation

Performance comparison is the key point in this analysis. Each one of the compensation schemes will be optimized for the same unity gain frequency, using the same method.

The small-signal elements $g_{m(1-3)}$, $g_{o(1-3)}$ and $c_{o(1-3)}$ represent transconductance, output conductance and output capacitance, respectively. Each amplifier has two compensation capacitors: $C_{m1/a}$ and C_{m2}. The load capacitor is represented by C_L and the output conductance g_{o3} includes the output resistive load.

The overall power consumption for each amplifier will be the sum of all transconductances, each one of them multiplied by an integer to take into account its transistor level implementation. For a folded cascode pair, a value of 4 will be used: 2 branches with the input transistors and 2 branches with the cascode. A positive gain stage will contribute 2 times. The output stage usually has only 1 branch, so the output transconductance contributes only once to the power consumption.

Table 3.1: Optimization Parameters

Parameter	Value
UGF	1 MHz
PM	>60°
c_{o1}	0.05 pF
c_{o2}	0.2 pF
C_L	100 pF

Considering that each amplifier has its own characteristics, comparing them requires a figure of merit that weight the tradeoff between the bandwidth, load capacitance, slew rate and power consumption. Two of them will be used: one for small-signal (FM_S) and one for large-signal (FM_L) performance. A higher value for (3.1) and (3.2) indicates a more efficient amplifier. For different load capacitors, supply voltages and UGF, the figures of merit have to be redefined.

$$FM_S = 1/Power_{[\mu A/V]} \tag{3.1}$$

$$FM_L = \frac{SR_{[\mu A/pF]}}{Power_{[\mu A/V]}} \tag{3.2}$$

The equations presented in each one of the papers describing each topology will be used as dimensioning criteria in conjunction with the optimization goals indicated in Table 3.1. For the DFCFC and AFFC, β [Leu00] is selected so the amplifier has at least the same FM_L as the NMC amplifier.

3.2.2 Unity Gain Frequency

3.2.2.1 NMC

The Nested Miller Compensation (NMC) amplifier has been discussed in Section 2.3. Only the sizing equations are thus given.

$$g_{m1} = C_{m1}UGF \tag{3.3}$$

$$g_{m2} = 2C_{m2}UGF \tag{3.4}$$

$$g_{m3} = 4C_LUGF \tag{3.5}$$

$$Power = 4g_{m1} + 2g_{m2} + g_{m3} \tag{3.6}$$

3.2.2.2 DFCFC

The structure of a Damping Factor Control Frequency Compensation (DFCFC) amplifier is shown in Fig. 3.1. The DFCFC [Leu00] tries to solve some of the inconveniences found in the NMC amplifier by using a feed-forward transconductance stage g_{mf} to increase the high frequency gain. A higher bandwidth is achieved by removing C_{m2} and

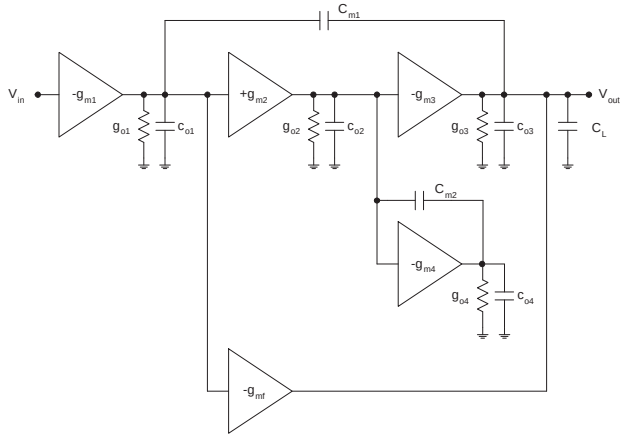


Figure 3.1: The DFCFC amplifier topology.

controlling the damping ratio of the nondominant complex pole by means of g_{m4} and C_{m2}. Slewing at the output is improved, as now g_{mf} can be used to implement a push-pull output stage. Another characteristic is that the poles and zeros do not depend on C_{m2}. A new degree of freedom, β appears. One drawback is that g_{m4} depends on a parasitic capacitor.

$$g_{m1} = C_{m1}UGF \tag{3.7}$$

$$g_{m2} = \frac{(2C_LUGF - g_{m3})g_{m4}}{g_{m3}} \tag{3.8}$$

$$g_{m3} = \frac{4C_LUGF}{\beta} \tag{3.9}$$

$$g_{m4} = g_{m3}\frac{c_{o2}}{C_L}\beta = 4c_{o2}UGF \tag{3.10}$$

$$\beta = g_{m3(NMC)}/g_{m3(DFCFC)} \tag{3.11}$$

$$Power = 4g_{m1} + 2g_{m2} + g_{m3} + 2g_{m4} \tag{3.12}$$

3.2.2.3 AFFC

The Active Feedback Frequency Compensation (AFFC) is proposed to improve the bandwidth of multistage amplifiers. A structure for the case of having three stages is depicted in Fig. 3.2 [Lee03a]. The feed-forward transconductance stage g_{mf} is used to increase the high frequency gain. The right half plane is removed as g_{ma} blocks the path from the output of the first stage to the output of the amplifier through C_a. In this topology, the parasitic capacitance c_{o1} is utilized to improve performance at the expense of a greater sensitivity.

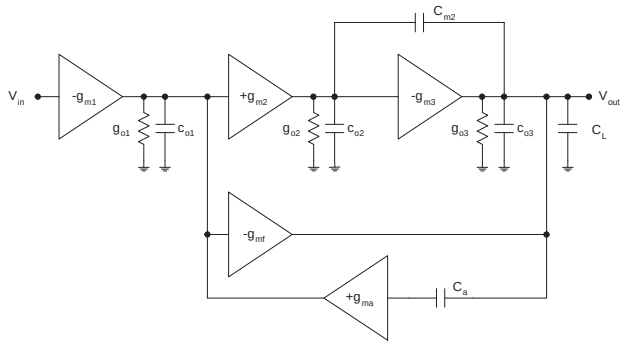


Figure 3.2: Structure of a AFFC amplifier.

$$
\begin{aligned}
g_{m1} &= C_a UGF && (3.13)\\
g_{m2} &= g_{m3} - \frac{g_{m3}^2 \beta^2 c_{o1}}{8 C_L g_{m1}} && (3.14)\\
g_{m3} &= \frac{4 C_L UGF}{\beta} && (3.15)\\
g_{ma} &= 4 g_{m1} && (3.16)\\
\beta &= g_{m3(NMC)} / g_{m3(AFFC)} && (3.17)\\
Power &= 4 g_{m1} + 2 g_{m2} + g_{m3} + g_{ma} && (3.18)
\end{aligned}
$$

3.2.2.4 TCFC

The three stage Transconductance with Capacitances Feedback Compensation (TCFC) amplifier in Fig. 3.3 is included in this set of comparisons because of its high figure of merit value as its design procedure does not follow the same method as all the others. The pair of complex conjugated poles do not necessarily have to have a damping ratio of $1/\sqrt{2}$, nor does the amplifier have to have third other Butterworth frequency response with unity gain feedback [Pen04]. Instead, all non-dominant poles are supposed to be sufficiently above the UGF. This guarantees not only that their varying effect on the PM is minimized but also advantages in the transient response.

Again, as done with the DFCFC and AFFC, the output transconductance is increased beyond the value given by (3.21) so that it has at least the same performance as the NMC amplifier. The design equations are given as follows

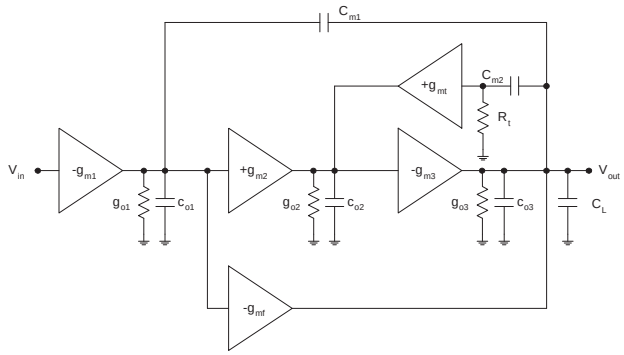


Figure 3.3: Block diagram of the TCFC amplifier.

$$
\begin{aligned}
g_{m1} &= C_{m1} UGF && (3.19)\\
g_{m2} &= 2g_{mt} && (3.20)\\
g_{m3} &= \frac{8C_L UGF}{(1+k_t)} \frac{c_{o2}}{C_{m2}} && (3.21)\\
g_{mt} &= \frac{(1+k_t)}{k_t} 2C_{m2} UGF && (3.22)\\
k_t &= 2 && (3.23)\\
Power &= 4g_{m1} + 2g_{m2} + g_{m3} && (3.24)
\end{aligned}
$$

3.2.2.5 NMCF

One characteristic common to all compensation topologies presented so far is the existence of a feed-forward transconductance g_{mf}. Its functionality is to bypass the last stages to provide a direct path to the output and as such to increase the bandwidth. As it is commonly used, this transconductance sits in the same branch as the g_{m3} which allows it to be produced without static power increase. It seems in this way interesting to compare the reference NMC topology with and without this frequency enhancement. It is named NMCF, short for NMC with feed-forward.

A procedure similar to Section 2.4 for the PFC amplifier is carried out. The open-loop small-signal transfer function of the Nested Miller Compensation with Feed-forward (NMCF) is firstly derived with sufficient accuracy. Once more, it is assumed that the amplifier has third order Butterworth frequency response with unity feedback. The design equations are then derived to be expressed as:

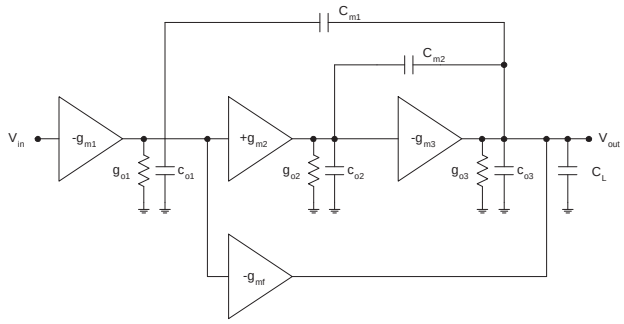


Figure 3.4: Three stage NMCF amplifier topology.

$$g_{m1} = C_{m1}UGF \quad (3.25)$$
$$g_{m2} = 4C_{m2}UGF \quad (3.26)$$
$$g_{m3} = 2C_LUGF \quad (3.27)$$
$$Power = 4g_{m1} + 2g_{m2} + g_{m3} \quad (3.28)$$

Comparing these equations with ones for the NMC amplifier, the effect of the feed-forward transconductance is reflected in the equations in a simple way: the transconductance of the second stage is multiplied by two while the transconductance from the last stage is divided by two. This affects the overall power consumption positively, as the output transconductance is the dominant source of power in the NMC.

3.2.2.6 PFC

The design equations for the Positive Feedback Compensation (PFC) topology are given in the set of equation (2.43)-(2.46) with the exception of the total power that is given as follows.

$$Power = 4g_{m1} + 2g_{m2} + g_{m3} \quad (3.29)$$

3.2.2.7 Simulation Results

The results of the equations describing each frequency compensation are presented in Table 3.2 and Table 3.3 for a fixed load capacitance and in Fig. 3.5 and Fig. 3.6 for a range of load capacitor values. Tabled values include the value of each transconductance and each node capacitance. With these values, the figure of merit given in this chapter are calculated. This procedure is repeated for each load capacitance value and presented graphically.

Comparing only the results presented in Table 3.2 with those given in Table 3.3, the effect in power saving is evident by simply using a lower value. Furthermore, efficiency

for the same topology also changes for a varying load capacitor. From Fig. 3.5 and Fig. 3.6 it seems clear that all topologies gain in terms of efficiency with an increasing load capacitor value.

From a slew rate perspective, the PFC and the NMCF are the topologies that allow to attain a more efficient design. If the slew rate is the limiting factor, the previous two compensation methods should be considered. On the other hand, if the constraint is maximizing the UGF, then the PFC, TCFC and DFCFC are the most efficient amplifiers, however, at a much lower improvement ratio than seen in Table 2.3. For smaller compensation capacitors, the previous group of three is reduced to only the TCFC and DFCFC as the PFC does not gain as much from a decrease in compensation capacitor C_{m1}.

A note on the performance improvement of the NMCF topology over the NMC is necessary. Without change in the circuit complexity, the power consumption is almost halved while the large-signal performance is significantly improved. If no limitation is found, this topology should replace the NMC amplifier as it is the one that most resembles it.

3.2.3 Settling Time

Equally important in an amplifier is the transient performance. Under the assumption that the input signal is small enough, the small-signal approximation of the amplifier closed loop transfer function is valid, and thus the amplifier response to a step input signal can be found.

The results for the case where the transconductances are taken from Table 3.2 and with C_{m1}=18 pF, C_{m2}=3 pF and C_L=100 pF is given in Fig. 3.7 for the six topologies presented so far. In this situation, the amplifiers have been sized for a given UGF. As such, evaluation of the transient behavior when sized is presented.

The first thing to note in Fig. 3.7 is that amplifiers having similar frequency bode diagrams have different step impulse responses. Visually, two types of responses exist: those that have overshot (underdamped) like the NMC, DFCFC and the PFC and those like the AFFC, TCFC and NMCF that have a critically damped or overdamped response.

From all the responses presented in Fig. 3.7, the TCFC is the fastest and the DFCFC the slowest. The others, in a decreasing order are: NMCF, NMC, AFFC and the PFC. The settling time worsens for the DFCFC for an increasing β. The first place as topology with the lowest power consumption for a given UGF comes at an expense of a poor transient response. The β has to increase (with a consequent increase in the power consumption) to improve the settling time.

Concerning the above, it must be added that the most power efficient topologies: DFCFC, TCFC and the PFC can have their settling time performance increased while still performing well in terms of power consumption for a UGF. This fact will be studied further in sub-section 3.3.1 where the NMC and PFC are optimized and it is shown that for the same power consumption, the PFC topology always has a higher UGF and also converges faster to the final value.

Table 3.2: Results for C_{m1}=18 pF, C_{m2}=3 pF and C_L=100 pF

	g_m (µA/V)	Node Cap. (pF)	FM_S $(\frac{1}{Power})$	FM_L $(\frac{SR}{Power})$
NMC	g_{m1}= 113 g_{m2}= 38 g_{m3}=2513	18 3 121	3.3	2.07
DFCFC β=14.7	g_{m1}= 113 g_{m2}= 32 g_{m3}= 171 g_{m4}= 5	18 3 118 3	14.3	2.08
AFFC β=2.7	g_{m1}= 113 g_{m2}= 927 g_{m3}= 931 g_{ma}= 452	c_{o1} 3 121 c_{o1}	2.7	2.08
TCFC	g_{m1}= 113 g_{m2}= 113 g_{m3}= 237 g_{mt}= 57	18 3 121 c_{o2}	10.9	2.14
NMCF	g_{m1}= 113 g_{m2}= 75 g_{m3}=1257	18 3 121	5.4	3.38
PFC	g_{m1}= 77 g_{m2}= 26 g_{m3}= 593	21 3 118	10.5	3.86

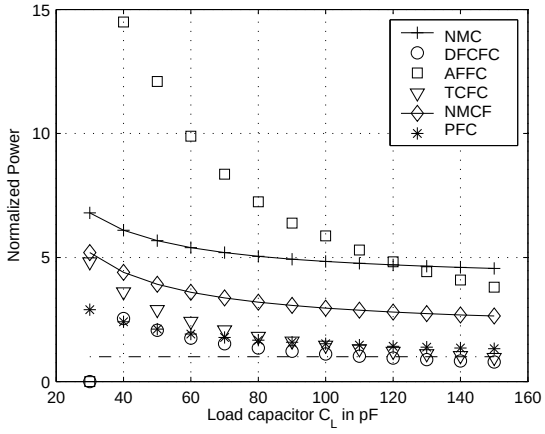


Figure 3.5: Normalized power as a function of the load capacitor, for C_{m1}=18 pF, C_{m2}=3 pF. Reference power consumption is the one transistor amplifier with $g_m = C_L UGF$.

Table 3.3: Results for C_{m1}=8 pF, C_{m2}=3 pF and C_L=100 pF

	g_m (µA/V)	Node Cap. (pF)	FM_S $(\frac{1}{Power})$	FM_L $(\frac{SR}{Power})$
NMC	g_{m1}= 50 g_{m2}= 38 g_{m3}=2513	8 3 111	3.6	2.25
DFCFC β=22.4	g_{m1}= 50 g_{m2}= 51 g_{m3}= 112 g_{m4}= 5	8 3 108 3	23.5	2.43
AFFC β=6.5	g_{m1}= 50 g_{m2}= 379 g_{m3}= 387 g_{ma}= 201	c_{o1} 3 111 c_{o1}	6.5	2.25
TCFC	g_{m1}= 50 g_{m2}= 113 g_{m3}= 154 g_{mt}= 57	8 3 111 c_{o2}	17.2	2.38
NMCF	g_{m1}= 50 g_{m2}= 75 g_{m3}=1256	8 3 111	6.2	3.91
PFC	g_{m1}= 34 g_{m2}= 23 g_{m3}= 665	11 3 108	11.8	3.68

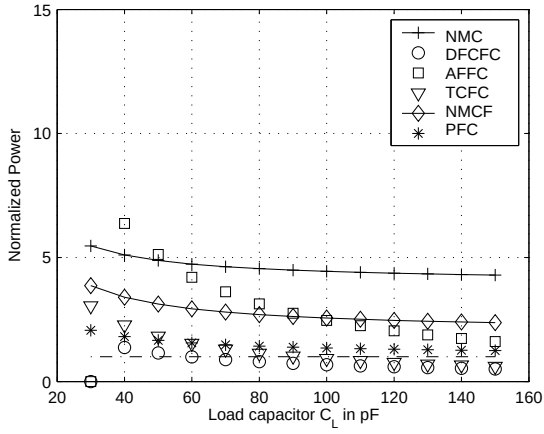
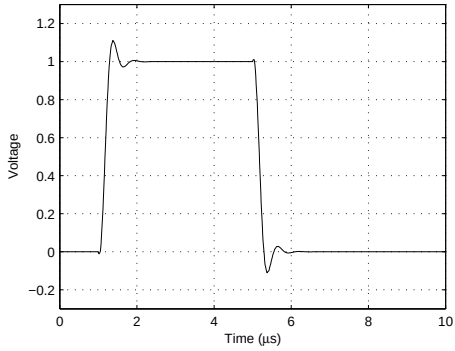
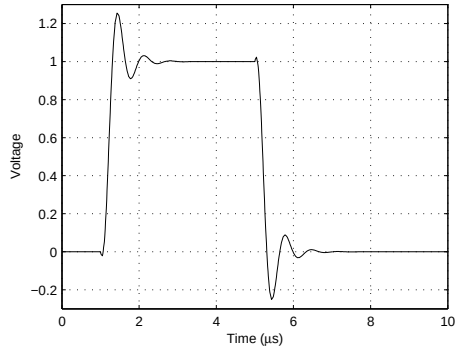


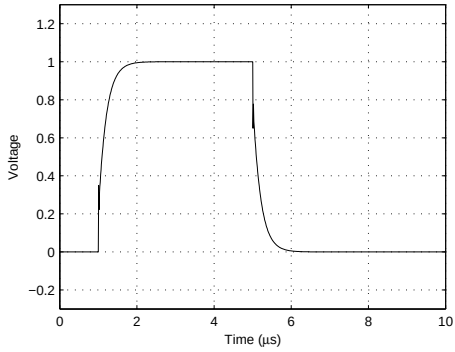
Figure 3.6: Normalized power as a function of the load capacitor, for C_{m1}=8pF, C_{m2}=3pF. Reference power consumption is the one transistor amplifier with $g_m = C_L UGF$.



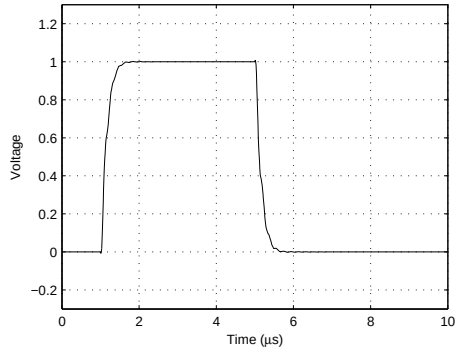
(a) NMC



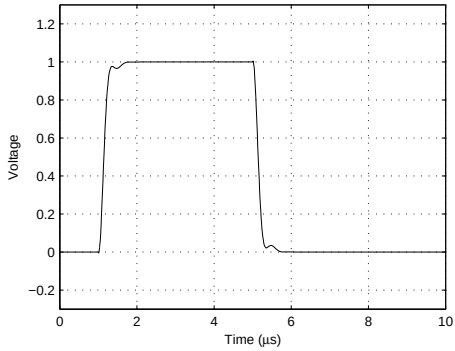
(b) DFCFC



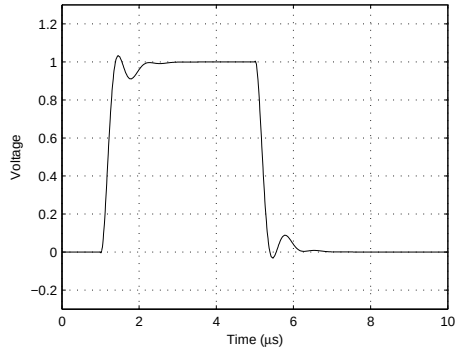
(c) AFFC



(d) TCFC



(e) NMCF



(f) PFC

Figure 3.7: Theoretical transient response of each topology to a step impulse.

3.2.4 Final Comments

From the results it can be concluded that different compensation or load capacitors have a significant impact on the result. Only under exactly the same conditions it is possible to compare the topologies.

If each topology is sized for the same conditions as suggested by Table 3.2, Table 3.3, Fig. 3.5 and Fig. 3.6, the performance of each frequency compensation method has only a minimum advantage over another. The two orders of magnitude for the figure of merit seen in Table 2.3 are a result of different compensation capacitors, different load capacitors and different supply voltages when calculating the figure of merit. Furthermore, with a simple change in the NMC topology, it is possible to significantly decrease the power consumption with the added advantage of an increased settling-time performance.

From what is seen in this chapter, the β definition in [Leu00, Lee03a] has to be updated to the definition given in sub-section 3.2.2: only a relation between the output stage transconductances.

The amplifier with the best power consumption for a given UGF is in this case the one with the poorest settling time performance. Each topology has to be carefully selected for the application in question.

The equations used have been derived by different designers and have different degrees of precision. This can also influence the results. The NMC equations have been verified to be the most accurate.

3.3 Optimization

In the previous section, each amplifier topology has been designed following the designers assumptions on what is the most efficient operating point to bias the amplifier. This assumption causes limitations as there is no justification that having third order Butterworth frequency response with unity gain feedback is the most efficient sizing criterion.

This section is organized as follows. Firstly, in sub-section 3.3.1 the NMC and the PFC amplifiers are compared in various characteristics and then optimized at the block level (Fig. 2.5 and Fig. 2.8(a)). Finally, in sub-section 3.3.2 the PFC topology is optimized at a transistor level without any previous assumption in the value of each transconductance given in Fig. 2.8(a). A minimum value for UGF is set and all transistors in Fig. 2.11 are then optimized for minimum power consumption.

3.3.1 Block Level: Comparison to the NMC

It is desirable to study each amplifier without exterior interferences, among others, different technology, design strategy or parasitics. For this purpose, a study under exactly the same conditions is described in this section.

The selection of the amplifier for a specific application is usually done by specifying its characteristics and then optimizing the current topology to accomplish the required

goals with the minimum power consumption possible. Depending on the application, more than one limiting performance may be specified. The unity gain frequency, slew rate and settling time tend to be the most frequently used. Consequently, they will be discussed next and compared in an application example.

3.3.1.1 Optimization Method

Given the fact that the number of optimization parameters is small and the constraints between them allow a simple mathematical description, a brute force approach is used.

Other optimization strategies are not really necessary as brute force guarantees the best solution for a given search space. Nowadays, fast CPU's allow the space exploration of small and simple circuits in a human acceptable period of time. With this in mind, a small-signal equivalent of the circuit from Fig. 2.6 and Fig. 2.11 is created in order to wrap around it an optimization routine in MATLAB®. In a couple of minutes, the surroundings neighborhoods of the values obtained in Section 2.3 and 2.4 is generated. It permits to try all possible values within a given interval, thus assuming (under certain limits) that all combinations are tested.

The selection of the best result is based on a figure of merit: (3.1) and (3.2). The higher it is, the more efficient the amplifier. Power estimation is based on the number of branches the transistor level circuit (Fig. 2.6 and Fig. 2.11) has, multiplied by the transconductance in that branch: $4g_{m1} + 2g_{m2} + g_{m3}$. A given transconductance can be achieved with more or less current depending on the overdrive voltage selected. As transconductance is proportional to the current ($g_m = 2I/(V_{GS} - V_T)$ and $V_{GS} - V_T$=const. $\therefore g_m \propto I$), it can be used as an accurate measurement of the relative power consumption between the amplifiers because it is desirable to study them under the same conditions: equal $V_{GS} - V_T$. Both amplifiers have the same capacitive and resistive loads.

3.3.1.2 Frequency Response

The goal is to get the optimal parameters for Fig. 2.5 and Fig. 2.8(a), for the case when both amplifiers have the same UGF and PM. The damping ratio ζ_o is set to be equal to $1/\sqrt{2}$. The result is the pole zero diagrams presented in Fig. 2.7 and Fig. 2.9, obtained after optimizing both amplifiers following the design criteria defined in Section 2.3 and 2.4.

All parameters are optimized with exception of capacitor C_{m2}, that had its value set to be only one order of magnitude higher than the parasitic capacitors of the nodes it connects to, with the purpose of maximizing the efficiency. C_{m1} is later optimized for minimum power consumption for the PFC scheme. Next, compensation capacitors for the NMC are set to be equal to the ones in PFC. Finally, the load capacitance (C_L) is swept. In this situation, as long as $C_L > C_{m1}, C_{m2}$, the PFC amplifier is more efficient than the NMC amplifier.

In conclusion, for the above stated conditions, the PFC scheme provides a more effective solution to reach the same UGF for less power consumption, and this is true for a wide range of load capacitors.

3.3.1.3 Slew Rate

In some applications, the maximum rate of change of the output of the amplifier can be the limiting factor of the performance of the circuit in which the amplifier takes part. The maximum rate at which the output node voltage can change is called slew rate (SR) and is given by

$$SR = \frac{I_c}{C} \tag{3.30}$$

The slew rate depends on the current (I_c) to charge and discharge the total capacitance (C) in the node. To increase the slew rate value, one can either increase the static current or decrease the node capacitance. Although it is possible, increasing the bias current will lead to an increase in the power dissipation, and so it is not the desirable solution. Furthermore, the selection of the appropriate output stage type, class AB instead of class A, is a practice that saves static power. This allows to increase the available current only when it is necessary: during large-signal transients. In other instance, decreasing the node capacitance can be achieved by either minimizing the parasitic capacitances or by attentive selection of the compensation capacitors. In an amplifier the load capacitor usually is not a design variable, and only the compensation capacitors can be optimized for the design goals.

The maximum rate at which the output node can change can also be expressed as a function of the derivative of the output voltage

$$SR = max\left(\frac{dy(t)}{dt}\right) \tag{3.31}$$

The PFC amplifier has a class AB output stage, hence the slew rate is not limited by the output stage, but rather by an internal node. It is also an internal node that is limiting the slew rate of the NMC amplifier, in this case, due to the high value of the output transconductance. It is in general much simpler and clearer to find which are the slew rate limiting nodes and check what is the rate of change on those for a large-signal perturbation.

After optimization, both amplifiers have the same UGF and PM. The numerical solution of (3.31) shows that the NMC slew rate is about 70 % higher than the value from PFC. In spite of this result, the higher slew rate is obtained at the expense of a higher power dissipation. The power dissipation in the last stage of the NMC is more than double the total power dissipation of the PFC amplifier. The figure of merit (3.2) gives a value 80 % higher for the PFC. Consequently, it is possible to obtain a more efficient solution using the PFC scheme.

3.3.1.4 Settling Time

The settling time is an important feature in amplifiers. In switched-capacitor applications this characteristic can in some cases be the limiting design criteria. In this case, it is desirable to optimize the amplifier for the best settling time possible required by

the application in question, instead of optimizing it for the highest possible unity gain frequency.

Although the method described in Section 2.4 provides a straightforward method to design for a desired specification, no guarantee exists that the most efficient solution will be indeed reached. In a two pole system, the phase margin is translated into a unique time domain response. In this case the settling time depends on the phase margin [Yan90], and thus, an open loop design strategy can be used to design for the desired closed loop characteristics, whereas this is not the case for a third order system with three poles. No unique relationship exists between settling time and phase margin (Fig. 3.9(b)). Consequently, a closed loop design approach is chosen to obtain the most efficient design from all the possible solutions.

Third order systems can be theoretically studied using the method presented in [Mar98]. The relative position of poles in the closed loop, after positioning for good slewing, is used to derive the position of the poles in the open loop transfer function.

In the large-signal regime, the inverse Laplace transform ($\mathscr{L}^{-1}$) of the closed loop transfer function (2.1) in response to the step function $U(s)$ does not accurately represent the output voltage

$$y(t) = \mathscr{L}^{-1}(G(s).U(s)) \tag{3.32}$$

However, if it is assumed that the input voltage step is small enough so that the small-signal approximation of the amplifier closed loop transfer function still holds reasonably well, the amplifier closed loop transfer function with a DC feedback f_{dc} can be modeled as in (2.1). This is the case of switched capacitor circuits between consecutive samples. For completeness, (2.1) is once more written as

$$G(s) = \frac{A\left(1+\frac{s}{z_1}\right)\left(1+\frac{s}{z_2}\right)}{\left(1+\frac{s}{\omega_d}\right)\left(1+2\zeta\left(\frac{1}{\omega_n}\right)s+\left(\frac{1}{\omega_n^2}\right)s^2\right)} \tag{3.33}$$

In the above equation, the DC gain is given by A and is approximately equal to 1. No simple representation exists to describe the closed loop dominant pole (ω_d), the undamped natural frequency (ω_n) or the damping ratio parameter (ζ) from the open loop parameters (ω_{do}, ω_{no} and ζ_o), although representation in the opposite direction is simple. Also, it is not always possible to place the closed loop poles and zeros in the desired location. Sometimes their relative location is not independent of each other. Another difficulty is that a set of unwieldy equations make this method inappropriate for a theoretical study. To avoid this situation, a different approach is followed. The open loop parameters from Fig. 2.5 and Fig. 2.8(a) are used to obtain the values, in (2.7) and (2.22), respectively. Afterwards, the time response is obtained by numerically solving the inverse Laplace transform (3.32). Finally, all relevant criteria are obtained, and plotted in Fig. 3.8(a) and Fig. 3.8(b). To better compare both topologies, in all

Table 3.4: Sizing results for sub-section 3.3.1 with C_{m1}=18 pF, C_{m2}=3 pF and C_L=100 pF and comparison with equations from Chapter 2.

Parameter	Unit	NMC		PFC	
		Section 2.3	here	Section 2.4	here
g_{m1}	µA/V	113	113	77	77
g_{m2}	µA/V	38	55	26	43
g_{m3}	µA/V	2513	2149	593	564
UGF	MHz	1.00	1.05	0.99	1.25
PM	°	60	67	60	61
SR	V/µs	0.63	0.63	0.37	0.37
ST	µs	0.73	0.59	1.57	1.41
$Power$	µA/V	3041	2711	953	958
$p_{2,3}$	MHz	-2.00 $\pm$ j2.00	-1.67$\pm$ j2.66	-0.78 $\pm$ j0.78	-0.68 $\pm$ j1.27
ζ_o		$1/\sqrt{2}$	0.531	$1/\sqrt{2}$	0.469

Table 3.5: Performance Improvement

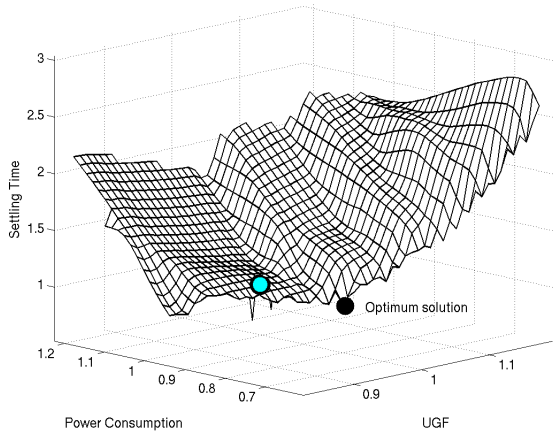
Parameter	NMC	PFC
Unity Gain Frequency	1.05	1.26
Settling Time	0.81	0.89
Power Consumption	0.88	1.00

cases the DC gain (A_o) and open loop dominant pole (ω_{do}) obtained previously is kept unaltered. Final settling time error is 1 %.

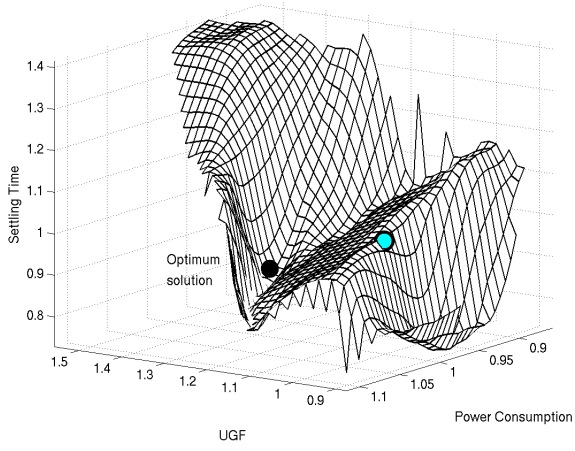
When comparing the values given by equations in Section 2.3 for the case of the NMC, the current sizing approach yields an amplifier 19 % faster while consuming 12 % less (Table 3.4). A marginal increase of 5 % in the unity gain frequency can also be seen. For the PFC (see Table 3.4), the amplifier will be 11 % faster and with an increase of 26 % in the UGF for the same power consumption than those given by equations in Section 2.4. The final improvements are presented in Table 3.5.

The final values obtained for the NMC makes it more than two times faster than the PFC. Yet, once the power is taken into account, the PFC is a better choice for a more efficient compensation topology, despite the fact that z_1 in Fig. 2.9, close to the non-dominant poles, influences the settling time. Analysis has shown that z_1 is always smaller than $\zeta_o\omega_{no}$. Another conclusion is that despite the NMC having a smaller settling time, it is not possible for the same power consumption to be as fast as the PFC.

Optimization has led to a faster amplifier, with the advantage of an increased bandwidth without an increase in power consumption. Some of the PFC characteristics are given in Fig. 3.9.

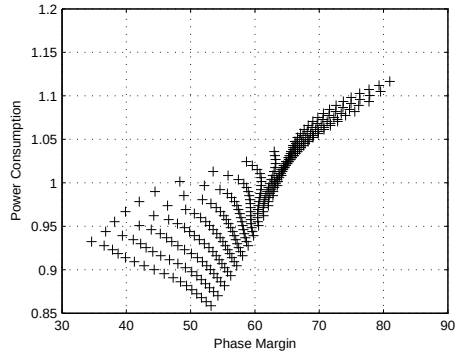


(a) Example for the NMC amplifier.

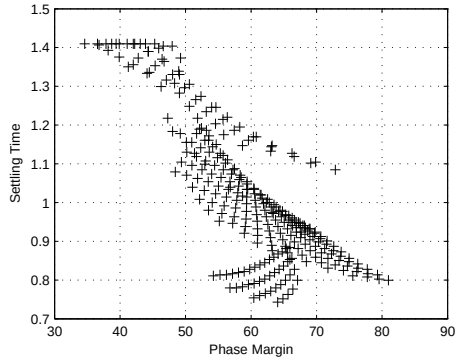


(b) Solutions for the PFC topology.

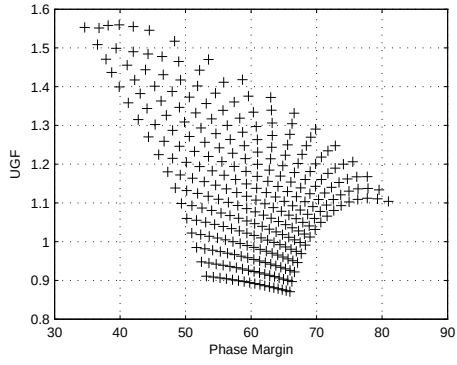
Figure 3.8: Normalized settling time as a function of power dissipation and UGF. The light point is the result as given by equations from Chapter 2 and the dark point is the new optimized point.



(a) Power consumption as a function of the phase margin.



(b) Settling time (to 1 %) versus phase margin.



(c) Variation of the UGF for different values of the phase margin.

Figure 3.9: Different performances of the PFC amplifier.

Table 3.6: Simulated performances for the original amplifier design (Section 2.4) and three automatically sized experiments using DANCE (including some optimization statistics).

Parameter	Unit	Ref. (Section 2.4)	DANCE (cont.)	DANCE (grid1)	DANCE (grid2)
DC Gain	dB	122	146	122	132
UGF	MHz	3.15	3.16	3.24	3.15
Phase Margin	deg	51.8	54.6	52.2	51.9
Slew Rate (calc.)	V/µs	0.777	0.948	1.014	0.957
Power Cons.	µW	572	416	440	504
time	hh:mm		13:44	14:15	16:57
iterations	–		500	500	500
evaluations	–		30061	30186	30657
simulations	–		19611	20677	25095

3.3.2 Transistor Level

The work done in the previous sub-section requires the modeling of the amplifier which is a long and tedious procedure. Furthermore, at the end each transistor and capacitor has to be chosen so each transconductance in the block diagram has an equivalent at the device level. This work can be simplified if an automated design procedure is followed. This is the purpose of this section which follows the work of [Fra03].

In order to automatically size the amplifier circuit, the netlist of the PFC amplifier is parameterized using 21 design variables (one bias voltage and current, two compensation capacitors, seven transistor lengths and ten transistor widths). The number of transistor geometry variables is somewhat reduced by taking standard analog design constraints (e.g. the matching of differential input pairs and current mirrors) into account. However, constraints on the operating point of the circuit are not included, only the performance specifications are given as input to the tool. On the one hand, this makes the design space much more complex, but on the other hand this doesn't require specific circuit knowledge. Three experiments are carried out (on an Intel Pentium 4 2.6 GHz PC) and the resulting performances together with some optimization statistics have been collected in Table 3.6.

For all three experiments using DANCE [Fra03], the original performances are taken as constraints, except for the power consumption, which is requested to be minimized. A random starting point is used in combination with a population size of 60 for the evolutionary optimization algorithm. The number of iterations is arbitrary set to 500, which translated into long (overnight) optimization times for all three experiments. However, this allowed to verify that the optimal values are no longer changing significantly. In the first experiment all device sizes are allowed to take on continuous values. A solution is found that achieves better performances (gain +20 %, phase margin +5 % and slew rate +22 %) and additionally decreases power consumption considerably (-27 %). Because arbitrary device geometries cannot be obtained in practice, a second experi-

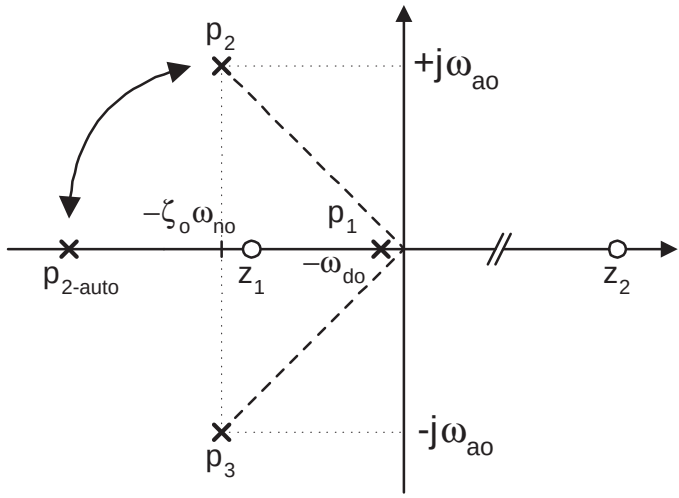


Figure 3.10: Pole-zero diagram for the sized amplifiers Table 3.7.

ment ('grid1') is set up in which all transistor sizes are only allowed to take on integer multiples of the manufacturing grid (0.05 µm for the technology used). Here also better performances are attained (unity gain frequency +3 %, phase margin +1 % and slew rate +31 %) combined with a reduction of the power consumption (-23 %). Finally, in the last experiment ('grid2'), all transistor lengths are constrained to be 0.7 µm (except M31 which is 0.5 µm) as in the original design. This results in a sized circuit which, compared to the original design, has better gain (+8 %) and SR (+23 %), while simultaneously consuming less power (-12 %). The device sizes, component values and biasing for the original circuit and the three experiments using DANCE are compiled in Table 3.7. Also, the g_m values of the three transistors marked in Fig. 2.8(a) and 2.11 have been listed in Table 3.7. The seven parameters for transistor lengths represent groups of identical length transistors: L_{G1} (M_{00}, M_{10}, M_{13}, M_{16}, M_{21}), L_{G2} (M_{11}, M_{12}), L_{G3} (M_{14}, M_{17}), L_{G4} (M_{15}, M_{18}), L_{G5} (M_{23}, M_{24}), L_{G6} (M_{20}, M_{30}) and L_{G7} (M_{31}). It can be seen that the g_m values can be vary considerably, but it is the ratio of g_{m1} over C_{m1} that determines the unity gain frequency of the amplifier. The difference of the g_m values can be further illustrated by the dominant poles and zeroes that are listed in Table 3.8 for each case. The pole-zero placement of the original circuit (which exhibits a conjugate pole pair) differs from that of the three automatically sized circuits (see Fig. 3.10). However, and this is most important, a stable amplifier results in all cases. Bode plots (magnitude and phase) for the original circuit and the last experiment in Table 3.7 are presented in Fig. 3.11. Finally, in Fig. 3.12, the minimum cost evolution for all three experiments using DANCE is shown. It can be seen that feasible solutions are quickly found and that the optimal solution is found after about 5000 evaluations (one sixth of the total number). This corresponds to approximately 2.5 hours.

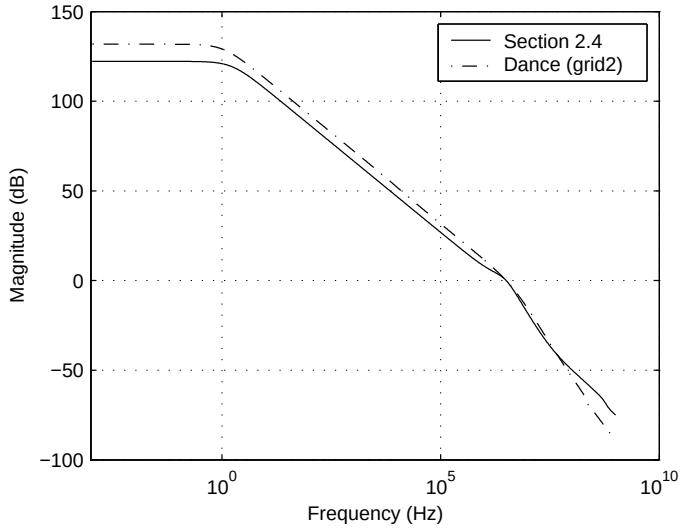
Table 3.7: Device sizes, component values and biasing for the original design (Section 2.4) and several automatically sized experiments using DANCE.

Parameter	Unit	Ref. (Section 2.4)	DANCE (cont.)	DANCE (grid1)	DANCE (grid2)
V_{BIAS}	V	1.25	1.87	2.56	1.93
I_{BIAS}	µA	5	3.09	3.44	3.63
C_{m1}	pF	15	4.3	4.7	5.7
C_{m2}	pF	3	13.5	7.9	14.2
L_{G1}	µm	0.7	2.87	6.50	0.7
L_{G2}	µm	0.7	0.83	0.50	0.7
L_{G3}	µm	0.7	0.81	0.45	0.7
L_{G4}	µm	0.7	0.36	0.45	0.7
L_{G5}	µm	0.7	2.43	0.45	0.7
L_{G6}	µm	0.7	0.38	0.35	0.7
L_{G7}	µm	0.5	0.66	0.65	0.5
W^*_{00-10}	µm	10	3.53	10.90	28.70
W_{11-12}	µm	40	24.03	12.85	23.70
W^*_{13-16}	µm	10	0.49	2.30	4.90
W_{14-17}	µm	6	0.39	0.65	1.20
W_{15-18}	µm	11	5.22	11.05	0.45
W_{20}	µm	15	2.80	1.05	0.75
W_{22}	µm	10	0.63	1.00	1.40
W_{21-23}	µm	2	2.07	26.10	5.40
W^*_{30}	µm	15	37.16	42.45	27.60
W^*_{31}	µm	15	40.86	4.85	4.95
g_{m1}	µS	211	120	127	143
g_{m2}	µS	76	9	5	3
g_{m3}	µS	1703	1947	1084	1350
g_{mf}	µS	1636	2268	2409	2215

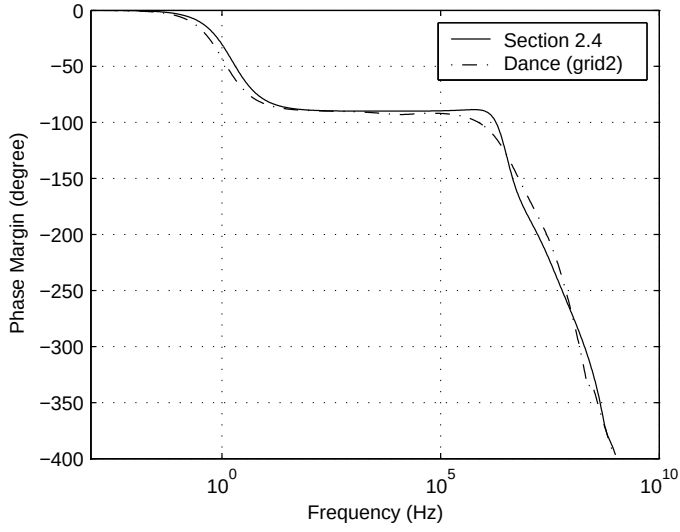
*: multiplication factors of transistors M_{10}, M_{13}, M_{16}, M_{30} and M_{31} are 6, 3, 3, 22 and 5 respectively.

Table 3.8: Dominant poles and zeroes for the sized amplifiers in Table 3.7. Unlike the reference which has explicitly set $g_{m3} = g_{mf}$, the previous transconductances can have different values in the proposed designs according to DANCE. The result is that the complex conjugate pole pair disappears and z_1 compensates for the other real nondominant pole p_3.

Parameter	Ref. (Section 2.4)	DANCE (cont.)	DANCE (grid1)	DANCE (grid2)
z_1 [Hz]	(-1.910E+6; 0)	–	–	–
p_1 [Hz]	(-1.700; 0)	(-0.184; 0)	(-3.646; 0)	(-1.095; 0)
p_2 [Hz]	(-2.063E+6; +2.315E+6)	(-5.242E+6; 0)	(-4.665E+6; 0)	(-4.584E+6; 0)
p_3 [Hz]	(-2.063E+6; -2.315E+6)	–	–	–
ζ	0.67	–	–	–
ω_n [Hz]	3.101E+6	–	–	–



(a) Magnitude



(b) Phase Margin

Figure 3.11: Magnitude and phase margin plots for the original circuit (Section 2.4) and for the last experiment in Table 3.7.

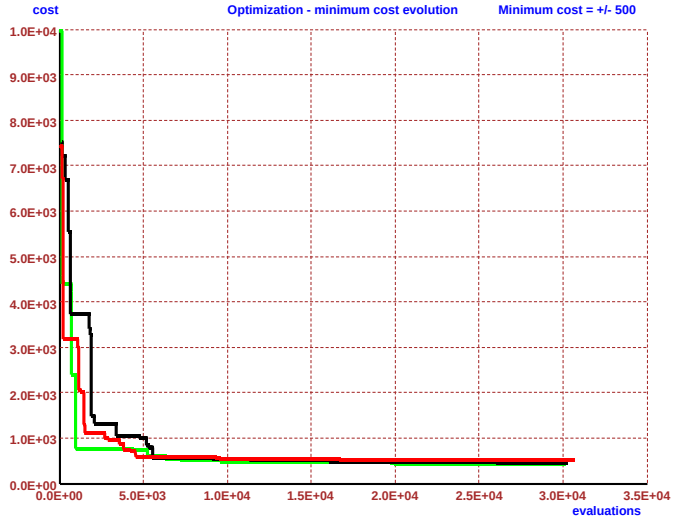


Figure 3.12: Minimum cost evolution for the experiments in Table 3.7.

3.4 Conclusion

A topology study of three stage amplifier frequency compensation has been presented. Two topics have been covered in this chapter. The first part has covered the comparison between the PFC amplifier and five other frequency compensation methods, while the second has dealt with the optimization of the PFC and the NMC to minimize power consumption.

A method for comparing the PFC with other three stage frequency compensation amplifiers has been given. Under the exact same conditions the performances of each one of the six amplifiers has been shown to be more similar than the measurement results suggest. The NMC topology can have its performance improved by adding a feed-forward stage. In this way, both the UGF and settling time efficiency is increased for the NMC with feed-forward.

Two different optimization strategies have been presented. In the block level comparison, the PFC topology is always more efficient in terms of UGF, slew rate and settling time than the commonly accepted reference: the NMC. Results of a transistor level optimization has been presented for the PFC amplifier. In this case a new operating point, where only real poles exist, has been chosen to be the best. This completely automated design strategy does not make assumptions on what might be the best design strategies, and in this way, different and possibly better designs are conceivable. Both optimization strategies suggest that the basic principle in stabilizing three stage amplifiers, which is assuming a third order Butterworth frequency response with unity gain frequency is not the most power efficient. A slight decrease in the open loop damping ratio parameter ζ_o can save up to 25 % in power.

Chapter 4

RF Power Amplifier Classification, Theory and Optimization

4.1 Introduction

Mobile electronics use a myriad of integrated circuits including a digital DSP and different analog integrated circuits. For the end user, autonomy of the battery is an important factor. For the equipment manufacturer, cost reduction allows to sell the product at a lower cost. It is of interest in this case to reduced the number of integrated circuits necessary to build a mobile phone or a PDA (Personal Digital Assistant). In this era of integration, CMOS is the technology that makes high integration of functions into a single chip possible. This prompts the desire to transfer circuits from other technologies to this mainstream and low-cost technology. Nowadays, it is also desirable to include some kind of communication link between the portable electronic and a base or with the Internet. Specifically, wireless communication demands higher performances without impairing autonomy. However, a considerable percentage of the power consumption goes towards creating this wireless path (Fig. 4.1). As such, it is of utmost importance to maximize the efficiency of this block: the Power Amplifier (PA).

Considerable effort has been put into research to integrate in CMOS full transceivers for mobile communications [Ste00]. The final objective is to include a system-on-chip with digital processing, analog functions and support blocks. On the analog side, the inclusion of the power amplifier is still a rare situation. The need for high efficiency and the required current flowing in it results in dedicated chips often being used. Some of the solutions can include: Gallium Arsenide (GaAs) [Sow95] and bipolar [Sim99]. Although they provide the required output power with high efficiencies, it is nevertheless desirable to replace them by solutions using CMOS [Su97, Yoo01, Sow03]. The advantages come mainly from the current trend to integrate all circuits in the same die for cost but also area reduction.

With more advanced CMOS technologies having a lower power supply it becomes extremely difficult to achieve the high output power required by communication standards. This is the reason why there is a lack of offer for the 3 W output power level range with integrated power amplifiers in CMOS. This is especially true for the cases where the previous block has an output power in the range of -10 to 0 dBm [Ste00]. As a result, a gain of more than 30 dB is required. In addition to this, higher currents are now required to achieve the same output power. The result is that wider interconnections are necessary to carry the current and avoid electromigration reliability issues.

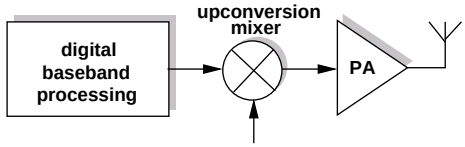


Figure 4.1: The transmitter part of a transceiver.

Moreover, with lower voltages, because of the higher current flowing, every small resistance has an increased importance.

The study of class E power amplifiers that follows, is divided in three parts. This chapter first presents the different amplifiers topology suitable for RF implementation which is followed by theory and a methodology for optimally designing power amplifiers for maximum efficiency. Chapter 5 later explores the possibility to increase the power amplifier efficiency by increasing the supply voltage and techniques to do it in commercial CMOS technologies. Finally Chapter 6 focuses on the underlying circuit details, design aspects and chip measurements.

A brief outline of this chapter is as follows. In Section 4.2 common metrics used to benchmark the power amplifier are given. Later, the two different categories for power amplifiers are described in Section 4.3, linear and non-linear amplifiers, respectively. This is followed by presenting in Section 4.4 the proposed design methodology to deal with circuit parasitics while optimizing the circuit for maximum efficiency. Some optimization experiments using the tool developed in Section 4.4 are given in Section 4.5. They include different optimization studies on the generalized class E PA and the optimization of a two stage amplifier including all relevant layout parasitics. The conclusions are finally presented in Section 4.6.

4.2 Definition of Performance Metrics

While in operation the PA converts a DC power from the supply voltage into a RF by sufficient amplification of the input RF signal. Only in the ideal case this conversion is lossless which means that the PA itself will drain more power than it delivers. The result with which the amplification takes place is called efficiency. Two other common metrics are the Drain Efficiency (DE) and Power Added Efficiency (PAE). Efficiency (η) is given by

$$\eta = \frac{P_{OUT,RF}}{P_{IN,RF} + P_{SUPPLY,DC}} \tag{4.1}$$

Drain Efficiency (DE), is a measure of how well the DC power is converted to RF signal. The other common metric used is the PAE which indicates how much power is necessary to drive the power stage. The two types of efficiency are defined as follows

$$DE = \frac{P_{OUT,RF}}{P_{SUPPLY,DC}} \tag{4.2}$$

$$PAE = \frac{P_{OUT,RF} - P_{IN,RF}}{P_{SUPPLY,DC}} \tag{4.3}$$

4.3 RF Power Amplifier Classification

Todays PAs can be said to operate in two different categories: linear (sub-section 4.3.1) or non-linear (sub-section 4.3.2). In the first the transistor acts as a current source while in the latter it acts as a switch. Each one of the categories is divided in classes of operation where one is defined by the way the output network shapes the output current and voltage (Fig. 4.2).

The choice the designer has to make is between linearity and efficiency. Linearity in an amplifier comes at the expense of efficiency because current is always flowing and is as such dissipating power in the form of heat. On the other hand, high efficiency is obtained in switched power amplifiers because losses in the active device are minimized by operating it as a switch. Ideally, the device has either zero voltage while conducting current or zero current when the voltage is non-zero.

4.3.1 Linear Power Amplifiers

4.3.1.1 Class A

A power amplifier working in class A is biased in the middle of the active range portion of the I-V curves. The output conducts during the entire cycle of the input signal. High linearity is only obtained when the device is working with small signals. In this case the maximum theoretical efficiency is 50 %. Higher efficiencies can still be obtained at the expense of linearity.

4.3.1.2 Class B

In class B usually two devices are operated in a push-pull configuration where only one of them is allowed to be conducting current at a time. Each one of the devices operates exactly half of the input signal cycle. Class B amplifiers have a theoretical maximum value for the efficiency equal to $\pi/4$=78.54 %. The extra time it takes to turn on and off results in distortion as a result of the crossover.

A good compromise between the high-linearity of class A and the high-efficiency of class B is to operate the devices in a mix of these two. In class AB as it is called, the output load conducts more than half of the cycle to avoid crossover distortion but not the entire cycle to maximize efficiency.

4.3.1.3 Class C

If the device is biased so that it conducts less than half of the input cycle, the amplifier is operating in class C. This class can be more efficient than class B at an expense of an even higher distortion. Theoretically efficiency can reach 100 % as the conducting cycle and output power go to zero.

For narrow band RF transmitters, distortion can be greatly reduced by the use of a tuned load which only resonates in the fundamental frequency. In this way, unwanted harmonics are suppressed.

4.3.2 Switching Power Amplifiers

4.3.2.1 Class D

Class D amplifiers use a technique called pulse width modulation where the devices are rapidly switched on and off several times for each cycle. The output has a strong harmonic content that has to be removed by means of a passive filter. Since there is current only when the voltage is zero, theoretically they do not dissipate power and can have 100 % efficiency.

Class D amplifiers have not succeeded for RF applications whereas they are successfully used for high power audio amplifiers and other low frequency applications where simplicity and efficiency are important.

4.3.2.2 Class E

First published in 1975 [Sok75], class E amplifiers have found extensive use in the RF and microwave field. It consists of a switching transistor driving a current through an inductor and then a resonant filter.

The switched mode class E amplifier can ideally achieve 100 % efficiency. The idea behind its operation is to have non overlapping output voltage and output current at the transistor drain. This is achieved by means of a tuned output filter that ensures that current and voltage are not present at the same time. For this high efficiency to occur three conditions have to be met [Sok75]. When the transistor turns on, the voltage across the transistor drain has to be zero and the slope of the drain voltage should be zero as well. The next criterion states that the device should be completely off before the rise of the voltage across the transistor. The main disadvantage of class E power amplifiers is the high-voltage at the drain terminal which under optimum condition is as high as 3.6 time the supply voltage.

4.3.2.3 Class F

The basic idea behind the class F tuned power amplifier is the addition of a third harmonic to the transistor output voltage waveform so it resembles more a square wave. In the limit, all even harmonics are added. This is switched-mode class-F operation, and can have 100 % efficiency.

4.3.3 Comparison

For the frequency range of interest (800-900 MHz) and in the field of CMOS technology, published measured results do not show a clear advantage of class E over class F. If it is the maximum peak drain voltage that puts a limitation, then class F is preferred with only $2V_{DD}$ compared to $3.6V_{DD}$ of the class E.

Again, with CMOS, the maximum published efficiency is in the range of 50 to 60 % for an output power in the order of 1 W. This efficiency has been reached not only by amplifiers operating in class E and F, but also operating as class AB [Fal01] and working even at higher frequency (1750 MHz).

Class E amplifiers however have some interesting attributes. The output shaping network is intrinsically simpler than class F. Insensitive in spite of parameter variations, it has show good robustness [Raa78a]. And in today's low-voltage circuits it is better suitable not only for the higher output impedance but also for the maximum achievable efficiency [Sow95].

To add to the above reasons for choosing class E topology, is the concept study introduced in Chapter 5. Research will be carried out on the possibility to have a more power efficient amplifier by using a higher voltage and techniques to reach this goal in commercial CMOS technologies.

4.4 Class E Design Methodology

A close inspection of the circuit of Fig. 4.3 reveals seven degrees of freedom. Although analog designers have clearly derived accurate equations for what appears to be more complex circuits, addressing the referred circuit has been a more difficult task. Different research has shown that apparently it cannot be explicitly solved. Equations do exist but do not address all tradeoffs between component values and amplifier efficiency.

Research on the topic has existed since the first paper from Sokal [Sok75]. In the work from Raab [Raa78b], simple parasitics in the transistor are taken into account and their effect on the final efficiency is presented.

Other research has focused on removing the limitation on the necessity to model inductor L_1 as a shunt-feed choke. Zulinski [Zul87] has shown that class E amplifier operating conditions could be met with a finite dc-feed inductance. Avratoglou [Avr89] later used this configuration and the new analysis further included the effect of the quality factor of the series-tuned (L_0-C_0) and the effect of the device on-resistance. Differential equations where solved using the Laplace transform and next represented in state variables. The complete formulation is rather complex but can nevertheless be considered a generalized method to analyze the class E amplifier.

In Smith's work [Smi90], a computer routine is used to partially alleviate all the complexity in the Avratoglou's art and determine the transforms and inverse Laplace transforms. Nevertheless, a much simpler formulation covering the same aspects as before, is given by Mandojana in [Man90].

Following this, Chudobiak [Chu94] has shown that the non-linear drain-bulk capacitance could be modeled and all derived circuit parameters and waveforms have been derived considering the previous effect. Although the output and current waveforms are not affected by the non-linear capacitance, the transistor peak drain voltage is increased in comparison with the linear case.

Equally important is that even for circuits with the complexity of the one given by Fig. 4.3, existing equations [Sok75, Raa78a, Raa78b] do not guarantee that the maxi-

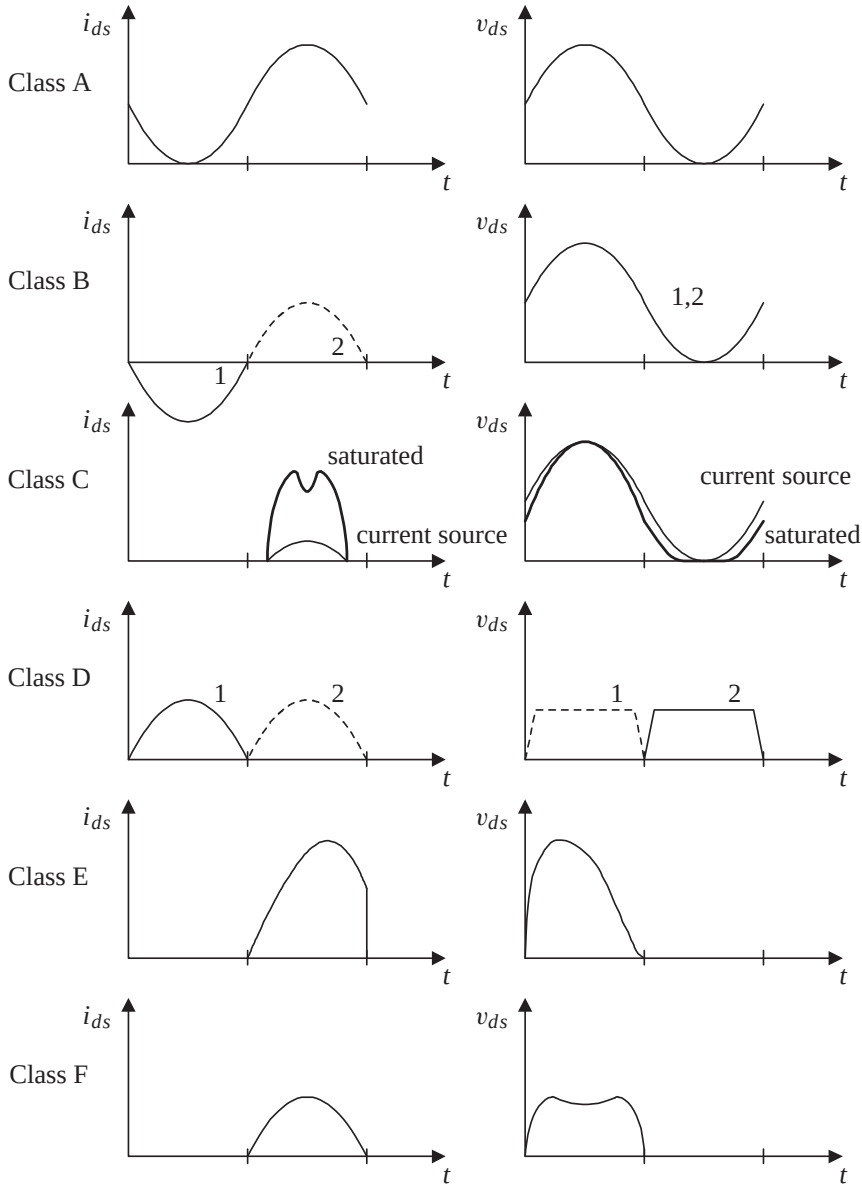


Figure 4.2: Ideal waveforms for the drain current and drain voltage for each operation class tuned for maximum power efficiency. Class B and D are usually implemented in push-pull configuration and have each waveform indicated by 1 and 2.

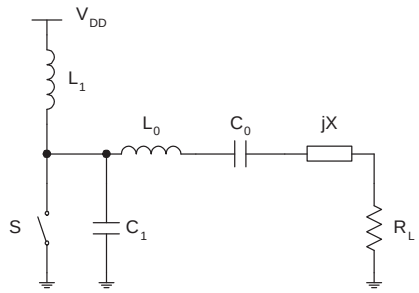


Figure 4.3: Ideal class E power amplifier. It consists of switch S, the finite dc-feed inductance L_1, the shunt capacitor C_1 the series-tuned (L_0-C_0) in series with a reactance jX and the load resistance R_L. Elements S and C_1 are a representation of the on-resistance and drain-bulk capacitance of a CMOS transistor.

mum efficiency is indeed reached, or that some of the initial assumptions made prior to equation derivation might not lead to the most power efficient design [Por93]. Furthermore, with inclusion of simple parasitics it becomes increasingly difficult to generate equations that are accurate enough.

Two possibilities will be given that partially address the aims of the analog designer:

- State-space model description
- Automated sizing using a simulation-based optimization

The state-space model allows to address the goal to have a set of partially derived equations. Nevertheless, at the end an optimization is necessary to find values that create a steady-state operation. On the other hand, initial conditions are not necessary when using a simulation-based optimization. Furthermore, any circuit can be optimized being limited by the capacity and accuracy of the simulator. The penalty clearly resides in the longer CPU time. The continuous increase in computing power alleviates this drawback partially. With a simulation-based optimization, a simulation tool is used to evaluate the parameter vectors proposed by the optimizer. As such, the inclusion of the parasitics is simplified. Also, as no preconceptions and restrictions are made prior to circuit sizing, the possibility exists that new and more interesting solutions can be found.

4.4.1 State-Space Model of the Power Amplifier

The operation of the power amplifier of Fig. 4.5 can be seen as linear time-invariant circuit [1] containing a periodically operated switch R_k with two states in the switching period T. In one of the states (τ_1) a small value represents the low on-resistance value.

[1] The notation of the work of Liou [Lio72] is followed.

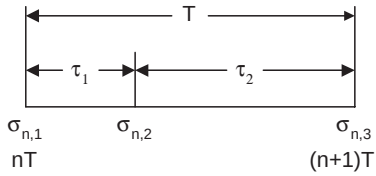


Figure 4.4: Notation for the *n*th switching period.

In the other one (τ_2), a high value represents the transistor in the off state (Fig. 4.4). Defining the switching instants $\sigma_{n,1} = nT$, $\sigma_{n,2} = nT + \tau_1$, and $\sigma_{n,3} = nT + \tau_1 + \tau_2 = (n+1)T$, a set of state equations of the form

$$\dot{x}_{n,k}(t) = A_k x_{n,k}(t) + B_k u(t) \tag{4.4}$$

$$y_{n,k}(t) = C_k x_{n,k}(t) + D_k u(t) \tag{4.5}$$

$$\sigma_{n,k} < t < \sigma_{n,k+1} \quad k = 1, 2$$

can be used to represent the linear circuit. The previous quantities $x_{n,k}(t)$, $y_{n,k}(t)$ and $u(t)$ are, respectively, the state, output and input vectors, and A_k, B_k, C_k, D_k are constant real matrices.

The standard continuous system in the form of

$$\frac{dx_{n,k}(t)}{dt} = A_k x_{n,k}(t) + B_k u(t) \tag{4.6}$$

in each portion of the *n*th switching period τ_k, k=1,2, has a solution in the form of

$$x_{n,k}(t) = e^{A_k(t-\sigma_{n,k})} x_{n,k}(\sigma_{n,k}) + e^{A_k t} \int_{\sigma_{n,k}}^{t} e^{-A_k \varphi} B_k u(\varphi) d\varphi \tag{4.7}$$

since

$$\int_{\sigma_{n,k}}^{t} e^{A_k(t-\varphi)} d\varphi = (A_k)^{-1}(e^{A_k(t-\sigma_{n,k})} - I) \tag{4.8}$$

(I is the identity matrix). The general solution of (4.6) to an input signal $u(t) = u$ at switching instant $\sigma_{n,k}$ is given by

$$x_{n,k}(t) = e^{A_k(t-\sigma_{n,k})} x_{n,k}(\sigma_{n,k}) + A_k^{-1}(e^{A_k(t-\sigma_{n,k})} - I) B_k u \tag{4.9}$$

Knowing that the values of $x_{n,k}$ at the end of the on-state are the initial values of the off-state, the complete system can be simulated. However, it is of interest to find the steady-state for a complete evaluation of the amplifier performance. The transient

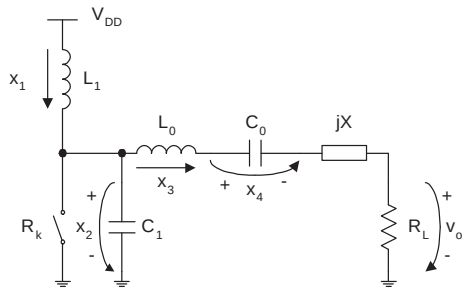


Figure 4.5: Ideal class E power amplifier as in Fig. 4.3 with the indication of the state-space variables. The switch S is now more conveniently represented by resistance R_k.

mode can be completely eliminated by setting the initial state vector $x_{0,1}(0)$ equal to Ju. Since the state doesn't change discontinuously at the switching instants, the complex matrix J can be simply written as

$$J = (I - M)^{-1}H \tag{4.10}$$

In the former expression, M is a real matrix and equal to:

$$M = e^{A_2\tau_2}e^{A_1\tau_1} \tag{4.11}$$

H is also a complex matrix expressed as:

$$H = e^{A_2\tau_2}A_1^{-1}(e^{A_1\tau_1} - I)B_1 + A_2^{-1}(e^{A_2\tau_2} - I)B_2 \tag{4.12}$$

4.4.1.1 State-Space Description

Consider the ideal class E power amplifier like the one presented in Fig. 4.5. The set of differential equations that describes the behavior of the amplifier is given by

$$L_1\frac{dx_1}{dt} = -x_2 + V_{DD} \tag{4.13}$$

$$C_1\frac{dx_2}{dt} = x_1 - \frac{x_2}{R_k} - x_3 \tag{4.14}$$

$$(L_0 + X)\frac{dx_3}{dt} = x_2 - R_L.x_3 - x_4 \tag{4.15}$$

$$C_0\frac{dx_4}{dt} = x_3 \tag{4.16}$$

with

$$\begin{cases} x_1 = I_{L_1}(t) \\ x_2 = V_{C_1}(t) \\ x_3 = I_{L_0}(t) \\ x_4 = V_{C_0}(t) \end{cases} \tag{4.17}$$

in this case, the state variables are the inductor L_1 current, the capacitor C_1 voltage, the inductor L_0 current and the voltage across C_0. During the off state the switch exhibits a high resistance R_{OFF}, while during the on state the value of R_k is reduced to R_{ON}.

Analyzing the circuit from Fig. 4.5, the A_k, B_k, $x_{n,k}$ and $u(t)$ matrices are respectively:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & \frac{-1}{L_1} & 0 & 0 \\ \frac{1}{C_1} & \frac{-1}{C_1 R_k} & \frac{-1}{C_1} & 0 \\ 0 & \frac{1}{L_0+X} & \frac{-R_L}{L_0+X} & \frac{-1}{L_0+X} \\ 0 & 0 & \frac{1}{C_0} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} \frac{1}{L_1} \\ 0 \\ 0 \\ 0 \end{bmatrix} V_{DD} \tag{4.18}$$

The remaining matrices, $y_{n,k}$, C_k and D_k are respectively

$$V_o = [0 \quad 0 \quad 0 \quad R_L] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + [0] \quad V_{DD} \tag{4.19}$$

The above algorithm is implemented in MATLAB® and used to optimize the circuit from Fig. 4.3. The initial values for the components can be obtained using the formulas presented in [Sok75]. Now, it is possible to optimize the circuit for maximum efficiency. It is also possible to see the influence of each component separately by varying each one of them and observing how the losses change.

Each simulation takes about 35 ms on a Intel P4 processor at 2.4 GHz. The small simulation time makes it very convenient for inserting in an optimization routine that usually has to execute a couple of thousand iterations. After a few minutes the optimization routine ends. The resulting circuit sizes of such an optimization using Simulated Annealing is given in Table 4.1. A drain efficiency of 91.5 % is obtained for a supply voltage of 2 V. The state variable waveforms are given in Fig. 4.6(a). After some manipulation, the switch waveforms are easily obtained. The resulting drain-source voltage and drain current of the class E amplifier is shown in Fig. 4.6(b).

Although simple and easy in implementation, the lack of parasitics other than the switch-on resistance makes this method impractical. A new run of the optimization routine generates a completely different set of solutions, although giving the same Drain Efficiency (DE). Moreover, using the above equations, the influence of the series

Table 4.1: Optimized circuit sizes for the circuit in Fig. 4.5.

Frequency	R_{ON}	C_1	L_1	L_0+X	C_0	R_L
850 MHz	0.22 Ω	15.96 pF	10.46 nH	2.73 nF	18.05 pF	2.96 Ω

resistance in the inductors (L_1 and L_0) can be studied. The usefulness of this method, however, ends here if it is also necessary to

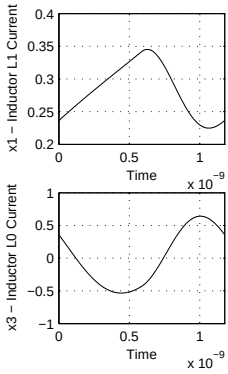
- Model C_1 as the non-linear C_{db} of a MOS Field Effect Transistor (MOSFET) [Chu94]
- Include a more complex model for the inductor [Nik98, Cao03]
- Fully optimize a multistage amplifier

One possibility to model the non-linear drain-bulk capacitance of a Metal Oxide Semiconductor (MOS) transistor is to make use of differential equations already derived and to make capacitor value C_1 voltage dependent. The resulting waveforms of one of such simulations on the same Intel P4 processor at 2.4 GHz computer is given in Fig. 4.6(c) and takes about 5.85 s which is an overhead of 167 times. Compared to less than 1 s for a Simulation Program with Integrated Circuit Emphasis (SPICE) simulation, with all the advantages it gives, plus the speed, the above methods do not seem particularly suitable for optimizing the circuit with the characteristics of a class E amplifier.

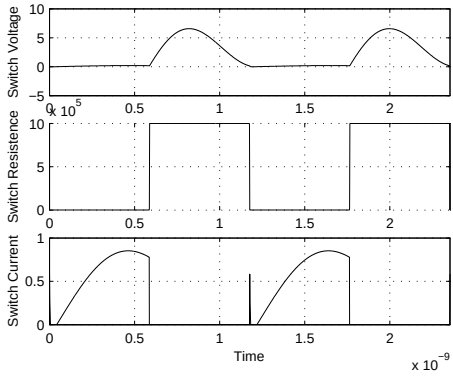
In fact, the other two items above can in an extreme case be solved using the state-space description. The major drawback is the huge effort in manual modeling necessary to reach such a solution. The difficulty to easily change from one topology to another or the difficulty to replace one device model by another can be considered, from a practical point of view, to be a disadvantage. One solution to reduce the quantity of hand-crafted work is to make use of already existing applications for circuit simulation (SPICE) and adding the missing functionality. Finally, gluing them together to form one tool can be an alternative. The developed tool suitable for RF power amplifiers is described in detail in the next sub-section.

4.4.2 Power Amplifier Automated Sizing

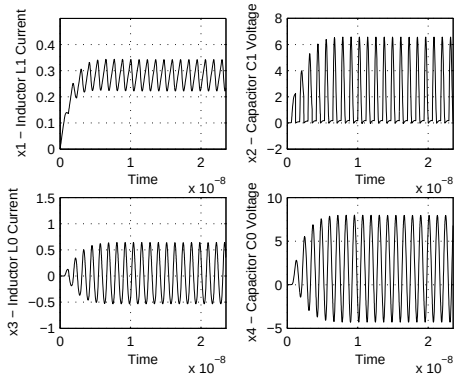
This sub-section presents a methodology for optimally designing power amplifiers for maximum efficiency. The tool methodology is firstly presented in 4.4.2.2 which allows to find optimal values for all components (transistors, passives) using realistic operating conditions, including the effect of their parasitics. A concise description on the automated analog sizing which is based on a simulation-based optimization is then given. This sub-section then ends with a description of the different device profilers that accurately extract device layout parasitics that are then used in user-defined device models.



(a) State variable waveforms.



(b) Switch waveforms.



(c) Time domain waveforms obtained by solving the differential equations set (4.13)-(4.16) with the same circuit sizes used in Fig. 4.6(a). After the initial transient time, the same steady state is reached.

Figure 4.6: Simulated waveforms of the circuit in Fig. 4.5 using two different approaches: by first finding the initial conditions of the state-space description and through solving the differential equations set.

4.4.2.1 Introduction

In previous work [Sok75, Man90, Chu94, Cha01], one or more of the following assumptions or simplifications (approximations of real behavior) are made:

- the transistor acts as an ideal switch
- the inductance of the DC feed inductor (RF choke) is infinite
- the quality factor of the resonant tank is infinite
- one or more passive components are linear
- circuit devices have no parasitics
- electromigration does not take place

In [Gup01] a CMOS RF power amplifier including on-chip matching networks is optimized using a simulated annealing based tool. For the inclusion of the parasitics in the Computer-Aided Design (CAD) optimization, a compact inductor model is created for a specific geometry and valid for a limited range of inductor values. In [Cho02] a CMOS class E power amplifier with bond wire and silicon inductors is optimized using simulated annealing.

All published solutions use a fixed model for the inductor, either limited in the physical geometry or in the lumped representation of the inductor. Square silicon inductors, although easily implementable and commonly used in any technology, do not provide the best performance. With the tool described next, the possibility to seamlessly add new coil geometries offers the advantage of being able to take benefit of new trends, such as the trend that foundries nowadays allow at least octagonal structures. Alternatively, the designer might want to use a broadband compact model [Cao03] for better representation over a wide frequency range or to have the flexibility to use their own model. Another important issue related to power amplifier design is the maximum current that can flow through the transistor and interconnect lines to achieve the desired output power. One of the problems for current values in the range of 500 mA is electromigration. This problem is not only related to the large currents that flow through the metal lines, but is also due to the fact that, at radio frequencies, the skin effect must be taken into account when designing an inductor. Accurate and efficient RF design automation therefore requires that all of the above issues can be taken into account by using an integrated tool flow such as the one presented in this sub-section.

The methodology (described in detail in 4.4.2.2) has been implemented in a software tool called PAMPER [2] and encompasses a simulation based circuit optimizer (DANCE [3] [Fra03], which is part of the developed software environment presented here) and device profilers that accurately extract device layout parasitics that are then used in user-defined device models within the DANCE tool. The DANCE module is discussed in detail in 4.4.2.3. The principle of device profilers is treated in 4.4.2.4.

[2] PAMPER stands for Power AMPlifiER.
[3] DANCE — Device-level ANalog Circuit Environment

Some of the features of the proposed methodology and tool implementation include:

- simulator independence (currently configured to work with Eldo™, HSPICE and MATLAB®)
- knowledge in the form of objectives and constraints is stored in a template which can later be reused
- replacement of any device by user-defined device models
- information on process variation of the inductors and other devices can be taken into account during the optimization
- integrated support for taking the maximum current that flows in the inductors into account during optimization (needed for meeting the electromigration constraint)

4.4.2.2 Tool Methodology

The proposed methodology encompasses a preparatory phase (which is represented by the shaded part at the top of Fig. 4.7) and the actual automatic design optimization where the start is a given circuit schematic. For power amplifiers the influence of coil parasitics has a severe impact on the overall efficiency of the PA. However, accounting for these parasitic effects in the design is tedious and time consuming. Especially in the case of silicon inductors, the resulting model is complex and difficult to analyze theoretically. Still, the presence of parasitics needs to be taken into account during simulation and sizing of the circuit schematic. To this end, an analog device-level circuit simulation and sizing tool (DANCE) that allows to take user-defined models for any selected device into account is developed. This piece of software – a tool on its own – is elaborated in the next section. However, the model parameters should also be made available to this tool. For the methodology presented here this means that a database of model parameters should be generated for the devices, including the effect of layout parasitics. This is done by device profilers (see 4.4.2.4).

One last input to the DANCE module is a rule-set containing design constraints (such as the desired output power, maximum acceptable voltage at turn-on, maximum drain voltage slope before turn-on, ...) and objectives (such as maximum efficiency) for the circuit. This can be either inserted by the (experienced) designer or a default template can be selected that already includes expert designer's knowledge. The DANCE tool, the optimizer and the device profiler are all written in standard C code resulting in fast execution times.

4.4.2.3 Automated Analog Circuit Sizing

Overview

Automated analog circuit sizing approaches can be classified into two broad categories [Car96]. A first class is the *knowledge-based* approach. The main advantage here is the computational speed. Drawbacks of this method are the flexibility and the time needed to develop the knowledge plan. Moreover, this plan usually has to

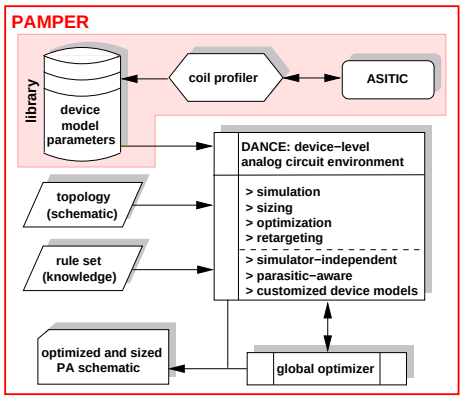


Figure 4.7: Flow diagram of the PAMPER tool.

be adjusted to support new process technologies. Another possible disadvantage is that the accuracy is proportional to the complexity of the plan. In a second class, the *optimization-based* approach, two subcategories can be found: equation-based optimization – which suffers from similar accuracy limitations – and simulation-based optimization. The big advantages of the latter are its flexibility and accuracy. It has the same accuracy as the simulator normally used in hand-crafted sizing. Here an optimizer iterates over simulations for different values of the device parameters to tune the circuit performance. However, the penalty clearly resides in the longer CPU time. The continuous increase in computing power alleviates this drawback partially.

Especially if parasitics (of active and passive components) are taken into account, an analytical solution for finding the optimal parameters of analog (RF) circuits becomes difficult. Therefore, the tool that is capable of:

- performing parameterized simulations (through the use of a GUI if desired) using any circuit simulator (can be configured without changing the code)
- sizing a given circuit topology from scratch
- automatic replacement of any device, active or passive, by customized device models (taking e.g. parasitics into account)
- retargeting a sized schematic:
  - to a new process technology
  - toward new performance specifications
  - to fine tune (optimize) an existing design

Sizing Methodology

The implemented sizing methodology is a simulation-based optimization approach using a differential-evolution optimization algorithm [Sto96] (the *global optimizer* block

in Fig. 4.7). The key property of this optimization algorithm is that it generates new parameter vectors by adding the weighted difference vector between two population members to a third member. If the resulting vector yields a lower objective function value than a predetermined population member, the newly generated vector replaces the vector with which it is compared. For each vector $\mathbf{x}_{i,G}$ of generation G, a perturbed vector $\mathbf{v}_{i,G+1}$ is generated as follows:

$$\mathbf{v}_{i,G+1} = \mathbf{x}_{r_1,G} + F \cdot \left(\mathbf{x}_{r_2,G} - \mathbf{x}_{r_3,G}\right) \tag{4.20}$$

The indexes r_1, r_2 and r_3 indicate three randomly chosen individuals of the population. They are mutually different integer indexes ($\in [0, (N-1)]$) and also differ from the running index i. The (real) constant factor F ($\in [0, 2]$) controls the amplification of the differential variation. The vector $\mathbf{x}_{r_1,G}$ that is being perturbed has no relation to the vector $\mathbf{x}_{i,G}$ that will potentially be replaced. To increase the potential diversity of the perturbed parameter vectors, crossover is introduced. More information about the algorithm and details of several variants or strategies for constructing new parameter vectors can be found in [Sto95, Sto96]. In addition, this algorithm has been altered to include parameter bounding, stop criteria and mixed continuous/discrete parameter support.

Any circuit variable (device sizes, component values, bias inputs, ...) can be selected as optimization parameters. Furthermore, one or more optimization *objectives* (minimize, maximize) can be specified as well as a number of (performance) *constraints* (e.g. $A_{LF} > 60$dB or $V_{node1} < 0.1$V). All these requirements (n objectives and m constraints) are combined into a single cost function which can be evaluated by the optimizer:

$$\begin{aligned} Cost &= W_{obj} \cdot \sum_{i=1}^{i=n} P_{sim_i} \\ &+ W_{con} \cdot \max_{j \in [1,m]} \left(\frac{P_{spec_j} - P_{sim_j}}{P_{spec_j}} \right) \end{aligned} \tag{4.21}$$

with W_{obj} and W_{con} the weights for the cost due to the objectives and constraints, respectively, and where P indicates performances, either simulated or specified. With properly scaled weights (which is very easily accomplished), the optimizer will first try to find feasible solutions (satisfying the constraints) and then further tunes the parameters to optimize the objectives. This scaling can easily be adjusted manually after a "*dry-run*" (which could also be automated) and only requires altering the order of magnitude of one of the weights depending on the cost values which are logged. In order to deal with complex problems with many constraints, a minimax problem formulation is used in (4.21). When the genetic algorithm proposes bad combinations of parameters (e.g. out of bound), a "high" cost is assigned (e.g. 10^8) to such solutions.

In order to facilitate the automated optimization of specific circuit classes (Op-Amps, comparators, ...), constraint and objective templates can be loaded. These could have

been stored for re-use by the designer himself or provided by another expert designer. Furthermore, in order to realistically take parasitics of active or passive devices into account, customized device models can be input to the tool. Fig. 4.8 illustrates such model for an inductor. In the PAMPER tool these models are entered using standard SPICE syntax. The actual values for the components that are part of a specific device model are derived beforehand in a tabulated form by device profilers (that can steer external tools). The effective values of the devices in the model are calculated by the optimizer depending on the optimization parameters and by interaction with the device profiler (see Fig. 4.7).

4.4.2.4 Device Profilers

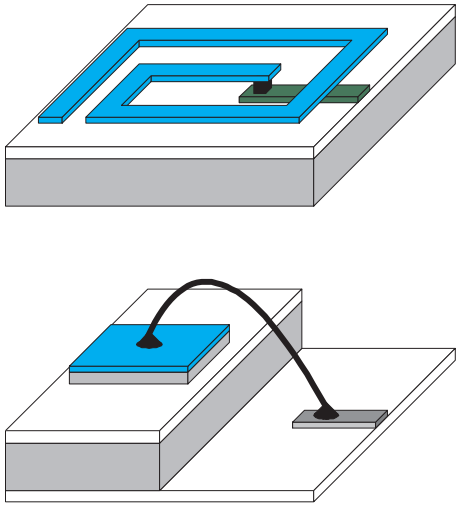
In principle, a profiler for any device could be implemented. However, for devices such as capacitors, pads, interconnections and transistors, the parasitics dependent on the layout implementation can, with limited complexity, be manually extracted by the designer or have already been analyzed and included in a knowledge library. Coil parasitics, on the other hand, need more attention and for this purpose an automated extraction tool has been included that can interact with external tools.

Inductor Parasitics

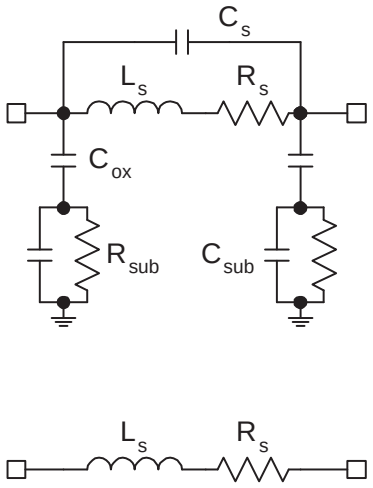
For the power amplifier design optimization tool, a device profiler for the coils is developed in order to generate the parameters for the coil model, including the effects of layout parasitics. To this end, an external tool is called for calculating the model parameters, which are then stored in a table (by the coil profiler) to be accessed later during sizing (by the DANCE module). The external tool used for analyzing silicon inductors is ASITIC [Nik98]. The coil profiler should be run only once for each technology process and takes a range of inductor values as input together with:

- the technology process
- the circuit operating frequency and
- the inductor geometry

Then tables are generated (for different currents) by ASITIC, using the selected model (see Fig. 4.8) and parsing the ASITIC output file. The information returned by ASITIC is the required inductance together with the device model parameters and layout information (number of turns and turn spacing). For each coil model and maximum allowed current a table is thus generated, containing a discrete list of different inductor values within a specified range. This range typically includes about 10 values around an anticipated value for the target application. Since the tables are generated only once for each technology the computational time is less important. Typically, a table contains about ten entries which each take around 20 minutes to generate. Other values for the inductors are then linearly interpolated. This technique is experimentally verified (by comparing interpolated values with the result obtained with ASITIC for the exact values) to be accurate enough. Hence, no expert knowledge is needed to select the range of inductor values and sampling density.



(a) Integrated spiral and bonding wire inductor.



(b) Equivalent lumped circuit model including inductor parasitics.

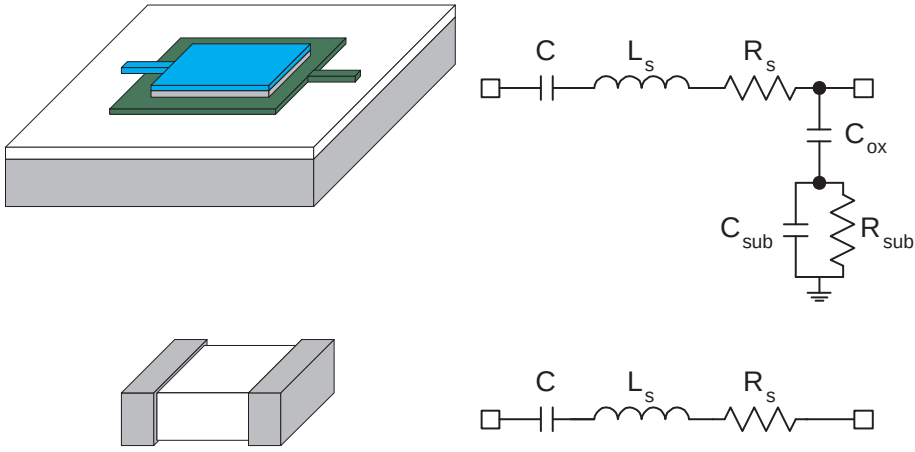
Figure 4.8: Example schematics of possible device models for an inductor serving as input to the DANCE tool. Any customized model representing more complex inductor layouts (octagonal, hexagonal, etc.) or a more accurate representation over frequency can be used.

Bonding wires (Fig. 4.8(a)) are still commonly used to get the best performance due to their higher quality factor. It has been shown that the machine-bonded bonding wires have less than 5 % inductance variation and less than 6 % quality factor variation [Lee98], making them suitable for integration. The inclusion of a resistor in series with the inductor is a simple way of representing its losses (Fig. 4.8(b)).

Capacitor Parasitics

At RF frequencies, the use of poly-poly capacitors is usually avoided. The high sheet resistance significantly degrades the quality factor of the capacitor. This effect is more noticeable as the frequency increases and the associated losses can be considerable. Metal-metal capacitors (Fig. 4.9(a)), existing in new technology processes with more emphasis on analog RF circuits have a relatively good quality factor, making it suitable for high-performance circuits. Care must be taken with the wiring. The series inductance (L_S) limits the maximum cut-off frequency of the device. Only for frequencies of interest much lower than this value is the high quality assured.

For frequencies below 1 GHz, discrete SMD (Fig. 4.9(a)) capacitors can still be used if their value is in the range of a few picofarads. For higher values of capacitance or



(a) On-chip metal-metal and SMD capacitor.

(b) Equivalent capacitor circuit model including parasitics.

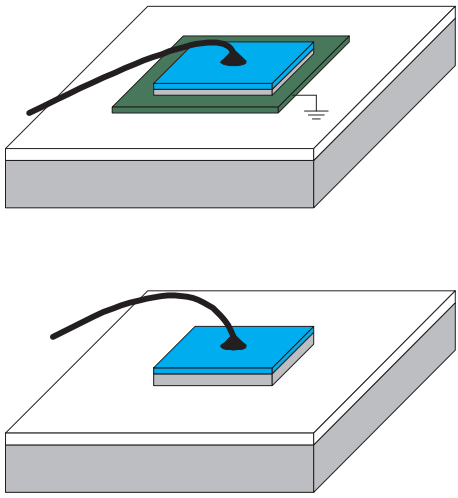
Figure 4.9: Possible capacitor elements including parasitics for accurate representation over frequency as used by the DANCE tool.

frequency of operation, the only possible solution is to use a completely integrated solution. If metal-metal capacitors are not available in the technology, capacitors with combined lateral and vertical field components are a practical solution [Apa02]. Some structures can even provide higher oxide capacitance density than the more expensive to manufacture metal-metal capacitors.

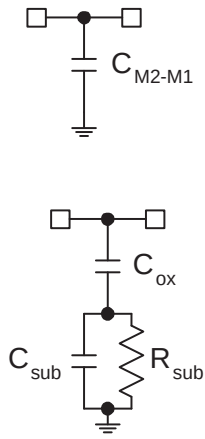
The lumped model from Fig. 4.9(b) can be obtained with a generic electromagnetic analysis program by considering the oxide between the capacitor plates as the via of an inductor and then curve fitting the results to a capacitor equivalent model. ASITIC takes this approach, making it equally suitable for analyzing capacitors. Parasitics to substrate can be minimized by either using a high-resistivity substrate or by placing a ground shield underneath. The latter technique can also be used to prevent the substrate resistance from injecting noise currents. The model for a SMD capacitor can be obtained from the manufacturer datasheet or easily measured with a vector network analyzer.

Pad and Interconnection Parasitics

In any circuit that has to connect to the outside world, a pad is necessary (Fig. 4.10). A pad can be simply a stack of metal layers connected through vias. Where a low capacitance value is required only the top metal layer can be used. Also, some applications require isolation from the noisy substrate or a well defined impedance. The solution in this case is to place a ground connected metal shield underneath. The resonance



(a) Shielded (top) and unshielded (bottom) bond pad. A lower metalization layer is used to effectively shield from substrate losses.



(b) RC representation of the devices on the left.

Figure 4.10: Lumped circuit models of a shielded and unshielded pad. The DANCE tool can use these and other equivalent models.

frequency decreases as the capacitance now has a higher value. The advantage is that the capacitance has in this situation a well defined value.

Interconnections can be modeled as a simple pad model. Depending on the frequency, it might be more accurate to also include in the model the series resistance and inductance. In this case the model will resemble the one used for the integrated capacitor of Fig. 4.9.

Transistor Parasitics

For the transistor, a simple switch S can represent its off- and on-resistance, although a more complete model can also include losses due to non-zero switching time and lead inductance [Raa78b]. For CMOS integrated circuits, the BSIM model is usually available. This rather complex model includes effects such as modeling of noise and temperature effects. The new version (BSIM4) is tailored for sub-100 nm devices and includes, for example, the substrate network necessary for accurate RF design.

In a class E amplifier, power is dissipated due to the switch-on resistance (R_{ON}) (Fig. 4.11). The efficiency can be approximately given by [Raa78b]

$$\eta \simeq \frac{R_L}{R_L + 1.365 R_{ON}} \tag{4.22}$$

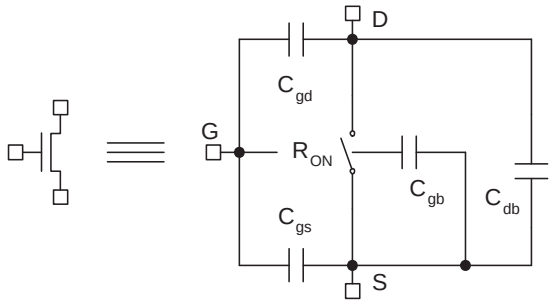


Figure 4.11: Possible equivalent circuit for a transistor in a class E power amplifier. The model includes the off- and on-resistance and the capacitances between the gate and the other three terminals. It gives satisfactory results when applied to a class E power amplifier.

where

$$R_L = \frac{8}{\pi^2 + 4} \frac{V_{DD}^2}{P_{OUT}} \tag{4.23}$$

In the above equation and for a constant output power we have

$$R_L \propto V_{DD}^2 \tag{4.24}$$

which leads to the conclusion that the higher the voltage the higher the load resistance R_L. This in turns means that more power is available to the load for a fixed loss in the parasitics of the series-tuned output filter.

Other Parasitics

Care has been taken so that any structure that can have a SPICE representation can be optimized. As an example it is necessary to consider the delay of a transmission line or the losses associated with other parasitics where the die will be attached. Up to certain limits the non-dissipative parasitic elements (inductors and capacitors) can be included and masked in the overall amplifier characteristics without significant decrease in performance.

4.5 Optimization Experiments

As mentioned, one of the difficulties in the design of class E power amplifiers resides in the lack of precise analytical equations for analysis and synthesis. As such, simulations will be used to get insight on the influence of parasitics on the circuit performances.

The next four experiments will take the generalized class E power amplifier and optimize it for the maximum DE for various operation conditions. The influence of different circuit elements parasitics or different design objectives is presented in a tabled

form together with a figure to better illustrate the results. In the last experiment, a two stage class E power amplifier is optimized.

In the first example, the effect of different supply voltage and inductor parasitics in the DE and PAE is given. While the second experiment deals with different output levels and constant frequency, the third covers an amplifier with different transistor sizes for a constant output power. In the last in this series, the influence of different frequencies of operation with a constant output power is presented. Finally, an application example for a two stage differential class E power amplifier is given. In this example, the limiting parasitics for passive and active components are included. Their accounting during optimization is essential for maximizing the circuit's efficiency. The amplifier from Fig. 4.12 is studied under the following conditions:

- class E operation conditions are always met (see 4.3.2.2)
- transistor and inductors have parasitics
- 0.35 µm CMOS technology
- square wave input signal with amplitude equal to the supply voltage

The simulator that is used in the following experiments is Eldo™ [Eldo]. However, the tool works with any (circuit) simulator (or general-purpose program) as long as it outputs its results to a file. The simulation configuration is done once by simply editing a form in the user-interface and is pre-configured for Eldo™, HSPICE, and MATLAB®.

To account for driving power and in order not to change the results obtained for the DE, the average power necessary to drive a capacitive load (C_{gg}) can be approximated by

$$P_{IN} = \alpha . C_{gg} . V_{DD}^2 . f \tag{4.25}$$

The constant α (the value of 1 will be used) represents a proportionality factor representing an inverter chain driving the transistor total input gate capacitance C_{gg}. Finally, each one of the simulations is optimized for maximum DE (4.2). The results are now presented for different case studies.

4.5.1 Influence of the Supply Voltage and Inductor Quality Factor

In order to show the relation between the inductor quality factor (Q_L) and the supply voltage, the circuit from Fig. 4.12 is simulated considering a bonding wire inductor (Q_L=40) and an on-chip spiral inductor (Q_L=7). For reference, simulations assuming an inductor without losses are also given.

The results of these simulations, with the inputs given in Table 4.2, are presented in Fig. 4.13(a). From what can be seen, the drain efficiency is almost equal and independent of the supply voltage, being much more dependent on the inductor quality factor. Note that in the case of lower supply voltages (imposed by technologies with smaller geometries) it is not possible to achieve the desired output power.

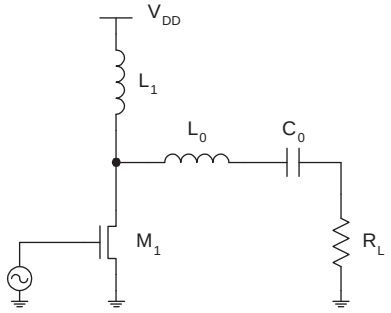


Figure 4.12: Class E amplifier schematic with the switching device replaced by a NMOS transistor. In this way, the effect of the different capacitances between the gate and the other terminal are considered. Capacitor C_1 in Fig. 4.3 is now the drain-bulk capacitance of the MOS transistor and reactance jX is merged with the inductor L_0.

A higher supply voltage does not allow to get a higher DE (Fig. 4.13(a)). This can be justified by the fact that [Sok75]

$$R_L = \frac{8}{\pi^2 + 4} \frac{V_{DD}^2}{P_{OUT}} \tag{4.26}$$

$$C_1 = \frac{P_{OUT}}{\pi \omega V_{DD}^2} \tag{4.27}$$

have opposite relationship to the power supply voltage. As the transistor capacitance decreases, the load resistance increases with an increasing supply voltage. The ideal is to balance the losses in a smaller transistor (higher on-resistance) with a higher load resistance.

Regarding the power added efficiency (PAE, Fig. 4.13(a)), going to higher supply voltages has some advantages. This is due to the fact that the transistor is smaller (4.27) and if the supply voltage on the previous stage is kept constant, it is now possible to save on driving power, therefore, increasing the PAE with an increasing supply voltage (Fig. 4.13(b)). Moreover, a lower current is necessary to achieve the same output power, hence, reducing the electromigration reliability issues.

In this situation the parasitics are relatively small, leading to an efficiency higher than 90 %. Basic equations already predicted that for MOSFET [4] it is normal for efficiency to vary only very slightly with V_{DD} [Sok75]. Results in Fig. 4.13 show that this is valid

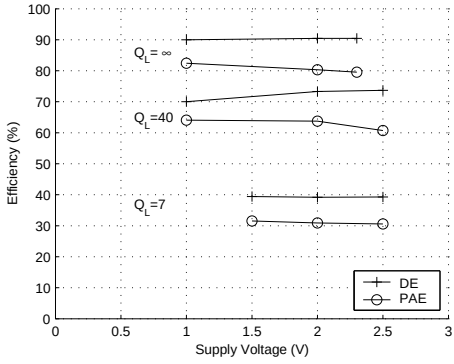
[4] BJTs have $V_{CE(sat,offset)}$, which reduces the effective value of V_{DD}, for producing RF output power from a switching-mode RF PA (e.g., class E or D). The effective value of V_{DD} is [actual V_{DD} - $V_{CE(sat,offset)}$]. But the value of V_{DD} from which the DC input power is calculated, is the actual V_{DD}, as $P_{SUPPLY,DC} = V_{DD} I_{DD}$. By comparison, MOSFETs do not have a saturation offset voltage, so the full value ofV_{DD} is used in the equation that predicts RF output power.

Table 4.2: Inputs for the first experiment: influence of the supply voltage and inductor quality factor.

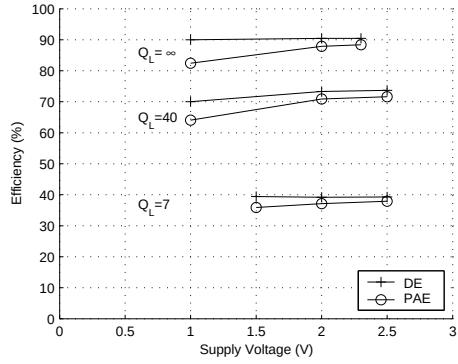
Simulation Conditions	Variables
Input frequency: 850 MHz	Supply voltage
Output power: 0.5 W	Inductor quality factor $\in [7, \infty]$

Table 4.3: Effect of the supply voltage and inductor quality factor on the circuit sizes.

Q_L	V_{DD}	M_1	L_1	L_0	C_0	R_L
∞	2.3 V	11134 µm	11.75 nH	2.51 nH	24.7 pF	3.54 Ω
	2.0 V	13715 µm	8.94 nH	2.18 nH	25.8 pF	2.86 Ω
	1.0 V	41203 µm	2.71 nH	1.16 nH	36.7 pF	0.81 Ω
40	2.5 V	13832 µm	6.39 nH	1.55 nH	65.0 pF	3.03 Ω
	2.0 V	16000 µm	3.82 nH	1.44 nH	48.2 pF	2.79 Ω
	1.0 V	66467 µm	0.91 nH	0.50 nH	103.2 pF	0.54 Ω
7	2.5 V	17399 µm	1.59 nH	1.29 nH	64.7 pF	3.28 Ω
	2.0 V	25986 µm	1.00 nH	0.86 nH	88.8 pF	2.21 Ω
	1.5 V	43435 µm	0.62 nH	0.45 nH	169.8 pF	1.05 Ω



(a) With V_{DD} in (4.25) equal to the supply voltage.



(b) With V_{DD} in (4.25) equal to 1 V.

Figure 4.13: Drain Efficiency and Power Added Efficiency as a function of supply voltage for different inductor quality factors Q_L.

even in the presence of important parasitics. In addition, Fig. 4.13(a) reveals similar behavior for the PAE.

Consequently, if the goal is to have a totally integrated amplifier, on-chip spiral inductors limit the maximum drain efficiency to about 40 %. However, if the intention is to have an amplifier with an efficiency in excess of 50-60 %, the use of high-quality inductors (bonding wires) is mandatory. This fact is illustrated in Fig. 4.13(a). Finally, circuit sizes are given in Table 4.3.

4.5.2 Influence of Different Output Power Values

When designing a power amplifier for a maximum given output power, the question can arise on whether the usage of one single amplifier or multiple parallel amplifiers (with the same total maximum output power) is more beneficial or not. Three simulations with three different output power levels are performed with the objective of studying this effect. The circuit from Fig. 4.12 is optimized for highest drain efficiency under the conditions summarized in Table 4.4. The results for output power from 2 W until 0.25 W are given in Fig. 4.14. Optimized component values are also summarized in Table 4.5. The results show that the DE remains constant and is independent of the output power level. The situation is somehow different for the PAE where one amplifier at a higher output power has a transistor size slightly smaller than the sum of two amplifiers having each one of them half the output power. The consequence of a reduced input capacitance is the reduction of the power needed to drive the amplifier. This can be seen in Fig. 4.14 where the PAE decreases with a decrease in the output power. Device circuit sizes in Table 4.5 show the expected from theory [Sok75]: the reduction of the inductance size and output load resistance along with the increase of the output filter capacitance. Also, because capacitor C_1, representing the transistor output capacitance, is proportional to the output power (4.27), an increase in the transistor size is expected.

The above results tend to support the idea that, for the case of maximum drain efficiency and the highest output power level, there is no difference in building one big amplifier or use smaller amplifiers in parallel. For the case of parallel amplifiers, it is also necessary to consider the losses in the combiner circuit which can in some cases make one topology more efficient that the other one.

4.5.3 Influence of Different Transistor Sizes

For the specific case of a class E power amplifier (Fig. 4.12), it is commonly assumed that having a higher switch-on resistance necessarily implies a less efficient amplifier [Raa78b]. This is true in case the shunt capacitor (C_1 in Fig. 4.12) is not the parasitic drain-source capacitance and thus can have its value selected independently of the transistor width. To evaluate this effect, a simulation of a typical amplifier (Fig. 4.12 under the conditions from Table 4.6 is performed. Table 4.7 and Fig. 4.15 present the results for the case of having transistor sizes ranging from 4 to 24 mm. Above the threshold of 16 mm, the DE stays constant for an increasing transistor width. Therefore, a decrease in the switch-on resistance does not translate to an increase in the drain efficiency if

Table 4.4: Inputs for the second experiment: influence of different output power values.

Simulation Conditions	Variables
Supply voltage: 2 V	Output power $\in$ [0.10 W, 2.00 W]
Input frequency: 850 MHz	
Inductor quality factor: 40	

Table 4.5: Circuit sizes for different output power values.

P_{OUT}	M_1	L_1	L_0	C_0	R_L
2.00 W	66987 µm	0.93 nH	0.44 nH	128.5 pF	0.64 Ω
1.00 W	28498 µm	1.87 nH	0.85 nH	70.9 pF	1.55 Ω
0.50 W	16000 µm	3.82 nH	1.44 nH	48.2 pF	2.79 Ω
0.25 W	9681 µm	7.49 nH	2.84 nH	23.1 pF	4.25 Ω
0.10 W	3664 µm	18.03 nH	6.74 nH	10.7 pF	11.94 Ω

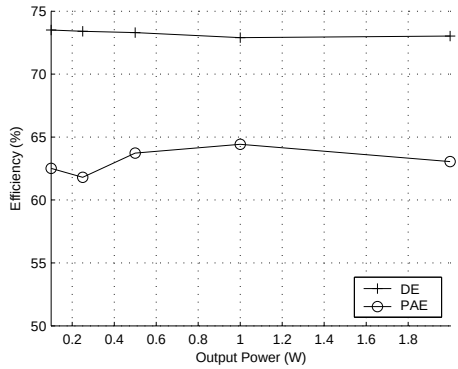


Figure 4.14: Efficiency for different output power values.

the shunt capacitor is the transistor output capacitance. It is the coupling between the output capacitance and the transistor width that causes this effect.

Despite the highest DE being obtained for a transistor length of about 16 mm, this does not reflect in the PAE. A smaller transistor, although having a higher on-resistance, can be used to make a more efficient power amplifier. As such, a transistor size in the range of 8 mm will indeed produce the PA with the highest PAE (Fig. 4.15). Nevertheless, a slight change in the circuit is necessary because of the relation between the output power and shunt capacitance (4.2). This capacitance must be increased, either by increasing the drain-bulk capacitance or by using an extra capacitor in parallel. The results where this capacitor is increased are represented by a dashdot line in Fig. 4.15 and by the symbol † in Table 4.7.

In conclusion, increasing the device size beyond the minimum required, although reducing the switch-on resistance, does not improve the drain efficiency.

4.5.4 Influence of Different Frequencies of Operation

The model in Fig. 4.12 has a good agreement with a more complex one where the transistor is modeled with its BSIM3 model. In addition, for frequencies much below the transistor cut-off frequency (f_T), the capacitance can be considered constant in terms of frequency. As a consequence this allows the PA efficiency to be tested in terms of frequency.

In the following set of simulations, the bonding wire inductor quality factor (Q_L) is considered a constant on all frequencies, in spite of a three-dimensional inductance extraction program (FastHenry), showing that inductor series resistance, due to the skin effect, increases with frequency. In this way, only the effect of a different frequency is seen in the simulations. Table 4.8 summarizes the optimization conditions.

Nevertheless, if the corrected resistance is included in the simulations, the efficiency will slightly increase for frequencies below 850 MHz, and decrease for higher frequencies. Although this may be true, the changes are not large enough to change the relation seen in Fig. 4.16 and the graphic has the same bearing. As such, DE still increases with an increasing frequency for values in the neighborhood of 850 MHz. The optimized component sizes are given in Table 4.9.

Fig. 4.16 presents the DE and PAE results obtained as a function of input signal frequency. As can be seen, both the drain and power added efficiency increase for an increasing frequency. These results are an indication that comparison of different power amplifiers has to be done at a similar frequency of operation. In fact, different CMOS technologies with different characteristics and other parameters do influence circuit performance. This once more supports the assertion that the comparison of different circuits is a research field on its own.

Table 4.6: Inputs for the third experiment: influence of different transistor sizes.

Simulation Conditions	Variables
Supply voltage: 2 V	Transistor width $\in$ [4000 µm, 24000 µm]
Input frequency: 850 MHz	
Output power: 0.5 W	
Inductor quality factor: 40	

Table 4.7: Optimized circuit sizes considering a fixed transistor size.

M_1	L_1	L_0	C_0	R_L
24000 µm	5.10 nH	0.98 nH	79.9 pF	1.35 Ω
20000 µm	3.87 nH	1.25 nH	58.1 pF	2.03 Ω
16000 µm	3.82 nH	1.44 nH	48.2 pF	2.79 Ω
†12000 µm	3.79 nH	1.40 nH	51.0 pF	2.89 Ω
†8000 µm	3.29 nH	2.08 nH	25.5 pF	3.44 Ω
†4000 µm	1.88 nH	2.65 nH	17.1 pF	3.38 Ω

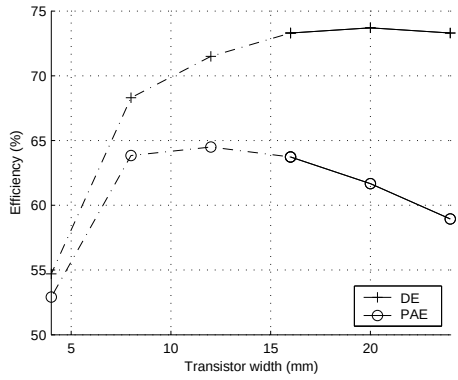


Figure 4.15: DE and PAE for an amplifier with different transistor sizes.

Table 4.8: Inputs for the fourth experiment: influence of different frequencies of operation.

Simulation Conditions	Variables
Supply voltage: 2 V	Input frequency $\in$ [500 MHz, 1200MHz]
Output power: 0.5 W	
Inductor quality factor: 40	

Table 4.9: Resulting device values optimized for different frequencies of operation.

Frequency	M_1	L_1	L_0	C_0	R_L
1200 MHz	12024 µm	3.15 nH	1.07 nH	30.2 pF	2.45 Ω
1000 MHz	13225 µm	3.44 nH	1.42 nH	31.3 pF	2.80 Ω
850 MHz	16000 µm	3.82 nH	1.44 nH	48.2 pF	2.79 Ω
700 MHz	20274 µm	4.02 nH	1.74 nH	58.8 pF	2.76 Ω
500 MHz	22274 µm	4.83 nH	2.92 nH	59.6 pF	3.65 Ω

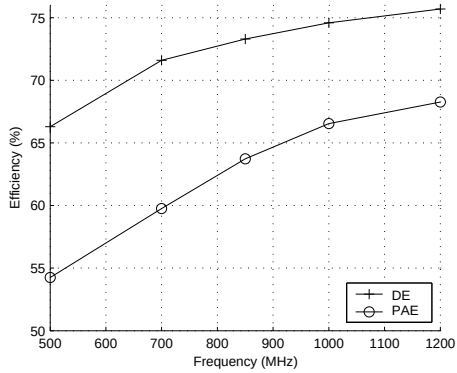


Figure 4.16: Variation of the DE and PAE with varying operation frequency.

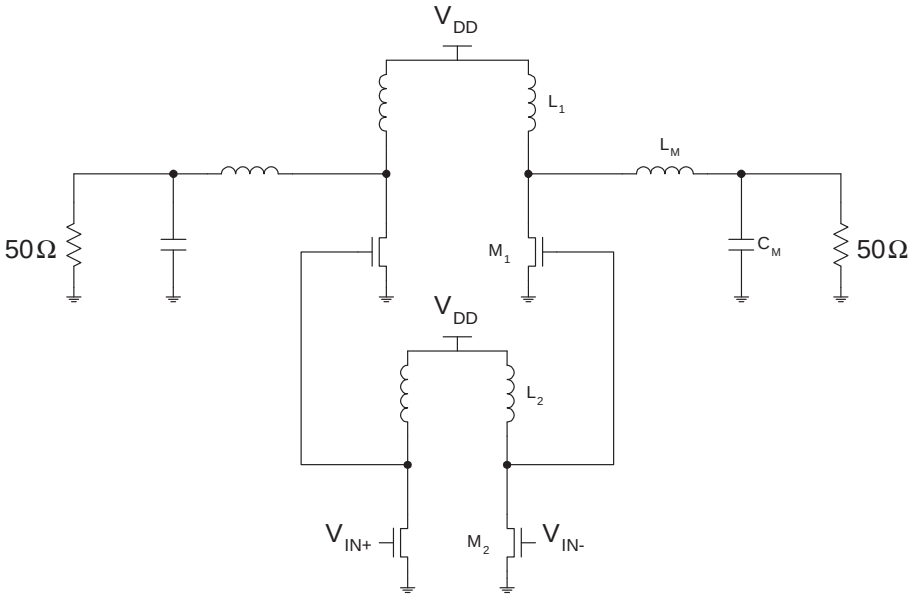


Figure 4.17: Circuit schematic of the differential two stage class E power amplifier. Appropriate parasitics are included for better representation of the losses. Measured efficiency should thus be similar to the optimized value obtained here.

4.5.5 Two Stage Amplifier

Of course, more complex and more functional circuits can also be optimized and studied. As an example, the circuit of Fig. 4.17 is optimized for highest drain efficiency in a 0.35 µm CMOS technology process to meet class E operation under the following conditions:

- supply voltage: 2 V
- input frequency: 850 MHz
- input power: 10 dBm
- output power: 1 W

After including all relevant device parasitics to the netlist to better represent the silicon implementation, the circuit is optimized.

The resulting drain efficiency, including the effect of the inductor parasitics, optimized by the PAMPER tool is 63 % (power added efficiency is 62 %). Here, the optimized result will be close to the measured one because the performance-limiting parasitics

Table 4.10: Constraint template.

Block	Constraint / objective	Value
	Power Supply Dissipation	minimize
	Inductor current (RMS, MAX)	$<$ 500 mA
Driver Stage	V_{DS} at turn-on	$<$ 2.0 V
	V_{DS} at turn-on	$>$ 0.0 V
	V_{DS} at turn-off	$< V_{TH}$
Output Stage	V_{DS} at turn-on	$<$ 0.1 V
	Output Power	$=$ 1 W
	V_{DS} in the last half period	$>$ -0.2 V
	V_{DS} slope at turn-on	$<$ 9e9 V/s

(from the transistor, pad, etc.) are taken into account. Besides the inductors, all other parasitics are modeled from the layout using SPICE subcircuits where the model values are calculated by the tool during sizing. The advantage is that the sizing is completely automatic. Simulation results of the optimized circuit are shown in Fig. 4.18. Both the (optimized) drain efficiency and the power added efficiency as a function of the input power are plotted in Fig. 4.19. Finally, Fig. 4.20 illustrates how the results of the optimization algorithm evolve by plotting the minimum cost versus the number of iterations. It can be clearly seen that functional circuits (complying with all constraints) are found after about 2000 iterations (resulting in a large cost decrease). This corresponds to one hour of optimization time on a multi-user, single processor (Intel XEON 2.4 GHz) PC. Although one of the stop criteria (minimum cost variation threshold) is met, the optimization process is deliberately continued (for a total optimization time of more than three hours) to verify the optimality of the obtained solution which did not further improve.

The constraint (and objective) template that has been used for this experiment is illustrated in Table 4.10. This information is based on the designer's knowledge and reflects realistic operation conditions. When the constraints for the driver stage are omitted, the amplifier has a 5 % higher efficiency (i.e., a drain efficiency of 68 %). The automated tool methodology allows to quickly evaluate these kinds of new trials before investing precious time in a more in-depth analysis. The impact of constraints like these is hard to derive theoretically and could not have been found by a manually handcrafted design. Both circuits sizes are given in Table 4.11. Attention has to be drawn to the relatively small value for the driver stage transistor width when no constraints are used for this stage. Although the simulation waveforms have the expected class E behavior, it is found that the amplifier works on the edge of stability. Any slight variation in the circuit sizes lead to a nonfunctional amplifier. As such, the maximum value of 68 % for the DE is to be seen as the maximum possible value.

Table 4.11: Optimized circuit sizes for the circuit in Fig. 4.17 for the situation where the constraints are applied to both stages or only to the output stage.

Constraint	M_2	L_2	M_1	L_1	L_M	C_M
both stages	2136 µm	1.40 nH	10139 µm	1.76 nH	3.93 nH	8.45 pF
output only	260 µm	2.10 nH	8978 µm	2.20 nH	3.90 nH	8.60 pF

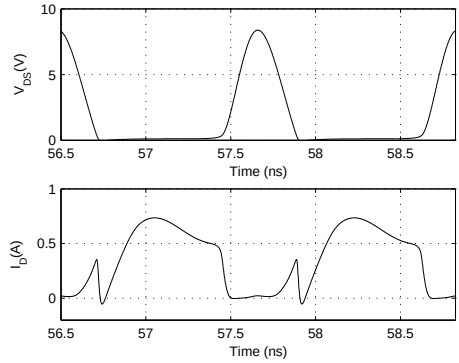


Figure 4.18: Simulated waveforms of the optimized class E power amplifier: drain current and drain voltage of the output transistor obtained with a sinewave with a 1 V peak voltage (+10 dBm input power). Due to the switching nature of the amplifier, sharper edges on the current waveform are obtained for a higher peak voltage which when on-chip can be obtained without the need to be matched to a 50 Ω resistor. And thus saving in the RF input power. Nevertheless, the transistor/switch current will never be similar to Fig. 4.6(b) due to the reason that the output stage is not driven by a perfect square wave.

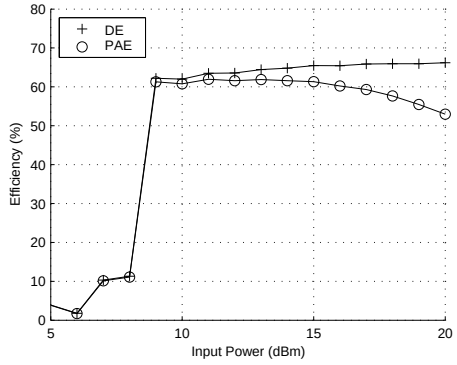


Figure 4.19: Drain efficiency and power added efficiency of the optimized circuit as a function of the input power.

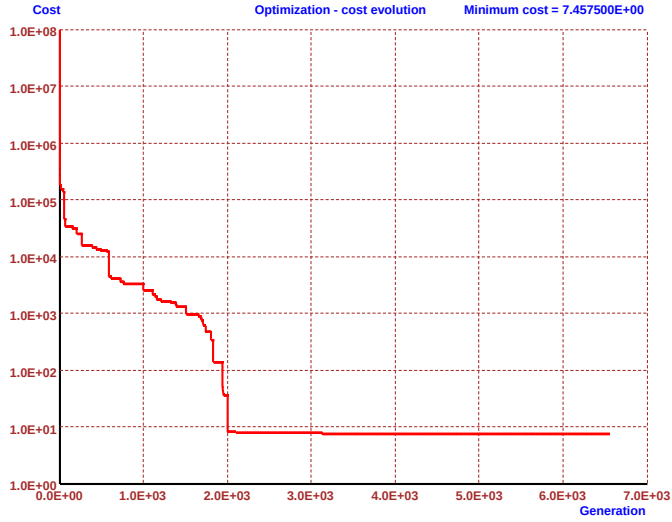


Figure 4.20: Minimum cost evolution during optimization. The optimization is forced to continue after meeting one of the stop criteria.

4.5.6 Comparison with Previous Approaches

In previous approaches the designer is limited by (1) a fixed topology and (2) the impossibility to include parasitics [Cha01]. Parasitics are taken into account in [Gup01] and [Cho02]. Although [Gup01] included inductor parasitics in the design, they are limited to square inductors. Also, the equations describing the inductor are inflexible to include more complex inductor layouts (octagonal, hexagonal, etc.) or models such as the ones given in [Nik98, Cao03]. In [Cho02], the possibility to also simulate bonding wire inductors is added.

In the PAMPER tool described here, new and more complex circuit topologies can be easily simulated and, at the same time, new SPICE device models can be included. These models can describe more complex geometries or a more accurate representation over frequency. For the specific case of power amplifiers, the current flowing through the circuit branches is high and thus it is of vital importance to include the maximum allowed current density. Process variations are also addressed by generating different models for each corner. To the authors knowledge, the proposed software tool is the first to tackle all these problems (including electromigration constraints) simultaneously during sizing [Ram04].

4.5.7 Final Remarks

Other models than the ones that are used in 4.4.2.4 can equally well be used by simply providing the equivalent netlist as input. An equivalent model for the transistor as shown in Fig. 4.11, where the transistor is replaced by a simple resistor and the most important parasitic capacitances further reduces the optimization time from one hour

to less than 15 minutes. The model values for the selected (user-defined) transistor model are calculated by combining layout information (the transistor size and geometry) with information from the foundry (the model file). The information from the layout, such as transistor folding, enables to accurately map the major parasitic capacitances and its resistance. Although the transistor model is only accurate within 5 % of the real transistor (in transient simulations), the same optimal efficiency is found for the synthesized PA.

4.6 Conclusion

Methods to optimized the class E power amplifier have been presented. The state-space description has been shown to be a simple and quick way to optimize the solution. However, the difficulty resides in the fact that it lacks flexibility. Moreover, adding parasitics or modeling some non-linear devices strongly increases the simulation time.

To overcome the above limitation a parasitic-aware power amplifier design optimization tool has been presented. The modular software architecture that is developed to this end seamlessly integrates a highly efficient (genetic) optimization program, an analog circuit sizing tool and a device profiler. By rendering these last two modules parasitic-aware, an accurate RF design optimization tool is obtained. One of the device profilers (namely the coil profiler) makes use of an external program (ASITIC/FastHenry). The profiler can also be configured to interact with other tools for extracting this kind of information.

The PAMPER tool has been used to gain insight on the generalized class E power amplifier. The influence of different design variables is given. It is seen that in order to have high efficiency, inductors with a high quality factor and transistors with small input capacitance are a must. In a last example, the PAMPER tool has been used to optimize the drain efficiency of a class E power amplifier for mobile communications including all the performance-limiting parasitics of active and passive components. The complete sizings are obtained in less than one hour CPU time. The total optimization time is further reduced from one hour to less than 15 minutes by including a simple transistor model.

The methodology is not specific for switched type amplifiers. Linear amplifiers can be optimized and thermal aspects and stability of the transistors or ACPR can be considered by a proper user defined constraint template. Although designed for RF power amplifiers, PAMPER is general enough to be useful for other applications where parasitics must be taken into account.

The next chapter presents techniques to increase the supply voltage in commercial CMOS technologies beyond the maximum stated by the foundry. Technological and circuit solutions are given alongside the advantages and drawbacks of such an approach. Applicability to the design of class E power amplifiers is then discussed.

Chapter 5

High-Voltage Operation

5.1 Introduction

In the previous chapter, the basic class E amplifier topology has been extensively optimized with the inclusion of basic parasitics to better represent final circuit implementation. In the different aspects of the design it has always been assumed that the switch characteristics are fixed, i.e., without considering the possibility of different active devices, such as the Lateral DMOS (LDMOS). This chapter as such complements the previous study in this respect.

More advanced CMOS technologies have the advantage of smaller geometries and distances which reflects in lower parasitics. Most of these characteristics hold true for the majority of the circuits. However, for power driving circuits, each new technology means that the same output power which is dictated by the standards, has to be achieved with an ever decreasing supply voltage. This poses difficulties, as it implies that an increasing current must flow in metallic conductors that have now smaller sections. Higher current also means that every small resistance has an increased importance in terms of power dissipation.

With the above in mind, this chapter will explore possibilities to increase the supply voltage beyond the stated maximum supply voltage of standard CMOS low-voltage technologies. The advantage is that this relatively high voltage is readily available as it is already present in todays 3.6 V batteries of mobile devices.

The outline of this chapter is as follows.

In Section 5.2 several alternatives the designer can use to design high-voltage circuits are briefly considered. Circuit and technological solutions are given and their advantages and disadvantages are discussed.

In Section 5.3 a LDMOS that can be seamlessly integrated on standard CMOS technologies is presented. Device isolation techniques on commercial CMOS processes are briefly discussed. Common characteristics of two of the most common isolation techniques are then used to design a compatible transistor for both. The chapter ends with suggestions on how to improve the breakdown voltage while at the same time decrease the input capacitance.

Experimental results are then presented in Section 5.4. Transistors following foundry design rules, the LDMOS described in the previous chapter and other high-voltage transistors are tested their characteristics compared. DC measurements with different

geometries are included to support conclusions on their effect while AC measurements are given to show its applicability to the RF range of current mobile operated standards.

Based on measured results from the previous section, Section 5.5 addresses the advantages of using high voltage for applications like RF class E power amplifier, either by reverting to an older CMOS technology or using a custom made LDMOS device. The conclusion is that power driving circuits do not necessarily gain from going to technologies with smaller geometries. More advanced technologies offer reduced input capacitance and on-resistance but have the drawback of reduced breakdown voltage and higher electromigration restrictions.

Finally, in Section 5.6 a general overview of the presented work is given and some final conclusions are drawn.

5.2 High-Voltage Solutions in CMOS

Low breakdown voltage in CMOS results from the high field near the drain region edge which is enhanced by the high doping concentration at the surface. A variety of solutions already exist that allow the designer to increase the supply voltage without causing circuit malfunction or device destruction. They can be divided in two categories:

- Circuit breakthroughs, where the goal is to correctly define the operating point such as the voltage of one transistor terminal to another, does not reach a critical value. The objective here is to equally stress the transistor without impairing its reliability.
- Customized silicon technologies. Although more expensive to manufacture it is possible to have breakdown voltages high enough to be useful for the applications in question.

5.2.1 Circuit Breakthroughs

5.2.1.1 Cascode

The easiest possibility is to simply use the cascode stage (Fig. 5.1(a)). An example of this applied to a class E RF power amplifier can be found in [Yoo01]. The main disadvantage is that now the voltage swing on the gate-drain of the cascode transistor is larger than the common-source transistor. To make them equal, dynamic biasing of the drain voltage of the cascode transistor is necessary.

5.2.1.2 Self-Biased Cascode

As proposed by Sowlati [Sow03], the biasing is now implemented using $Rb - Cb$ with the increased advantage of not requiring any extra bond pad (Fig. 5.1(b)).

For class E, the positive voltage swing around the supply voltage is larger than the negative swing. In this case the circuit from Fig. 5.1(c) is more convenient, where $M1$ and $M2$ have the same voltage swing at the gate-drain. Thus, the hot-carrier effect

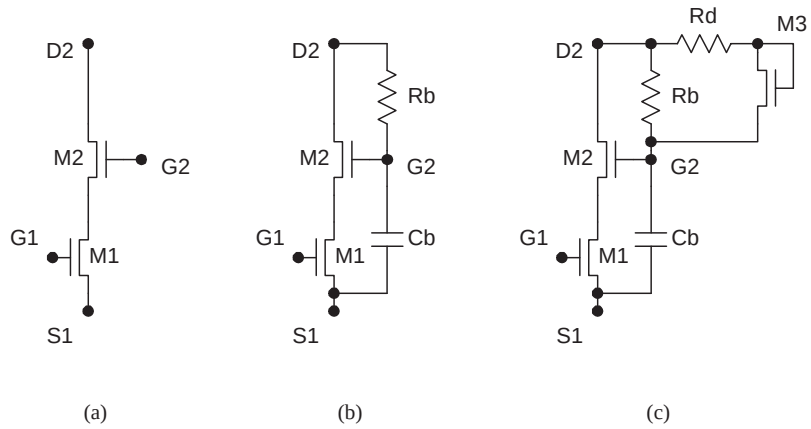


Figure 5.1: Various cascode configurations: (a) Conventional, (b) self-biased and (c) bootstrapped.

is relaxed. As a result, the amplifier works under maximum output power without showing performance degradation.

5.2.1.3 Non Zero Drain Voltage at Turn-On

The last example does not require any change in the circuit [Por93]. Instead, it properly selects another condition for operation. The amplifier electrical diagram is thus equal to the ideal class E amplifier in Fig. 4.3.

In contrast to [Sok75] that requires both the voltage across the transistor and the slope of the drain voltage to be zero to achieve high efficiency (5.2(a)), the amplifier is now designed so that the switch transition from the off to the on state has a substantial voltage step (Fig. 5.2(b)).

If sized as described, for the same output power it is possible to decrease the maximum switch voltage. Another advantage is that it has the effect of lowering the average current through the switch, thus lowering conduction losses for the same conduction angle. Experimental results in [Por93] have shown that the introduction of the voltage step improves efficiency.

A verification of the above claims can be done with the tool that is described in subsection 4.4.2. Two circuits (Fig. 4.12) are optimized for maximum DE, the only difference between them being that in the optimization constraint template, the transistor drain voltage in one does not have to be zero at turn-on. The other remaining two class E conditions must be met. As such, only the differences of not having zero voltage in the drain are seen. Because the optimization procedure is not deterministic, for assurance of the results, multiple simulations are done and the trends in the results collected.

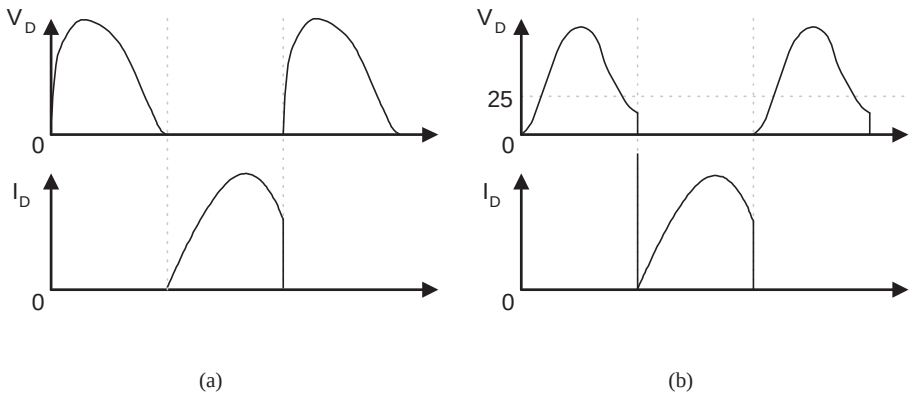


Figure 5.2: Voltage and current waveforms at the switch. (a) As commonly designed [Sok75, Raa78b] and (b) with non zero voltage at turn-on [Por93].

Also, for broadening the possible conclusions, inductors with different parasitics per unit length and different supply voltages are simulated. The conclusions are now given:

- The maximum drain voltage decreases up to a maximum of 25 % from the previous peak voltage.
- The drain efficiency can be 3 to 6 % better

This simple procedure does not alter the circuit schematic of the class E power amplifiers. Nevertheless it provides advantages, such as an increase in the drain efficiency while reducing the stress on the active device due to a lower drain voltage. It is worthwhile to stress that the above improvements are a function of circuit parasitics. In general, the lower the inductor quality factor, the greater the improvements, i.e., the drain efficiency increases and the reduction in the drain voltage will be more significant for inductors having higher losses per unit length.

5.2.2 Customized Silicon Technologies

5.2.2.1 Lightly Doped Drain: LDD

Lightly Doped Drain (LDD) [Tsi99], common amongst current CMOS technologies, is designed to lower the maximum electric field by having part of the depletion region inside the drain. This drain engineering method is commonly used to minimize hot-carrier effects of submicrometer devices and is achieved by reducing the doping gradient at the gate edge. This in turn lowers the electric field in the neighborhood of the drain. A two step process, where first a light to moderate implant forms the LDD region that is followed by a heavy implant after the spacer formation for making drain and source junctions and poly-doping for decreased sheet resistance (Fig. 5.3).

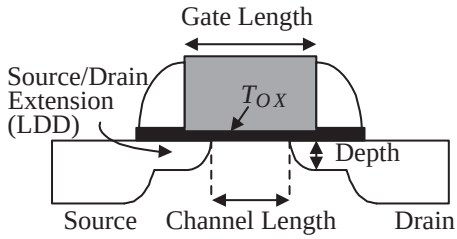


Figure 5.3: Device with lightly doped drain implant (LDD, n) surrounded by a high-doped region ($n+$) to implement the source and drain region.

Unfortunately, a lightly doped drain is also a lightly doped source. This extra resistance increases the channel resistance and also degrades the maximum frequency at which a circuit can operate. The latter is especially important for circuits designed for high speed. One possibility to minimize this effect and as such, to maximize current drive capability is to use an asymmetrical LDD. Yet, in a recent study [Che99], these devices show a shorter hot-carrier lifetime at a fixed supply voltage but with a higher I_{dsat}. Nevertheless, working at lower supply voltages, circuits with these transistors have an improved circuit speed and power consumption without sacrificing reliability [Che99]. Consequently, they are an affordable solution for increasing circuit performance.

5.2.2.2 Thick Oxide

The usage of a thicker oxide can be a method to increase the supply voltage without causing transistor disruption. The new device can be seen as a device from a previous generation implemented in an advanced technology. By having dual oxide transistors, the I/O circuits can use the higher voltage to connect to the outside world while high speed and low power can be obtained with the use of a lower supply voltage.

An example of such a device is given in Fig. 5.4. A similar structure to Fig. 5.3, it has a higher breakdown voltage but has nevertheless a lower cut-off frequency (f_T). For a switching type power amplifier operating in current mobile applications, the latter is usually not a limiting factor since the transistor works like a switch and does not have to provide current gain. On the other hand, the input capacitance has a larger value which will increase the total power necessary to drive the transistor.

5.2.2.3 Reduced Surface Field: RESURF

Invented by Appels and Vaes [App79] in 1979, it can be applied in different layers to equally distribute the applied voltage laterally across the silicon surface in the drift region of the device. The purpose is to move the breakdown from the surface to the substrate by depleting the layer where the breakdown is to occur. Symmetrical electric field distribution at surface occurs when the doping (N_{epi}) and the thickness (d_{epi}) follow: $N_{epi}.d_{epi} = 10^{12}$ at/cm^2.

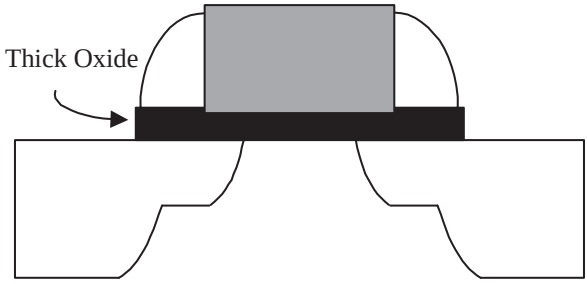


Figure 5.4: In dual gate oxide technology, for applications where a high breakdown voltage is desired, transistors featuring a thick oxide are used, whereas the other devices are used for low-voltage operation.

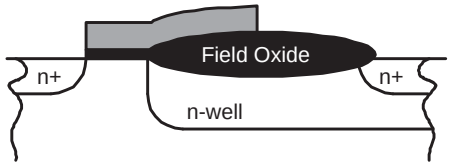


Figure 5.5: Cross section of the high-voltage transistor with the poly acting as field plate [Par87].

5.2.2.4 Fully Customized Technology: EZ-HV™

Taking the previous approach to its limit, extremely high voltages are possible [Phi99]. The possibility to fully control the number of layers, its doping and its thickness can drive the specifications beyond what is commonly expected.

It is a clever way to solve the problem of hot-electrons. With a thick layer of silicon, electrons can be accelerated to a point where they have enough energy to cause avalanche breakdown. A better way is to simply create a device where electrons cannot move far enough to acquire the necessary energy to generate current by impact ionization [Phi99]. In addition, latch-up immunity is a consequence of the oxide isolation used between devices on the chip. Breakdown values exceeding 600 V with an on-resistance of 7.6 Ω/mm^2 have been measured [Phi99].

5.2.2.5 Lateral Double-Diffused MOSFET: LDMOS

Perhaps the most common way to increase the breakdown voltage in current technologies, is to implement an LDMOS transistor [Par87, Bal98, Cas04]. The polysilicon overlaps the drift region, acting as a field plate to reduce the electric field near the gate edge [Gro67, Par87], thereby increasing the breakdown voltage (Fig. 5.5). If changing the process flow is an option, or the use of more complex and more expensive technologies is an option, a few other possibilities exist [Ehw01, Bak02].

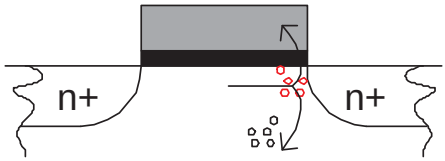


Figure 5.6: Electrons with high kinetic energy upon reaching the drain lattice can generate substrate current or destroy the thin oxide. They are the cause of severe reliability problems in semiconductor devices. The inclusion of LDD implants can reduce electrons speed and minimize their effect.

5.2.3 Reliability Concerns

To maintain performance over time, reliability challenges must be addressed [Tsi99, Gro01, Moe04]. Although the most common failure mechanisms are not of concern to the designer if foundry design rules are followed, it is nevertheless of interest to know how some of these mechanisms influence and indeed limit the designer's freedom.

5.2.3.1 Time Dependent Dielectric Breakdown: TDDB

The time (TDDB) is the time needed to break an oxide stressed with a high electric field. The oxide is stressed as a result of the high voltage between the gate and the channel underneath. Some of its characteristics are: (a) it has a strong dependence on temperature (exponentially), (b) strongly depends on the area (linearly), (c) on the duty-cycle, (d) on the applied voltage and (e) it is a statistic process.

5.2.3.2 Hot-Carrier Injection: HCI

As a result of the high electric field between the source and drain, the carriers are accelerated towards the drain until they become ballistic. When reaching the drain, they periodically destroy the drain high doped region. The inclusion of the LDD partially solves this issue. Furthermore, if the gate terminal has a high voltage value, the carriers can be deflected and literally thrown towards the thin oxide. In this case three things can happen: (a) if they have enough energy, oxide penetrations can happen. As a consequence variations in V_{TH} can occur. Moreover, this can happen slowly over days, weeks or even months, depending on the energy; (b) if they have sufficiently high energy, the oxide can be immediately and permanently destroyed, or (c) they can be deflected to the substrate and cause considerable substrate currents. More information on hot-carrier degradation can be found in [Gro01, Moe04].

In Fig. 5.6 a schematic representation of the hot-electrons effect is presented. Methods to provide adequate reliability in this regard include the control of the doping dose and profile of the drain and source under the gate edges and gate sidewalls. This is commonly done with an LDD implant.

5.2.3.3 Electromigration

With integrated circuits becoming progressively more complex, requiring more current and at the same time being manufactured in technologies with smaller features, means that metal electromigration is becoming an increasing form of circuit malfunction.

Electromigration is an effect caused by a large number of electrons colliding with metal ions, causing them to gradually drift with the electric current. Thermal energy produces scattering by causing atoms to vibrate. This is the source of resistance of metals. Semiconductors do not suffer from electromigration unless they are so heavily doped that they exhibit metallic conduction.

The increase of the temperature or an increase in local current density can decrease the useful lifetime of a circuit. The latter has an exponential relationship to temperature. Some of the causes can include: (a) an increased interconnection resistance, (b) a short-circuit with neighboring metal or (c) creating an open-circuit.

5.2.3.4 Junction Breakdown

Junction breakdown is a nondestructive effect as long as no large current flows in the junction and overheats it. By adhering to foundry design rules this phenomenon can be disregarded.

5.2.4 Short Channel Effects

In general short channel effects are not destructive but mainly alter transistor characteristics over time. A good source of information is the book by Tsividis [Tsi99].

Common technologies suffer mainly from:

- Drain Induced Barrier Lowering (DIBL)
- Punchthrough
- Subthreshold Current
- Hot-Carrier

5.2.4.1 Drain Induced Barrier Lowering: DIBL

For devices with long channel lengths, depletion of the channel is solely achieved by the gate. But in a short channel transistor, drain and source voltage also influence this depletion. The consequence is that less gate voltage is now required to create a low resistivity path for the electron flow, hence decreasing the barrier for electron injection into the channel (Fig. 5.7). Furthermore, this channel barrier is reduced when the drain voltage increases (Fig. 5.7). This effect is known as Drain Induced Barrier Lowering (DIBL).

This dependence on the drain voltage is one of the problems in CMOS scaling. There are implications in the carriers' velocity and also in the efficiency of carriers injection from the source to the channel. This effect can be reduced by reducing the size of the

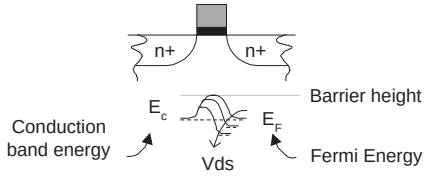
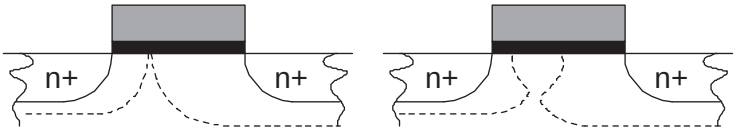


Figure 5.7: Effect of DIBL on threshold voltage.



(a) Surface Punchthrough.

(b) Bulk Punchthrough.

Figure 5.8: Punchthrough occurs when the depletion region from the drain reach the source side and reduces the barrier for electron injection. It can either occur at the surface (a) or within the bulk (b).

depletion region for a given voltage which can be achieved by proper control of the doping in the channel. A retrograde doping profile is a relatively simple solution where the increase in the doping has the effect of decreasing the size of the depletion region.

5.2.4.2 Punchthrough

Punchthrough occurs in CMOS transistors when at high drain-source voltage, the drain space charge region expands over the channel width and comes in contact with the source depletion region. This can occur either at the surface (Fig. 5.8(a)) or within the bulk (Fig. 5.8(a)). In this situation a large current can flow causing device failure. In short channel devices the maximum allowed drain voltage is usually determined by punchthrough. Again, the reduction of the depletion regions can be obtained by means of a retrograde doping profile in the channel that selectively increases the vertical doping which decreases the depletion region, consequently decreasing the voltage at which punchthrough would occur.

5.2.4.3 Subthreshold and Leakage Current

Even with a gate-source voltage equal to 0 V, uncontrolled current still flows in the transistor. It can either be leakage if due to current flowing across reverse-biased junctions or subthreshold current when in the transistor a very small current flows for V_{GS} lower than V_{TH}.

The parasitic-diode leakage current shows strong dependence on the type and quality of the process as well as temperature, whereas the subthreshold current has the characteristic that it does not scale with process technology. The presence of leakage current and subthreshold current detracts from the ideal switch model, causing static power dissipation.

More information can be obtained from the subthreshold current. The slope of the transistor current in the subthreshold region is a measure of how good a device turns off. A value of 100 mV/dec is a good value for the subthreshold slope. The minimum value of 60 mV/dec is usually accepted and CMOS technologies usually have a value around 80 mV/dec.

5.3 STI/LOCOS Compatible LDMOS Structure in Standard CMOS

The possibility of achieving high breakdown voltages on current advanced CMOS processes has increased importance as the maximum operating supply voltage decreases. They are important for I/O circuits [Ann00] where the operating voltage is not reduced, for handling the high programming voltage for an Electrically Erasable Programmable Read Only Memory (EEPROM) in a CMOS technology [tE97] and for RF LDMOS transistors to be used in integrated power amplifiers for mobile communications. Furthermore, in current technologies the preferred method for device isolation is to use a silicon trench around the transistor, to the effect that it is no longer possible to form a channel under the gate. Consequently, CMOS compatible high-voltage transistors using extended drain and designed for Local Oxidation of Silicon (LOCOS) technologies [Par87] cannot be integrated with Shallow Trench Isolation (STI).

This section starts by giving an overview of the two most common methods of device isolation in CMOS. Afterwards, a novel high-voltage transistor structure compatible with STI and LOCOS is described. This device, which can be implemented in a standard CMOS process, is capable of handling high voltages without destruction. And finally, a new structure which features lower input capacitance is proposed based on simulation results.

5.3.1 Device Isolation Techniques

The main advantages of integrated circuits over discreet components result from the compaction provided by current technologies that allow multiple devices to lie in the same substrate. As a consequence, interaction from one device with another is possible which in turns change device characteristics. This can either be in the form of parasitic conduction from one device to another or latch-up. The latter occurs in CMOS, when under certain bias conditions a thyristor like device is formed between a transistor and substrate contact junctions. This may cause the circuit to malfunction or self-destruct. Device isolation is used to avoid or ease these effects.

The main characteristics of two of the most widely used device isolation methods are briefly summarized below: LOCOS and STI.

(a) In LOCOS isolation, thick oxide is used so that wiring above it cannot easily invert the channel underneath. For the same reason, the p_+ doping implant is used to increase the threshold voltage of the parasitic transistor.

(b) A trench is etched into the semiconductor substrate and filled with oxide. Electrical isolation is achieved by such a trench within adjacent devices.

Figure 5.9: LOCOS vs. STI comparison.

5.3.1.1 Local Oxidation of Silicon: LOCOS

Local Oxidation of Silicon (LOCOS) is the main method of isolation used in technologies of 0.35 µm and above. As a result of its simplicity a low cost is possible. However, a disadvantage resides in the lateral oxidation of the silicon during the thick oxide growth, making the edge of the field oxide resemble the shape of a bird's beak (Fig. 5.9(a)). Other drawbacks include field implant encroachment and a non-planar surface, creating depth of focus problems during subsequent lithographic patterning of the polysilicon layer.

With decreasing feature size the requirements on allowed isolation area become stringent and LOCOS consume large amounts of silicon space because oxidation expands the isolation region laterally in proportion to its depth (Fig. 5.9(a)), thus becoming a critical problem at 0.35 µm design geometries. To achieve a better dimensional control and packing density, a new method is used that allows more functionality and speed per unit area.

5.3.1.2 Shallow Trench Isolation: STI

From a processing point of view, STI (Fig. 5.9(b)) is more complicated. Nevertheless, when compared to LOCOS, a STI structure relatively reduces lateral encroachment, offers better trench depth and width control, low junction capacitance, and greater packing density. Furthermore, STI is also more adequate to prevent punchthrough and latch-up. These characteristics have made STI the preferred isolation scheme for CMOS technologies at or below the 0.25 µm technology design node and has virtually replaced LOCOS as the method for isolating active areas in semiconductor devices.

The process can be seen as etching a trench in the silicon substrate that is later filled with a thick oxide. The next step consists of planarization, using Chemical Mechanical Planarization (CMP). In this way, near zero field encroachment is possible. Moreover, the sidewalls are now nearly vertical, this angle being limited by the technological limit of the oxide capability used in the trench. Different foundries provide different, al-

though similar structures to the one presented in Fig. 5.9(b). The trench depth depends on the etch time. As such, circuits with higher latch-up immunity and better junction characteristics can be obtained with a deeper trench. Nevertheless, the maximum depth is limited by the trench fill capability of the oxide used. It is important to note that STI structures strongly depend on CMP. In fact, they could not be manufactured without it and are an object of research.

As a consequence of isolation trench width being defined by the lithography step, STI can be scaled with each technology generation. However, STI stress has impact on the junction characteristics: diode junction leakage and junction capacitance. This effect also influences MOS electrical characteristics. Nevertheless, it is possible to include these effects in standard compact models without a considerable effort.

5.3.2 High-Voltage Device Structure

STI means that the new transistor has to be completely planar. Furthermore, the lack of thick oxide means that the field plate technique [Gro67, Par87] cannot be used to reduce the electric field intensity near the gate edge. A cross section of the proposed LDMOS structure is shown in Fig. 5.10. The entire region between the drain and source terminal has to be defined as an active area, otherwise an oxide trench is created within the drift region. This is in contrast to the procedure in [Par87]. The n^+ implant on the right side of the self-aligned gate is a result that is necessary to dope the gate in order to reduce its resistance. The drain is taken apart to decrease the electric field within the thin oxide connecting to the gate terminal. Finally, the poly over the *n*-well is used to create a low doping drift region by blocking the n^+ implant over all the *n*-wells.

The device uses a few different methods to achieve a higher working voltage. To decrease the electric field near the gate edge on the high doping drain side, a low doping layer is used. In addition this layer also serves the purpose of decreasing the speed with which the electrons reach the drain lattice. One advantage is that the drift region is not contaminated by the thick oxide growth [Kim99]. Also, the drift region can now be made smaller. Both characteristics lead to a lower on-resistance than in [Par87].

5.3.2.1 Improving Breakdown Voltage

Extending the *n*-well towards the source junction decreases the speed of the electrons reaching the high doped region at the gate's right side (Fig. 5.11). When comparing with the device presented previously, the new transistor has two advantages: increase of the breakdown voltage and reduction of the on-resistance. Another factor is that now current can flow in the device in the saturation/strong inversion region almost up to the breakdown voltage (Fig. 5.20(a)).

5.4 Experimental Results

Some of the principles for increasing the supply voltage have been discussed. Their implication in transistor reliability have been addressed alongside its advantages and disadvantages. Based on this, a LDMOS transistor is studied, layed-out and manufactured in a commercial 0.35 µm CMOS technology without change to the process flow.

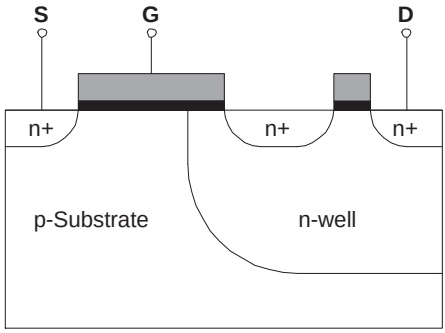


Figure 5.10: Cross sectional representation of the STI/LOCOS compatible high-voltage structure in standard CMOS, LDMOS M_{5_03}.

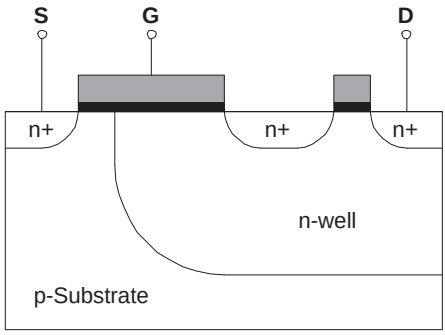


Figure 5.11: Starting with the structure from Fig. 5.10, an improvement in the breakdown voltage and a reduction in the on-resistance is obtained by extending the *n*-well towards the source junction. This creates the above M_{A_01} transistors.

Its characterization is now presented. Firstly, the characterization is done at DC. The on-resistance and the breakdown voltage is in this way extracted. Afterwards, an AC characterization is done for extraction of the device capacitances. With this, a simple transistor model suitable for SPICE simulation is built.

Different variations in the design of each transistor are introduced for assertion during measurements. Also, each die included different structures that serve the purpose of technology characterization. Moreover, this makes possible to compare the device given in Section 5.3 with the foundry transistors.

5.4.1 DC Experimental Results

High-voltage n-channel MOSFET (NMOS) transistors with different channel and drift lengths are fabricated in a standard 0.35 µm CMOS technology. To avoid the edge

effects of high-voltage breakdown transistors are layed-out with a circular geometry. Its effectiveness, however, will later be commented on based upon measurement results from different dies (but within the same wafer) and different manufacturing runs.

As much of the information concerning doping levels, doping profiles and thicknesses for this technology are considered a secret, the study is done using typical values freely available on the Internet. Typical technology parameters are then simulated using the Medici™ device simulator. The simulations are used to explore possible solutions, all dependent on the real parameters for this technology, which are unknown. Measurements on the manufactured device are used for certification of device functionality.

Different solutions have been investigated. Fig. 5.13(b) shows the die containing nine large families of test devices with each one of them containing a different arrangement of the masks to control the electric field and the doping. Each one of the families is later subdivided horizontally and vertically. Now, only two layers are changed. As a consequence, during measurements it is possible to have device characteristics close to its optimum value. The complete list of the text structures are included. Each device will as such be referred to as M_{α_β}, where α is the family (1 until 9) and β is its subdivision within each family (01 until 25). The floor-plan of the test chip from run723 is shown in Fig 5.13(a). Each of the nine rectangles represents each one of the families whereas the zoomed part represents one family with each one of the 25 individual transistors. A microphotograph is shown in Fig. 5.13(b).

- Family 1 is used for technology characterization. It includes the foundry transistors $M_{1\_01,06,11,16,21}$ (Fig. 5.3), circular shaped transistors ($M_{1\_02,07,12,17,22}$), square shaped transistors ($M_{1\_03,08,13,18,23}$) and n-well implants layed-out in multi-side polygons with and without METAL1 acting as field plate. All transistors within the same row have a different gate length.

- Family 2 and 3. As in the conventional LDMOS in Fig. 5.5, a n-well is used as a lightly doped drift region for the high-voltage NMOS. A different n-well to channel overlap is used for each family.

- Family 4 to 9 contain different approaches that have been tried to achieve a higher breakdown voltage than the foundry transistor, while being able to be fully compatible with STI and LOCOS technologies. The best results are obtained when using a low doping layer to decrease the electric field near the gate edge on the high doping drain side and at the same time to decrease the speed of the electrons reaching the drain lattice (Fig. 5.10).

As another run within a reduced time frame was available (Fig. 5.14), it was decided to use another test chip in order to further fine-tune the results already measured. This also allows to compare the measurements from run to run and not only from different dies from the same wafer and run. As much of the test structure study is already attained, only a minimum of variations in the design structures are made. The floor-plan of run815 is shown in Fig. 5.14(a) and the chip microphotograph is presented in Fig. 5.14(b). The test chip is also organized similarly to run723 where each of the rectangles representing one of the seven families, is divided into 25 unique devices.

- Family 1, 5 and 3 are used for comparison with the previous run. Measured results have not shown significant variations from run to run.
- Family A and C, intended for comparison with family 5 include slight variations in the *n*-well implant to channel overlap. Measured results have shown an on-resistance reduction and breakdown voltage increase with the structure depicted in Fig. 5.11.
- Family B and D, likewise, are included for gaining further insight to family 2 and 3. These new devices have a different *n*-well to channel overlap.

Once the device has been chosen based upon its DC operation, a new test chip is designed. Now the objective is to measure the RF performance. To minimize parasitics influence during measurement, the new transistor has a total gate width of 5.6 mm.

- LDMOS2 transistor is built upon 180 M_{A_01} parallel transistors (Fig. 5.15).

The measurement setup used to characterize the transistors performance is described.

5.4.1.1 Measurement Setup

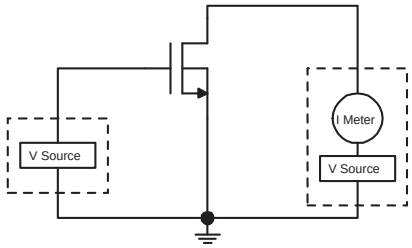
The complete measurement setup used to evaluate the DC performance of the transistor is presented in Fig. 5.12. Each die is glued to a thick film ceramic substrate for chip manipulation easiness. All four terminals (including substrate) are connected to individual source-monitor unit plug-ins (SMU) for full characterization. Each one of the SMU is connected to each one of the transistors in the die by means of a thin analytical probe needle.

Each one of the transistors is measured under the following conditions:

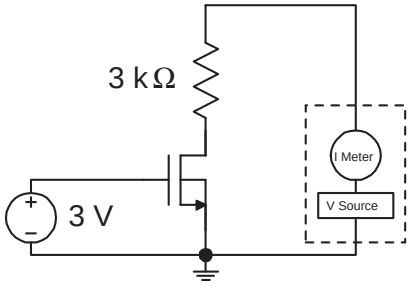
- The minimum and maximum gate-source voltage is -0.5 and 3 V, respectively.
- The drain-source voltage is swept from 0 V until 3 V or to transistor breakdown.
- Both the source and the bulk terminals are grounded.

For each V_{GS}, the drain voltage is swept from its minimum to its maximum. As such, it is possible that during this procedure the transistor is destroyed due to the high speed electrons reaching the drain, although the gate-source voltage is still within the manufacturer safety region. Another reason for destruction is the high voltage between the gate and the drain (Fig. 5.16(e)). In general, one phenomenon can induce others, the latter being the cause of the destruction. Therefore, it is difficult to identify the exact cause of malfunction as it can be the result of a simultaneous sequence of destructive events. Long term reliability degradation is not studied and is an important issue [Gro01]. The reduced number of devices available from the foundry are not enough to obtain results that are supported by the statistical analyses of the measurements.

The transistor is said to be in breakdown if the measured drain current is above 1 µA. For the measurement of a source or drain junction, the same value of 1 µA is used.



(a) Instrument setup to measure NMOS transistor I-V curves.



(b) Arrangement to measure transistor on-resistance.

Figure 5.12: Schematic diagram of the experimental arrangement used for DC characterization.

From a measurement perspective, non-destructive transistor malfunction can be due to punchthrough or junction breakdown. For minimum transistor length, punchthrough is usually the cause for breakdown.

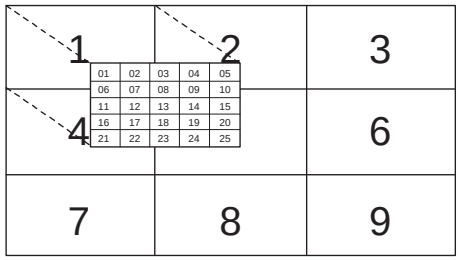
The DC experimental results obtained with the fabricated test circuits are now given.

5.4.1.2 Foundry Transistors

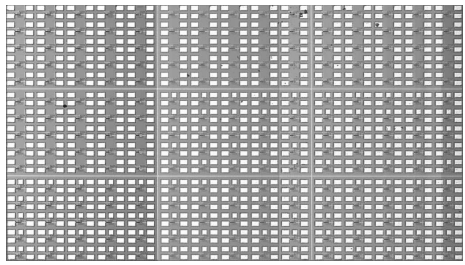
Transistors following foundry design rules are measured for characterization purposed and for effective measurement of the improvements with the new structures. In total, five different devices are layed-out having a gate width of 15 µm. Device M_{1_01} (also referred to as Low Voltage MOS (LVMOS)) has the minimum gate length possible in the technology: 0.35 µm. This length is duplicated for each of the other devices until a total gate length of 5.6 µm is reached for transistor M_{1_21}. Transistors $M_{1\_01,06,11,16,21}$ are constituents of the foundry devices.

The output, input and breakdown characteristic for one of these devices are shown in Fig. 5.16(a), 5.16(b) and 5.16(c). The subthreshold characteristic is given in Fig. 5.16(d). An analysis shows that the channel current is effectively controlled by the gate. This is the result of the small shift between both curves and the flattening only for values lower than V_{TH}. The low leakage current even for high drain voltages is an indication that the device M_{1_01} is far from punchthrough.

The breakdown voltage is situated around 9 V which is close to the junction breakdown voltage part of the transistor. Note that device M_{1_01}, from die to die has a measured breakdown voltage within the range of 6 and 9 V. This characteristic is not observed with any of the other devices, suggesting that it is specific for the minimum gate length transistor.



(a) Floor-plan of the test chip.



(b) Microphotograph.

Figure 5.13: Designed test chip from run723 including nine large groups subdivided in 25 individual elements. Each of the 225 individual devices is numbered between M_{1_01}–M_{9_25}.

If we now increase the gate voltage while not limiting the drain voltage to 3 V as is done in Fig. 5.16(a), the device can be tested on the Hot Carrier Injection (HCI) reliability.

- Transistor M_{1_01} is destroyed as soon as current flows in the channel. Although the V_{GS} is smaller than the V_{TH}, the high V_{GD} is the reason for its destruction. To better support this claim, a new measurement is done in another sample, but now with V_{GS}=3 V. As show in Fig. 5.16(e), the transistor stops to work properly with a drain-source voltage of approximately 6 V but normal operation can be recovered once this voltage is again decreased. The measurement instrument limits the maximum current that can flow in the drain and thus avoids destruction due to overheating. Fig. 5.16(e) shows that the device works properly at least twice.

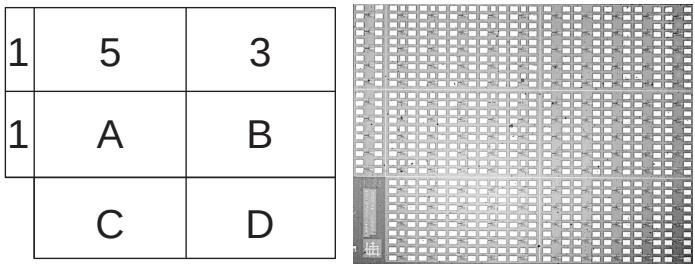
- Transistors $M_{1\_06,11}$ are destroyed when drain-source voltage, with $V_{GS} < V_{TH}$, reaches 9 V.

- Transistors $M_{1\_16,21}$ are measured to be always operational. They cease to operate due to breakdown of source-bulk junction. Nevertheless, if the junction is allowed to breakdown, they are useful until 13 V when $V_{GS} \approx V_{TH}$.

In general, no variations between values measured from one run to another run are observed, suggesting that the process used has indeed reach a stable point.

Comparison with the SPICE Model from the Foundry

Comparison will be carried out only for device M_{1_01} (LVMOS). The results from the foundry's BSIM3 model superimposed on experimental results are given in Fig. 5.17. The measured results have a high resemblance to the simulated waveforms. This is especially true for the case of Fig. 5.17(a) and 5.17(b). The leakage current looks



(a) Run815 floor-plan. (b) Microphotograph.

Figure 5.14: Test chip designed in run815. It includes three families equal to run723 included for run to run comparison: M_{1_xx}, M_{5_xx} and M_{3_xx}. Also included are four new groups including 100 unique transistors numbered between M_{A_01}–M_{D_25}.

like a shifted waveform (Fig. 5.17(c)). Again, this effect is once more seen in the subthreshold wave-shape which is accurate only for values above V_{TH} (Fig. 5.17(d)). A reason to justify the previous observations is the 4 years old model file. During the production time of a technology, as the process stabilizes, it is expected that the final parameters will drift. As such, these variations are somehow expected. This model is today not only inappropriate for high performance analog circuits but also for the digital designer, where an accurate value of the leakage current is a prerequisite. But sometimes the limitation is imposed by the mathematical model describing transistor operation. In this case, the BSIM3 model does not include the effect of breakdown as seen in Fig. 5.17(e). Once more, this is not of designers' concern as most of the designs are intended to be at 3.6 V and below.

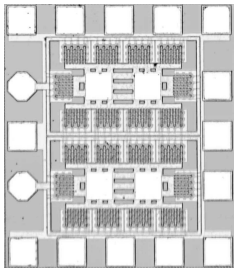
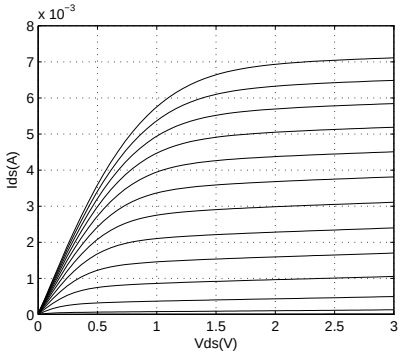
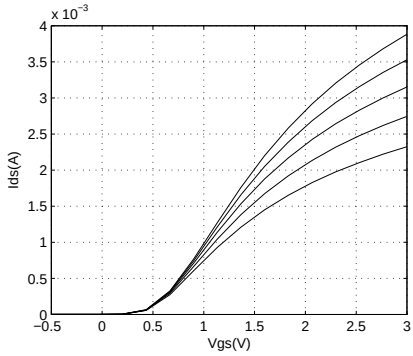


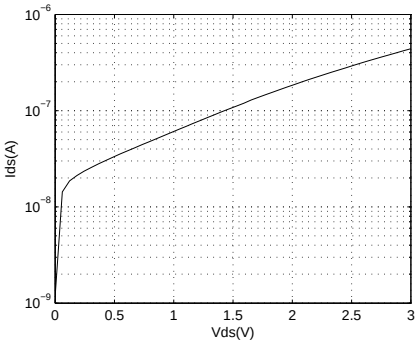
Figure 5.15: Microphotograph from run876. Transistor LDMOS2 is composed of 180 M_{A_01} parallel transistor. Its large size (5.6 mm) minimizes the effect of parasitics during AC characterization.



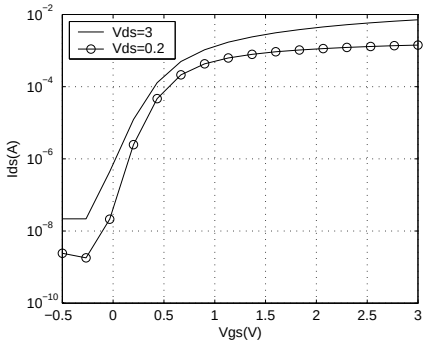
(a) Measured current-voltage characteristics of transistor M_{1_01} with -0.5 V< V_{GS} <3.0 V.



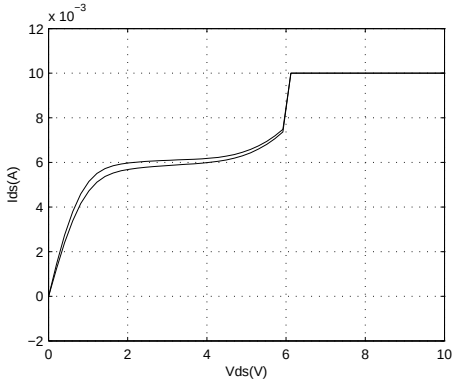
(b) Input characteristic measured for drain-source voltage within 0.3 V< V_{DS} <0.6 V.



(c) Measured transistor current for the off-state (V_{GS} =0 V).



(d) Measured subthreshold characteristics showing the expected behavior.



(e) Two consecutive measurements from the same transistor in another sample with V_{GS}=3 V. The drain current is limited to avoid destruction due to overheating.

Figure 5.16: Set of DC measurements for transistor M_{1_01}.

Table 5.1: Comparison between measured and simulated on-resistance for transistors with W=15 µm operating with V_{GS}=3 V and V_{DS}=0.1 V.

	M_{1_01} L=0.35 µm	M_{1_06} L=0.7 µm	M_{1_11} L=1.4 µm	M_{1_16} L=2.8 µm	M_{1_21} L=5.6 µm
run723	146 Ω	214 Ω	363 Ω	644 Ω	1207 Ω
run815	145 Ω	216 Ω	373 Ω	662 Ω	1234 Ω
SPICE	122 Ω	199 Ω	346 Ω	637 Ω	1220 Ω

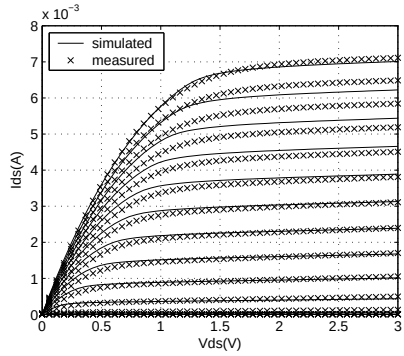
In Table 5.1, a comparison between measured values for the on-resistance for different runs and SPICE is presented for all types of foundry transistors tested. While the error between runs is kept below 3 % for all cases, the model has an error for small length transistors. This error can be as high as 18 %. One conclusion that can be drawn from the results, is that the error between measurements and the SPICE model decreases with an increasing transistor length. Nevertheless, for high-speed applications minimum transistors sizes are often used. As such, deviations from simulation can occur if the designer strongly relies on an accurate value of the on-resistance value.

5.4.1.3 Multi-Side Polygon Junctions and Transistors.

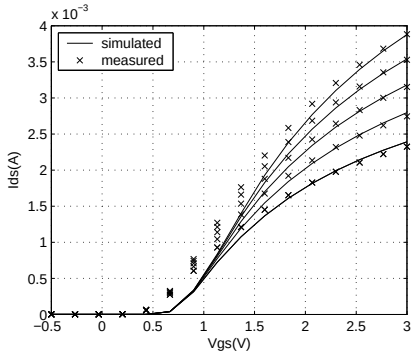
The maximum breakdown possible achieved with this technology without using any new art is in the *n*-well to substrate junction. An average value of 30 V is measured. With the METAL1 field plate voltage between ±15 V, it is possible to change this value by ±1 V. This low control is the result of the considerable distance between the first metalization layer and the *n*-well, and as such only an ineffective operation as field plate is achieved. The other important junction is the source or drain implant. This junction is operational until it has 10 V at its terminals.

Square and circular transistors with the same gate length than the foundry transistors are measured for comparison and to allow conclusions on the geometry influence in the breakdown voltage. With the same purpose, circular, square, triangular and multi-side polygon are layed-out. The breakdown voltage that is measured for the transistors does not greatly differ from those that are measured for the foundry transistor (LVMOS). Nevertheless, the variation in the breakdown voltage that is measured (between 6 and 9 V), is not observed. This suggests that the variation might be due to the geometry.

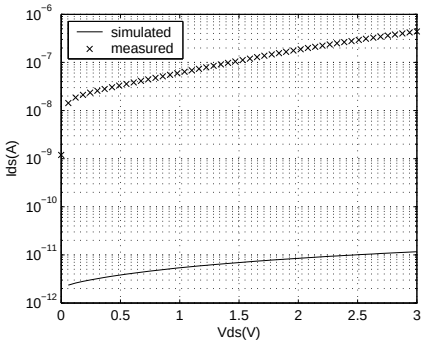
From all measurements, results do not support the assertion that edges must be avoided to achieve a higher breakdown voltage. The same implant but with different layout does not appear to break first if it has sharp edges. These are valid for the technology studied and might not be necessarily true for technologies where a higher breakdown voltage is desirable, or in other words, that the edge induced breakdown is weak in this 0.35 µm CMOS technology. The above conclusions are based on multiple samples from the same run only.



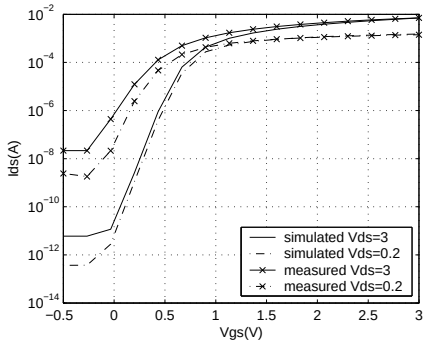
(a) Measured and simulated output characteristic with -0.5 V< V_{GS} <3.0 V.



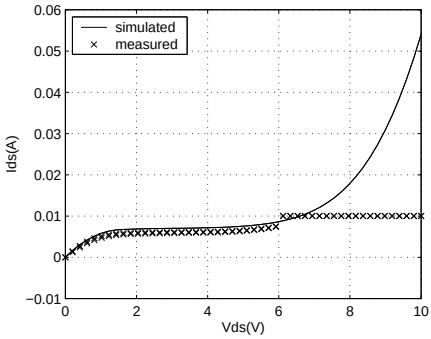
(b) Drain current as a function of gate-source voltage with 0.3 V< V_{DS} <0.6 V.



(c) Off current as given by measurement and simulation with V_{GS} =0 V.



(d) Subthreshold characteristics.



(e) Comparison between measurements and simulations in the breakdown region when V_{GS}=3 V.

Figure 5.17: Comparison between measured and simulated waveforms for transistor M_{1_01}.

5.4.1.4 LDMOS M_{2_02}

The high-voltage NMOS transistor from Fig. 5.5 is designed in a standard 0.35 µm technology. Current-voltage characteristics are shown in Fig. 5.18. The high breakdown voltage measured (27 V) is close to the breakdown voltage of the *n*-well implant. Referring to Fig. 5.18(a) and 5.18(b), the observed behavior is most likely due to the drain series resistance in *n*-well to channel interface resistance as a result of the improper blocking of p^+ channel stop implant. On the contrary, the threshold voltage is close to the value for the foundry transistor. As seen in Fig. 5.18(d), the channel current is in effect controlled by the gate terminal, with result that the flattening only happens for gate voltages lower than the threshold voltage. This device exhibits the highest measured breakdown voltage (Fig. 5.18(c) with a value of 27 V. A summary of its performance is given in Table 5.2. The two semicolon separated values represent the gate length and the oxide thickness over the thin and thick oxide.

5.4.1.5 LDMOS M_{5_03}

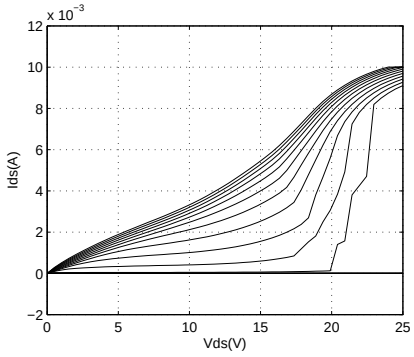
The device is designed and manufactured in a standard 0.35 µm technology without any change in the process flow or with any layer not available to all designers. A summary of the measured electrical parameters is given in Table 5.2. A duplication in the breakdown voltage has been measured for the LDMOS M_{5_03} (Fig. 5.10) in comparison with the typical transistor (LVMOS). The measured current-voltage characteristics are presented in Fig. 5.19(a) and Fig. 5.19(b). Referring to the subthreshold curves in Fig. 5.19(d), the small shift between both curves and the flattening only for values lower than the V_{TH} is evidence that the channel current is effectively controlled by the gate. Also, no significant leakage current up to breakdown is measured (Fig. 5.19(c)), an indication that the LDMOS M_{5_03} is far from punchthrough. Even with this uncommon structure, measurement results from run to run have shown good agreement, further supporting that this process has stabilized.

5.4.1.6 LDMOS M_{A_01}

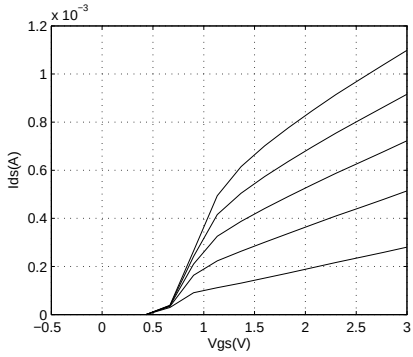
The output, input and breakdown characteristics for one of these devices are shown in Fig. 5.20(a), 5.20(b) and 5.20(c). The subthreshold characteristic is given in Fig. 5.20(d). An analysis shows that the channel current is effectively controlled by the gate. This is the result of the small shift of both curves and the flattening only for values lower than V_{TH}. The low leakage current even for high drain voltages is an indication that the device M_{A_01} is far from punchthrough. Table 5.2 lists a summary of its characteristics from where it is possible to see that this device simultaneously achieves a higher breakdown voltage while featuring a lower on-resistance than device M_{5_03}.

5.4.1.7 LDMOS2

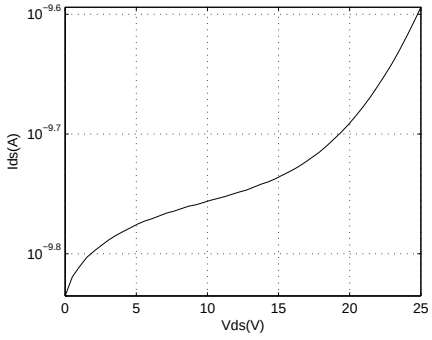
Using a commercial 0.35 µm CMOS technology, the proposed device is designed, manufactured and tested. In Table 5.2 the main electrical parameters are summarized. The output characteristic is shown in Fig. 5.21. Measurements, reproducible from sample to sample within the same run, are done without external cooling. As a consequence, in Fig. 5.21 a self-heating phenomenon is visible. Power dissipation per unit area is



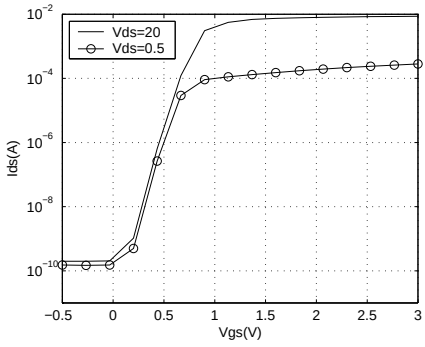
(a) Measured current-voltage characteristics of LDMOS transistor M_{2_02} with -0.5 V< V_{GS} <3.0 V. The transistor current appears not to saturate. This resulted from the improper blocking of the field implant in the drift region.



(b) Input characteristic measured for drain-source voltage within 0.5 V< V_{DS} <2.5 V.

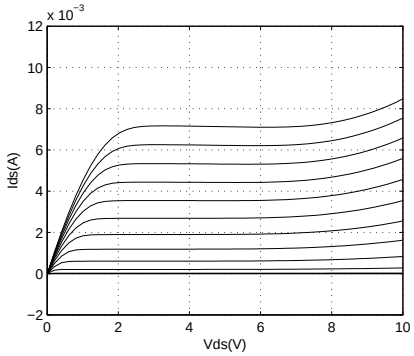


(c) Measured transistor current for the off-state (V_{GS} =0 V).

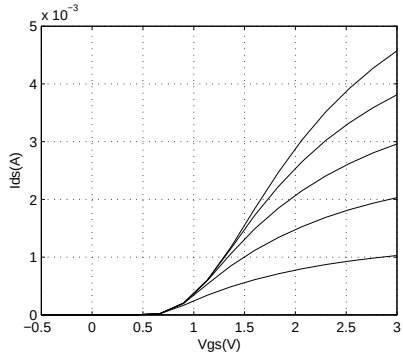


(d) Measured subthreshold characteristics.

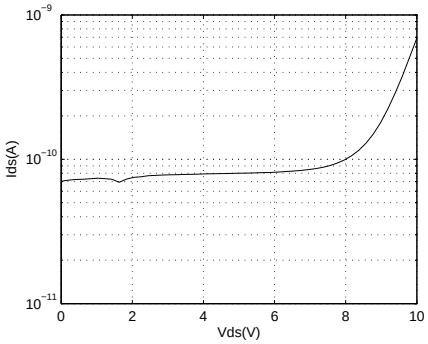
Figure 5.18: Set of DC measurements for transistor M_{2_02} .



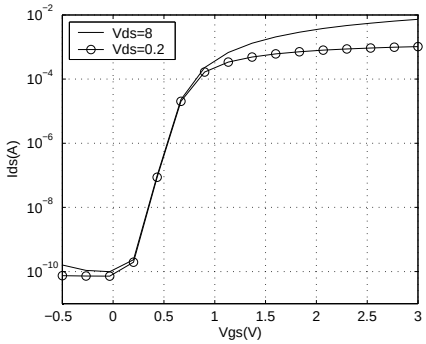
(a) Measured current-voltage characteristics of LDMOS transistor M_{5_03} with -0.5 V< V_{GS} <3.0 V.



(b) Input characteristic measured for drain-source voltage within 0.2 V< V_{DS} <1.0 V.

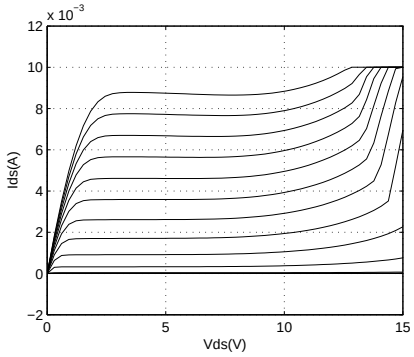


(c) Measured transistor breakdown characteristic (V_{GS} =0 V).

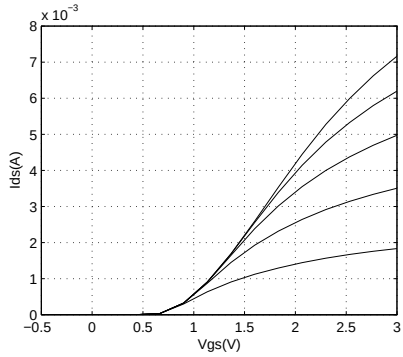


(d) Measured subthreshold characteristics.

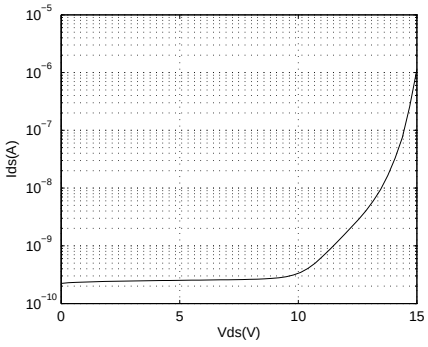
Figure 5.19: Set of DC measurements for transistor M_{5_03}.



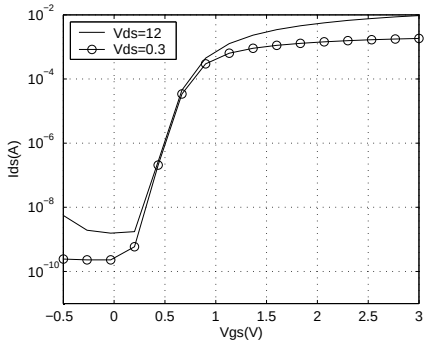
(a) Output characteristic of the transistor M_{A_01} with -0.5 V< V_{GS} <3.0 V. Current can flow in the device almost up to the breakdown voltage.



(b) Input characteristic measured for drain-source voltage within 0.2 V< V_{DS} <1.0 V.



(c) Breakdown characteristics of transistor M_{A_01} (V_{GS} =0 V).



(d) Measured subthreshold characteristics showing no significant leakage current. Device M_{A_01} has a subthreshold slope of 100 mV/dec@Vds=12 V.

Figure 5.20: Set of DC measurements for transistor M_{A_01}.

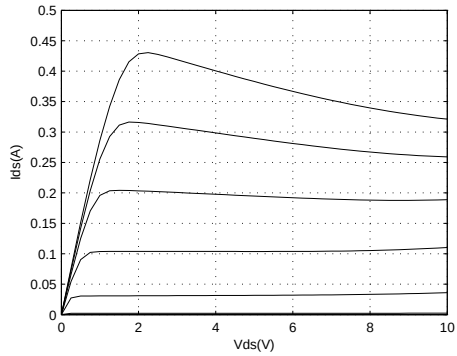


Figure 5.21: LDMOS2 output characteristics with gate voltage from 0 V to 1.6 V. This device is made of 180 M_{A_01} unity transistors giving a total length of 5.6 mm. As a result of the high power dissipation in the device, self-heating phenomenon is observed in the saturation region.

Table 5.2: Measured Electrical Parameters of Various Devices

Parameter	Unit	LVMOS	M_{2_02}	M_{5_03}	M_{A_01}
W	µm	15	24	27.5	29.7
L_{DRIFT}	µm	—	1.75	0.35	0.35
L_{GATE}	µm	0.35	0.7; 1.4	1.4	1.4
T_{OX}	nm	7.4	7.4; 380	7.4	7.4
V_{BR}	V	7.5	27	14	14.7
R_{ON}	Ω/mm	2.2	19.4	5	4.8

greater than in the most advanced microprocessors. The transistor named LDMOS2 is composed of 180 parallel M_{A_01} devices.

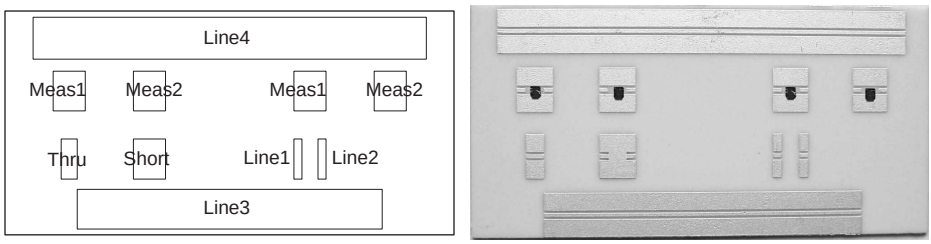
5.4.2 AC Experimental Results

The device from Fig. 5.15 is specifically designed to ease RF measurements. For this reason, the transistor has a total gate width of 5.6 mm, large enough to minimize parasitic influence.

5.4.2.1 Measurement Setup

The transistor is attached to a ceramic substrate by means of a non-conductive glue. Some characteristics of the used aluminum oxide ceramic substrate are given in Table 5.3. For board parasitic de-embedding, several structures are included. As shown in Fig. 5.22, they are:

- *Line1-4*, the *Thru* and the *Short* are used for parasitic de-embed. Furthermore, coplanar lines *Line3* and *Line4* allow the characterization of the substrate propa-



(a) (b)

Figure 5.22: Photograph of the dies on a ceramic substrate used for device characterization.

gation characteristics.

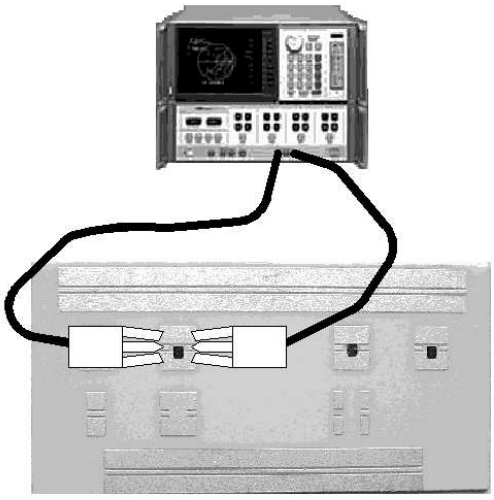
- In *Meas1*, also *open-ground connected*, the transistor is mounted for proper measurement by an RF probe with ground-signal-ground (GSG) configuration (Fig. 5.23(b)).
- For the bonding wire de-embedding, structure *Meas2* is used (Fig. 5.23(c)). A similar arrangement to *Meas1* with bonding wires having similar sizes but where the on-die ground plane short-circuits both. By subtracting the ground plane effect, the bonding wire model is obtained.

The completed measurement setup used to characterize the performance of the transistor LDMOS2 is presented in Fig. 5.23(a). A Cascade Microtech 9000 analytical probe station is used along with Picoprobe 40A GSG 800 µum pitch probes and Picoprobe calibration substrate CS-10 for instrument calibration. The active device performance is measured by use of a vector network analyzer (Agilent 8510C). The RF probes are placed near the end of the coplanar lines which are connected to the circuit via the bonding wires.

Since the bare device had to be mounted in a test fixture and connected by means of bonding wires in order to do the measurements (Fig. 5.22), special care has been taken to include all the parasitics presented in these measurements in the final model. The resultant model is the one indicated in Fig. 5.26. Additionally to the typical intrinsic and extrinsic elements of a MOS device, the silicon substrate losses have also been included in this model to get a higher accuracy.

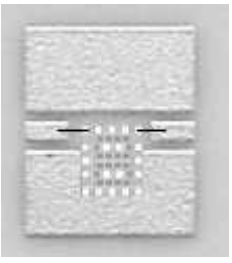
5.4.2.2 LDMOS2: Measurements and Characterization of the Device

The procedure followed to extract the extrinsic and intrinsic elements of the device consisted of a first calculation of each of these component values and then a gradient optimization process in order to tune the values such that the *s* simulated output fitted



(a) Schematic diagram of the experimental arrangement used for AC measurements.

(b) Section of the arrangement for measurement of chip mounted on the substrate for device characterization.



(c) Diagram of the experimental arrangement used for bonding wire de-embedding.

Figure 5.23: Setup used in the AC measurements of the transistor LDMOS2.

Table 5.3: Aluminum oxide (Al_2O_3) ceramic substrate properties.

Parameter	Unit	Value
Purity	%	99.6
Thickness	µm	600
Dielectric Strength	kV/mm	>14
Dielectric Constant (at 1 MHz)		$\approx 9.8\varepsilon_0$
Loss Tangent (at 1 MHz)	$\tan\delta$	0.0001
Volume Resistivity (at 20 °C)	Ω.cm	$> 10^{14}$
Top Aluminum Cladding	µm	8 ± 0.5
Bottom Aluminum Cladding	µm	≈ 10

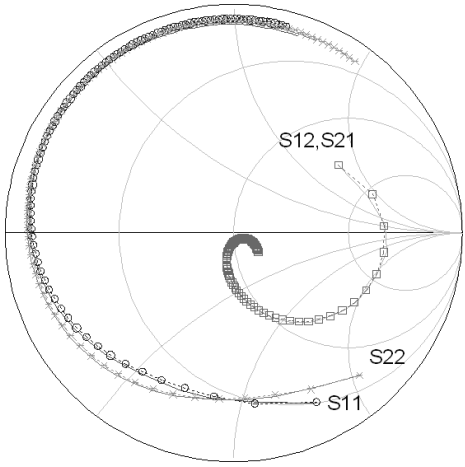


Figure 5.24: Smith chart representation of measured (○, □, and ×) and modeled (–) *s* parameters of the LDMOS device at V_{DS}=0 V and V_{GS}=0 V. Values measured over a range of frequencies from 0.1 to 5 GHz.

with the *s* measured data. This procedure succeeds as far as a good estimation of the initial values is done. The initial estimation of most of the extrinsic values are done at V_{DS}=0 V and V_{GS}=0 V, which usually is known as the cold-fet method [Lov94]. Although there are more precise and accurate methods to extract the extrinsic values [Ves02], the described procedure has shown to be sufficient for this purpose. The model with the extracted values shows a good fitting with the measured data over a range of frequencies from 0.1 to 5 GHz (Fig. 5.24), which is enough to calculate the cut-off frequency (f_T), the maximum oscillation frequency (f_{max}) of this device and for extracting the intrinsic components values as well.

5.4.2.3 Estimation of the Initial Values of the Components

A valid model tuned by optimization can be obtained as far as a good estimation of the initial values of the extrinsic element values is obtained, i.e., bonding wires, pad connections and drain, gate, and source parasitics. For the estimation of the bonding wire values, the typical values of 1 nH/mm and 0.125 Ω/mm are used. The initial values for the pad capacitances are calculated theoretically.

In order to estimate the initial values of the drain, gate and source parasitics (resistance and inductance) a simplification of the complete model had to be done at V_{DS}=0 V and V_{GS}=0 V. Basically, this simplification involves neglecting the pad parasitics and the silicon substrate losses since their impedances values are smaller compared to the intrinsic values of the device. Then a simpler model can be obtained as indicated in Fig. 5.25.

By applying the traditional cold-fet method, [Lov94] initial values of the resistance elements can be obtained (for instance bonding wire and ground plane resistance in Fig. 5.26), then subtracting the values of the bonding wire resistance (know from the previous step), initial values for the drain, gate and source parasitic resistance part can be obtained. To estimate the initial values of the inductance part, a first estimation of the device intrinsic capacitance should be done first. This is done by considering that at low frequencies the reactance values of the capacitances are much larger than the inductance reactance values. Hence the inductance can be neglected initially, and a first approach of these capacitance values can be obtained at lower frequencies. Then, the values for the inductances are estimated by subtracting the capacitance reactance from the total reactance at high frequencies, taking into account the initial values of the bonding wires inductances. Initial values for the substrate losses are obtained from previous reported values [Fra88].

Once all these initial values are obtained, an optimization process with a gradient method is applied to tune the model component values in such a way that the model simulation output fits the actual device measurements. As can be seen in Fig. 5.24, a good fitting is obtained from 0.1 GHz to 5 GHz. These tuned values are used to extract the intrinsic values of the device at different bias conditions.

5.4.2.4 Determination of the cut-off Frequency and the Maximum Oscillation Frequency

To obtain the f_T from the scattering parameters of the device, these parameters values are first de-embedded using the extracted model and then a transformation from scattering (s) to hybrid (h) parameters is made. As can be seen in Fig. 5.27, the typical characteristic of 20 dB/dec slope is shown in h_{21}.

To determine the f_{max}, several criteria are available. In order to apply the figures of merit unilateral gain (U) or maximum transducer power gain (G_{TUmax}) the device has to be unilateral, which is not the case for this device over the frequencies of interest. This narrows the criteria to the Maximum Available Gain (MAG) or Maximum Stable Gain (MSG), but it is determined that this device is conditionally stable over the frequency range of interest, so the only criterion applicable is the maximum stable gain.

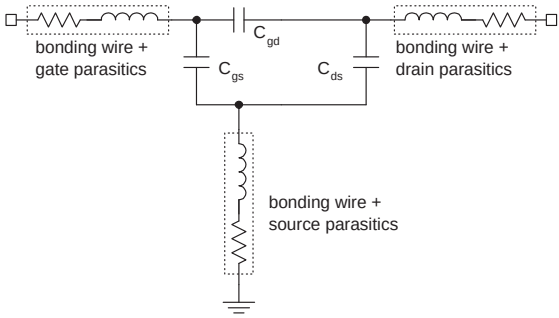


Figure 5.25: Simplified transistor model used to estimate initial values for the indicated components.

Based on this criterion the f_{max} could be determined (Fig. 5.28).

Two values are obtained, one given by the zero-dB cross and the other one by the extrapolated value from the 10 dB/dec MSG slope line. It is worth noting that among all the criteria for calculating f_{max}, the MSG is the one giving the highest values; therefore this value could be overestimated when it is calculated in this way.

The remaining figures present the different small signal capacitances. For the off-state, the capacitances C_{gs} and C_{gd} are given in Fig. 5.29 whereas Fig. 5.30 presents them for the on-state. The non-linear capacitance between the drain and the bulk is given in Fig. 5.31. These results are summarized in Table 5.4. For comparison of the LDMOS2 with the minimum length transistor, simulated values are given in Table 5.5.

5.4.2.5 Other Measurements

At frequencies where the lumped model no longer applies, the use of transmission lines in the form of coplanar waveguides or microstrips is a generalized method for transmitting a signal. In this case it is of interest to know the losses the signal suffers when passing through one of these structures. For such lines on the ceramic substrate described in Table 5.3m the insertion loss (S_{21}) is:

- microstrip: 0.102 dB.cm.GHz
- coplanar waveguide: 0.0148 dB.cm.GHz

For *Line4* with 45 mm this represents a total loss of 1 dB at 15 GHz while a microstrip with the same length would have 1 dB loss at 2.2 GHz.

The thin film process also makes available the use of vias between the top and bottom aluminum layers. A comparison between simulation and measurements show that the vias have the expected behavior. This is confirmed for frequencies below 5 GHz. Nevertheless, the manufacturing process does not always produce a good quality via. In this case a number of vias in parallel are necessary to compensate the defective ones.

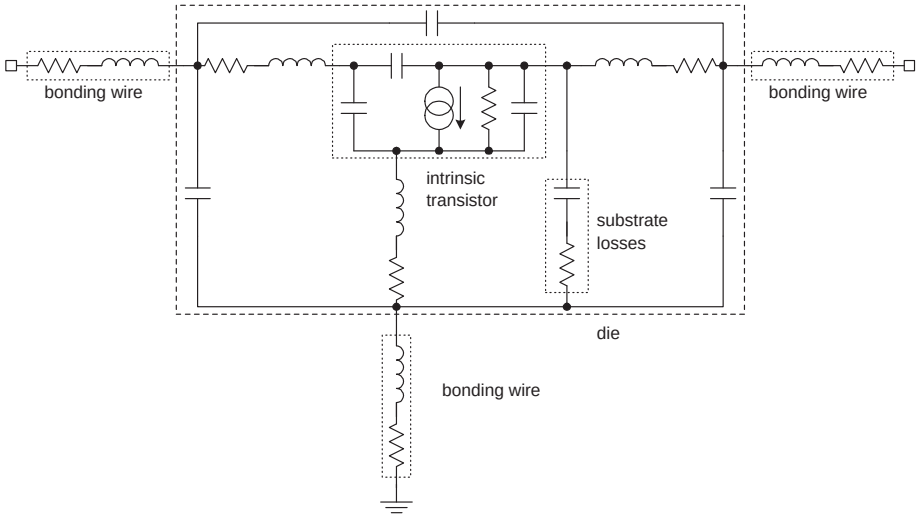


Figure 5.26: Transistor equivalent model including both the intrinsic and extrinsic elements. Substrate parasitics are included to increase accuracy of the fitting parameters to the measurements.

5.4.2.6 RF Transistor Model for Class E Power Amplifier

In a class E power amplifier, the active device does not simultaneously have current and voltage. Therefore, modeling requirements are simply reduced to the on- (triode) and off-region. A simple yet accurate model for the application in question is shown in Fig. 4.11, which includes a resistor and all its relevant capacitances with the values given by Table 5.2 and 5.4.

As show in Table 5.1, the simulated on-resistance for transistor M_{1_01} has an error of 18 % of the measured value. Concerning device accuracy influence in a circuit implementation, simulations using both models are carried out:

- BSIM3 model from the foundry
- Transistor M_{1_01} model from Fig. 4.11 with the capacitances and on-resistance values given respectively by Table 5.5 and Table 5.2

For the circuits given in Fig. 4.3 and 4.17 no significant difference is observed (Fig.5.32). Although the transistor M_{1_01} model is only accurate within 5 % of the BSIM3 transistor model (in transient simulation), the same optimal efficiency is found for the synthesized PA. The conclusion here is twofold: or the class E power amplifier is relatively insensitive to variations in the transistor on-resistance, or the other elements in the model somehow compensate the transistor on-resistance influence.

Table 5.4: Measured capacitances per millimeter length.

Device	C_{gs} (pF)		C_{gd} (pF)		C_{dbo} (pF)
	OFF	ON	OFF	ON	
LDMOS2	1.61	4.30	0.54	2.92	0.66

Table 5.5: Simulated capacitances per millimeter length.

Device	C_{gs} (fF)		C_{gd} (fF)		C_{gb} (fF)		C_{dbo} (fF)
	OFF	ON	OFF	ON	OFF	ON	
M_{1_01}	192	802	192	764	377	76	1594

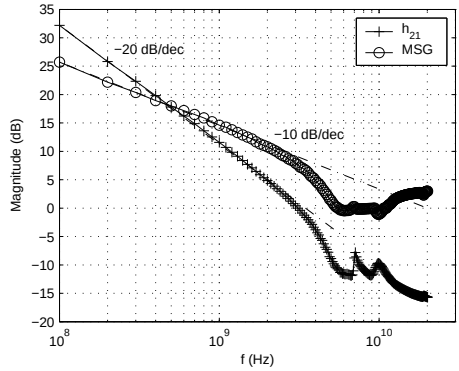


Figure 5.27: Device cut-off frequency and f_{max} at V_{DS}=3 V and V_{GS}=1 V.

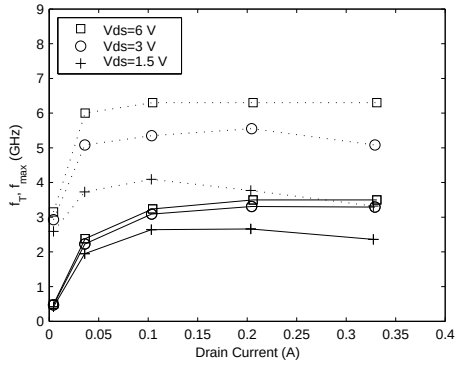


Figure 5.28: f_T (solid line) and f_{max} (dashed line) versus drain current for three different V_{DS}. A maximum value of 3.5 GHz is measured for the f_T. The maximum f_{max} is measured to be 6.3 GHz.

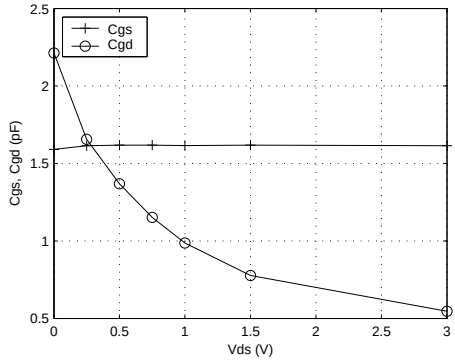


Figure 5.29: C_{gs} and C_{gd} per millimeter length at 1 GHz with V_{GS}=0.

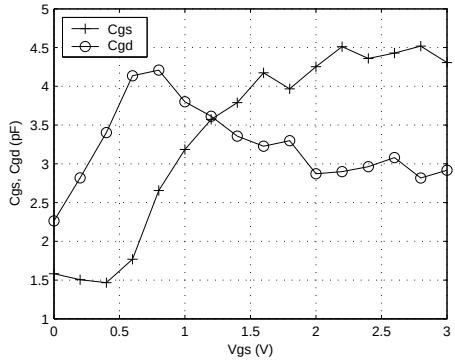


Figure 5.30: C_{gs} and C_{gd} per millimeter length at 1 GHz with V_{DS}=0.

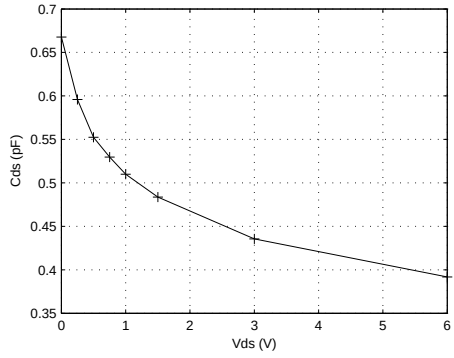
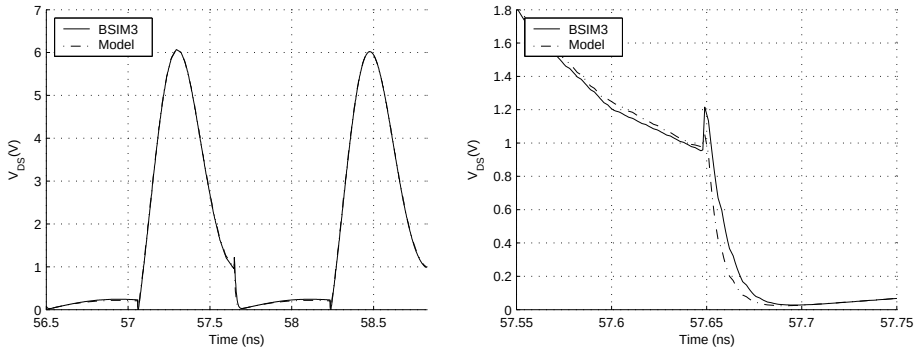


Figure 5.31: C_{ds} per millimeter length at 1 GHz with V_{GS}=0.



(a) Transient response.

(b) Zoom in of (a) for accuracy comparison.

Figure 5.32: Comparison between the SPICE BSIM3 model and the model from Fig. 4.11. Values for the capacitances and on-resistance are taken from Table 5.2 and Table 5.5. Because the substrate is short-circuited with the source, the measured C_{gs} value includes the C_{gs} and C_{gb}.

5.4.3 Further Improvements

When designing a circuit, it is important to minimize the power consumption. As the gate length is not at a minimum, the input capacitance will increase the power on the previous driving stage. Using typical published parameters for a 0.35 µm process, the device simulator Medici™ is used to investigate alternative designs. A solution exists but requires the use of an extra mask: n^-. The new schematic diagram is given in Fig. 5.33 [Ram03b]. A minimum length transistor is now possible with some advantages. First, the gate capacitance is smaller. Secondly, the manufacturer transistor model can be used to model the channel region below the gate. Thirdly, the drain extension that is added requires minor modeling effort. For some simple applications it might be enough to know the capacitance model and the switch-on resistance. Also, the new device continues to use the poly as n^+ implant blocking layer in the drift region.

Usually the breakdown first occurs near the surface [Gro67]. In this case the implant depth for the new layer n^- can be made shallower as it only needs to avoid surface breakdown. However, it can include the drain implant n^+ if it is necessary to increase the operating voltage above the high doping drain bottom breakdown voltage.

A simple manufacturing process can be used where the implant dose to create the n-well is unaltered, but where the implant energy is lowered to reduce the penetration depth. When both the implant energy and the implant dose can be controlled, a higher breakdown voltage or a lower on-resistance can be obtained.

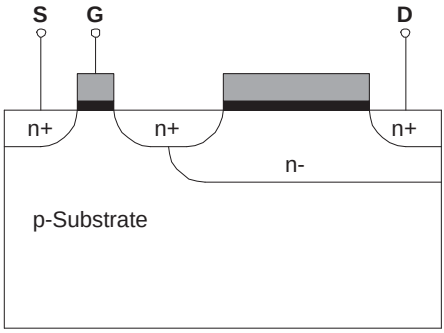


Figure 5.33: Cross section of improved LDMOS.

5.5 Design for High-Voltage: an Advantage?

In the previous sub-sections solutions to work at high voltages and reliability issues and their impact on the performance of transistors have been discussed. Now, with the measurement results, the other aspect to consider is how it influences the results of the class E power amplifier.

In sub-section 5.5.1 the effect of the switch-on resistance in the performance of the class E amplifier is presented. In sub-section 5.5.2 the impact of device scaling is discussed. In particular, the effect of the reduced supply voltage with the consequent increase in the circuit current is given. In sub-section 5.5.3, the effect of the characteristics of the LDMOS transistor presented in sub-section 5.3.2 is given, focusing in particular on the effect of the total input capacitance.

5.5.1 Influence of the on-resistance in the DE and PAE

The measured on-resistance of the LDMOS in Fig. 5.11 is approximately 2.2 times greater than the foundry transistor. To see its influence in the class E performance a similar study to the one in sub-section 4.5.1 is done. Note that the same circuit and the same capacitors per unit length are used. This makes it easier to draw conclusions on the real effect of an increase in the switch-on resistance. Nonetheless, the inclusion of the correct capacitance values from Table 5.5 is required for implementing the final amplifier.

Optimization results are given in Fig. 5.34. For easiness, the results for the case of the LVMOS are given alongside. The maximum DE, shown in Fig. 5.34(b), sees its value decreased due to the increased resistance. The penalty is proportional to the inductor quality factor (Q_L) and ranges from 2 % to 8 %. The PAE also reveals a similar behavior. In conclusion, it is not possible to achieve better performances with a transistor if it has a higher switch-on resistance, as is the case of an LDMOS and more specifically, the one presented in Fig. 5.34. The possibility of a better PAE, as show in Fig. 5.34(c) exists if the driving stage makes use of a lower supply voltage.

Table 5.6: Optimized circuit sizes for same conditions as in Table 4.3 with the exception that the switch-on resistance is 2.2 times greater.

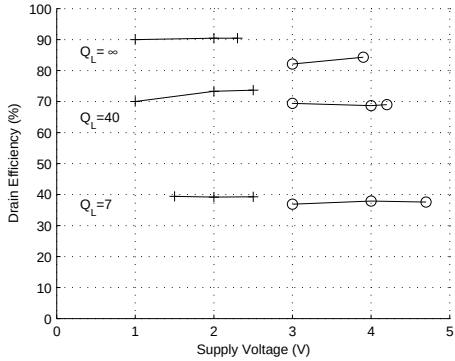
Q_L	V_{DD}	W	L_1	L_0	C_0	R_L
∞	3.9 V	4656 µm	37.4 nH	4.64 nH	20.6 pF	9.18 Ω
	3.0 V	6396 µm	14.5 nH	3.91 nH	16.8 pF	6.98 Ω
40	4.2 V	5796 µm	16.38 nH	3.24 nH	77.0 pF	8.04 Ω
	4.0 V	6610 µm	18.22 nH	2.75 nH	82.1 pF	6.41 Ω
	3.0 V	10533 µm	7.19 nH	2.04 nH	59.1 pF	4.36 Ω
7	4.7 V	6908 µm	5.65 nH	3.16 nH	40.6 pF	7.69 Ω
	4.0 V	9956 µm	4.30 nH	1.99 nH	79.9 pF	4.74 Ω
	3.0 V	15095 µm	2.13 nH	1.51 nH	61.5 pF	3.43 Ω

Table 5.6 gives the final optimization results. Comparing them with Table 4.3, the use of a higher supply voltage results in considerably smaller transistor lengths. The main advantage is having a higher optimum output load. As resistive losses in the matching network will always exist, more power will be available as they are now a smaller fraction of the total resistance. As a result, the design of the amplifiers is eased. Equally important is the reduction of the current that flows in the circuit.

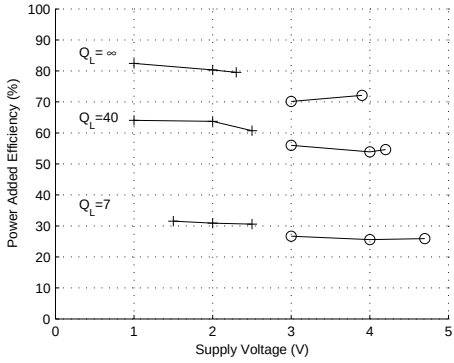
5.5.2 Design on a More Advanced Technology

As seen in 4.4.2.4, there are different sources of losses in a class E power amplifier. For simplicity, only switch-on resistance is now considered. When scaling the transistor sizes, new and more advanced CMOS technologies impose constraints in design, namely, the maximum supply voltage that can be used. Assuming a scaling factor k and for a constant electric field strength, arises that maximum supply voltage must be decreased. With scaling, transistor input capacitance ($WL.C_{OX}$) is also linearly decreased with this factor. Moreover, scaling also reduces channel resistance which in turns reduces losses for a fixed current. Nevertheless, for power driving circuits, as is the case with a power amplifier, a fixed output power and a reduced supply voltage requires that more current must flow in the circuit. With three variables to look at: V_{DD}, input capacitance and channel resistance, some considerations can be made.

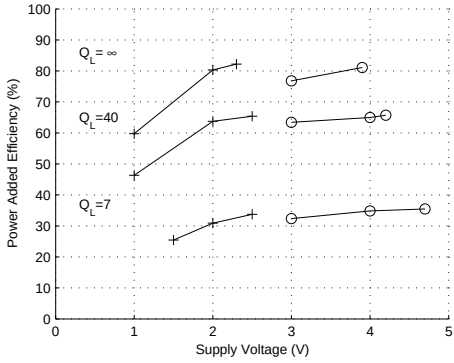
- The power amplifier gains from going to more advanced technologies in terms of DE and PAE. Although the results in Fig. 5.34 are for the case where the input capacitance per unit length is the same for both devices, all the other characteristics can represent one technology and its scaled version. As such, the input capacitance is higher than the value used to generated the PAE line. Consequently, the input power increases and the PAE value will have its value decreased.
- Older technologies can sustain a higher supply voltage and are hence allowed to have smaller transistor widths. However, the higher input capacitance per unit length added to the fact of a higher on-resistance of the switch has the effect that a similar input capacitance is reached. This further supports the inference in the previous item that the DE and PAE is higher in more advanced technologies.



(a) Drain Efficiency.



(b) Power Added Efficiency with V_{DD} in (4.25) equal to the supply voltage.



(c) Power Added Efficiency with V_{DD} in (4.25) equal to 2 V.

Figure 5.34: Drain Efficiency and Power Added Efficiency as a function of supply voltage for different inductor quality factor Q_L. Transistor on-resistance for the LDMOS M_{A_01} (○) is 2.2 times greater than for the LVMOS M_{1_01}(+).

- The effect of an increased current in the circuit for the case of the scaled technology cannot be overlooked. Because electromigration is proportional to the current density, electromigration will increase by a factor of $1/k$.

An immediate consequence of the last observation is that the decision of going to more advanced technologies or not in terms of amplifier efficiency is limited by how serious the electromigration problem is and how limiting to the amplifier efficiency are the losses in the interconnections and other passives as a result of the increased current. This is a consequence of the reduction in the on-resistance and input capacitance provided by a scaled technology with the consequent increase in efficiency. Still, previous technologies have a lower price tag which can in the same case make one solution more attractive as it produces a lower-cost final product.

5.5.3 Design with an LDMOS Transistor in the same Technology

The device in Fig. 5.11 is designed with a class E power amplifier in mind. Without changes in the process flow or any additional mask it is possible to duplicate the breakdown voltage of the 0.35 µm CMOS technology. Hence, a supply voltage of 4 V can be used, resulting in a higher load resistance R_L and a narrower transistor. The gain in the transistor size is not enough to compensate for the increase in the input capacitance per unit width. Consequently, the power necessary to drive this output stage will increase slightly. Nevertheless, the current flowing in the circuit is significantly lower.

More complex technologies are required to lower the input capacitance and as such profit from the advantages given by a higher supply voltage. Minimum process changes in CMOS (sub-section 5.4.3 [Ram03b]) or the use of more advanced technologies [Ehw01] do allow a reduction in the input capacitance. Furthermore, this has to be accompanied by a reduction in the supply voltage of the driving state so that the PAE increases with an increasing supply voltage (Fig. 5.34(a)).

5.6 Conclusion

In this chapter, the design and measurement results of a lateral diffusion MOS high-voltage devices that can be seamlessly integrated into commercial CMOS processes without any additional mask have been presented. Because the LDMOS relies on common technological steps, integration with more advanced technologies using STI as their preferred device isolation method is possible.

Designed in a 0.35 µm technology and using standard 7.4 nm gate oxide, the device exhibits a blocking voltage of 14.7 V, nearly double those available in the technology. It features a R_{ON} of 4.8 Ω/mm. A maximum value of 3.5 GHz is measured for the f_T while the maximum f_{max} is measured to be 6.3 GHz. The application can range from I/O circuits interfacing the outside world; handling the high programming voltage in a CMOS compatible EEPROM; RF LDMOS transistors to be used in integrated power amplifiers for mobile communications and where a change in the process flow is unwanted. Based on device simulations, a new structure is shown. At an expense of an extra mask it provides reduced input capacitance and higher breakdown voltage.

Guidelines for minimizing some of the transistor issues as a result of the high electric fields have been outlined. They range from circuit design to technologic advances that increase chip performance.

The correct selection of the technology should value the small input capacitance and on-resistance offered by more advanced technologies as they provide higher Drain Efficiency (DE) and Power Added Efficiency (PAE). On the other hand, with an older technology or with an LDMOS transistor, resistive losses in the output network will weigh less due to a higher equivalent output resistance as a results of the higher supply voltage used. Furthermore, if a high output power level is required, the LDMOS or an older technology might be the most appropriate as there is a limit to how much output power can be obtained with an advanced low-voltage process. Some of the advantages of a high power supply include: (a) lower current and lower losses in the interconnections, (b) transistors with smaller widths, and (c) the possibility to have a higher output power.

The next chapter provides, in conjunction with this chapter and Chapter 4, the specific circuit level implementation of a class E power amplifier for mobile communications. A relatively detailed analysis and design of a 30 dBm class E power amplifier operating at 850 MHz is presented.

Chapter 6

A 850-MHz, 30 dBm Class E Power Amplifier

6.1 Introduction

The necessity for low-power consumption for reasons of battery autonomy but also for less stringent cooling strategies calls for higher efficiency circuits. For wireless circuits, the power amplifier is responsible for a large slice of the total power budget. As such, integration of this block in the full transceiver is seen as one solution to decrease power consumption, by eliminating the need to match the output of the up-converter to be matched to 50 Ω resistor, and probably, by creating circuits where some of the functionality is merged within the last stages before the antenna. In particular, it is advantageous if they can be designed in standard CMOS technology.

The theoretical analysis of the class E power amplifier from Chapter 4 needs to be verified in silicon. The design of such a circuit is presented in this chapter. The difficulty in this case resides not only in the fact that other technologies have better RF characteristics but also in the fine tuning of all circuit elements. This is the reason why there is a lack of circuits with a Power Added Efficiency (PAE) above 50 % for circuits designed in CMOS.

The design presented here is a class E power amplifier working at 850 MHz with a maximum PAE of 66 % implemented in a standard 0.35 µm CMOS technology.

Firstly, Section 6.2 briefly delineates some of the characteristics of the Global System for Mobile Communications (GSM) standard for mobile communications. In Section 6.3 the design considerations to maximize circuit efficiency are described. This is followed in Section 6.4, by an explanation on how the circuit has been designed. Next, in Section 6.5 the main measurements are reported and a comparison of the designed PA with existing state-of-art CMOS power amplifiers is presented. Finally, the conclusion of this chapter is given in Section 6.6.

6.2 Global System for Mobile Communications (GSM)

GSM is today the most popular standard for mobile phones in the world. The universality of the GSM has lead to the situation where a user from one network can use their phone in an area where the phone provider does not have a network. This can either be in a region or within different countries and is usually known by the term ”roaming” in telecommunications.

Although GSM has been evolving with time, it has retained backward compatibility with the original GSM phones. This not only decreases the cost of operation but also makes evolution easier. In total, GSM has had two phases of evolution. During Phase 1, the initial delivery of the standard was made available. Phase 2 added non-voice services like the SMS and enriched functionality such as caller information.

GSM exist in four main versions from a total of 12 [3GPP], based on the band they use. GSM-900 and GSM-1800 used in most of the world and GSM-850 and GSM-1900 in the USA and Canada are the most common ones. Some of its technical characteristics are now given.

6.2.1 Specifications

In its specification the GSM-850 uses two bands. One is used for the uplink at 824-849 MHz (mobile station to base station) and the other at 869-894 MHz is used for the downlink (base station to mobile station).

GSM is a radio network based on FDMA/TDMA technology. Each group of eight users transmit through one channel (200 kHz), sharing transmission time (TDMA) with a slow frequency hopping between channels. Each of the 25 MHz bands are subdivided into 124 channels (FDMA), each one of them 200 kHz wide. At each side of the band, a guard frequency of 100 kHz is used, with a 200 kHz spacing existing between each carrier. The edge channels (1 and 124) are optional for the operator.

A Gaussian shift-key modulation is used which gives mobile devices a better battery life because it encodes the data by varying the frequency and not the amplitude. In this way, an efficient nonlinear power amplifier can be used at a high output power level without distorting the transmitted information. The tradeoff is that this type of modulation is also inefficient in terms of spectral efficiency which means that each user consumes more bandwidth than what would be necessary with a more efficient modulation scheme.

In Phase 1 there are five classes of mobile stations (MS) defined according to their peak transmission power rated as given in Table 6.1. Again, to conserve power (and minimize co-channel interference) each device operates at the lowest possible power level that maintains an acceptable signal quality. As an example, a class 4 with a RF power capability of 2 W transmits 250 mW on average if using a single time slot per frame. Mobile devices that have experienced the biggest success and have an higher market quota can emit up to 2 W. The desire for smaller and lighter battery operated devices has pushed the value down to as low as 0.8 W [3GPP]. Output power level can be dynamically adjusted in 2 dB steps from the peak power for the class down to a minimum of 13 dBm (20 mW).

Phase 2, also known as E-GSM-900, added the capacity for additional channels in the GSM-900 band by increasing the up-link and down-link bandwidth by 10 MHz which boils down to 50 additional channels. The minimum power level transmitted by the mobile station has also been decreased to 5 dBm (3 mW).

Smaller power levels have been also added for the base station (BS) to accommodate a new feature. The GSM Cordless Telephony System (CTS) which allows subscribers to

use their GSM mobile phones to send and receive calls over the fixed public network.

GSM signal transmissions, by being a TDMA signal, have the characteristic that they are time limited. This requires the transmitted signal to have a strict control over time and in frequency. In Fig. 6.1 a power-time mask to be applied against the transmitter signal is shown. The spectral mask due to Gaussian Minimum Shift Keying (GMSK) modulation is presented in Fig. 6.2. Compliance with both ensures minimal interference with users in adjacent time slots and channels and must be strictly followed.

6.3 Design for Maximum Drain Efficiency

Maximum circuit efficiency is obtained when all power sources are minimized with respect to the output power. In an integrated circuit this includes not only the supply voltage but also the driving signal necessary to properly switch the amplifier. For a discrete power amplifier, the RF input signal has to be brought off-chip and drive the 50 Ω on board. In a fully integrated solution, only the voltage swing is necessary and less power is thus dissipated.

For the power levels intended in the current application, a 10 dBm input signal has minimum influence in overall PAE at maximum output power. This simplifies the design as the circuit can be designed to only minimize the power consumption from the power supply. This and other optimization characteristics are now described.

6.3.1 Basic Architecture

The class E power amplifier, of which the implementation is described shortly, is a differential two stage amplifier (Fig. 4.17). The proposed solution is defined by:

- A nonlinear, high efficiency class E power amplifier well suited for the application GSM-850 standard. This is because information is encoded only in frequency variations and not in amplitude.
- In a differential architecture, substrate noise is reduced as current is discharged to the substrate twice per cycle. This noise frequency is now at twice the desired signal frequency which can be considered to be a common-mode signal, to which the differential architecture is more immune.
- A two stage amplifier is a compromise between a simple design while being suitable for full integration as now the previous block in the transmitting path, the VCO, sees the much smaller transistor from the driving stage.
- Although on-chip spiral inductors are desirable for circuit integrability, their low quality factor limits the maximum attainable efficiency to about 40 % (Fig. 4.13).

The next two sub-sections now describe the optimization procedure of the circuit with the above characteristics. In sub-section 6.3.2 a final design exploration on the variation of the drain-bulk capacitance per transistor unit width is investigated for the circuit in Fig. 4.17. Finally, in sub-section 6.3.3, the circuit optimization of the final differential

Table 6.1: GSM Mobile Station Types.

Class	Output Power (W)
1	20
2	8
3	5
4	2
5	0.8

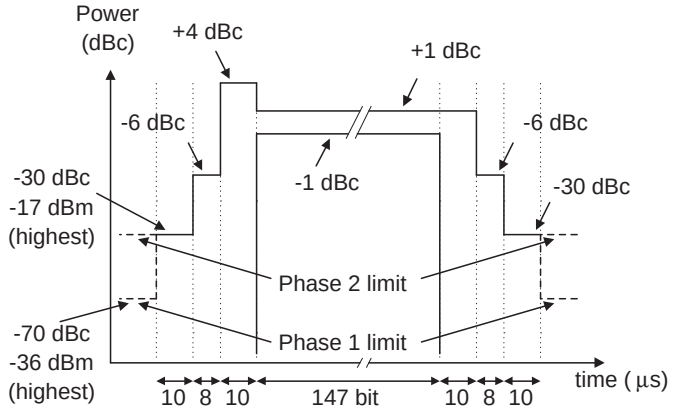


Figure 6.1: Mask limits for the GSM burst as specified by the standard.

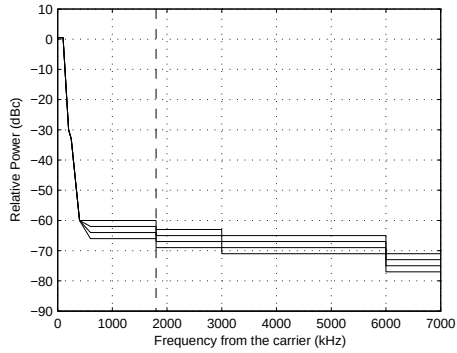
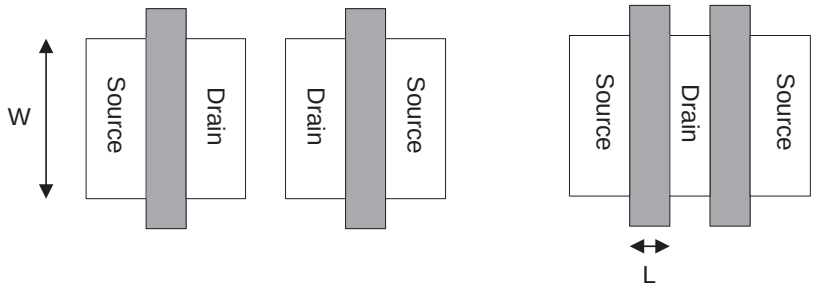


Figure 6.2: Mobile station GSM spectral mask due to GSMK modulation. A 30 kHz measurement bandwidth is used for frequencies below 1800 kHz and 100 kHz otherwise. A higher attenuation is required for higher transmitted power (Table 6.1), resulting in the different horizontal lines.



(a) GEO=0: drain junctions are not shared with other stripes of the same transistor. This is used to increase the transistor drain-bulk capacitance.

(b) GEO=1: drain junctions are shared. The drain-junction capacitance (and the necessary area used) is kept to its minimum for a given transistor width.

Figure 6.3: Source/Drain transistor junctions layout.

two stage power amplifier, including all relevant circuit and board parasitics to better describe the circuit measurement performance, is given.

6.3.2 Two Stage Class E Power Amplifier Optimization

In a fully integrated solution, the necessary power to drive the PA must be as low as possible to maximize the PAE of the amplifier (4.25). According to (4.3), the smaller the size of the input stage, the less power consumed by the previous stage in providing the required signal drive. It is thus advantageous to size the full amplifier where the tradeoff of a more (or less) efficient output stage is weighed against a less (or more) efficient driving stage. Equally important is the need to consider the limitation imposed on the transistor breakdown voltage which can result in some cases that a less efficient solution must be used because it has a lower maximum drain voltage.

Anther variation that can be included in the basic circuit of Fig. 4.17, is the variation of the parallel capacitor C_{db} value per unit width. This extra degree of freedom can be exploited to further tune the design goals. The results for the case when this capacitance has been doubled or halved are represented in Table 6.2 by 2.0 and 0.5, respectively. Although the drain-bulk capacitance has been considered here, this can also be an additional linear metal-metal capacitor, which has the advantage of causing a lower peak voltage across the transistor [Chu94], thus causing lower stress to the device.

In Table 6.2, this reference (1.0), doubling (2.0) and halving (0.5) represent, respectively:

- A transistor where the drain-bulk junction is not shared and is layed-out with a minimum diffusion size (Fig. 6.3(a)).

Table 6.2: Efficiency for different transistor drain-bulk capacitance per unit width for the circuit in Fig. 4.17 and 1 W output power.

$C_{db}(W_2)$	$C_{db}(W_1)$	DE	$V_{DS}(max)$	W_2	W_1
1.0	1.0	68.0 %	7.95 V	260 µm	8978 µm
0.5	0.5	69.6 %	8.60 V	100 µm	12077 µm
0.5	2.0	65.3 %	7.05 V	491 µm	6826 µm
2.0	0.5	69.8 %	8.45 V	105 µm	13760 µm
2.0	2.0	65.8 %	6.90 V	696 µm	7413 µm

- As above, but now the drain-bulk junction is increased so that the transistor has the double drain-bulk capacitance per unit width as before.
- The arrangement of (Fig. 6.3(b)) where the drain-bulk capacitance is shared with the same stripes of the same transistor and is layed-out with minimum diffusion size.

From the results given in Table 6.2 a few conclusions can be drawn. First, the higher DE values are obtained for the lower values of C_{db} in the output stage. The value for the driving stage seems somehow to have a lower influence on the overall efficiency. Secondly, the required transistor size for the driving stage is considerably smaller than the output stage. Although the results presented suggest values in the range of a few hundred microns, further analysis reveals that the amplifier works as in sub-section 4.5.5, on the edge of stability. This occurs in spite of the simulation output revealing the expected class E waveforms. As a consequence, the size of the transistor W_2 must increase to have a functional amplifier. Therefore, Table 6.2 gives the best case estimation for the DE. Thirdly, the width of the output transistor, as expected, is inversely proportional to the capacitance per unit width (4.27). This is the consequence of a certain output power requiring a specific shunt capacitance value. Once this capacitance per unit width is decreased, the transistor size must be proportionally increased to compensate for this reduction. And finally, the maximum drain-source in voltage in the output stage ($V_{DS}(max)$) is decreased for the cases where the unit length capacitance is increased.

From the above, a few choices must be made prior to chip implementation.

- The higher efficiency is obtained with the largest size for the transistor in the output stage, as a consequence requiring more silicon area.
- The lower stress in the output stage transistor occurs when the transistor drain-bulk capacitance per unit width is increased, consequently increasing the chip area.

Although a higher efficiency is desirable, it comes at the expense of a higher drain-source voltage. As such, the maximum supply voltage must decrease in order for the circuit to operate in a safe region. The consequence is that the same output power level

Table 6.3: Optimized circuit sizes for the circuit in Fig. 6.4 considering all circuit and ceramic substrate parasitics.

	W_2	L_2	W_1	L_1	L_M	C_M
PA-LVMOS1	1500 µm	4.17 nH	4750 µm	2.08 nH	5.34 nH	6.01 pF
PA-LVMOS2	2000 µm	2.43 nH	8000 µm	2.30 nH	4.87 nH	5.86 pF

must be obtained from a lower supply voltage, hence, with a higher current flowing in the circuit. The stress on the transistor is considered to be a more critical design issue. Accordingly, an implementation featuring a lower maximum drain-source voltage has been selected to design the PA presented in this work.

6.3.3 Fully Parasitic-Aware Power Amplifier Optimization

In the previous sub-section a relatively accurate power amplifier has been studied. The present section further expand this study by including the effect of the transmission lines, inductor couplings, pads, on- and off-chip interconnections and capacitor parasitics. The schematic of a such circuit is presented in Fig. 6.4. Unlike what has been commonly used so far, the series tuned (L_0-C_0) is replaced by a more convenient output L-matching (L_M-C_M) network. The 50 Ω output resistance (representing the antenna) is in this way transformed to the previous R_L value. Nevertheless, this necessary change to make testing possible does not come without a cost. Under the same conditions, the change of filter topology decreases the maximum achievable efficiency by about 4 %.

The schematic diagram of Fig. 6.4 accurately represents a possible test setup. The input transmission lines are considered to be lossless and their effect mainly accounts for the peak input voltage variation. The coupling between each of the parallel inductors is extracted with FastHenry, a 3D inductance extraction program. The coupling factor is iteratively changed in the SPICE simulation and the optimized value used again as the input for FastHenry. The power supply trace in the ceramic substrate is also modeled. Decoupling capacitors are then added to maximize the DE. The parasitics of the SMD capacitor used in the output low-pass matching filter is modeled as given in Fig. 4.9. Bonding wires and the output pads are modeled as shown in Fig. 4.8 and Fig. 4.10, respectively.

The correct placement of capacitor C_M along the output transmission line have shown to be a manner to further tune the impedance seen by the output transistor. This effect is represented by TD in Fig. 6.4. An increase of about 2-3 % in the DE is possible by placing the capacitor no further than a couple of millimeters apart from the bonding wire end on top of the transmission line.

The final optimized values for circuit elements are given in Table 6.3. Both the amplifiers PA-LVMOS1 and PA-LVMOS2 have the input transistor layout with GEO=1 (Fig. 6.3(b)) while the output transistor has GEO=0 (Fig. 6.3(a)). The difference is that the output capacitance per unit transistor width for the PA-LVMOS1 is double that of the PA-LVMOS2 i.e., in Table 6.2 $C_{db}(W_1)$ is equal to 2.0 and 1.0, respectively. Both

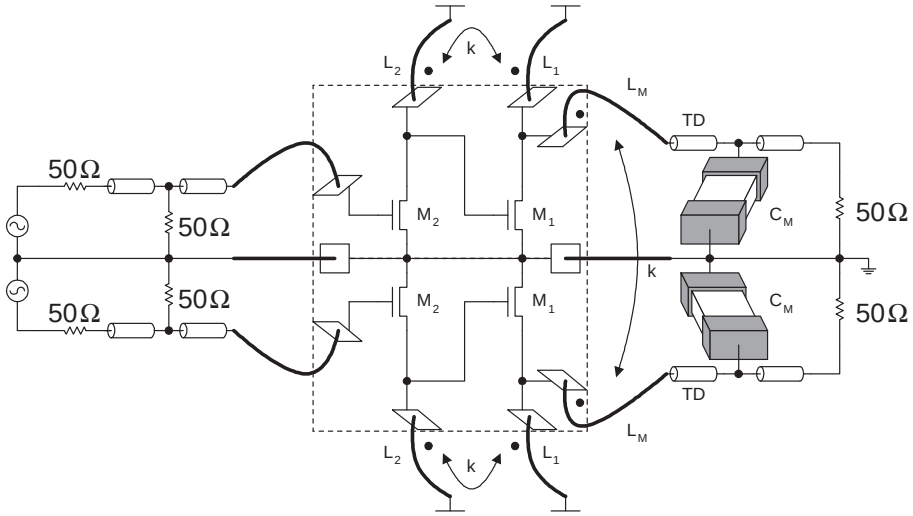


Figure 6.4: Complete two stage fully differential class E PA circuit schematic including all parasitics. The performance-limiting parasitics are taken into account during optimization: pad, inductor, discrete SMD capacitor and transistor. Bonding wires connecting to the ground are considered to be enough in order not to limit the performance. To better describe the testing environment, coupling between the inductors is included. Furthermore, the transmission line delay (TD) between L_M and C_M is fine-tuned to maximize efficiency. The die is indicated by a dashed line.

have an approximated DE of 66 %. The difference is that the PA-LVMOS1 has a higher on-resistance. It has, nevertheless, a lower input capacitance which produces a lower transistor for the driving stage, consequently reducing the power necessary to drive the power amplifier.

In the next section the implementation of the power amplifier in silicon is presented.

6.4 Implementation of the Class E Power Amplifier

The layout of the final amplifier is mainly limited by the need to provide a good ground connection at 850 MHz and having interconnections wide enough to avoid being affected by electromigration. For a given current, the width of the metal interconnections strongly depends on the temperature. Due to this exponential dependence, the less the power that is dissipated due to a more efficient amplifier, the more compact it can be made. For each stage in the amplifier, these and other layout limiting characteristics are now given.

6.4.1 Output Stage

Being the largest stage, special care has been taken to facilitate the large current flow. Supply and output pad are surrounded by parallel smaller transistors, each one having a total transistor width of 250 µm, which are added together until the total output transistor size is reached. In this way, the current flows from the center of the chip to the periphery which forms a ring of pads connecting to the ground. This provides not only a symmetrical layout but also an even thermal gradient distribution along the chip.

The basic building block transistor, having a width of 250 µm, is layed-out with ground connection using only the bottom metal layer. Sufficient substrate contacts are placed around each basic block so that any point in the fingered transistor is not further than 10 µm from one substrate contact. This is supposed to create a good and stable grounding in terms of frequency and total current in the circuit. A total of 20 transistors, each one having a width of 12.5 µm, compose the cell. Two things define this width: the RC time constant of the transistor and the number of vias that are necessary for the total current in each block. In this case the number of vias are over-dimensioned for certainty while the RC of the transistor is supposed to be much smaller than the period of the 850 MHz carrier signal.

Connection from the transistor drain to the output and supply pad are layed-out using the remaining four metal layers. All interconnections have their size decreased to the minimum necessary.

6.4.2 Driver Stage

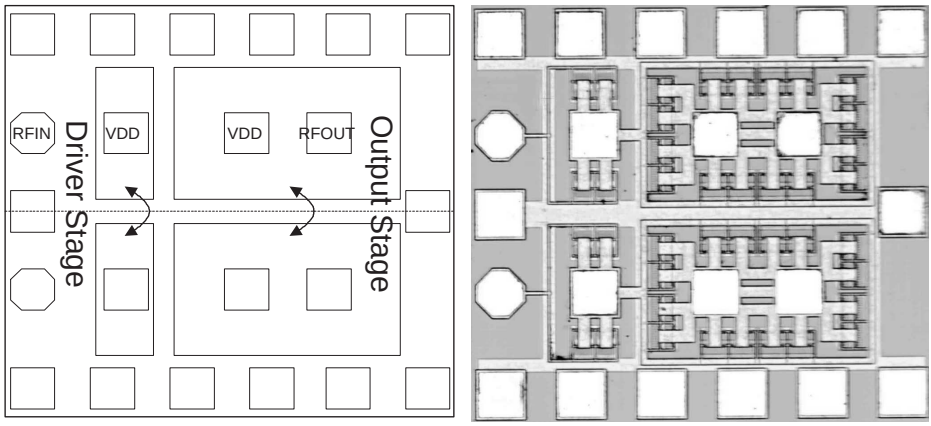
The smaller driver stage is also built using a transistor of 250 µm, but different unit cell. The layout follows the same guidelines of the output stage and no further information is thus required.

6.4.3 Interstage Interconnection

The signal enters the chip and travels from the driving stage to the output stage through metal lines. These have been sized considering the current flow, resistance and inductance. FastHenry is used to model each section of the interconnection that is later simulated in SPICE. The goal is to minimize the resistance and inductance of the interstage interconnection.

6.4.4 Chip Layout

The chip microphotograph is shown in Fig. 6.5(b). From left to right, the driving stage and output stage can be clearly distinguished. The hexagonal bondpad is used to bring the signal into the chip; two other bondpads form the drain connections of the transistor, and the last one in the central part is used to take the signal out. The top circuit is then vertically mirrored, in this way forming a differential amplifier. With the exception of the two circular RF input pads on the left, the periphery is used only for the 14 ground connected bond pads.



(a) Floor-plan of the amplifier.

(b) Microphotograph.

Figure 6.5: The RF signal enters the chip on the left side through the hexagonal pad that connects directly to the driver stage. The amplified signal is further amplified in the output stage before leaving the chip via the RFOUT pad. The final fully differential amplifier is constructed by a mirror and copy operation. All the remaining pads are used for connecting to the ground.

6.5 Experimental Evaluation

In sub-section 4.4.2 a methodology for optimally designing power amplifiers for maximum efficiency has been presented. Different simulations are then given to show its usefulness. In this chapter more simulations are given. This section presents the measurement results of one power amplifier designed with PAMPER following the proposed methodology being a validation of the tool described in sub-section 4.4.2. The fully differential two stage class E power amplifier has been produced in a standard 0.35µm CMOS technology. The experiments are now presented. The outline of this section is as follows.

In sub-section 6.5.1 the measurement setup used to characterize the amplifier performance is given.

The measurement of the designed power amplifier is probably the most difficult task in the overall design frame. The correct design of the ceramic substrate, the quality of the ground plane and the reduction of all parasitics has proved to be a time-consuming task. Afterwards, measuring has been a pleasant and an enriching experience. The measurement results are presented in sub-section 6.5.2.

Finally, in sub-section 6.5.3, the performance of the implemented amplifier is compared to the performance of other recently published power amplifiers. The conclusions of this section are then drawn.

6.5.1 Measurement Setup

The measurement setup used to characterize the class E power amplifier is presented in Fig. 6.6(a). A photo of the die attached to the ceramic substrate with the bonding wires, filter capacitor (C_M) and decoupling capacitors is shown in Fig. 6.6(b). The substrate is inside a copper-beryllium box to shield the circuit from external interference sources. The bare die is directly glued to the substrate by means of a conductive epoxy. All inductors are implemented using 25 µm in diameter aluminum bonding wires.

Decoupling capacitors are used to give a stable power supply voltage. A high number of vias are used to provide a low-inductive path from the top ground cladding metal layer to the full bottom grounding plane. Each individual metal section parasitic is modeled with FastHenry. Determination of their influence is verified with SPICE. All SMD components, transmission lines, ground quality and output L-matching networks are then measured with a Vector Network Analyzer (RS ZVM) to verify all the assumptions.

The input of the PA is on the right side of the board. A commercial power splitter (Mini-Circuits ZAPDJ-2, 1-2GHz 2 Way-180° Power Splitter/Combiner) is used to provide the necessary differential input signal coming from the Vector Signal Generator (RS SMIQ 06B). On board, 50 Ω transmission lines, terminated on a 50 Ω resistor provide the necessary input signal for the operation of the power amplifier. The output of the PA exits from the right side in Fig. 6.6(b) through a DC blocker. The differential signal is then transformed into single ended signal via Power Combiner (HP 11667B). A Power Meter (RS NRVS) with a Power Sensor (RS NRV-Z32) is the used to measure the power. Spectrum diagrams are obtained with the help of a Signal Analyzer (RS FSIQ 26). Compensation of the losses in the coaxial cables and power splitter/combiner is done to obtain the final values. The measurement results obtained with the described setup are now given.

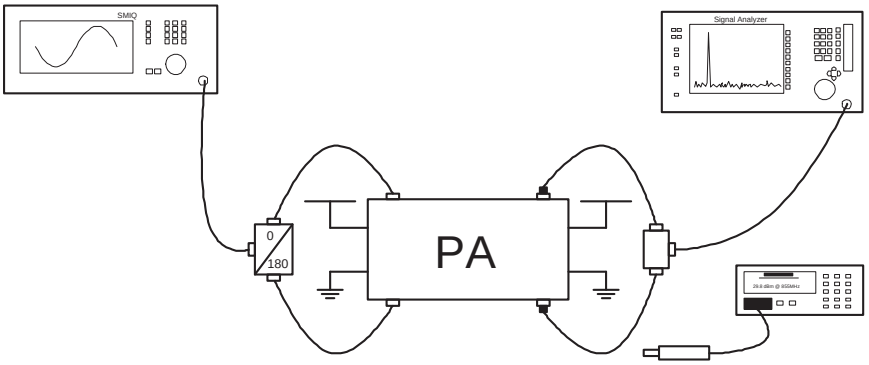
6.5.2 Measurement Results

6.5.2.1 PA-LVMOS2

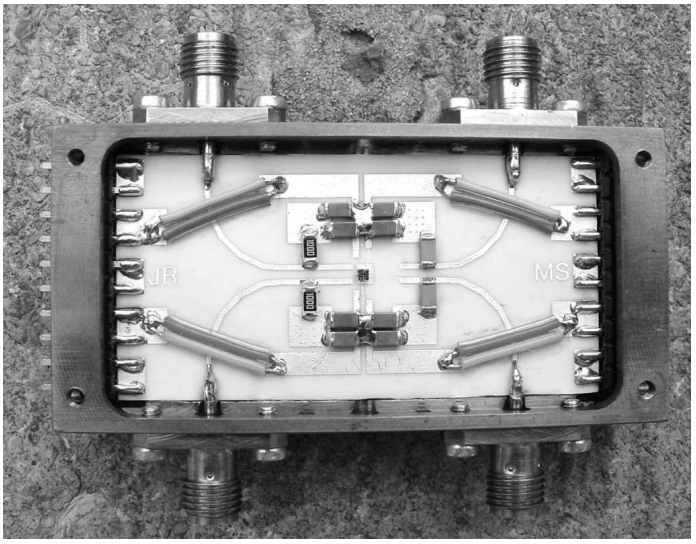
The following measurement results have been obtained without external cooling for the designed 0.35 µm CMOS power amplifier. Circuit losses in the power splitter/combiner and coaxial cables are calibrated. As such, only the losses inside the copper-beryllium box are considered for the final circuit efficiency. The power signal of the signal generator, $P_{IN,RF}$ is equal to +11 dBm.

Fig. 6.7 presents the output spectrum of the converter for one sinewave input signal with 855 MHz. The spectral purity is close to the quality of the signal from the generator itself.

In Fig. 6.8, the output power DE and PAE as a function of the supply voltage is shown. As expected, the dependence of the output power level is approximately proportional to V_{DD}^2. For an output level above 158 mW, or $V_{DD} > 1\ V$, the PAE is always greater than 60 %, reaching a maximum of 66 % (DE=67 %) when the supply voltage is equal to 2.0 V. A maximum output power level equal to 955 mW (29.8 dBm) is measured



(a) The signal coming from the Vector Signal Generator (top left) is divided in two opposite phase signals by the power splitter. This signal reaches the amplifier driving stage inside the copper-beryllium box (center). The differential output signal is then added by a power combiner and measured on a Power Meter (bottom right) with the help of a power sensor. The Signal Analyzer is represented on the top right and is used to obtain the different spectrum graphics.



(b) Copper-Beryllium box photograph of the die mounted on a ceramic substrate.

Figure 6.6: Setup used in the measurements of the PA-LVMOS2 power amplifier.

at 2.26 V. Above a supply voltage of 2 V, nondestructive junction breakdown starts to occur due to impact ionization currents and the PAE decreases because a current larger than predicted by common device models flows. This phenomenon can be alleviated by using a linear capacitor which decreases the maximum drain voltage for the same output power [Chu94], and replacing the output stage transistor junction to have a shared drain junction (Fig. 6.3(b)). The linear capacitor can be implemented using lateral flux which also saves area [Apa02] in comparison with more traditional metal-metal capacitors.

Fig. 6.9 shows measurements versus frequency done with a 2 V supply voltage. Two peaks in efficiency are clearly seen. At 810 MHz and 850 MHz a maximum output power level in excess of 720 mW is measured. The PAE of at least 59 % is maintained over the frequency range from 780 MHz to 850 MHz with a maximum value of 70 % at 790-800 MHz.

The amplifier is designed to be used with a constant envelope modulation scheme. More precisely, it is intended to fulfill the GSM specification limits for spectrum due to modulation. In Fig. 6.10, the amplifier output signal when the input is a modulation type GMSK signal is given. The Gaussian filter used in GMSK is generally specified by its BT product, where B is the 3 dB bandwidth of the filter and T is the symbol duration. In this case a value of 0.3 is used. A signal frequency of 855 MHz is used. As seen, the output spectrum falls within the spectral mask of the GSM specifications.

To verify the proper operation of the differential architecture in suppressing the second harmonic, a measurement up to 2.8 GHz is performed. At the maximum output power level and with an input signal frequency of 855 MHz, the output spectral content in Fig. 6.11 shows the suppression of the second harmonic which is 45 dB below the transmitted signal.

Measurements from chip-to-chip of dies within the same run has shown similar results. The same frequency range of operation is obtained and with similar efficiency. On a whole, the small differences measured can be attributed to uncertainty as a result of the manual bonding of the inductors and expected variations from die-to-die and between the SMD capacitors.

The measurements summarized on Table 6.4 are in very good agreement with the simulation results from sub-section 6.3.3 where a maximum DE of 66 % is expected from simulations and a value of 67 % is measured in the laboratory. This accuracy is well within the circuit modeling uncertainty. Nevertheless, this corroborates the simulation results and supports the statement that all limiting performance parasitics are included during the optimizations. A maximum value of 30 dBm is expected whereas a maximum value of 29.8 dBm is measured. This difference is attributed to the breakdown of the transistor and the inexistence of a design safety margin. Nevertheless, at each measured voltage step, measured current consumption and output power show good agreement with the simulations. The measurement is thus a validation of the PAMPER tool and the methodology it incorporates in this design for the GSM-850 band.

Table 6.4: Summary of the class E power amplifier performance.

Parameter	Unit	PA-LVMOS2
Supply Voltage	V	2.26
Operating Frequency	MHz	855
Maximum DE	%	67
Maximum PAE	%	66
Output Power	mW	955
Die Area	µm × µm	865 × 785
CMOS Technology	—	0.35 µm

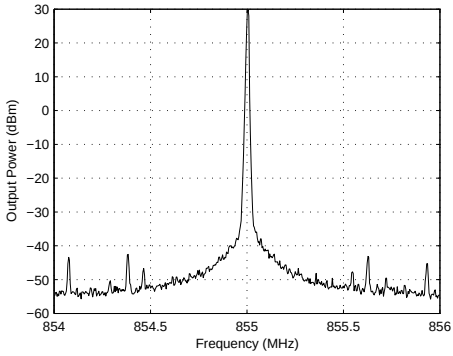


Figure 6.7: Output spectrum for a sinewave input signal measured with a RS FSIQ 26 signal analyzer with a resolution bandwidth of 30 kHz.

6.5.2.2 Effect in Efficiency of the Number of Bonding Wires to the Ground

It is commonly accepted that a larger number of bonding wires in parallel provide a low-inductance connection to the ground board. The quality of on-chip ground is strongly dependent on the quality of the connection to the ceramic substrate and from the ceramic substrate to the box.

The circuit from Fig. 6.5(b) has 12 bond pads that are used for connecting to the ceramic substrate through a bondwire. The two remaining bondpad included for ground connection cannot be used due to space constraints. The supply voltage of the circuit is chosen so that the amplifier works at the maximum output power level. The number of bonding wires is decreased each time by two until only two are left. At each step the output power is measured. At maximum output power level, it has been observed that the output power decreases by less than 100 mW for the case when only two bonding wires connect the circuit to the ground. The circuit becomes nonfunctional if none of the bonding wires are present which is a sign that the connection to the ground is not made from the chip to the ground through the conductive glue attaching the die to the ceramic substrate.

Reducing the number from 12 to 10 has no effect on the maximum output power, a

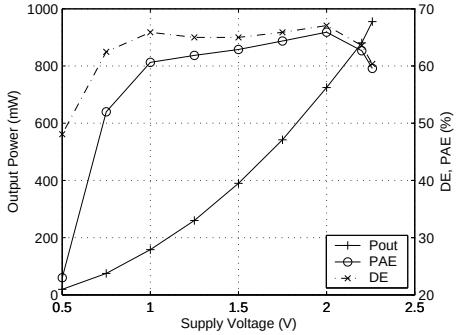


Figure 6.8: Measured output power, DE and PAE over supply voltage.

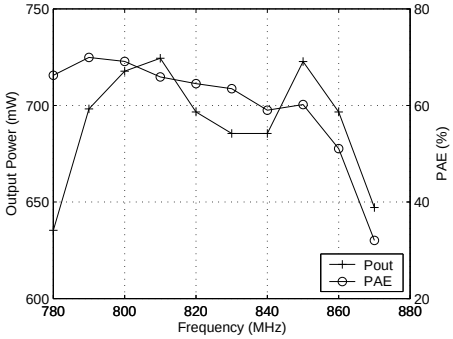


Figure 6.9: Output power, DE and PAE *versus* frequency measured with V_{DD}=2 V.

sign that at these frequencies and output power level, no more than the number of bonding wires already in use are necessary. In addition, when only two bonding wires are connected, a current in the range of 350 mA flows in each bonding wire without observed output power variation at the cent-dB level over time.

6.5.3 Performance Comparison

In the work presented so far, the design, optimization, implementation and measurement of a differential two stage class E power amplifier in CMOS has been given. At this point, a comparison with state-of-the-art is presented.

An overview of state-of-the-art power amplifiers published is given in Table 6.5 for circuits with a different class of operation, frequency and maximum output power. Although the DE and PAE of the presented PA outperforms all other referenced amplifiers, for a better comparison, restriction to a similar frequency band and output power level is necessary, as it minimizes other influences. In conclusion, there are no published results for a circuit implemented in CMOS that for a similar output power level

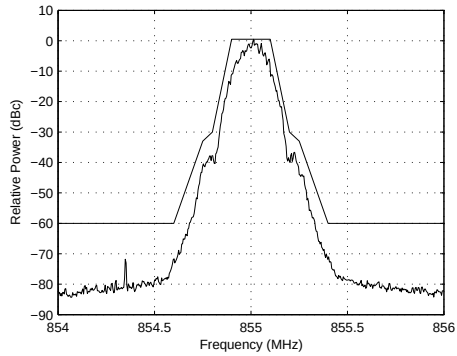


Figure 6.10: GSM spectrum with the output at the maximum output power level at an input signal of 855 MHz and a filter parameter BT=0.3. The resolution bandwidth of the signal analyzer (RS FSIQ 26) is 30 kHz.

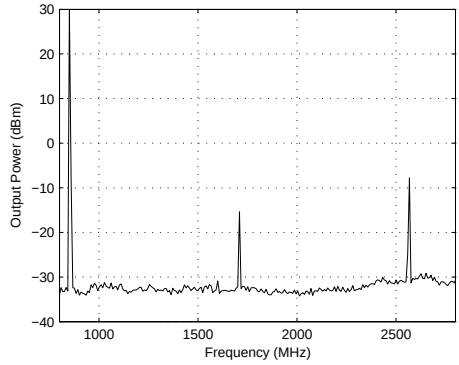


Figure 6.11: Frequency domain output of the differential amplifier from 800 MHz up to 2800 MHz with a 855 MHz input signal frequency.

and frequency of operation show a better performance. This is the result of proper modeling of parasitics and circuit optimization.

Other strong points of the presented PA are

- The power amplifier maintains at least 60 % efficiency for an output power level above 158 mW. This is especially important as the power amplifier is not always working at the maximum power level. This characteristic means that a high efficiency is possible over a wide operational range.

- The design is fully automated after including all parasitics. The sizing and layout of a new power amplifier with similar characteristics can be made within a few days.

Open literature supports somehow that the class E power amplifier is often preferred to the other topologies. However, measured results do not show a clear advantage of one topology over the other. This is especially true if only class E and F are involved. The simplicity of class E is sometimes the reason to use this topology. Another advantage is the relative insensitivity to component variation [Raa78a].

6.6 Conclusion

A class E power amplifier has been presented. The PAMPER tool has been used during the optimization of the circuit for the optimal sizing of one RF power amplifier for maximum efficiency. Careful analysis and inclusion of all parasitics during the design phase has lead to close agreement between simulation and measurements. A maximum value of 66 % for the Drain Efficiency (DE) has been simulated versus 67 % measured. There are no published results for a circuit implemented in CMOS that for a similar output power level and frequency of operation shows a better performance.

The designed amplifier has been manufactured in a commercial 0.35 µm 5M2P CMOS process. It occupies a total area of 0.68 mm^2. Measurement results show that it works at 855 MHz and has a maximum output power of 955 mW at 2.26 V. It has a high efficiency over a broad range of output power levels, the PAE is always greater than 60 % for an output power above 158 mW and a maximum PAE of 66 % is achieved. The above characteristics make the designed circuit suitable for GSM-850 mobile station class 5.

Table 6.5: Overview of CMOS power amplifiers.

	[Su97]	[Tsa02]	[Kuo01]	[Yoo01]	[Fal01]	[Mer02]	[Shi02]	This Work
Supply Voltage	2.5 V	2.0 V	1.8/3.0 V	1.8 V	3.4 V	2.3 V	1.5 V	2.3 V
Frequency	850 MHz	1900 MHz	900 MHz	900 MHz	1750 MHz	700 MHz	1400 MHz	850 MHz
DE	62 %	–	–	46 %	–	–	–	67 %
PAE	42 %	48 %	43 %	41 %	55 %	62 %	49 %	66 %
Output Power	1 W	1 W	1.5 W	900 mW	1096 mW	1 W	304 mW	955 mW
Class	D	E	F	E	AB	E	F	E
Die Area	1.5 mm^2	1.6 mm^2	2 mm^2	4 mm^2	1.1 mm^2	2.64 mm^2	0.43 mm^2	0.68 mm^2
Technology	0.8 µm	0.35 µm	0.2 µm	0.25 µm	0.35 µm	0.35 µm	0.25 µm	0.35 µm

Chapter 7

Conclusions

7.1 Research Overview and Contributions

For each application, complexity and performance must be carefully weighed against power consumption. Even though the desire is to integrate as much as possible, one must remember that this might not always lead to a more power efficient circuit. Each one of the available technological options, when properly combined, offer not only the desired increased battery autonomy but also a lower cost. The research work that is now completed has contributed, amongst others, the following results:

- The analysis, under the same conditions, of six multistage frequency compensation schemes has shown that they have more in common than what could be concluded from the measurement results and the used figures of merit. There are different factors that influence their performance while others can be used to artificially decrease the power consumption. After all, this means that the so far used figures of merit are of no practical use.

  The topology that performs best in terms of a higher Unity Gain Frequency (UGF) can do it at the expense of a poorer transient response or lower slew rate. There is no best compensation topology that suits all applications. Instead, it should be chosen for the particular application in question. A simple variation of the NMC scheme can save up to 50 % of the total power.

  Considering more efficient design criteria, the assumption that the amplifier has a third order Butterworth frequency response in a unity gain feedback configuration should be somehow ignored. Reducing the loop damping ratio parameter (ζ_o) from the typical $1/\sqrt{2}$ can decrease the power consumption by about 25 %.

- The methodology presented to optimize class E power amplifiers for maximum efficiency has been validated and has shown good agreement with measurements. It has been demonstrated that it is impossible to account for all circuit parasitics using only equations and still maximize efficiency. The methodology has been incorporated in a tool to automatically size CMOS power amplifiers. It has been shown to be sufficiently flexible to optimize different variations to the generalized class E power amplifier.

  High voltage in CMOS has been explored as a way to achieve a high output power level. It has been shown that high breakdown voltage is possible without

any change in the design process flow. The device can be used where the requirement for a high voltage is mandatory and the increased input resistance and capacitance have minimal influence.

The automated design of a high-efficiency class E power amplifier in CMOS has been presented. It has demonstrated the necessity to include not only all device parasitics but also board parasitics for maximizing the Drain Efficiency (DE).

To conclude: the analysis, simulation, optimization and measurement of high efficiency amplifiers in CMOS has been discussed in this thesis. A novel frequency compensation topology for three stage operational amplifiers has been presented. It has been shown that it is possible to achieve high breakdown voltage in commercial CMOS technologies without any change in the process flow. Finally, a differential two stage class E power amplifier has been optimized and has shown good agreement with simulations and high efficiency.

7.2 Possibilities for Future Work

Future research can continue, either incrementally from the point where the presented study has left off or by pursuing one of the following suggestions. High performances can be obtained in a relatively inexpensive technology like CMOS but the main advantages arise from the possibility of integrating a full system on a chip. Nevertheless, I would like to stress the fact that the most difficult thing in any design like the ones presented in this work is measuring. Without measurement results our work is unfinished but this is sometimes a disregarded issue.

Advancements can come both from integration, where the final product design is eased, or by increasing the performance of a single block. Several issues can be addressed and different research paths can be followed of which the following seem interesting to dedicate more time to.

- The design of multistage amplifiers, although a field with a considerable amount of work done by the academia, still has to be accepted in the industry. Most of the designs presented so far include only the amplifier. Without more complex systems showing good performance, it is likely that the use of the well established NMC topology is to continue. Other characteristics such as PSRR and CMRR have increased importance in systems where the injected digital switching noise has results in performance degradation. Equally important is the amplifier intrinsic noise. All the previous characteristics eventually limit the circuit performance. Comparing each topology on a theoretical level makes it possible to determine the advantages of each amplifier.
- Differential circuits have advantages in terms of better noise immunity and increased dynamic range. However, today's circuits commonly use a single ended antenna. Hence differential signal to single ended conversion must exist. Higher power can thus be generated from the battery if this conversion can be made more efficient. This might lie in a similar approach to the work presented in [Pau03] where it is even possible to perform this conversion on-chip instead of on-board.

- The measured performance of the class E power amplifier presented in this work agrees well with the performance expected from simulations. This justifies that the RF performance can be precisely determined beforehand. This automated sizing can be extended to include the auxiliary circuits necessary by an autonomous PA: GSM burst shaping, power control and input-output matching networks.

- Technologies including the Gallium Arsenide (GaAs), bipolar and Silicon Germanium (SiGe) are predominant in implemented power amplifiers mainly resulting from the required output power level and power added efficiency necessary for the application. The prototype presented in this work and in other publications show that CMOS can be used as well, although only for the less stringent mobile station type. Due to the maximum voltage decrease for more advanced technologies, new approaches must be derived. In this field, further investigation is worthwhile.

- The level of performance for the PA presented in this work were only possible because an automated design and optimization environment has been used. The simulation programs and the numerous design considerations could be used and extended... without having to reinvent the wheel once more.

- Parasitic losses in the PA strongly limit the maximum power efficiency attainable. More advanced CMOS technologies do not necessarily offer passives with a higher quality factor than those obtained by a much cheaper thin-film technology. This is especially true in the case of the inductors. Considerable work could be done in merging the best of each technology to obtain a simultaneously higher performance and a lower cost.

- The power amplifier presented in this thesis cannot be used for applications that also transmit information by varying the amplitude on top of the frequency. Different linearization schemes exist. Whether a linearization scheme or the use of a linear power amplifier is advantageous is yet to be seen. Measurement results do not show any clear trend.

- It is expected that in the near future full transceivers including the PA will be available in one chip implemented in CMOS. Once the knowledge for each of the different blocks has been gathered, full integration is an interesting technological achievement to pursue.

Now that I'm finishing my studies I would like to think that the reasons that made me come here are still valid. I wish you all the best!

List of Publications

- J. Ramos and M. Steyaert, "Three Stage Amplifier With Positive Feedback Compensation Scheme," in *Proceedings Custom Integrated Circuits Conference (CICC)*, Orlando, Florida, May 12–15, 2002, pp. 333–336.

- J. Ramos, X. Peng, M. Steyaert, and W. Sansen, "A Comparative Study of Three Stage Amplifier Frequency Compensation," in *Proceedings European Conference on Circuit Theory and Design (ECCTD)*, vol. III, Cracow, Poland, Sept. 1–4, 2003, pp. 385–388.

- J. Ramos, X. Peng, M. Steyaert, and W. Sansen, "Three Stage Amplifier Frequency Compensation," in *Proceedings European Solid-State Circuits Conference (ESSCIRC)*, Estoril, Portugal, Sept. 16–18, 2003, pp. 365–368.

- J. Ramos and M. Steyaert, "STI/LOCOS compatible LDMOS structure in standard CMOS," *IEE Electronics Letters*, vol. 39, no. 19, pp. 1417–1419, Sept. 2003.

- J. Ramos and M. Steyaert, "High voltage devices for RF power amplifiers: an advantage?" in *Proceedings Workshop on Advances in Analog Circuit Design (AACD)*, Montreux, Switzerland, Apr. 6–8, 2004.

- J. Ramos, K. Francken, G. Gielen, and M. Steyaert, "Knowledge- and Optimization-Based Design of RF Power Amplifiers," in *Proceedings IEEE International Symposium on Circuits and Systems (ISCAS)*, Vancouver, Canada, May 23–26, 2004.

- J. Ramos and M. Steyaert, "Positive Feedback Frequency Compensation for Low-Voltage Low-Power Three Stage Amplifier," *IEEE Transactions on Circuits and Systems— I: Regular Papers*, vol. 51, no. 10, pp. 1967–1974, Oct. 2004.

- J. Ramos and M. Steyaert, *High voltage devices for RF power amplifiers: an advantage?* Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004, ch. 9 in *Analog Circuit Design (Sensor and Actuator Interface Electronics, Integrated High-Voltage Electronics and Power Management, Low-Power and High-Resolution ADC's)*, edited by *Johan H. Huijsing, Michiel Steyaert and Arthur van Roermund*, pp. 177–200.

- J. Ramos and M. Steyaert, "Design of a Class E Power Amplifier with LDMOS in Standard CMOS," *Analog Integrated Circuits and Signal Processing*, accepted for publication.

- J. Ramos, K. Francken, G. G. E. Gielen, and M. S. J. Steyaert, "An Efficient, Fully Parasitic-Aware Power Amplifier Design Optimization Tool," *IEEE Transactions on Circuits and Systems— I: Regular Papers*, accepted for publication.

Bibliography

[3GPP] 3^{rd} *Generation Partnership Project*, TS 45.005 [Online]. Available: `http://www.3gpp.org/ftp/Specs/html-info/45005.htm`.

[Ann00] A.-J. Annema, G. Geelen, and P. de Jong, "5.5V Tolerant I/O in a 2.5V 0.25µm CMOS Technology", In *Proc. Custom Integrated Circuits Conference (CICC)*, pages 417–420, May 2000.

[Apa02] R. Aparicio and A. Hajimiri, "Capacity Limits and Matching Properties of Integrated Capacitors", *IEEE J. of Solid-State Circuits*, 37(3):384–393, March 2002.

[App79] J. A. Appels and H. M. J. Vaes, "High Voltage Thin Layer Devices (RESURF Devices)", In *Proc. International Electron Devices Meeting (IEDM)*, pages 238–241, December 1979.

[Avr89] C. P. Avratoglou, N. C. Voulgaris, and F. I. Ioannidou, "Analysis and Design of a Generalized Class E Tuned Power Amplifier", *IEEE Trans. on Circuits and Systems*, 36(8):1068–1079, August 1989.

[Bak02] B. Bakeroot, M. Vermandel, P. Moens, J. Doutreloigne, and D. Bolognesi, "Cost Effective Implementation of a 90 V RESURF P-type Drain Extended MOS in a 0.35 μm Based Smart Power Technology", In *Proc. European Solid-State Device Research Conference (ESSDERC)*, pages 291–294, September 2002.

[Bal98] H. Ballan and M. Declercq, *High Voltage Devices and Circuits in Standard CMOS Technologies*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[Bul90] K. Bult and G. J. G. M. Geelen, "A Fast-Settling CMOS Op Amp for SC Circuits with 90-dB DC Gain", *IEEE J. of Solid-State Circuits*, 25(6):1379–1384, December 1990.

[Cao03] Y. Cao, R. A. Groves, X. Huang, N. D. Zamdmer, J.-O. Plouchart, R. A. Wachnik, Tsu-Jae King, and C. Hu, "Frequency–Independent Equivalent–Circuit Model for On–Chip Spiral Inductors", *IEEE J. of Solid-State Circuits*, 38(3):419–426, March 2003.

[Car96] R. Carley, G. Gielen, R. Rutenbar, and W. Sansen, "Synthesis Tools for Mixed–Signal ICs: progress on frontend and backend strategies", In *Proc. Design Automation Conference (DAC)*, 33rd, pages 298–303, June 3–7, 1996, Las Vegas, USA.

[Cas04] H. Casier, P. Moens, and K. Appeltans, "Technology Considerations for Automotive", In *Proc. European Solid-State Circuits Conference (ESSCIRC)*, pages 37–41, September 2004.

[Cha01] C. K. T. Chan and C. Toumazou, "Design of a Class E Power Amplifier With Non–Linear Transistor Output Capacitance and Finite DC–Feed Inductance", In *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, volume 1, pages 129–132, Sydney, Australia, May 6–9, 2001.

[Che99] J. F. Chen, J. Tao, P. Fang, and C. Hu, "Performance and Reliability Comparison Between Asymmetric and Symmetric LDD Devices and Logic Gates", *IEEE J. of Solid-State Circuits*, 34(3):367–371, March 1999.

[Cho02] K. Choi, D. J. Allstot, and S. Kiaei, "Parasitic-Aware Synthesis of RF CMOS Switching Power Amplifiers", In *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, volume 1, pages 269–272, Scottsdale, Arizona, May 26–29, 2002.

[Chu94] M. J. Chudobiak, "The Use of Parasitic Nonlinear Capacitors in Class E Amplifiers", *IEEE Trans. on Circuits and Systems—Part I: Fundamental Theory and Applications*, 41(12):941–944, December 1994.

[Dor04] R. C. Dorf and R. H Bishop, *Modern Control Systems*, Prentice–Hall, 2004.

[Ehw01] K.-E. Ehwald, B. Heinemann, W. Roepke, W. Winkler, H. Rücker, F. Fuernhammer, D. Knoll, R. Barth, B. Hunger, H. E. Wulf, R. Pazirandeh, and N. Ilkov, "High Performance RF LDMOS Transistors with 5nm Gate Oxide in a 0.25μm SiGe:C BiCMOS Technology", In *Proc. International Electron Devices Meeting (IEDM)*, pages 40.4.1–40.4.4, December 2001.

[Eldo] *Eldo User's Manual, Document Number 310101 for Software Version v5.3_1.1*, Mentor Graphics®Corp. [Online]. Available: `http://www.mentor.com/eldo/overview.html`.

[Esc92] R. G. H. Eschauzier, L. P. T. Kerklaan, and J. H. Huijsing, "100-MHz 100-dB Operational Amplifier with Multipath Nested Miller Compensation Structure", *IEEE J. of Solid-State Circuits*, 27(12):1709–1717, December 1992.

[Fal01] C. Fallesen and P. Asbeck, "A 1W CMOS Power Amplifier for GSM-1800 with 55% PAE", In *IEEE MTT-S International Microwave Symposium Digest (IMS)*, volume 2, pages 911–914, May 2001.

[Fon91] J. Fonderie and J. H. Huijsing, "Operational Amplifier with 1-V Rail-to-Rail Multipath-Driven Output Stage", *IEEE J. of Solid-State Circuits*, 26(12):1817–1824, December 1991.

[Fra88] A. Fraser, R. Gleason, and E. W. Strid, "GHz On-Silicon-Wafer Probing Calibration Methods", In *Bipolar Circuits and Technology Meeting*, pages 154–157, September 1988.

[Fra03] K. Francken, *A Framework for Analyzing and Synthesizing High-Level and Circuit-Level Analog Blocks*, PhD thesis, K. U. Leuven, Belgium, September 2003.

[Gri93] C. Van Grieken and W. Sansen, "A testset for automatic characterisation of opamps in the frequency domain", In *IEEE International Conference on Microelectronic Test Structures, Barcelona, Spain*, volume 6, pages 83–88, March 1993.

[Gro67] A. S. Grove, O. Leistiko, Jr., and W. W. Hooper, "Effect of Surface Fields on the Breakdown Voltage of Planar Silicon p-n Junctions", *IEEE Trans. on Electron Devices*, ED-14(3):157–162, March 1967.

[Gro01] G. V. Groeseneken, "Hot Carrier Degradation and ESD in Submicrometer CMOS Technologies: How Do They Interact?", *IEEE Trans. on Device and Materials Reliability*, 1(1):23–32, March 2001.

[Gup01] R. Gupta, B. M. Ballweber, and D. J. Allstot, "Design and Optimization of CMOS RF Power Amplifiers", *IEEE J. of Solid-State Circuits*, 36(2):166–174, February 2001.

[ITRS] *International Technology Roadmap for Semiconductors* [Online]. Available: `http://public.itrs.net/`.

[Kim99] J. Kim, S.-G. Kim, T. M. Roh, H. S. Park, J.-G. Koo, and D. Y. Kim, "Characteristics of P-channel SOI LDMOS Transistor with Tapered Field Oxides", *ETRI Journal*, 21(3):22–28, September 1999.

[Kuo01] T. C. Kuo and B. B. Lusignan, "A 1.5 W Class-F RF Power Amplifier in 0.2µm CMOS Technology", In *Proc. International Solid-State Circuits Conference (ISSCC)*, volume XLIV, pages 154–155, February 2001.

[Lak94] K. R. Laker and W. M. C. Sansen, *Design of Analog Integrated Circuits and Systems*, McGraw–Hill, 1994.

[Lee98] Y.-G. Lee, S.-K. Yun, and H.-Y. Lee, "Novel high-Q bondwire inductor for MMIC", In *Proc. International Electron Devices Meeting (IEDM)*, pages 548–551, December 1998.

[Lee03a] H. Lee and P. K. T. Mok, "Active-Feedback Frequency Compensation for Low-Power Multistage Amplifiers", *IEEE J. of Solid-State Circuits*, 38(3):511–520, March 2003.

[Lee03b] H. Lee, K. N. Leung, and P. K. T. Mok, "A Dual-Path Bandwidth Extension Amplifier Topology With Dual-Loop Parallel Compensation", *IEEE J. of Solid-State Circuits*, 38(10):1739–1744, October 2003.

[Leu00] K. N. Leung, P. K. T. Mok, W. H. Ki, and J. K. O. Sin, "Three-Stage Large Capacitive Load Amplifier with Damping-Factor-Control Frequency Compensation", *IEEE J. of Solid-State Circuits*, 35(2):221–230, February 2000.

[Lio72] M.-L. Liou, "Exact Analysis of Linear Circuits Containing Periodically Operated Switches with Applications", *IEEE Trans. on Circuit Theory*, 19(10):146–154, March 1972.

[Lov94] D. Lovelace, J. Costa, and N. Camilleri, "Extracting Small-Signal Model Parameters of Silicon MOSFET Transistors", In *IEEE MTT-S International Microwave Symposium Digest (IMS)*, pages 865–868, May 1994.

[Man90] J. C. Mandojana, K. J. Herman, and R. E. Zulinski, "A Discrete/Continuous Time-Domain Analysis of a Generalized Class E Amplifier", *IEEE Trans. on Circuits and Systems*, 37(8):1057–1060, August 1990.

[Mar98] A. Marques, Y. Geerts, M. Steyaert, and W. Sansen, "Settling Time Analysis of Third Order Systems", In *International Conference on Electronics, Circuits and Systems (ICECS)*, pages 2153–2156, September 1998.

[Mer02] K. Mertens and M. Steyaert, "A 700–MHz 1–W Fully Differential CMOS Class–E Power Amplifier", *IEEE J. of Solid-State Circuits*, 37(2):137–141, February 2002.

[Moe04] P. Moens, G. Van den Bosch, and G. Groeseneken, "Hot-Carrier Degradation Phenomena in Lateral and Vertical DMOS Transistors", *IEEE Trans. on Electron Devices*, 51(4):623–628, April 2004.

[Ng99] H. T. Ng, R. M. Ziazadeh, and D. J. Allstot, "A Multistage Amplifier Technique with Embedded Frequency Compensation", *IEEE J. of Solid-State Circuits*, 34(3):339–347, March 1999.

[Nik98] A. M. Niknejad and R. G. Meyer, "Analysis, Design, and Optimization of Spiral Inductors and Transformers for Si RF ICs", *IEEE J. of Solid-State Circuits*, 33(10):1470–1481, October 1998.

[Par87] Z. Parpia, C. A. T. Salama, and R. A. Hadaway, "Modeling and Characterization of CMOS-Compatible High-Voltage Device Structures", *IEEE Trans. on Electron Devices*, ED-34(11):2335–2343, November 1987.

[Pau03] S. A. Paul and T. J. Dupuis, "Power Amplifier Circuitry and Method Using an Inductance Coupled to Power Amplifier Switching Devices" U.S. Patent 6,549,071, April 15, 2003

[Pen02] X. Peng and W. Sansen, "Nested feed-forward Gm-stage and nulling Resistor plus Nested-Miller Compensation for Multistage Amplifiers", In *Proc. Custom Integrated Circuits Conference (CICC)*, pages 329–332, May 2002.

[Pen04] X. Peng and W. Sansen, "Transconductance with Capacitance Feedback Compensation for Multistage Amplifiers", In *Proc. European Solid-State Circuits Conference (ESSCIRC)*, pages 143–146, September 2004.

[Per93] S. Pernici, S. Nicollini, and R. Castello, "A CMOS Low-Distortion Fully Differential Power Amplifier with double Nested Miller Compensation", *IEEE J. of Solid-State Circuits*, 28(7):758–763, July 1993.

[Phi99] *EZ-HV™ LDMOS*, Philips Electronics N.V. [Online]. Available: `http://www.semiconductors.philips.com/news/backgrounders/bg9901/`.

[Por93] R. M. Porter, Jr. and M. L. Mueller, "High power switch-mode radio frequency amplifier method and apparatus", U.S. Patent 5,187,580, February 16, 1993.

[Raa78a] F. H. Raab, "Effects of Circuit Variations on the Class E Tuned Power Amplifier", *IEEE J. of Solid-State Circuits*, 13(2):239–247, April 1978.

[Raa78b] F. H. Raab and N. O. Sokal, "Transistor Power Losses in the Class E Tuned Power Amplifier", *IEEE J. of Solid-State Circuits*, 13(6):912–914, December 1978.

[Ram03a] J. Ramos, X. Peng, M. Steyaert, and W. Sansen, "Three Stage Amplifier Frequency Compensation", In *Proc. European Solid-State Circuits Conference (ESSCIRC)*, pages 365–368, Estoril, Portugal, September 16–18, 2003.

[Ram03b] J. Ramos and M. Steyaert, "STI/LOCOS compatible LDMOS structure in standard CMOS", *IEE Electronics Letters*, 39(19):1417–1419, September 2003.

[Ram04] J. Ramos, K. Francken, G. Gielen, and M. Steyaert, "Knowledge- and Optimization-Based Design of RF Power Amplifiers", In *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, Vancouver, Canada, May 23–26, 2004.

[Shi02] A. Shirvani, D. K. Su, and B. A. Wooley, "A CMOS RF Power Amplifier With Parallel Amplification for Efficient Power Control", *IEEE J. of Solid-State Circuits*, 37(6):684–693, June 2002.

[Sim99] W. Simbürger, H.-D. Wohlmuth, P. Weger, and A. Heinz, "A Monolithic Transformer Coupled 5-W Silicon Power Amplifier with 59% PAE at 0.9 GHz", *IEEE J. of Solid-State Circuits*, 34(12):1881–1892, December 1999.

[Smi90] G. H. Smith and R. E. Zulinski, "An Exact Analysis of Class E Amplifiers with Finite DC-feed Inductance at Any Output Q", *IEEE Trans. on Circuits and Systems*, 37(4):530–534, April 1990.

[Sok75] N. O. Sokal and A. D. Sokal, "Class E-A New Class of High-Efficiency Tuned Single-Ended Switching Power Amplifiers", *IEEE J. of Solid-State Circuits*, 10(3):168–176, June 1975.

[Sow95] T. Sowlati, C. A. T. Salama, J. Sitch, G. Rabjohn, and D. Smith, "Low Voltage, High Efficiency GaAs Class E Power Amplifiers for Wireless Transmitters", *IEEE J. of Solid-State Circuits*, 30(10):1074–1080, October 1995.

[Sow03] T. Sowlati and D. M. W. Leenaerts, "A 2.4-GHz 0.18 μm CMOS Self-Biased Cascode Power Amplifier", *IEEE J. of Solid-State Circuits*, 38(8):1318–1324, August 2003.

[Ste00] M. S. J. Steyaert, J. Janssens, B. de Muer, M. Borremans, and N. Itoh, "A 2-V CMOS Cellular Transceiver Front-End", *IEEE J. of Solid-State Circuits*, 35(12):1895–1907, December 2000.

[Sto95] R. Storn and K. Price, "Differential Evolution – A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces", Technical report, Technical Report TR-95-012, ICSI, March 1995.

[Sto96] R. Storn, "On the Usage of Differential Evolution for Function Optimization", In *NAFIPS*, pages 519–523, 1996.

[Su97] D. Su and W. McFarland, "A 2.5-V, 1-W Monolithic CMOS RF Power Amplifier", In *Proc. Custom Integrated Circuits Conference (CICC)*, volume May, pages 189–192, 1997.

[tE97] F. O. 't Eynde and C. Zorio, "An EEPROM in a standard CMOS technology", In *Proc. European Solid-State Circuits Conference (ESSCIRC)*, pages 264–267, September 1997.

[Tsa02] K.-C. Tsai and P. R Gray, "A 1.9-GHz, 1-W CMOS Class-E Power Amplifier for Wireless Communications", *IEEE J. of Solid-State Circuits*, 34(7):962–970, July 2002.

[Tsi99] Y. Tsividis, *Operation and Modeling of the MOS Transistor*, McGraw–Hill, New York, 2nd edition, 1999.

[Ves02] L. Vestling and J. Ankarcrona, "A General Small-Signal Series Impedance Extraction Technique", *IEEE Microwave and Guided Wave Letters*, 12(7):1531–1309, July 2002.

[Yan90] H. C. Yang and D. J. Allstot, "Considerations for Fast Settling Operational Amplifiers", *IEEE Trans. on Circuits and Systems*, 37(3):326–334, March 1990.

[Yoo01] C. Yoo and Q. Huang, "A Common-Gate Switched 0.9-W Class-E Power Amplifier with 41% PAE in 0.25-μm CMOS", *IEEE J. of Solid-State Circuits*, 35(5):823–830, May 2001.

[You97] F. You, S. H. K. Embabi, and E. Sanchez-Sinencio, "Multistage Amplifier Topologies with Nested $G_m - C$ Compensation", *IEEE J. of Solid-State Circuits*, 32(12):2000–2011, December 1997.

[Zul87] R. Zulinski and J. Steadman, "Class E Power Amplifiers and Frequency Multipliers with finite DC-Feed Inductance", *IEEE Trans. on Circuits and Systems*, 34(9):1074–1087, September 1987.