Patrick Reynaert
Michiel Steyaert

ACSP
Analog Circuits And Signal Processing

RF Power Amplifiers for Mobile Communications

Springer

RF POWER AMPLIFIERS FOR MOBILE COMMUNICATIONS

ANALOG CIRCUITS AND SIGNAL PROCESSING SERIES

Consulting Editor: Mohammed Ismail. *Ohio State University*

Related Titles:

IQ CALIBRATION TECHNIQUES FOR CMOS RADIO TRANCEIVERS
Chen, Sao-Jie, Hsieh, Yong-Hsiang
ISBN: 1-4020-5082-8

CMOS CURRENT-MODE CIRCUITS FOR DATA COMMUNICATIONS
Yuan, Fei
ISBN: 0-387-29758-8

ADVANCED DESIGN TECHNIQUES FOR RF POWER AMPLIFIERS
Rudiakova, A.N., Krizhanovski, V.
ISBN 1-4020-4638-3

CMOS CASCADE SIGMA-DELTA MODULATORS FOR SENSORS AND TELECOM
del Río, R., Medeiro, F., Pérez-Verdú, B., de la Rosa, J.M., Rodríguez-Vázquez, A.
ISBN 1-4020-4775-4

Titles in former series International Series in Engineering and Computer Science:

SIGMA DELTA A/D CONVERSION FOR SIGNAL CONDITIONING
Philips, K., van Roermund, A.H.M.
Vol. 874, ISBN 1-4020-4679-0

CALIBRATION TECHNIQUES IN NYQUIST A/D CONVERTERS
van der Ploeg, H., Nauta, B.
Vol. 873, ISBN 1-4020-4634-0

ADAPTIVE TECHNIQUES FOR MIXED SIGNAL SYSTEM ON CHIP
Fayed, A., Ismail, M.
Vol. 872, ISBN 0-387-32154-3

WIDE-BANDWIDTH HIGH-DYNAMIC RANGE D/A CONVERTERS
Doris, Konstantinos, van Roermund, Arthur, Leenaerts, Domine
Vol. 871 ISBN: 0-387-30415-0

METHODOLOGY FOR THE DIGITAL CALIBRATION OF ANALOG CIRCUITS AND SYSTEMS: WITH CASE STUDIES
Pastre, Marc, Kayal, Maher
Vol. 870, ISBN: 1-4020-4252-3

HIGH-SPEED PHOTODIODES IN STANDARD CMOS TECHNOLOGY
Radovanovic, Sasa, Annema, Anne-Johan, Nauta, Bram
Vol. 869, ISBN: 0-387-28591-1

LOW-POWER LOW-VOLTAGE SIGMA-DELTA MODULATORS IN NANOMETER CMOS
Yao, Libin, Steyaert, Michiel, Sansen, Willy
Vol. 868, ISBN: 1-4020-4139-X

DESIGN OF VERY HIGH-FREQUENCY MULTIRATE SWITCHED-CAPACITOR CIRCUITS
U, Seng Pan, Martins, Rui Paulo, Epifânio da Franca, José
Vol. 867, ISBN: 0-387-26121-4

DYNAMIC CHARACTERISATION OF ANALOGUE-TO-DIGITAL CONVERTERS
Dallet, Dominique; Machado da Silva, José (Eds.)
Vol. 860, ISBN: 0-387-25902-3

ANALOG DESIGN ESSENTIALS
Sansen, Willy
Vol. 859, ISBN: 0-387-25746-2

DESIGN OF WIRELESS AUTONOMOUS DATALOGGER IC'S
Claes and Sansen
Vol. 854, ISBN: 1-4020-3208-0

MATCHING PROPERTIES OF DEEP SUB-MICRON MOS TRANSISTORS
Croon, Sansen, Maes
Vol. 851, ISBN: 0-387-24314-3

LNA-ESD CO-DESIGN FOR FULLY INTEGRATED CMOS WIRELESS RECEIVERS
Leroux and Steyaert
Vol. 843, ISBN: 1-4020-3190-4

SYSTEMATIC MODELING AND ANALYSIS OF TELECOM FRONTENDS AND THEIR BUILDING BLOCKS
Vanassche, Gielen, Sansen
Vol. 842, ISBN: 1-4020-3173-4

LOW-POWER DEEP SUB-MICRON CMOS LOGIC SUB-THRESHOLD CURRENT REDUCTION
van der Meer, van Staveren, van Roermund
Vol. 841, ISBN: 1-4020-2848-2

WIDEBAND LOW NOISE AMPLIFIERS EXPLOITING THERMAL NOISE CANCELLATION
Bruccoleri, Klumperink, Nauta
Vol. 840, ISBN: 1-4020-3187-4

RF POWER AMPLIFIERS FOR MOBILE COMMUNICATIONS

by

Patrick Reynaert

Katholieke Universiteit Leuven, Belgium

and

Michiel Steyaert

Katholieke Universiteit Leuven, Belgium

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-5116-6 (HB)
ISBN-13 978-1-4020-5116-6 (HB)
ISBN-10 1-4020-5117-4 (e-book)
ISBN-13 978-1-4020-5117-3 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Printed in the Netherlands.

Contents

Preface

Since the early nineties, mobile communication systems have entered our daily life. The main reason for this unprecedented wireless revolution, is the high integration level that can be achieved with CMOS. This allowed the integration of enormous amounts of digital functionality on one single chip. As such, it became feasible to introduce digital coding and digital signal processing in wireless communication systems which resulted in the powerful mobile networks of today. Another reason for the successful wireless development, is the low cost of the user equipment which in turn is due to the low cost of CMOS.

The evolution of mobile communication systems continues and today, telephony, television, internet, e-mail, radio broadcast, ... are all being merged together. They have become services, rather than stand-alone systems, that users can access through one single mobile device. Putting all this functionality into one small mobile device, at a reasonable cost, requires a higher integration level. For the comfort of the user, it also requires an increased battery lifetime and thus a low power consumption.

Mobile phones and wireless network equipment both require a power amplifier to amplify the radio signal before it can be transmitted through the antenna. The power amplifier should amplify the radio signal to the desired output level, as accurately as possible, but without consuming too much power itself as this would reduce the battery lifetime. In other words, besides the required output power, the power amplifier should have sufficient linearity and a high efficiency.

The overall goal of this work is to provide circuit design techniques that allow the reader to design a power amplifier that (1) meets the output power and linearity requirements of a mobile communication system, (2) has a high efficiency and gain, (3) is integrated in CMOS and (4) requires no expensive off-chip components. To achieve this goal, a theoretical foundation is developed first. It investigates the consequences of CMOS integration with respect

to power amplification. Impedance transformation and power combining are crucial to achieve sufficient output power in a low-voltage CMOS technology and is subsequently covered. Combining efficiency and linearity leads to the development of a polar modulation architecture.

To validate the developed theory, two amplifiers were successfully designed, fabricated and measured. The first amplifier is designed for GSM-EDGE in a $0.18\,\mu\mathrm{m}$ CMOS technology and operates at $1.75\,\mathrm{GHz}$. To efficiently amplify the non-constant envelope EDGE signal, a polar modulation architecture was developed. The amplifier achieves a peak output power of $27\,\mathrm{dBm}$ with an overall efficiency of $34\,\%$. When transmitting EDGE signals, the amplifier achieves an overall efficiency of $22\,\%$ at an output power of $23.8\,\mathrm{dBm}$ or $240\,\mathrm{mW}$. The second amplifier is integrated in a $0.13\,\mu\mathrm{m}$ CMOS technology, operates at $2.45\,\mathrm{GHz}$, is fully differential and has a single-ended output. To achieve sufficient output power in the $0.13\,\mu\mathrm{m}$ technology, a lattice-type LC power combining network is integrated on the CMOS chip, that allows the parallel connection of four amplifiers. The amplifier achieves an output power of $23\,\mathrm{dBm}$ with an overall efficiency of $29\,\%$. The power combining network allows for both a discrete power control and an efficiency improvement.

Chapter 1

INTRODUCTION

1.1 Wireless Communication

Wireless and mobile communication systems have become ubiquitous in our daily life and it has changed our society and our way of living in a tremendous way. For sure, the desire for mobility and communication are natural human properties and society is always changing. But it is only recently that the possibilities have taken a steep flight upwards, it is only recently that people are able to *see and hear* things that are not nearby. Thanks to the invention of wireless (radio-)communication in the nineteenth century and television in the twentieth century, humanity has extended its own sensorial capabilities; an unprecedented change.

It is generally accepted that mobile communication was born in 1897, when Guglielmo Marconi gained a patent for his wireless telegraph[1]. In those days, radio-communication was merely transmitting the dots and dashes of the Morse code. Slowly, communication equipment enhanced and radio-communication was used for navigation and to keep contact with ships and airplanes. But still, there was as strong need for new technologies that could manipulate, amplify and decode the weak electronic radio signals.

The invention of the vacuum tube in 1906 by Lee De Forest made it possible to amplify and process the received radio signals. Armstrong was the first to develop radio receivers and is well-known for his invention of the regenerative receiver in 1913 , the invention of the super-heterodyne receiver during World War I and his successful demonstration of frequency modulation in 1933. Although a brilliant radio amateur, Armstrong eventually lost all his patents.

[1] Nikola Tesla is now credited with having inventing modern radio; the Supreme Court overturned Marconi's patent in 1943 in favor of Tesla.

The invention of the transistor in 1947 by John Bardeen, Walter Brattain and William Shockley resulted in a tremendous size and weight reduction of most electronic equipment and increased the reliability; the *transistor radio* is a nice example of this. Another important technological step was made in 1958, when Jack Kilby invented the integrated circuit. He conceived and built the first electronic circuit in which all of the components, both active and passive, were fabricated in a single piece of semiconductor material half the size of a paper clip. The successful laboratory demonstration of that first simple microchip on September 12, 1958, made history.

Apart from the wireless evolution, the invention of the integrated circuit also enabled engineers to design large digital systems at a relative low cost. This in turn gave rise to a rapid growth of the number of personal computers and a shift from the old telephone systems to digital networks. The Internet, invented in 1973 and laid out in 1983 and the World Wide Web, developed in 1989, were the logical consequences.

Since the introduction of GSM in Europe in 1991, the mobile telephony market is growing rapidly. In 2003, the number of global mobile subscribers exceeded the number of fixed lines for the first time and it is expected that by 2010, there will be over 23 billion individual wireless subscribers worldwide [Deut04]. Besides the mobile phone networks, the development of wireless data networks, like W-LAN and Bluetooth, followed quickly and they became very popular to make a wireless link between all kinds of devices and for wireless internet access.

Today, telephony, television, internet, e-mail, radio broadcast, . . . are all being merged together. They have become services, rather than stand-alone systems, that users can access through one single mobile device. The design of such a single mobile device requires a high level of integration and miniaturization, a low power consumption and a low production cost. This is the point were CMOS pops up.

1.2 CMOS Technology and Scaling

1.2.1 Moore's Law

Device scaling aims to integrate more transistors per unit area. This requires less silicon area for the same functionality and a lower production cost. Gordon Moore has predicted this trend already in 1965 [Moor65]. He observed that the number of transistors on a single chip doubles every year. In 1975, he updated his prediction to once every two years. While originally intended as a rule of thumb in 1965, it has become the guiding principle for the industry to deliver ever-more-powerful semiconductor chips at proportionate decreases in cost. To achieve such high integration levels, the size of each individual transistor has to shrink, and to reduce the cost, the yield has to go up.

In 1962, Steven Hofstein and Fredric Heiman at the RCA research laboratory in Princeton, New Jersey, invented a new family of devices called metal-oxide semiconductor field-effect transistors, or MOSFET. CMOS circuits were invented in 1963 by Frank Wanlass at Fairchild Semiconductor. The first CMOS integrated circuits were made by RCA in 1968 by a group led by Albert Medwin. Since the eighties, CMOS is pre-eminently *the* digital technology of choice. The success of CMOS in the digital semiconductor market has resulted in huge technological investments to shrink the transistors and to increase the production yield. CMOS scaling has followed *Moore's Law* for over 40 years and nowadays, several billions of transistors can be integrated on a single chip. As such, CMOS has also become the cheapest technology available today. Roughly, the cost of a SiGe technology is two to three times the cost of CMOS and GaAs is about five to ten times the cost of CMOS [Jaco].

The influence of Moore's Law on our every day life can not be overestimated. The semiconductor industry is the only industry that achieves a cost reduction every two years. In the 2005 annual report of the Semiconductor Industry Association [SIA05], it is formulated as follows: *"... in 1978, a commercial flight between New York and Paris cost 900 USD and took seven hours. If the principles of Moore's Law were applied to the airline industry, that flight would now cost about a penny and take less than one second ..."* an impressive thought.

1.2.2 RF-CMOS: Moore meets Marconi

The main trigger for the tremendous growth of the mobile phone market, was the introduction of digital coding and signal processing in wireless communications [Reyn03b]. The development and scaling of CMOS allowed the integration of enormous amounts of digital functionality on one single chip. This *digital power* enabled the use of sophisticated modulation schemes, complex demodulation algorithms, high quality error detection and correction, and allowed to obtain high data rate communications.

For a consumer, performance is only one aspect, he or she also wants a low cost mobile device with a high battery lifetime. In other words: low cost and low power consumption. The digital circuitry, typically integrated on one or two CMOS chips, already fulfills this requirement to a great extend. It is only recently that the radio frontend, i.e. the analog interface between the antenna and the digital baseband circuitry, is being integrated in CMOS [Abid04]. For this, it took the persistence of some academic institutions [Stey98] and some

pioneering firms [Silb] to prove the feasibility of CMOS design at radio frequencies[2] (RF).

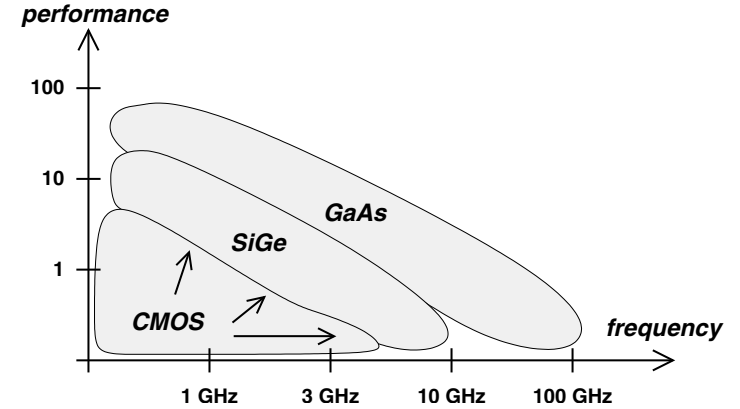
CMOS is a digital technology and was originally not developed for high-frequency or microwave design. It is thanks to the scaling, dictated by Moore's Law, that CMOS became able to operate at GHz frequencies and this triggered researchers to investigate the possibilities to do analog RF design in CMOS. Surely, better RF performance can be achieved with a dedicated RF technology like GaAs, SiGe or InP. However, the real strength of RF-CMOS is the low cost and the possibility to use digital signal processing to improve the performance of the RF frontend. Furthermore, in a highly integrated solution, the signals stay on-chip. Driving off-chip RF components requires more power and makes the system prone to noise pick-up. A highly integrated RF frontend in a CMOS technology thus results in a low power consumption, a better noise immunity and a low cost solution.

A lot of controversy still exists around RF design in CMOS. After all, CMOS can operate at high frequencies, but it is not a dedicated RF or microwave technology. Yet, the cost reduction and high integration level are the main motivations. If extreme high performance is needed, like in military or space applications, no doubt that other technologies are preferable. However, for medium performance applications and especially if low cost and high production volumes are an issue, CMOS is unbeatable. To illustrate this view, figure 1.1 shows a traditional technological view and a market driven view of CMOS, compared to SiGe and GaAs [Jaco]. The operating frequency and performance of CMOS has improved over the last decades; this is the technological view. Today, an entire mobile phone at $1.8\,\mathrm{GHz}$ can be integrated in CMOS [Silb] and research is done to integrated circuits at $24\,\mathrm{GHz}$ and even $60\,\mathrm{GHz}$ in CMOS [Komi04, Doan04]. The market-driven view on the other hand shows for which applications CMOS is of importance. If it has to be cheap and large quantities are required, CMOS is the only viable solution. But of course, the performance has to be met and no doubt that for some applications, CMOS will never be good enough.

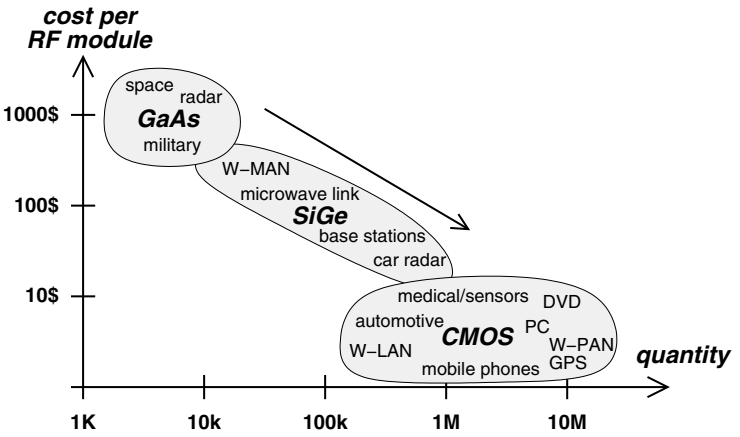
1.3 The Research Work

The RF power amplifier is a vital part of any wireless transmitter as it has to amplify the electric radio signal before it can be transmitted through the antenna. Wireless communication systems are of course very broad, submarine communications at $18\,\mathrm{kHz}$ are wireless, a microwave link at $60\,\mathrm{GHz}$ is also wireless. ... and yet they both require a power amplifier. However, this

[2] RF stands for Radio Frequencies and is rather general term since radio waves can have frequencies from a few hertz up to several hundreds of gigahertz. However, RF has become a synonym for frequencies roughly above 1 GHz.



(a) technological view



(b) market driven view

Figure 1.1. Position of RF-CMOS compared to SiGe and GaAs.

research is focussed on the design of RF power amplifiers in mobile user devices, like mobile phones.

Four keywords characterize the electrical performance of a power amplifier: *output power*, *efficiency*, *gain* and *linearity*. Output power and linearity are performance figures that are set by the requirements of the wireless system. If these specifications are not met, the power amplifier is useless. This is different for efficiency and gain. A high efficiency and a large gain results in a power amplifier that consumes little power from the battery while amplifying and transmitting radio signals. In other words, these figures are related to the battery lifetime. A higher efficiency gives a longer battery lifetime and thus a longer talk time of the mobile device. The product cost, maybe the most important figure for many customers, is related to the technology in which the

amplifier is integrated and the number of external components that are needed to have a functional power amplifier. As said before, CMOS in large volumes undoubtedly has a production cost advantage over competing technologies like SiGe and GaAs.

Several research institutes have already demonstrated the CMOS integration of RF transceivers for mobile telephony, Bluetooth and WLAN. Nowadays, these research efforts become visible in the many commercially available products [Ath, Silb, STM, Axi, Bro, RFMa] and the research focus on RF transceivers in CMOS has shifted to higher frequencies [Komi04, Doan04] and to low power consumption [Otis05].

The integration of power amplifiers in CMOS seems to follow a slightly different story. Although CMOS RF power amplifiers become commercially available [Sila, Axi], many research institutes and companies continue to investigate the possibilities to combine linearity, efficiency and output power in one single CMOS RF power amplifier. Many alternative architectures and approaches exist in literature and, in contrast to integrated receiver architectures, a clear road or solution has not arisen yet.

1.4 Outline of the Work

Combining the previous thoughts, the aim of this work is to develop design techniques for an integrated CMOS RF power amplifier. These techniques should allow the design a CMOS RF power amplifier that meets the output power and linearity requirements of a mobile communication system, that has a high efficiency and gain, that is integrated in CMOS and that requires no expensive off-chip components. The outline of this work is shown in figure 1.2. It is divided in two major parts: *theory* and *implementations*.

- Chapter 2 aims to welcome the reader in the world of power amplification. It gives a general overview of digital modulation and it will introduce some important definitions and figures that characterize a power amplifier. Next, a classification of RF power amplifiers is given together with a discussion on how to combine efficiency and linearity in one power amplifier.

- The aim to achieve both a high efficiency and a high integration level in a low cost CMOS technology is the basic idea that will lead us to chapter 3. It first presents a tool to analyze and design the Class E amplifier with the inclusion of all power losses. The influence of the parasitic capacitances on the transistor sizing is demonstrated with the design tool. Next, the impact of technology scaling, device stacking and the shift towards a higher frequency are investigated. The chapter concludes with some CMOS layout aspects.

- The next problem to tackle is the low supply voltage of current CMOS technologies, and this is covered in chapter 4. First the L-match network is discussed, followed by the lattice-type LC balun network. The latter network allows to achieve a higher output power, can easily be integrated in CMOS and can be merged with the Class E amplifier. The LC balun also allows to implement a discrete form of power control.

- Modern communication systems, like W-LAN and CDMA, allow both amplitude and phase modulation of the RF carrier to increase the datarate of a wireless link. Hence, the amplifier must have sufficient amplitude linearity. Polar modulation of the Class E amplifier allows to combine a high efficiency and a high integration level together with the required linearity and output power specifications. Chapter 5 presents a thorough discussion on polar modulation. The architectural issues and the distortion mechanisms of polar modulation are covered in this chapter and are expanded towards full digital linearization.

The theoretical aspects of this research were also put into practice and resulted in the design, fabrication and measurement of two integrated CMOS RF power amplifiers.

- Chapter 6 will discuss the design, the implementation and the measurement results of a polar modulated power amplifier for the GSM-EDGE mobile phone system. The amplifier is integrated in a 0.18 μm CMOS technology and requires no expensive RF components. First, the system level aspects of EDGE are covered, which leads to the design requirements of the integrated amplifier. The circuit implementation of the RF amplifier and the amplitude modulator are covered and followed by an extensive discussion on the measurement results.

- Chapter 7 will discuss the design, implementation and measurement results of an integrated power amplifier for Bluetooth. The differential amplifier is fully integrated in a 0.13 μm CMOS technology, has a single-ended output and is capable to efficiently control its output power. First, a brief discussion of the Bluetooth system is given and followed by the circuit level implementation issues. The measurement results clearly demonstrate the efficiency improvement of the amplifier.

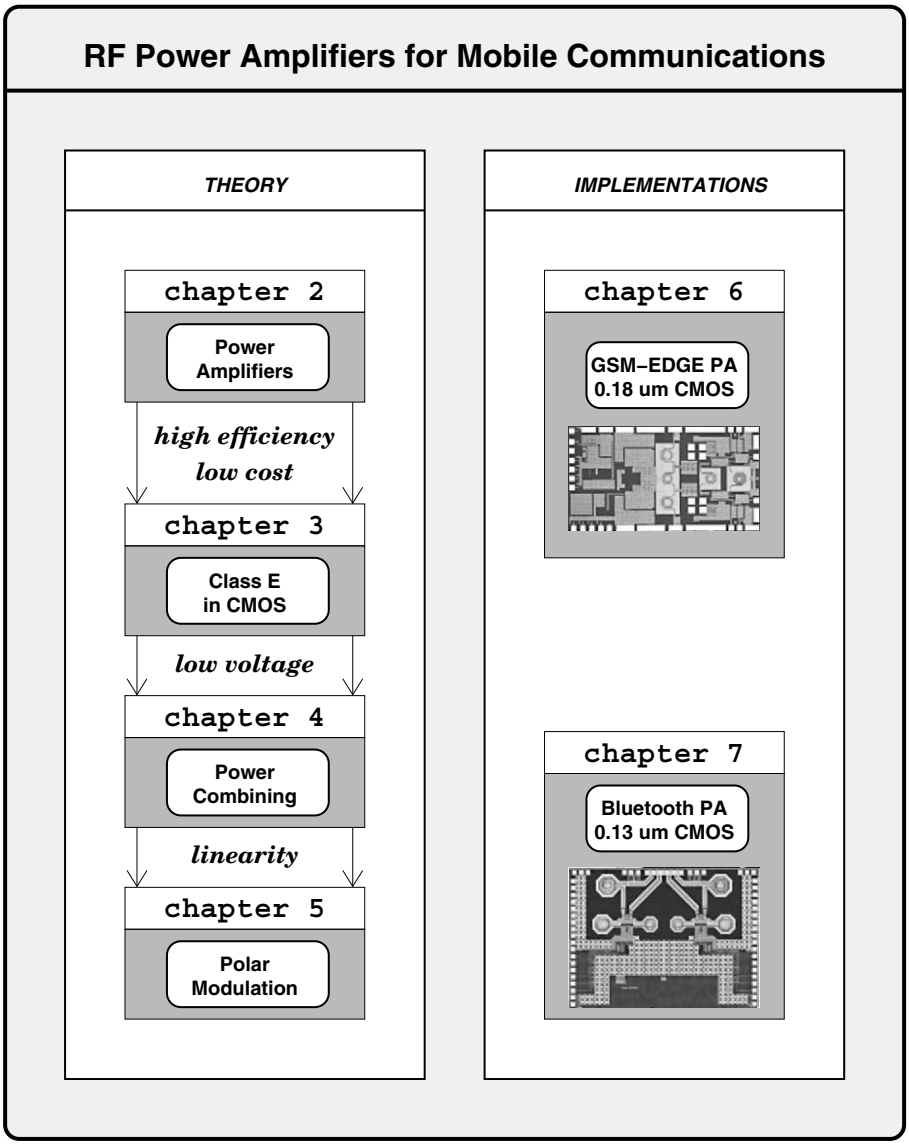


Figure 1.2. Outline of the Work.

Chapter 2

MOBILE COMMUNICATION SYSTEMS AND POWER AMPLIFICATION

2.1 Introduction

Amplifying an electrical signal is the sole purpose of a power amplifier. Though trivial at first sight, several conditions will impede the design and the implementation. For a power amplifier that is designed for a mobile or wireless communication system, *output power*, *efficiency*, *gain* and *linearity* are the most important properties, and they can easily be quantified. On the other hand *cost* and *reliability* are not as easy to quantify but their importance should not be underestimated.

To gain better insight in the different tradeoffs, this chapter will first discuss some system level aspects of mobile communication systems and the properties of the signals that need to be amplified. Next, some key parameters of an RF power amplifier are defined, as they will frequently be used throughout this text. In section 2.4, a classification of power amplifiers is given based on the classical theory of conduction angle, overdrive level and harmonic termination at the output. The classification is focussed on the difficulties regarding CMOS implementation and integration. Finally, in section 2.5 the tradeoff between efficiency and linearity is clarified and some efficiency improvement and linearization techniques are discussed.

2.2 Mobile Communication Systems

The very first step in designing a power amplifier for wireless or mobile communication is a good knowledge of the communication system itself and the signals that needs to be amplified. Therefore, this first section will review some important concepts of digital modulation and some signal properties that are important for power amplifier are defined.

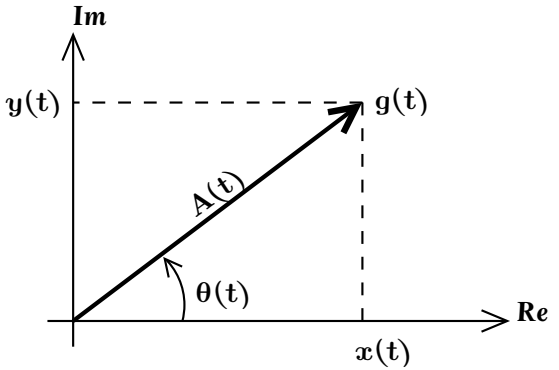


Figure 2.1. Representation of the complex envelope in the complex plane.

2.2.1 Modulated Bandpass Signals

A modulated bandpass signal can be represented as [Couc97]

$$v(t) = \mathrm{Re}\left\{g(t)e^{j\omega_c t}\right\} \tag{2.1}$$

with $f_c = \omega_c/2\pi$ the carrier frequency and $g(t)$ the *complex envelope* of $v(t)$. The complex function $g(t)$ thus modulates the phasor $e^{j\omega_c t}$. Since $g(t)$ is a complex function, its instantaneous value can be represented in the complex plane, see figure 2.1.

A single point in the complex plane can also be represented by Cartesian and polar coordinates.

$$g(t) = x(t) + jy(t) = A(t)e^{j\theta(t)} \tag{2.2}$$

Using the Cartesian and polar representation of $g(t)$, the modulated signal $v(t)$ can now be expressed as

$$v(t) = x(t)\cos(\omega_c t) - y(t)\sin(\omega_c t) \tag{2.3}$$

$$v(t) = A(t)\cos(\omega_c t + \theta(t)) \tag{2.4}$$

Looking at above equation, $A(t)$ carries the amplitude modulation and $\theta(t)$ contains the phase information. In short, $A(t)$ is called the *envelope signal* or *amplitude signal* and $\theta(t)$ is called the *phase signal*. For similar reasons, $x(t)$ is the *in-phase* or $I(t)$ signal and $y(t)$ is the *quadrature* or $Q(t)$ signal. All these signals are baseband signals with a relatively low bandwidth, at least compared to ω_c. The conversion between the two equivalent representations is as follows

$$A(t) = \sqrt{x(t)^2 + y(t)^2} \tag{2.5}$$

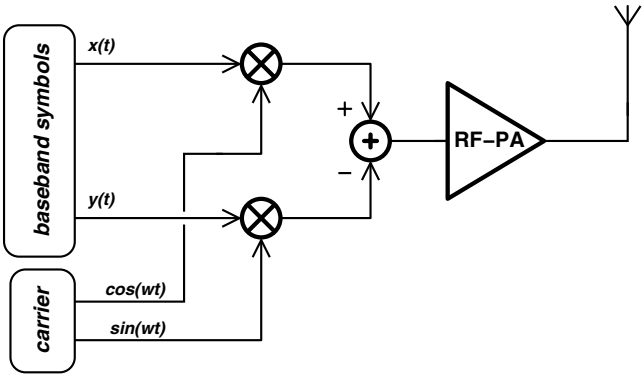


Figure 2.2. Transmitter architecture based on a Cartesian representation of a modulated signal.

$$\theta(t) = \arctan\left(\frac{y(t)}{x(t)}\right) \tag{2.6}$$
$$x(t) = A(t)\cos(\theta(t)) \tag{2.7}$$
$$y(t) = A(t)\sin(\theta(t)) \tag{2.8}$$

In most wireless transmitter architectures, the Cartesian representation of the signal is directly converted in a circuit diagram, as shown in figure 2.2.

In figure 2.3, the example of a two-tone signal is given. From this signal, the envelope and phase signal can be calculated as

$$v(t) = \sin(\omega_{LF}t)\cos(\omega_c t) \tag{2.9}$$
$$= |\sin(\omega_{LF}t)| \cdot sign\left[\sin(\omega_{LF}t)\right] \cdot \cos(\omega_c t) \tag{2.10}$$
$$= |\sin(\omega_{LF}t)| \cos\left(\omega_c t + \pi/2 - \pi/2 \cdot s(\omega_{LF}t)\right) \tag{2.11}$$
$$= A(t)\cos\left(\omega_c t + \theta(t)\right) \tag{2.12}$$

It can easily be seen that

$$A(t) = |\sin(\omega_{LF}t)| \tag{2.13}$$
$$\theta(t) = \pi/2 - \pi/2 \cdot s(\omega_{LF}t) \tag{2.14}$$
$$x(t) = \sin(\omega_{LF}t) \tag{2.15}$$
$$y(t) = 0 \tag{2.16}$$

with $s(\omega_{LF}t) = sign\left[\sin(\omega_{LF}t)\right]$ being a ± 1 switching function having the same sign as $\sin(\omega_{LF}t)$. To conclude, the envelope signal is a rectified sine wave and the phase signal is a square wave between 0 and π, as indicated by figure 2.3. Therefore, a two-tone signal exhibits both amplitude and phase modulation.

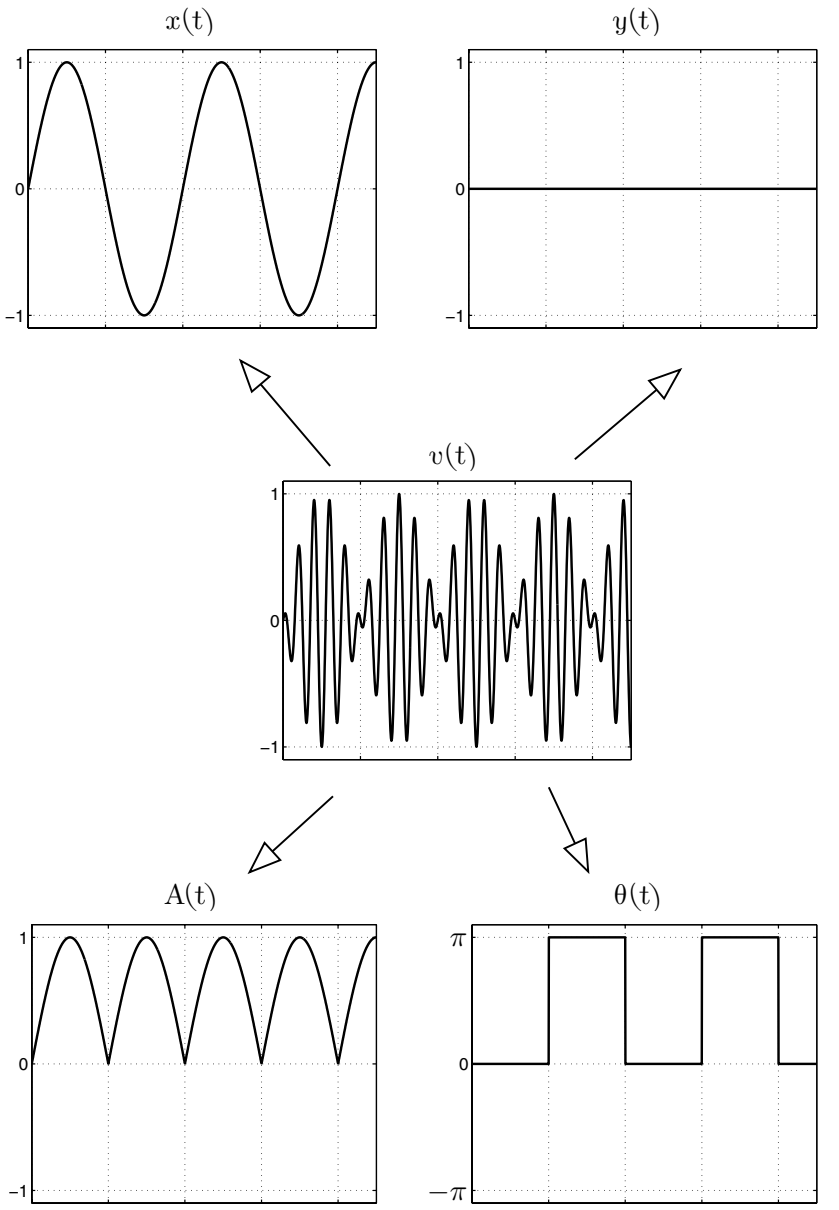


Figure 2.3. Two-tone signal and the corresponding envelope and phase signals.

To summarize, table 2.1 gives an overview of the different signals and their symbols as used throughout this text.

Table 2.1. nomenclature of modulated signals

symbol	name
$v(t)$	modulated RF signal
$g(t)$	complex envelope signal, complex baseband signal
$A(t)$	envelope signal, amplitude signal
$\theta(t), P(t)$	phase signal
$e^{j\theta(t)}, e^{jP(t)}$	complex phase signal
$v_P(t) = \cos(\omega t + P(t))$	RF phase signal
$x(t), I(t)$	in-phase signal
$y(t), Q(t)$	quadrature signal

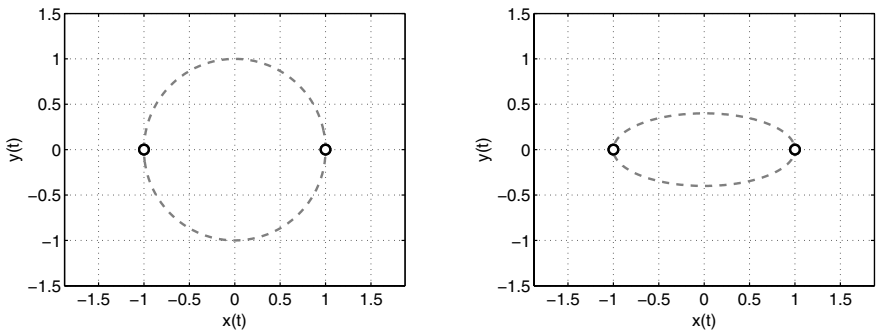
2.2.2 Digital Modulation

In digital modulation systems, a digital signal, i.e. a time discrete signal with a finite set of amplitudes, is mapped on a finite number of points in the complex plane, and these points are called *constellation points*. The corresponding graph showing all possible constellation points is named a *constellation diagram*.

One should realize that a constellation diagram is only half the story. How the complex envelope moves from one constellation point to another will determine the bandwidth of the transmitted RF signal. The complex envelope will not move instantaneously to another constellation point, as this would require an infinite bandwidth. Rather, the transition from one point to another constellation point is smoothed by applying a baseband filter on both the in-phase signal $x(t)$ and the quadrature signal $y(t)$.

The trajectory from one constellation point to another will also determine the envelope variations of the output RF signal. It will be shown later that these variations have a large impact on the design of the power amplifier. As a first example, consider the case of a Binary Phase Shift Keying (BPSK) constellation diagram, consisting of 2 constellation points. If the trajectory between the two points follows a circle, the amplitude of the complex envelope signal, i.e. the magnitude of the complex vector $g(t)$ and thus the amplitude signal $A(t)$, will not change and the RF output signal that needs to be transmitted has a *constant envelope*. In other words, $A(t)$ does not change in time. This is depicted in figure 2.4(a).

However, the path between the two constellation points of the previous example can be shortened by using an ellipse or a straight line between the two points. If a straight line is chosen, the trajectory would then go through the origin of the complex plane. The length of the complex vector will continuously change and this can be seen as a variation of the amplitude of the RF output



(a) constant envelope (b) non-constant envelope

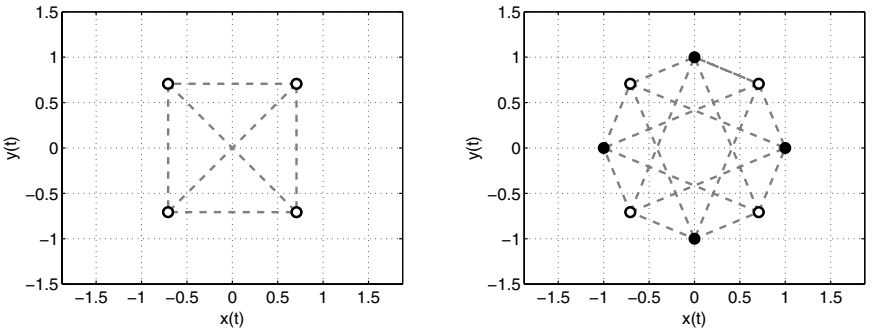
Figure 2.4. BPSK modulation.

signal. The RF signal at the output will have a changing envelope, which is denoted as a *non-constant envelope signal*. In short, $A(t)$ will change in time. This is demonstrated in figure 2.4(b).

The same holds for Quadrature Phase Shift Keying (QPSK) modulation, depicted in figure 2.5(a). A common technique to reduce the variation of the envelope signal is to use two constellation schemes that are rotated to each other. The trajectory is continuously switching between the two constellation diagrams. As an example, consider the $\pi/4$-QPSK of figure 2.5(b). The two QPSK constellation diagrams are rotated 45 degrees to each other. As such, the origin is avoided and the amplitude variations of the complex envelope signal are less severe compared to the QPSK example. Therefore, the amplitude variations of the RF signal that needs to be transmitted will also be less. In other words, the modulation depth of the amplitude modulation is reduced.

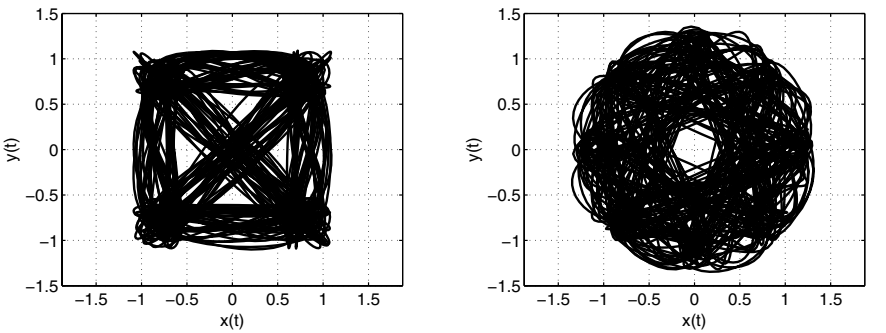
It can be concluded that envelope variations of the RF output signal are not only caused by the fact that constellation points are not lying on a circle. The transitions between the constellation points are as important to determine the envelope variations of the RF output signal. Even if the constellation points are lying on a circle, the output RF signal can still have envelope variations.

Furthermore, the transitions between the constellation points are filtered by a baseband filter, to limit the bandwidth of the transmitted signal, and this filtering operation will also effect the variation of the envelope signal. As an example, the QPSK and $\pi/4$-QPSK modulation schemes of figure 2.5 is filtered by a commonly used root-raised cosine filter with a roll-off of 0.35. Figure 2.6 shows the resulting filtered constellation diagram. Depending on the sequence of the transmitted symbols, the overshoot and undershoot of the



(a) QPSK (b) $\pi/4$-QPSK

Figure 2.5. Unfiltered complex envelope signal for (a) QPSK and (b) $\pi/4$-QPSK modulation. The trajectory avoids the origin in $\pi/4$-QPSK modulation.



(a) QPSK (b) $\pi/4$-QPSK

Figure 2.6. Filtered complex envelope signal for (a) QPSK and (b) $\pi/4$-QPSK modulation. A root-raised cosine baseband filter (r=0.35) has been applied.

baseband filter will add or subtract and thus the dynamic range of the complex signal is increased. Also, the *eye* or opening of the $\pi/4$-QPSK modulation around the origin, becomes smaller due to the baseband filtering.

2.2.3 Probability Density Function of the Envelope Signal

The location of the constellation points and the transitions or trajectories between these points, will both determine how the transmitted RF signal will look like. Nevertheless, even for the most complex (multi-carrier) modulation

scheme, one can still think of it as a carrier that is modulated both in amplitude and phase.

The design of the power amplifier will be constrained by (1) the bandwidth of the phase signal, (2) the bandwidth of the amplitude signal and (3) the variation of the envelope or amplitude signal. If no amplitude modulation is present, one denotes this as a constant envelope signal or system and amplitude linearity is of no concern. As such, the design of the power amplifier is facilitated.

A more complete description of the amplitude modulation of the carrier is given by the *probability density function of the envelope signal* (*PDF*). It gives the relative amount of time the envelope spends at a certain value. Besides the probability density function, one can also define the *cumulative density function* (*CDF*) . It describes the probability that $A(t)$ is lower than a certain value.

Signals that only use phase or frequency modulation, do not have a varying envelope signal. Therefore, as the envelope is at a constant value, the corresponding probability density function will be a Dirac impulse. It should be stressed that, for RF communication, it is more common to look at the envelope *PDF* . In baseband amplifiers and line drivers, it is more common to look at the *PDF* of the actual output signal.

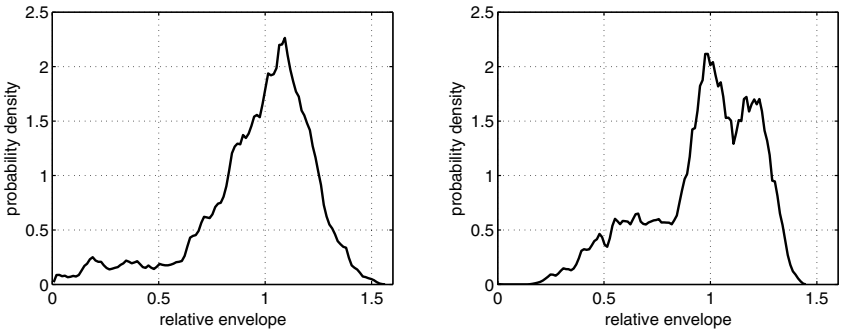
The envelope *PDF* is important for the optimization of the amplifier. After all, the *PDF* tells the designer what signal the amplifier will have to transmit, most of the time. As such, one could optimize the power amplifier in the region where the *PDF* is high.

As an example, the QPSK and $\pi/4$-QPSK modulation schemes can be used again. The envelope waveform can be obtained from figure 2.6 and the corresponding envelope probability density function is shown in figure 2.7. For QPSK, the envelope has a peak value of 1.58, an average value of 0.96 and an rms value of 1. The crest factor is $3.97\,\mathrm{dB}$. Notice that the envelope for which the probability density is maximum, does not correspond to either the average or rms value of the envelope. For $\pi/4$-QPSK, the envelope has a peak value of 1.5, an average value of 0.97 and an rms value of 1. The crest factor is $3.28\,\mathrm{dB}$. As said before, the $\pi/4$-QPSK modulation scheme avoids the origin, and this can clearly be seen in the corresponding PDF.

2.3 Some Aspects of Power Amplification

2.3.1 Output Power

Consider the basic circuit of figure 2.8, which shows a power amplifier connected to an antenna. The output power is defined as the active power, delivered by the power amplifier and flowing into the antenna. Inside the antenna, the power is *dissipated* under the form of a radiated electromagnetic wave. In most cases, the antenna impedance Z_{ant} is designed to be purely resistive at



(a) QPSK (b) $\pi/4$-QPSK

Figure 2.7. Probability density function of the envelope signal after applying a root-raised cosine baseband filter (r=0.35).

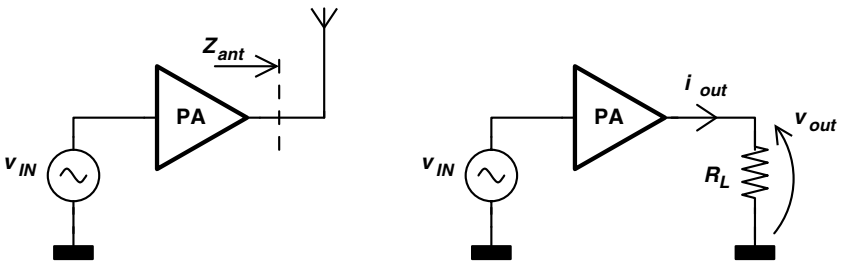


Figure 2.8. Definition of output power.

the frequencies of interest. Therefore, at these frequencies, the antenna can be represented by a single load resistor R_L. The power dissipated in R_L (under the form of heat) is, by definition, equal to the power in the electromagnetic wave, transmitted by the antenna.

In RF and Microwave, it is common to *design towards* $50\ \Omega$. Antennas, antenna filters and other microwave components typically have single-ended input and output impedances of $50\,\Omega$. However, it will be shown in chapter 4 that it is possible to convert the $50\,\Omega$ antenna impedance to a lower or higher value. Therefore, it is assumed for now that R_L can virtually have any desired value.

The *instantaneous output power* is defined as

$$p_o(t) = v_{out}(t) \cdot i_{out}(t) \tag{2.17}$$

The *total* or *average* output power $P_{o,tot}$ is defined as

$$P_{o,tot} = \langle p_o(t) \rangle = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} p_o(t) dt \tag{2.18}$$

with $\langle \cdot \rangle$ the time average operator. If the output voltage is a sine wave with frequency f_c en period T_c, previous equation simplifies to

$$P_{o,tot} = \langle p_o(t) \rangle = \frac{1}{T_c} \int_{-T_c/2}^{Tc/2} p_o(t) dt \tag{2.19}$$

Under the assumption of a resistive load,

$$P_{o,tot} = \langle v_{out}(t) \cdot i_{out}(t) \rangle = \frac{\langle v_{out}^2(t) \rangle}{R_L} \tag{2.20}$$

$$= \frac{V_{o,rms}^2}{R_L} \tag{2.21}$$

with

$$V_{o,rms} = \sqrt{\langle v_{out}^2(t) \rangle} \tag{2.22}$$

the well known RMS value of the output voltage.

Although trivial at first sight, these basic textbook definitions are not always as useful for a power amplifier. The amplifier will not only generate power at the frequency of interest, but also at integer multiples of the fundamental frequency f_c. In most cases, only the power at the fundamental frequency is wanted and the harmonic power has to be filtered or suppressed at the output. Therefore, it is more useful to define a *fundamental average output power*, P_{o,f_c}, equal to the output power at the fundamental frequency only

$$P_{o,f_c} = \frac{V_o^2}{2R_L} \tag{2.23}$$

with V_o the amplitude or peak value of the sinusoidal output voltage at frequency f_c. This value can be obtained from a Fourier Series expansion of $v_{out}(t)$.

Throughout this text, the general term *output power*, denoted as P_o, will be used to indicate the average output power at the fundamental frequency. Thus,

$$P_o = P_{o,f_c} \tag{2.24}$$

2.3.2 Peak Output Power and Crest Factor

For modulated signals, the profile of the envelope or amplitude signal $A(t)$ becomes important. One can proof that the average output power of a bandpass

modulated signal is given by [Couc97]

$$P_o = \frac{\left\langle |g(t)|^2 \right\rangle}{2R_L} = \frac{\langle A^2(t) \rangle}{2R_L} = \frac{(A_{rms})^2}{2R_L} \tag{2.25}$$

If the output RF signal is only phase modulated, the output power is the same as if the signal was not modulated at all. Hence, for constant envelope systems, the presence of the modulation will not change the output power.

Clearly, the average output power will be dependent on the behavior of the envelope signal. One could define a *fixed envelope output power* $P_o(A)$ as the average output power, dissipated in the load, if the envelope or amplitude signal is kept constant at a specific value A

$$P_o(A) = \frac{A^2}{2R_L} \tag{2.26}$$

which is basically the same as definition 2.23.

One can also define the *peak envelope output power*, *PEP* . It is the power, dissipated in the load, if the envelope or amplitude signal is kept at its maximum value [Couc97].

$$PEP = \max\{P_o(A(t))\} = \frac{(\max\{A(t)\})^2}{2R_L} = \frac{A_{max}^2}{2R_L} \tag{2.27}$$

The peak output power value is of great importance in PA design since the amplifier must be designed to transmit these high peaks. However, the average output power can be much lower than the peak output power.

In this regard, an important parameter is the *Peak to Average Power Ratio*, *PAPR* , defined as

$$PAPR = \frac{PEP}{P_o} \tag{2.28}$$

The peak to average ratio can also be calculated based on voltages. This is expressed by the *Crest Factor* (CF), defined as the ratio of the peak value to the rms value of the RF output voltage.

$$CF = \frac{V_{o,max}}{V_{o,rms}} = \frac{A_{max}}{A_{rms}} \tag{2.29}$$

Note that

$$PAPR = CF^2 \tag{2.30}$$

A signal with a high *PAPR* , will require a lot of headroom for the power amplifier. Most of the time, the amplifier will operate at a relatively low output power level, but from time to time, the power amplifier has to be able to transmit relatively large power peaks. This time-domain behavior of the power

amplifier is determined by the characteristics of the RF signal that needs to be transmitted. One such characteristic is the probability density function of the envelope signal, as already discussed in section 2.2.3. For systems with only phase modulation, the peak output power will be the same as the average output power, since the envelope signal $A(t)$ is a constant and equal to $V_{o,max}$. Therefore, the *PAPR* and *CF* will both be equal to 1 and thus $0\,\mathrm{dB}$.

In the general case where multiple tones or carriers are present, the average output power of the combined signal will be [Krau80]

$$\begin{aligned} P_o &= P_{o,1} + P_{o,2} + \ldots + P_{o,k} && (2.31) \\ &= \frac{1}{2R_L}\left(A_1^2 + A_2^2 + \ldots + A_k^2\right) && (2.32) \end{aligned}$$

with $P_{o,i}$ the average output power of the *i-th* carrier and A_i the amplitude of the *i-th* carrier. If it is assumed that all carriers are independent from each other, they will at some point in time simultaneously reach their positive maximum values, resulting in a peak envelope output power of

$$PEP = \frac{1}{2R_L}\left(A_1 + A_2 + \ldots + A_k\right)^2 \tag{2.33}$$

Furthermore, if all carriers have the same amplitude A

$$\begin{aligned} P_o &= \frac{kA^2}{2R_L} && (2.34) \\ PEP &= \frac{(kA)^2}{2R_L} && (2.35) \end{aligned}$$

and thus

$$PAPR = \frac{PEP}{P_o} = k \tag{2.36}$$

2.3.3 Input Power and Power Gain

In order to drive the power amplifier, a certain amount of RF input power is required. An oscillator can be regarded as an exception to this, since it only requires DC input power.

The power gain, usually expressed in dB, is defined as

$$G_{P,dB} = 10\log_{10}\left(\frac{P_o}{P_{in}}\right) \tag{2.37}$$

2.3.4 Efficiency

An *efficient* power amplifier aims to deliver a certain amount of power to the load, without consuming too much power itself. The DC power consumption,

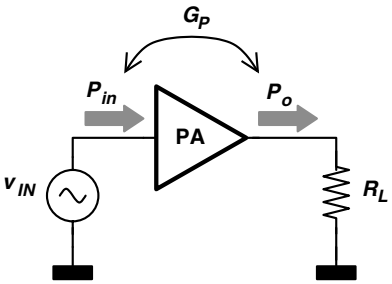


Figure 2.9. Definition of input power and power gain.

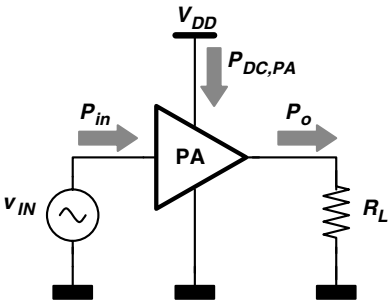


Figure 2.10. Definition of DC power consumption.

$P_{DC,PA}$ in figure 2.10, will always be larger than P_o. The *drain* or *collector efficiency* η_d is defined as

$$\eta_d = \frac{P_o}{P_{DC,PA}} \tag{2.38}$$

Equation 2.38 can also be regarded as a *fundamental* efficiency, since it only takes into account the output power at the fundamental frequency.

One can also define a *conversion* efficiency η_{conv} which is the ratio of the total RF output power to the DC power consumption of the amplifier.

$$\eta_{conv} = \frac{P_{o,tot}}{P_{DC,PA}} \tag{2.39}$$

$P_{o,tot}$ includes the fundamental output power P_o as well as the (unwanted) output power at the higher harmonics. Therefore, η_{conv} can be regarded as an indication how well DC power is converted to RF power, hence the name conversion efficiency.

In almost any transmitter chain, several driver stages are required between the signal source or upconversion mixer and the last amplifier stage. These

driver stages will also consume DC power, but it is not as easy to define input and output power, since the impedance levels at the input and output of each stage will be different and will normally be composed of both a real and complex part. Taking the DC power consumption of the driver stages into account, the *overall efficiency* of the power amplifier can be defined as [Raab02]

$$\eta_{oa} = \frac{P_o}{P_{DC,PA} + \sum_{i=0}^{n} P_{DC,DRV,i}} \tag{2.40}$$

The more driver stages are added, the higher the power gain of the entire amplifier, but the lower the overall efficiency η_{oa}.

The Power Added Efficiency (*PAE*) takes into account the RF input power P_{in} and it is defined as

$$PAE = \frac{P_o - P_{in}}{P_{DC,PA} + \sum_{i=0}^{n} P_{DC,DRV,i}} \tag{2.41}$$

By using equation 2.37, the PAE can be rewritten as

$$PAE = \eta_{oa}\left(1 - \frac{P_{in}}{P_o}\right) = \eta_{oa}\left(1 - \frac{1}{G_P}\right) \tag{2.42}$$

However, some authors define the *PAE* as

$$PAE = \frac{P_o - P_{in}}{P_{DC,PA}} = \eta_d\left(1 - \frac{1}{G_P}\right) \tag{2.43}$$

which clearly results in a better performance. Anyhow, if the gain is sufficiently high, the *PAE* defined in equation 2.41 becomes equal to the overall efficiency.

The drain efficiency of an amplifier can, at least in theory, reach $100\,\%$. However, the overall efficiency will always be smaller than $100\,\%$, even in the ideal case. The reason is clear; the power that is consumed by the driver stages will not flow to the output but is dissipated at the input of the next stage in line. As such, even if the power amplifier and the driver stages operate at a drain efficiency of $100\,\%$, the overall efficiency will not be $100\,\%$, simply because of the definition of overall efficiency.

The question remains which definition of efficiency to use. From a circuit level point of view, drain efficiency and *PAE* seems the best, especially if for a stand-alone power amplifier with an input matched to $50\,\Omega$. From a system point of view and for a system designer, the PA is *everything after the upconverter*. Therefore, the overall efficiency η_{oa} is a better indication of how

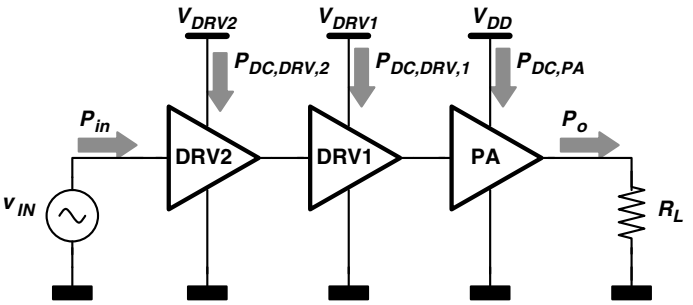


Figure 2.11. Definition of DC power consumption, including driver stages.

much power is needed to amplify the signal, relatively to the output power itself. Of course, one could argue if the power consumption of the DSP and other digital and analog block shouldn't be included as well. After all, these blocks, just like the driver stages, are also required to modulate and transmit a signal. This, however, would lead us to definitions that have little to do with *conversion* of DC power to an electromagnetic wave, but rather with *consumption* of energy to transmit information.

2.3.5 Efficiency and Modulated Signals

Regardless of what type of modulation is used, the definition of drain efficiency is still equal to the average output power divided by the average DC power consumption.

However, in almost any power amplifier, drain efficiency and output power are closely related to each other. In section 2.3.2, the fixed envelope output power was defined as the average output power, if the envelope signal is kept constant at a specific value A.

$$P_o(A) = \frac{A^2}{2R_L} \tag{2.44}$$

With this envelope output power, one can define a drain efficiency for a specific constant envelope signal. In other words, the drain efficiency is made dependent on the envelope signal and one can define a *fixed envelope drain efficiency* as.

$$\eta_d(A) = \frac{P_o(A)}{P_{DC,PA}(A)} \tag{2.45}$$

If the envelope signal changes in time, the fixed envelope efficiency will also change. The average efficiency can now be interpreted as the time average of

the fixed envelope efficiency.

$$\eta_d = \langle \eta_d(A(t)) \rangle = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} \eta_d(A(t)) dt \tag{2.46}$$

The calculation of the average efficiency can also be done, based on the envelope probability density function. This does not require knowledge of the the time behavior of the envelope signal. Only the probability density function is required.

$$\eta_d = \int_0^{A_{max}} \eta_d(A) \cdot p(A) \cdot dA \tag{2.47}$$

The product $\eta_d(A) \cdot p(A)$ can be seen as the *probability density function of the drain efficiency*.

For the Class A power amplifier (see section 2.4.1), the relationship between efficiency and envelope signal is given by

$$\eta_d = \frac{P_o}{P_{DC,PA}} = \frac{1}{2} \cdot \left(\frac{A}{V_{DD}} \right)^2 \qquad A = [0 \ldots V_{DD}] \tag{2.48}$$

If the supply voltage V_{DD} is normalized ($V_{DD} = 1\,\mathrm{V}$), the relationship becomes

$$\eta_d = \frac{A^2}{2} \qquad A = [0 \ldots 1] \tag{2.49}$$

which reveals a quadratic dependency of η_d on the amplitude of the output signal A. Figure 2.12 shows the calculation of the efficiency, based on equation 2.47. The efficiency versus envelope curve is first multiplied with the envelope probability density function. The resulting graph is then integrated to obtain the average efficiency of the amplifier.

2.3.6 Power Control

Most wireless systems utilize some form of power control to either avoid interference with other users or to conserve battery power. After all, it is of no use to transmit more power than actually needed to preserve a communication link. Normally the highest output power levels are only needed in a worst-case scenario. In the case of wireless telephony, this may occur when the mobile station is distant from the nearest base station. Other factors are: attenuation by buildings, multi path fading and orientation of the mobile antenna. In modern systems, the average output power is constantly adopted to the lowest possible level in order to increase the battery lifetime. This variation of the average output power can also be interpreted as an amplitude modulation and thus as a variation of the envelope signal $A(t)$.

It should be stressed that power control is also strongly related to the actual system. In CDMA systems [Sahu04] , all users transmit at the same time, at the same frequency. This may cause a so called near-far problem if no

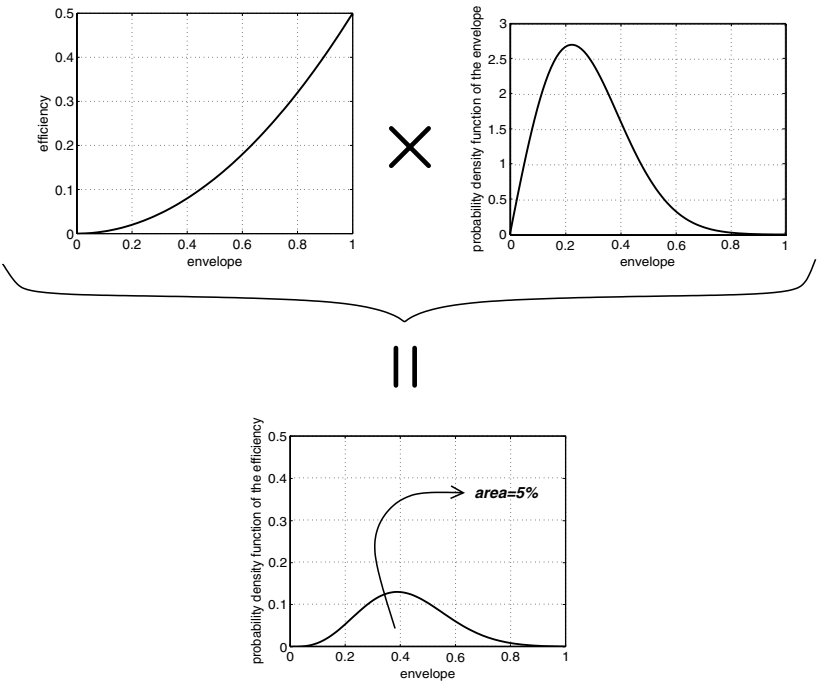


Figure 2.12. Calculation of the efficiency, based on the probability density profile of the envelope and the efficiency versus envelope curve of the amplifier.

precautions are taken. At the base station, the signal of a distant user will arrive at a much lower power level compared to nearby users. As such, nearby users may *overpower* the weak signal of a distant user, since both signals occur at the same time and at the same frequency. For a maximum capacity in a CDMA system, power control is crucial and the system will try to make the received level of each user equal at the base station. Consequently, the power control in each mobile terminal needs to be accurate and adjusted frequently. In systems that employ TDMA or FDMA , each user either transmits in another time slot or on another frequency. For these systems, power control is not crucial for maximum capacity and it is mainly used to increase the battery lifetime of the mobile terminal.

Based on actual measurements, a probability density function of the average output power can be obtained. Such a function will heavily depend on both the environment and the system in use. For a TDMA and FDMA system in a rural environment, the average output power will be fairly constant and only dependent on the distance to the base station. Base station density is rather low in rural areas, and therefore, the mobile terminal will operate at higher output

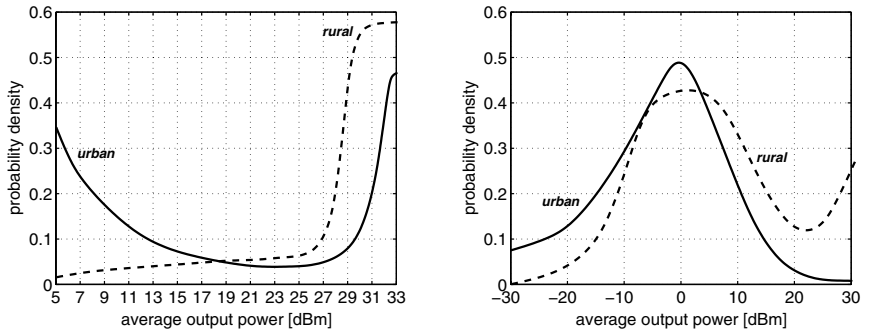


Figure 2.13. Examples of a probability density function of the average output power for (a) TDMA/FDMA and (b) CDMA.

power levels. In an urban environment, the attenuation can raise to very high peaks, as buildings may create a strong shading of the antenna. Such strong attenuation peaks will raise the output power towards its peak levels. On the other hand, the base station density is much higher in urban environment than in a rural environment. This will enable the mobile terminal to operate at lower power levels from time to time. For CDMA, similar observations can be made. From these qualitative descriptions, the probability density function of the average output power could have a profile as illustrated in figures 2.13(a) and 2.13(b) [Hani99], [Raab02] and [Sahu04].

The time-variant behavior of the envelope signal will thus have two causes, (1) amplitude modulation of the carrier and (2) the variation of the output power due to power control. From this point of view, a *long-term average efficiency* could be defined that takes into account the variation of $A(t)$ due to power control. However, calculating such a long-term average efficiency requires accurate knowledge of the environment and the behavior of the user, which will be difficult to predict.

2.3.7 Linearity

Linearity of a power amplifier can be defined in two ways: phase linearity and amplitude linearity. Phase linearity is easy to achieve as long as the bandwidth of the modulated signal is small compared to the carrier frequency. Phase nonlinearity or phase distortion is denoted as PM-PM distortion. Amplitude linearity is harder to achieve, and in general results in a lower efficiency as will become clear in section 2.4. Amplitude nonlinearity or amplitude distortion is denoted as AM-AM distortion. In this regard, constant envelope systems, like GSM and Bluetooth, have an efficiency advantage. The amplifier only

needs to have phase linearity and amplitude linearity is of no concern. Recent years have shown that most new communication systems (EDGE, W-CDMA, W-LAN) have varying envelope signals. To amplify these signals correctly, the power amplifier must thus have sufficient amplitude linearity.

The varying envelope of the RF signal may also induce phase errors, indicated as AM-PM. Likewise, PM-AM are amplitude or envelope errors caused by the phase or frequency modulation of the carrier. AM-AM, AM-PM, . . . are in fact a DC approximation of the power amplifier nonlinearity. In reality, the distortion curves are not fixed but can change slowly due to temperature changes or other environmental effects. This is denoted as memory-effects.

Different means exist to quantify the linearity of a power amplifier. The third order intercept point (IP3) is a well known characteristic of low noise amplifiers and mixers [Crip99]. It is based on a two-tone signal test and gives the (virtual) output power for which the third order intermodulation term becomes as large as the actual output power. Modern communication systems create highly complex signals and the relative simple IP3 measurement is often not suited as a good indicator for the required linearity. Furthermore, communication standards do not specify the required IP3 of the transmitter but rather, terms like *spectral mask*, *error vector* and *adjacent channel power* are used.

The distortion of RF power amplifiers and the linearity requirements will be further discussed in chapters 5, 6 and 7.

2.3.8 Inductors, Capacitors and Quality Factor

As inductors and capacitors play a crucial role in RF amplification, some basic metrics are briefly reviewed in this section.

2.3.8.1 Un-tuned Network

In reality, inductors and capacitors are never ideal. They always exhibit some loss, which can be modeled by adding a resistance either in series or parallel.

Figure 2.14(a) shows the very basic circuit of an inductor with a series resistance. One can define a quality factor of this little circuit as being the ratio of the average reactive power to the average power dissipated in the resistor. For the circuit of figure 2.14(a), this becomes

$$Q = \frac{X_S}{R_S} = \frac{\omega L_S}{R_S} \tag{2.50}$$

For a capacitor with a series resistance, the formula becomes

$$Q = \frac{X_S}{R_S} = \frac{1}{\omega C_S R_S} \tag{2.51}$$

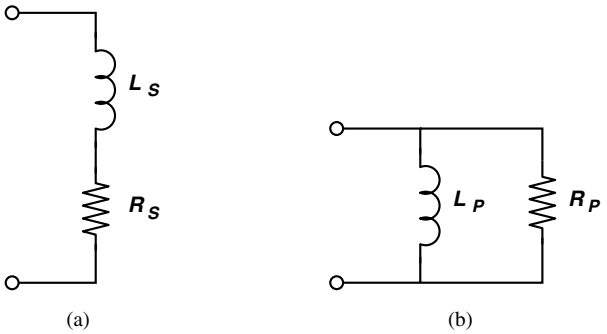


Figure 2.14. Inductor with loss. (a) Series model and (b) parallel model.

This little circuit can be converted in a parallel network, as given by figure 2.14(b). The quality factor of the parallel network can be calculated as

$$Q = \frac{R_P}{X_P} = \frac{R_P}{\omega L_P} \tag{2.52}$$

For a capacitor with a parallel resistance, the formula becomes

$$Q = \frac{R_P}{X_P} = R_P\,\omega C_P \tag{2.53}$$

The series and parallel representation are equivalent to each other. A relationship between L_S, L_P, R_S and R_P can easily be obtained by calculating the impedance of both networks and equalizing the real and imaginary part. This results in following equivalences:

$$L_P = L_S\left(1 + \frac{1}{Q^2}\right) \approx L_S \tag{2.54}$$

$$R_P = R_S\left(1 + Q^2\right) \tag{2.55}$$

Equivalent formulas can be obtained for a capacitance

$$C_P = C_S\left(\frac{1}{1 + 1/Q^2}\right) \approx C_S \tag{2.56}$$

$$R_P = R_S\left(1 + Q^2\right) \tag{2.57}$$

2.3.8.2 Tuned Network

Figure 2.15 depicts a tuned RLC network with losses. In most cases, the quality factor of a capacitor will be much higher than the quality factor of an

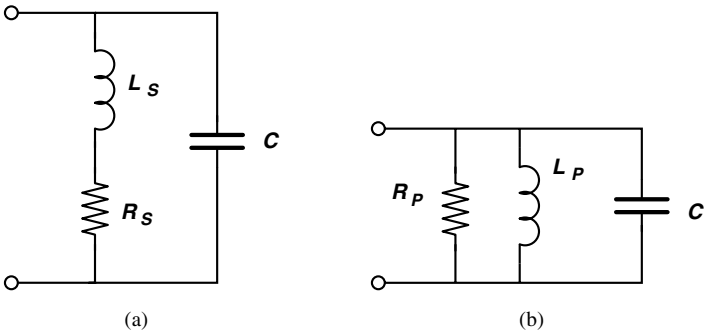


Figure 2.15. RLC network with inductor loss.

inductor. Therefore the quality factor of the entire network will mainly be determined by the quality factor of the inductor alone.

For figure 2.15(b), the resonance frequency will be equal to

$$\omega_0 = \frac{1}{\sqrt{L_P C}} \tag{2.58}$$

and the quality factor of that circuit becomes

$$Q(\omega_0) = R_P \sqrt{\frac{C}{L_P}} \tag{2.59}$$

The circuit of figure 2.15(a) has the same resonance frequency since both circuits are electrically equivalent. After some calculations, one can find that

$$\omega_0 = \frac{1}{\sqrt{L_P C}} = \sqrt{\frac{1}{L_S C} - \frac{R_S^2}{L_S^2}} \approx \frac{1}{\sqrt{L_S C}} \tag{2.60}$$

and therefore

$$Q(\omega_0) \approx \frac{1}{R_S} \sqrt{\frac{L_S}{C}} \tag{2.61}$$

If, in a first order approximation and at a fixed frequency, the series resistance is assumed to be proportional to the value of the series inductance, i.e.

$$R_S = \alpha L_S \tag{2.62}$$

then, the impedance of the tuned network at the resonance frequency ω_0 will become purely resistive and equal to R_P

$$R_P = R_S(1 + Q^2) \approx R_S Q^2 \approx \frac{1}{\alpha C} \tag{2.63}$$

From previous equation it can be seen that, although the capacitor is *tuned* by the inductor, the higher the capacitance, the lower the equivalent resistance seen at the (same) resonant frequency.

2.4 Power Amplifier Classification

Power Amplifiers are traditionally divided in several classes, depending on how the transistor is driven and the harmonic content or time behavior of the drain voltage [Crip99]. This classification dates back from the early days of electronics, during the glory days of triode and pentode tubes. Although most real-world solid-state amplifiers are situated in between the different classes, a strict classification allows to obtain a better view on the tradeoff between output power, gain, efficiency and linearity.

For wireless communication, the most important distinction is that between linear and nonlinear amplifiers. The latter refers to a power amplifier that only has phase linearity but no amplitude linearity. The main motivation for using a nonlinear amplifier, is the higher efficiency that can be achieved. The nonlinear behavior doesn't necessarily has to be a drawback. A lot of wireless systems and standards use only phase modulation and the corresponding RF waveforms do not have amplitude variations. As a consequence, the power amplifier only needs to have phase linearity and the amplitude linearity is of no concern.

As this work is focused on CMOS, MOS transistors and their formulas will be used. But of course, the classification is independent of technology.

2.4.1 Class A

Figure 2.16 show the simplest circuit that can operate as a Class A amplifier, using an nMOS transistor. At the gate, a DC bias voltage ensures that a quiescent current I_Q is flowing through the transistor and inductor L_{DC}. A DC-blocking capacitor C_{BL} ensures that no DC current is flowing through the load resistor R_L.

The main requirement for Class A operation, is that the transistor conducts current all the time. The maximal output voltage, equal to V_{DD}, occurs when the transistor drain current almost goes to zero. The minimal output voltage is equal to $V_{DD} - I_Q R_L$. The output swing is thus maximized if the quiescent current through the transistor is chosen equal to V_{DD}/R_L.

If the AC input signal is a sine wave, the output voltage will also be a sine wave. and the output power will be equal to

$$P_o = \frac{V_o^2}{2R_L} \tag{2.64}$$

with V_o the amplitude or peak value of the sinusoidal output voltage $v_o(t)$. The maximum amplitude of the output voltage is equal to V_{DD}. The power

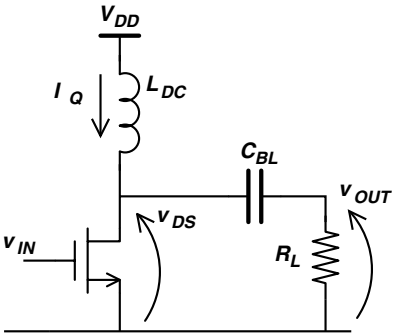


Figure 2.16. Simplified circuit of a Class A power amplifier.

delivered by the DC voltage source is equal to

$$P_{DC} = I_Q \cdot V_{DD} = \frac{V_{DD}^2}{R_L} \tag{2.65}$$

where it is assumed that the quiescent current I_Q is chosen equal to V_{DD}/R_L From these equations, the well known formula for the efficiency of the Class A power amplifier can be obtained.

$$\eta = \frac{P_o}{P_{DC}} = \frac{1}{2} \cdot \left(\frac{V_o}{V_{DD}}\right)^2 \tag{2.66}$$

Above equation reveals the quadratic dependency of the efficiency on the output voltage amplitude V_o and the maximum sinusoidal efficiency of $50\,\%$.

If the Class A power amplifier is *normalized*, i.e. $V_{DD} = 1\,\mathrm{V}$ and $R_L = 1\,\Omega$, the maximum normalized output power is equal to $0.5\,\mathrm{W}$, the maximum drain-source voltage becomes $2\,\mathrm{V}$ and the maximum drain-source current is $2\,\mathrm{A}$. Although trivial, the normalized voltage and current waveforms and the normalized power dissipation in the transistor are shown in figure 2.17 for completeness.

Some authors also define an *output power capability* as

$$C_P = \frac{P_o}{v_{DS,max} \cdot i_{DS,max}} \tag{2.67}$$

It originates from the fact that discrete power transistors are one of the most expensive components of a transmitter. This means that the devices have to be used as close as possible to their maximum voltage and current ratings. For CMOS, the main restriction is the breakdown voltage of the transistor, rather than the current limits of the technology. In this regard, one can define a

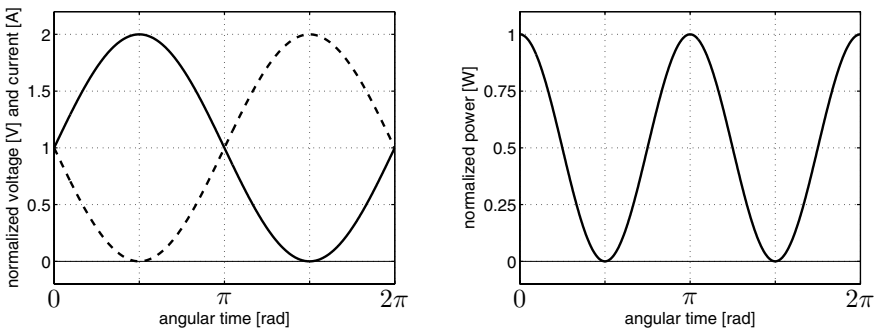


Figure 2.17. (a) Normalized drain voltage and drain current (dashed line) and (b) normalized power dissipation for a Class A power amplifier.

maximum output power, related to the maximum drain-source voltage as

$$P_{o,max} = \frac{1}{8} \cdot \frac{v_{DS,max}^2}{R_L} = 0.125 \cdot \frac{v_{DS,max}^2}{R_L} \tag{2.68}$$

The fact that the nMOS transistor of a Class A amplifier will continuously conduct current, has a large influence on the reliability and this will be discussed in section 2.4.8.

It should be stressed that the $50\,\%$ efficiency of the Class A amplifier only occurs at the maximum (idealized) output swing. If an amplitude modulated signal is applied, the amplitude of the RF sinewave at the output V_o will change according to the envelope signal $A(t)$. Depending on the probability density function of $A(t)$, the average efficiency will be much lower (see section 2.3.5).

Furthermore, the assumption of a maximal amplitude of V_{DD} will no longer hold if the *knee*-region[1] of the transistor is taken into account. For an nMOST, the drain-source voltage should remain higher than $V_{GS} - V_{th}$ as the transistor would otherwise be over-driven and no longer behave as a current source[2]. As an example, consider a $0.18\,\mu\mathrm{m}$ CMOS technology. The supply voltage is chosen to be $0.9\,\mathrm{V}$, which makes the maximal drain voltage equal to $1.8\,\mathrm{V}$. If a $V_{GS} - V_{th}$ of $200\,\mathrm{mV}$ is chosen, the output voltage will have a maximum amplitude of $700\,\mathrm{mV}$ and the maximum efficiency will only be $30\,\%$, not taking into account the power loss of the inductor. Thus, it is necessary to reduce

[1] which is the linear region for a MOST and the saturated region for a BJT.
[2] in section 2.4.3, the Class A amplifier is revisited under overdrive conditions.

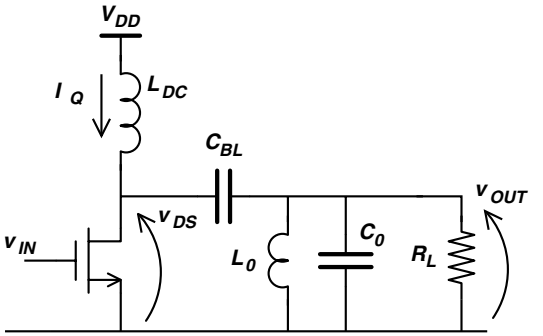


Figure 2.18. Simplified circuit for reduced conduction angle operation.

the value of $V_{GS} - V_{th}$. In a (*very*) first order approximation,

$$I_{DS} = K\frac{W_g}{L_g}\left(V_{GS} - V_{th}\right)^2 \tag{2.69}$$

and thus for a given current,

$$\left(V_{GS} - V_{th}\right) \sim \frac{1}{\sqrt{W_g}} \tag{2.70}$$

$$g_m \sim \sqrt{W_g} \tag{2.71}$$

$$C_{gs} \sim W_g \tag{2.72}$$

To increase the output voltage swing, it is beneficial to choose a large transistor width. In order to halve $V_{GS} - V_{th}$, the transistor width W_g has to be multiplied by four. which will also double the gain, equal to $g_m R_L$. Multiplying the transistor width by four will, again in a first order approximation, also increase the input capacitance by four, and thus the input impedance is divided by four, even if this capacitance is tuned by an inductor (see section 2.3.8.2). For the same gain in the driver stage, the DC current of the driver therefore has to be increased by four. On the other hand, the gain of the power stage itself is now two times as large. Thus, the gain of the driver can be halved, but even then, the DC current of the driver stage has to be increased by two. In any case, a larger transistor will always increase the power consumption of the driver stage. This tradeoff is fundamental and present in any power amplifier.

2.4.2 Reduced Conduction Angle: Class AB, B and C

In a Class A amplifier, the transistor will always conduct the quiescent current. This leads to the low efficiency of the Class A amplifier, especially if the effect of the knee region is included. A higher efficiency can be obtained if the

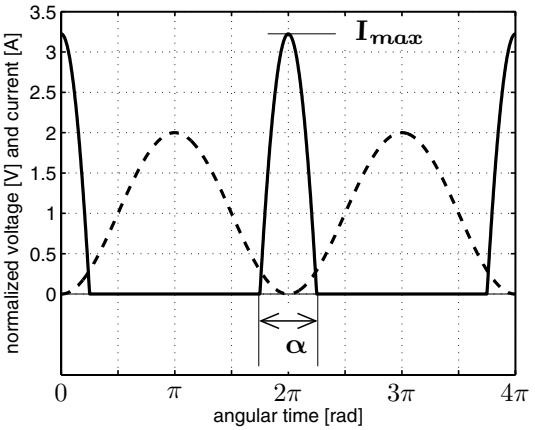


Figure 2.19. Reduced conduction angle waveforms: drain current (solid line) and drain-source voltage (dashed line). Normalization: $V_{DD} = 1\,\mathrm{V}$ and $R_L = 1\,\Omega$.

transistor does not have to conduct current all the time. This has lead to the development of more efficient, but less linear power amplifiers. Turning of the transistor can easily be obtained by reducing the DC bias at the gate.

If the transistor is turned off, the shape of the voltage and current waveforms can no longer be determined by the transistor itself. It is the external passive circuitry that now will determine the shape of current and voltage. Consequently, it will also be the external passive circuitry that will have a great impact on the efficiency. After all, *shaping* the voltage and current waveforms also means *shaping* the voltage-current overlap, which directly affects the efficiency of the PA.

Turning off the transistor will also generate higher harmonics. How these harmonics are being terminated will be crucial for the shaping of the voltage and current waveforms, again having a large impact on the efficiency. In this section, it is assumed that all higher harmonics are shorted to ground by a high-quality LC-tank, as indicated in figure 2.18. Because of the high quality factor, the output voltage must be a sine wave, and so will be the drain-source voltage. The maximum amplitude of the output sine wave is again equal to V_{DD}.

The drain voltage and current are depicted in figure 2.19. From this figure, it is clear that the overlap between the current through and the voltage across the transistor is reduced. Basically, this is the reason for the higher efficiency that can be achieved as the conduction angle is reduced. The current through the transistor can be written as [Crip99]

$$i_{DS}(\theta) = \begin{cases} I_{DC} + I_{pk} \cdot \cos(\theta) & -\alpha/2 < \theta < \alpha/2 \\ 0 & \text{elsewhere} \end{cases} \tag{2.73}$$

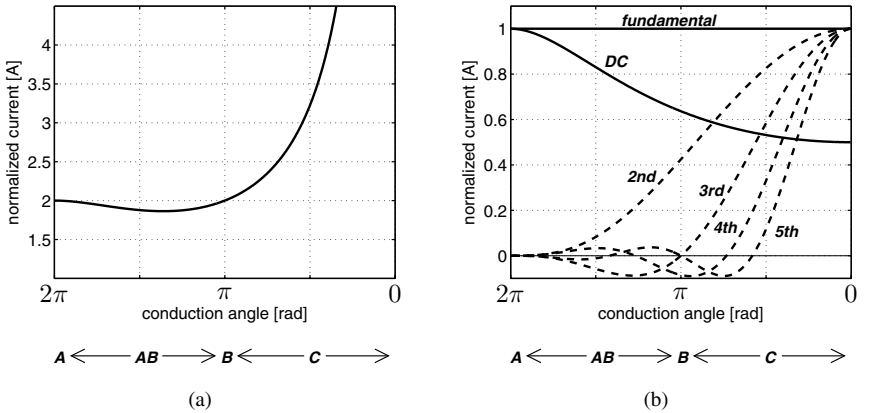


Figure 2.20. Analysis of the reduced conduction angle mode, (a) required value of I_{max} and (b) Fourier analysis of the drain current. Normalization: $V_{DD}=1\,\mathrm{V}$ and $R_L=1\,\Omega$.

where

$$\cos(\alpha/2) = -\frac{I_{DC}}{I_{pk}} \qquad \text{and} \qquad I_{pk} = I_{max} - I_{DC} \tag{2.74}$$

and thus

$$i_{DS}(\theta) = \frac{I_{max}}{1-\cos(\alpha/2)}\left(\cos(\theta) - \cos(\alpha/2)\right) \tag{2.75}$$

The value of the DC current and the harmonic current can be found by expanding the Fourier Series of $i_{DS}(\theta)$. The DC current is given by

$$I_{DC} = \frac{1}{2\pi}\int_{-\alpha/2}^{\alpha/2} \frac{I_{max}}{1-\cos(\alpha/2)}\left[\cos(\theta) - cos(\alpha/2)\right] d\theta \tag{2.76}$$

$$= \frac{I_{max}}{2\pi}\frac{2\sin(\alpha/2) - \alpha\cos(\alpha/2)}{1-\cos(\alpha/2)} \tag{2.77}$$

The magnitude of the n-th harmonic current is given by

$$I_n = \frac{1}{\pi}\int_{-\alpha/2}^{\alpha/2} \frac{I_{max}}{1-\cos(\alpha/2)}\left[\cos(\theta) - \cos(\alpha/2)\right]\cos(n\theta)d\theta \tag{2.78}$$

which, for the first harmonic, or fundamental current, results in

$$I_1 = \frac{I_{max}}{2\pi}\frac{\alpha - \sin(\alpha)}{1-\cos(\alpha/2)} \tag{2.79}$$

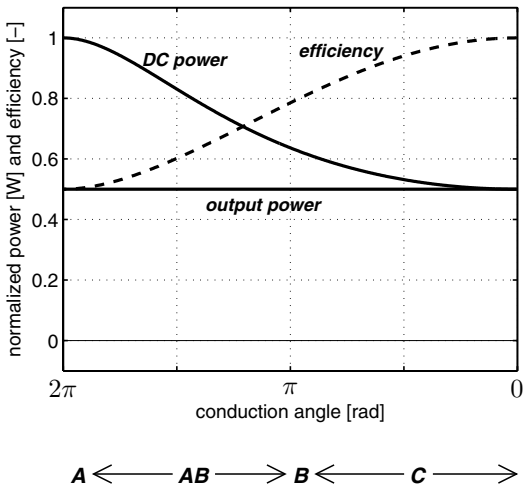


Figure 2.21. Output power, DC power consumption and efficiency versus conduction angle. Normalization: $V_{DD} = 1\,\mathrm{V}$ and $R_L = 1\,\Omega$.

The DC current I_{DC} will flow through the large inductor L_{DC}. The fundamental RF current (I_1) will flow trough the load resistor and the higher harmonic current (I_2,I_3 etc.) will flow through the capacitance C_0 of the harmonic trap. Hence, the fundamental output power is given by

$$P_o = R_L \left(\frac{I_1}{\sqrt{2}} \right)^2 \tag{2.80}$$

The DC power consumption of the amplifier is equal to

$$P_{DC} = V_{DD} I_{DC}. \tag{2.81}$$

and the efficiency can now be calculated as

$$\eta = \frac{P_o}{P_{DC}} \tag{2.82}$$

The output voltage will be a sine wave with an amplitude of $R_L I_1$ and this amplitude can maximally be equal to V_{DD}, as the transistor would otherwise start to clip. In real life, the amplitude will be smaller than V_{DD} because of the knee region of the transistor.

It is important to realize that for each conduction angle α, the value of I_{max} has to be adjusted in order to meet the maximum output swing. The required value for I_{max} can be obtained from

$$I_{max} = 2\pi I_1 \frac{1 - \cos(\alpha/2)}{\alpha - \sin(\alpha)} = 2\pi \frac{V_{DD}}{R_L} \frac{1 - \cos(\alpha/2)}{\alpha - \sin(\alpha)} \tag{2.83}$$

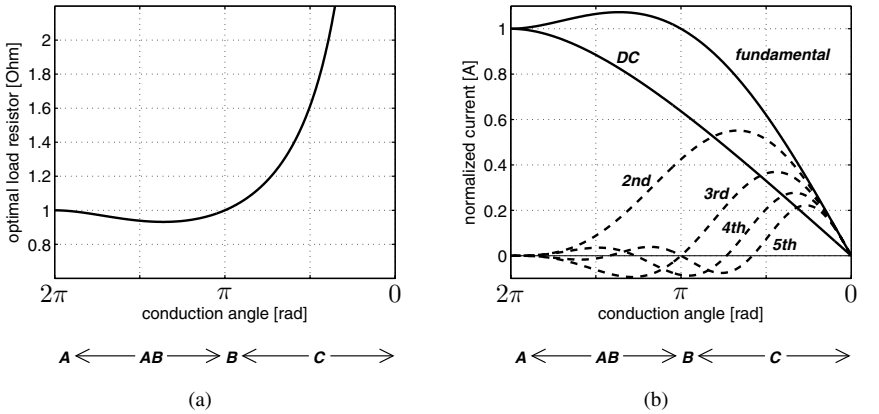


Figure 2.22. Analysis of the reduced conduction angle mode, (a) required value of R_L and (b) Fourier analysis of the drain current. Normalization: $V_{DD} = 1\,\mathrm{V}$ and $I_{max} = 2\,\mathrm{A}$.

To complete the analysis, the amplifier is normalized, i.e. $V_{DD} = 1\,\mathrm{V}$ and $R_L = 1\,\Omega$. Figure 2.20(a) shows the required value of I_{max} versus the conduction angle α. For very small conduction angles, the current will consist of a short but very high peak and for $\alpha = 0$, the current will be a Dirac pulse. Figure 2.20(b) depicts the Fourier analysis of the drain current. A conduction angle of 2π resembles a Class A amplifier. It can indeed be seen that for Class A, $I_{max} = 2\,\mathrm{V}$, $I_{DC} = 1\,\mathrm{A}$, $I_1 = 1\,\mathrm{A}$ and no harmonics are generated.

Figure 2.21 shows the output power, DC power consumption and efficiency versus conduction angle. Again, I_{max} is adjusted for a maximum voltage swing at the output and therefore the output power will be constant. Class B corresponds to a conduction angle of π, or a $50\,\%$ duty cycle. In Class B, the same maximum peak current as in Class A occurs, and the efficiency is higher ($\eta = \pi/4 \approx 78.5\%$) which clearly shows the benefit of reduced conduction angle operation. Reducing the conduction angle further towards zero (Class C) will further increase the efficiency, but the peak drain current rises to extremely high values and this will limit the operation at low conduction angles.

Previous calculations assumed that R_L is constant and I_{max} is adjusted to achieve the maximum permissible amplitude of V_{DD} at the output. This, however, will push the transistor to its current capability limits. Furthermore, the input signal has to be made very large to achieve these high current peaks. Therefore, it is more common to keep I_{max} fixed at the level of a Class A amplifier and to increase the load resistance in order to obtain a maximal voltage swing at the output. This is the approach which can be found in most textbooks, but one should be aware that the load resistor has to be adopted for every conduction angle.

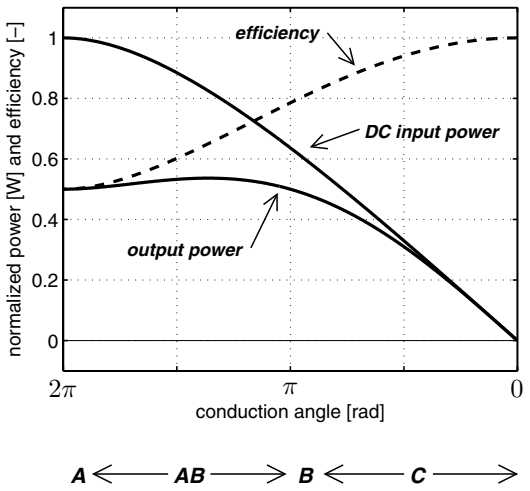


Figure 2.23. Output power, DC power consumption and efficiency versus conduction angle. Normalization: $V_{DD} = 1\,\mathrm{V}$ and $I_{max} = 2\,\mathrm{A}$.

The analysis remains the same, but the normalization is done as $V_{DD} = 1\,\mathrm{V}$ and $I_{max} = 2\,\mathrm{A}$, and the required value of R_L can be calculated as

$$R_L = \frac{V_{DD}}{I_1} = 2\pi \frac{V_{DD}}{I_{max}} \frac{1 - \cos(\alpha/2)}{\alpha - \sin(\alpha)} \tag{2.84}$$

The value of R_L is depicted in figure 2.22(a), and figure 2.22(b) shows the Fourier analysis of the drain current.

The output power, DC power consumption and efficiency can be found in figure 2.23. Now, for small values of α (i.e. Class C), the output power goes to zero. The reason for this is clear: the output voltage amplitude is still equal to V_{DD} but the load resistance goes to infinite. Also note that in Class C, both P_o and P_{DC} fall together which explains the seemingly contradicting 100% efficiency although no output power is delivered. The low output power of a Class C will make it less attractive in low-voltage applications, despite its high efficiency. Again, the Class B seems more favorable as its output power is equal to a Class A, and the efficiency is increased to about 78.5%

The maximum drain voltage of a reduced conduction angle amplifier is independent of the conduction angle and equals $v_{DS,max} = 2 \cdot \mathrm{V_{DD}}$ Hence, the maximum output power is written as

$$P_{o,max} = \frac{1}{8} \cdot \frac{v_{DS,max}^2}{R_L} = 0.125 \cdot \frac{v_{DS,max}^2}{R_L} \tag{2.85}$$

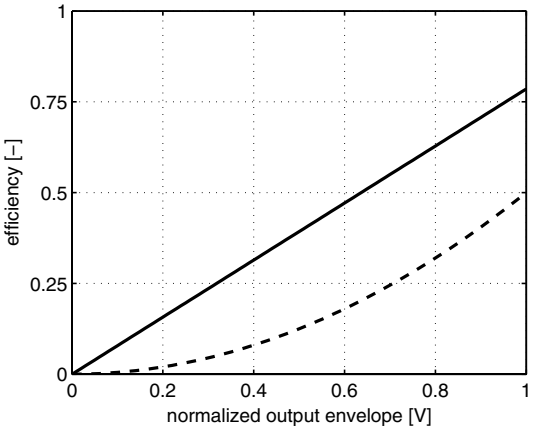


Figure 2.24. Efficiency versus output envelope for Class B (solid line) and Class A (dashed line).

The Class B seems the most promising reduced conduction angle amplifier and is therefore discussed in more detail. For a Class B amplifier $(\alpha = \pi)$, the output power is equal to

$$P_o = \frac{V_o^2}{2R_L} \tag{2.86}$$

and the DC power consumption equals

$$P_{DC} = V_{DD} I_{DC} = V_{DD} \frac{I_{max}}{\pi} \tag{2.87}$$

It can be derived that

$$I_1 = \frac{I_{max}}{2} = \frac{V_o}{R_L} \tag{2.88}$$

and thus

$$P_{DC} = \frac{2}{\pi} \frac{V_{DD} V_o}{R_L} \tag{2.89}$$

From these equations, the efficiency of a Class B can be obtained as

$$\eta = \frac{\pi}{4} \frac{V_o}{V_{DD}} \tag{2.90}$$

This efficiency curve is depicted in figure 2.24 and compared to the Class A efficiency curve. Besides the fact that Class B achieves a higher peak efficiency, more important is that the efficiency is linearly dependent on V_o, in contrast to the quadratic dependency for Class A.

To conclude, the Class B amplifier requires twice the amount of voltage swing at the gate of the transistor compared to a Class A amplifier. In other

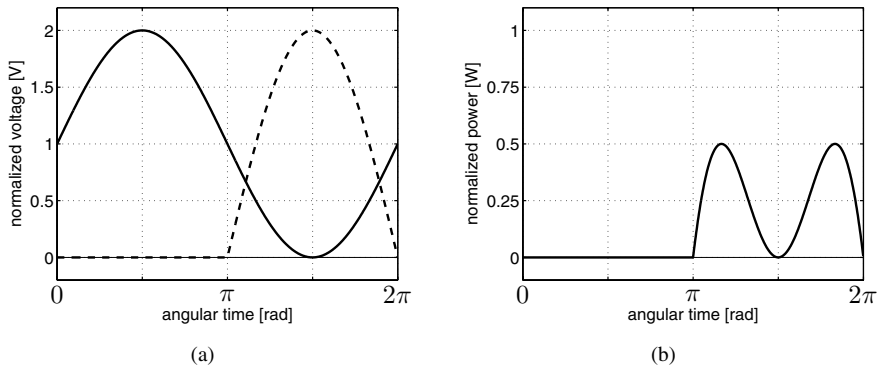


Figure 2.25. The Class B amplifier: (a) normalized drain voltage and drain current (dashed line) and (b) normalized power dissipation.

words, the gain of a Class B amplifier is only half the gain of a Class A. This effect becomes more pronounced if the conduction angle is further reduced to Class C operation. As already indicated in section 2.3.4, high efficiency and high gain are, to some extend, contradicting objectives.

As another remark, one should be aware that the high quality harmonic trap at the output of the amplifier is crucial for correct reduced conduction angle operation. For a high voltage vacuum tube amplifier, the load resistor will be relatively high for a certain output power and this makes it easier to achieve a high quality LC network. In low voltage technologies, the load resistor will also be low to achieve sufficient output power. As a consequence, the quality factor of the LC network will be reduced which makes the design of a Class B or Class C power amplifier a difficult task.

Finally, the push-pull Class B amplifier is not tackled here, but an elaborate discussion can be found in [Crip99] and [Keni00]. As an RF amplifier, the push-pull is less suited since two large transistors need to be driven. Furthermore, the pMOST has less transconductance for the same DC current compared to an nMOST, and the width of the pMOST typically has to be three times larger. Hence, the input capacitance of the push-pull amplifier will approximately be four times as large compared to the Class B amplifier discussed here.

2.4.3 Saturated Class A

For a reduced conduction angle amplifier, the transistor will be turned off for a certain amount of time, depending on the bias as well as the amplitude

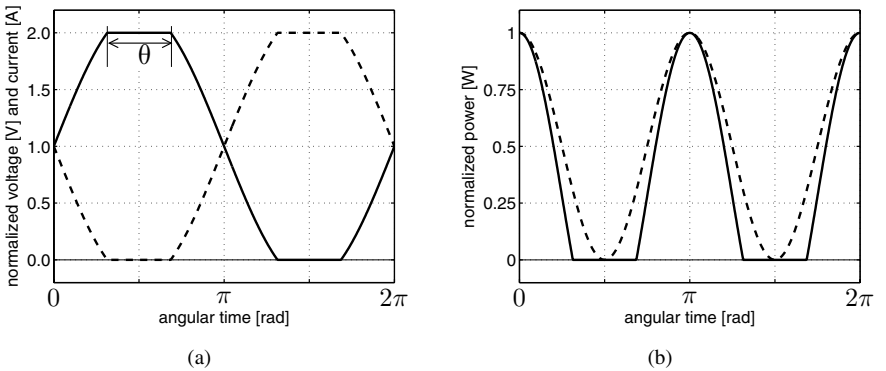


Figure 2.26. Normalized waveforms of a saturated Class A power amplifier: (a) drain voltage (solid line) and drain current (dashed line), (b) normalized power dissipation of a saturated Class A (solid line) and a linear Class A (dashed line)

of the RF input voltage. This reduces the voltage-current overlap resulting in a higher efficiency. The drawback is a reduction of the output power.

If the bias of a Class A amplifier is unaffected, but the amplitude of the input voltage is increased, the output waveform will no longer be a sinusoid as it would start to clip to the supply voltage and to the ground. This *re-shaping* will also reduce the overlap between the current through the transistor and de drain-source voltage across the transistor. A higher efficiency will be the result, but the input-output amplitude linearity is now distorted.

Figure 2.26(a) shows the resulting current and voltage waveforms of an over-driven or saturated Class A amplifier. To demonstrate the increase in efficiency, figure 2.26(b) depicts the power dissipation in the transistor for both Class A and over-driven Class A. It is clear that, due to the clipping behavior, the transistor now dissipates during a smaller amount of time.

From the waveforms of figure 2.26, the harmonic content of the output voltage and output power can be calculated as a function of the overdrive angle θ and the results are shown in figure 2.27 [Snid67]. The higher the overdrive angle, the higher the output power. For an overdrive angle of π, which means that the output voltage looks like a square wave, the normalized output power has increased from $0.5\,\mathrm{W}$ to $0.81\,\mathrm{W}$. The voltage and current waveforms are still symmetrical, and therefore the DC power consumption will not change as the amplifier is pushed into saturation. Therefore, the efficiency will also increase from $50\,\%$ to $81\,\%$.

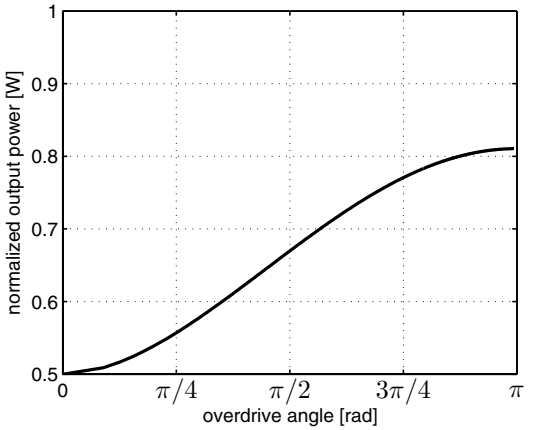


Figure 2.27. Normalized output power of an over-driven Class A power amplifier versus overdrive angle.

The overdrive angle can be converted to a normalized input power, and the result of this is shown in figure 2.28. For low input signals, the output voltage is proportional to the input voltage. Again, the theoretical efficiency at maximum linear output power is 50%. In the saturated or over-driven region, the output power will slightly increase and a higher efficiency will be the result.

At very high input power levels, the MOS transistor will switch between the off-state and the linear region, just like a digital buffer. If the switch is assumed to be ideal ($R_{on} = 0\,\Omega$), no power will be dissipated in the transistor. The efficiency is however not equal to 100% as one might expect. The reason is the presence of the higher harmonics in the output waveform.

The calculation is as follows. The output voltage is assumed to be a square wave with an amplitude of V_{DD}. Therefore, the total output power (including the power in the higher harmonics) is equal to

$$P_{o,tot} = \frac{V_{DD}^2}{R_L} \tag{2.91}$$

Since the transistor is acting as an ideal switch, no current-voltage overlap is present, the conversion efficiency must be 100% and therefore, the DC power consumption must equal to $P_{o,tot}$. The fundamental output voltage has an amplitude equal to

$$V_o = V_{DD} \cdot 4/\pi \tag{2.92}$$

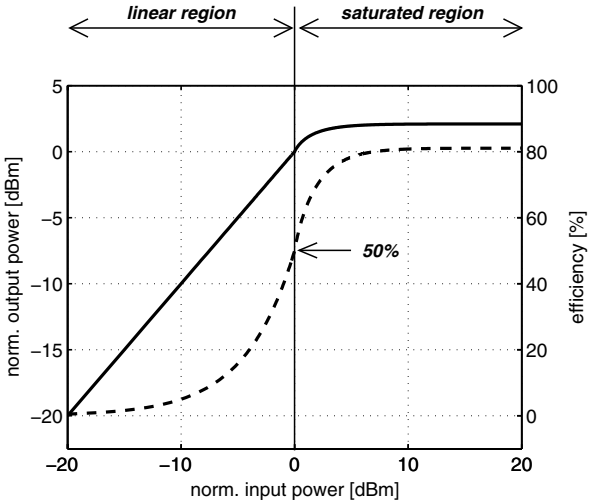


Figure 2.28. Output power (solid line) and efficiency (dashed line) of a Class A power amplifier versus input power.

as given by the Fourier expansion of a square wave, and the corresponding output power at the fundamental frequency is equal to

$$P_o = \frac{(V_{DD} \cdot 4/\pi)^2}{2R_L} \tag{2.93}$$

The efficiency of the amplifier can now be calculated as

$$\eta = \frac{(4/\pi)^2}{2} = 8/\pi^2 \approx 81\% \tag{2.94}$$

Although the saturated Class A has a conversion efficiency of 100%, the actual PA efficiency is only 81% in the ideal case. This is also the reason why a *simple* digital switching buffer, although this circuit achieves a 100% DC-to-RF conversion efficiency, is not well suited as an RF power amplifier. The theoretical efficiency is only 81%, while other switching amplifiers are capable of achieving a 100% fundamental efficiency.

The maximum drain voltage of a saturated Class A amplifier equals $2 \cdot V_{DD}$ and thus, the maximum output power can also be written as

$$P_{o,max} = \frac{2}{\pi^2} \cdot \frac{v_{DS,max}^2}{R_L} \approx 0.2026 \cdot \frac{v_{DS,max}^2}{R_L} \tag{2.95}$$

A drawback of saturating a power amplifier is the reduction of the power-gain, defined as the ratio of the output power to the input power. To drive a

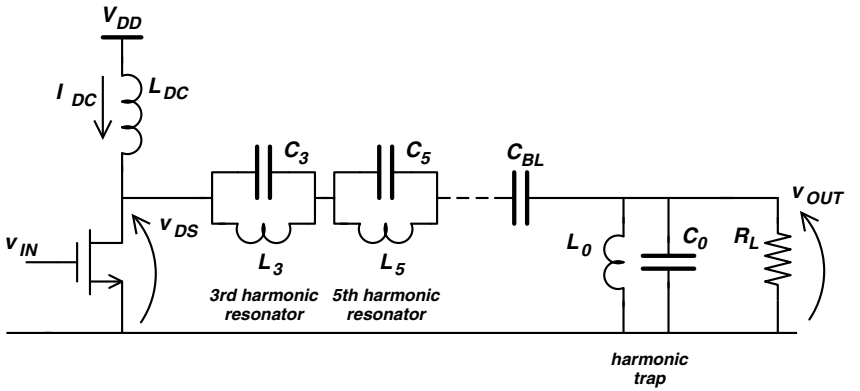


Figure 2.29. Class F power amplifier with harmonic resonator tuning.

power amplifier in saturation, the input power has to be increased. The output power on the other hand more or less stays the same and the power-gain of the amplifier will be reduced. In other words, a saturated Class A amplifier combines both high output power with a decent efficiency, but the power gain will be low.

2.4.4 Harmonic Tuning for Improved Efficiency: Class F

From section 2.4.2, it was found that the efficiency can be increased by reducing the conduction angle. In particular the Class B operation mode is very promising, as it is able to generate the same amount of output power as the Class A amplifier, and the efficiency is increased from 50% to 78.5%. Further reducing the conduction angle shows no benefit since the output power shows a drastic reduction.

From the previous sections, it was found that the efficiency of the Class A amplifier can be increased by pushing the amplifier in overdrive. However, this also requires more RF input power and it will reduce the gain. The *flattening* of the sine wave can also be accomplished by external passive elements. The transistor does not have to be over-driven and the gain will remain high.

To combine the best of both worlds, a Class B amplifier, with a 50% conduction time, is extended with some harmonic resonators that further reduce the voltage-current overlap. This operation mode is designated as Class F and figure 2.29 depicts the basic circuit for Class F operation. The shape of the output voltage is still determined by the harmonic trap across the load resistor R_L. If the quality factor of this harmonic trap is high enough, the output voltage will be a sine wave. The drain-source voltage v_{DS} is now equal to the output voltage, plus the voltage across each of the harmonic resonators, that

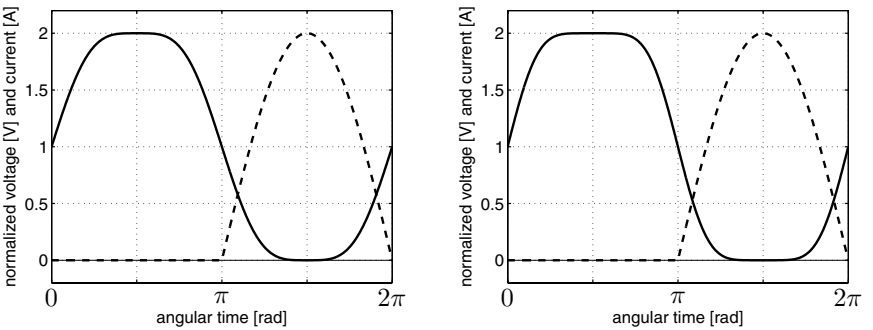


Figure 2.30. Normalized drain voltage (solid line) and drain current (dashed line) of a Class F amplifier with (a) third and (b) third and fifth harmonic peaking.

Table 2.2. Class F performance overview

included harmonic		3	5	∞
Class	B	F3	F5	D
peak efficiency	78.5%	88.4%	92.0%	100%
normalized output power [W] ($V_{DD} = 1V$ and $R_L = 1\Omega$)	0.5000	0.6328	0.6866	0.8106
maximum output power [W] ($v_{DS,max} = 1V$ and $R_L = 1\Omega$)	0.1250	0.1582	0.1717	0.2026

are tuned to odd multiples of the fundamental frequency. The addition of the odd harmonics makes the drain-source voltage look like a square wave and the voltage-current overlap is therefore reduced. The drain current is still a half sine wave, just like in the Class B amplifier. Figure 2.30(a) depicts the drain voltage and drain current of a Class F with addition of the third harmonic and figure 2.30(b) shows the result for both third and fifth harmonic flattening. It can clearly be seen that the voltage-drain overlap, and hence the dissipation in the transistor, is reduced compared to a Class B amplifier.

Of course, the amplitudes of the higher harmonics must be well selected in order to maximize the output power and efficiency. Normally, these amplitudes are selected to achieve a maximally flat waveform at the drain, like in figure 2.30 [Raab97].

Table 2.2 summarizes the performance of a Class F amplifier. Addition of the third harmonic resonator (F3) increases the efficiency from 78.5% to 88.4%. The normalized output power has increased from $0.5\,\mathrm{W}$ to $0.63\,\mathrm{W}$. If the fifth harmonic is included as well (F5), the efficiency further increases to 92% and the normalized output power becomes $0.69\,\mathrm{W}$ [Raab01]. If all odd harmonics would be added, the drain-source voltage would be a perfect square wave, resulting in a 100% efficiency and a normalized output power of $8/\pi^2 \approx 0.81\mathrm{W}$. This power is exactly the same as the normalized output power of the the saturated Class A amplifier, but the harmonics are not dissipated in the load resistor and therefore the efficiency does achieve 100%. Such an amplifier can also be regarded as a (harmonic) Class D amplifier (see section 2.4.6).

The maximum drain voltage in Class F mode is, like all previous amplifiers, equal to $2 \cdot V_{DD}$. The maximum output power of a Class F with third harmonic peaking is therefore

$$P_{o,max} = \frac{(9/8)^2}{8} \cdot \frac{v_{DS,max}^2}{R_L} \approx 0.1582 \cdot \frac{v_{DS,max}^2}{R_L} \tag{2.96}$$

and for third and fifth harmonic peaking:

$$P_{o,max} = \frac{(75/64)^2}{8} \cdot \frac{v_{DS,max}^2}{R_L} \approx 0.1717 \cdot \frac{v_{DS,max}^2}{R_L} \tag{2.97}$$

If all harmonics are included, the output power becomes

$$P_{o,max} = \frac{2}{\pi^2} \cdot \frac{v_{DS,max}^2}{R_L} \approx 0.2026 \cdot \frac{v_{DS,max}^2}{R_L} \tag{2.98}$$

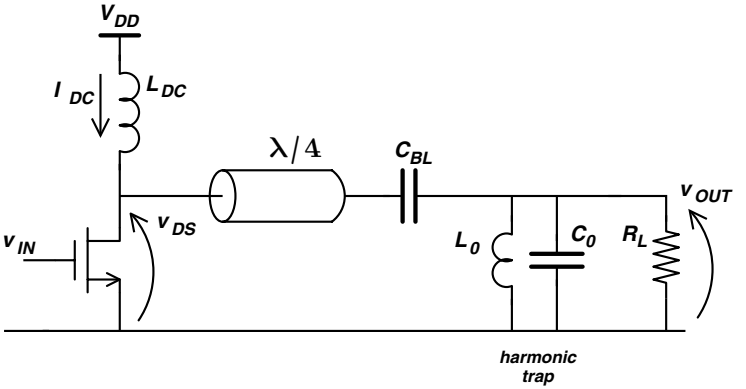
which is the same result as an over-driven Class A amplifier.

To have an ideal Class F amplifier, an infinite number of harmonics need to be summed. In practice, this can be achieved by using a $\lambda/4$ transmissionline as shown in figure 2.31. The use of a $\lambda/4$ transmissionline will, however, increase the required area to implement the amplifier. The length of the stripline can be calculated from

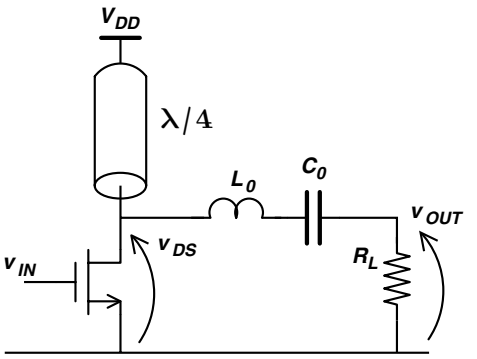
$$\frac{\lambda}{4} = \frac{c}{4\sqrt{\epsilon_r}\, f} \tag{2.99}$$

For example, at $1\,\mathrm{GHz}$ and with a relative permitivity of $\epsilon_r = 9$, the length of the stripline is $25\mathrm{mm}$. A huge area compared to chip-scale dimensions. This makes the (ideal) Class F amplifier less attractive if one strives for a high level of integration [Kuo01].

The Class F amplifier is sometimes designated as a *switching* amplifier. However, only when an infinite number of harmonics are added (i.e. harmonic Class D), the transistor can be replaced by the switch. Otherwise, the transistor has the same drive requirements as a Class B amplifier [Raab75].



(a)



(b)

Figure 2.31. Class F power amplifier with transmissionline tuning.

Inverted Class F, or Class F $^{-1}$ is derived from the *original* Class F and makes use of even-harmonic resonators. In such an amplifier, the drain-source voltage will tend to a half sine wave and the drain current will look like a square wave. In other words, the roles of current and voltage in a Class F amplifier can be reversed [Fort01],[Raab97]. The only drawback of an inverse Class F is the higher peak drain voltage and thus a lower maximum output power.

$$v_{DS,max} = \pi \cdot V_{DD} \tag{2.100}$$

$$P_o = \frac{\pi^2}{8} \cdot \frac{V_{DD}^2}{R_L} \approx 1.2337 \cdot \frac{V_{DD}^2}{R_L} \tag{2.101}$$

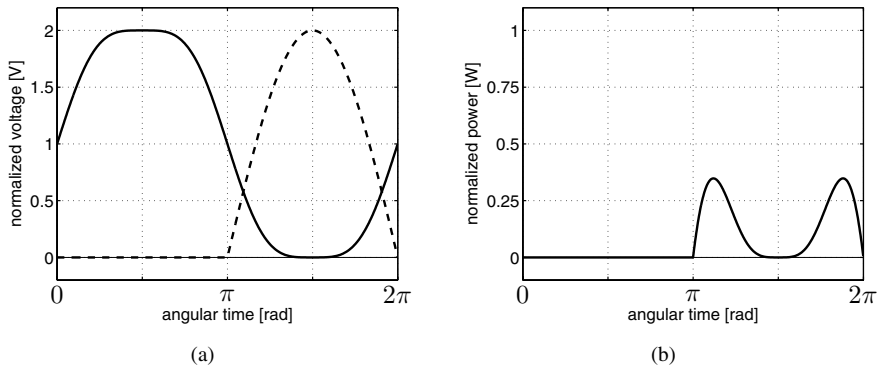


Figure 2.32. The Class F3 amplifier: (a) normalized drain voltage (solid line) and drain current (dashed line) and (b) normalized power dissipation.

$$P_{o,max} = \frac{1}{8} \cdot \frac{v_{DS,max}^2}{R_L} \approx 0.125 \cdot \frac{v_{DS,max}^2}{R_L} \tag{2.102}$$

The advantage of an inverse Class F with only second harmonic peaking compared to a Class F with third harmonic peaking, is the lower operating frequency of the resonator; only two times ω_0 instead of three times.

Although the Class F amplifier looks attractive at first sight, several issues make the integration in CMOS difficult. As stated before, the length of a $\lambda/4$ transmissionline is too high to be integrated, unless one moves to frequencies above $10\,\mathrm{GHz}$. On the other hand, if only third and fifth harmonic peaking is used, the (theoretical) efficiency already increases to 88% and 92% respectively. However, in CMOS the major issue is the large drain-source parasitic capacitance. This capacitance will provide a low-impedance path for the higher harmonics, and the *squaring* will be less pronounced, again reducing the benefit of Class F. If the $\lambda/4$ transmissionline is approximated with lumped elements, the parasitic drain capacitance can become part of the stripline approximation and as such, its influence can partially be reduced [HH04].

Finally, figure 2.32 shows the voltage, current and transistor dissipation of Class F3 amplifier with third harmonic peaking.

2.4.5 Switching Amplifiers

In a switching amplifier, as the name suggests, the transistor acts as a simple switch. This behavior was already encountered in the over-driven Class A amplifier of section 2.4.3. Although the switch itself has no losses and dissipates no heat, the drain efficiency of that simple amplifier did not reach 100%. The

conversion efficiency is indeed 100%, but harmonic power is lost in the load resistor. Therefore, like with the Class F amplifier, a tuning network is needed to avoid harmonic power being wasted. This will lead us to the Class D and Class E power amplifiers.

A major difference between a switching amplifier and the Class A, AB, B, C and F amplifiers discussed before, is the hard nonlinearity of the amplifier. Although Class AB and C are nonlinear, still some relationship exists between the input amplitude and the output amplitude. A switch, on the other hand, can only be on or off. In other words, as long as the input voltage is large enough to switch on the transistor, the output will have a constant amplitude which will only be dependent on the supply voltage.

2.4.6 Class D

The basic Class D power amplifier is depicted in figure 2.33(a). The input signal v_{IN} and the output signal v_{DS} are both square waves. As already pointed out, if this square wave is directly applied to the load resistor R_L, a lot of harmonic power is wasted. However, the series tank consisting of L_0 and C_0, only allows a sinusoidal current through R_L. Therefore, the drain efficiency will be 100%, at least in the ideal case.

Figure 2.33(b) depicts the drain voltage and drain current waveform of an ideal Class D amplifier. The drain voltage will be a square wave with an amplitude of $V_{DD}/2$ and a DC value of $V_{DD}/2$. The first harmonic of that square wave has an amplitude of $V_{DD} \cdot 2/\pi$ and therefore the output current has an amplitude of

$$I_o = \frac{2/\pi\, V_{DD}}{R_L} \tag{2.103}$$

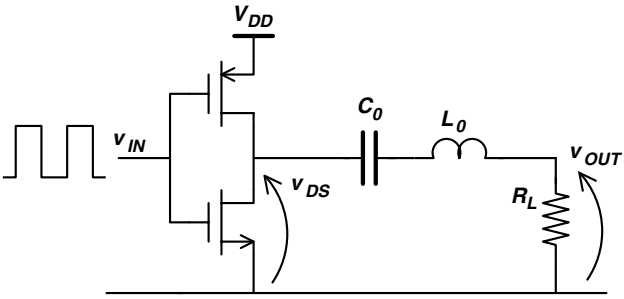
and the output power will be equal to

$$P_o = \frac{2}{\pi^2}\frac{V_{DD}^2}{R_L} \tag{2.104}$$

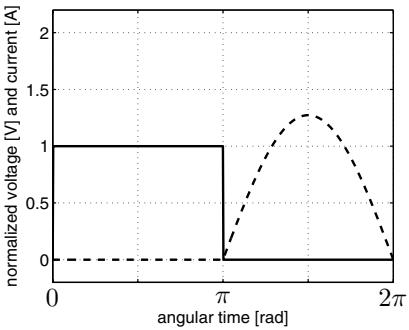
The peak drain voltage is equal to V_{DD} and therefore, the maximum output power becomes

$$P_{o,max} = \frac{2}{\pi^2} \cdot \frac{v_{DS,max}^2}{R_L} \approx 0.2026 \cdot \frac{v_{DS,max}^2}{R_L} \tag{2.105}$$

It is not so easy to integrate the Class D amplifier in CMOS at higher frequencies. A first drawback is the *hard switching* property of the amplifier, since the switch will close while the voltage across the switch in not equal to zero. In CMOS, the switch will have a large parasitic drain-source capacitance. If the transistor is turned on, the charge on the parasitic capacitance will correspond



(a)



(b)

Figure 2.33. The Class D power amplifier: (a) schematic and (b) normalized drain voltage (solid line) and drain current (dashed line).

to an energy of

$$E_C = \frac{CV^2}{2} \tag{2.106}$$

and this energy is dissipated in the switch. Therefore,

$$P_{diss} = \frac{1}{2} f C V^2 \tag{2.107}$$

Another important loss mechanism is the finite turn-on and turn-off times of the transistor [Keni00], again resulting in an efficiency degradation.

A third drawback of this switching Class D amplifier is the high drive power that is required. Indeed, two switches need to be driven. Especially the pMOS transistor will degrade the overall efficiency due to its lower on resistance. Typically, the size of the pMOST has to be three times larger than the size of

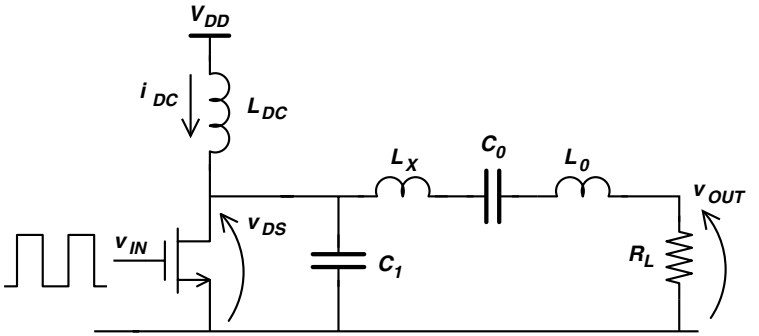


Figure 2.34. Class E power amplifier.

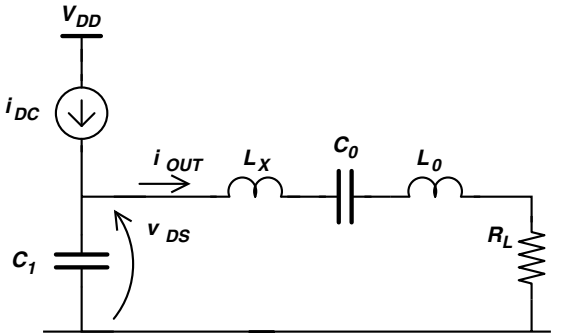


Figure 2.35. Simplified circuit to analyze the Class E waveform.

the nMOST. The resulting high input capacitance makes the Class D amplifier less attractive for RF applications.

2.4.7 Class E

Like Class D, the Class E power amplifier is also capable of achieving a 100% efficiency. The basic circuit of a Class E amplifier is depicted in figure 2.34 [Soka75]. Due to the tuned series tank (L_0 and C_0) the output voltage will be sinusoidal, and no harmonic power will be dissipated. The nMOS transistor acts as a switch.

The Class E amplifier is entirely designed in the time domain. If the switch is closed, the DC current from inductor L_{DC} will flow through the switch. If the switch opens, the DC current through minus the sinusoidal output current will be dumped in the capacitor C_1. This will result in a second order voltage response across C_1, which is equal to the voltage across the switch.

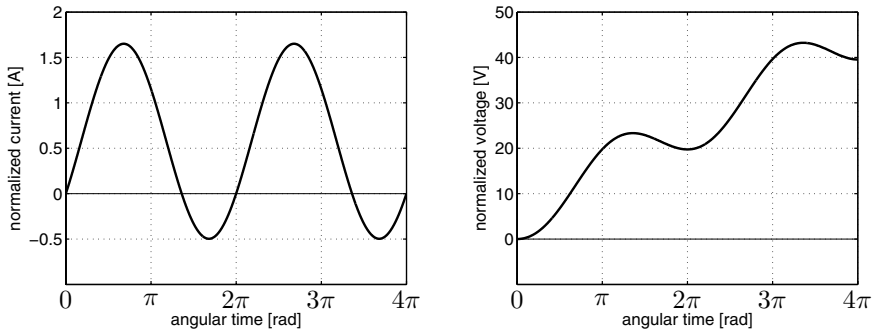


Figure 2.36. Waveforms of the simplified circuit of figure 2.35. (a) capacitor current and (b) capacitor voltage.

To better understand the principles of a Class E amplifier, one can leave out the switching transistor and assume that the output current is a pure sine wave [Fran00]. The resulting schematic is depicted in figure 2.35. The current through capacitor C_1 will now be equal to

$$i_{C_1}(t) = I_{DC} - I_o \sin(\omega t + \varphi) \tag{2.108}$$

and the resulting voltage across the capacitor will become

$$\begin{aligned} v_{C_1}(t) &= \frac{1}{C_1} \int_0^{\infty} (I_{DC} - I_o \sin(\omega t + \varphi))\, dt \\ &= \frac{1}{C_1} \left(I_{DC} \cdot t + \frac{I_o}{\omega} \cdot \cos(\omega t + \varphi) \right) + K_0 \end{aligned} \tag{2.109}$$

in which K_0 represents the voltage across the capacitor at $t = 0$ and can be chosen equal to zero. The angle of the output current (φ) will depend on the value of L_x.

Figure 2.36(a) depicts the current through C_1 and figure 2.36(b) shows the resulting voltage across C_1. In these figures, the normalization of $V_{DD} = 1\,\mathrm{V}$, $R_L = 1\,\Omega$ is selected and all other parameters are calculated using the standard Class E design equations [Raab77].

The voltage across the capacitor, which is equal to the voltage across the switch, shows an interesting shape between π and 2π as indicated on figure 2.37(b). If the transistor is switched on from 0 to π, the voltage at time π will be equal to zero. Mathematically, this can be accomplished by choosing another value for the integration constant K_0 in equation 2.109. The most interesting

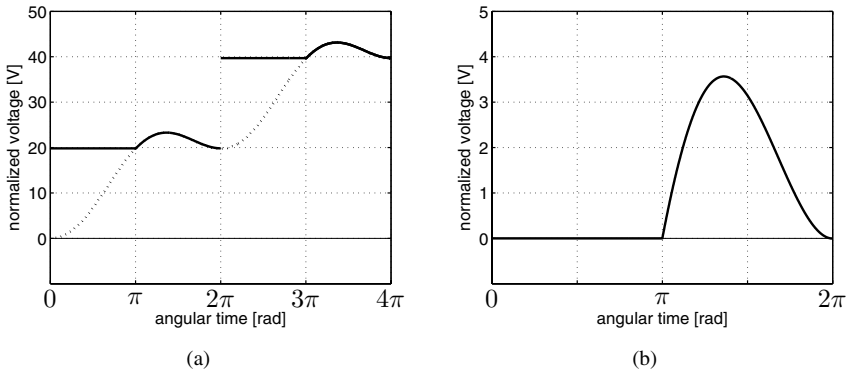


Figure 2.37. Waveforms of the simplified circuit of figure 2.35. (a) capacitor voltage and (b) magnified version where it is assumed that the switch is closed from 0 to π and from 2π to 3π.

property of this waveform occurs at 2π, since at that point the switch voltage and it first derivative both become zero again. Thus, if the switch is turned on again at 2π, the switching losses have been made equal to zero.

This is *the* big difference compared to a hard-switching Class D, where a parasitic switch capacitance will always lead to power dissipation and a reduction of the efficiency. In Class E, the switch is closed when the voltage becomes zero and therefore no switching losses occur. This is also referred to as *soft switching* or *zero voltage switching, (ZVS)* . In fact, the parasitic capacitance of the switch can become part of C_1. In other words, the parasitic drain-source capacitance can become part of the amplifier circuit, which is a huge advantage, especially in CMOS.

The Class E theory, as stated by Sokal and Sokal in 1975 [Soka75], thus requires that both the switch voltage and its first derivative are zero when the switch closes. In other words:

$$\text{Class E} \Leftrightarrow \begin{cases} v_{DS}(t=t_1) &= 0 \\ \dfrac{dv_{DS}(t)}{dt}\Big|_{t=t_1} &= 0 \end{cases} \tag{2.110}$$

The requirement for a zero first derivative is not crucial to achieve a 100% efficiency. However, this property makes the amplifier less sensitive to component variations and the output power is maximized for a given peak drain voltage [Raab77].

In the *original* Class E power amplifier, inductor L_{DC} is assumed to be large, the current through the load resistor is assumed to be a pure sine wave

and no losses are included. Under these conditions, analytical design equations can be derived and are given by

$$C_1 = \frac{8}{\pi(\pi^2+4)}\frac{1}{\omega R_L} \approx 0.1836\frac{1}{\omega R_L} \tag{2.111}$$

$$L_x = \frac{\pi(\pi^2-4)}{16}\frac{R_L}{\omega} \approx 1.1525\frac{R_L}{\omega} \tag{2.112}$$

$$P_o = \frac{8}{\pi^2+4}\frac{V_{DD}^2}{R_L} \approx 0.5768\frac{V_{DD}^2}{R_L} \tag{2.113}$$

The major drawback of the Class E amplifier is the high drain voltage that occurs when the switch is open. This value is, in the ideal case, given by

$$v_{DS,max} = 2\pi\left[\frac{\pi}{2} - \arctan\left(\frac{\pi}{2}\right)\right]\cdot V_{DD} \approx 3.5620 \cdot V_{DD} \tag{2.114}$$

and the maximum output power can now be written as

$$P_{o,max} = 0.0455 \cdot \frac{v_{DS,max}^2}{R_L} \tag{2.115}$$

This value is much lower compared to Class D, Class F and overdriven Class A. On the other hand, since the transistor is a switch, it will not conduct current during the high voltage peak and when the transistor is switched on, the drain-source voltage is zero. Hence, the value of $v_{DS,max}$ can be much larger compared to Class D, F, B and A.

In an actual implementation, the current through inductor L_1 will not be constant, but this is not an obstruction to meet the Class E requirements. It is also possible to design a Class E amplifier with only one inductor and one capacitor, i.e with a combination of a tuned network, a load resistor and a switch [Soka81] [Kazi87a]. Such a simplified circuit is depicted in figure 2.38(a). If the switch is closed, the current through the inductor will start to increase. If the switch opens, that current will be dumped in the capacitor and the resistor. Due to the second-order nature of the network, the switch voltage will have a shape as given by figure 2.38(b). One can define the quality factor of the network as

$$Q = R_L\sqrt{\frac{C_1}{L_1}} \tag{2.116}$$

For a high value of Q, the switch voltage will show a lot of oscillations. For a very low value of Q, the voltage will only have one or even no oscillations at all. For a specific value of Q, the switch voltage and its first derivative become zero at a certain time (the solid line in figure 2.38(b)). The switch can

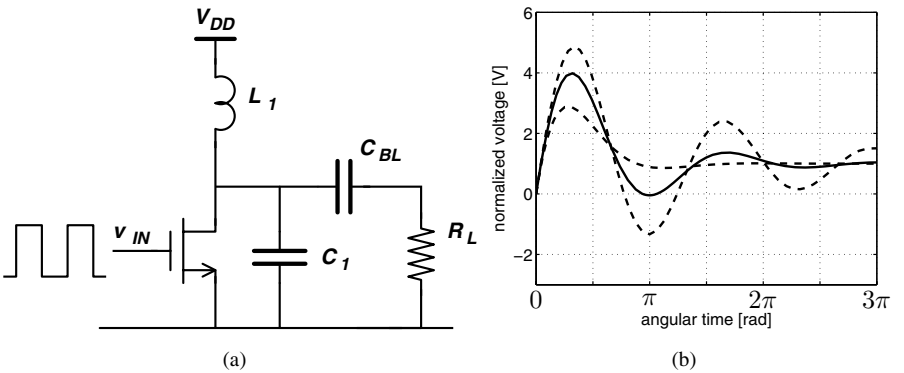


Figure 2.38. The Class E amplifier with only one inductor and one capacitor: (a) schematic and (b) switch voltage.

be closed at this time, no switching loss occurs and the Class E requirements are met. However, for the simple circuit of figure 2.38(a), harmonic power is lost in the load resistance and thus the efficiency, for this example, will not be 100%. Further details regarding the design and CMOS integration of a Class E amplifier, are discussed in chapter 3.

As it is a switching amplifier, the Class E amplifier is inherently nonlinear and therefore not suited to amplify amplitude modulated signals. However, for phase modulated and thus constant envelope signals, the Class E amplifier can be applied. The only means to change the output power, is by changing the supply voltage of the amplifier. The efficiency of the Class E amplifier is independent of the supply voltage and thus, changing the power supply of a Class E amplifier is an efficient means to change the amplitude of the output RF signal. This idea is further elaborated in chapter 5.

Finally, the application of the Class E amplifier is not limited to RF amplification. It can also used at lower frequencies where it serves as an efficient DC-to-AC power converter to feed implantable and wearable biomedical devices, as shown in [Schu98] and [Catr04].

2.4.8 Reliability

Power amplification of electrical signals requires both a high voltage and a high current. As a consequence, in a power amplifier the transistor is pushed to its voltage and current limits. Focusing on CMOS, the current and voltage limits of the nMOS transistor are mainly determined by following mechanisms

- hot carrier injection

- junction breakdown and punch-through
- time dependent dielectric breakdown

An elaborate discussion on this topic can be found in [Ramo05] and [Mert05] and will not be repeated here. The general conclusion from both is that hot carrier injection plays a dominant role for the reliability and for the calculation of the nMOS transistor lifetime.

The generation of hot carriers is as follows. In an nMOS transistor, electrons travel through the channel, from the source to the drain. If a high lateral electric field is present in the channel, these electrons may gain sufficient energy to cause impact ionisation, due to collisions with the Silicon lattice at the drain side. Impact ionisation will in its turn generate high energetic electrons/hole pairs, also denoted as *hot carriers*. The hot electrons can tunnel through the gate oxide, or may get trapped inside the gate oxide. These trapped charges will cause a shift in the nMOS parameters, like threshold voltage and saturation current, and on a longer term the trapped electrons may initiate a breakdown of the oxide.

From the previous discussion, it is clear that hot carriers are only generated if a large amount of electrons in the channel is accelerated by a high lateral electric field. In other words, it requires the presence of both a high drain current and a high drain voltage. In current submicron and nanometer technologies, the presence of a lightly-doped drain alleviates the hot carrier problem and in most cases reliable operation can be obtained if the drain voltage is kept below the nominal supply voltage of the technology. In some extreme cases, the gate length has to be increased to extend the lifetime of the transistor. With this in mind, the different power amplifier classes are now reviewed.

2.4.8.1 Class A

In a Class A amplifier, the transistor is continuously conducting current and the maximum drain voltage rises to two times the applied supply voltage. As such, the supply voltage of the Class A amplifier has to be kept below half the nominal supply voltage of the actual CMOS technology. The maximum output power of the Class A amplifier is therefore

$$P_{o,max,A} = 0.125 \cdot \frac{V_{DD,nom}^2}{R_L} \tag{2.117}$$

2.4.8.2 Class B

As said before, the Class B amplifier only conducts current half the time and figure 2.25 depicts the normalized voltage and current waveforms. When the nMOS transistor starts to conduct current, the drain voltage is equal to the supply voltage of the amplifier. To avoid hot carrier generation, the supply voltage

of the power amplifier thus has to be equal to the nominal supply voltage of the technology. As such, the maximum drain voltage of the Class B can go up to two times the nominal supply voltage of the technology without damaging the transistor, at least with respect to hot carriers. The maximum output power of the Class B amplifier thus becomes

$$P_{o,max,B} = 0.5 \cdot \frac{V_{DD,nom}^2}{R_L} \tag{2.118}$$

2.4.8.3 Saturated Class A

In the saturated Class A, the drain voltage and drain current will approximate a square wave. The drain voltage is equal to two times the supply voltage when the transistor starts to conduct current. Therefore, to avoid hot carrier injection, the supply voltage of the saturated Class A amplifier has to be lower than half the nominal supply voltage of the technology. Thus,

$$P_{o,max,Sat.A} = 0.2026 \cdot \frac{V_{DD,nom}^2}{R_L} \tag{2.119}$$

2.4.8.4 Class F

The transistor in the Class F amplifier starts to conduct current when the drain voltage is equal to the supply voltage of the amplifier. The supply voltage of the amplifier can thus be make equal to the nominal supply voltage of the technology. Therefore, the maximum output power of a Class F with third harmonic peaking becomes

$$P_{o,max,F3} = 0.6328 \cdot \frac{V_{DD,nom}^2}{R_L} \tag{2.120}$$

and for third and fifth harmonic peaking:

$$P_{o,max,F5} = 0.6868 \cdot \frac{V_{DD,nom}^2}{R_L} \tag{2.121}$$

2.4.8.5 Class D

Because of the hard-switching nature of the Class D amplifier, the supply voltage of the amplifier is limited to the nominal supply voltage of the technology. Therefore, the maximum output power can be written as

$$P_{o,max,D} = 0.2026 \cdot \frac{V_{DD,nom}^2}{R_L} \tag{2.122}$$

2.4.8.6 Class E

A key property of the Class E amplifier is time separated drain voltage and drain current. High voltage and high current never coincide and when the

Table 2.3. Power amplifier performance overview: linear and saturated

Class	A	B	F3	Sat. A
peak efficiency	50.0%	78.5%	88.4%	81.0%
normalized output power [W] ($V_{DD} = 1V$ and $R_L = 1\Omega$)	0.5000	0.5000	0.6328	0.8106
maximum output power [W] ($v_{DS,max} = 1V$ and $R_L = 1\Omega$)	0.1250	0.1250	0.1582	0.2026
maximum output power [W] (reliable operation,$V_{DD,nom} = 1V$ and $R_L = 1\Omega$)	0.1250	0.5000	0.6328	0.2026

transistor starts to conduct current, the drain voltage is close to zero due to the external passive network. In a Class D amplifier, the drain voltage and drain current do not coincide either, but it is the transistor itself that brings the drain voltage to zero (hard switching). As such, the Class E amplifier is limited by the oxide breakthrough and junction breakthrough but not by hot carrier injection.

The breakdown drain voltage of a CMOS technology, for a zero drain current, is typically two to three times the supply voltage. Taking this into account, the maximum output power of the Class E amplifier becomes

$$P_{o,max,E} = 0.0455 \cdot \frac{\left((2\ldots3)\cdot V_{DD,nom}\right)^2}{R_L} = (0.1820\ldots0.4095)\cdot\frac{V_{DD,nom}^2}{R_L} \tag{2.123}$$

To conclude this short overview on RF power amplifiers, tables 2.3 and 2.4 summarize the relevant properties of the discussed amplifiers. Looking at the efficiency, the Class E amplifier is clearly the ideal candidate. Looking at the output power, the Class B and F are the better ones. However, Class F requires more components for the harmonic tuning, which makes the Class B a better choice for CMOS integration. On the other hand, in the Class E amplifier the parasitic drain-source capacitance can become part of the amplifier network. As such, a Class E or a mixture of Class E and Class B seems the ideal CMOS power amplifier. The design and the implementation of this amplifier in CMOS, will be discussed in the next chapter.

2.5 Efficiency and Linearity

Efficiency and linearity are two contradicting goals in the design of a power amplifier. In section 2.4, it was already clear that the shift from Class A to Class C is accompanied with a linearity decrease. Moving further to the highly efficient switching amplifiers like Class D and Class E, results in an amplifier that is entirely nonlinear for amplitude modulated signals. On the other hand

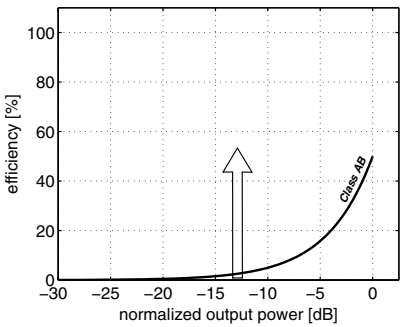
Table 2.4. Power amplifier performance overview: switching

Class	D	E
peak efficiency	100%	100%
normalized output power [W] ($V_{DD} = 1V$ and $R_L = 1\Omega$)	0.8106	0.5768
maximum output power [W] ($v_{DS,max} = 1V$ and $R_L = 1\Omega$)	0.2026	0.0455
maximum output power [W] (reliable operation,$V_{DD,nom} = 1V$ and $R_L = 1\Omega$)	0.2026	$0.1820 \ldots 0.4095$

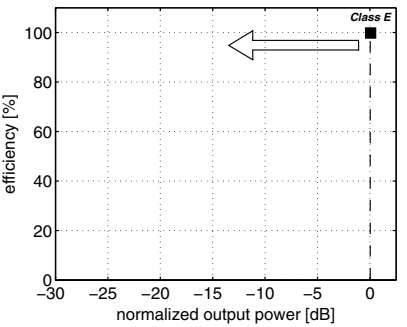
the Class E amplifier is an ideal candidate for a fully integrated CMOS power amplifier. It is therefore often stated that such an amplifier is well suited for the amplification of phase modulated or constant envelope signals. However, every mobile communication system requires that the output power of the amplifier can be regulated or controlled, which is designated as *power control*. Two forms of power control can be distinguished. First, power control that is not mandatory for proper operation of the communication system, but used to reduce the power consumption and extend the battery lifetime of the mobile terminal. This can be seen in TDMA systems, like GSM. Secondly, power control is also used to ensure that the signals of all mobile users reach the base station with equal amplitudes. This is used in CDMA systems like UMTS . Power control is a relative low-frequency process and the output power typically changes at intervals ranging from several microseconds up to milliseconds, resulting in bandwidths from one kilohertz to several hundred kilohertz. Thus all mobile communication systems require that the output power that can be reduced or increased, which can be considered as a low-frequency amplitude modulation of the output RF signal.

Apart from power control, a wide variety of systems employ both amplitude and phase modulation. Examples range from military SSB transmissions on shortwave to Wireless LAN at 5 GHz. Since the amplitude modulation is explicitly present in the modulation format, these systems inherently have a varying or non-constant envelope RF signal, apart from any additional power control. Typically, the bandwidth of the envelope signal is much higher compared to the bandwidth of the power control signal.

From the above, it can be concluded that every power amplifier must have a means to regulate its output power. In other words, a modulation scheme can produce constant envelope signals but a constant envelope power amplifier is useless for mobile communication systems.



(a) **efficiency improvement:** *a linear amplifier has a low efficiency at lower power levels; efficiency improvement techniques allow to increase the efficiency at lower power levels.*



(b) **linearization:** *a Class E amplifier can only transmit one discrete power level; linearization techniques allow to amplify amplitude modulated signals.*

Figure 2.39. Combining both efficiency and power.

For linear amplifiers, like Class A or Class B, the control of the output envelope is straightforward since these amplifiers have an amplitude linearity between input and output. Thus, reducing the amplitude of the RF input signal will also reduce the output power. This allows both power control and amplitude modulation of the output signal. The major drawback is the low efficiency of these amplifiers. For a nonlinear Class E amplifier, the amplitude linearity is completely lost, and one has to apply a linearization technique to obtain an amplitude modulated signal at the output.

Two main strategies exist two combine both efficiency and linearity in one amplifier. Starting from an amplifier with sufficient linearity, Class A or AB, one can try to improve the efficiency. The alternative is to linearize an efficient but nonlinear amplifier. Both approaches are depicted in figure 2.39. Several linearization and efficiency improvement techniques exist in literature and are briefly discussed in this section. A profound overview can be found in [Raab02], [Keni00] and [Crip99].

2.5.1 Efficiency Improvement of Linear Amplifiers

In the discussion on linear amplifiers, the Class B came out as the best candidate regarding output power and efficiency. Indeed,the maximum normalized output power under reliable operation is $0.5\,\mathrm{W}$ and the peak efficiency equals

78.5 %. If the non-idealities of the transistor are included, the amplitude linearity of the Class B amplifier is insufficient for most wireless standards. Therefore, most linear RF power amplifiers operate as Class AB and are biased rather close to Class A.

A Class AB amplifier is still capable to combine both a high output power and a decent efficiency. However, if the output signal contains amplitude modulation, the average efficiency will be dependent on the profile of the efficiency versus envelope signal curve, as discussed in section 2.3.5. The average efficiency can be obtained by evaluating following expression:

$$\eta_d = \int_0^{A_{max}} \eta_d(A) \cdot p(A) \cdot dA \tag{2.124}$$

For the amplification of amplitude modulated signals, the profile of the efficiency versus output envelope is thus more important the actual peak efficiency that can be achieved. For a Class B, the normalized relationship between efficiency and output envelope is given by

$$\eta_d(A) = \frac{\pi}{4} \cdot A \tag{2.125}$$

Similar, for a Class A amplifier, the relationship equals

$$\eta_d(A) = \frac{1}{2} \cdot A^2 \tag{2.126}$$

For a Class AB, the actual profile is dependent on the biasing and will be lying between Class A and B. The major drawback of a Class AB amplifier is thus the unfavourable profile of the efficiency versus envelope curve.

Furthermore, the linearity of the amplifier will degrade severely near the point of maximum output power and maximum efficiency, which is denoted as *compression*. The *1-dB compression point* is defined as the input power for which the output is 1-dB below the expected output level. Systems with a high *PAPR* must be biased sufficiently below the 1-dB compression point, which is denoted as *output power back-off*. Because of this compression, the amplifier must be biased further away from the maximum efficiency point, again resulting in a low efficiency.

To increase the efficiency of a linear amplifier, the supply voltage and bias point of the Class AB amplifier can be adopted according to the instantaneous value of the envelope signal. This group of efficiency improvement techniques is denoted as dynamic biasing, envelope tracking, dynamic supply or Class H [Yang99, Hani99, Sahu04]. A Class G amplifier is similar to this, but uses two or more discrete supply voltages instead of a continuous varying supply voltage. Both Class G and Class H topologies were originally developed for audio amplifiers and line drivers.

Another approach is to push the 1-dB compression point further away. In general, this is done by using an additional or auxilary power amplifier that is

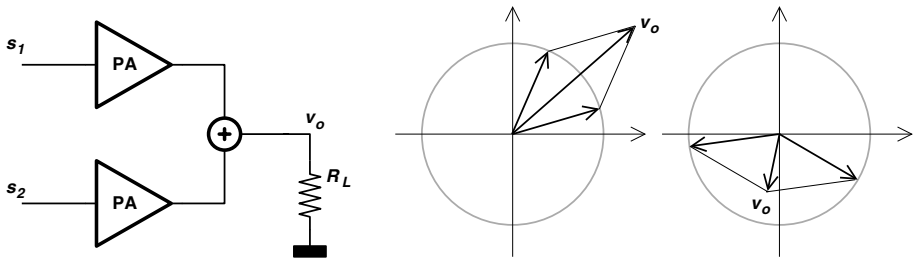


Figure 2.40. Principle of Outphasing and LINC.

activated at higher output power levels [Ding05]. This principle is similar to the *Doherty amplifier* [Dohe36], which is actually an active load pull technique [Crip99].

Finally, error correction can be applied on a linear power amplifier to increase the linearity of the amplifier which can then be operated closer to the 1-dB compression point, where the efficiency is high. Predistortion and feedback are systems that use error correction at the input of the RF amplifier, whereas feedforward applies error correction at the output of the power amplifier. However, these are all linearization techniques that only marginally increase the efficiency. A detailed discussion of these techniques can be found in [Crip99] and [Keni00].

2.5.2 Linearization of Nonlinear Amplifiers

Starting from a nonlinear but efficient amplifier that delivers a constant envelope RF signal, two major techniques exist to create an amplitude modulated signal.

2.5.2.1 Outphasing

Outphasing [Chir35] or LINC (Linear amplification with nonlinear components) [Cox74] combines the output of two nonlinear amplifiers. The two output signals can be represented as two constant envelope vectors, as shown in figure 2.40. By changing the phase difference between these two vectors, the sum of the two can have any amplitude from zero up to the sum of the two envelopes.

The drawback of this technique is that it requires an efficient power combiner at the output, which is difficult to integrate in CMOS. Furthermore, the signals at the input of this combiner are not fully differential. The phase difference between the two signals constantly changes and each signal may distort

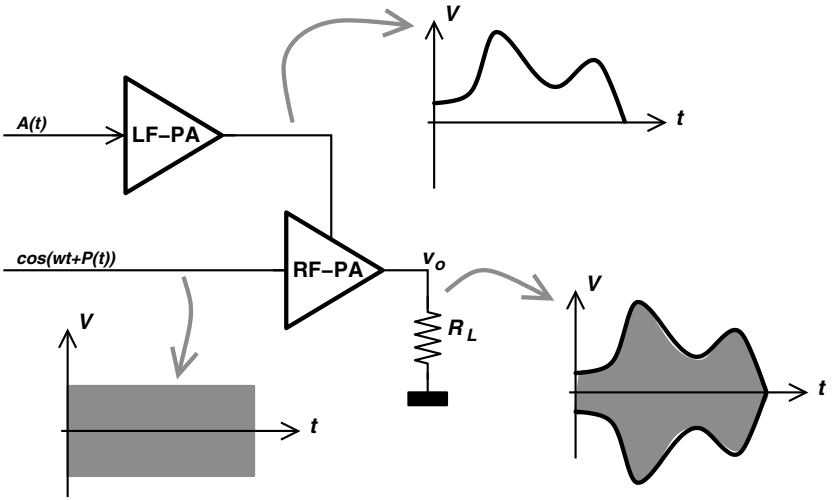


Figure 2.41. Principle of Polar Modulation.

the phase of the other. On the other hand, this technique is able to transmit wideband signals [HH04].

2.5.2.2 Polar Modulation

Polar modulation uses a nonlinear RF amplifier and a linear low frequency amplifier that delivers the supply voltage of the nonlinear RF amplifier. This technique results in an efficient RF amplifier since the efficiency of the Class E amplifier is independent of the supply voltage, even if the losses of the inductors and the loss of the switch are included. Changing or modulating the supply voltage of a nonlinear amplifier is denoted as *polar modulation* and will be discussed in detail in chapter 5.

One could argue what would be the best; increasing the efficiency of a Class AB or increasing the linearity of a Class E. Both techniques require that the supply voltage is adjusted according to the envelope signal. However, the Class AB requires that the entire RF transmit path is linear.

Furthermore, changing the supply voltage or bias point of the Class AB will degrade the linearity. As such, it may be necessary to bias the RF amplifier closer to Class A, with a lower efficiency as a result, or one has to include a linearity improvement like feedforward, feedback or predistortion. A polar modulated Class E amplifier on the other hand, can still operate as a switching amplifier. As such, the driver stages can also be nonlinear and this will always be more efficient compared to a complete linear RF architecture. Finally, in polar modulation the linearity requirements are shifted to a low frequency. This

makes it easier to achieve the needed linearity, and results in a low power consumption as well.

2.6 Conclusion

In this chapter, some important topics related to the amplification of RF signals have been reviewed and defined. First, the generation of a modulated RF signal, and some signal properties that are relevant for or have a large impact on the design of a power amplifier, have been discussed.

Next, some key properties of the power amplifier itself have been defined. Output power and efficiency, although they seem fundamental, may take many forms, especially if the amplifier needs to transmit amplitude modulated signals. Therefore, an in-depth discussion and some additional comments and remarks regarding efficiency were given.

In section 2.4, the classification of power amplifiers has been made with a strong emphasis on RF and CMOS. The reliability of the different amplifier was discussed as well. Taking into account the circuit simplicity, the beneficial use of the parasitic drain-source capacitance and the superior performance regarding hot carriers, CMOS seems to be the natural habitat of the Class E amplifier. Class E indeed does achieve a high efficiency, but the amplitude linearity is completely lost in this amplifier.

Combining efficiency and linearity was the topic of section 2.5. The efficiency improvement of a linear amplifier was compared with the linearization of a nonlinear amplifier like the Class E amplifier. Polar modulation is suggested to combine both efficiency and linearity in a fully integrated CMOS power amplifier.

To conclude, the Class E amplifier is designed in the time domain by solving a set of differential equations. If all losses are taken into account, an analytical solution becomes infeasible. The design of a Class E amplifier in CMOS and the different tradeoffs that exist will be the topic of the next chapter.

Chapter 3

ANALYSIS AND DESIGN OF THE CLASS E POWER AMPLIFIER IN CMOS

3.1 Introduction

Although the Class E circuit, presented in section 2.4.7, seems relatively simple, the design of the different elements requires some tedious mathematical calculations. As such, the design of a this amplifier was the topic for a lot of theoretical publications in the eighties. The main reason for the difficult calculations is the fact that a Class E amplifier is imposed by two time-domain equations. This inherently requires that one has to solve a set of differential equations and, more important, the correct initial conditions have to be found in order to obtain the steady-state solution of the amplifier.

This chapter presents a new technique and design methodology based on a state-space description of the Class E amplifier [Reyn03a]. This technique enables to design a Class E amplifier, including the different losses. Section 3.3 will focus on the CMOS implementation of a Class E amplifier. The different tradeoffs will be made visible and design guidelines for the different circuit parameters are given. Finally, in the last section, some layout aspects regarding the integration of RF power amplifiers in CMOS are discussed.

3.2 A Theoretical Study of the Class E Amplifier

3.2.1 The Class E Requirements

An amplifier that operates in Class E has to fulfill two time-domain constraints:

$$\text{Class E} \Leftrightarrow \begin{cases} v_{DS}(t_1) & = & 0 \\ \dfrac{dv_{DS}(t)}{dt}\Big|_{t=t_1} & = & 0 \end{cases} \tag{3.1}$$

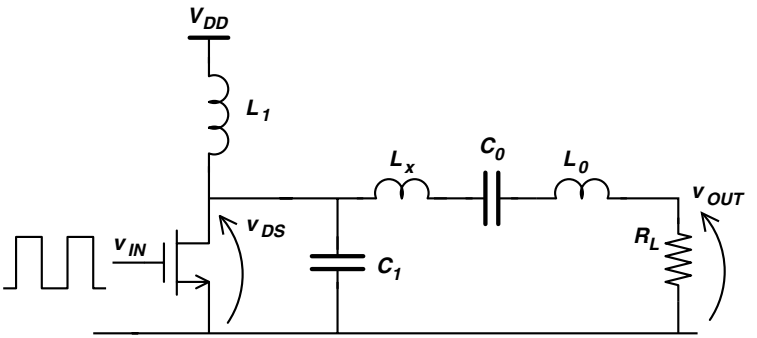
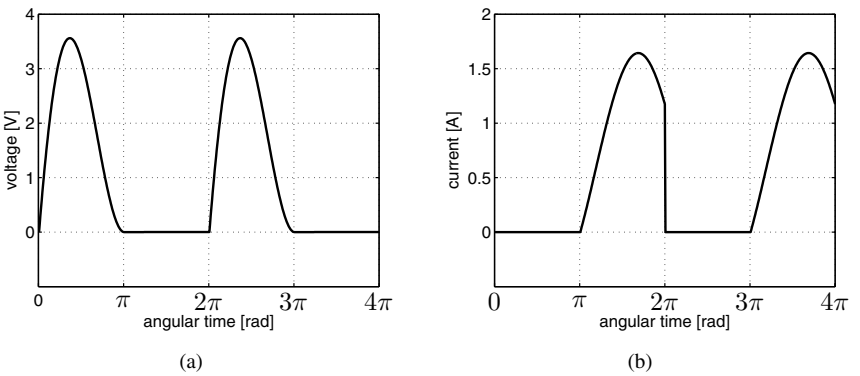


Figure 3.1. Class E power amplifier.



(a) (b)

Figure 3.2. Class E power amplifier waveforms (a) switch voltage and (b) switch current.

whereas t_1 represents the instant at which the switch closes, and $v_{DS}(t)$ represents the transistor or switch voltage.

The basic circuit that can fulfill the Class E conditions is given in figure 3.1. Since the first publication of the Class E amplifier in 1975 [Soka75], only very little circuits were developed that also work as Class E [Soka81]. Figure 3.2 depicts the current and voltage waveform of a normalized Class E amplifier ($V_{DD} = 1\,\mathrm{V}$ and $R_L = 1\,\Omega$).

The basic Class E circuit excels in simplicity, but most of all it is the presence of the capacitor C_1 that makes this amplifier attractive if a CMOS integration is aimed for. A CMOS transistor has a relatively large drain-source parasitic capacitance that, in Class E, can become part of the required capacitor C_1. As such, the parasitic capacitance is no longer a disadvantage, like in

Class B and Class F where it shorts the harmonics that are crucial to shape the drain waveform and thus to increase the efficiency.

If the transistor is driven by a square wave, or another large signal, the transistor can be replaced by a switch with an equivalent on-resistance r_{on}. Class E requires that the voltage across C_1 equals zero when the switch closes. This is the so-called *zero voltage switching* requirement. If a nonzero voltage is present across C_1 at transistor switch-on, the corresponding charge would be lost in the switch, resulting in an energy loss of

$$E = \frac{1}{2} C_1 v_{DS,t_1}^2 \tag{3.2}$$

Therefore, in order to have a lossless amplifier with a $100\,\%$ efficiency, the Class E operation requires a zero switch voltage when the switch is closing. The other requirement, a zero first derivative, is not crucial for $100\,\%$ efficiency, but it makes the amplifier less sensitive to component variations.

In figure 3.1, and under the assumption of an ideal switch, the six passive components are constrained by following general considerations. The combination of L_0 and C_0 forms a harmonic filter that is tuned to the switching frequency of the amplifier. The values of L_0 and C_0 will be determined by the switching frequency and the allowed harmonic distortion of the output signal, which is related to the quality factor of this LC filter. The value of the load resistance R_L will mainly be determined by the wanted output power. In a *traditional* Class E design, the DC-feed inductance L_1 is chosen large so it acts like a current source. Finally, since the Class E working conditions are given by two equations, two components of the circuit can be chosen in such a way that the amplifier fulfills the Class E working conditions. For the circuit of figure 3.1, these two components are the *shunt capacitor* C_1 and the *excess inductance* L_x. In this way, all components are determined, under the assumption that equation (3.1) can be solved.

If one implements the Class E amplifier in a deep submicron or nanometer CMOS technology, several obstacles will appear. As the supply voltage decreases, the current in the circuit increases for the same output power and the losses in the DC-feed L_1 can no longer be neglected. By lowering the inductance value of L_1, the power loss in the latter will also decrease and as a result, the efficiency of the amplifier will increase. Therefore, it is necessary to take the influence of a small value of L_1 into account. Furthermore, a small inductor is easier to integrate in CMOS. The switch is implemented as an nMOS transistor, and can be represented by an equivalent switch resistance or on-resistance r_{on}. This resistor will have a very large influence on the operation of the amplifier and cannot be neglected as efficiency will mainly depend on this resistor. The output filter consisting of $L_0\,C_0$ will have a finite quality factor, especially if the inductors are integrated. The non-sinusoidal output

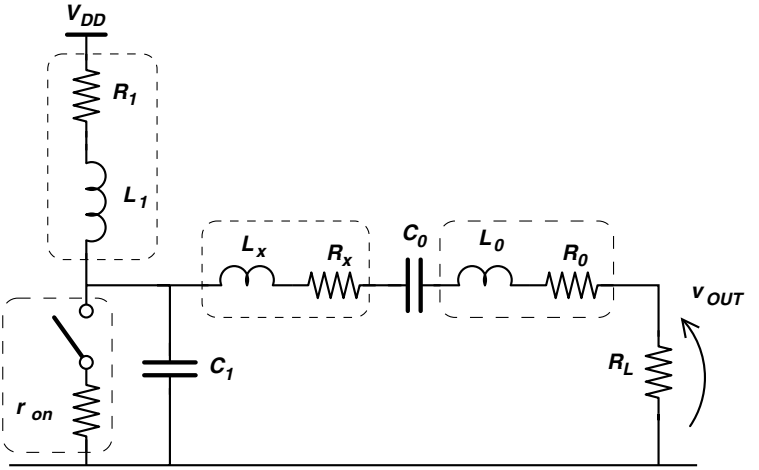


Figure 3.3. Model of the Class E power amplifier, including all losses.

current and the loss in the output filter will also influence the Class E operation. From these observations, the Class E amplifier can be represented by an equivalent model, that includes all important losses. This model is shown in figure 3.3.

3.2.2 Existing Methods to Solve the Class E Equations

Since the Class E conditions are specified in the time-domain, a time-domain expression for the switch voltage $v_{DS}(t)$ and switch current $i_{DS}(t)$ must be found. Once these expressions are derived, they can be substituted in equation 3.1. The two Class E requirements then lead to two equations, from which the values of L_x and C_1 can be solved. The most difficult task in this approach is to find the expressions $v_{DS}(t)$ and $i_{DS}(t)$.

If the quality factor of the $L_0\,C_0$ output filter is assumed to be infinite, the output current will have a sinusoidal shape. Under this assumption, and neglecting all the losses, a straightforward analytical expression for $v_{sw}(t)$ and $i_{sw}(t)$ can be found. F. H. Raab was the first to derive expressions for the Class E amplifier in 1977 [Raab77]. He also assumed an infinite value of L_1 and no losses were taken in account.

Kazimierczuk [Kazi86] and [Kazi87b] derived Class E design equations under the assumption of a non sinusoidal output current. He applied the inverse Laplace transformation to obtain the voltage and current waveforms. Still, no losses were considered and the value of L_1 is assumed to be infinite.

Zulinski [Zuli86], [Zuli87] and Li [Li94] repeated the analysis of Raab, under the assumption of a finite value of L_1. But again, no losses were considered

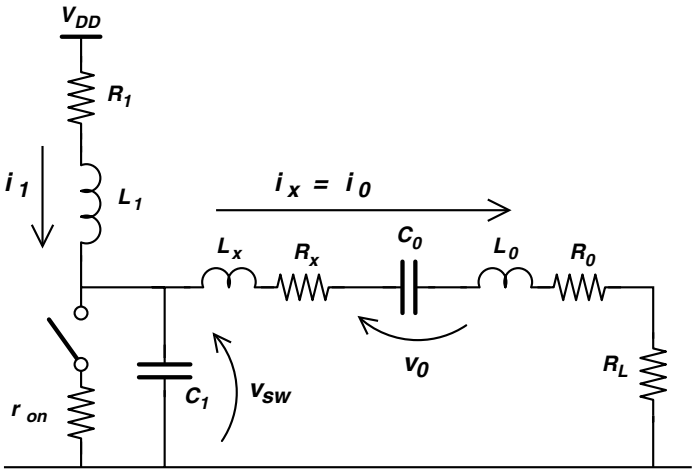


Figure 3.4. Model of the Class E power amplifier, including all losses and indicating the state variables.

and the output current is assumed to be sinusoidal. The method was expanded to non sinusoidal output current [Smit90] but still without losses.

The first attempt to analyze the influence of the switch resistance can be found in [Avra89]. Alinikula [Alin99] analyzed the Class E amplifier with a nonlinear capacitance C_1, but still under the assumption of an infinite DC-feed inductance, a sinusoidal output current and a lossless amplifier.

None of the abovementioned techniques include the losses of the inductors, although these losses have a large influence of the efficiency and design of the amplifier, especially if integrated in CMOS. The presence of the switch resistance can not be neglected either and this loss will interact with the losses in the inductors.

3.2.3 A State-Space Model of the Class E Power Amplifier

In this section, a new technique is presented to find the steady-state solution of the Class E power amplifier. All losses are taken into account, as well as the effect of the non sinusoidal output current and a finite or small DC-feed inductance L_1.

3.2.3.1 Derivation of the State-Space Equations

Figure 3.4 depicts the proposed model of the Class E amplifier. Notice that all important losses are taken into account. When the transistor is in the on-state, it is modeled as a linear resistor r_{on}. The state variables of this circuit are the so-called energy-storing elements, i.e. the voltage across a capacitor and

the current through an inductor. They are grouped together in the state matrix $\mathbf{q}(t)$:

$$\mathbf{q}(t) = \begin{bmatrix} i_1(t) \\ i_x(t) \\ v_{sw}(t) \\ v_0(t) \end{bmatrix} \tag{3.3}$$

The current through inductors L_x and L_0 are the same and therefore only one state variable is required. The drain-source voltage of the transistor is represented by $v_{sw}(t)$.

The state equations can be obtained from the first derivative of each state variable. This leads to following set of equations when the switch is closed:

$$L_1 \frac{di_1(t)}{dt} = V_{DD} - v_{sw}(t) - R_1 i_1(t) \tag{3.4}$$

$$(L_x + L_0) \frac{di_x(t)}{dt} = v_{sw} - (R_x + R_0 + R_L)\, i_x(t) - v_0(t) \tag{3.5}$$

$$C_1 \frac{dv_{sw}(t)}{dt} = i_1(t) - i_x(t) - \frac{v_{sw}(t)}{r_{on}} \tag{3.6}$$

$$C_0 \frac{dv_0(t)}{dt} = i_x(t) \tag{3.7}$$

When the switch is open, equation 3.6 changes to

$$C_1 \frac{dv_{sw}(t)}{dt} = i_1(t) - i_x(t) \tag{3.8}$$

These equations can now be put in a matrix format

$$\frac{d\mathbf{q}(t)}{dt} = \mathbf{A}\mathbf{q}(t) + \mathbf{B} \tag{3.9}$$

The matrices $\mathbf{A}$ and $\mathbf{B}$ can be formulated for the two different states of the switch. This leads to two sets of state-space equations of which the first describes the system with the switch open and the second describes the system with a closed switch. As such, one obtains the matrices $\mathbf{A_1}$, $\mathbf{B_1}$, $\mathbf{A_2}$ and $\mathbf{B_2}$. Notice that all these matrices are time independent.

3.2.3.2 Solution of the State-Space Equations

The general solution of a state-space equation is given by:

$$\mathbf{q}(t) = e^{\mathbf{A}t}\mathbf{q_0} + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}d\tau \tag{3.10}$$

in which $\mathbf{q_0}$ represents the initial states or initial conditions of the state variables. Since $\mathbf{A}$ is time independent, the integral in equation 3.10 can be written

as:

$$\int_0^t e^{-\mathbf{A}\tau} d\tau = \mathbf{A}^{-1}\left(\mathbf{I} - e^{-\mathbf{A}t}\right) \tag{3.11}$$

Substituting this expression in 3.10 leads to

$$\mathbf{q}(t) = e^{\mathbf{A}t}\mathbf{q_0} + e^{\mathbf{A}t}\left(\mathbf{A}^{-1}\left(\mathbf{I} - e^{-\mathbf{A}t}\right)\right)\mathbf{B} \tag{3.12}$$

As the matrix multiplication $e^{\mathbf{A}t} \cdot \mathbf{A}^{-1}$ is commutative, the above expression can further be simplified as

$$\mathbf{q}(t) = e^{\mathbf{A}t}\mathbf{q_0} + \mathbf{A}^{-1}\left(e^{\mathbf{A}t} - \mathbf{I}\right)\mathbf{B} \tag{3.13}$$

This derivation can be made for the two states of the switch. In state $\mathbf{1}$ the switch is open, in state $\mathbf{2}$ the switch is closed.

$$\mathbf{q_1}(t) = e^{\mathbf{A_1}t}\mathbf{q_{01}} + \mathbf{A_1^{-1}}(e^{\mathbf{A_1}t} - \mathbf{I})\mathbf{B_1} \tag{3.14}$$

$$\mathbf{q_2}(t) = e^{\mathbf{A_2}t}\mathbf{q_{02}} + \mathbf{A_2^{-1}}(e^{\mathbf{A_2}t} - \mathbf{I})\mathbf{B_2} \tag{3.15}$$

In order to obtain the time-domain steady-state solution of the state variables, it is necessary to find an expression for $\mathbf{q_{01}}$ and $\mathbf{q_{02}}$, the steady-state initial conditions.

3.2.3.3 Steady-State Initial Conditions $\mathbf{q_{01}}$ and $\mathbf{q_{02}}$.

A state variable is associated with energy and therefore its value cannot change instantaneously. At the end of state $\mathbf{1}$ (switch open), the state variables should have the same as at the beginning of state $\mathbf{2}$ (switch closed), and vice versa. This can be written as

$$\begin{cases} \mathbf{q_{02}} \equiv \mathbf{q_1}(t_1) = e^{\mathbf{A_1}t_1}\mathbf{q_{01}} + \mathbf{A_1^{-1}}(e^{\mathbf{A_1}t_1} - \mathbf{I})\mathbf{B_1} \\ \mathbf{q_{01}} \equiv \mathbf{q_2}(t_2) = e^{\mathbf{A_2}t_2}\mathbf{q_{02}} + \mathbf{A_2^{-1}}(e^{\mathbf{A_2}t_2} - \mathbf{I})\mathbf{B_2} \end{cases} \tag{3.16}$$

The instant at which the switch closes is denoted with t_1 and is equal to δT, with δ the duty cycle and T the period of the switching frequency. Note that the solution of state $\mathbf{2}$ is also starting from $t = 0$ and not from δT. Therefore, the switch opens at time $t_2 = (1 - \delta)T$. In this set of two equations, only two variables are undefined ($\mathbf{q_{01}}$, $\mathbf{q_{02}}$) and thus an expression for the steady-state initial conditions can be found.

A final difficulty lies in the evaluation of $e^{\mathbf{A}t}$. However, this expression only has to be evaluated at two instants, i.e. t_1 and t_2. Evaluating such a matrix can easily be done numerically (e.g. with MATLAB) or by determining the eigenvalues of $\mathbf{A}$.

Once the steady-state initial conditions $\mathbf{q_{01}}$ and $\mathbf{q_{02}}$ are known, they can be substituted in expressions (3.14) and (3.15). This enables to analyze the

time behavior of each state variable. From this solution, the other variables such as the output voltage and the switch current, can easily be obtained from Kirchoff's current and voltage laws. This approach not only allows to evaluate Class E power amplifiers, but also deviations from the Class E conditions.

3.2.3.4 Calculating Power and Efficiency

The output power of the circuit of figure 3.4 is given by

$$\begin{aligned} P_{out} &= R_L \cdot \frac{1}{T} \cdot \int_0^T i_x^2(t)dt \\ &= R_L \cdot \frac{1}{T} \cdot \left(\int_0^{t_1} i_x^2(t)dt + \int_0^{t_2} i_x^2(t)dt \right) \end{aligned} \tag{3.17}$$

By defining the following matrices

$$\mathbf{W_1} = \int_0^{t_1} \mathbf{q_1 q_1^T} dt \tag{3.18}$$

$$\mathbf{W_2} = \int_0^{t_2} \mathbf{q_2 q_2^T} dt \tag{3.19}$$

the integrals in equation 3.17 can also be obtained as

$$\int_0^{t_1} i_x^2(t)dt = \mathbf{W_1}[2,2] \tag{3.20}$$

$$\int_0^{t_2} i_x^2(t)dt = \mathbf{W_2}[2,2] \tag{3.21}$$

and therefore, the output power can now be written in a matrix form as

$$P_o = \frac{1}{T} \mathbf{C}_{\mathrm{R_L}} \left(\mathbf{W_1} + \mathbf{W_2} \right) \mathbf{C}_{\mathrm{R_L}}^{\mathbf{T}} \tag{3.22}$$

in which $\mathbf{C}_{\mathrm{R_L}}$ equals

$$\mathbf{C}_{\mathrm{R_L}} = \left[\begin{array}{cccc} 0 & \sqrt{R_L} & 0 & 0 \end{array} \right]$$

To evaluate the loss of the inductors L_1, L_x and L_0, one can apply a similar method. As such, the loss of L_1, denoted as P_{R_1}, can be written as

$$P_{R_1} = \frac{1}{T} \mathbf{C}_{\mathrm{R_1}} \left(\mathbf{W_1} + \mathbf{W_2} \right) \mathbf{C}_{\mathrm{R_1}}^{\mathbf{T}} \tag{3.23}$$

and

$$\mathbf{C}_{\mathrm{R_1}} = \left[\begin{array}{cccc} \sqrt{R_1} & 0 & 0 & 0 \end{array} \right]$$

The DC power consumption of the amplifier is equal to

$$P_{DC} = \frac{1}{T} \int_0^T V_{DD}\, i_1(t)dt = \frac{1}{T} V_{DD} \int_0^T i_1(t)dt \tag{3.24}$$

which, using the same technique as above, can be written in a matrix form as

$$P_{DC} = \frac{1}{T}\,\mathbf{C}_{\mathrm{DC}}\left(\int_0^{t_1}\mathbf{q_1}dt + \int_0^{t_2}\mathbf{q_2}dt\right)\mathbf{C}_{\mathrm{DC}}^{\mathbf{T}} \tag{3.25}$$

and $\mathbf{C}_{\mathrm{DC}}$ equals

$$\mathbf{C}_{\mathrm{DC}} = \begin{bmatrix} \sqrt{V_{DD}} & 0 & 0 & 0 \end{bmatrix}$$

Finally, the power dissipated in the switch resistance r_{on} only has to be evaluated during the second phase.

$$\begin{aligned} P_{sw} &= \frac{1}{T}\int_0^{t_1}\frac{v_{sw}^2(t)}{r_{on}}dt & (3.26)\\ &= \frac{1}{T}\,\mathbf{C}_{\mathrm{sw}}\mathbf{W_2}\mathbf{C}_{\mathrm{sw}}^{\mathbf{T}} & (3.27)\end{aligned}$$

with

$$\mathbf{C}_{\mathrm{sw}} = \begin{bmatrix} 0 & 0 & 1/\sqrt{r_{on}} & 0 \end{bmatrix}$$

All the above equations need a value or expression for $\int_0^{t_i}\mathbf{q}dt$ and $\int_0^{t_i}\mathbf{q}\mathbf{q}^{\mathbf{T}}dt$. Integrating the basic state-space equation and considering the fact that $\mathbf{A}$ and $\mathbf{B}$ are time independent leads to

$$\int_0^{t_i}\mathbf{q}\,dt = \mathbf{A}^{-1}\left(\mathbf{q}(t_i) - \mathbf{q}(0) - \mathbf{B}\,t_i\right) \tag{3.28}$$

and

$$\int_0^{t_i}\mathbf{q}^{\mathbf{T}}dt = \left(\mathbf{q}^{\mathbf{T}}(t_i) - \mathbf{q}^{\mathbf{T}}(0) - \mathbf{B}^{\mathbf{T}}\,t_i\right)\left(\mathbf{A}^{\mathbf{T}}\right)^{-1} \tag{3.29}$$

The derivation of $\int_0^{t_i}\mathbf{q}\mathbf{q}^{\mathbf{T}}dt$ is somewhat more cumbersome. It starts by taking following derivative:

$$\begin{aligned} \frac{d\mathbf{q}\mathbf{q}^{\mathbf{T}}}{dt} &= \frac{d\mathbf{q}}{dt}\mathbf{q}^{\mathbf{T}} + \mathbf{q}\frac{d\mathbf{q}^{\mathbf{T}}}{dt}\\ &= \mathbf{A}\mathbf{q}\mathbf{q}^{\mathbf{T}} + \mathbf{B}\mathbf{q}^{\mathbf{T}} + \mathbf{q}\mathbf{q}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}} + \mathbf{q}\mathbf{B}^{\mathbf{T}} & (3.30)\end{aligned}$$

Integrating this expression leads to

$$\begin{aligned} \mathbf{q}\mathbf{q}^{\mathbf{T}}\Big|_0^{t_i} &= \mathbf{q}\mathbf{q}^{\mathbf{T}}(t_i) - \mathbf{q}\mathbf{q}^{\mathbf{T}}(0)\\ &= \mathbf{A}\int_0^{t_i}\mathbf{q}\mathbf{q}^{\mathbf{T}}dt + \mathbf{B}\int_0^{t_i}\mathbf{q}^{\mathbf{T}}dt + \int_0^{t_i}\mathbf{q}\mathbf{q}^{\mathbf{T}}dt\,\mathbf{A}^{\mathbf{T}} + \int_0^{t_i}\mathbf{q}\,dt\,\mathbf{B}^{\mathbf{T}} & (3.31)\end{aligned}$$

By using expressions (3.28) and (3.29), the above equation can be simplified to

$$\mathbf{A}\mathbf{W} + \mathbf{W}\mathbf{A}^{\mathbf{T}} + \mathbf{G} = 0 \tag{3.32}$$

in which $\mathbf{G}$ equals

$$\begin{aligned}\mathbf{G} &= \mathbf{A}^{-1}\left(\mathbf{q}(t_i) - \mathbf{q}(0) - \mathbf{B}t_i\right)\mathbf{B}^{\mathbf{T}} + \\ &\quad \mathbf{B}\left(\mathbf{q}^{\mathbf{T}}(t_i) - \mathbf{q}^{\mathbf{T}}(0) - \mathbf{B}^{\mathbf{T}}t_i\right)\left(\mathbf{A}^{\mathbf{T}}\right)^{-1} - \\ &\quad \left(\mathbf{q}\mathbf{q}^{\mathbf{T}}(t_i) - \mathbf{q}\mathbf{q}^{\mathbf{T}}(0)\right)\end{aligned}$$

Equation (3.32) is the Lyapunov equation and can be solved for the unknown variable $\mathbf{W}$ since the matrices $\mathbf{A}$ and $\mathbf{G}$ are completely determined. The Lyapunov equation can be solved e.g. in MATLAB by using standard matrix techniques. Once $\mathbf{W}$ is obtained, it can be substituted in the different power expressions and from this one can evaluate the drain efficiency of the power amplifier.

3.2.3.5 Designing for Class E conditions

In order to have a Class E power amplifier, both the switching voltage and its first derivative should be zero at the beginning of the conducting phase. Since the value of $v_{sw}(t_1)$ is incorporated in $\mathbf{q_{02}}$, this requires no further calculation. The derivative of $\mathbf{q_{02}}$ can easily be found by applying equation 3.9.

The values of $v_{sw}(t_1)$ and $dv_{sw}(t)/dt|_{t_1}$ can be minimized by changing the values of L_x and C_1. This minimization process can be implemented in various ways. A standard Least Square Error technique was used to combine both numerical stability and computational speed. The method will vary the values of L_x and C_1 until $v_{sw}(t_1)$ and $dv_{sw}(t)/dt|_{t_1}$ are both *close enough* to zero. A starting point for this optimization process is provided by the simple formulas of [Raab75].

3.2.4 Limitations of the State-Space Approach

The presented method allows to design and evaluate a Class E amplifier but it also allows to analyze the behavior of a switching amplifier that deviates from the Class E conditions [Mert01]. The power loss of the inductors and the dissipation in the switch are included. However, it does not support a time-dependent behavior of the switch resistor or a nonlinear behavior of the capacitors.

Although other design and optimization techniques have been developed lately, based on a Spice circuit simulator and accurate device models for both the transistor and the inductors [Ramo04] [Gupt01], the presented state-space technique is more related to the theoretical Class E behavior. Furthermore, if all parasitic effects are taken into account and a strong nonlinear transistor model is used, the dependency of output power and efficiency on a certain parameter is not always clear.

Of course, an optimizer that uses accurate device models will always be an advantage to the circuit designer, but one should be aware that such tools, in a

first place, should help the designer in gaining insight in the tradeoffs. As such, the state-space approach is still useful to visualize the influence of component variations and to predict the impacts of e.g. lower operation voltage. It has, of course, its limitations regarding circuit optimization, since the used device models are kept simple.

3.3 Design of the Class E Amplifier in CMOS

The presented state-space approach can be applied to investigate the influence of using a CMOS technology with regard to the design of the Class E amplifiers. Four effects will be dominant and they will have a large impact on the design tradeoffs

1 low supply voltage

2 power loss in the inductors

3 on-resistance of the transistor

4 parasitic drain and gate capacitance of the transistor

In fact, it is exactly because of these reasons why most amplifiers are still designed in a GaAs or a SiGe technology, that have low parasitic capacitances and are capable of handling large voltage swings. It also explains why most amplifiers use high quality off-chip inductors or external striplines.

3.3.1 Design of the Load Resistor

In RF and microwave design, it is habitual to *design towards* $50\,\Omega$, i.e each individual block has an input and output impedance of $50\,\Omega$. Therefore, most antennas and antenna filters are designed to have an input or output impedance of $50\,\Omega$. On the other hand, the supply voltage of CMOS decreases with each new technology generation, and the supply voltage of a state of the art 45 nanometer CMOS technology is below $1\,\mathrm{V}$. One should realize that a sinewave with an amplitude of $1\,\mathrm{V}$ across a load resistor of $50\,\Omega$ means *only* $10\,\mathrm{mW}$ of output power.

A higher output power can be obtained by inserting an *impedance transformation network* between the $50\,\Omega$ antenna connection and the output of the amplifier. The transformed load resistance, *seen* by the power amplifier is much lower and thus the actual output power, dissipated in the $50\,\Omega$ load, will be higher. The design of such a network will be the topic of chapter 4. For now, it will be assumed that such a network can be designed and for clarity, the losses of this network are not included hereafter.

Going back to the Class E circuit of figure 3.1, the value of the shunt capacitor (C_1) and the excess inductance (L_x) will be fixed by the two Class E

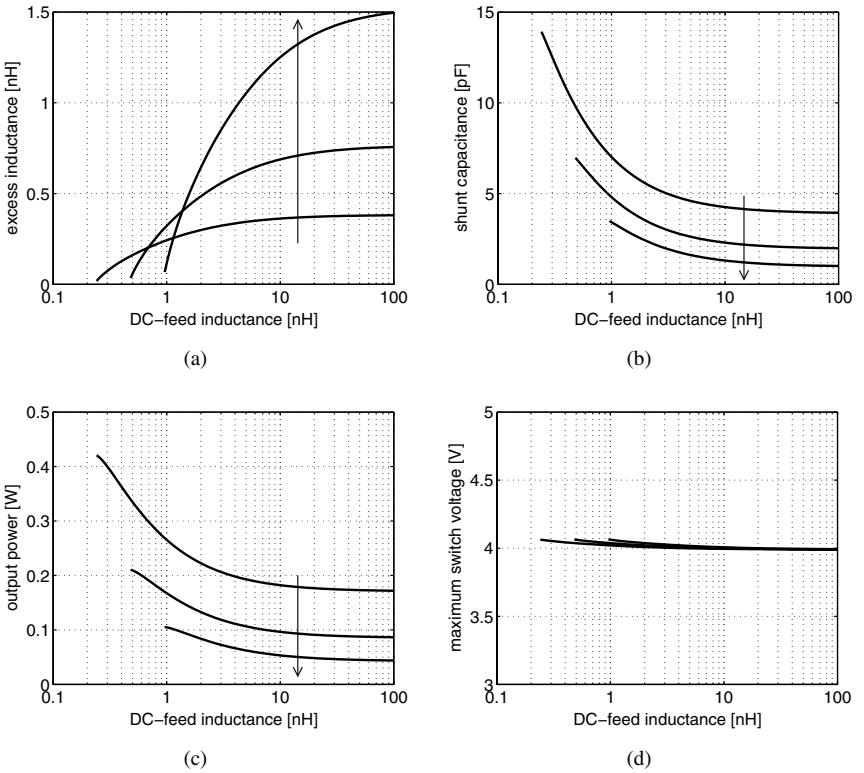


Figure 3.5. Lossless Class E design versus DC-feed inductance L_1 for $R_L = 4, 8$ and $16\,\Omega$. The arrow indicates the direction of higher R_L.

requirements. Thus, only the DC-feed inductance L_1 and the on-resistance of the switch r_{on} have to be chosen.

In what follows, the state-space approach is employed to clarify how the circuit designer should handle these two parameters and what tradeoffs are present. A *generic* technology, based on the values of a $0.18\,\mu\mathrm{m}$ technology, will be used as a benchmark.

3.3.2 Design of the DC-feed Inductance

3.3.2.1 Lossless design example

As a first experiment, the Class E amplifier is designed with the state-space approach for various values of the DC-feed inductance L_1 and the amplifier is assumed to be lossless. For each value of L_1, the optimizer calculates the required values of the shunt capacitance C_1 and the excess inductance L_x. The zero-current breakdown voltage of the $0.18\,\mu\mathrm{m}$ CMOS technology is $4\,\mathrm{V}$ and

Table 3.1. Comparison between infinite L_1 and minimum value of L_1 for a lossless Class E amplifier.

	shunt capacitance C_1 (pF)		output power P_o (mW)	
R_L	$L_1 = \infty$	$L_1 = min.$	$L_1 = \infty$	$L_1 = min.$
$4\,\Omega$	3.7	14.2	175	409
$8\,\Omega$	1.8	7.1	87	204
$16\,\Omega$	0.9	3.6	44	102

therefore, the supply voltage is $1.12\,\mathrm{V}$. The duty cycle is $50\,\%$ which results in the highest output power [Raab77] and the frequency is $2\,\mathrm{GHz}$. The simulation is done for three different values of R_L: $4, 8$ and 16Ω.

Figure 3.5 shows the results for three different values of R_L. If the DC-feed is large, the value for C_1 and L_x become independent of L_1 and equal to the theoretical values derived by Raab [Raab77]. When the value of the DC-feed is reduced, the excess inductance also decreases and becomes zero at a certain point. As its value decreases, so will the voltage drop across L_x. Since the excess inductance is in series with the load resistor, the output voltage will increase and so will the output power. The shunt capacitance C_1 will also increase for a smaller DC-feed inductance. The maximum switch voltage remains more or less the same.

Table 3.1 summarizes the effect of a small DC-feed in a Class E amplifier and compares it with the values of an infinite DC-feed. The larger value of C_1 is beneficial for CMOS integration, since the parasitic drain capacitance and the interconnect capacitance can become part of C_1. The larger output power is also an advantage in low-voltage technologies. As such, it is always the best to design the Class E in such a way that the excess inductance becomes zero.

3.3.2.2 Including inductor and switch losses

The benefit of a smaller DC-feed remains valid if losses are taken into account. To simulate these losses, it is assumed that inductors L_x and L_1 have a series resistance whose value is proportional to the inductance value. A loss of $0.75\,\Omega/\mathrm{nH}$ is chosen as this value is close to an actual implementation of an inductor in CMOS at the switching frequency of $2\,\mathrm{GHz}$. For the switch resistance, a value of $0.25\,\Omega$ is chosen, which corresponds to an nMOS gate-width of $2200\,\mu\mathrm{m}$ and a gate-length of $0.18\,\mu\mathrm{m}$. It will be shown later that this is a realistic choice. The maximum switch voltage will also be dependent on the losses and thus also on the value of the DC-feed inductor. Therefore, the supply voltage is adjusted as to keep the maximum switch voltage equal to $4\,\mathrm{V}$, which is the breakdown voltage of the nMOS transistor.

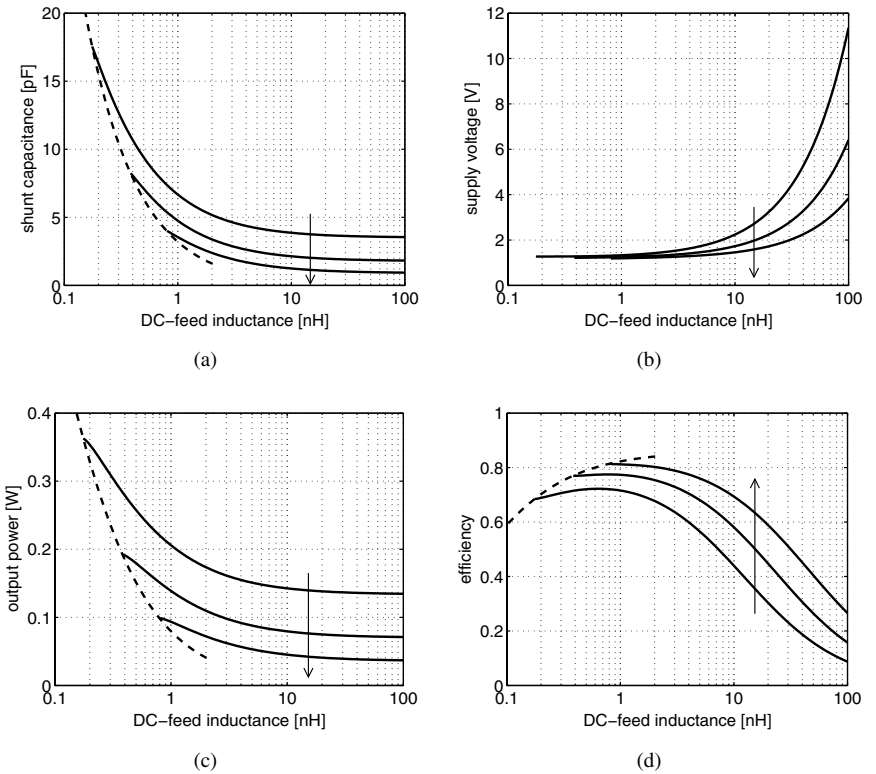


Figure 3.6. Class E design versus DC-feed inductance L_1, including an inductor loss of $0.75\,\Omega/\mathrm{nH}$ and switch resistance of $0.25\,\Omega$. The load resistance equals $R_L = 4, 8$ and $16\,\Omega$. The arrow indicates the direction of higher R_L.

The results of the Class E amplifier with all these losses included are shown in figure 3.6. The efficiency typically ranges from $60\,\%$ to $80\,\%$, but again, the output filter is assumed to be lossless and the loss of the impedance transformation network is not included. The dashed line in figure 3.6 indicates a Class E design with a zero excess inductance.

Reducing the DC-feed inductance not only results in more output power, but also in a higher efficiency for a given output power. To further clarify this relationship, figure 3.7 depicts the efficiency versus output power.

An important consequence of a reduced DC-feed, and to some extend it can be regarded as a drawback, is the increased current swing in the DC-feed, and thus an increased AC-current that has to be delivered by the power supply. To demonstrate this, figure 3.8(a) depicts the current through the DC-feed for three values of R_L. A negative DC-feed current means that current is flowing back to

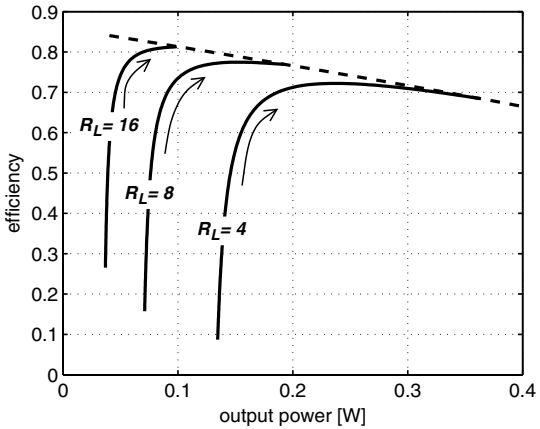


Figure 3.7. Efficiency versus output power for a Class E power amplifier with inductor and switch loss. The arrows indicate the direction to a smaller DC-feed inductance.

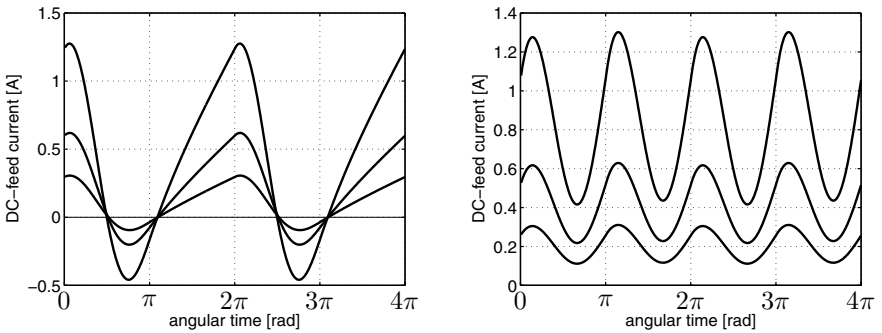


Figure 3.8. Current through the DC-feed inductance for a load resistance of $16\,\Omega$, $8\,\Omega$ and $4\,\Omega$. (a) Single-ended current and (b) differential current.

the DC power supply. Clearly, the designer has to allow these currents to flow, either through on-chip or off-chip decoupling capacitance. The current swing can be reduced by using a differential Class E amplifier and the corresponding supply current waveforms are shown in figure 3.8(b). More on this topic can be found in section 3.4.2

3.3.3 Design of the nMOS switch

In CMOS, the on-resistance of an nMOS transistor in the linear region is given by

$$r_{on} = \frac{1}{\mu_n C_{ox} W_g / L_g \left(V_{GS} - V_{th,n} - V_{DS}/2\right)} \sim \frac{1}{W_g} \tag{3.33}$$

In the Class E amplifier, the transistor should act as a switch and the gate-source voltage is a large switching waveform that goes up to the supply voltage. Therefore, using a minimal gate length and large width are the only means to reduce the on-resistance of the nMOS switch. The lower the on-resistance, the higher the efficiency, until the loss in the switch can be neglected compared to the loss in the inductors. From that point on, the efficiency is dominated by the loss of the inductors, and making the nMOS transistor wider has then little or no influence.

However, the width of the transistor is also restrained by the parasitic drain and gate capacitance and the gate length is limited by the maximum drain current and drain voltage. When a large drain current is accompanied by a large drain voltage, highly energetic or so called *hot carriers* are generated near the drain. These carriers, electrons in the case of an nMOS transistor, can be injected in the gate oxide, resulting in a shift of transistor parameters and on a longer term, they might cause breakdown of the gate oxide. This problem is alleviated in the Class E amplifier since current and voltage are separated in time. As such, the gate length is, in most cases, made equal to the minimal gate length of the technology.

As such, the transistor width is the only remaining parameter of the transistor but needs to be carefully selected. The different tradeoffs concerning the sizing of the nMOS width are discussed in the next paragraphs [Reyn02].

3.3.3.1 Influence of the gate capacitance on the nMOS transistor sizing

The gate capacitance needs to be charged to V_{DD} and discharged to zero. If this is done with a digital buffer, as indicated in figure 3.9, this charging and discharging roughly requires a DC power of

$$P_{DRV,1} = f\, C_g\, V_{DD}^2 \tag{3.34}$$

The gate capacitance C_g will consist of the gate-source capacitance and the gate-drain capacitance, and the latter will be enlarged by the Miller effect. The transistor acts as a switch and thus the gate and drain capacitances will change drastically in time. Nevertheless, one can still define an average or equivalent gate capacitance, and this capacitance will be proportional to the gate-width of the transistor. Therefore,

$$P_{DRV,1} \sim C_g \sim W_g \tag{3.35}$$

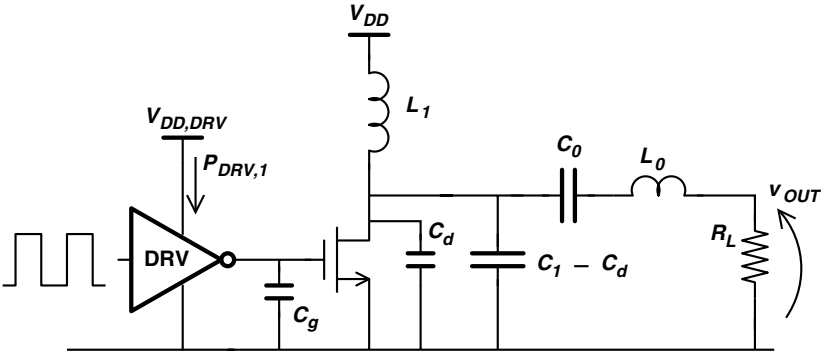


Figure 3.9. Class E power amplifier with digital inverter as driver.

From previous equations, it can be concluded that a large transistor width is beneficial for the efficiency of the amplifier, i.e. the drain efficiency of the last stage only, but not for the power consumption of the driver stage. If the transistor width is made too large, the drain efficiency will not improve any further, since the efficiency will be dominated by the loss of the inductors. On the other hand, the driver stages will dissipate more power and the overall efficiency, in this case defined as

$$\eta_{oa} = \frac{P_o}{P_{DC} + P_{DRV,1}} \tag{3.36}$$

will show an optimum in terms of transistor gate width.

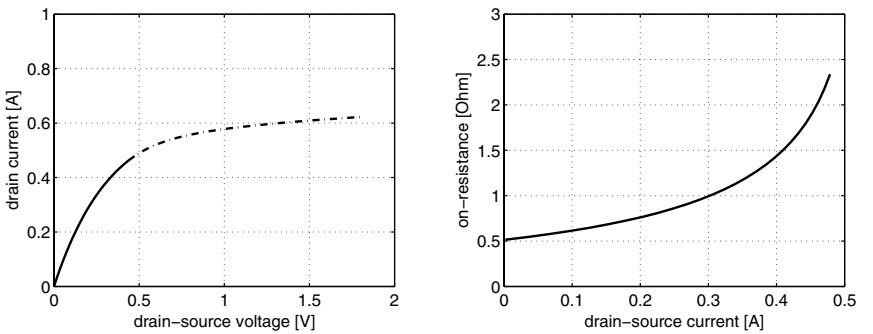
3.3.3.2 Influence of the drain capacitance on the nMOS transistor sizing

The parasitic drain capacitance C_d can become part of the required Class E shunt capacitance C_1, as shown in figure 3.9. This is one of the major benefits of the Class E amplifier. Like the gate capacitance, the parasitic drain capacitance will change during the switching cycle, but still an average that is proportional to the width of the transistor can be thought of, and thus $C_d \sim W$.

The width of the transistor can be increased until the parasitic drain capacitance exceeds the value of the required shunt capacitance. At that point, the shunt capacitance virtually disappears as it is entirely taken up by the parasitic transistor capacitance.

3.3.3.3 Exploration of the design space

To demonstrate the dependencies between the on-resistance, the drain capacitance and the gate capacitance, the values from a $0.18\,\mu\mathrm{m}$ CMOS technology will now be used.



(a) (b)

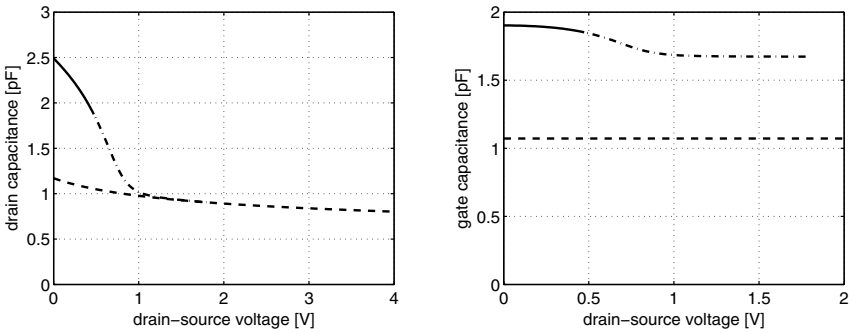
Figure 3.10. (a) Drain current and (b) on-resistance of a $0.18\,\mu\mathrm{m}$ nMOS transistor versus drain voltage. $W_g = 1000\,\mu\mathrm{m}$, $L_g = 0.18\,\mu\mathrm{m}$ and gate tied to $V_{DD} = 1.8\,\mathrm{V}$. The solid line indicates the linear region, dashed-dotted line indicates the saturation region.

Figure 3.10(b) shows the resistance of a $0.18\,\mu\mathrm{m}$ nMOS transistor with a minimal gate length and a gate width of $1000\,\mu\mathrm{m}$. The gate is tied to the supply voltage of $1.8\,\mathrm{V}$ to simulate a switch. For low values of v_{DS}, the nMOS will be in the linear region with a low on-resistance (solid line in figure 3.10(a)). If the drain voltage increases, the drain current will flatten and the on-resistance will thus increase. If the on-resistance at zero drain voltage is taken, one could say that

$$r_{on,0.18\mu m} = 1000\,\mu\mathrm{m} \cdot 0.55\Omega \cdot \frac{1}{W_g} \tag{3.37}$$

It should be clear that this is a rather optimistic view since the on-resistance increases for higher current levels.

The total drain capacitance of the $1000\,\mu\mathrm{m}$ nMOS transistor is shown in figure 3.11(a) If the gate voltage is made equal to V_{DD}, in this case $1.8\,\mathrm{V}$, the drain capacitance is large, especially in the linear region. However, it is the drain capacitance in the off-state that is important. This capacitance is given by the dashed line in figure 3.11(a). The parasitic drain capacitance partly consists of the drain-bulk junction capacitance, which explains the voltage dependency and the lower capacitance at a higher drain voltage. For the analysis, the average of the drain capacitance in the off-state, between $0\,\mathrm{V}$ and $4\,\mathrm{V}$, which is the breakdown voltage of the technology, is taken. This capacitance has to be increased by the parasitic interconnection capacitance, which is typically about



(a) (b)

Figure 3.11. (a) Total drain capacitance and (b) gate capacitance of a $0.18\,\mu\mathrm{m}$ nMOS transistor versus drain voltage. $W_g = 1000\,\mu\mathrm{m}$ and $L_g = 0.18\,\mu\mathrm{m}$. The solid line indicates the linear region, the dashed-dotted line indicates the saturation region and the dashed line is the transistor in the off-state.

$1\ \mathrm{pF}/1000\ \mu\mathrm{m}$. All this results in

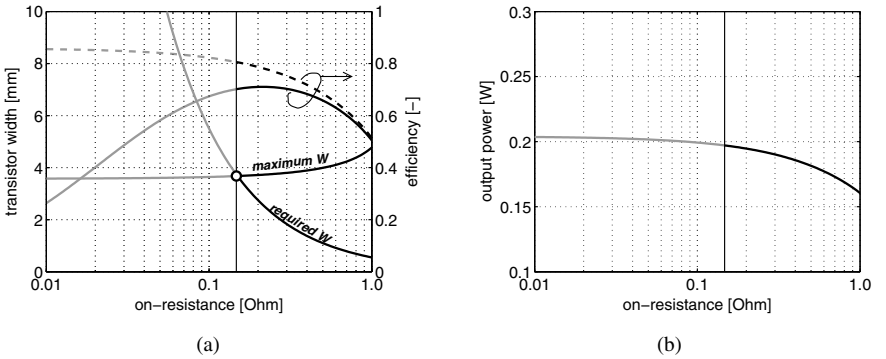
$$C_{d,0.18\mu m} = \frac{2.2\ pF}{1000\ \mu m} \cdot W_g \tag{3.38}$$

Finally, the gate capacitance can be obtained from figure 3.11(b). For the analysis here, a worst case value of $1.9\ \mathrm{pF}/1000\ \mu\mathrm{m}$ is taken. Thus,

$$C_{g,0.18\mu m} = \frac{1.9\ \mathrm{pF}}{1000\ \mu\mathrm{m}} \cdot W_g \tag{3.39}$$

The efficiency of the Class E amplifier can now be evaluated for different values of r_{on}. The on-resistance is first translated to a transistor gate width, which can be seen as the required transistor width. The required value of the shunt capacitance C_1 for Class E operation is not fixed, but will be dependent on the value of on-resistance. The value of the shunt capacitance can also be translated to a transistor width, if one assumes that this capacitance is entirely taken up by the parasitic drain capacitance of the nMOS transistor. Hence, the corresponding transistor width is in fact a maximum transistor width.

Figure 3.12 shows the results of such a calculation. A load resistor of $8\,\Omega$ is chosen and the amplifier is designed for zero excess inductance since it was demonstrated in the previous section that this results in the optimal point regarding power and efficiency. The switching frequency is $2\,\mathrm{GHz}$. For each value of r_{on}, the supply voltage is adjusted to ensure that the maximum drain voltage remains below $4\,\mathrm{V}$.



(a) (b)

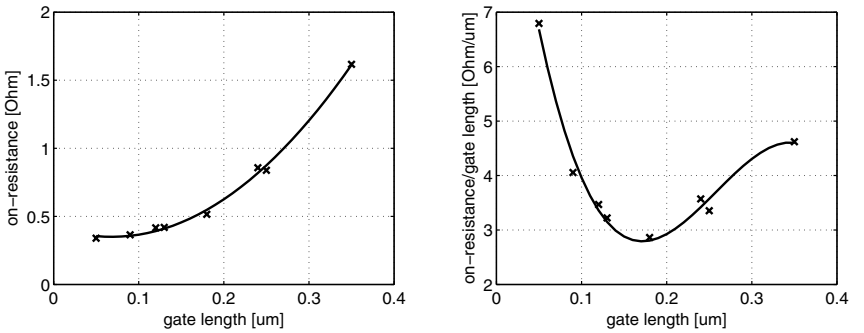
Figure 3.12. Class E design versus on-resistance with a load resistance of $8\,\Omega$. (a) Required and maximum transistor width, drain efficiency (dashed line) and overall efficiency. (b) Output power.

A smaller on-resistance clearly increases the drain efficiency of the power amplifier itself (dashed line in figure 3.12(a)), but the overall efficiency will have an optimum for a specific transistor width. In this case, that optimum is reached at an on-resistance of $0.22\,\Omega$, corresponding to a transistor gate width of about $2500\,\mu\mathrm{m}$ and an overall efficiency of $72\,\%$. The maximum transistor width is $3800\,\mu\mathrm{m}$, since at that point the parasitic capacitance of the transistor is as large as the required shunt capacitance. In other words, the region at the right of the intersection (indicated by the circle) is the design space for the width of the nMOS transistor.

It is, of course, important that the optimum overall efficiency point can be reached. If a technology with more parasitic drain capacitance would be used, the curve of *maximum W* would lower, and the intersection between the two curves would shift to the right. As such, it might occur that the optimal point can no longer be reached.

3.3.4 Technology Scaling

Technology scaling in CMOS means that the gate length of the nMOS transistor becomes smaller. A first consequence is that the on-resistance of the switch reduces, which is of course an advantage. Scaling to the next technology node means that the gate length is reduced by $1/\sqrt{2}$. In other words, the on-resistance is reduced by $\sqrt{2}$. This is shown in figure 3.13(a), in which the on-resistance of several submicron and nanometer CMOS technologies is compared In fact, between $0.35\,\mu\mathrm{m}$ and $0.18\,\mu\mathrm{m}$, the on-resistance decreases more



(a) (b)

Figure 3.13. On-resistance for several submicron and nanometer technologies for a $1000\,\mu\mathrm{m}$ wide transistor.

than the decrease in gate length, as can be seen by the ratio r_{on}/L_g, plotted in figure 3.13(b).

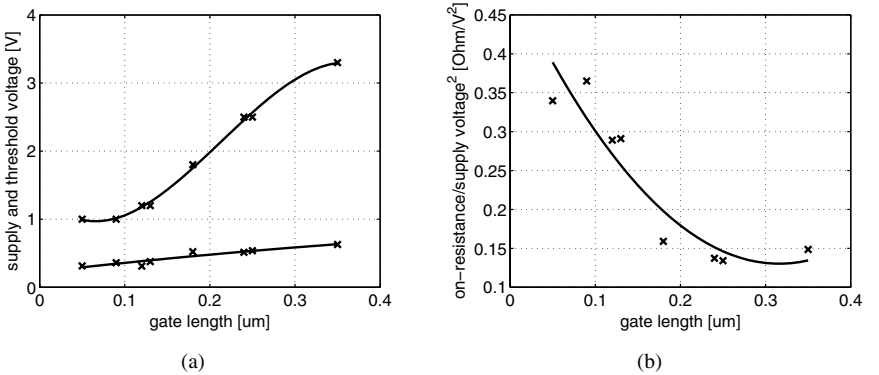
Another consequence of technology scaling is that the supply voltage is reduced as well, and scaled by the same factor $\sqrt{2}$. Figure 3.14(a) depicts the supply voltage and threshold voltage versus gate length. The supply voltage scales by $\sqrt{2}$ but the threshold voltage only shows a moderate decrease. The decrease in supply voltage requires that the load resistance is divided by 2 in order to achieve the same output power.

Some first order calculations [Raab78] [Yoo01] reveal that the drain efficiency can be written as

$$\eta \approx \frac{1}{1 + 1.4 \cdot \frac{r_{on}}{R_L}} \tag{3.40}$$

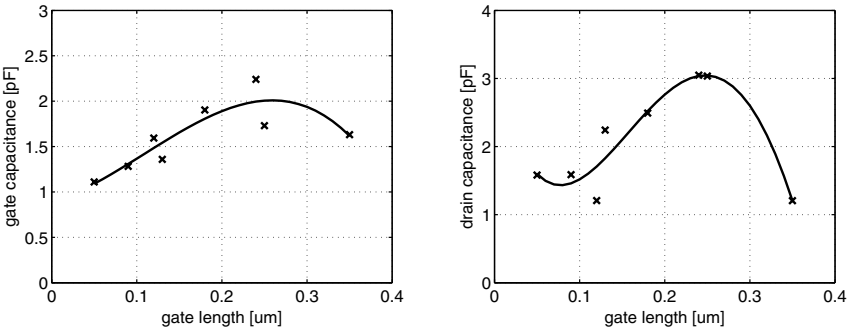
Technology scaling will reduce the on-resistance by $\sqrt{2}$ but for the same output power, the load resistance is reduced by 2 and thus the ratio r_{on}/R_L increases by $\sqrt{2}$. The ratio r_{on}/R_L can also be written as r_{on}/V_{DD}^2 if one assumes the same output power. This ratio is plotted in figure 3.14(b).

To accommodate for the increase of r_{on}/R_L, the width of the nMOS transistor has to be increased, approximately by a factor of $\sqrt{2}$. This will in turn increase the input capacitance of the transistor and might cause an increase in the power consumption of the driver stages. On the other hand, the voltage swing at the gate will be smaller, because of the lower supply voltage of the scaled technology. As such, and if one assumes a digital driver stage, the power consumption of the driver is reduced by $\sqrt{2}$. Furthermore, the gate capacitance



(a) (b)

Figure 3.14. Supply voltage and threshold voltage for several submicron and nanometer technologies.



(a) (b)

Figure 3.15. Gate and drain capacitance of a $1000\,\mu\mathrm{m}$ wide transistor as a switch for several submicron and nanometer technologies.

slightly decreases with technology scaling as shown in figure 3.15(a). This will further reduce the DC power consumption of the driver stages.

It is rather difficult to draw conclusions for the drain capacitance regarding technology scaling. Figure 3.15(b) depicts the drain capacitance of a $1000\,\mu\mathrm{m}$ wide nMOS transistor, and a decrease could be recognized at least compared to $0.25\,\mu\mathrm{m}$. In any case, a lower parasitic drain capacitance will broaden the design space of the Class E amplifier.

From the previous, CMOS technology scaling might not seem to give a lot of problems for power amplifiers. The transistor needs to be made larger to keep

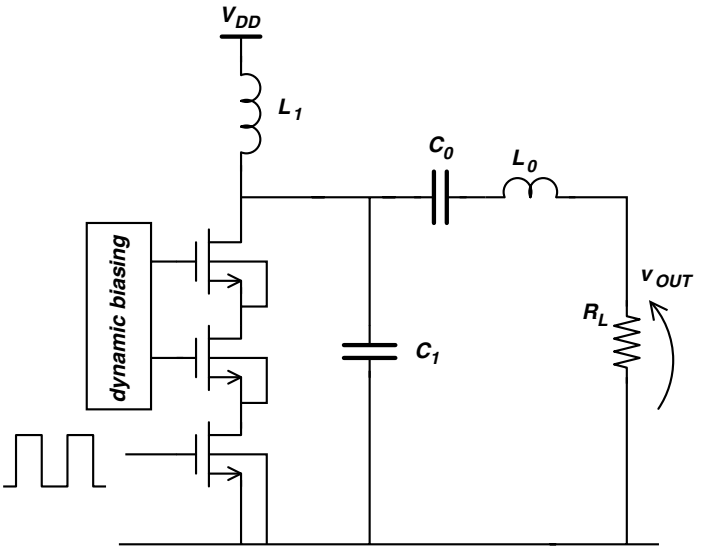


Figure 3.16. Class E power amplifier using stacked devices.

the ratio r_{on}/R_L constant, but even then, the dissipation in the driver stage can be reduced. However, the major obstacle is the reduction of the required load resistance R_L in order to achieve sufficient output power. Indeed, the load resistance must be scaled by V_{DD}^2, or in other words, divided by two with every new technology node. It will be shown in chapter 4 that this requirement strongly reduces the efficiency of the amplifier.

3.3.5 Device Stacking

As said before, the supply voltage of CMOS decreases with each new technology generation. To achieve sufficient output power at a low supply voltage, a transformation network is used which makes the impedance, *seen* by the amplifier, much smaller. Although this issued is tackled in chapter 4, it should be noted that such a network will also have losses and these losses are proportional to the impedance transformation ratio. The higher the impedance transformation, the lower the efficiency. Transforming 50Ω to 1Ω will be far less efficient than transforming 50Ω to 10Ω.

Another approach to increase the output power is to place multiple transistors *on top* of each other [Anne01], [Sern05] as demonstrated in figure 3.16. Only the bottom transistor is driven by the driver stage. If this transistor is turned off, the drain-source, drain-bulk and drain-gate voltage of each transis-

tor can be the same as if only one transistor is used and the supply voltage can be increased. For figure 3.16, the supply voltage could be made three times as high, provided that the gates of the two stacked transistors are properly driven. Also notice that the bulk of each nMOS transistor is connected to the source and not to the ground. This requires a triple well process, where the p-type bulk of the nMOS is separated from the p-type grounded substrate.

This technique is very interesting for low-voltage CMOS transistors as it allows to design a power amplifier that achieves a high output power although it uses a set of low-voltage transistors. On the other hand, if the same technology is used, stacked transistors or stacked devices enable to achieve a certain output power, without the use of an impedance transformation network, or with a less severe impedance transformation. Roughly, one could say that if n transistors are stacked, the supply voltage can be increased n times, and the load resistor can be made n^2 times as high to achieve the same output power. But there is more; if the supply voltage is made n times as high, the current in the circuit will be n times lower. Therefore, the dissipation in the transistor, which is the power dissipated in r_{on}, will be divided by n^2 and a higher efficiency will be the result.

The question arises which circuit parameters of the Class E amplifier are changing if the supply voltage or the load resistor changes. First of all, the value of the passive components of the Class E amplifier (L_1, C_1 and L_x) are independent of the supply voltage. An increase of the supply voltage will therefore not change the Class E amplifier. The efficiency of the amplifier will also be independent of the supply voltage, although the output power increases quadratically with the supply voltage.

The load resistance on the other hand will have a great impact on the values of the inductors and capacitors. To investigate this effect, figure 3.17(a) shows the required shunt capacitance versus load resistance. The output power decreases, solid line in figure 3.17(b), but the efficiency increases because less current is flowing through the switch.

If the supply voltage is adopted as

$$V_{DD}^{\text{new}} = \sqrt{\frac{R_L^{\text{new}}}{R_L}}\, V_{DD} \tag{3.41}$$

the effect of device stacking is simulated. The adjusted supply voltage is given by the dashed line in figure 3.17(c) and the corresponding output power stays more or less the same as depicted by the dashed line in figure 3.17(b).

The decrease of the required shunt capacitance will have a large influence on the selection of the on-resistance as it will lower the maximum transistor width that can be implemented. On the other hand, due to the smaller currents in the circuit, the transistor width can become smaller for the same efficiency.

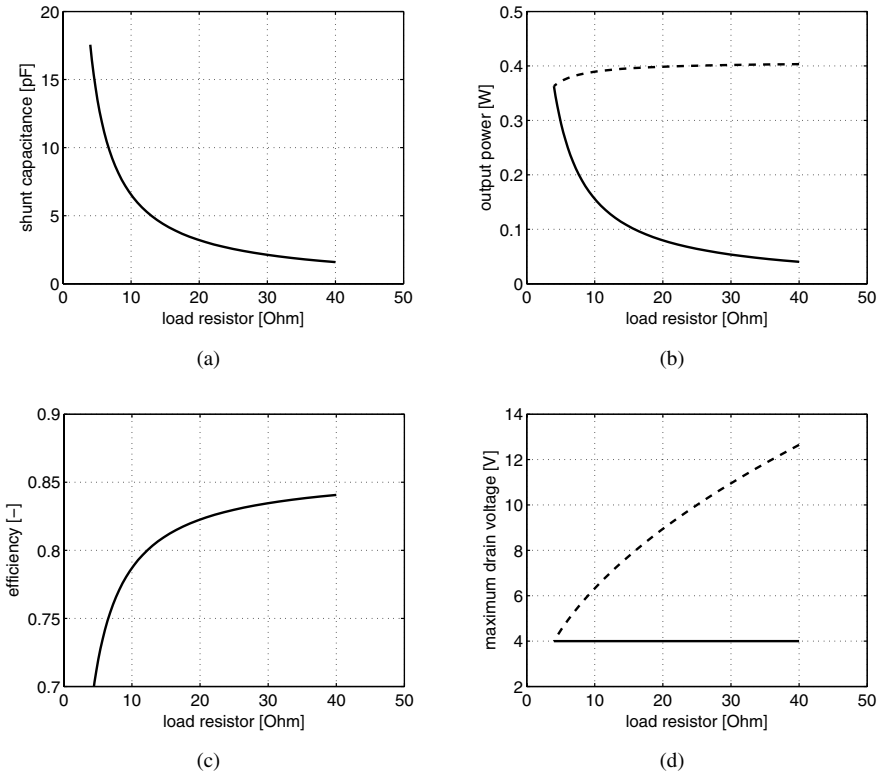


Figure 3.17. Class E design versus load resistance.

The estimation of the parasitic drain capacitance is not so clear for stacked devices. As an example, the circuit of [Sowl02], given in figure 3.18, is used. The lower transistor is the actual switch, the upper transistor serves to divide the high voltage swing across the two transistors as to maintain safe operation of the amplifier. The upper nMOST has its own separate p-well. This p-well is lying in an nwell, to isolate the device, and the nwell is connected to a higher voltage to prevent latch-up. As such, a large junction capacitance is present from the bulk of the upper transistor to the ground of the circuit, indicated by C_w. The drain capacitance of the upper nMOS transistor is in series with the parallel connection of the drain capacitance of the lower nMOST, the well junction capacitance C_w and the parasitic interconnection capacitance of the lower transistor. The value of the junction capacitance is dependent on the total area of the p-well including well contacts. All this makes an estimation of the parasitic drain capacitance hard, but for the analysis here, it is assumed that

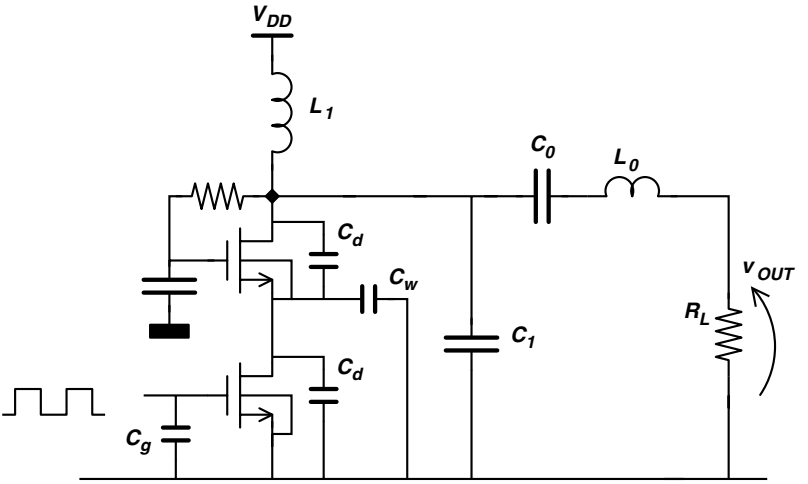


Figure 3.18. Class E power amplifier with two stacked transistors [Sowl02].

C_d equals C_w. Therefore, the total parasitic drain capacitance, now becomes

$$C_d = \frac{1.9\text{pF}}{1000\mu\text{m}} \cdot W_g \tag{3.42}$$

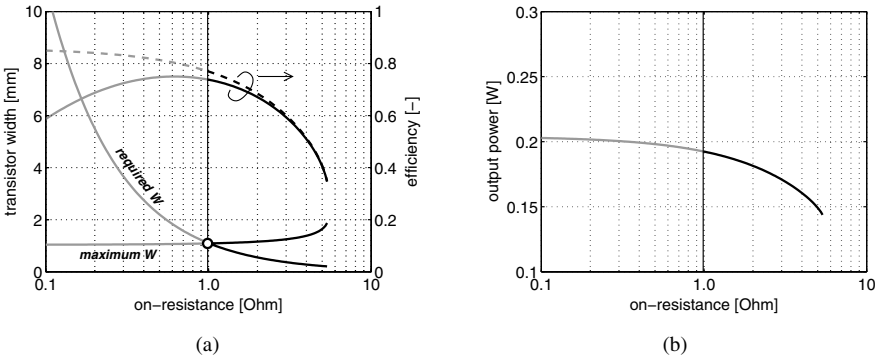
including the same parasitic interconnection capacitance as before ($1\ \text{pF}/1000\ \mu\text{m}$) for the two transistors.

The total resistance of the switch is now two times as large, since two transistors are connected in series and thus

$$r_{on} = 2 \cdot 1000\ \mu\text{m} \cdot 0.550\Omega \cdot \frac{1}{W_g} \tag{3.43}$$

The performance of such an amplifier versus the total on-resistance of the stacked transistors is given in figure 3.19. In this figure, it is assumed that two transistors are stacked and the load resistance is equal to $32\,\Omega$, which results in almost the same output power as in figure 3.12. It can be seen that the required on-resistance to achieve an $80\,\%$ efficiency is $0.7\,\Omega$ whereas it is $0.18\,\Omega$ in figure 3.12. However, the smaller value of the shunt capacitance narrows the design space. The maximum overall efficiency is obtained for an on-resistance of $0.5\,\Omega$, but this point can not be reached due to the presence of the parasitic drain capacitance. The maximum achievable overall efficiency is $75\,\%$. Although the maximum overall efficiency can not be obtained, the maximum achievable overall efficiency is still higher compared to the previous case.

Of course, whether the optimal performance can be reached will strongly depend on the technology and also on the layout of the transistors. But in any



(a) (b)

Figure 3.19. Class E design versus on-resistance for two stacked transistors and the same output power as in figure 3.12. The load resistance is 20 Ω. (a) Required and maximum transistor width, drain efficiency (dashed line) and overall efficiency, and (b) output power.

case, a higher load resistance will decrease the required shunt capacitance and this phenomenon will put a strong restriction on the applicability of stacked devices for the design of RF power amplifiers. On the other hand, the impedance transformation network will be less demanding since the required load resistance for the stacked case is four times as large compared to the single transistor case.

The second scenario, using stacked devices to achieve more output power within the same technology, is depicted in figure 3.20. Now, the supply voltage is doubled and the value of the load resistance and shunt capacitance remains the same. Because of the lower parasitic drain capacitance in the stacked case, the curve of maximum transistor width is higher than in figure 3.12. However, the curve of the required transistor width is lying two times as high since the two switches are placed in series. Again, the point of maximum overall efficiency cannot be reached.

It can be concluded that device stacking is, in a first place, beneficial to increase the efficiency of the amplifier and, though not tackled here, it will also increase the efficiency of the impedance transformation network. However, a higher supply voltage means a higher load resistance and a lower shunt capacitance for the same output power. Therefore, *too much* stacking will make the design space very small and it may occur that optimal operation cannot be obtained.

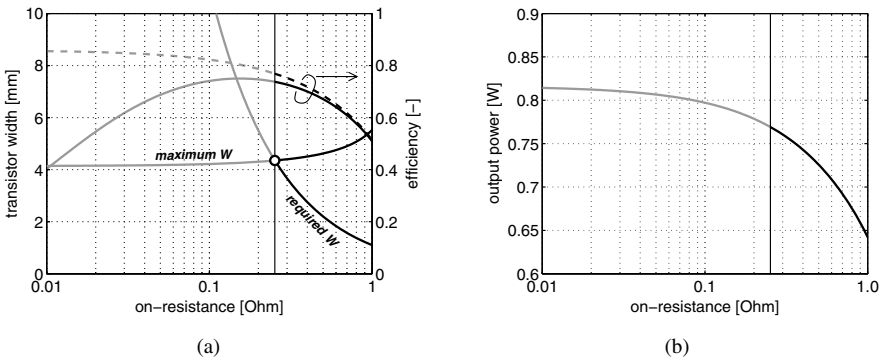


Figure 3.20. Class E design versus on-resistance for two stacked transistors to achieve more output power. The load resistance is $5\,\Omega$. (a) Required and maximum transistor width, drain efficiency (dashed line) and overall efficiency, and (b) output power.

3.3.6 Increasing the Operating Frequency

If the frequency of the Class E amplifier is increased, the shunt capacitance will also decrease and this will reduce the available design space and the maximum transistor width of the nMOS switch.

Figure 3.21 illustrates this effect and depicts the required value of the DC-feed inductance and shunt capacitance versus frequency. The decreased value of L_1 will be beneficial for the efficiency of the amplifier, which slightly increases. This increase is, however, counteracted by the skin effect which results in a higher series resistance at higher frequencies.

The lower value of C_1 will indeed make the design space very narrow, as can be seen in figure 3.22, which shows the design space of a Class E design at $10\,\mathrm{GHz}$. The width of the transistor has to be kept below $785\,\mu\mathrm{m}$ which corresponds to an on-resistance of $0.7\,\Omega$. At that point, the drain efficiency of the amplifier is $70\,\%$. The driver stage will consume a lot more power, since the frequency is now five times higher. Therefore, the overall efficiency is only $62\,\%$.

From this example, it can be concluded that a higher operating frequency will push the Class E amplifier to its limits with respect to the parasitic drain capacitance and optimum operation may not be achieved. If a high operating frequency is combined with stacked devices, things only gets worse.

Before leaving this section, a final note on the power consumption of the driver stage is required. In the previous analysis, a digital inverter or Class D amplifier is used to drive the gate capacitance of the nMOS switch. One could also use an inductor to make the gate impedance higher and as such, less

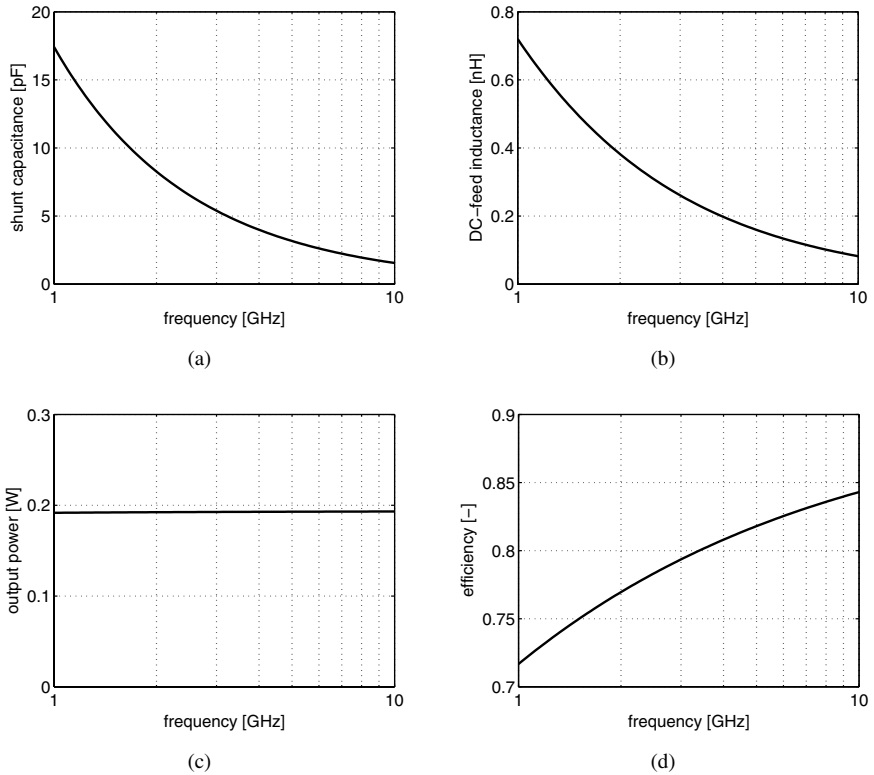
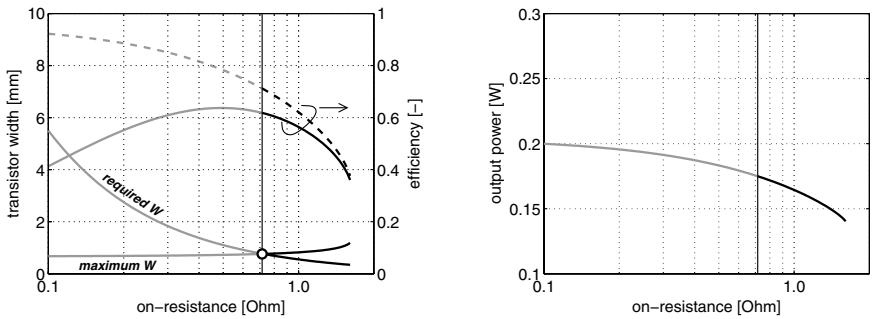


Figure 3.21. Class E design versus frequency.

driver power is required. Such an approach will, in a first place, increase the overall efficiency of the amplifier. But, looking back to figures 3.12, 3.19, 3.20 and 3.22, the point of maximum overall efficiency will shift to the left and this point might fall outside the design space. It depends on the parasitics of the nMOS switch whether this region of higher efficiency can actually be reached. Furthermore, integrated inductors consume a large silicon area and a tuned driver stage will therefore inherently increase the area cost of the amplifier.

3.3.7 Deviation from Class E: Class BE

Till now, only the Class E switching conditions were examined. For a lossless amplifier, Class E will indeed result in a $100\,\%$ efficiency. However, if losses are included, one could question whether the Class E conditions still guarantee the highest efficiency.



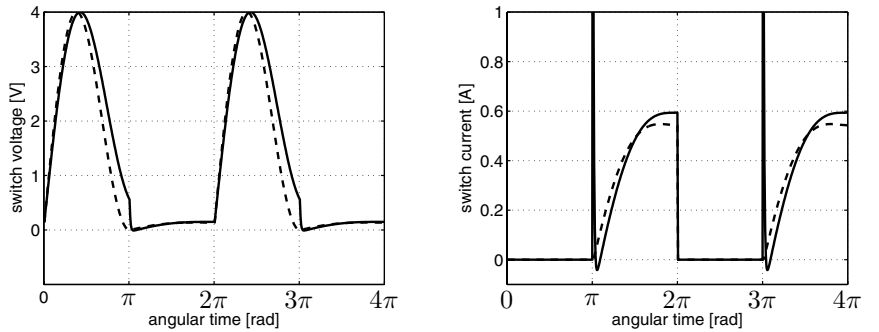
(a) (b)

Figure 3.22. Class E design versus on-resistance at 10 GHz with a load resistance of 8 Ω. (a) Required and maximum transistor width, drain efficiency (dashed line) and overall efficiency and (b) output power.

Table 3.2. Class E and deviated Class E design results.

parameter	Class E	deviated Class E
frequency	2 GHz	
maximum switch voltage	4 V	
load resistor	8 Ω	
switch resistance	0.25 Ω	
supply voltage	1.21 V	1.36 V
shunt capacitance	8.3 pF	10 pF
DC-feed inductance	0.38 nH	0.38 nH
output power	192 mW	221 mW
efficiency	76 %	78 %

To analyze this, the Class E design with a load resistor of 8 Ω, a switch resistance of 0.25 Ω and zero excess inductance, is recapitulated in table 3.2. If the value of the shunt capacitor is increased to 10 pF, the Class E conditions are no longer fulfilled. Extra losses are introduced when the switch is closed, and this will give rise to a current peak at the beginning of the conducting phase. On the other hand, the peak voltage across the switch will be lower, and thus the supply voltage can be increased from 1.21 V to 1.36 V in order to have the same maximum drain voltage of 4 V. The corresponding switch voltage and current are depicted in figure 3.23.



(a) (b)

Figure 3.23. Switch voltage and switch current for a Class E design (dashed line) and a deviated Class E design (solid line).

The output power is now $221\,\mathrm{mW}$ and the efficiency equals $78\,\%$. It can therefore be concluded that, although the amplifier does no longer fulfill the Class E switching conditions, the output power and efficiency are better than the *pure* Class E amplifier. A higher shunt capacitance will reduce the peak switch voltage and thus, the supply voltage can be augmented which in turn will increase the output power. This increase will in turn result in a higher efficiency and output power.

It may come as a surprise that *deviating* from Class E results in a better efficiency. The main cause for this phenomenon is the presence of the switch resistance r_{on} and the losses of the inductors. The Class E amplifier indeed results in the highest efficiency if no losses are included.

To further clarify this technique, figure 3.24 shows the output power and efficiency of the *deviated* Class E design, as well as the *pure* Class E design, which is indicated by a circle. If the shunt capacitance is slightly increased, the output power increases and the power dissipated in the switch stays almost the same. As such, both efficiency and output power increase.

Besides the shunt capacitance, also the DC-feed inductance can be varied to *boost* the performance of the Class E amplifier. This phenomenon can better be seen on the efficiency versus output power graph, given by figure 3.25.

During switch-on, the nMOS transistor will not act as a resistor but will be in its saturation region and thus will act as a current source. However, in the mathematical model, it is still assumed that the transistor can be modeled as a small resistance, which explains the large current peak that occurs in figure 3.23. In reality, this current peak will not be as large, but will be given by the

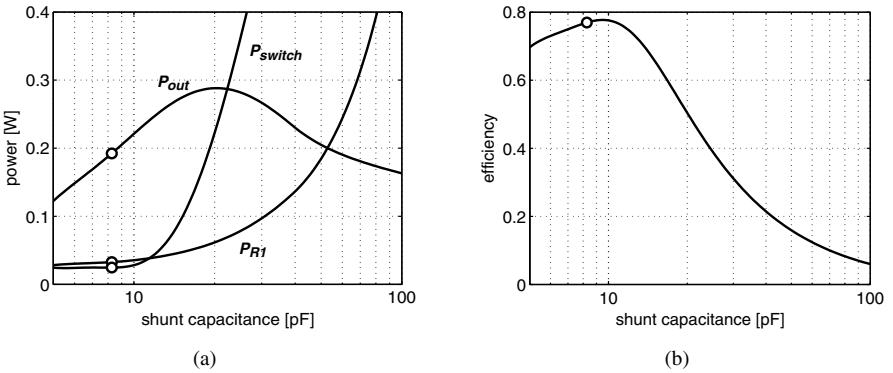


Figure 3.24. Power and efficiency for a deviated Class E design.

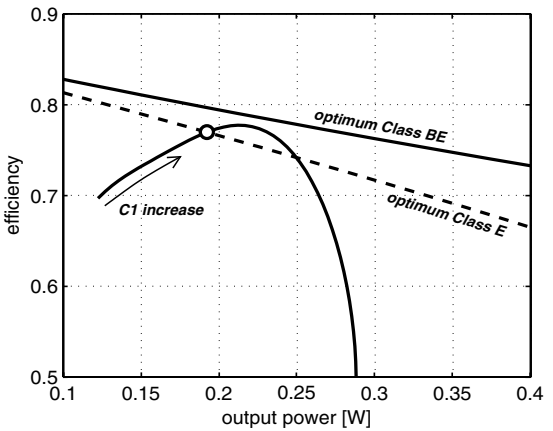


Figure 3.25. Efficiency versus output power for a deviated Class E power amplifier. Dashed line indicates the performance of the Class E design.

transconductance of the transistor. This will reduce the high current peak when the transistor starts to conduct current, but the benefit of both higher efficiency and output power are still valid.

In a *deviated* Class E amplifier, the transistor partly acts as a current source and partly as a switch. Therefore, the *deviated* Class E amplifier is sometimes referred to as a *Class CE* amplifier [Kazi89, Mert05]. However, the conduction angle is still $50\,\%$, and therefore it is more suitable to denote this amplifier as a *Class BE* amplifier.

The Class BE amplifier will also be beneficial regarding the design of the switch transistor. In the previous sections, it was found that the design space was limited due to the large parasitic drain capacitance of the nMOS switch. In this section, it is found that the shunt capacitance can be made larger than the value needed to fulfill Class E conditions. Hence, this will broaden the available design space which is particularly useful at higher frequencies or stacked devices.

Finally, this mixture between Class E and Class B has already been suggested at the end of section 2.4.7. Looking back to tables 2.3 and 2.4 should remind the reader why a mixture between B and E gives an excellent power amplifier.

3.4 CMOS Layout Aspects

3.4.1 Integrated Inductors

A Class E amplifier at gigahertz frequencies requires inductor values below $1\,\mathrm{nH}$. Such low values can no longer be made with multi turn rectangular or circular inductors. A recent trend [Aoki02b, HH04] is the use of *slab* or *trace* inductors to achieve high quality inductors with inductance values below $1\,\mathrm{nH}$. At first sight, it may seem strange to define the inductance of a wire, as opposed to the inductance of a loop. Of course, current will always flow in a loop, but if the loop is made large enough, the magnetic flux around a specific part of the inductor is only determined by the current in that part of the inductor and not by the current in other parts of the current loop.

The quality factor of an inductor is defined as [Yue98]

$$Q = 2\pi \cdot \frac{\text{energy stored}}{\text{energy loss in one oscillation cycle}} \tag{3.44}$$

Three factors will deteriorate the quality factor: the substrate loss, the self-resonance and the skin effect.

3.4.1.1 Substrate Loss

A lumped electrical model of a slab inductor is shown in figure 3.26. It consists of an inductance L_S with a series resistance R_S and a parasitic oxide capacitance $C_{par,ox}$ with a resistance R_{Si} in series to model the losses in the silicon substrate. The substrate loss is due to the penetration of the electric field inside the substrate.

To reduce the substrate loss, a patterned ground shield [Yue98] can be placed underneath the inductor to stop the electrical field from penetrating into the substrate. The influence on the magnetic field can be neglected, since the slots in the shield avoid the generation of induced currents. In the electrical model, the presence of the ground shield will short-circuit the substrate resistance R_{Si}.

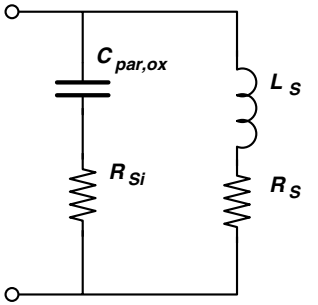


Figure 3.26. Lumped electric model of a slab inductor above a silicon substrate.

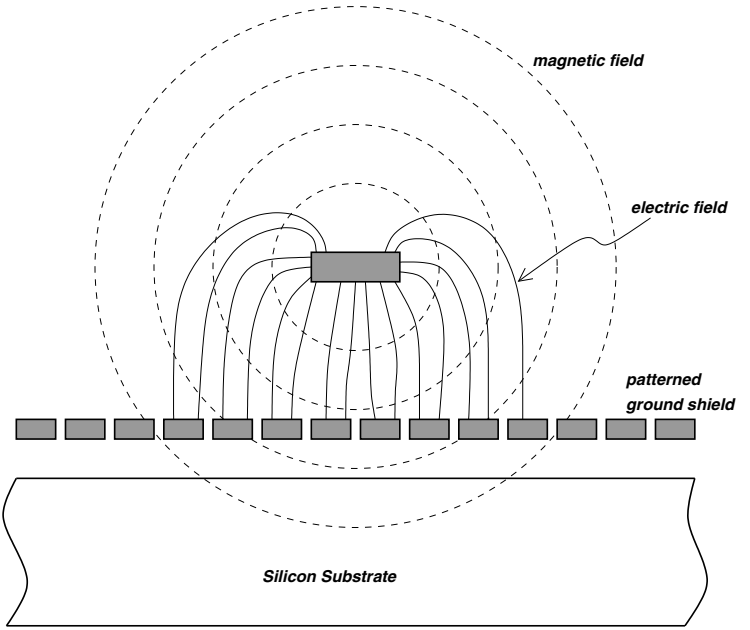


Figure 3.27. Sketch of the electric and magnetic field of a slab inductor above a patterned ground plane.

Figure 3.27 is a sketch of the magnetic and electric field distribution of a metal trace above a patterned ground shield.

3.4.1.2 Self-Resonance

The self-resonance is due to the parasitic capacitance from the inductor to the substrate or patterned ground shield. For an inductor, only the energy stored

in the magnetic field is of interest. Therefore, the electric energy present in the parasitic capacitance, has to be subtracted from the magnetic energy. At the self-resonant frequency, the energy in the magnetic field is equal to the energy in the electric field and the quality factor of the inductor becomes zero. For frequencies above the self-resonance frequency, no net magnetic energy is available from an inductor to any external circuit. With this in mind, the quality factor of a slab inductor above a patterned ground shield can be written as [Yue98]

$$Q = 2\pi \cdot \frac{\text{peak magnetic energy} - \text{peak electric energy}}{\text{energy loss in one oscillation cycle}} \tag{3.45}$$

$$= \frac{\omega L_s}{R_s} \cdot \left[1 - \left(\frac{\omega}{\omega_0}\right)^2\right] \tag{3.46}$$

$$= \frac{\omega L_s}{R_s} \cdot \left[1 - \frac{R_S^2 C_{par,ox}}{L_S} - \omega^2 L_S C_{par,ox}\right] \tag{3.47}$$

with ω_0 the self-resonance frequency of the inductor.

3.4.1.3 Skin Effect

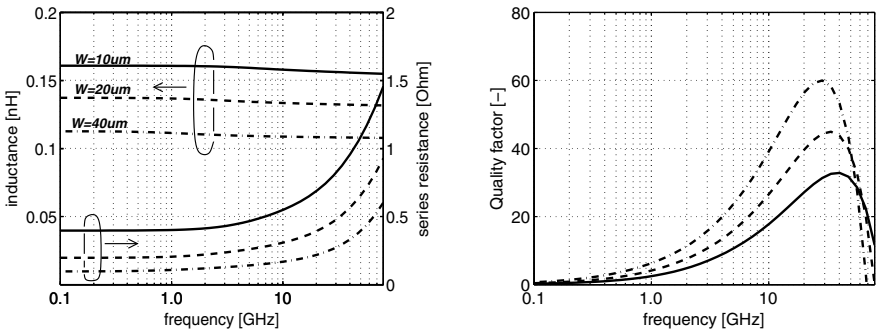
Another important aspect is the presence of a magnetic field inside the metal trace. If the conductor carries an AC current, the flux inside the conductor will change. Currents will be induced inside the inductor and they will try to counteract the magnetic flux changes. The initial current together with the induced current will cause a redistribution of the current through the inductor, and as such, the current will mainly flow near the surface of the conductor. This is well known as the *Skin Effect* and the induced currents are referred to as *Eddy currents* or *Foucault currents*.

Because of this, the effective area through which current will flow, is reduced. The higher the frequency, the more the current will be pushed outwards and the higher the resistance of the conductor. The current in the conductor will be maximum at the surface of the conductor and it will decrease exponentially towards the middle of the conductor. The *Skin Depth* is defined as the depth from the surface of the conductor, at which de current has decreased by $1/e$ and can be calculated as

$$\delta = \sqrt{\frac{2\rho}{\omega\mu}} \tag{3.48}$$

3.4.1.4 Slab Inductor Design

In general, a wider trace will result in a higher quality factor, but the self-resonance frequency will be lower. Also, a wider trace will have more substrate losses, but these can be reduced by the use of a patterned ground shield.



(a) (b)

Figure 3.28. A single metal trace. (a) Inductance and series resistance and (b) quality factor for different trace widths.

To demonstrate the discussed effects, the inductance of a metal trace in a $0.18\,\mu\mathrm{m}$ CMOS process is simulated. The silicon substrate has a conductivity of $10\,\Omega\mathrm{cm}$. The conductor consists of a metal trace with a thickness of $2\,\mu\mathrm{m}$ and a sheet resistance of $20\,\mathrm{m}\Omega\square$. The conductor is located $8\,\mu\mathrm{m}$ above the silicon substrate. The skin depth is $3.18\,\mu\mathrm{m}$ at $1\,\mathrm{GHz}$, $1\,\mu\mathrm{m}$ at $10\,\mathrm{GHz}$ and only $0.41\,\mu\mathrm{m}$ at $60\,\mathrm{GHz}$. The impedance of the trace can then be obtained from a finite element simulator like FASTHENRY.

Figure 3.28(a) gives the inductance and resistance for a $200\,\mu\mathrm{m}$ long metal trace versus frequency, for three different trace widths. The quality factor is calculated as well and given by figure 3.28(b). The increment of the series resistance is proportional to the square root of the frequency. The inductance remains more or less constant and thus the quality factor of the inductor, approximated by

$$Q = \frac{X_S}{R_S} = \frac{\omega L_S}{R_S} \tag{3.49}$$

will increase with frequency. At higher frequencies, the self-resonance effect will be responsible for a decrease of the quality factor and at the self-resonance frequency of the inductor, the quality factor becomes zero.

In a nanometer CMOS process, the maximum width of a metal trace is limited by the design rules to about $10\ldots20\mu\mathrm{m}$ and thus the maximum quality factor is about $20\ldots40$. Also, at higher frequencies the simple model of a lumped inductor with a lumped capacitor is no longer valid and one should use a distributed model. Nevertheless, the discussed effects remain valid.

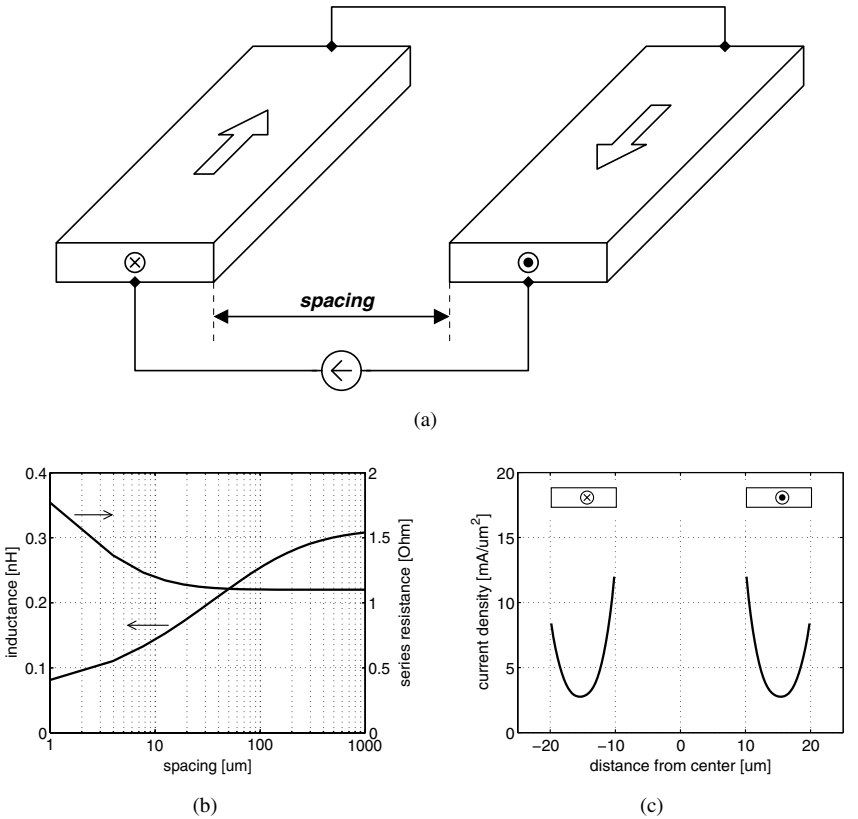


Figure 3.29. Two adjacent metal traces, placed in series.

Figure 3.30. Two metal traces placed in series. (b) Inductance and series resistance versus the spacing between the two traces (c) current distribution.

3.4.1.5 Proximity Effect

The presence of nearby currents will change the flux in and around the conductor and it will cause mutual coupling between nearby conductors. This is called the *Proximity Effect*. Two simple cases can be distinguished and will be discussed here: two parallel conductors with the same current, but flowing in opposite direction and two conductors with the same current, flowing in the same direction.

3.4.1.6 Two traces in series

For the first case, the magnetic field of the two currents will partially cancel each other, and the total magnetic field will be reduced. This will also reduce

the inductance of each of the wires. To demonstrate this effect, the total inductance of two metal traces in series is calculated. The structure is shown in figure 3.30. The connection between the end of the two wires and the connection to the current source is again assumed to be infinitely small. The same $0.18\,\mu\mathrm{m}$ technology as before was used, the conductor is $10\,\mu\mathrm{m}$ wide and the frequency is $10\,\mathrm{GHz}$. At that frequency, the self-resonance effect is not important.

Figure 3.30(b) shows the simulated inductance of the two wires versus the spacing between the two. If only one trace is simulated, see figure 3.28(b), a value of about $0.15\,\mathrm{nH}$ is obtained for a $10\,\mu\mathrm{m}$ wide trace at $10\,\mathrm{GHz}$. For two metal lines in series, one would thus expect a value of about $0.30\,\mathrm{nH}$. However, it can be seen that this value is only obtained if the two metal traces are very far apart from each other. If the two traces are close to each other, their magnetic fields will partially cancel and thus the total inductance becomes smaller. For the series resistance of the two traces, one would expect a value of about $1\,\Omega$, which is twice the value obtained from figure 3.28(b). If the traces are placed close to each other, closer that $20\,\mu\mathrm{m}$, one can see that the series resistance increases above this value. This can be understood if one looks at the current distribution inside the traces. The simulated current distribution for two traces with a width of $10\,\mu\mathrm{m}$, spaced $20\,\mu\mathrm{m}$ apart and conducting a total current of $100\,\mathrm{mA}$, is shown in figure 3.28(c). Due to the Proximity Effect, it seems as if the current is attracted by the other trace. The effective area, through which current will flow, becomes smaller and as such, the resistance increases.

The case of two conductors with the same current, but flowing in opposite directions, can be found in a single-turn inductor. It has the advantage that the inductor ends are close to each other, which is clearly not the case for a single-trace inductor. The disadvantage is a lower quality factor if the two traces are close to each other.

3.4.1.7 Two traces in parallel

The other case, two metal traces in parallel, is shown in figure 3.31. Again, the interconnection between the two traces is made ideal. The simulated inductance is shown in figure 3.31(b). Because of the parallel nature, one would expect an inductance of $75\,\mathrm{pH}$. This value is only obtained if the two traces are far enough from each other, since the magnetic field of the two conductors will add if they are close to each other. The same observation can be made for the parallel resistance of the two traces.

The current distribution for two metal traces in parallel, spaced $20\,\mu\mathrm{m}$ apart is given in figure 3.31(c). This distribution looks as if one would have only one conductor that is twice as wide. Since the oxide between the two conductors has the same magnetic permeability as the aluminum, the magnetic field sees

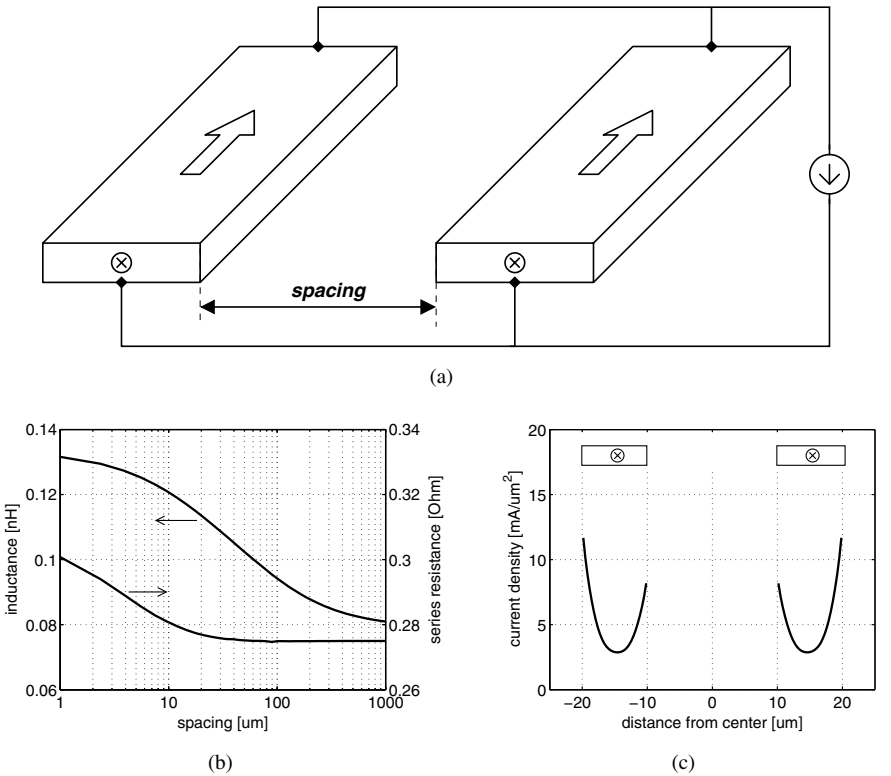


Figure 3.31. Two metal traces placed in parallel. (a) Inductance and series resistance versus the spacing between the two traces (b) current distribution.

no difference between the oxide and the metal. Therefore, the current is pushed outwards, just as if only one conductor is present.

3.4.2 Decoupling and Bondwires

In the discussion on the Class E amplifier, it was found that concluded that the DC-feed inductance will enhance the performance of the amplifier. The only drawback is the large current swing in the DC-feed, see figure 3.8, and this large AC current has to be delivered by the power supply.

Of course, the voltage source itself will not be able to deliver these high frequency AC currents. An obvious solution is to place some large decoupling capacitors close to the packaged chip. But this requires high quality capacitors, that are still capacitive up to the frequency of interest. The main problem

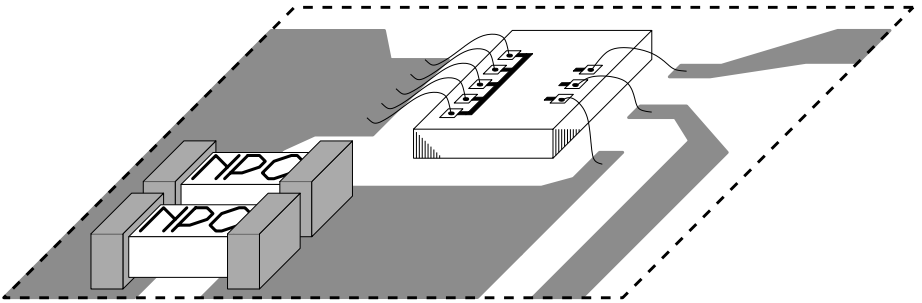


Figure 3.32. Example of a chip, mounted on a ceramic substrate or printed circuit board with bondwire connections. On the left, multiple bondwires are placed in parallel to reduce the total parasitic inductance of the connection to the chip.

with such off-chip capacitors is the parasitic inductance of their package, in series with the capacitor itself. So-called *NPO* capacitors achieve a parasitic inductance of only $0.6\,\mathrm{nH}$. But even then, at $1\,\mathrm{GHz}$, only capacitors with a value smaller than $20\,\mathrm{pF}$ are effective. After all, a capacitor of $20\,\mathrm{pF}$ in series with an inductance of $0.6\,\mathrm{nH}$ will be resonant at $1.5\,\mathrm{GHz}$. Roughly, at frequencies above $1\,\mathrm{GHz}$, off-chip capacitors larger than $20\,\mathrm{pF}$ will thus behave as an inductance and are of little or no use for decoupling at RF frequencies.

A typical solution for this problem is to place multiple small-value capacitors near or on top of each other, as shown in figure 3.32. However, the parasitic inductance of the metal traces from the decoupling capacitors to the chip and especially the inductance of the bondwires will also limit the frequency up to which this off-chip decoupling is still effective. From this it can be concluded that off-chip decoupling is useless at frequencies above $1\,\mathrm{GHz}$, unless one can make the parasitic inductance of the bondwires very small and one is able to place multiple capacitors with a small value very close to the chip.

The total inductance of the bonding can be reduced by placing multiple bondwires in parallel. On the other hand, in the previous section it was found that placing multiple metal traces in parallel is only efficient if the spacing is made wide enough. The same holds for the bondwires that are used to connect the chip with the package, see figure 3.32. For a typical aluminum bondwire, with a length of $3\,\mathrm{mm}$ and a diameter of $25\,\mu\mathrm{m}$, figure 3.33(a) gives the resulting inductance if one or more bondwires are placed in parallel, at $1\,\mathrm{GHz}$. For a typical bondpad spacing of $120\,\mu\mathrm{m}$, the decrease of the total inductance is indeed rather modest. On could question what would be the best; less bondwires spaced further apart or more bondwires close to each other. In figure 3.33(b), it can be seen that, for a given length over which one can place bondwires,

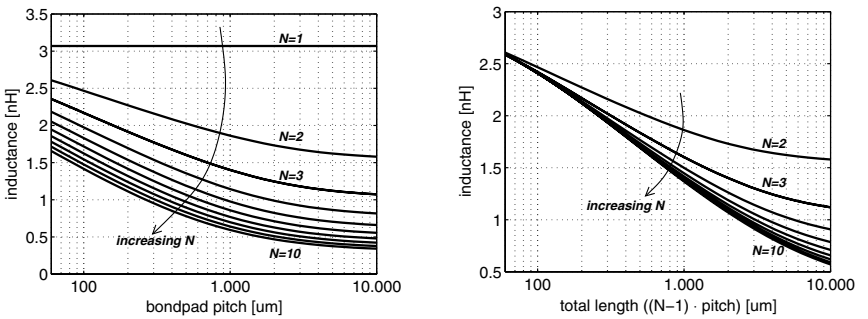


Figure 3.33. Total inductance of parallel bondwires at 1 GHz. (a) Inductance versus bondpad pitch and (b) inductance versus length, consumed by the bondwires.

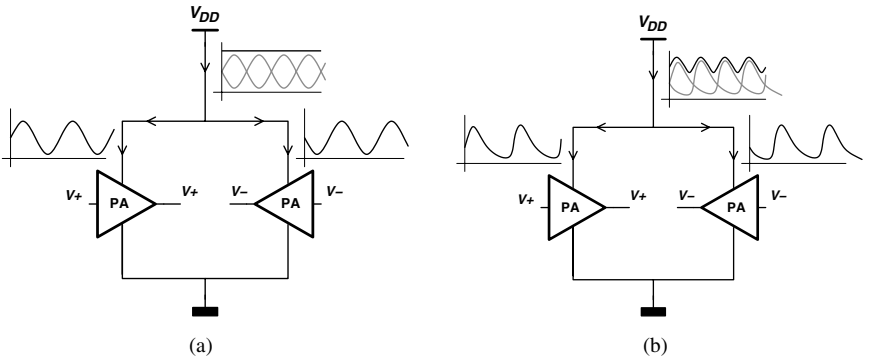


Figure 3.34. AC supply current: (a) a differential amplifier will have no AC supply current, (b) a switching differential circuit will still have some AC supply current.

it is always the best to place as many as possible to decrease the inductance. But again, the benefit is rather small. In fact, the inductance of the bondwires will easily exceed the parasitic inductance of the capacitor, and thus the maximum frequency up to which off-chip decoupling is effective will be lower than 1 GHz. Even worse, the combination of the bondwire inductance and the off-chip decoupling capacitors will have one ore more resonant peaks. At these frequencies, the supply rail will have a large impedance and oscillations may occur. Techniques to suppress these oscillations can be found in [Inge97].

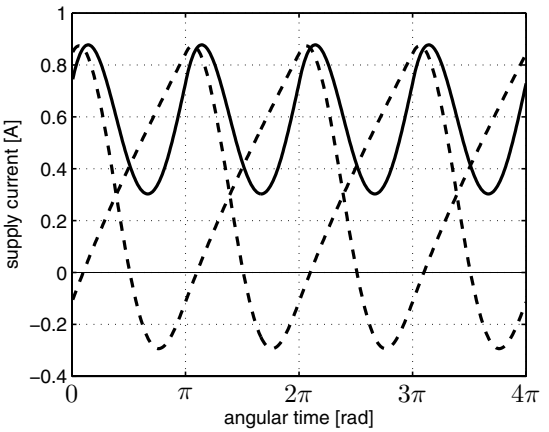


Figure 3.35. Supply current for two differential Class E amplifiers (dashed line) and the total supply current (solid line).

One way to really solve the problem of the AC power supply currents is the use of differential structures combined with on-chip decoupling. For a differential operation, two amplifiers are placed on the same chip, close to each other, and are driven by a signal that is 180 degrees shifted in time. If the supply currents are symmetrical, the current through the bondwire will be DC, as indicated in figure 3.34(a). In reality, such a situation only occurs for circuits with a relative small AC current, like small-signal amplifiers or operational amplifiers. If switching circuits, like digital buffers are used in a differential manner, the total supply current will not be a pure DC current. It all depends on the symmetry of the current waveform. For the Class E amplifier, the current through the DC-feed is also not fully differential, but the total AC current of a differential Class E amplifier will be considerably less compared to its single-ended version. This is demonstrated in figure 3.34(b). Figure 3.35 shows the current through the DC-feed for two Class E amplifier that operate in a differential manner. Clearly, the sum of the two currents has a much lower swing and looks more like a sinewave compared to its single-ended counterpart. Another important fact is that the frequency of the supply current will be twice the operating frequency of the amplifier. Therefore, the requirements of the decoupling can be divided by two and one could state that a differential circuit only requires less than half the decoupling of the same but single-ended circuit. Furthermore, the supply noise, injected in the substrate, will occure at twice the operating frequency and thus any circuitry tuned to the operating frequency will pick-up less noise.

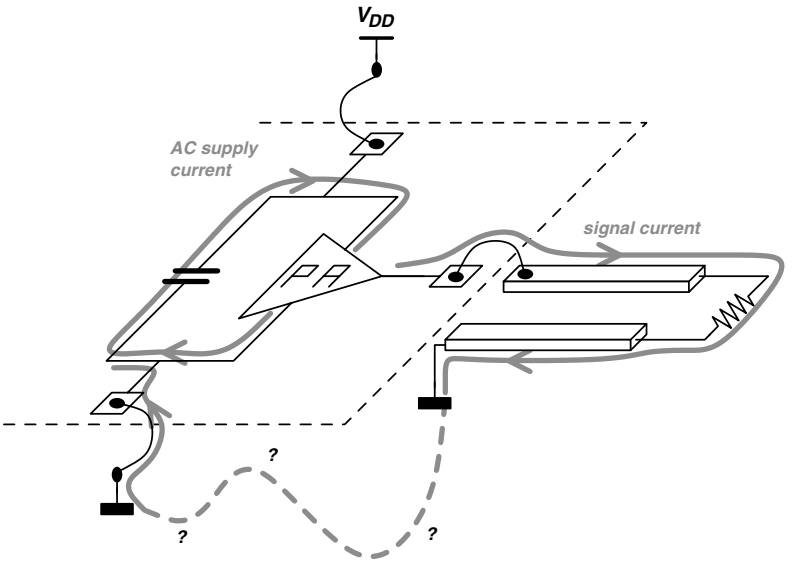


Figure 3.36. Integrated power amplifier without a dedicated return path for the signal current.

In such a *pseudo-differential* amplifier, still some AC current has to be delivered. As mentioned before, this AC current can not be delivered by off-chip decoupling capacitors and some on-chip decoupling is required. The benefits of on-chip decoupling are clear: the self-resonance frequency of on-chip capacitors is very high and the capacitors can be placed extremely close to the amplifier. If sufficient decoupling is placed on-chip, there is no need to minimize the ground and power supply inductance. The area that would be used for ground and supply bonding pads can better be filled with on-chip decoupling.

By using on-chip decoupling and differential structures, the flow of the AC supply currents can be controlled and one can ensure that these currents remain on-chip. But of course, the output signal, which is also a large AC current, has to get off the chip sooner or later. Again, on has to carefully control the flow of this AC current, by using a dedicated current return path. Figure 3.36 depicts a common situation. The AC supply current will be delivered by the on-chip decoupling and thus the supply and ground bondwires are, at first sight, not critical. However, the AC return current of the output signal will also flow through the ground bondwires. Therefore, this inductance has to be made as small as possible. An even bigger problem is the return current from the load, back to the chip. If no special attention is paid to this path, a lot of parasitic inductance might be present between the ground of the load or measurement equipment, and the ground of the chip. Therefore, it is better to create a dedicated ground connection from the chip to the ground of the load. In such a

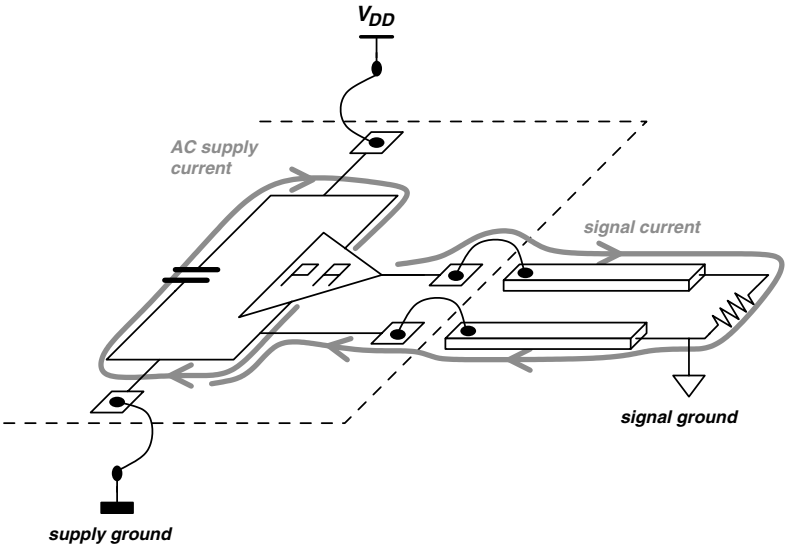


Figure 3.37. Integrated power amplifier with dedicated bondwires for the signal current.

case, depicted by figure 3.37, the power supply and ground bondwires do not have to be small. The output signal AC current will flow through a dedicated return path and the corresponding bondwires have to be carefully designed and are the critical ones. Note that the current through the two signal bondwires will flow in opposite directions and by placing them close to each other, their total inductance will be reduced because of the mutual coupling. Of course, one can also place many bondwires in parallel to reduce the inductance of the signal path. The best approach is to incorporate the signal bondwires in the design or the use them in an output matching or filtering network.

The inductance of the return path may also cause problems on-chip. Figure 3.38 depicts an example of an nMOS transistor with a trace or slab inductor at the drain. Assume that this trace inductor is $15\,\mu\mathrm{m}$ wide and $200\,\mu\mathrm{m}$ long. From a circuit point of view, the source of the transistor is connected to the ground, and so does the bottom plate of the decoupling capacitance. In a layout, however, these two points are physically at a different location, approximately $200\,\mu\mathrm{m}$ apart from each other. If one would connect them with a metal wire with a width of $15\,\mu\mathrm{m}$, an inductance would be present between these two points, equal to the drain inductance. To avoid such a condition, a metal ground plane can be used all over the chip area, which ensures a low-inductive path between the different ground connections. Underneath the slab inductor, the metal ground plane should be slotted to avoid the generation of Eddy currents.

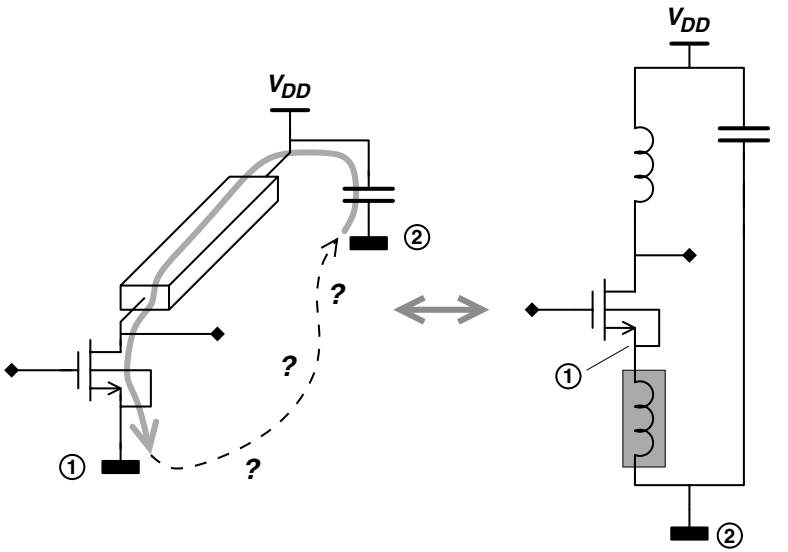


Figure 3.38. Example of return-path currents on-chip.

3.5 Conclusion

This chapter has focussed on the design of the Class E amplifier in a CMOS technology. In section 3.2, the Class E operating principles were reviewed and the difficulties to obtain analytical design equations were discussed. A new design methodology, based on a state-space description of the Class E amplifier, was developed.

In section 3.3, the state-space technique was then applied to discuss the design aspects of the Class E amplifier in CMOS. Four effects will impede the design: the low supply voltage, the power loss of the inductors, the on-resistance of the transistor and the parasitic capacitances of the transistor.

First, the effect of the DC-feed inductance on the design of the Class E amplifier was investigated. It was found that a smaller DC-feed results in an improved performance in terms of output power and efficiency. Next, the sizing of the nMOS switch is discussed. In this regard, the relative large parasitic capacitances of CMOS play an important role and they will restrict the design space. The gate capacitance needs to be charged and discharged and the larger the switch, the higher the power consumption of the driver stages. This will result in an optimal transistor width for a maximum overall efficiency. The parasitic drain capacitance will limit the maximum size of the nMOS switch. In this regard, the Class E amplifier has an advantage compared to the Class B and Class F amplifiers. After all, the parasitic drain capacitance can be absorbed by the required shunt capacitance of the Class E network. The width of the nMOS

can be increased until the value of the parasitic drain capacitance exceeds the value of the required shunt capacitance. The influence of technology scaling has been investigated as well. Technology scaling results in better switches, but not good enough to counteract the effect of the reduced supply voltage. As such, the major obstacle for the design of an RF power amplifier in nanometer technologies and beyond, is situated in the impedance transformation network. This will be the topic of the next chapter. A promising solution for the low supply voltage is the use of stacked devices. It was found that this enables to achieve the same or even better performance in a smaller technology. On the other hand, a higher load resistance requires a lower shunt capacitance and as such, device stacking may narrow the design space. A higher operating frequency will also narrow the design space because of a lower required shunt capacitance. Finally, the deviation from Class E, denoted as Class BE has been investigated. A mixture between Class E and Class B results in a higher output power and higher efficiency. The only limitation for Class BE is the drain voltage of the transistor when the latter starts to conduct current. One has to ensure that this voltage remains below the nominal supply voltage of the technology to avoid the generation of hot carriers that might be injected in the gate oxide.

In section 3.4, some layout aspects of CMOS power amplifier design were discussed. First, the use of a metal trace as an inductor was covered and the design and tradeoffs have been clarified. To conclude this chapter, the issues of on-chip and off-chip decoupling and the influence of the bondwires have been discussed and some guidelines for *good* PA design were outlined.

The design of the impedance transformation network is clearly of great importance and crucial for the design of a power amplifier in a low voltage technology. Therefore, the next chapter is entirely devoted to this topic.

Chapter 4

IMPEDANCE TRANSFORMATION AND POWER COMBINATION

4.1 Introduction

In RF and Microwave, it is common to use a characteristic impedance of $50\,\Omega$. Antennas, antenna filters and other microwave components typically have single-ended input and output impedances of $50\,\Omega$. If a power amplifier, designed in a low-voltage CMOS technology, is directly connected to this $50\,\Omega$ resistance, the output power will be low; a sine wave with an amplitude of $1\,\mathrm{V}$ across a $50\,\Omega$ resistor results in an output power of only $10\,\mathrm{mW}$.

To achieve sufficient output power an *impedance transformation network* is used. At a specific frequency, the $50\,\Omega$ load is converted to a smaller value, by means of inductors and capacitors. The power amplifier will thus deliver more power to the impedance transformation network, and if the latter is lossless, all this power will reach the output. Though many solutions exist in litterature, not all of them can be integrated. Two networks are discribed in this chapter; the L-match network and the lattice-type LC balun network, and they are both excellent candidates for CMOS integration.

One should not confuse these networks with *impedance matching networks* e.g. for low noise amplifiers. It is not the purpose to match the $50\,\Omega$ load with the output impedance of the amplifier, but one tries to transform the load to a smaller value, in a power efficient way.

4.2 L-match Impedance Transformation

One of the most popular impedance transformation networks is the L-match, shown in figure 4.1. This network, as given here, is able to convert the $50\,\Omega$ load resistance R_L to a lower value R_{in}. In practice, the load resistance can be the impedance of the antenna, the input impedance of an antenna filter or the input impedance of the measurement equipment.

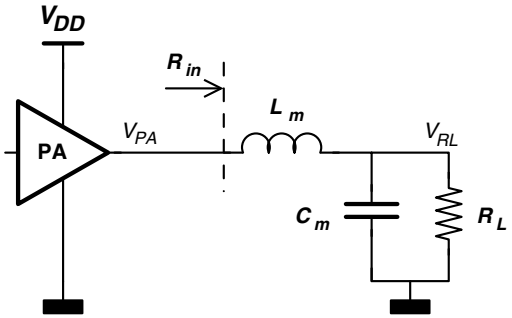


Figure 4.1. L-match impedance transformation network.

4.2.1 Basic Equations

To get a better insight in the principles of a transformation network, the transformation from the load impedance R_L to the input impedance R_{in} is depicted in figure 4.2. It uses the equivalent circuit ideas of section 2.3.8. First, the parallel circuit of C_m and R_L is converted to a series equivalent of C_S and R_m, and they can be calculated as

$$R_m = \frac{R_L}{1+Q_C^2} \approx \frac{R_L}{Q_C^2} \tag{4.1}$$

$$C_S = C_m \cdot \left(1+\frac{1}{Q_C^2}\right) \approx C_m \tag{4.2}$$

with

$$Q_C = \frac{R_L}{1/(\omega C_m)} = \frac{1/(\omega C_S)}{R_m} \tag{4.3}$$

To create a resistive impedance at the operating frequency ω_0, inductor L_m is designed to resonate with capacitor C_S.

$$\omega_0 = \frac{1}{\sqrt{L_m C_S}} \tag{4.4}$$

Thus, at resonance and without losses, the resistive input impedance R_{in} of the L-match network becomes equal to R_m. The *impedance transformation ratio, r* of the network is defined as

$$r \equiv \frac{R_L}{R_{in}} = \frac{R_L}{R_m} \tag{4.5}$$

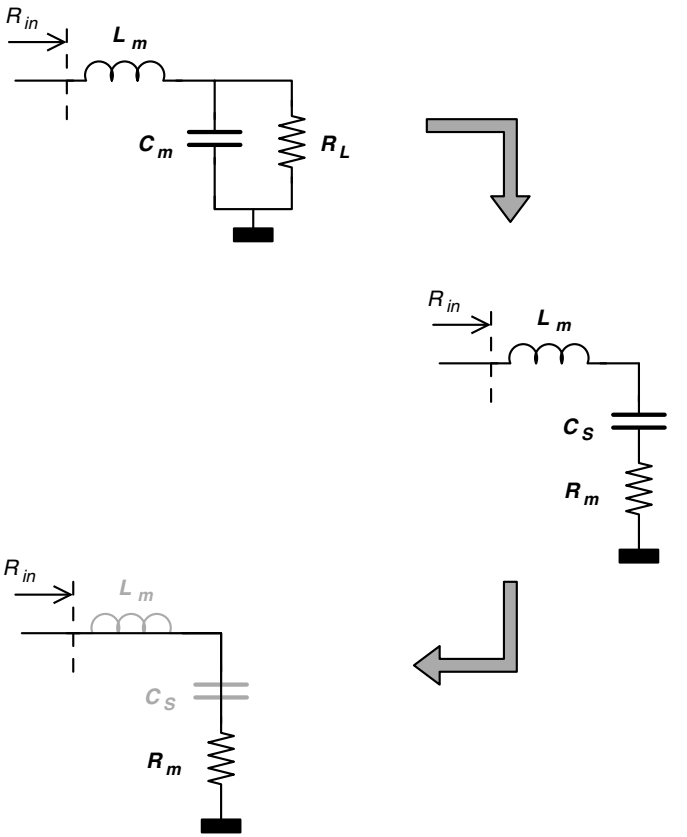


Figure 4.2. L-match impedance transformation network.

The design equations can be obtained from previous equations. First, the quality factor of the transformation network is defined

$$Q_m = \sqrt{\frac{R_L}{R_m} - 1} = \sqrt{r - 1} \tag{4.6}$$

From this, the exact values of L_m and C_m can easily be obtained as

$$C_m = \frac{Q_m}{\omega R_L} \tag{4.7}$$

$$L_m = \frac{Q_m R_m}{\omega} \tag{4.8}$$

By using following equality

$$B = \omega L_m \tag{4.9}$$

The input impedance of the L-match network can also be written as

$$R_{in} = R_m \approx \frac{B^2}{R_L} \tag{4.10}$$

and the impedance transformation ratio becomes

$$r \approx \left(\frac{R_L}{B}\right)^2 \tag{4.11}$$

An impedance transformation network can be interpreted as a transformer. The output voltage is higher than the input voltage, but the output impedance is also higher than the input impedance. In other words, power is conserved, just as with a transformer, and the *voltage gain* of the impedance transformation network is equal to

$$\frac{V_o}{V_{PA}} = \sqrt{\frac{R_L}{R_m}} = \sqrt{r} \tag{4.12}$$

Without the impedance transformation network, the output power would only be

$$P_{o,0} = \frac{V_{PA}^2}{R_L} \tag{4.13}$$

were it is assumed that V_{PA} is an RMS value. With the impedance transformation network in place, the output power becomes

$$P_o = \frac{V_{PA}^2}{R_m} \tag{4.14}$$

An important parameter is the *power enhancement ratio, E*, defined as the ratio of P_o to $P_{o,0}$ i.e. the output power with the impedance transformation network divided by the output without the network [Aoki02a].

$$E \equiv \frac{P_o}{P_{o,0}} \tag{4.15}$$

For the lossless L-match network, the power enhancement ratio is equal to the impedance transformation ratio.

$$E = \frac{R_L}{R_m} = \frac{R_L}{R_{in}} = r \approx \left(\frac{R_L}{B}\right)^2 \tag{4.16}$$

4.2.2 Inductor Loss and Efficiency

The loss of the L-match network will be dominated by the loss of the inductor L_m. This can be modeled by a series resistance R_{Lm} as indicated in figure 4.3. The quality factor of the inductor is defined as

$$Q_L = \frac{\omega L_m}{R_{Lm}} \tag{4.17}$$

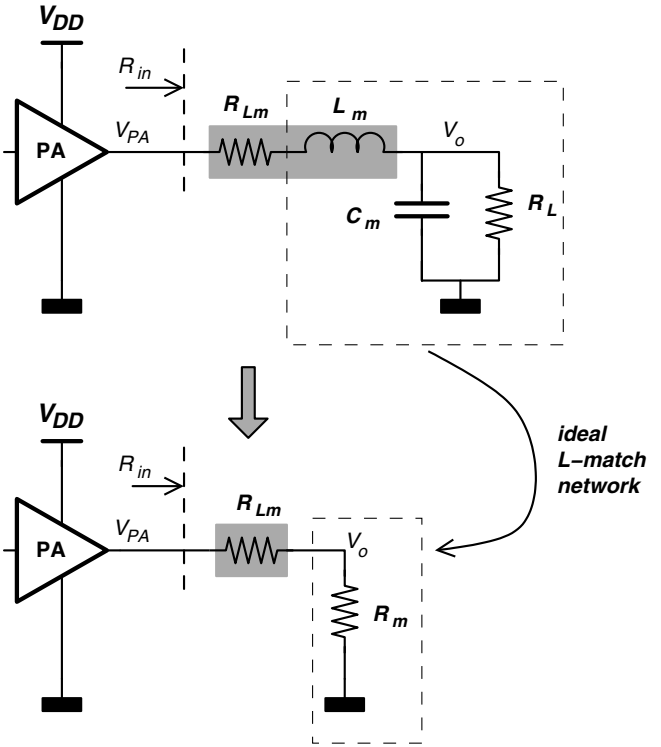


Figure 4.3. L-match impedance transformation network with series inductor loss.

The loss resistance R_{Lm} can virtually be shifted out of the impedance transformation network, and as such it is in series with an ideal L-match network. At the resonance frequency,R_{Lm} and R_m are thus in series.

$$R_{in} = R_m + R_{Lm} \tag{4.18}$$

The impedance transformation ratio for an L-match network with loss becomes

$$r = \frac{R_L}{R_{in}} = \frac{R_L}{R_m + R_{Lm}} \tag{4.19}$$

Since the L-match network is still assumed to be ideal, the calculation remains the same as before

$$Q_m = \sqrt{\frac{R_L}{R_m} - 1} \tag{4.20}$$

$$C_m = \frac{Q_m}{\omega R_L} \tag{4.21}$$

$$L_m = \frac{Q_m R_m}{\omega} \tag{4.22}$$

The output power is still equal to the power dissipated in the resistance R_m.

$$P_o = \left(V_{PA} \frac{R_m}{R_{Lm} + R_m} \right)^2 \frac{1}{R_m} = V_{PA}^2 \frac{R_m}{(R_{Lm} + R_m)^2} \tag{4.23}$$

The total power flowing into the impedance transformation network is

$$P_{in,L} = \frac{V_{PA}^2}{R_m + R_{Lm}} \tag{4.24}$$

With these equations, the efficiency of the transformation network can be calculated

$$\eta_L = \frac{P_o}{P_{in,L}} = \frac{R_m}{R_{Lm} + R_m} \tag{4.25}$$

Using equations 4.17 and 4.22, the efficiency can also be written as

$$\eta_L = \frac{Q_L}{Q_L + Q_m} \tag{4.26}$$

and the impedance transformation ratio becomes

$$r = \eta_L \cdot \frac{R_L}{R_m} \tag{4.27}$$

It is the purpose of the impedance transformation network to increase the output power. However, because of the series loss of the inductor, less power will now reach the output. Therefore, the impedance transformation ratio is not that important, since a high impedance transformation ratio with a low efficiency still results in a low output power. The power enhancement ratio is a better figure, that takes these losses into account.

$$E = \frac{P_o}{P_{o,0}} = \frac{R_m R_L}{(R_m + R_{Lm})^2} = \eta_L \cdot r \tag{4.28}$$

Using equations 4.26 and 4.28, the relation between the efficiency and the power enhancement ratio can be derived:

$$\eta_L = \sqrt{E \cdot \frac{R_m}{R_L}} \tag{4.29}$$

It is important to realize that the efficiency of the transformation network will be dependent on both the quality factor of the inductor Q_L and the power

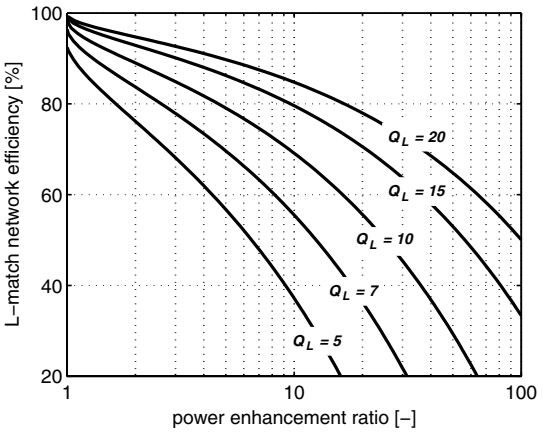


Figure 4.4. L-match efficiency versus power enhancement ratio for different values of inductor quality factor.

enhancement ratio. An approximate analysis can be made as follows. The quality factor of the transformation network is approximated as

$$Q_m \approx \sqrt{\frac{R_L}{R_m}} \tag{4.30}$$

and thus,

$$L_m = \frac{Q_m R_m}{\omega} \sim \sqrt{R_m} \tag{4.31}$$

On the other hand, the series loss resistance of the inductor is proportional to

$$R_{Lm} = \frac{\omega L_m}{Q_L} \sim L_m \sim \sqrt{R_m} \tag{4.32}$$

As an example, assume that the transformed resistor R_m is made two times as small to increase the output power. This however will only reduce the loss resistance R_{Lm} by $\sqrt{2}$ and the efficiency will decrease. This example indicates that the efficiency of the transformation network will be low at higher values of E. To analyze this effect, various values for R_m can be chosen and the corresponding values of E and η_L for a given inductor quality factor, can be calculated. The result is given in figure 4.4, which depicts the efficiency of the transformation network versus the power enhancement ratio for several values of inductor quality factor.

As an example, the amplifier given in table 3.2 can be used. An output power of $200\,\mathrm{mW}$ is achieved with a load impedance of $8\,\Omega$. If the load would have been $50\,\Omega$, the output power would only be $32\,\mathrm{mW}$. Thus, a power enhancement ratio of about 6.25 is required. If the inductor can be made with a

quality factor of 10, figure 4.4 shows that the efficiency of the transformation network alone for $E = 6.25$ equals 75%. The efficiency of the amplifier itself is about 78%, see table 3.2, and thus the total drain efficiency becomes 59%, not including the power consumption of the driver stages. If an output power of $1\,\mathrm{W}$ is aimed for, a power enhancement ratio of at least 32 is required and the efficiency of the transformation network alone, for a Q_L of 10, is only 45%. Furthermore, the power amplifier itself will have a lower efficiency at such high output power.

To obtain a better result, either the quality factor of the inductor has to be increased or the power enhancement ratio has to be lowered. The latter is achieved if the supply voltage can be increased while maintaining a reliable operation. Figure 4.4 is thus an argument to use stacked devices as discussed in section 3.3.5. If the amplifier can operate at a higher voltage, E may be lower, resulting in a higher efficiency.

Figure 4.4 is also an argument to use *slab* or *trace* inductors with a high quality factor. However, such inductors are only feasible if low inductance values are required, which happens when the operating frequency is increased. As such, low-voltage and high-frequency are complementary and, to some extend, indispensable to each other.

To achieve sufficient output power at a low supply voltage, a multi segment L-match network can be used to minimize the power losses. Also, a transformer can be used to combine several amplifiers in order to achieve sufficient output power [Aoki02b]. In the next section, another alternative will be demonstrated that inherently allows to control the output power and can easily be merged with the Class E amplifier.

4.3 Power Combination

A high output power at a low supply voltage requires a high power enhancement ratio E. If an L-match network is used, this would result in a low efficiency as demonstrated by figure 4.4. An obvious solution is to use two differential amplifiers and a differential power combiner or external RF transformer, as shown in figure 4.5. Besides the benefits of a differential amplifier, discussed in section 3.4.2, the output power is two times as high. Each power amplifier can have a lower power enhancement ratio, and thus a higher efficiency. This idea can be further expanded by placing multiple power amplifiers in parallel, to achieve an even higher output power. However, a microwave power combiner or an RF transformer is difficult to integrate in CMOS.

Another approach is demonstrated in figure 4.6. The output of a second power amplifier is connected to the capacitor of the L-match impedance transformation network. The output of this amplifier is 180 degrees shifted in time. The inductor will give a positive phase shift and the capacitor will give an equal

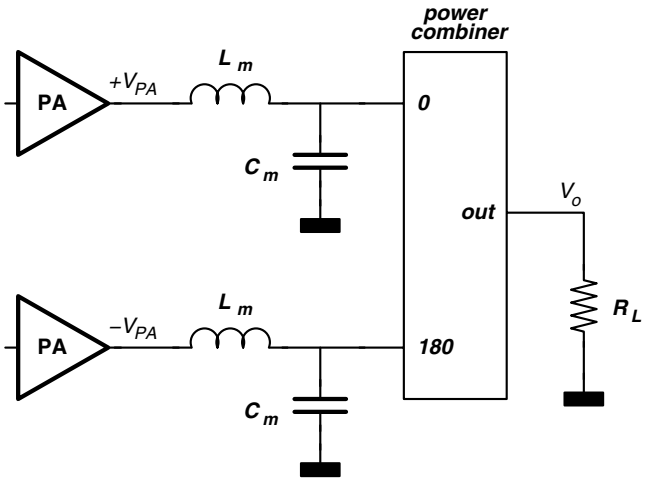


Figure 4.5. Two differential power amplifiers with a differential power combiner.

but negative phase shift. Therefore the currents at the output will recombine in phase and a higher output power will be the result.

The network of figure 4.6 has two interesting properties; first it performs the impedance transformation function required to achieve sufficient output power in a low voltage technology. Second, the two outputs of a differential amplifier are combined to a single-ended output. This circuit thus allows for both impedance transformation and power combining with a single-ended output. Since it converts a differential or **bal**anced signal to a single-ended or **un**balanced signal, this network is also denoted as a balun. Balun networks can also be made with transmission lines, but to stress the use of lumped inductors and capacitors, the network is named *Lattice-type LC Balun* or *lumped element Balun* [Baka02, Crip99]. It has been derived here starting from an L-match network, but the lattice-type LC balun stems from antenna theory and was already known since 1932 as an antenna balun [BT32].

4.3.1 Basic Equations

The calculation of the lattice-type LC balun of figure 4.6 is as follows. The RMS output voltage of the two amplifiers is assumed to be identical and equals V_{PA}. An RMS value is always positive, but the notation $+V_{PA}$ and $-V_{PA}$ is used here to indicate that the two output voltages are 180 degrees out of phase. The current flowing in each of the branches is equal to

$$I_L = \frac{V_{PA} - V_o}{j\omega L_m} \tag{4.33}$$

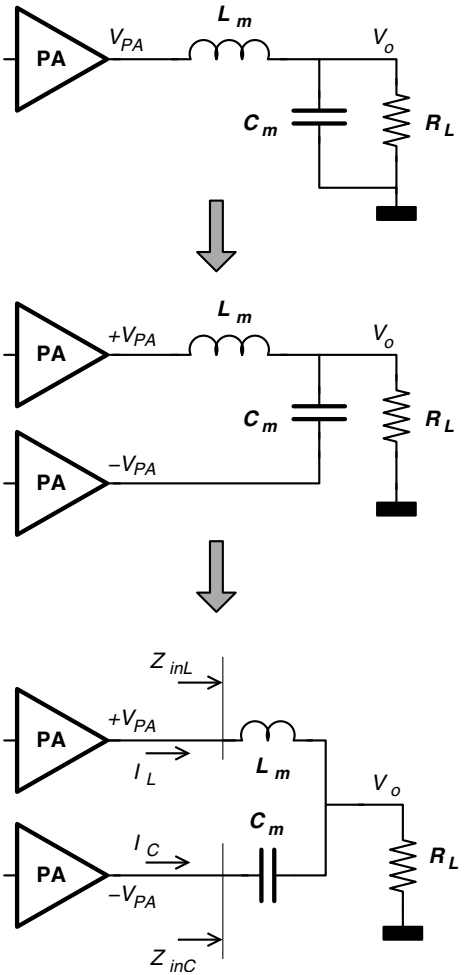


Figure 4.6. Evolution from an L-match network to a differential power combining and impedance transformation network with a single-ended output.

$$I_C = \frac{-V_{PA} - V_o}{1/j\omega C_m} \tag{4.34}$$

and the output voltage is obtained from Kirchoff's current law

$$V_o = (I_L + I_C) \cdot R_L \tag{4.35}$$

From these, the output voltage can be calculated as

$$V_o = -jV_{PA}R_L \cdot \frac{\omega C_m + 1/\omega L_m}{1 + jR_L\left(\omega C_m - 1/\omega L_m\right)} \tag{4.36}$$

The output voltage reaches a maximum at the resonance frequency $\omega_0 = 1/\sqrt{L_m C_m}$. At that specific frequency,

$$\omega_0 L_m = \frac{1}{\omega_0 C_m} = B \tag{4.37}$$

and thus

$$V_o = -2\,j\,V_{PA}\,R_L\sqrt{\frac{C_m}{L_m}} = -2\,j\,V_{PA}\,\frac{R_L}{B} \tag{4.38}$$

from which the voltage gain of the circuit is obtained.

$$\frac{V_o}{V_{PA}} = -2\,j\,\frac{R_L}{B} \tag{4.39}$$

The input impedances can be derived from equations 4.33, 4.34 and 4.38. Looking into the inductor, the input impedance at the resonance frequency is given by

$$Z_{inL} = \left(\frac{B^2}{2R_L}\right) \parallel (+jB) \tag{4.40}$$

and the input impedance when looking in the capacitor is given by

$$Z_{inC} = \left(\frac{B^2}{2R_L}\right) \parallel (-jB) \tag{4.41}$$

Besides the real or resistive part, the input impedances also have a complex part. This was not the case with the L-match transformation network. To make the input impedance purely resistive at the resonance frequency, the complex conjugate element is placed at the input of each port. This is demonstrated in figure 4.7 and the input impedances now become

$$R_{inL} = R_{inC} = R_{in} = \frac{R_L}{2}\cdot\left(\frac{B}{R_L}\right)^2 \tag{4.42}$$

By choosing B smaller than R_L, the input impedance can be made smaller than R_L and this results in an impedance transformation. Also notice that this circuit is able to create an input impedance that is larger than the load impedance. The impedance transformation ratio equals

$$r = \frac{R_L}{R_{in}} = 2\,\left(\frac{R_L}{B}\right)^2 \tag{4.43}$$

To obtain an expression for the power enhancement ratio, first the output power is calculated

$$P_o = \frac{V_o^2}{R_L} = 4\,\frac{V_{PA}^2}{R_L}\,\left(\frac{R_L}{B}\right)^2 \tag{4.44}$$

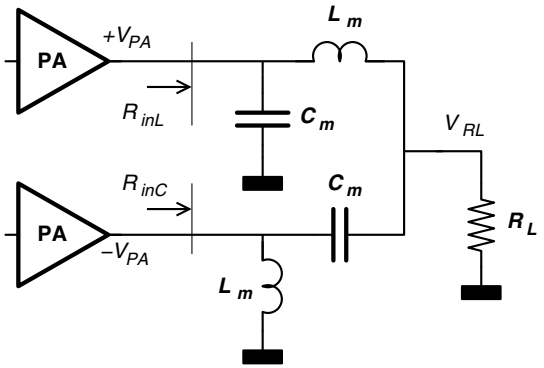


Figure 4.7. Lattice-type LC balun with a resistive input impedance.

As with the L-match network, the power enhancement ratio is calculated, based on the output power of a single power amplifier with a load resistor R_L. In that case

$$P_{o,0} = \frac{V_{PA}^2}{R_L} \tag{4.45}$$

and

$$E = \frac{P_o}{P_{o,0}} = 4\left(\frac{R_L}{B}\right)^2 \tag{4.46}$$

Compared to the L-match network, see equation 4.15, the lattice-type LC balun achieves a higher power enhancement ratio for a given value of $B = \omega L_m$. This suggests that the efficiency of the lattice-type LC balun will be higher for a given value of E. Also, the impedance transformation ratio is twice the value obtained for an L-match network. By connecting the second amplifier to the capacitor, the input impedance is thus divided by two, without changing the inductor and capacitor values. To employ this property even more, multiple sections can be connected in parallel to obtain an even higher power enhancement ratio. This will be discussed in section 4.3.3.

4.3.2 Inductor Loss and Efficiency

Like the L-match network, the lattice-type LC balun will also have some power loss. Again, the main loss is due to the finite quality factor of the inductor L_m which can be modeled by a series resistor R_{Lm} as indicated in figure 4.8. The dissipated power is equal to

$$P_{diss} = \left[R_{Lm} \cdot I_L^2 + R_{Lm}^* \cdot I_C^{*2}\right] \tag{4.47}$$

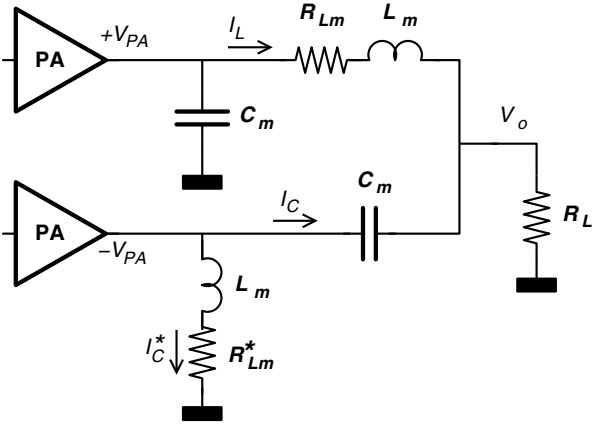


Figure 4.8. Lattice-type LC balun with power loss.

with

$$I_L = \frac{V_{PA} - V_o}{R_{Lm} + j\omega L_m} \tag{4.48}$$

$$I_C^* = \frac{-V_{PA}}{R_{Lm}^* + j\omega L_m^*} \tag{4.49}$$

The dissipation in R_{Lm}^* will be much smaller compared to the dissipation in R_{Lm}, because the voltage across the latter is larger. The output voltage of the network equals

$$V_o = V_{PA} \cdot \frac{R_L(1 + \omega^2 C_m L_m - j\omega C_m R_{Lm})}{R_L + R_{Lm} - \omega^2 R_L C_m L_m + j\omega R_L R_{Lm} C_m + j\omega L_m} \tag{4.50}$$

The resonance frequency will no longer be exactly equal to $1/\sqrt{L_m C_m}$, but the difference will be very small. Therefore, we will still assume that the maximal output voltage occurs at $\omega_0 = 1/\sqrt{L_m C_m}$. At that frequency, the output power, the efficiency and the power enhancement ratio can be calculated.

$$P_o = \frac{V_o^2}{R_L} \tag{4.51}$$

$$\eta = \frac{P_o}{P_o + P_{diss}} \tag{4.52}$$

$$E = \frac{P_o}{P_{o,0}} = \left(\frac{V_o}{V_{PA}}\right)^2 \tag{4.53}$$

One can now evaluate the performance of the lattice-type LC balun for several values of inductor quality factor. It was already pointed out that the lattice-type LC balun achieves a higher output power and power enhancement ratio

for the same inductor value, compared to the L-match network. Therefore, one can expect that the efficiency of the lattice-type LC balun will be higher. Of course, it is also possible to place two differential power amplifiers with two L-match networks on a single chip and combine their outputs with an external power combiner or balun. First of all, such a solution requires more board area because of the external combiner. Furthermore, the combiner has some power loss, given by the insertion loss IL_{dB} and this loss will add to the loss of the L-match network itself. Therefore, if two single-ended power amplifiers are combined by an off-chip power combiner, the output power becomes

$$P_o = 2\, V_{PA}^2\, \frac{R_m}{(R_m + R_{Lm})^2}\, \frac{1}{IL} \tag{4.54}$$

with

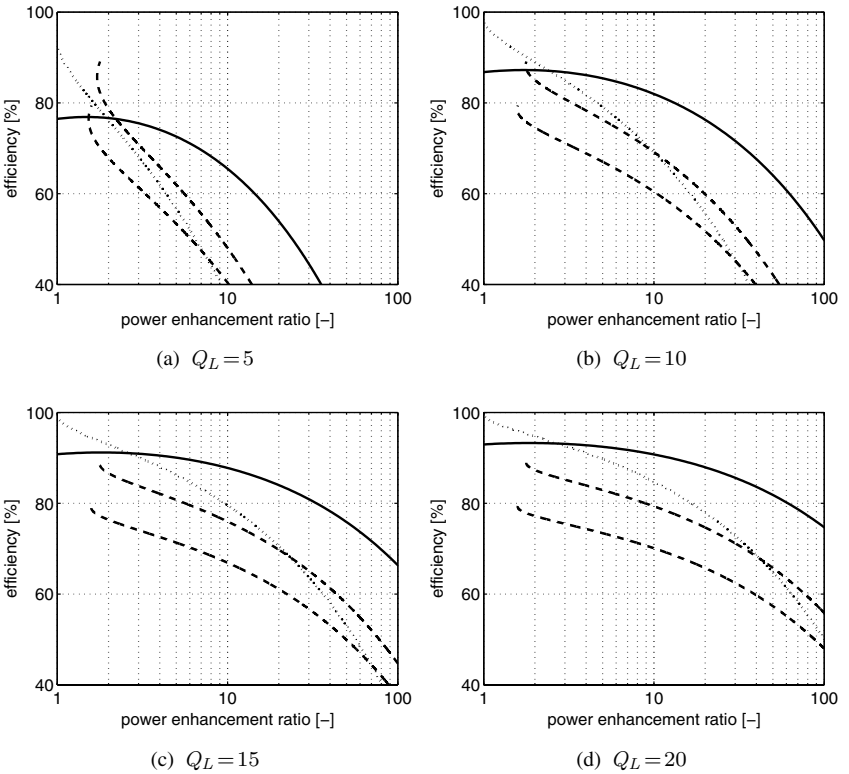
$$IL_{dB} = 10 \log_{10}(IL) \tag{4.55}$$

The efficiency of the impedance transformation now has to be multiplied by the efficiency of the external combiner, and the latter equals

$$\eta = \frac{1}{IL} \tag{4.56}$$

Typical values for insertion loss are between $0.5\,\mathrm{dB}$ and $1\,\mathrm{dB}$, which gives efficiencies from $80\,\%$ to $90\,\%$.

The comparison between the different solutions is shown in figure 4.9. In each plot, the comparison with the L-match network is made. The dotted line is the curve for one power amplifier with one L-match network. The two dashed lines represent the case were two differential amplifiers are combined with an external power combiner or balun. In the upper curve, an insertion loss of $0.5\,\mathrm{dB}$ and in the lower curve an insertion loss of $1\,\mathrm{dB}$ is chosen. Because of the loss in the external combiner, it is sometimes better to use only one amplifier with one L-match network, than a differential amplifier with an external combiner. However, one should realize that a differential amplifier requires less on-chip decoupling but on the other hand more silicon area. For low power enhancement ratios, one amplifier with one L-match network will always be the best. Of course, for $E = 1$, the L-match can be left out and the efficiency of the impedance transformation network goes to 100%. It should not be a surprise that the single L-match network outperforms the lattice-type LC balun for low power enhancement ratios. If a power enhancement ratio below two is required with the lattice-type LC balun, the impedance transformation must be done to an impedance that is larger than $50\,\Omega$. For higher power enhancement ratios, the lattice-type LC balun clearly outperforms the other solutions. Note that a high power enhancement is, to some extend, equivalent to a low supply voltage.



(a) $Q_L = 5$ (b) $Q_L = 10$

(c) $Q_L = 15$ (d) $Q_L = 20$

Figure 4.9. Efficiency versus power enhancement ratio. The solid line is the (integrated) lattice-type LC balun network, the dotted line is the L-match network with only one amplifier and the dashed lines are two differential amplifiers and two L-match networks, combined by an external power combiner with insertion loss. For the upper dashed line, an insertion loss of $0.5\,\mathrm{dB}$ and for the lower dashed line, an insertion loss of $1\,\mathrm{dB}$ is taken.

Finally, figure 4.9 should be handled with care as it shows only the efficiency of the transformation network. To obtain the efficiency of the entire power amplifier, one has to take the losses of the amplifier itself into account. In the previous chapter, it was demonstrated that the drain efficiency of the amplifier is strongly related to the equivalent load impedance that it must drive. This impedance will be higher for a differential amplifier that delivers the same output power as its single-ended counterpart. Therefore, despite the additional loss in the external differential power combiner, a differential solution might still be better. The verdict clearly depends on a lot of parameters and a general conclusion would be disrespectful to the complex tradeoff of integrated PA

design. However, the techniques presented here aim to identify these tradeoffs and to make a correct conclusion.

4.3.3 Multi Section Lattice-Type LC Balun

The lattice-type LC balun network of figure 4.7 thus allows to combine the outputs of two differential amplifiers to a single-ended output. Besides power combining, it also performs an impedance transformation. By selecting the proper values of L_m and C_m, the impedance seen by the power amplifier can be made larger or smaller.

The power enhancement ratio and impedance transformation ratio can also be increased by placing multiple sections in parallel [Paul03][1]. This idea is demonstrated in figure 4.10. With the assumption that each section is identical, the equations become

$$R_{inL} = R_{inC} = \frac{R_L}{2} \cdot \left(\frac{B}{R_L}\right)^2 \cdot \frac{1}{N} \tag{4.57}$$

$$V_o = -j\,2\,N \cdot V_{PA} \cdot \frac{R_L}{B} \tag{4.58}$$

$$P_o = 4\,N^2 \cdot \frac{V_{PA}^2}{R_L} \cdot \left(\frac{R_L}{B}\right)^2 \tag{4.59}$$

with N the number of differential sections. From the above equations, the impedance transformation ratio and power enhancement ratio can be obtained.

$$r = 2\,N \cdot \left(\frac{R_L}{B}\right)^2 \tag{4.60}$$

$$E = 4\,N^2 \cdot \left(\frac{R_L}{B}\right)^2 \tag{4.61}$$

Clearly, two means exist to increase the power enhancement ratio. Either the number of sections N or the ratio R_L/B can be increased and they both have a quadratic effect on P_o and on E. However, the effect on the impedance transformation ratio will be different; N has a linear effect and R_L/B has a quadratic effect.

From the previous section it was found that a large impedance transformation ratio will always lead to a low efficiency. Increasing the number of stages will quadratically increase the output power but the impedance transformation ratio will only increase linearly. Therefore, increasing the number of stages will be beneficial for the efficiency.

[1] to some extend, the series equivalent of this solution is the distributed transformer architecture of [Aoki02a]

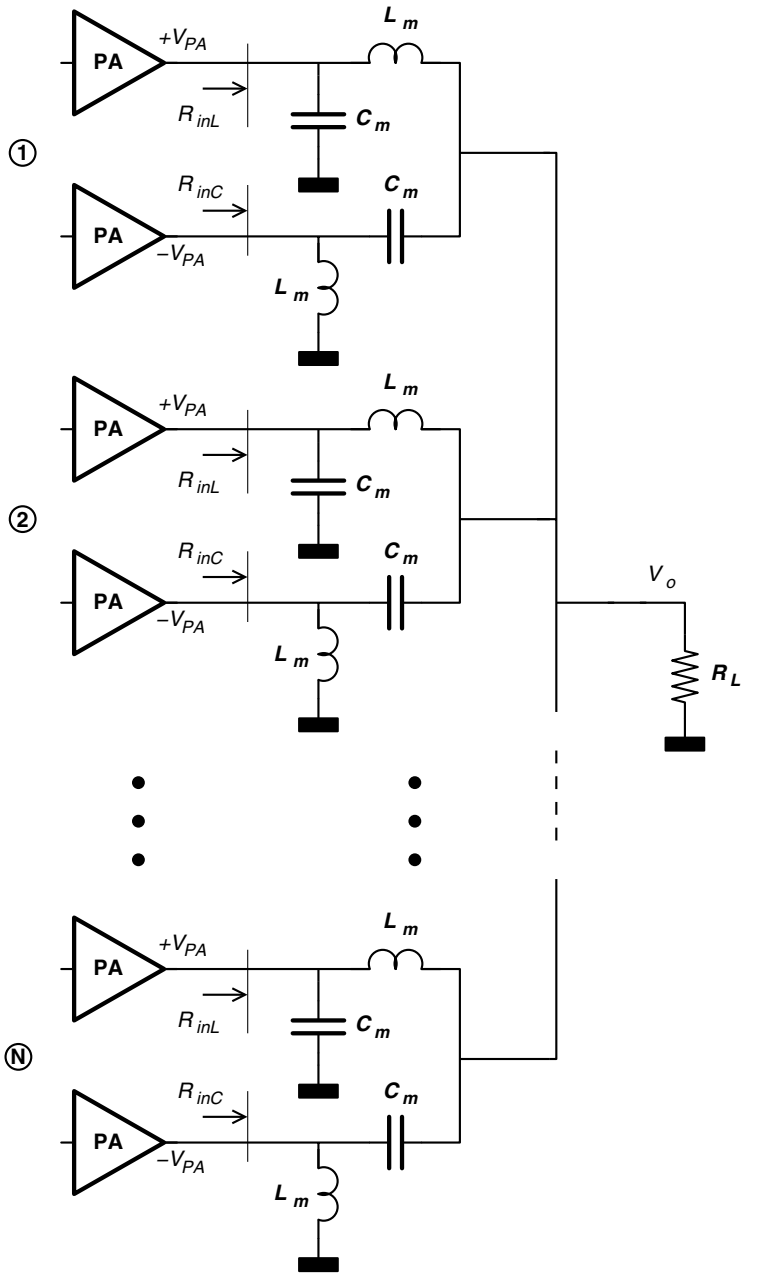
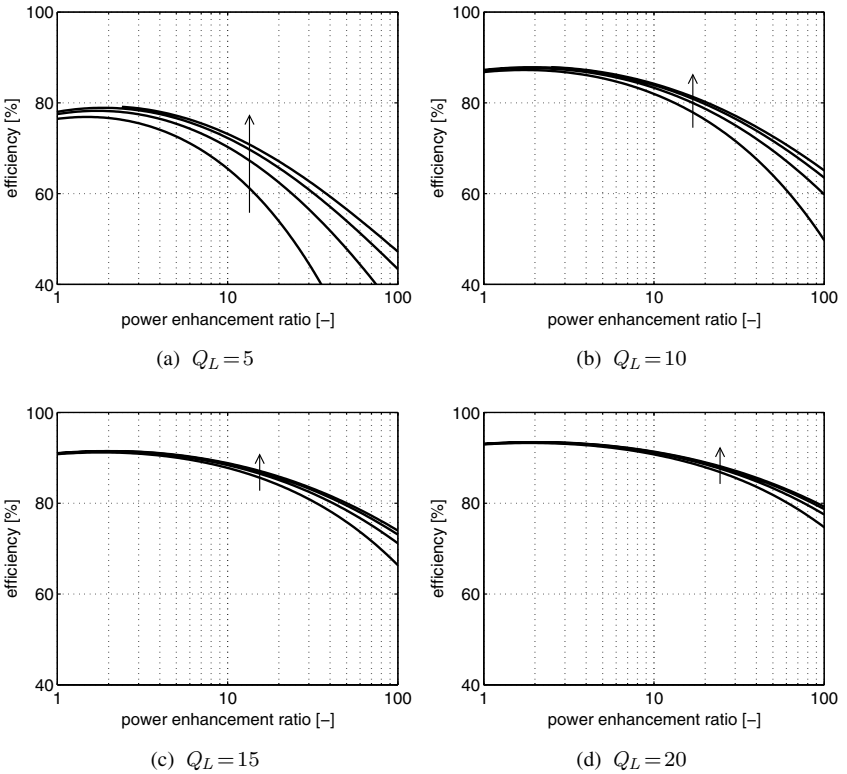


Figure 4.10. Lattice-type LC balun consisting of N sections.



(a) $Q_L = 5$ (b) $Q_L = 10$

(c) $Q_L = 15$ (d) $Q_L = 20$

Figure 4.11. Efficiency versus power enhancement ratio of the lattice-type LC balun for different inductor quality factors. In each plot, four different values of N (1, 2, 4 and 8) are used. The arrow indicates the direction of a higher N.

Figure 4.11 depicts the efficiency versus power enhancement ratio for the multi section lumped element balun network. Clearly, for a given power enhancement factor and inductor quality factor, the efficiency always increases when more sections are placed in parallel. On the other hand, for a reasonable inductor quality factor between 10 and 15, the highest improvement is achieved by placing two sections in parallel. Two sections already means four amplifiers and one should realize that more sections also means more silicon area and thus a higher cost. Therefore, two sections seems the best compromise to combine both high efficiency and high output power.

4.3.4 Power Control

If two differential power amplifiers of a specific section are not operational and their outputs behave as an AC ground, the respective section behaves as a high-impedance parallel LC tank, connected with the output. This idea is

Table 4.1. Comparison between an L-match network and a lattice-type LC balun network with N sections.

	L-match	power combining network
R_{in}	$\approx \frac{B^2}{R_L}$	$\frac{1}{2\,N} \cdot \frac{B^2}{R_L}$
r	$\approx \left(\frac{R_L}{B}\right)^2$	$2N \left(\frac{R_L}{B}\right)^2$
E	$\approx \left(\frac{R_L}{B}\right)^2$	$4\,N^2 \cdot \left(\frac{R_L}{B}\right)^2$

depicted in figure 4.12, in which section 2 is assumed to be inactive. The corresponding section will not contribute to the total output power, N is reduced to $N-1$ and because of the high-impedance nature, the corresponding output network will, in the ideal case, not influence the operation of the other amplifiers.

The multi section lattice-type LC balun can thus be used to implement a discrete form of *power control*. In other words the output power can easily be controlled by turning on and off different sections. The power control step depends on the total number of stages of the entire structure.

The output voltage is proportional to the number of active stages N but the input impedance of each section of the power combining network is proportional to $1/N$. Therefore, changing the number of active stages will also change the load impedance seen by all other amplifiers. This is in fact a discrete form of an active load-pull technique, as used in a Doherty Amplifier [Crip99, Dohe36].

The power control technique can be employed in two means; either as a means to efficiently control the output power, discussed in section 5.4, or as a digital linearization technique. The latter is discussed in section 5.5.

The power control step depends on the number of parallel sections. From the previous section it was found that $N = 2$ is sufficient for a good tradeoff between efficiency and output power. If the power combining network is used as a means of power control or as digital linearization, more sections can be used to increase the accuracy.

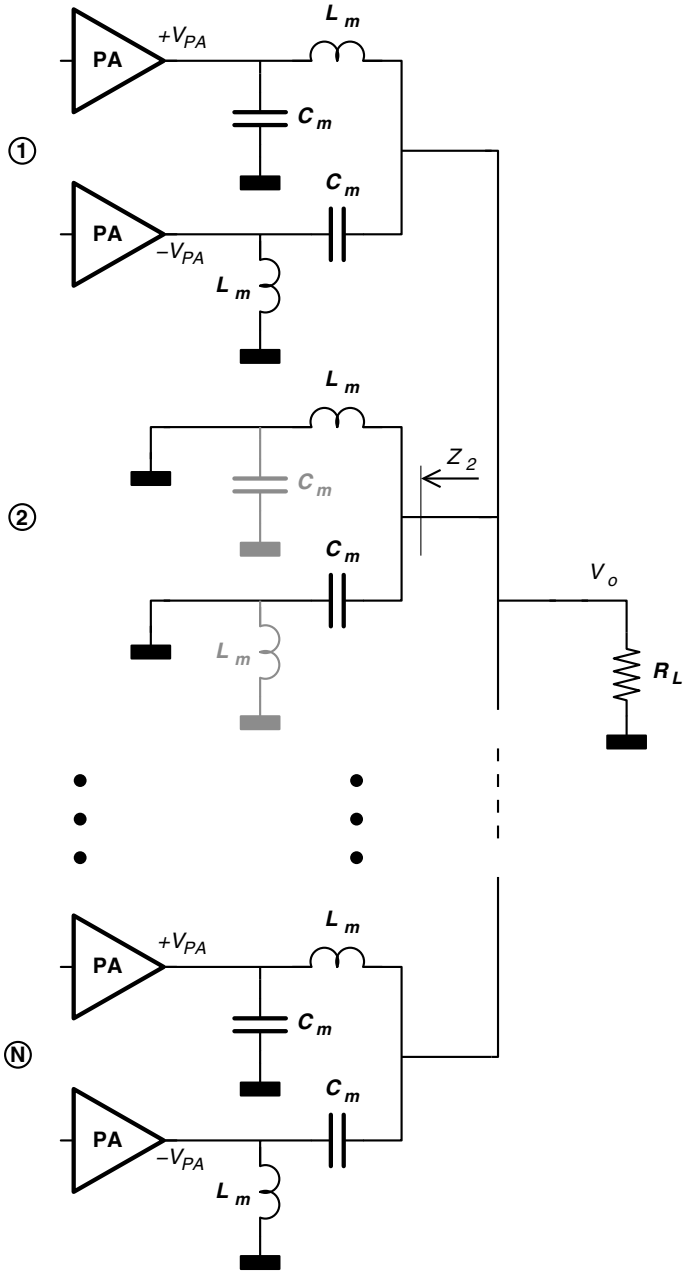


Figure 4.12. Differential to single-ended impedance transformation and power combining network, consisting of N sections. The second section is not active and creates a high-impedance LC-tank.

4.3.5 Multi Section LC Balun with Non-Identical Sections

Till now, it is assumed that all sections are identical. This is the best approach to distribute the total required output power amongst all sections.

If it is the goal to control the output power over a wide range, it might be better to implement different sections, each one with a specific impedance transformation ratio. In other words, each section will have a different value of B. To analyze this, the formulas are first rewritten in a more general way for section i.

$$R_{inL,i} = R_{inC,i} = \frac{1}{2} \cdot \frac{B_i}{R_L} \cdot \frac{1}{\left(\sum_{k=1}^{N} \frac{1}{B_k} \right)} \tag{4.62}$$

$$V_o = -j\,2 \cdot V_{PA} \cdot R_L \cdot \sum_{k=1}^{N} \frac{1}{B_k} \tag{4.63}$$

$$P_o = 4 \cdot \frac{V_{PA}^2}{R_L} \cdot \left(\sum_{k=1}^{N} \frac{R_L}{B_k} \right)^2 \tag{4.64}$$

To have the highest accuracy with the least number of stages, a binary scaling can be employed. In that case,

$$\begin{aligned} R_L/B_1 &= \beta \\ R_L/B_2 &= 2\beta \\ R_L/B_4 &= 4\beta \end{aligned}$$

By turning on and off the sections, the output voltage will behave in a binary manner. Therefore, such a partition can also be interpreted as an RF D-to-A power converter.

A point of attention is the transformed load impedance. If only the first stage is working, the input impedance, and thus the load impedance of the power amplifier, is equal to

$$R_{in,1} = \frac{1}{2} \cdot B_1 \cdot \frac{1}{\beta} \tag{4.65}$$

If all stages are working, that impedance becomes

$$R_{in,1} = \frac{1}{2} \cdot B_1 \cdot \frac{1}{7\beta} \tag{4.66}$$

in other words, seven times as small. Clearly, the power amplifier has to be able to deal with this small load impedance.

4.3.6 Merging the Class E Amplifier and the LC Balun

The very basic Class E amplifier consists of an nMOS switch with a drain-capacitance to ground and a DC-feed inductance. If this type of amplifier is used in conjunction with a lattice-type LC balun network, two elements can be merged together as demonstrated by figure 4.13. The parallel capacitor C_m, necessary to create a pure resistive input impedance, is added to the Class E drain capacitance. The parallel inductor L_m, normally connected to the ground, can be placed in parallel with the DC-feed inductance L_1. After all, the supply voltage can be considered as an AC ground, if sufficient decoupling is provided. This reduces the total number of inductors and thus, the silicon area and cost is also reduced. In fact, this network requires less inductors than its differential L-match implementation.

4.4 Conclusion

A transformation of the $50\,\Omega$ load impedance to a smaller value is required to achieve sufficient output power in a low-voltage CMOS technology. This topic of impedance transformation has been discussed in this chapter.

The L-match network excels in simplicity and is easy to integrate in a CMOS technology. Of course, the network does not work at a $100\,\%$ efficiency and less power than intended will reach the output. A high impedance transformation ratio is needed for a high output power, but this will be accompanied by a low efficiency of the transformation network alone, which restricts the applicability of the L-match. Furthermore, most integrated amplifiers operate in a differential manner to decrease the required on-chip decoupling and to increase the output power. Unless a differential antenna is available, a BalUn or RF-transformer is required.

The lattice-type LC balun network was presented in section 4.3 and is able to convert the differential output signal of the amplifier to a single-ended signal across the load. As such, no external balun or RF-transformer is required. The LC balun is also able to perform the impedance transformation function, needed to increase the output power. In addition, the LC balun also allows that multiple sections are connected in parallel, and this is a second means to increase the output power. Like the L-match network, the lattice-type LC balun will also exhibit power loss. However, the performance of the latter is superior, especially if the number of parallel sections is increased. Besides the benefit of a higher efficiency, the LC balun can also be merged with the Class E amplifier, reducing the total number of inductors and the silicon area.

Another advantage of the lattice-type LC balun network, is that one or more sections can easily be turned off. This will reduce the output power and thus allows to efficiently implement some form of output power control or digital linearization. To further exploit this property, the different parallel sections

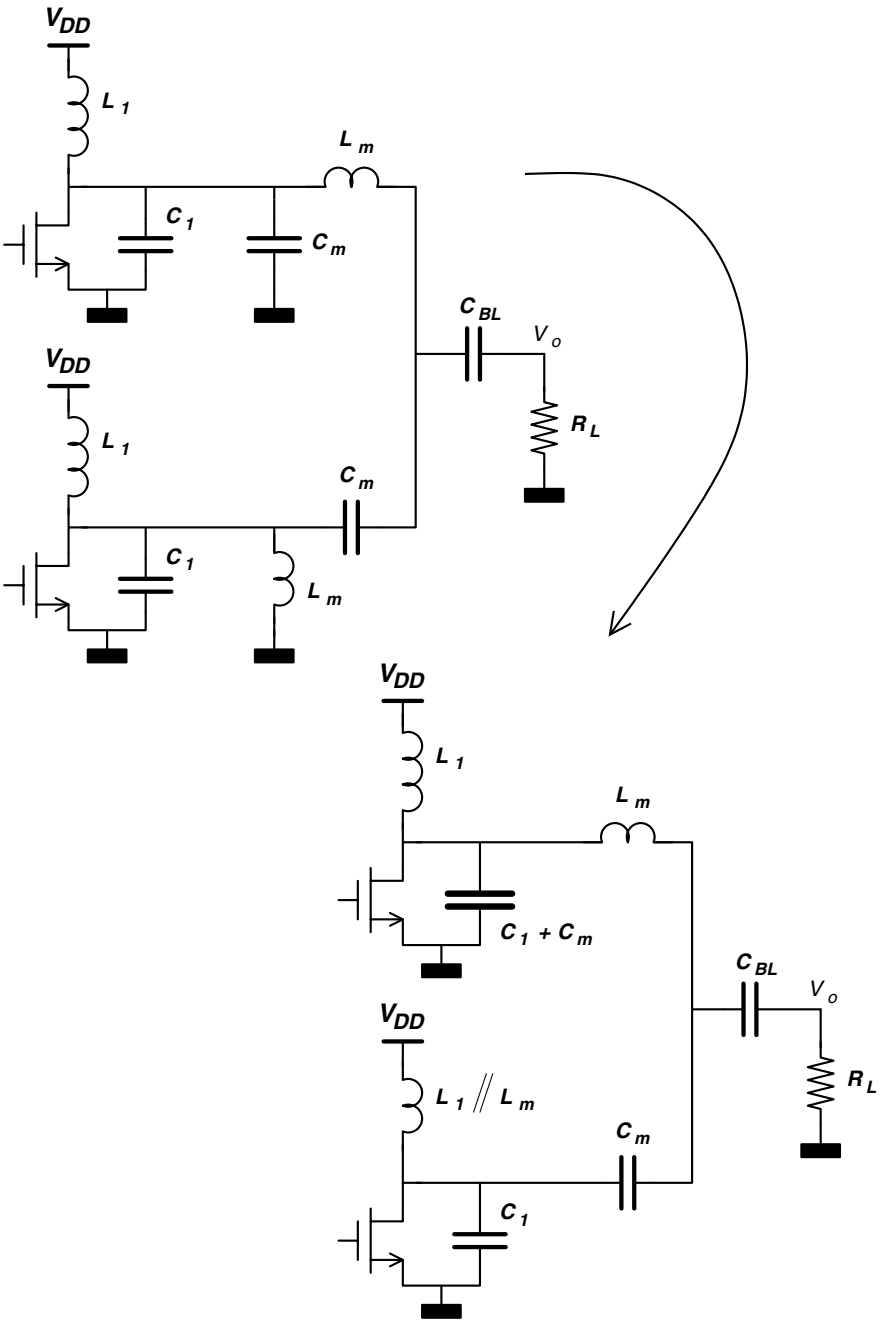


Figure 4.13. The differential to single-ended impedance transformation and power combining network can easily be merged with a differential Class E power amplifier.

can be designed with a different impedance transformation ratio. If a binary partitioning is chosen, one could think of this as an RF D-to-A power converter, enabling the way towards full digital linearization techniques.

Chapter 5

POLAR MODULATION

5.1 Introduction

In chapter 2, polar modulation was introduced to linearize the efficient Class E amplifier. The major benefits of this appraoch is that the RF path can be nonlinear, the linearity is shifted to a low frequency and a high integration level can be achieved. This chapter will give a profound analysis of the polar modulation architecture.

Section 5.2 will introduce some general principles and ideas on polar modulation as well as some architectural aspects and tradeoffs. In section 5.3, the nonlinearities of polar modulation are discussed and some linearity improvement techniques are discussed. An efficiency improvement technique for polar modulated power amplifiers is presented in section 5.4. Finally, techniques that expand the polar modulation technique towards full digital linearization are discussed in section 5.5.

5.2 The Polar Modulation Architecture

5.2.1 Basic Equations

The Class E amplifier can only amplify phase modulated signals, and the envelope of the RF output signal is proportional to the value of the DC supply voltage V_{DD}. The output voltage of a Class E amplifier can thus be written as

$$v(t) = \gamma \cdot V_{DD} \cdot \cos(\omega_c t + P(t)) \tag{5.1}$$

in which the phase modulation is represented by $P(t)$ and γ corresponds to the ratio of the output envelope to the supply voltage.

The next step is to modulate the value of the supply voltage V_{DD}. In this regard, the Class E amplifier acts as a double sideband mixer. Assume that the

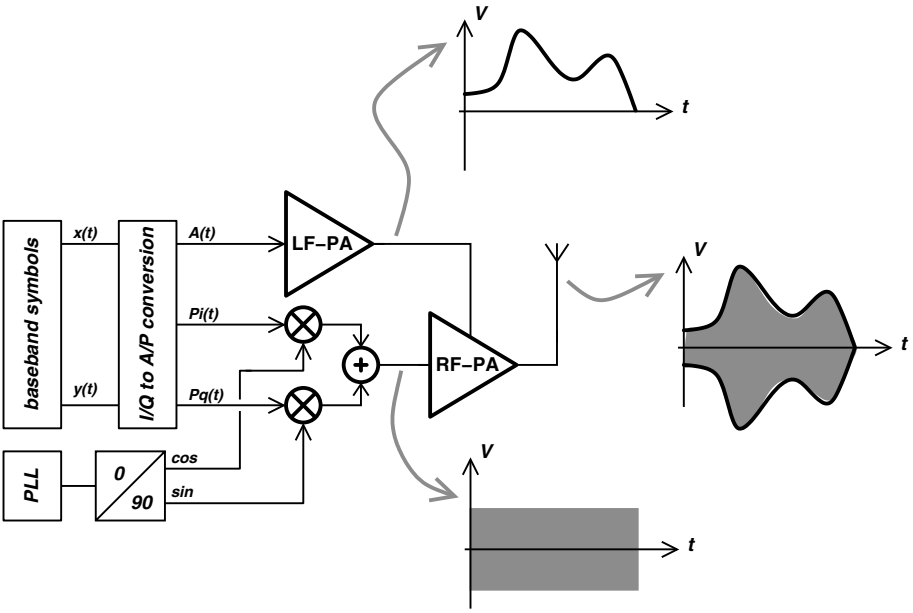


Figure 5.1. Linearized transmitter based on polar modulation of a nonlinear RF amplifier.

supply voltage can be written as

$$V_{DD}(t) = \nu \cdot A(t) \tag{5.2}$$

The *mixing* operation can now explicitly be expressed as

$$\begin{aligned} v(t) &= \nu \cdot \gamma \cdot A(t) \cdot \cos\left(\omega_c t + P(t)\right) \\ &\sim A(t) \cdot \cos\left(\omega_c t + P(t)\right) \end{aligned} \tag{5.3}$$

Equation 5.3 is a polar representation of a bandpass modulated signal, see section 2.2.1, which clarifies the names *polar modulation*, *polar linearization* and *polar modulated power amplifier*. The phase modulation is contained in $P(t)$ and $A(t)$ represents the amplitude modulation. It may be useful to recall that

$$v(t) = A(t) \cdot \cos(\omega_c t + P(t)) \tag{5.4}$$

is denoted as the *modulated RF signal* and contains both AM and PM. The amplitude modulation is present in $A(t)$, which is denoted as the *amplitude* or *envelope signal*. The phase modulation is present in $P(t)$ or $\theta(t)$, and denoted as the *phase signal*. The signal

$$v_P(t) = \cos(\omega_c t + P(t)) \tag{5.5}$$

Table 5.1. Nomenclature of modulated signals.

symbol	name
$v(t)$	modulated RF signal
$A(t)$	envelope signal, amplitude signal
$\theta(t), P(t)$	phase signal
$e^{j\theta(t)}$, $e^{jP(t)}$	complex phase signal
$v_P(t) = \cos(\omega_c t + P(t))$	RF phase signal

is an RF signal that is only phase modulated by $P(t)$, and this RF signal has a constant envelope. It is denoted as the *RF phase signal*.

The system level implementation of polar modulation is depicted in figure 5.1. At baseband, the in-phase signal $x(t)$ or $I(t)$ and quadrature signal $y(t)$ or $Q(t)$ are generated. First, these signals are converted to an amplitude signal $A(t)$ and a phase signal $P(t)$ or $\theta(t)$. The phase signal is also split in an in-phase component $Pi(t)$ and quadrature component $Pq(t)$ in order to have a correct upconversion. The amplitude signal is amplified by a linear low frequency power amplifier (LF-PA) and the constant envelope phase modulated signal is amplified by the nonlinear RF Class E amplifier (RF-PA). It should be noted that other high efficiency amplifiers can be used instead of a Class E, as long as the output envelope is proportional to the supply voltage. The low frequency linear power amplifier is also denoted as an *amplitude modulator* or *supply voltage modulator*, and should have sufficient linearity. However, since it operates at a lower frequency, linearity is easier to achieve. The Class E amplifier is still operating as an efficient nonlinear amplifier, and the amplitude linearity requirement is completely shifted to a baseband block.

Figure 5.1 is an indication of the trend in CMOS RF power amplifiers. In CMOS, digital signal processing is widely available at a relative low cost and low power consumption. This availability makes a polar linearization scheme feasible. It also indicates that the amplifier can no longer be regarded as *everything after the up-converter*. Indeed, if one compares figure 5.1 with figure 2.2, it suggests that the entire transmit architecture has to be adopted in order to get either the best performance or to achieve the highest level of integration. A similar trend is visible in the shift from the heterodyne receiver architecture to the low- and zero-IF receivers which also enables a higher integration level in a CMOS technology [Stey98], [Stey00], [Reyn04b].

5.2.2 Envelope Elimination and Restoration

The idea of polar modulation originates from the *Envelope Elimination and Restoration, (EER)* technique, first proposed by Kahn in 1952 [Kahn52]. Hence,

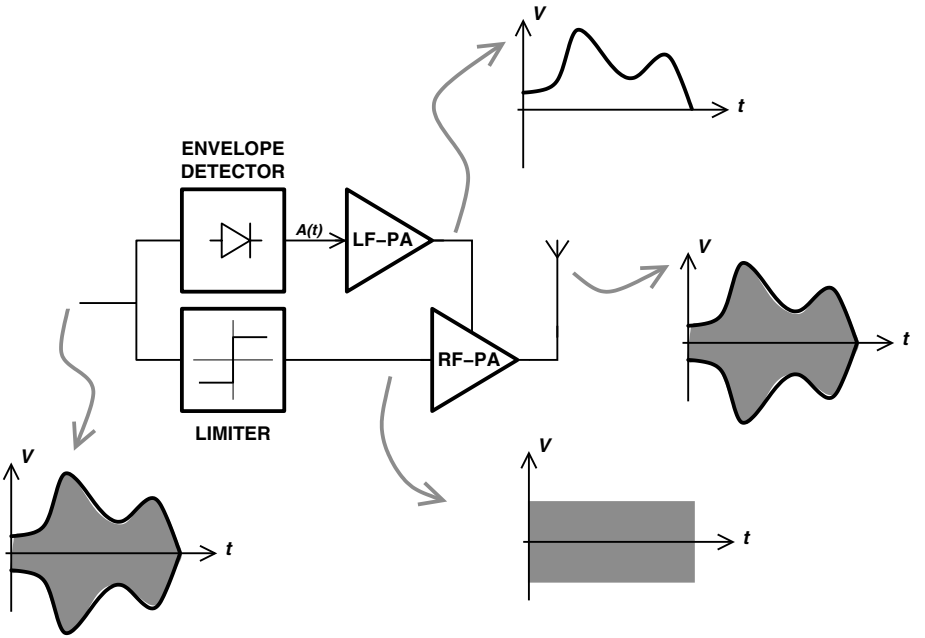


Figure 5.2. Linearized amplifier, based on Envelope Elimination and Restoration or Kahn Technique Transmitter.

it is sometimes referred to as the Kahn Technique Transmitter. In such a linearized amplifier, a signal that contains both amplitude and phase modulation is converted by a limiter into a constant envelope phase modulated signal, hence the name *envelope elimination*. Likewise, an envelope detector extracts the low frequency envelope signal. Both signals are combined, and thus the amplitude modulation is *restored*, in the last stage of a nonlinear power amplifier, as demonstrated by figure 5.2.

The Kahn technique transmitter assumes that the linear or linearized power amplifier is a separate block, not integrated with the upconversion mixers. Of course, in a fully integrated transmitter, there is no need to first create the RF signal, splitting it up into a phase and envelope signal, and then recombining it in the power amplifier. Instead, the digital signal processing block can deliver the amplitude and phase signal separately, which facilitates the design.

5.2.3 Influence of the Driver Stages on the Overall Efficiency

The drain efficiency of the Class E amplifier remains constant if the supply voltage is modified. As such, the drain efficiency versus output power curve is a horizontal line and linearity is then combined with efficiency. However, the overall efficiency, defined as

$$\eta_{oa} = \frac{P_o}{P_{DC} + \sum_{i=0}^{n} P_{DRV,i}} \tag{5.6}$$

will not be constant when the supply voltage of the Class E amplifier is varied. Indeed, the overall efficiency can be rewritten as

$$\eta_{oa} = \eta_d \cdot \frac{P_{DC,PA}}{P_{DC,PA} + \sum_{i=0}^{n} P_{DC,DRV,i}} \tag{5.7}$$

When the supply voltage of the Class E amplifier is reduced, the DC power consumption of the Class E amplifier also reduces, but the dissipation of the driver stages is not affected and remains constant. Therefore, for a lower output power, the overall efficiency will be reduced due to the power consumption of the driver stages. As such, the overall efficiency becomes dependent on the envelope signal.

Assume that the total power consumption of the driver stages can be expressed as a fraction of the peak DC power consumed by the Class E amplifier. In that case

$$\sum_{i=0}^{n} P_{DC,DRV,i} = \delta P_{DC,PA,max} \tag{5.8}$$

and the overall efficiency can be writtten as

$$\eta_{oa}(A) = \eta_d \cdot \frac{1}{1 + \delta \left(\frac{A_{max}}{A}\right)^2} \tag{5.9}$$

Figure 5.3 gives the overall efficiency of an ideal Class E amplifier, with a drain efficiency of 100%, for several values of δ. Clearly, dispite the fact that the Class E amplifier is ideal, the overall efficiency of the transmitter is not 100%.

However, it should be clear to the reader that the overall transmitter efficiency can never reach 100%, even if the power amplifier and the driver stages all have drain efficiencies of 100%. The power consumed by the driver stages is always lost, since this power does not flow towards the output.

The obvious solution is to reduce the power consumption of the driver stages at lower output power. For a Class E amplifier, the driver stages will consist of

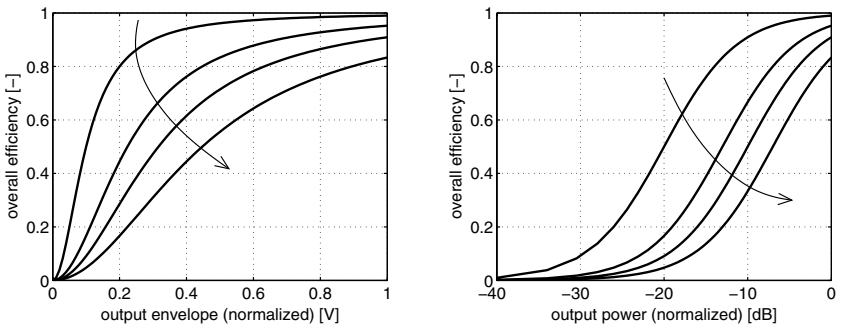


Figure 5.3. Overall efficiency of the Class E amplifier (a) versus relative output envelope and (b) versus relative output power, for $\delta = 0.01, 0.05, 0.1$ and 0.2. The arrows indicate the direction of a larger δ.

nonlinear switching or saturated amplifiers with a high drain efficiency. Thus, one might think that the power consumption of the drivers can be reduced by simply reducing their supply voltage. However, this can not be done since the drivers must still switch the transistor of the Class E amplifier completely on and off. If that transistor is not completely switched on, the drain efficiency of the Class E amplifier would quickly start to degrade, counteracting the initial goal. Furthermore, the amplifier will no longer work as a Class E amplifier, and the output envelope will no longer be dependent on the supply voltage of the Class E amplifier alone, but also on the supply voltage of the driver stages. This will degrade the linearity of the Class E amplifier. A better way to reduce the consumption of the driver stages will be presented in section 5.4.

5.2.4 Implementation of the Amplitude Modulator

The linearity of the polar modulated Class E amplifier will, to a great extend, depend on the linearity of the amplitude modulator or low frequency power amplifier, indicated as LF-PA in figure 5.1. Since it operates at a lower frequency, linearity is easier to achieve with less power consumption, compared to a linear RF amplifier.

Besides linearity, the amplitude modulator should also have a high efficiency. After all, the power that is wasted in the amplitude modulator will degrade the efficiency of overall linearized amplifier. A such, yet another definition of *efficiency* appears. Based on figure 5.4, following definitions can be

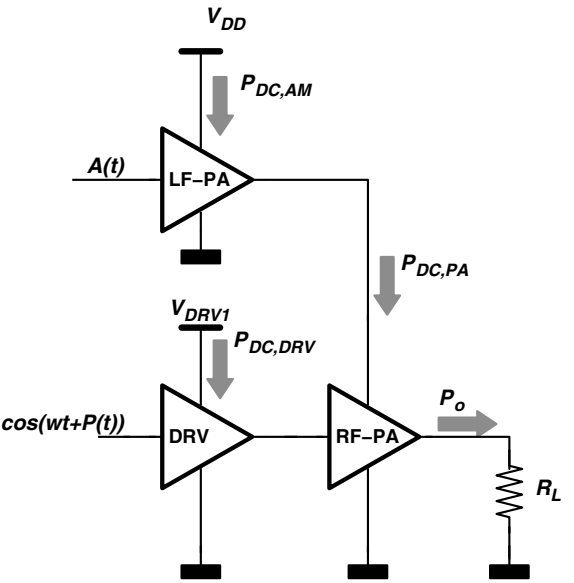


Figure 5.4. Power consumption in a polar modulated power amplifier.

made. The drain efficiency of the Class E amplifier remains equal to

$$\eta_{d,E} = \frac{P_o}{P_{DC,PA}} \tag{5.10}$$

and this efficiency becomes 100% in an ideal Class E amplifier. If the power consumption of the RF drivers is included, the overall efficiency can be defined as

$$\eta_{oa,E} = \frac{P_o}{P_{DC,PA} + P_{DC,DRV}} \tag{5.11}$$

As stipulated before, the overall efficiency can never reach 100% since the power consumed by the driver stages does not flow to the output. One can also define the overall efficiency of the amplitude modulator as

$$\eta_{oa,AM} = \frac{P_{DC,PA}}{P_{DC,AM}} \tag{5.12}$$

Finally, the overall efficiency of the polar modulated amplifier (PMA) is defined as

$$\eta_{oa,PMA} = \frac{P_o}{P_{DC,AM} + P_{DC,DRV}} \tag{5.13}$$

$$= \eta_{d,E} \cdot \eta_{oa,AM} \cdot \frac{P_{DC,AM}}{P_{DC,AM} + P_{DC,DRV}} \tag{5.14}$$

Clearly, to achieve a high overall efficiency of the entire polar modulated amplifier, the overall efficiency of the amplitude modulator has to be high as well. Again, as in section 2.3.4, one could argue whether the power dissipation of the driver stages and the power consumption of the additional circuitry in the amplitude modulator have to be included in the calculation of the overall efficiency.

The overall efficiency of the amplitude modulator ($\eta_{oa,AM}$) will, in general, depend on the power that it has to deliver. Two types can be distinguished and will be discussed; the linear regulator and the switching regulator.

5.2.4.1 The Class E power amplifier as a load

In a polar modulation scheme, the Class E amplifier actually becomes the load of the amplitude modulator. In order to properly design the latter one, a simplified representation of the Class E amplifier as a load is required. However, the Class E amplifier is a highly nonlinear switching amplifier, that cannot be represented by a small-signal equivalent.

An important difference between the Class E amplifier and the amplitude modulator is the large difference between the time constants of the two. In general, one could say that the RF amplifier operates in the gigahertz range and the amplitude modulator operates in the megahertz range. Therefore, the supply voltage of the Class E amplifier stays fairly constant within one RF period, which is denoted as a *quasi-static* approximation. As such, the Class E amplifier can be represented by its equivalent DC load resistance, which can be obtained from a transient analysis.

5.2.4.2 Linear Amplitude Modulator

The basic circuit of a linear amplitude modulator is depicted in figure 5.5. Although the pMOS transistor has a relative high output impedance, the voltage feedback creates a low impedance at the output node and as such, the circuit acts as a voltage source. If the gain of the opamp is sufficiently high, the transfert function between the input voltage of the amplitude modulator and the output, the voltage $V_{DD,PA}$, becomes $1/H(j\omega)$. The output impedance of the amplitude modulator is equal to

$$Z_{out,AM} = \frac{r_o}{1 + T(j\omega)} \tag{5.15}$$

with r_o the output impedance of the pMOST and $T(j\omega)$ the loopgain of the amplitude modulator, which can be written as

$$T(j\omega) = G(j\omega) \cdot H(j\omega) \cdot g_m r_o \tag{5.16}$$

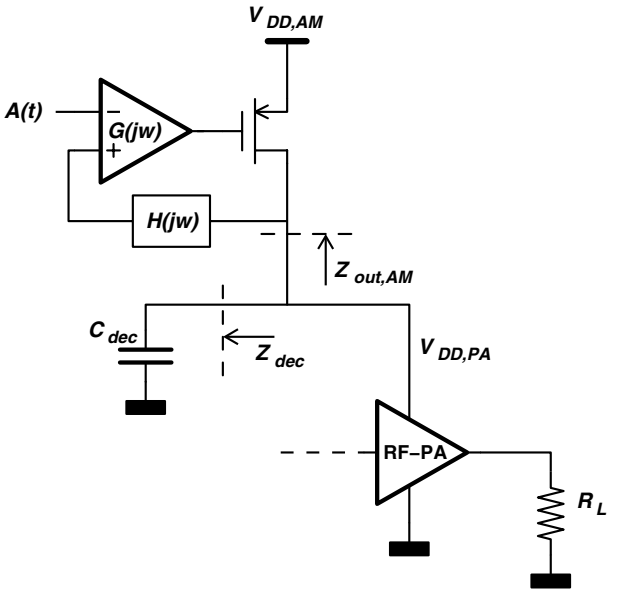


Figure 5.5. Basic diagram of a linear amplitude modulator.

If one assumes that the loopgain has a simple first order frequency response, given by

$$T(j\omega) = g_m r_o \cdot G_0 H_0 \cdot \frac{1}{1 + j\omega/\omega_d} = T_0 \cdot \frac{1}{1 + j\omega/\omega_d} \tag{5.17}$$

the output impedance can be calculated as

$$Z_{out,AM} \approx \frac{1}{G_0 H_0 g_m} \cdot \frac{1 + j\omega/\omega_d}{1 + j\frac{\omega}{\omega_d(1+T(0))}} \tag{5.18}$$

At DC and low frequencies, the output impedance can be approximated by

$$Z_{out,AM} \approx \frac{1}{G_0 H_0 g_m} \tag{5.19}$$

At higher frequencies, the output impedance has an inductive behavior between ω_d and $\omega_d(1+T(0))$, and for frequencies higher than $\omega_d(1+T(0))$, the output impedance is equal to r_o. Therefore, at the operating frequency of the Class E amplifier, the high output impedance of the amplitude modulator has to be compensated by sufficient decoupling capacitance, which has to be placed on chip since it has to operate at RF frequencies. The decoupling capacitance can interact with the inductive output impedance and the total impedance of the

amplitude modulator ($Z_{out,tot}$) might have a resonance peak which is harmfull for the stability of the system. However, the power amplifier will act as a low resistance in parallel with the decoupling capacitance and as such it will strongly attenuate this resonance behavior.

A major difficulty in the design of a linear amplitude modulator is the large current swing through the pMOST in figure 5.5. This will inevitably change the transconductance and output resistance of the transistor and this will impede the design of the feedback loop, as this loop has to be stable over the entire dynamic range of the envelope signal. On the other hand, the RF power amplifier resembles a low ohmic load, and this will shift the pole on the output node of the amplitude modulator to higher frequencies. As such, the dominant and second pole of the open loop system can easily be located within the operational amplifier and stable operation can be guaranteed, despite the large change in gain and output impedance of the pMOS transistor.

The efficiency of the linear amplitude modulator is proportional to its output voltage $V_{DD,PA}$ and can be written as

$$\eta_{AM} = \frac{V_{DD,PA}}{V_{DD,AM}} \tag{5.20}$$

in which the power consumption of the opamp is neglected. Furthermore, since the envelope of the RF signal at the output of the RF amplifier is proportional to the modulated supply voltage, the efficiency of the linear amplitude modulator can be written as a function of the envelope signal.

$$\eta_{AM}(A) = \frac{A}{A_{max}} \tag{5.21}$$

Together with equation 5.14, the overall efficiency of the polar modulated power amplifier with a linear amplitude modulator can now be obtained and is equal to

$$\eta_{AM}(A) = \eta_{d,E} \cdot \frac{A}{A_{max}} \cdot \frac{1}{1 + \delta \frac{Amax}{A}} \tag{5.22}$$

The efficiency of the amplitude modulator, the overall efficiency of the Class E amplifier and the efficiency of the polar modulated amplifier are depicted in figure 5.6 for an ideal Class E amplifier with 100% drain efficiency.

One could argue that the linear amplitude modulator has a rather poor efficiency versus output envelope curve. After all, the polar modulated amplifier will have a linear drain efficiency curve, which is the same as if an RF Class B amplifier was used. However, the theoretical peak efficiency of a Class B only reaches 78.5%, and the driver stages should be linear as well, which increases the power consumption of the driver stages. Furthermore, the linearity of a Class B is not enough for most wireless standards, and thus the RF amplifier

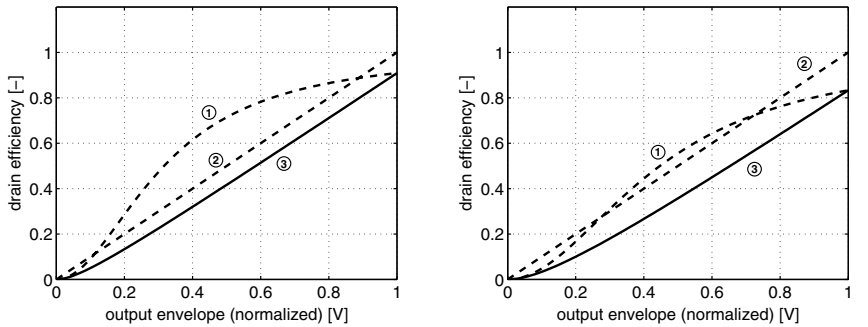


Figure 5.6. Overall efficiency of (curve 1) the Class E amplifier, (curve 2) the linear amplitude modulator and (curve 3) the entire amplitude modulated RF power amplifier, for (a) $\delta = 0.1$ and (b) $\delta = 0.2$, versus normalized output envelope.

needs to be biased closer to Class AB, with an even lower efficiency and a *more quadratic* efficiency curve.

On the other hand, in a polar modulated amplifier the linearity is shifted to baseband frequencies and the entire RF path can operate as highly efficient and nonlinear transmitter. Furthermore, if one looks to the overall efficiency of the polar modulated amplifier, the power dissipation of the driver stages already *pulls down* the efficiency versus output envelope curve. This effect can not be compensated for, even if the driver stages operate at 100% drain efficiency. Hence, the drawback of the linear efficiency curve of the amplitude modulator is less apparent in the overall efficiency curve of the polar modulated amplifier. A final argument to select the linear amplitude modulator, and maybe the most important one from a cost perspective, is the high integration level that can be achieved with it. Compared to a switching amplitude modulator, which will be discussed in paragraph 5.2.4.4, the linear amplitude modulator can be fully integrated on a single chip and requires no additional off-chip components. This will be demonstrated in chapter 6.

5.2.4.3 Linear Amplitude Modulator with External Resistors

The efficiency of the linear amplitude modulator is proportional to the voltage that it supplies to the RF power amplifier.

$$\eta_{AM} = \frac{V_{DD,PA}}{V_{DD,AM}} \tag{5.23}$$

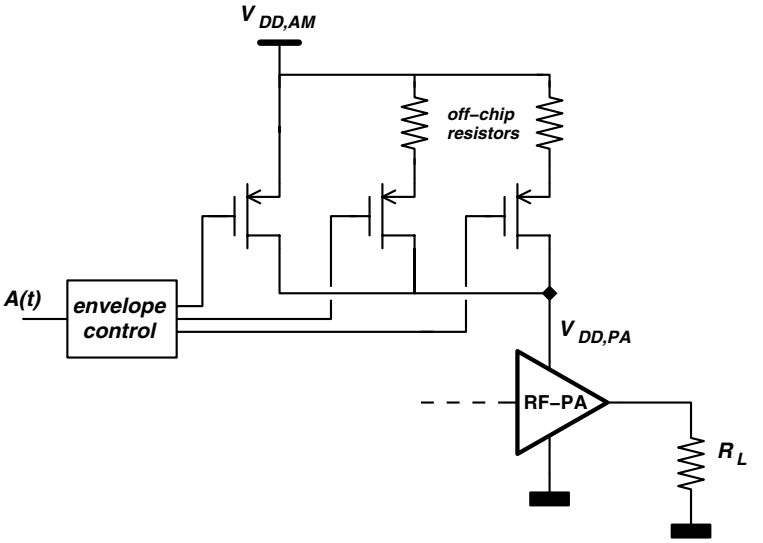


Figure 5.7. Linear amplitude modulator with external (off-chip) resistors.

The power dissipated in the pMOST of figure 5.5 is equal to

$$P_{diss} = \frac{V_{DD,PA}^2}{R_{L,PA}} \cdot \left(\frac{V_{DD,AM} - V_{DD,PA}}{V_{DD,PA}} \right) \tag{5.24}$$

and when normalized as $V_{DD,AM} = 1V$ and $R_{L,PA} = 1\Omega$ this becomes

$$P_{diss} = V_{DD,PA} \left(1 - V_{DD,PA}\right) \tag{5.25}$$

The dissipated power is converted into heat in the pMOST. As such, for high power levels the chip might get too hot, depending on the power rating of the RF power amplifier. In order to alleviate this problem, the heat dissipation can be shifted off-chip by using external resistors, as shown in figure 5.7 [Dupu02]. Depending on the required RF output power, another pMOST transistor can be selected to ensure that most of the heat generation occurs off-chip.

5.2.4.4 Switching Amplitude Modulator: Class S

A switching amplitude modulator can be made from a Class D amplifier with a low loss low-pass filter. Such an amplifier is denoted as Class S and the basic circuit is depicted in figure 5.8. The envelope signal first passes through a pulse-width modulator, a sopa modulator [Pies01] or a delta-sigma modulator. The resulting pulses are efficiently amplified by the Class D amplifier and the low-loss low-pass filter restores the envelope signal, which is then the supply voltage of the RF amplifier.

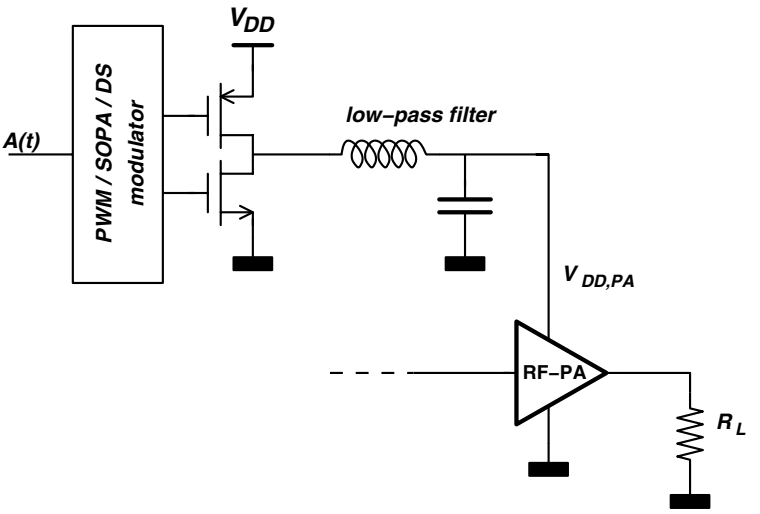


Figure 5.8. Basic diagram of a Class S amplifier as amplitude modulator.

At first sight, the Class S is capable to achieve a 100% efficiency, as it is a switching amplifier. However, the power loss in the parasitic drain-source capacitances of the output transistors will degrade the drain efficiency as discussed in section 2.4.6. Furthermore, the additional power loss of the drive circuitry and especially the power loss in the low-pass filter can not be neglected. The Class S amplifier is well suited for CMOS integration, except for the low-pass filter. The relative large inductor (μH range) and capacitor (nF range) can not be integrated because of their high values. Therefore, a fully integrated solution cannot be achieved with standard CMOS. Furthermore, the inductor of the low-pass filter should have a high quality, a high self-resonance frequency and a high current capability. Combining these three requirements inevitably results in a high cost component or even wishfull thinking.

Another drawback is the bandwidth limitation of the Class S amplifier. The corner frequency of the low-pass filter should be at least as high as the bandwidth of the envelope signal. The switching frequency of the Class S amplifier must be at least one decade higher in order to have at least $40\,\mathrm{dB}$ suppresion of the switching noise at the output, and most wireless systems require more than 60dB of suppression. This easily pushes the switching frequency above $100\,\mathrm{MHz}$. However, the higher the switching frequency, the higher the power loss in the Class S amplifier, making the efficiency advantage of a switching topology less obvious. Furthermore, it is difficult to fabricate a high quality inductor with a self-resonant frequency above $100\,\mathrm{MHz}$ and capable of conducting sufficient current.

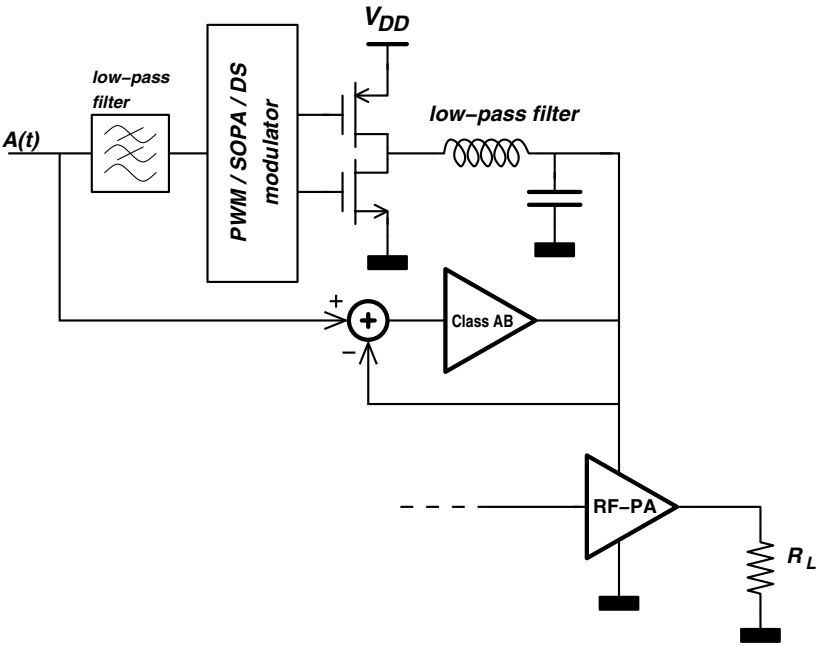


Figure 5.9. Beauty and the Beast: a highly efficient low frequency Class S amplifier with a high frequency low power Class B correction amplifier.

Therefore, most designs and publications use the Class S amplifier not to linearize a nonlinear amplifier, but rather to improve the efficiency of a linear amplifier [Yang99], [Hani99] and [Sahu04]. In such a case, the bandwidth of the Class S amplitude modulator can be reduced, as it is sufficient if the modulator can *slowly* follow the envelope signal, and the linearity demand of the amplitude modulator is less stringent. Altogether, the Class S amplifier can be used as an efficient modulator if the bandwidth of the envelope signal is below *a few megahertz* and a fully integrated solution is not aimed for.

Another approach is the *beauty and the beast* solution, depicted in figure 5.9. It is based on the fact that most of the energy of the envelope signal is located around a relative small frequency band around DC, and the amplitudes of the higher frequencies in the spectrum of the envelope signal are much lower. As such, one could make a solution were a switching Class S amplifier delivers most of the low frequency energy, and a low-power linear amplifier with sufficient bandwidth delivers the high frequency power. The linear amplifier will be less efficient, but delivers less power and thus its influence on the overall efficiency is less pronounced. Of course, such a solution still requires an

off-chip inductor and capacitor and switching noise may deteriorate the output spectrum. Examples of this approach can be found in [Schl04] and [Wang05]

5.3 Distortion in a Polar Modulated Power Amplifier

The major benefit of the polar modulated architecture, is the shift of the linearity requirement from RF to baseband. At lower frequencies, well known techniques such as feedback can easily be applied to improve the linearity of the amplitude modulator. Therefore, the distortion of a polar modulated power amplifier is mainly dominated by the nonlinearities of the Class E amplifier.

First, the modelling of the RF amplifier nonlinearities is discussed. The next sections describe the different causes of nonlinearities in CMOS Class E power amplifiers [Reyn04a].

5.3.1 Nonlinear Polar Modulated Power Amplifier Models

5.3.1.1 AM-AM and AM-PM Distortion

AM-AM distortion in a polar modulated power amplifier is a nonlinear relationship between the envelope signal $A(t)$ at the input of the amplitude modulator, and the envelope of the modulated RF signal at the output of the RF power amplifier. In contrast to the AM-AM distortion that occurs in the Class AB amplifier, the AM-AM distorion discussed here is between a baseband signal and the envelope of an RF signal. Therefore, it should be denoted as AM_{BB}-AM_{RF} whereas in a Class AB, it is rather AM_{RF}-AM_{RF}. The cause of AM-AM distortion in a polar modulated power amplifier can be situated both in the amplitude modulator and in the Class E amplifier. The amplitude distortion of the amplitude modulator, which is in fact AM_{BB}-AM_{BB}, is easy to control since it is a low frequency block. On the other hand, the AM_{BB}-AM_{RF} distortion of the Class E amplifier is less easier to solve and even the cause of such distortion in not always clear. Furthermore, as the Class E amplifier is a highly nonlinear block, it is not always possible to derive analytical formulas.

AM-PM distortion is an unwanted rotation or phase modulation of the RF signal at the output of the RF power amplifier, caused by changing the envelope signal at the input of the amplitude modulator. As with AM-AM, one should denote this distortion as AM_{BB}-PM_{RF}. Likewise, the AM-PM of the Class E amplifier will dominate over the AM-PM distortion of the amplitude modulator.

To simulate or measure the AM-AM and AM-PM distortion of the Class E amplifier, the power supply is slowly varied over the entire dynamic range. A fast Fourier transformation of the captured output voltage can then be calculated and from this, the amplitude and phase of the fundamental output voltage is obtained. These two numbers versus the supply voltage of the Class E amplifier then gives the AM-AM and AM-PM distortion curves of the Class E

amplifier, which is the same as the AM-AM and AM-PM of the polar modulated power amplifier if the distortion of the amplitude modulator can be neglected. Clearly, the AM-AM and AM-PM distortion curves that are obtained in such a manner are based on a DC simulation or measurement of the Class E amplifier. In other words, one assumes that the output envelope and phase are only dependent on the actual value of the envelope signal.

5.3.1.2 PM-AM and PM-PM Distortion

PM-AM distortion is, in most cases, not a concern in RF amplifiers. As already mentioned in section 2.3.2, phase modulation does not change the peak or average output power and thus the biasing of the amplifier does not change by a phase or frequency modulation.

Only for wideband systems, PM-AM distortion may arise. After all, the Class E amplifier is a tuned amplifier and theoretically works at only one specific frequency. Of course, and due to the finite quality factor of the LC networks in a Class E amplifier, the Class E amplifier will have the same performance for frequencies *near* the carrier frequency. In most systems, the bandwidth of the modulated carrier is small compared to the carrier frequency itself. As such, slightly changing the carrier frequency, in other words applying phase modulation to the carrier, will not change the output power of the amplifier and PM-AM distortion is of no concern. However, for wideband systems, like ultra wideband, PM-AM should be considered as well.

From the previous reasoning, one can conclude that PM-PM distortion is not a concern either. As long as the bandwidth is small compared to the RF carrier, the phase will have a linear behavior, meaning a constant group delay which is a requirement for distortionless transmission of phase modulated signals [Couc97].

5.3.1.3 Memory effects

The AM-AM and AM-PM distortion of the amplifier can be obtained from a DC measurement or simulation, as already clarified before. However, the actual amplitude and phase of the RF output voltage will not only depend on the instantaneous value of the envelope signal, but also on the previous values. This effect is denoted as *memory effect*.

A typical example of a memory effect is the thermal behavior of the power amplifier. If the amplifier is working at a higher output power, more power is dissipated an thus the amplifier heats up. If the output power is then reduced, the higher temperature of the amplifier will change the AM-AM and AM-PM distortion curves.

Understanding memory effects is still an issue in power amplifier design. To model these effects accuratly, nonlinear models with memory have been developed, but for a circuit designer little or no insight can be gained from

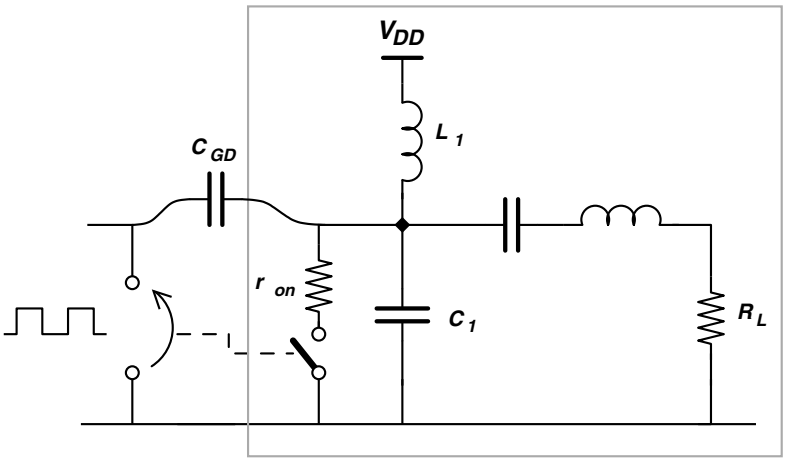


Figure 5.10. Simplified model of the Class E amplifier to extract the AM-AM and AM-PM distortion.

them. In the simulation or measurements, the presence of the memory effects will *smear out* the AM-AM and AM-PM distortion curves of the amplifier.

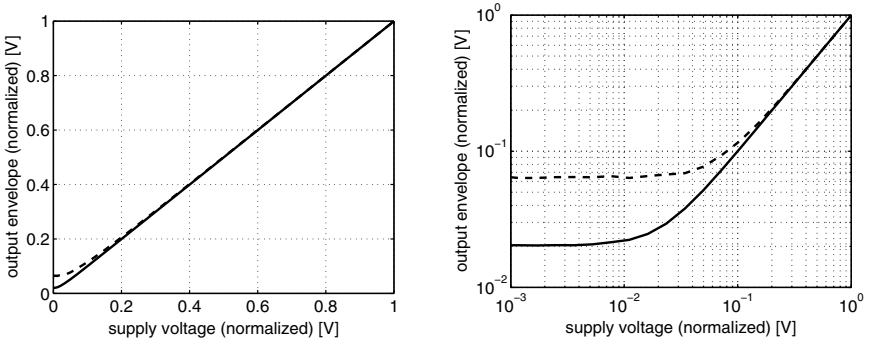
5.3.2 Feedforward

Feedforward from the RF driver stage to the Class E output stage, that flows through the gate-drain capacitance of the Class E nMOS transistor, is the major cause of both AM-AM and AM-PM distortion [Kazi84]. Actually, the feedforward itself is not causing the AM-AM and AM-PM distortion. Rather, if the supply voltage of the Class E amplifier is reduced, the feedforward is more pronounced and it is that effect that actually causes the AM-AM and AM-PM distortion. If the supply voltage of the Class E amplifier is made equal to zero, the feedforward from the driver stage will result in an output voltage that is not equal to zero. On the other hand, if the supply voltage is high enough, the amplitude and phase of the RF output voltage will be dominated by the Class E amplifier itself and the feedforward is then of less importance.

As an example, an idealized Class E amplifier with losses is simulated with a Spice-like simulator. The transistor is replaced by a voltage controlled switch with a series resistance, as shown in figure 5.10. Other parameters are summarized in table 5.2 and are obtained from the design tool discussed in chapter 3. The AM-AM and AM-PM distortion curves are obtained by calculating the fast Fourier transformation of the RF output voltage for different supply voltages. This gives the amplitude and phase of the fundamental output voltage versus the supply voltage of the Class E amplifier.

Table 5.2. Class E PA design values

supply voltage	$0 \ldots 1.25\,\mathrm{V}$
frequency	$2\,\mathrm{GHz}$
switch resistance	$0.25\,\Omega$
inductor loss	$0.75\,\Omega/\mathrm{nH}$
L_1	$0.23\,\mathrm{nH}$
C_1	$13.7\,\mathrm{pF}$
load resistance	$5\,\Omega$
peak envelope output power	$300\,\mathrm{mW}$



(a) (b)

Figure 5.11. AM-AM distortion of a Class E amplifier with a gate-drain capacitance of $1\,\mathrm{pF}$ (solid line) and $3\,\mathrm{pF}$ (dashed line).

The resulting AM-AM curve is shown in figure 5.11. Remember that these curves are a DC approximation. In reality, the normalized supply voltage on the x-axis will be modulated by the amplitude or envelope signal $A(t)$. Two values of gate-drain capacitance were chosen: $1\,\mathrm{pF}$ (solid line) and $3\,\mathrm{pF}$ (dashed line). One can see that the amplitude linearity between power supply and output envelope is well preserved. Only for a supply voltage below $200\,\mathrm{mV}$, the curve starts to deviate for the $3\,\mathrm{pF}$ case. Also, a higher gate-drain capacitance results in more feedforward and thus an increased AM-AM distortion.

The AM-AM distortion only becomes important for relative low supply voltages. Section 2.2.2 showed that some modulation schemes create envelope signals that avoid the origin, or in other words, that limit the dynamic range of the envelope signal. As such, the region of high AM-AM distortion

is avoided. The feedforward from the driver stages to the output stage also restricts the dynamic range of the output signal. This was already suggested by figure 5.11; the minimum output envelope is determined by the feedforward. This may cause problems if one requires a large power control range. The easiest way to reduce the feedforward, is to reduce the supply voltage of the driver stage as well [Sand03]. However, such an approach will also deteriorate the linearity of the amplifier.

The feedforward through the gate-drain capacitance will also cause a rotation of the output carrier. Again, the rotation itself is not a problem, but the rotation will change if the supply voltage of the Class E amplifier is modulated. Therefore, a variation of the envelope signal will induce a time-varying rotation or phase modulation of the carrier, which is denoted as AM-PM. As before, it should actually be written as AM_{BB}-PM_{RF}.

To better understand the cause of AM-PM distortion in a polar modulated amplifier, figure 5.12 depicts a simple model. Of course, the Class E amplifier is a switching amplifier and as such, a small signal model can not be used. However, an intuitive reasoning can still be made. Assume there is no feedforward at all. In that case, the signal at the drain of the nMOS transistor will be approximatly 180 degrees out of phase compared to the signal at the gate. However, the amplitude of the drain voltage is not dictated by the amplitude of the gate voltage, but rather by the amplitude of the supply voltage. The feedforward current through the gate drain capacitance will lead the gate-drain voltage with 90 degrees. This current will flow to the drain of the nMOS transistor and will be converted into a voltage. At the drain, the impedance at the switching frequency is almost resistive since it is a tuned amplifier. The drain voltage will thus rotate, and the amount of rotation will depend on the amount of feedforward. If the supply voltage is reduced, so will the drain voltage and thus the rotation becomes more pronounced, as can be seen on figure 5.12. This intuitive model also predicts that the carrier would rotate 90 degrees if the supply voltage is made equal to zero, which is is exactly the same as the positive zero in a Miller opamp.

Figure 5.13 shows the spice simulation of the same amplifier that was used for the AM-AM simulation. The AM-PM distortion is obtained from the phase of the fundamental Fourier coefficient of the output voltage. From these graphs, it can be concluded that the carrier indeed makes a rotation of 90 degrees if the supply voltage goes to zero. Furthermore, the larger the gate-drain capacitance, the sooner the phase rotation starts. Also, compared to the AM-AM distortion, the AM-PM is more pronounced. From this, it can be expected that the AM-PM distortion is more important than AM-AM distortion in a polar modulated power amplifier.

The AM-AM and AM-PM distortion of a polar modulated Class E amplifier can also be drawn in the complex plane as shown in figure 5.14. Though

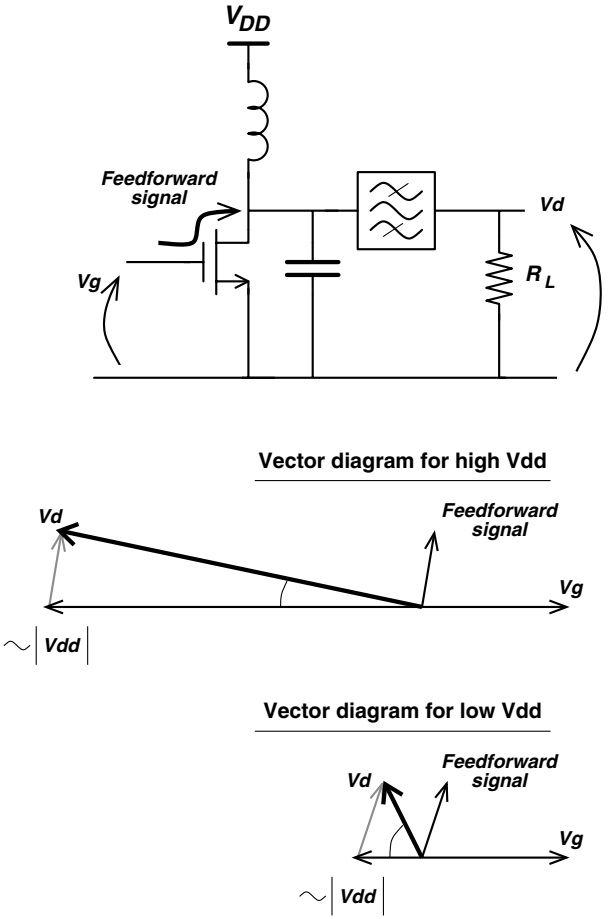


Figure 5.12. Simple model to understand the AM-PM distortion of a switching amplifier.

the distortion seems neglectable in such a representation, it will be shown in chapter 6 that even this relative small rotation can be catastrophic for the power amplifier, since the specifications of most cellular standards are quite high.

The drain-gate capacitance will not only give rise to feedforward. Feedback will occur as well. The feedback signal will slightly change the waveform at the gate of the Class E amplifier, which might also change the phase of the output signal. However, this effect is rather small and can be neglected compared to the influence of the feedforward.

Finally, it should be noted that the AM-AM and AM-PM distortion of the RF amplifier can not be reduced by using a differential RF amplifier.

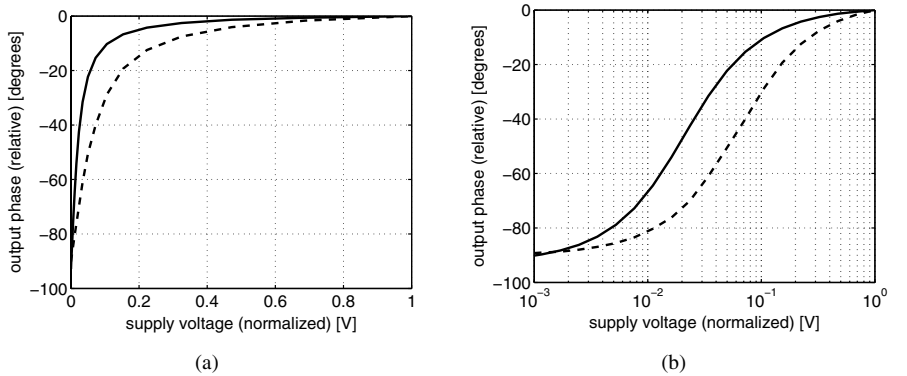


Figure 5.13. AM-PM distortion of a Class E amplifier with a gate-drain capacitance of 1 pF (solid line) and 3 pF (dashed line).

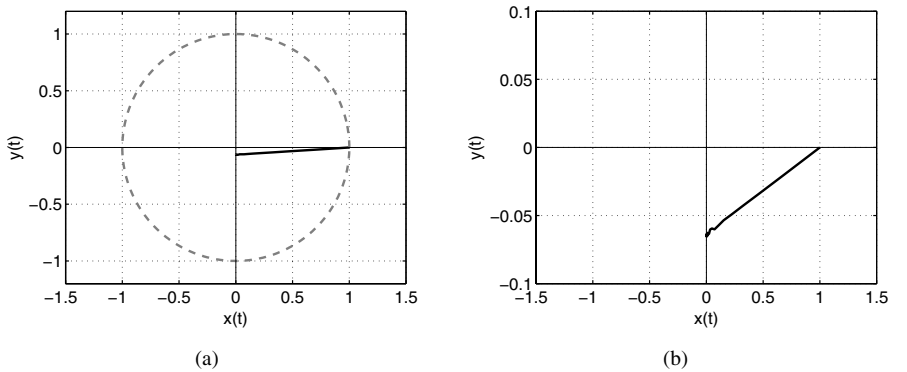


Figure 5.14. AM-AM and AM-PM distortion of a Class E amplifier with a gate-drain capacitance of 3 pF, represented in the complex plane.

5.3.3 Nonlinear on-resistance

In the previous section, the nMOS transistor of the Class E amplifier is modelled as a switch. However, the on-resistance of this switch is not constant, but its value will increase for a higher current. This can clearly be seen in figure 5.15 which shows the well known I-V characteristic of a transistor as well as the on-resistance of the nMOS transistor in the linear region. Of course, the actual value of the on-resistance will change during one RF period, but one can

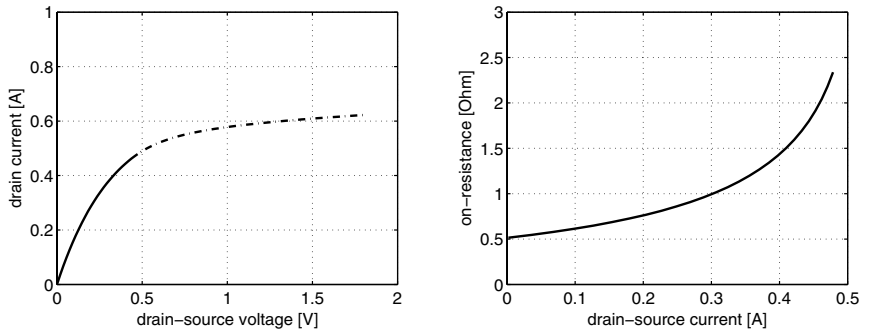


Figure 5.15. (a) drain current and (b) on-resistance of a $0.18\ \mu\mathrm{m}$ nMOS transistor versus drain voltage. $W_g = 1000\ \mu\mathrm{m}$, $L_g = 0.18\ \mu\mathrm{m}$ and gate tied to V_{DD}. The solid line indicates the linear region, dashed-dotted line indicates the saturation region.

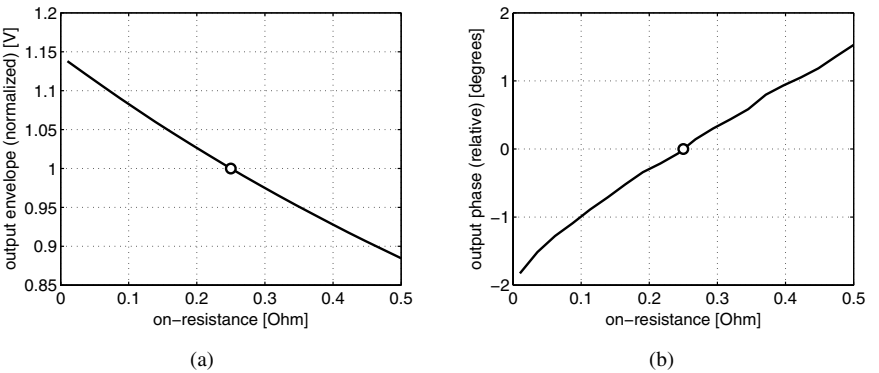
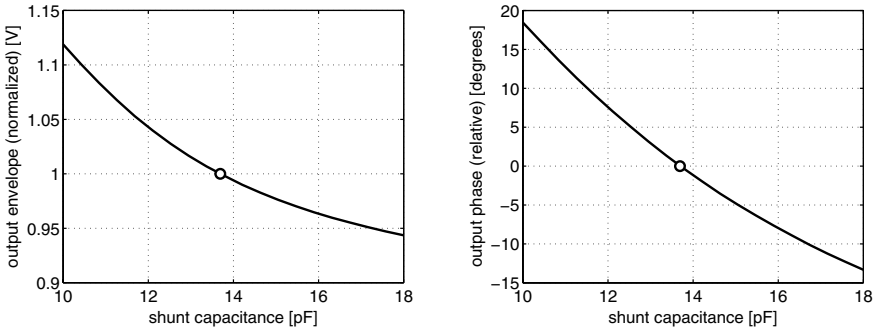


Figure 5.16. AM-AM and AM-PM distortion of a Class E amplifier versus on-resistance.

still think of an average on-resistance that is dependent on the average drain current through the switch The average drain current will in turn depend on the time-varying supply voltage of the amplifier. If the supply voltage is reduced, the average drain current through the transistor will be smaller and thus the average on-resistance will decrease. As a consequence, the efficiency of the amplifier will increase and more power than expected will reach the output. This can be modeled as AM-AM distortion.



(a) (b)

Figure 5.17. AM-AM and AM-PM distortion of a Class E amplifier versus shunt capacitance.

If the on-resistance changes, the shape of the drain voltage will also change and this will in turn change the phase of the sinusoidal output voltage, which is thus a form of AM-PM distortion.

Figure 5.16 shows the AM-AM and AM-PM distortion when the on-resistance is varied. The distortion is normalized to the amplitude and phase of the output voltage for an on-resistance of 0.25Ω. Clearly, the influence of the nonlinear on-resistance is rather small on both AM-AM and AM-PM.

5.3.4 Nonlinear drain-bulk junction capacitance

The drain-bulk capacitance of the nMOS switch in the Class E amplifier has a nonlinear behavior since it partly consists of junction capacitance. It is well known that the value of this junction capacitance decreases as the drain voltage increases. Again, one could think of an average drain capacitance that thus increases when the supply voltage is reduced. A change of the drain capacitance also means a change of the Class E waveform at the drain. Normally, these changes are small and the efficiency of the Class E amplifier remains high.

However, if the shape of the drain voltage slightly changes, so will the amplitude and phase of the output voltage. As such, the nonlinear drain capacitance will cause both AM-AM and AM-PM distortion. The change of the junction capacitance is typically within $\pm 25\%$ of the average value. Figure 5.17 shows the resulting AM-AM and AM-PM for such a variation of C_1. Clearly, the AM-AM is negligible, but the influence on the AM-PM distortion can not be neglected.

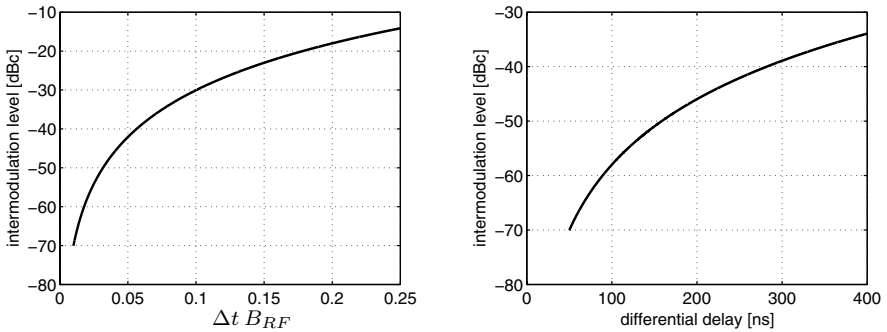


Figure 5.18. Intermodulation distortion of a two-tone test versus delay between envelope and phase, after [Raab96]. In (b), an RF bandwidth of 200 kHz is used.

5.3.5 Differential Delay

In a polar modulated power amplifier, the envelope and phase signals flow through different paths, and recombine only in the last stage of the RF amplifier. Therefore, both signals may experience a different delay and as such, the envelope signal gets recombined with the *wrong* phase signal. A delay of the envelope signal and phase signal on itself is not an issue, rather it is the delay difference between the two signals that causes distortion. As for AM-AM and AM-PM, it is difficult or even impossible to translate this distortion mechanism to a linearity degradation in general. After all, the impact of a distortion mechanism is dependent on the type of modulation and on the requirements of a specific wireless standard.

Nevertheless, the linearity degradation for a differential delay can be calculated for the case of a two-tone signal [Raab96]. The derivation is not repeated here, but the approximated formula for the intermodulation equals

$$IMD = \pi \left(\Delta t \; B_{RF}\right)^2 \tag{5.26}$$

with Δt the delay between amplitude and phase and B_{RF} the bandwidth of the modulated RF signal.

Equation 5.26 is plotted in figure 5.18. In order to have an intermodulation that is better than $-50\,\mathrm{dBc}$, the $\Delta t \; B_{RF}$ product should be smaller than 0.032. For an RF channel bandwidth of $200\,\mathrm{kHz}$, this means that the delay between amplitude and phase should be kept below $150\,\mathrm{ns}$. Of course, a two-tone signal is a rather poor approximation of a complex digital modulated signal, but nevertheless equation 5.26 can still be used as a guideline.

5.3.6 Envelope Filtering

The bandwidth of the envelope signal $A(t)$ will be limited by the bandwidth of the amplitude modulator. If the latter has a wide bandwidth it will consume more power, which reduces the overall efficiency of the linearized amplifier. On the other hand, if the bandwidth of the envelope signal is reduced, the modulated RF signal at the output of the Class E amplifier will be distorted. Thus, as is the case with linear amplifiers, a *linearity-efficiency* tradeoff appears.

An important issue regarding this envelope filtering is the fact that the bandwidth of the envelope signal is considerably larger than the bandwidth of the modulated RF signal. A two-tone signal can be used to explain this phenomenon. Figure 5.19(a) shows the time waveform of a $1\,\mathrm{MHz}$ baseband sinewave, multiplied by a $1\,\mathrm{GHz}$ carrier[1]. The resulting spectrum is shown in figure 5.19(b) and the modulated RF signal occupies a bandwidth of $2\,\mathrm{MHz}$. The envelope signal $A(t)$ and the corresponding spectrum of $A(t)$ are shown in figures 5.19(c) and 5.19(d). The spectrum of the envelope signal is clearly much broader than the $2\,\mathrm{MHz}$ bandwidth of the modulated RF signal. The same holds for the phase signal. The RF phase signal, defined as

$$v_P(t) = \cos\left(\omega_c t + P(t)\right) \tag{5.27}$$

is shown in figures 5.19(e) and 5.19(f). Clearly, the constant envelope RF phase signal, applied at the gate of the Class E amplifier, has a wider spectrum compared to the modulated RF signal. Also note that the first spectral line of the envelope signal is located at $2\,\mathrm{MHz}$, and not at $1\,\mathrm{MHz}$.

The fundamental reason for the wider bandwidth of the envelope and phase signals, is the nonlinear mathematical relationship between these signals and the in-phase and quadrature signals $x(t)$ and $y(t)$ [McCu03].

$$A(t) = \sqrt{x(t)^2 + y(t)^2} \tag{5.28}$$

$$P(t) = \arctan\left(\frac{y(t)}{x(t)}\right) \tag{5.29}$$

In most cases, and also in the case of a two-tone signal, $x(t)$ and $y(t)$ have a limited bandwidth, see figure 2.3, but $A(t)$ and $P(t)$ will have a wide bandwidth, due to the nonlinear mathematical operation (square root and arctan).

The question now arises how large the bandwidth of the envelope signal has to be, in order to avoid excessive linearity degradation. Similar to the differential delay distortion, the intermodulation distortion can be calculated for a simple case only, like a two-tone signal [Raab96]. A closed formula can not be obtained, but the numerical results are shown in figure 5.20. The figure shows

[1] for clarity, a carrier frequency of only $10\,\mathrm{MHz}$ is selected in figures 5.19(a) and 5.19(e)

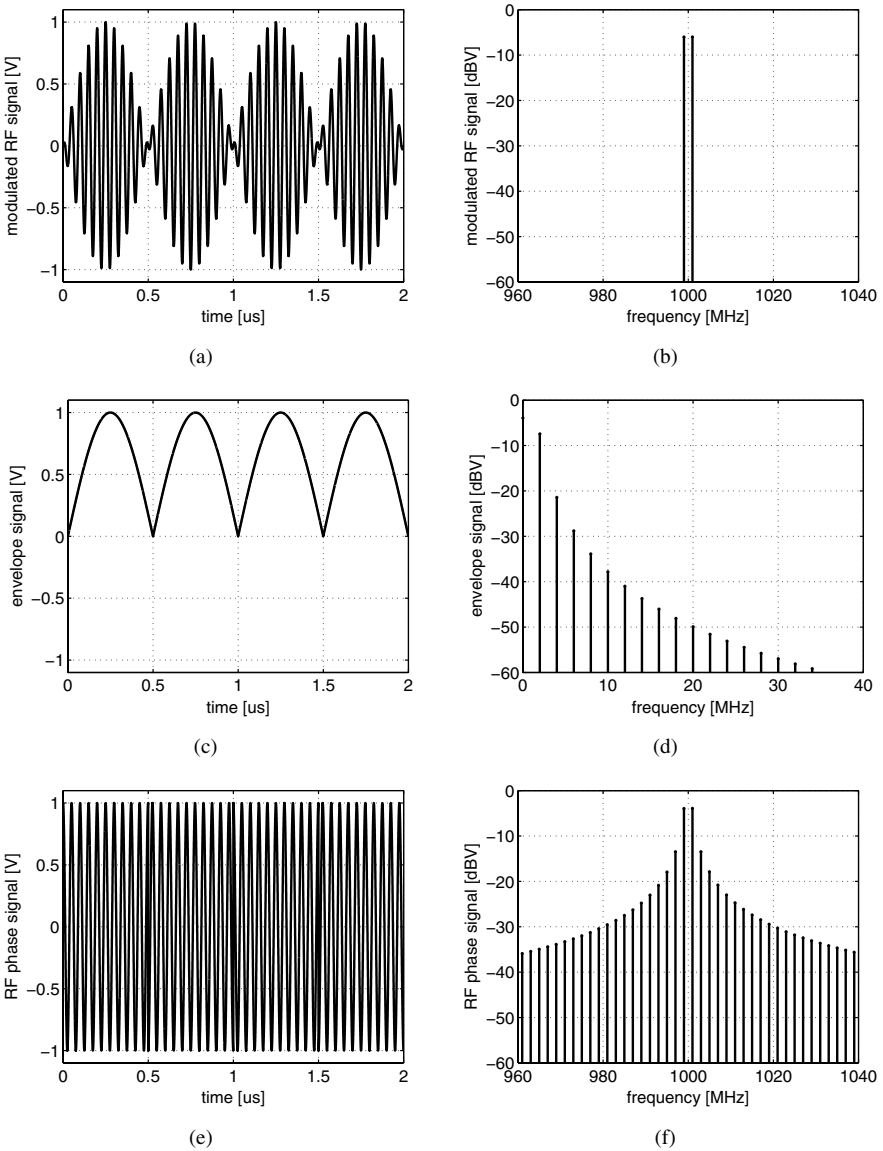


Figure 5.19. Time waveforms and spectra of a two-tone signal.

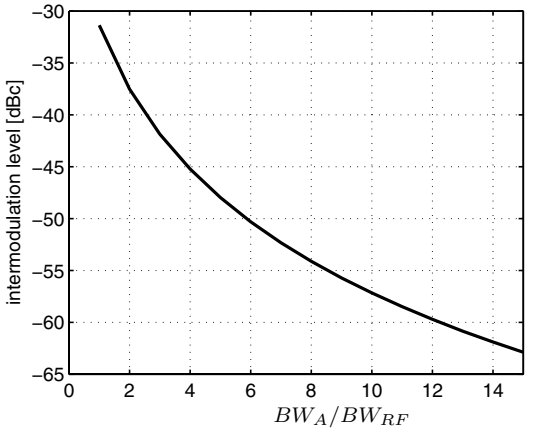


Figure 5.20. Approximated intermodulation distortion of a two-tone test versus envelope bandwidth, after [Raab96].

the intermodulation distortion of a two-tone signal versus the relative envelope bandwidth, with BW_{RF} the bandwidth of the RF signal and BW_A the bandwidth of the envelope signal $A(t)$. It can be seen that, in order to have better than $-50\,\mathrm{dBc}$ of intermodulation distortion, the bandwidth of the envelope path should be more than six times the bandwidth of the modulated RF signal.

The derivation in [Raab96] and the result in figure 5.20 assume an ideal brick-wall filter. Such a filter makes the harmonics higher than the cut-off frequency equal to zero, and it has a zero phase response. However, the phase response of the envelope filter will have a large impact on the distortion. To demonstrate this, the two-tone example of figure 5.19 can be used again. Figure 5.21 shows the effect when the envelope signal is filtered by a first order low-pass filter with a corner frequency of $8\,\mathrm{MHz}$, i.e. four times the bandwidth of the modulated RF signal. Figure 5.21(a) shows the spectrum of the original envelope signal and the filtered version and figure 5.21(b) shows the time waveform of the filtered envelope signal. In the latter, the sharp peaks of the envelope signal that go to zero, are gone. Also, the envelope signal is a little bit delayed by the phase response of the envelope filter. In figure 5.21(c), the spectrum of the modulated RF signal at the output is depicted, and the intermodulation products, due to the $8\,\mathrm{MHz}$ envelope filtering, are clearly visible. These products are generated by both the amplitude and phase response or time delay of the $8\,\mathrm{MHz}$ envelope filter, despite the fact that this filter is a linear block.

The delay of the envelope filter can be compesated by inserting an *appropriate* delay in the phase path. For this specific example, the best intermodulation

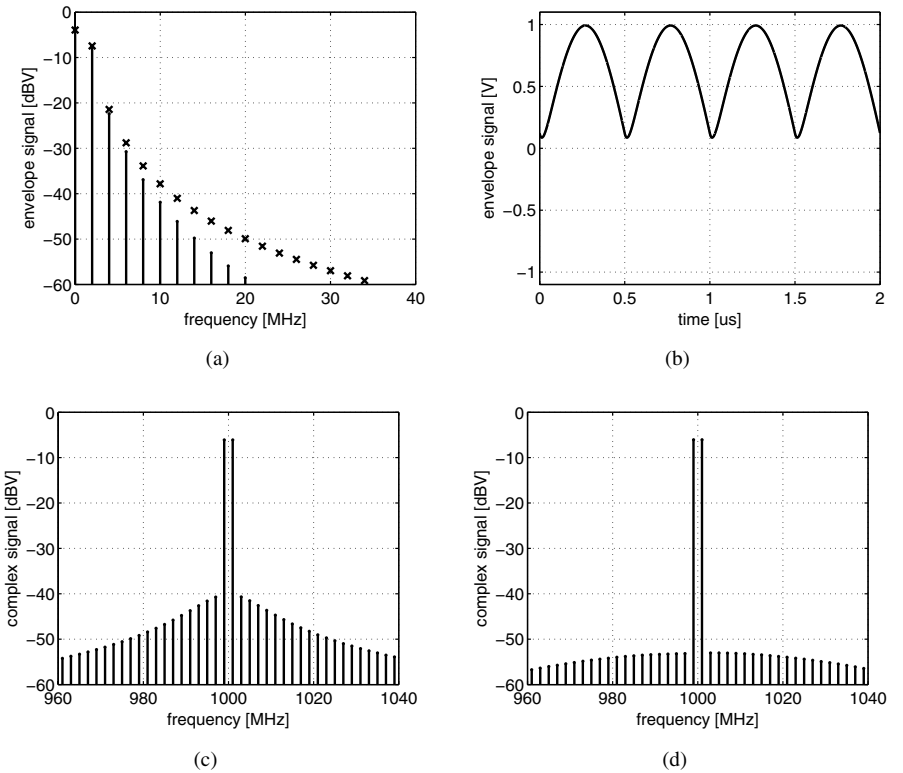


Figure 5.21. Time waveforms and spectra of a two-tone signal with an $8\,\mathrm{MHz}$ low-pass envelope filter. In (d), the phase signal is delayed by $20\,\mathrm{ns}$.

is obtained if a delay of $20\,\mathrm{ns}$ is inserted in the phase path. The resulting spectrum is shown in figure 5.21(d). Without the delay insertion in the phase path, the intermodulation distortion is only $-34\,\mathrm{dBc}$. With the delay of $20\,\mathrm{ns}$, the intermodulation distortion drops to $-47\,\mathrm{dBc}$, which is close to $-45\,\mathrm{dBc}$, predicted by figure 5.20 and [Raab96].

The phase response of the envelope filter thus causes a delay of the envelope signal, and this was not taken into account in [Raab96]. In reality, there is no *real* time delay between input and output signals, but only a phase difference. The phase difference can be converted into a time delay between the unfiltered and filtered envelope signal. In order to have a distortionless transmission, the amplitude response of the filter should be flat and the phase response should be a linear function of the frequency [Couc97]. In a first order low-pass filter,

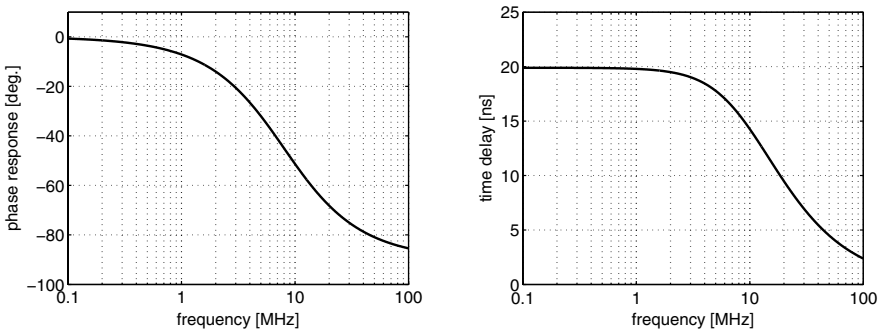


Figure 5.22. Phase response and time delay of a low-pass filter with a $8\,\mathrm{MHz}$ corner frequency, used to filter the envelope signal in figure 5.21.

both requirements are clearly not met. For a first order low-pass filter with transfert function

$$H(j\omega) = \frac{1}{1 + j\omega/\omega_{3dB}} \tag{5.30}$$

the phase response is equal to

$$\theta(\omega) = \tan^{-1}\left(-\omega/\omega_{3dB}\right) \tag{5.31}$$

and from this, the time delay of the filter can be obtained

$$T_d = -\frac{\theta(\omega)}{\omega} = \frac{1}{\omega} \cdot \tan^{-1}\left(\omega/\omega_{3dB}\right) \tag{5.32}$$

If the phase response is a linear function of frequency, the time delay will be constant over the entire frequency range. For a low-pass filter however, the time delay will change with frequency. For low frequencies, well below the corner frequency of the filter, the time delay is constant and equal to

$$T_d = \frac{1}{\omega_{3dB}} \tag{5.33}$$

and for higher frequencies, the delay goes to zero. The phase response and time delay of the $8\,\mathrm{MHz}$ low-pass filter used in the previous example, is shown in figure 5.22.

The group delay, defined as

$$T_g = -\frac{1}{2\pi} \cdot \frac{d\theta(\omega)}{d\omega} \tag{5.34}$$

is of no importance here. A constant group delay is a requirement for a distortionless transmission of a bandpass signal. However, the envelope signal is a baseband signal and the requirement of a constant time delay is more strict than the requirement of a constant group delay.

In reality, the envelope signal consists of many frequency components, and the delay of the high frequency components will be less than the delay of the low frequency content of the envelope signal. If the corner frequency of the envelope filter is high enough to ensure that most of the envelope spectrum falls within the pass band of the low-pass filter, the envelope signal is delayed by the low frequency time delay of the filter, which is equal to

$$T_d = \frac{1}{\omega_{3dB}} \tag{5.35}$$

As said before, this delay has to be added in the phase path to realign the two signals and to ensure that the recombination in the power amplifier is correct.

The phase signal of a two-tone signal is a square wave, meaning that the carrier makes an abrupt rotation of 180 degrees. In the unfiltered modulated RF signal, this abrupts change of phase is not visible since the envelope signal is zero at that instant. However, the filtered envelope signal does not go to zero, and as such the instant change of the phase is visible in the RF signal and will cause intermodulation products. A detailed view of the filtered envelope signal is shown in figure 5.23. The original envelope signal in shown in grey, and clearly the filtered envelope signal does not go to zero and it is delayed. Also, the minimum of the filtered envelope signal does not occur at the sudden transition of the phase signal, shown by a dashed line in figure 5.23(a). As said before, for this specific example, the best intermodulation is obtained for a delay of $20\,\mathrm{ns}$. If a delay of $40\,\mathrm{ns}$ is used (see figure 5.23(e)), the intermodulation distortion increases.

The *appropriate* delay, i.e. the delay that causes the least amount of intermodulation, is of course related to the time delay of the filter but a clear relationship is not easily derived, since every frequency component of the envelope signal is delayed by a different amount. In the two-tone example, the optimal delay is $20\,\mathrm{ns}$ which indeed corresponds to the time delay of the $8\,\mathrm{MHz}$ filter for low frequencies. However, the optimal delay is not only dependent on the corner frequency of the envelope filter, but also on the actual time behavior of the envelope and phase signals.

It can be concluded that the bandwidth of the envelope path and thus the bandwidth of the amplitude modulator, has to be considerably larger than the bandwidth of the modulated RF signal at the output. However, it is possible to reduce the bandwidth of the envelope signal. The generated distortion can partially be overcome by inserting a delay in the phase path. This delay is related to, but smaller than, the time delay of the envelope filter. As a final

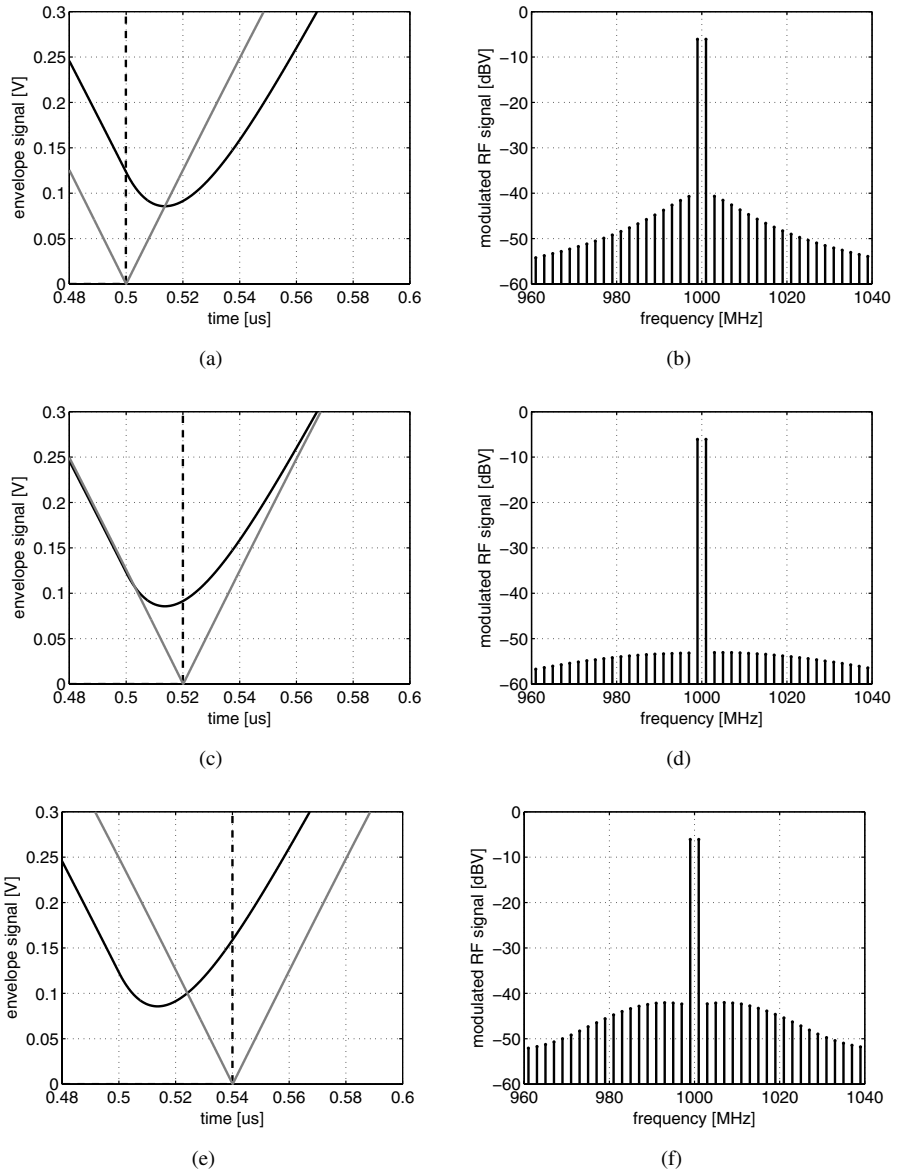


Figure 5.23. Time waveforms and spectra of a two-tone signal with a filtered envelope signal. In (a) and (b), no delay is applied, in (c) and (d) a dely of 20 ns is used and in (e) and (f) a delay of 40 ns is applied.

note, it is remarkable that a low-pass filtering of the envelope signal creates a wider bandwidth after upconversion.

5.3.7 Injection of the Phase Signal

The large gate-drain capacitance of the nMOS switch in the Class E amplifier, will feedforward the signal from the driver stage to the output. This feedforward will cause both AM-AM and AM-PM as already discussed.

However, the feedforward signal is only phase modulated and has a wider bandwidth compared to the RF signal, as demonstrated by the two-tone signal example in figure 5.19. The feedforward will thus not only cause AM-AM and AM-PM, but it will also inject a wideband signal at the output.

If the envelope signal has a DC offset, the phase signal also gets injected towards the output. Such an offset may occur in the D/A converter after the DSP as well as in the analog circuitry of the amplitude modulator. The injection due to the offset can easily be seen by following equation

$$\begin{aligned} v(t) &= (A(t) + V_{offset}) \cdot \cos(\omega_c t + P(t)) && (5.36) \\ &= A(t) \cdot \cos(\omega_c t + P(t)) + V_{offset} \cdot \cos(\omega_c t + P(t)) && (5.37) \end{aligned}$$

The second term, $V_{offset} \cdot \cos(\omega t + P(t))$ suggests that the RF phase signal is added to the output. Remember from figure 5.19 that this signal has a wide bandwidth. The injection of this signal at the output will thus give rise to a large number of spectral components, though these are not caused by a nonlinearity.

5.3.8 Linearity Improvement Techniques

The distortion mechanisms that were discussed in the previous sections will deteriorate the linearity of the polar modulation technique. This can be solved by using a linearity improvement technique. The two important ones are discussed in this section.

5.3.8.1 Predistortion

Predistortion first measures the non-idealities of the RF amplifier and applies the inverse curve to obtain a more linear behavior. In a polar modulated power amplifier, the correction can be applied both on the amplitude and phase path. Hence, predistortion can alleviate both AM-AM and AM-PM distortion. Delay compensation can also be accounted for and can also be regarded as predistortion. Predistortion can not solve the dynamic related distortion of the amplifier, such as filtering of the envelope signal or memory effects. [Kusu02], [Teik04], [Ceyl04] and [Andr99]. Recent results show that thermal memory effects, that have a slow enough time behavior, can also be compensated for [Boum03].

For predistortion to be effective, the AM-AM and AM-PM distortion should first be measured, stored in a look-up table, and applied on the actual transmitted data. A drawback of predistorion is the inability of the system to anticipate on any changes in the system behavior. Of course, the distortion curve could be updated frequently to accomodate for this problem, but still, short term memory effects and envelope filtering can not be overcome with predistortion [Teik04].

The advantage of predistortion is that instability can never occur, and it is relatively easy to implement at a low cost. As such, predistortion is frequently applied, not also in research, but also in commercially available products [Sand03].

5.3.8.2 Polar Feedback

Feedback is a well known technique in electronics and has been applied in conjunction with polar modulation as well. Either the amplitude signal, the phase signal or both signals could be fed back and compared with the input signal. In contrast to predistortion, feedback has the advantage to operate in real-time and as such, it can overcome distortion mechanisms that are related to the dynamic behavior of the RF amplifier like memory effects, envelope filtering and temperature or other environmental variations. Of course, a feedback loop might become unstable and the analog blocks that are placed in the feedback path should have little delay and low distortion. Polar feedback consists of two loops: a phase feedback loop and an amplitude feedback loop [Sowl04]. The two feedback loops should be accuratly matched to each other and stability is hard to maintain over the large dynamic range of modern communication systems.

Phase feedback seems relatively easy since the feedback path does not need amplitude linearity. As such, a simple limiter can be used to extract the phase signal from the RF output signal. However, the limiter may have a substantial amount of AM-PM distortion, especially if one takes into account that the RF signal at the input can have large envelope variations [Sowl97]. Also, for modulation schemes that pass through the origin of the complex plane, the RF output voltage may become equal to zero, meaning that there is no voltage to be fed back in the phase path. The same holds for amplitude feedback. An envelope detector is required that can operate over the entire dynamic range of the system and has a low AM-PM distortion [Su98].

5.4 Power Combination and Polar Modulation

In this section, it will be demonstrated how the power combining technique, discussed in section 4.3, can improve the efficiency and reduce the power consumption of the polar modulation architecture. In a polar modulated power amplifier, the overall efficiency reduces when the supply voltage of the Class

E amplifier is lowered. The cause of this reduction is the fixed dissipation of the driver stages, as discussed and demonstrated in section 5.2.3.

The Class E amplifier requires a switching waveform at the gate of the nMOS and thus the driver stages are implemented as switching or saturated amplifiers, requiring as less DC power as possible. Digital inverters or satured tuned amplifiers are most commonly used. For a digital driver, the DC power consumption can be written as

$$P_{DC,DRV} \sim f \cdot W_g \cdot V_g^2 \tag{5.38}$$

with W_g the width of the Class E nMOS transistor, V_g the maximal voltage at the gate and f the switching frequency. Clearly, reducing V_g has a quadratic effect and it is thus most tempting to reduce the voltage of the driver stages in order to increase the overall efficiency. On the other hand, this will increase the on-resistance of the Class E nMOS transistor. The required value of the latter is determined by the load resistance of the amplifier, which remains the same if the supply voltage of the Class E amplifier is reduced. Therefore, if the on-resistance of the Class E amplifier increases, the Class E amplifier would quickly leave the Class E operating regime and the efficiency would reduce dramatically, though an improvement was initially aimed for.

Another solution is to change the width of the nMOS transistor, W_g in equation 5.38. In practice, this would result in several parallel nMOS transistors that can be switched on and off. Placing less transistors in parallel would then reduce the physical width W_g of the Class E switch. The driver stages would consume less power, since they now have to charge and discharge a smaller capacitance, but again the on-resistance of the Class E amplifier increases. As indicated before, the amplifier will no longer operate as Class E and the drain efficiency would degrade.

In the two previous examples, the fixed value of the load impedance of the amplifier is the main cause of trouble. If the load impedance could be increased, either the size or the gate voltage of the Class E nMOS switch could be reduced and a reduction of the power consumption in the driver stages will be the result. This approach will keep the amplifier in its Class E operating regime, thus maintaining a high drain efficiency.

Electrically changing the impedance *seen* by the amplifier, is denoted as *active load pull*. A discrete form of this technique was already found in chapter 4, section 4.3.4. The active load-pull technique is used to generate a wide variety of load impedances, seen by the power amplifier. As such, one can look for the ideal load, i.e. the load for which the power transistor delivers the highest amount of output power, or achieves the highest efficiency, or achieves the best linearity, etc. The basic idea of active load pull is depicted in figure 5.24. Current source I_{PA} represents the power transistor and current source I_{ALP}, which has the same frequency but may have a different amplitude or

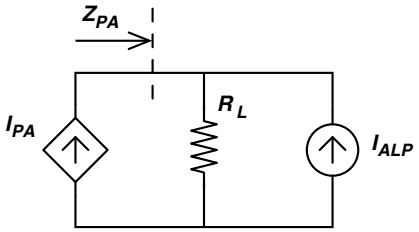


Figure 5.24. Principle of active load-pull.

phase, is the active load-pull current source. The impedance *seen* by the power transistor is equal to

$$Z_{PA} = R_L \left(1 + \frac{I_{ALP}}{I_{PA}}\right) \tag{5.39}$$

Clearly, by changing the phase and amplitude of I_{ALP} versus I_{PA}, virtually any impedance can be created. This technique is also applied in the Doherty amplifier [Crip99, Dohe36]. In the power combining technique, discussed in chapter 4, the impedance seen by each amplifier also changes, depending on the number of parallel stages as

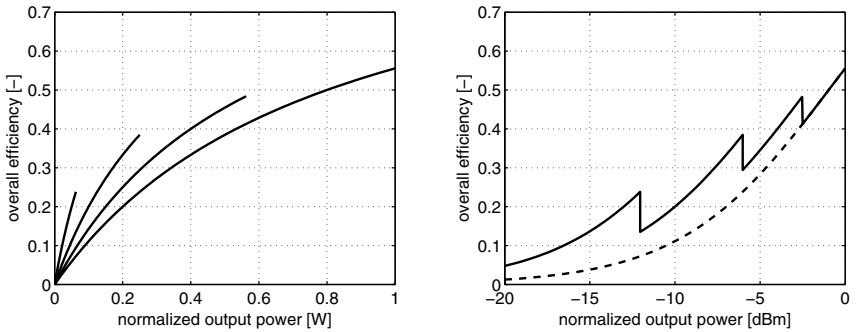
$$R_{in} = \frac{R_L}{2} \cdot \left(\frac{B}{R_L}\right)^2 \cdot \frac{1}{N} \tag{5.40}$$

and can thus be seen as a discrete form of active load pull.

The power combining technique will have a positive influence on both the drain efficiency of the Class E amplifier, and on the overall efficiency. First of all, the drain efficiency will strongly depend on the ratio of the load resistance to the on-resistance of the switch. An approximated formula is given by [Raab78] and [Yoo01].

$$\eta \approx \frac{1}{1 + 1.4 \cdot \frac{r_{on}}{R_L}} \tag{5.41}$$

By placing less sections in parallel, thus decreasing N in equation 5.40, the load resistance R_L will increase which in turn increases the drain efficiency of the remaining power amplifiers, according to equation 5.41. Secondly, if one or more power amplifiers are turned off, the driver stages of the corresponding sections can be shut down as well. Therefore, the total power consumed by the driver stages will be reduced. If a digital driver is used, the consumption of the driver stages will be proportional to the number of parallel sections N whereas the output power will be proportional to N^2. To demonstrate the latter effect, figure 5.25 shows the overall efficiency when one to four sections are placed in parallel. The power consumption of the driver stage is assumed to be one fifth (20%) of the peak power dissipation of each individual amplifier. When



(a) (b)

Figure 5.25. Overall efficiency versus output power for 1, 2, 3 and 4 parallel sections.

the supply voltage of all four amplifiers is reduced, the overall efficiency decreases as already pointed out in section 5.2.3. However, at lower power levels, sections can be turned off and this will clearly increase the overall efficiency. As such, one could select the curve that gives the best overall efficiency for the required output power. In figure 5.25, the theoretical increase of the drain efficiency due to the higher load resistance is not accounted for, so the actual overall efficiency curve is in fact a little higher, at least in theory.

From the previous discussion, one can conclude that the power combining technique of section 4.3 can be used together with polar modulation, to increase the overall performance of the architecture at lower output power levels. A CMOS implementation of the power combining technique will be demonstrated in chapter 7.

5.5 Full Digital Linearization

5.5.1 A single-bit RF D-to-A

The major drawback of the switching amplitude modulator of section 5.2.4.4, is the *bulky* off-chip LC filter. However, this filter can be *shifted* to the output of the RF amplifier, as depicted in figure 5.26. One could say that the lowpass filter is *upconverted* to a bandpass filter at the output. Now, the RF power amplifier is turned on and off by the pulses delivered by the Class D amplifier, which is controlled by a PWM, SOPA or $\Delta\Sigma$ modulator [Stas05, Wagh04, Raab02, Pies01]. This can be considered as *single bit digital polar modulation*.

It should be noted that the low frequency Class D amplifier still has to deliver a large amount of current. As such, this technique is not a solution for wideband systems that require an envelope bandwidth of several tens of Mega-

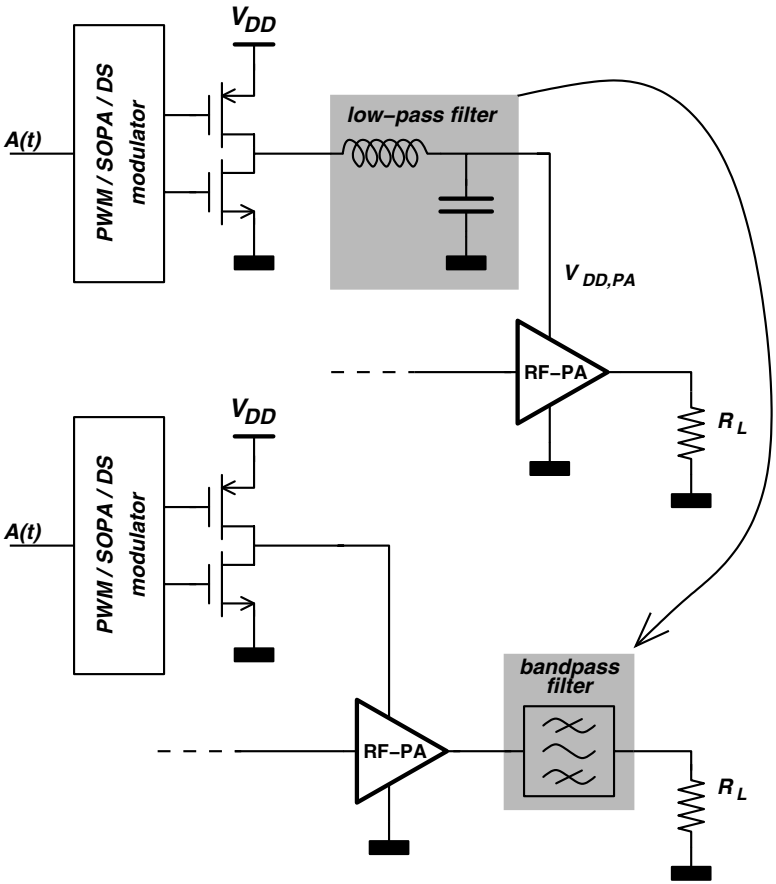


Figure 5.26. The low-pass filter of the Class S amplifier can be shifted to the output of the RF amplifier, where it becomes a bandpass filter.

hertz. Rather, it avoids the need of the expensive LC filter, but still an external RF bandpass filter, with a high efficiency, is needed.

Turning on and off the RF amplifier can also be done at the input of the RF amplifier, or the upconversion mixers can be turned on and off by the square-wave from the modulator. This technique is depicted in figure 5.27. If the output of the $\Delta\Sigma$ is lying between +1 and 0, the RF amplifier is turned on and off continuously, which may cause additional distortion. The $\Delta\Sigma$ modulator can also be designed that its output is lying be +1 and -1. In that case, the input of the RF amplifier has abrupt phase transitions, but it also has a constant envelope. In other words, the power amplifier will not be turned off, but has to amplify a phase-modulated carrier with a wide bandwidth.

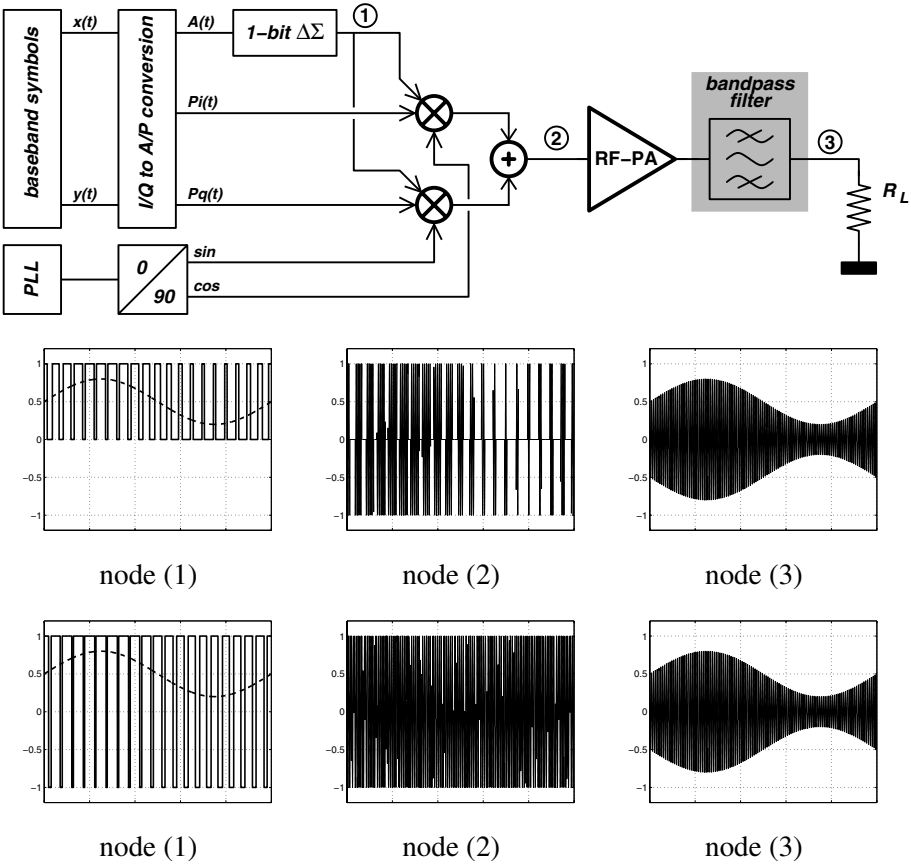


Figure 5.27. Full digital linearization.

5.5.2 The Lattice-type LC balun as a multi-bit RF D-to-A

The feasibility of these techniques will clearly depend on the performance of the RF bandpass filter at the output. After all, this filter has to be able to suppress the switching noise from the modulator and should have a very low power loss, since it is placed in series with the load. Both the linearity and the noise suppression are improved if one could use a multi-bit $\Delta\Sigma$ modulator. This requires an RF power amplifier that can efficiently switch between several discrete output power levels. Again, the power combining technique presents itself as a solution for this problem. In the latter, the output power of the RF amplifier can be controlled in a discrete manner by turning on and off parallel sections. One could think of an *enable-pin* that ensures that the output of the RF amplifier creates an AC ground to properly terminate the corresponding

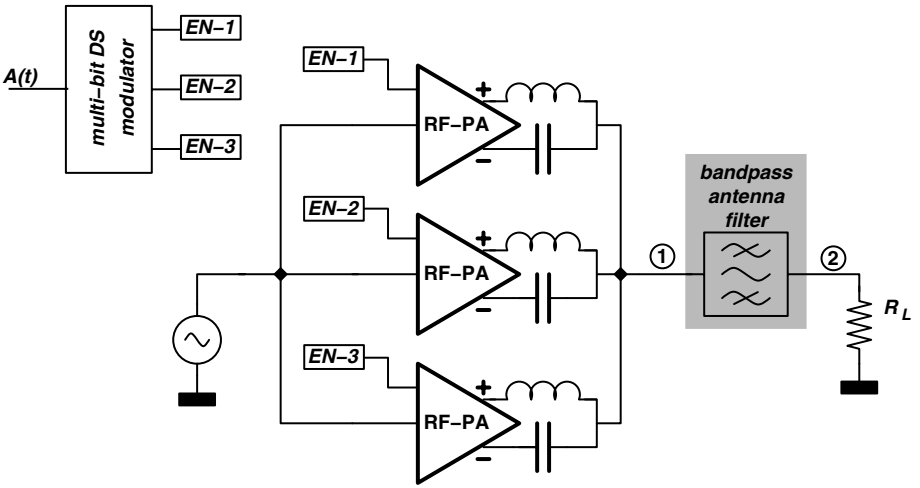


Figure 5.28. Lattice-type LC balun, used as an multi-bit RF D-to-A converter.

section of the lattice-type LC balun. As such, a multi-bit $\Delta\Sigma$ modulator could drive the *enable-pins* of several parallel power amplifiers, as shown in figure 5.28. Altogether, the architecture of figure 5.28 can be seen as an RF D-to-A converter and could be denoted as *multi-bit digital polar modulation*.

The different sections of the lattice-type LC balun can be scaled in a binary manner, as discussed in section 4.3.5. The advantage is a higher accuracy, but the active load-pull effect will make the amplifiers of the lower bits less efficient since they have to be designed to drive a low load impedance. Therefore, a linear scaling in which only one specific power amplifier is operational, may result in a higher efficiency since each amplifier can be optimized for a specific load resistance. For the same accuracy, a higher silicon area, mainly consumed by the inductors, will be the penalty. On the other hand, if one mitigates to higher frequencies, the inductors can be implemented as *slab* inductances [HH04, Aoki02b] and as such, the silicon area can be kept at a reasonable level.

For wideband systems, one should be able to switch on and off the different sections at a very high rate, without consuming too much power. The circuit implementation of the *mystical* enable pin is thus crucial for a wideband linearization. An obvious approach, shown in figure 5.29(a), is to place a pMOS switch in series with the power supply of each differential section to disconnect the power supply of that section. The two nMOS transistors of the corresponding differential Class E amplifier can then be switched on to connect L_m and C_m in parallel, which creates the high output impedance of that section. The

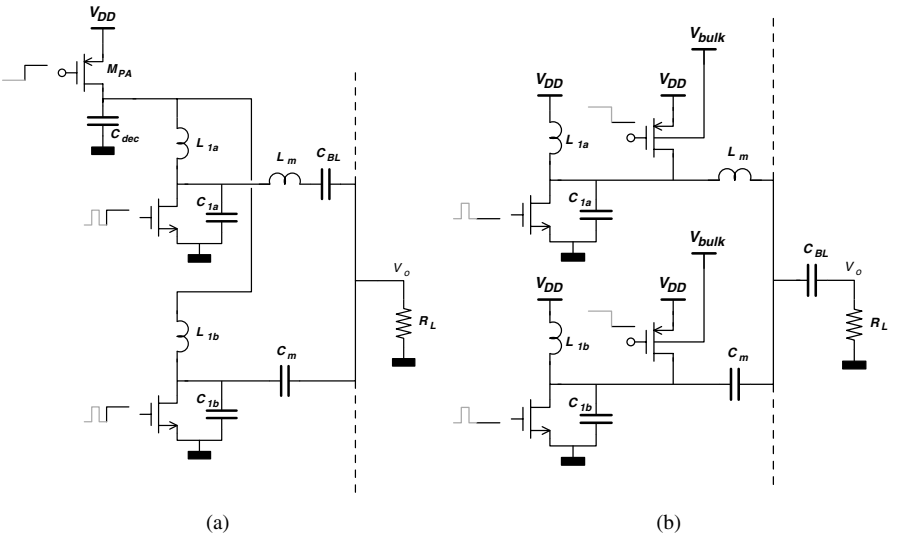


Figure 5.29. Implementation of the enable-pin in a Class E amplifier that uses the lattice-type LC balun.

major drawback of this technique is that the pMOS transistor needs to be large to reduce the drain-source voltage drop across that tranistor. This also means a high capacitance that needs to be driven by the $\Delta\Sigma$ modulator, and thus a large power consumption. On the other hand, if one can place many sections in parallel, each section only needs to deliver a small amount of power, and hence that pMOS transistor can now become smaller.

Another solution is to connect L_m and C_m not with the ground but with the power supply itself, which is an AC ground [Shir01]. This is shown in figure 5.29(b). The width of these two pMOS transistors can now be considerably smaller since they only need to provide a short circuit for the signal current. A problem is the bulk voltage of the two pMOS switches. When the section is activated, the drain voltage goes above the supply voltage and therefore the bulk voltage of the pMOS has to increase to avoid conduction of the drain-bulk diode of the pMOS. When the section is turned off, the bulk voltage of the pMOS can be made equal to V_{DD} in order to avoid the bulk effect and to have a low on-resistance.

5.6 Conclusion

In this chapter, the nonlinear RF Class E amplifier is linearized by modulating the supply voltage. This technique is denoted as *polar linearization*. The advantage of polar modulation is that the entire RF path can be nonlinear, which means a lower DC power consumption. Furthermore, the linearity

requirements are shifted to a low frequency, which again results in less power dissipation. Finally, a fully integrated solution becomes feasible. The basic equations, the architecture and the implementation of the amplitude modulator were covered in section 5.2.

The distortion of the polar modulated power amplifier was discussed in section 5.3. First, some nonlinear models are reviewed. AM-AM and AM-PM distortion are the most common and known techniques to discribe distortion in RF amplifiers. PM-PM and PM-AM are of less importance for narrowband communication systems. Memory effects are difficult to discribe and interprete and the corresponding models give little or no insight for the circuit designer. Feedforward from the driver stage to the Class E output stage is the dominant distortion mechanism and causes both AM-AM and AM-PM. A differential delay between the amplitude and phase signal will also result in a incorrect or distorted output signal. Finally, the third important distortion mechanism is low-pass filtering of the envelope signal. This distortion can partially be overcome by inserting a time delay in the phase path to correct for the phase response of the low-pass envelope filter.

In section 5.4, the power combining technique is suggested to improve the performance of the polar modulated power amplifier. The lattice-type LC balun enables to shut down one or more amplifier sections, see chapter 4, and this will increase the overall efficiency at lower output power levels due to the lower power consumption of the driver stages.

The last part of this chapter is devoted to the idea of full digital linearization. Instead of linearly changing the supply voltage of the Class E amplifier, the RF amplifier is fully switched on and off by a digital signal. The on and off switching can be done by a PWM, SOPA or $\Delta\Sigma$ modulator, and the architecture can be seen as a single-bit RF D-to-A converter. If the lattice-type LC balun is used, this idea can be extended to a multi-bit RF D-to-A converter.

Chapter 6

A CMOS POWER AMPLIFIER FOR GSM-EDGE

6.1 Introduction

This chapter will discuss the design and CMOS implementation of a linearized power amplifier [Reyn05a, Reyn05b]. The linearization is done by the polar modulation principle, discussed in the previous chapter.

The linearized amplifier is designed in a $0.18\,\mu\mathrm{m}$ CMOS technology and requires no expensive off-chip components. The presented solution is fully integrated and as such the bill of materials (BOM) is heavily reduced, resulting in a lower product cost. To demonstrate the industrial relevance of this research, the linearized amplifier targets and meets the GSM-EDGE specifications [rGPP01, ETSI01].

The EDGE system is first thoroughly discussed in section 6.2. A good knowledge of the system and signals that need to be amplified is crucial in power amplifier design. Section 6.3 will discuss the overall transmitter architecture and linearity requirements. For this, the linearity discussion of section 5.3 is repeated here, but now focussed on EDGE signals and the EDGE linearity requirements. From this discussion, guidelines are obtained for the design of the RF amplifier and the amplitude modulator. In section 6.4, the circuit design in the $0.18\,\mu\mathrm{m}$ CMOS technology is presented. Some layout aspects are covered in this section as well. The measurement setup, the packaging and the different measurements of the fully integrated amplifier are discussed in section 6.5. Based on the measurement results, some improvements of this solution are suggested in section 6.6 and final conclusions are summarized in section 6.8.

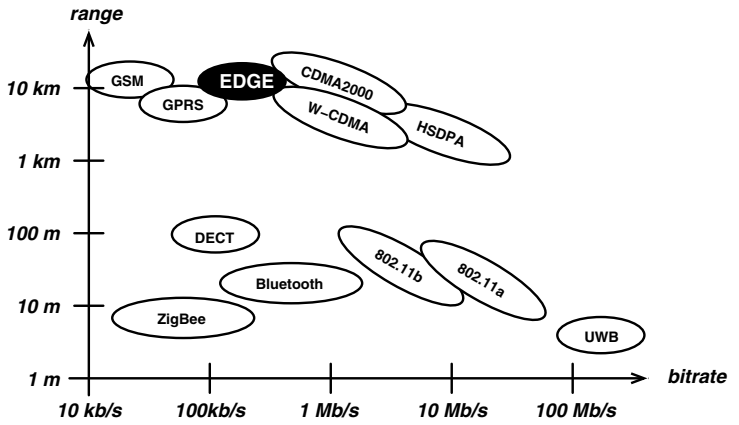


Figure 6.1. Position of the EDGE system between other wireless standards.

6.2 The EDGE System

This section will first discuss some key properties of the EDGE system and the generation of the EDGE signal. Next, the linearity and output power requirements for an EDGE mobile phone are given.

6.2.1 Enhanced Datarates for GSM Evolution

EDGE is an acronym for Enhanced Datarates for GSM Evolution, and is a follow up of the popular and well-known GSM system. Since the implementation of the first GSM network in 1991, the number of networks has increased rapidly and currently (2005) more than 200 countries have adopted this standard [Deut04].

The unexpected success of the Short Message Service (SMS) within the GSM standard was a clear indication of the need for other forms of communication. The mobile phone of today has become an intelligent device by which users can transmit and receive pictures and movies, read their e-mail and surf on the world wide web. To support these services, the network infrastructure and the mobile phone itself must be able to handle a much higher datarate compared to the $271\,\mathrm{kbps}$ raw datarate of GSM. Since 2002, telephone operators are implementing the Universal Mobile Telecommunication System (UMTS) which is called a third-generation system (3G). UMTS strives to reach a $2\,\mathrm{Mbps}$ datarate, which would enable streaming video and other multi-media applications. W-CDMA is the technology behind UMTS and is completely different from GSM. The large investment cost associated with the extension of an entirely new network and the huge prices that were paid for the UMTS li-

censes have delayed the implementation of the 3G networks in most European countries.

In this regard, the EDGE system gained large interest as it can re-use the existing GSM infrastructure and licenses. For backward compatibility, EDGE uses the same carrier frequency and channel spacing as GSM, but the datarate of EDGE is three times as high i.e. $812\,\mathrm{kbps}$ compared to GSM. As such, EDGE enables to quickly transmit and receive pictures, to do teleshopping and also video telephony is one of the capabilities. Therefore, EDGE is denoted as a 2.5G system, lying in between GSM (2G) and UMTS (3G). Some network operators even predict that, at least in some countries, the 3G UMTS system will not be implemented and some operators will jump from the 2.5G EDGE system to the 4G system, though 4G still requires standardization. EDGE thus enables operators to introduce 3G data applications and services into existing GSM networks. To emphasize this, some network operators denote EDGE as a 2.75G system. The number of commercial EDGE networks in operation worldwide reached 116 at the end of September 2005 and is available in 70 countries [Wir05].

The importance of EDGE is also reflected in the research community [McCu03, Hadj04, Sowl04, Elli04]. In the design of the mobile terminal, the backward compatibility between EDGE and GSM is an advantage as well, since an EDGE transmitter can be build from the existing GSM transmit architectures. Finally, figure 6.1 depicts the position of EDGE between other popular wireless standards.

6.2.2 Generation of the EDGE Signal

6.2.2.1 Symbol Mapping

The EDGE system uses an 8-PSK constellation diagram, as shown in figure 6.2(a). The serial bit stream is converted into 3-bit words and mapped on the 8-PSK constellation using Gray encoding. Adjacent symbols differ by only one bit and this minimizes the number of interpreted error bits when a symbol is incorrectly decoded as one of its nearest neighbors. The symbol rate of EDGE is the same as for GSM, i.e. $270.833\,\mathrm{kHz}$. The major difference between the two is that GSM uses a 1-bit per symbol GMSK encoding whereas EDGE maps three bits on one 8-PSK symbol.

The symbol mapping of EDGE can be written as

$$s_i = e^{j2\pi l/8} \tag{6.1}$$

where l is given by

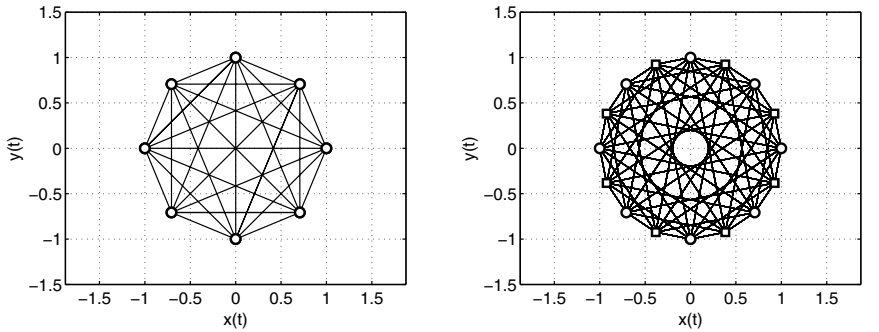


Figure 6.2. (a) 8PSK and (b) rotated 8-PSK modulation scheme.

modulating bits	l
(1,1,1)	0
(0,1,1)	1
(0,1,0)	2
(0,0,0)	3
(0,0,1)	4
(1,0,1)	5
(1,0,0)	6
(1,1,0)	7

6.2.2.2 Symbol Rotation

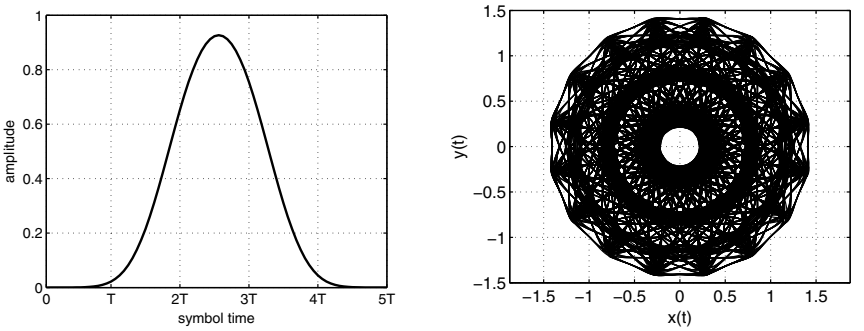
The EDGE symbols are continuously rotated with $3\pi/8$ radians per symbol before pulse shaping. The rotated symbols can be defined as

$$s_{i,r} = s_i \cdot e^{(j\, i\, 3\pi/8)} \tag{6.2}$$

This results in the rotated 8-PSK constellation diagram of figure 6.2(b). Due to the symbol rotation, the trajectories avoid the origin, and the dynamic range of the envelope of the transmitted RF signal is hereby reduced. Figure 6.2(b) can also be obtained if one uses two 8-PSK modulation diagrams that are rotated by an angle of $3\pi/8$. The signal is then alternately mapped on the two 8-PSK constellation diagrams.

6.2.2.3 Pulse Shaping

The transition from one constellation point to another does not occur immediately as this would require a very large bandwidth. A baseband filter is applied to limit the bandwidth of these transitions. For EDGE, the filter is



(a) (b)

Figure 6.3. (a) Impulse response of the EDGE pulse filter and (b) filtered EDGE constellation diagram.

a Gaussian filter [rGPP01] of which the impulse response is shown in figure 6.3(a). A symbol at time 0 is thus spread out over five consecutive symbol periods. In other words, severe inter-symbol interference will occur and in the receiver, the correct demodulation filter needs to be applied in order to cancel this ISI.

The instantaneous filtered baseband signal is thus the result of five consecutive symbols, and therefore, the filtered signal is difficult to interpret. Also note that the ISI will increase the dynamic range of the signal. In other words, the eye around the origin has become smaller due to the pulse shaping filter. The resulting filtered constellation diagram is shown in figure 6.3(b).

6.2.2.4 EDGE Signal Characteristics

To conclude the modulation process, the I(t) and Q(t) signals that represent the complex symbols, are upconverted to the RF carrier. In total, there are nine frequency bands defined by the GSM-EDGE Standard, ranging from $450\,\mathrm{MHz}$ up to $1.9\,\mathrm{GHz}$. The most popular ones are the primary GSM-900 band (P-GSM), the extended GSM-900 band (E-GSM[1]), the DCS-1800 band and the PCS-1900 band. The amplifier discussed in this chapter is designed for the DCS-1800 band, of which the corner frequencies are equal to

mobile transmit, base station receive	mobile receive, base station transmit
1710 MHz - 1785 MHz	1805 MHz - 1880 MHz

[1]E-GSM is often confused with EDGE. E-GSM only specifies the frequency band; within this band both constant envelope GMSK modulation or non-constant envelope 8-PSK modulation can be applied.

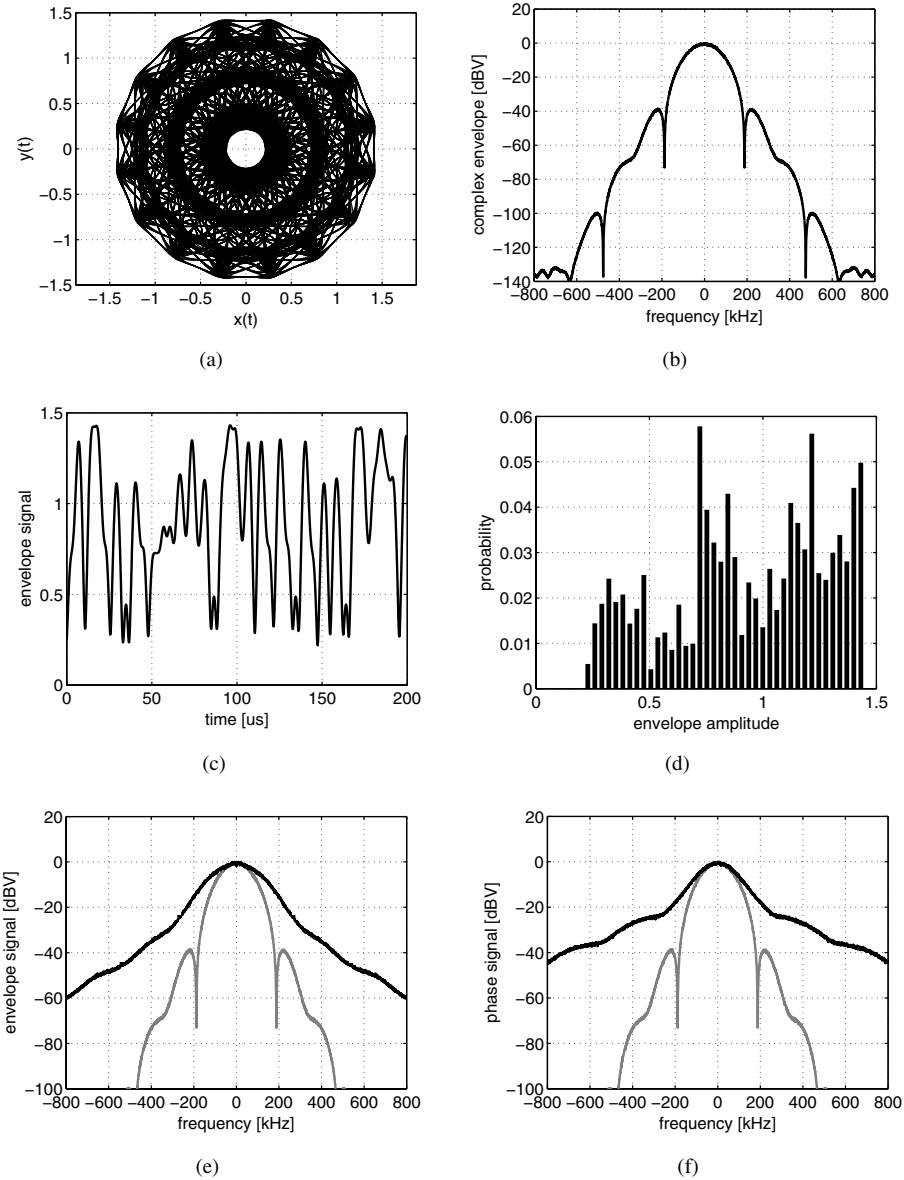


Figure 6.4. Overview of the EDGE signals: (a) filtered constellation diagram, (b) spectrum of the complex envelope, (c) time behavior of the envelope signal and (d) distribution of the envelope signal, (e) spectrum of the envelope signal and (f) spectrum of the complex phase signal.

Figure 6.4(a) depicts the filtered constellation diagram which is equal to the complex envelope defined in section 2.2.1. The spectrum of the complex envelope, shown in figure 6.4(b), is equal to the actual transmitted spectrum at the RF frequency. The envelope signal of EDGE is shown in figure 6.4(c), and one can clearly see that this is a non-constant envelope modulation. In section 2.2.3, the *PDF* of the envelope signal was introduced as an important signal property. The *PDF* can be approximated by the histogram, shown in figure 6.4(d). Note that the envelope signal avoids the origin. From this histogram, it can be calculated that the envelope of the EDGE signal has a dynamic range of $1/7$ or $17\,\mathrm{dB}$ and the PAPR or Crest Factor is equal to $3.4\,\mathrm{dB}$. Finally, figures 6.4(e) and 6.4(f) depict the spectrum of the envelope signal and the spectrum of the complex phase signal, defined as $e^{jP(t)}$. As already mentioned in section 5.3.6 and clarified by figure 5.19, the spectrum of these individual signals is much broader compared to the spectrum of the modulated RF signal of figure 6.4(b).

The channel spacing of EDGE is the same as for GSM, i.e. $200\,\mathrm{kHz}$. The EDGE spectrum is similar to the GSM spectrum and fits nicely within the $200\,\mathrm{kHz}$ spacing. However, the datarate of EDGE is three times higher compared to that of GSM, despite the same channel bandwidth. The higher datarate of EDGE is due to the 8-PSK modulation, the envelope variations of the output signal and the severe pulse shaping filter. The drawbacks are that a linear amplifier is needed in the transmitter and a more complex baseband filter is needed in the receiver to neutralize the inter-symbol interference.

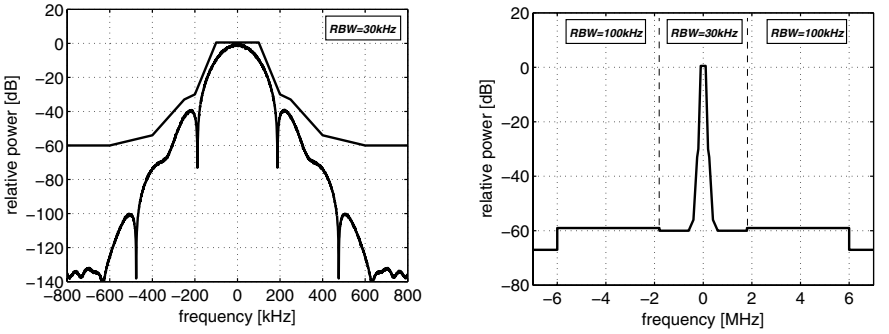
6.2.3 EDGE Transmitter Linearity Requirements

The envelope variation of the EDGE signal clearly requires a linear or linearized amplifier and the GSM-EDGE Standard specifies the amount of distortion or nonlinearity that can be tolerated. Unfortunately for the circuit designer, the GSM-EDGE Standard does not specify the minimal required IMD or HD of the amplifier, but it uses system level requirements to describe the maximal allowable distortion. The two most important requirements are given below.

6.2.3.1 Spectral Mask Requirement

A first linearity requirement is given by the spectral mask. The spectral mask requirement is specified at the power amplifier output and ensures that the transmitter does not corrupt or block the spectrum of neighboring channels.

For the DCS-1800 frequency band, this mask is shown in figure 6.5(a). Of course, the ideal EDGE spectrum falls well within this mask. The GSM-EDGE Standard also specifies a wideband spectral mask, shown in figure 6.5(b) for a $24\,\mathrm{dBm}$ average output power. The noise power measured outside the transmit band is specified as well.



(a) (b)

Figure 6.5. EDGE spectral mask requirements.

6.2.3.2 EVM Requirement

After demodulation by an ideal receiver, the received constellation point will not correspond exactly with the transmitted constellation point. The distortion and nonlinearities of the transmitter will create little clouds of demodulated symbols, located near the ideal constellation points. An example is shown in figure 6.6(a). The error vector is defined as the difference between the ideal constellation point and the actual transmitted constellation point, as shown in figure 6.6(b). The error vector can be written as

$$E(k) = T(k) - S(k) \tag{6.3}$$

in which $T(k)$ is the actual transmitted constellation point or symbol, $S(k)$ is the ideal symbol and $E(k)$ is the error vector that connects these two points. Note that these three variables are complex.

the Error Vector Magnitude (EVM) of symbol k is defined as

$$EVM = \sqrt{|E(k)|^2 \,/\, \left(\frac{1}{N} \cdot \sum_k |S(k)|^2\right)} \tag{6.4}$$

and the RMS value of the Error Vector Magnitude is defined as

$$EVM_{RMS} = \sqrt{\sum_k |E(k)|^2 \,/\, \sum_k |S(k)|^2} \tag{6.5}$$

The GSM-EDGE Standard specifies that the RMS value of the EVM should be lower than 9% and the peak value of the EVM should be lower than 30%.

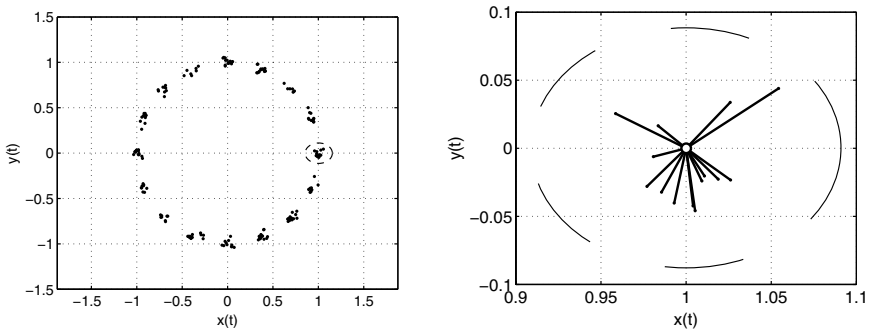


Figure 6.6. (a) Distorted EDGE constellation diagram. (b) Close-up of the error vectors around the ideal constellation point (1,0).

Note that the Error Vector Magnitude is specified after reception and demodulation by an ideal receiver and ensures a correct transmission within the channel, whereas the spectral mask requirement ensures that the transmitter does not interfere with neighboring channels.

6.2.4 EDGE Transmitter Output Power Requirements

The EDGE signal contains amplitude modulation and the output power is thus defined as the long term average output power over several symbols. Remember that the peak to average power ratio or crest factor of EDGE is equal to $3.4\,\mathrm{dB}$, whereas it is $0\,\mathrm{dB}$ for GSM. For the DCS-1800 frequency band, three power classes are defined:

power class	average output power
E1	$30\,\mathrm{dBm}$ - $1\,\mathrm{W}$
E2	$26\,\mathrm{dBm}$ - $400\,\mathrm{mW}$
E3	$22\,\mathrm{dBm}$ - $160\,\mathrm{mW}$

The GSM-EDGE Standard also requires that the output power can be regulated down to $0\,\mathrm{dBm}$ with $2\,\mathrm{dB}$ steps, in order to increase the battery lifetime.

6.3 A Polar Modulated Power Amplifier for EDGE

To amplify the EDGE signal, a linear or linearized amplifier is required. In this chapter, a Class E amplifier is linearized by means of polar modulation, a technique that was thoroughly discussed in chapter 5.

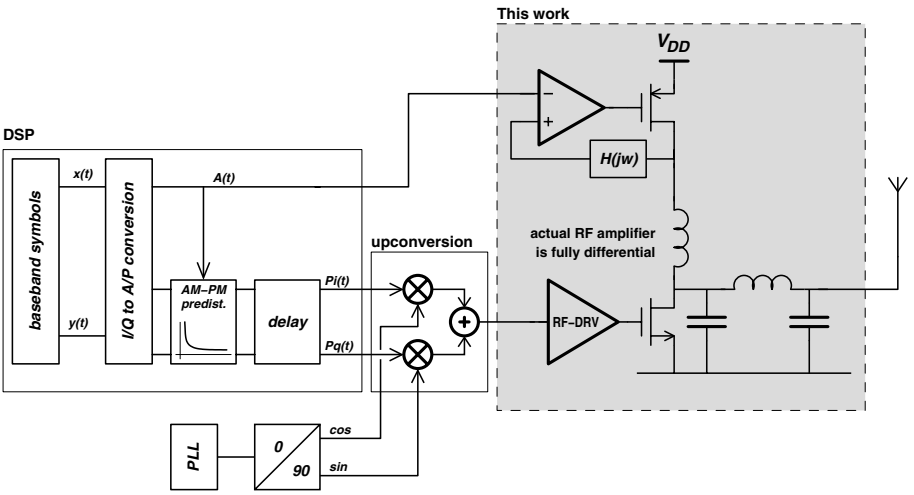


Figure 6.7. Architecture of the EDGE CMOS power amplifier.

The first section will discuss the advantages of the selected architecture. Next, the degradation of the EDGE signal, due to the different distortion mechanisms of polar modulation, is demonstrated.

6.3.1 Architecture

The EDGE system is backwards compatible with GSM. In fact, a mobile phone that is able to transmit and receive EDGE signals should also be able to handle constant envelope GSM signals. In this regard, polar modulation results in a linearized architecture that is backwards compatible with a constant envelope architecture. After all, a polar modulated power amplifier uses a constant envelope RF path, and thus for the phase modulation and RF upconversion, the same architecture as for GSM can be used [Elli04, Sowl04].

Figure 6.7 depicts the selected architecture to amplify the EDGE signal. The focus of this work is on the power amplifier itself, indicated by the gray box. The other blocks are not integrated on this chip since the CMOS integration of these blocks has already been demonstrated several years ago [Stey98, Stey00]. Also note that a polar modulated architecture easily allows the implementation of AM-PM predistortion in the DSP [Sand03].

The envelope of the EDGE signal has a peak to average ratio of about $3.4\,\mathrm{dB}$. The dynamic range or the peak to minimum ratio of the envelope, equals 7 or $17\,\mathrm{dB}$. Furthermore, if one looks to figure 6.4(b), it can be seen that the bandwidth of the envelope signal is much broader compared to that of

the complex signal. It will be shown in section 6.3.2.4 that the envelope bandwidth should be at least 3 to $5\,\mathrm{MHz}$, and this poses serious problems to use a switching amplitude modulator. After all, note that the spectral mask is $60\,\mathrm{dB}$ below the carrier. Roughly this requires more than $60\,\mathrm{dB}$ of attenuation of the switching frequency. For an LC output filter with a bandwidth of $4\,\mathrm{MHz}$, the switching frequency should be at least $160\,\mathrm{MHz}$ to have $60\,\mathrm{dB}$ of suppression. Delivering several watts of power at such a high switching frequency is not an obvious task. Furthermore, the losses associated with this high frequency will make the efficiency benefit of the switching modulator less obvious. From this, a linear amplitude modulator is selected that is able to combine both a high bandwidth and a large linear dynamic range. But the most important argument to select a linear amplitude modulator, is the high level of integration that can be obtained. In fact, the developed solution requires no expensive off-chip components. For the RF amplifier, a Class E amplifier is selected since this enables a fully integrated solution in CMOS.

6.3.2 Distortion

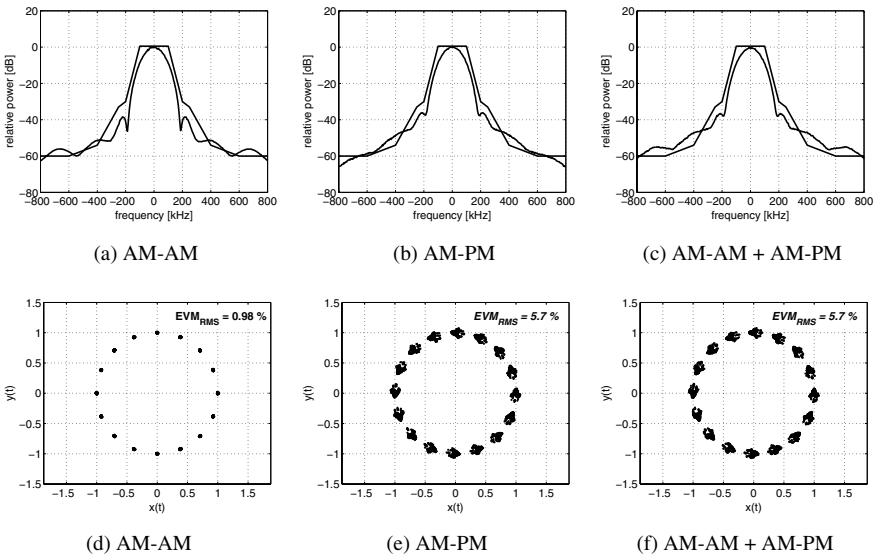
In section 5.3, the different distortion mechanisms of polar modulation were discussed. The influence of the nonlinearities are now investigated for the EDGE system.

6.3.2.1 AM-AM and AM-PM Distortion

The feedforward from the driver stage to the Class E amplifier stage will cause both AM-AM and AM-PM distortion. The AM-AM and AM-PM curves, obtained from the simplified network of figure 5.10, are used here again to investigate the influence on the EDGE signal. The resulting spectra and constellation diagrams are shown in figure 6.8. The corresponding value of the EVM_{RMS} is given in the constellation diagrams, and this value should be lower than 9%.

From these figures, it can be seen that the EVM_{RMS} is mainly dependent on the AM-PM distortion, and the spectral distortion is equally dependent on both the AM-AM and AM-PM distortion. Also, the spectral mask requirements are much harder to meet than the EVM_{RMS} requirement. This is not necessarily true for other wireless standards. The W-LAN standards (IEEE 802.11a,b,g) are typical examples for which the EVM requirement can be harder to meet than the spectral mask requirement [Zarg02], [Behz03], [Zhan03], [Zarg04].

It is also important to realize that the feedforward signal is only phase modulated and hence it has a wide bandwidth [Elli04]. Therefore, the feedforward signal itself will also broaden the output spectrum, though this is not caused by a nonlinearity. This effect is not included in figure 6.8.



(a) AM-AM (b) AM-PM (c) AM-AM + AM-PM

(d) AM-AM (e) AM-PM (f) AM-AM + AM-PM

Figure 6.8. Influence of AM-AM and AM-PM distortion for $C_{gd} = 3\,\mathrm{pF}$.

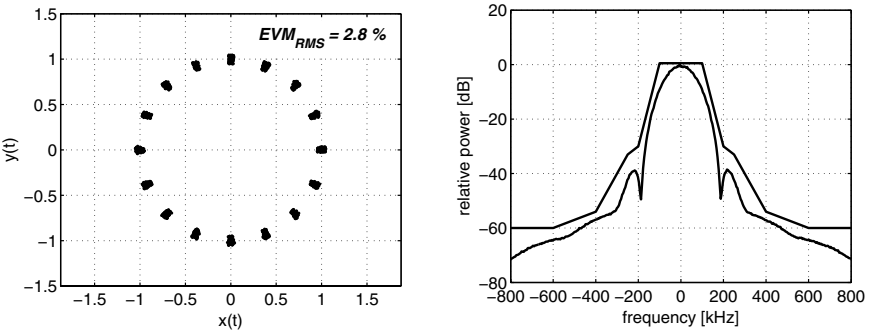
6.3.2.2 Maximum Differential Delay

Another source of distortion in this architecture is a delay between the amplitude and phase path. Such a delay will result in a misalignment when amplitude and phase are recombined in the RF power amplifier. Figure 6.9 depicts the result for the EDGE system with an arbitrary delay of $58\,\mathrm{ns}$ between the amplitude and phase. Clearly, the spectrum at $400\,\mathrm{kHz}$ and $600\,\mathrm{kHz}$ frequency offset are the most critical points and at these points, the relative output power should be $-54\,\mathrm{dB}$ and $-60\,\mathrm{dB}$.

Figure 6.10(b) shows the output spectrum at $400\,\mathrm{kHz}$ and $600\,\mathrm{kHz}$ frequency offset versus the delay and figure 6.10(a) depicts the simulated EVM_{RMS}. From the EVM requirement, a delay of $200\,\mathrm{ns}$ can be tolerated. The more stringent spectral mask requirement however, requires a delay of less than $75\,\mathrm{ns}$.

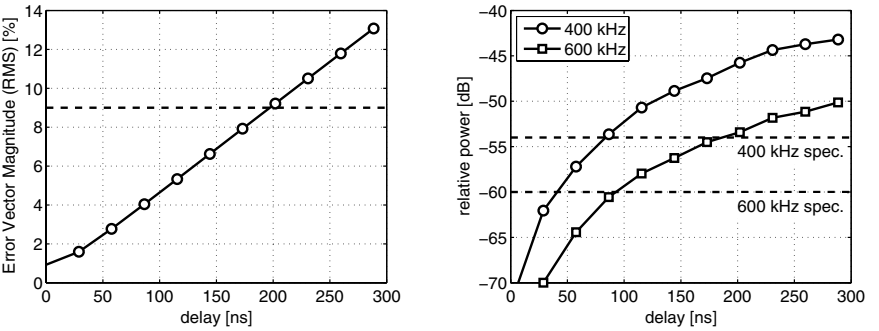
6.3.2.3 Spectral Asymmetry

An interesting effect occurs if AM-PM distortion and delay occur simultaneously. Figure 6.11(c) shows the output spectrum if these two effects occur together, and reveals an asymmetry in the output spectrum which violates the spectral mask. In 6.11(c), the envelope is delayed with respect to the phase, and this causes a higher spectrum at the left side. If the phase is delayed with respect to the envelope signal, the spectrum is higher on the right side.



(a) (b)

Figure 6.9. Influence of a 58 ns delay on (a) the constellation diagram and (b) the output spectrum.

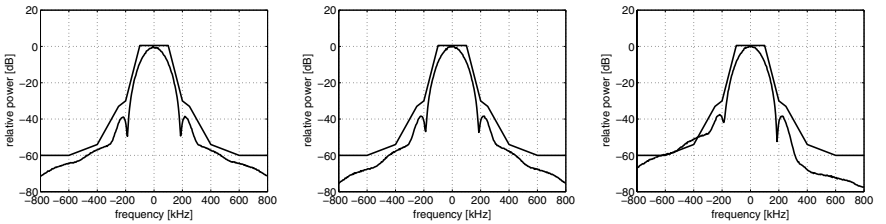


(a) (b)

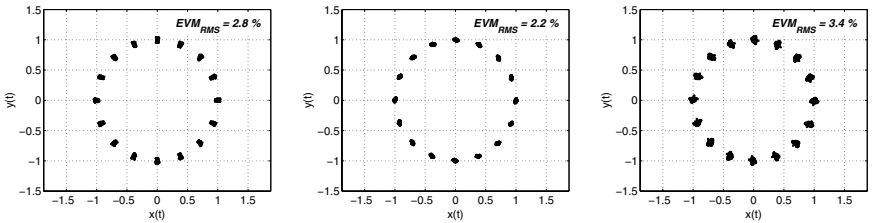
Figure 6.10. (a) RMS value of the Error Vector Magnitude and (b) relative power at 400 kHz and 600 kHz frequency offset, versus delay between amplitude and phase.

6.3.2.4 Amplitude Modulator Bandwidth

An important design specification is the required bandwidth of the amplitude modulator. It was already shown in figure 6.4(b) that the envelope signal has a much broader bandwidth compared to the RF signal. Figure 6.12(a) shows the simulated complex output spectrum when the envelope signal is filtered by a first order Butterworth low pass filter with a -3 dB frequency of 1 MHz. The output spectrum at 400 kHz and 600 kHz frequency offset are the most critical points and at these points the spectral mask is violated. The corresponding RMS value of the EVM is 6%, which is still below the required

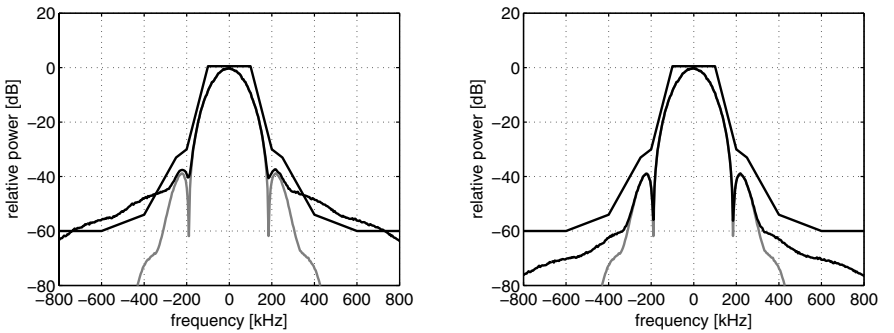


(a) 58ns delay (b) AM-PM (c) both



(a) 58 ns delay (b) AM-PM (c) both

Figure 6.11. The combination of both differential delay and AM-PM distortion causes an asymmetry in the output spectrum.



(a) (b)

Figure 6.12. EDGE spectrum after an envelope filter of 1 MHz, undistorted EDGE spectrum (gray) and spectral mask requirement. In (b) a delay of 144 ns is inserted in the phase path.

9%. Remember that the distortion of the envelope filtering can partially be overcome by adding an additional delay in the phase path, as demonstrated in section 5.3.6. As an example, figure 6.12(b) shows the output spectrum for the same 1 MHz envelope filter but with a delay of 144 ns in the phase path.

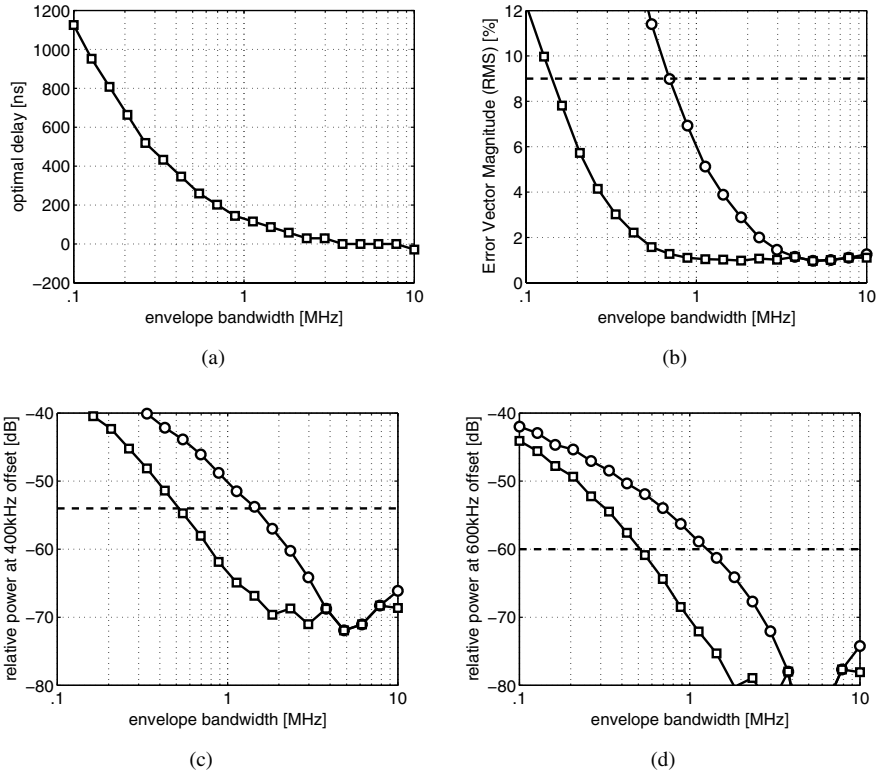


Figure 6.13. Influence of the envelope bandwidth on the EDGE signal, with (□) and without (○) a delay in the phase path; (a) optimal delay, (b) RMS value of the EVM, (c) relative power at 400 kHz offset and (c) relative power at 600 kHz offset.

This value was obtained by a simple optimization algorithm. With this *optimal* delay in the phase path, the spectral mask is completely met.

Figure 6.13 shows the *optimal* delay, the EVM_{RMS} and the spectral mask margin, versus the bandwidth of the amplitude modulator. Based upon the EVM requirement, a bandwidth of only 700 kHz would be sufficient. If a delay is inserted, the bandwidth can be reduced down to 150 kHz. However, the stringent spectral mask requirement sets the bandwidth to at least 1.5 MHz and can be reduced down to 500 kHz with the optimized delay. To provide sufficient margin for the other nonlinearities, the bandwidth of the amplitude modulator should be about 3 MHz.

6.4 Circuit Implementation

From the GSM-EDGE Standard and the distortion analysis of the previous section, the requirements of the different blocks can be derived. Some circuit details of the RF amplifier and the linear amplitude modulator are now discussed.

6.4.1 Design of the RF amplifier

The polar modulated power amplifier is designed in a $0.18\,\mu\mathrm{m}$ CMOS technology with analog and RF options. The analog option makes it possible to use low and zero V_T $0.18\,\mu\mathrm{m}$ AMOS transistors, thick gate-oxide transistors, that are similar to $0.35\,\mu\mathrm{m}$ transistors, and high density MiM capacitors ($1.1\,\mathrm{fF}/\mu\mathrm{m}^2$). The RF option means that the top metal aluminum layer has a thickness of $4\,\mu\mathrm{m}$ and a resistivity of only $20\,\mathrm{m}\Omega\square$. This is necessary to implement high quality inductors.

The GSM-EDGE Class E3 specifications require a maximum modulated output power of $22\,\mathrm{dBm}$. Since the peak to average power ratio for EDGE is $3.4\,\mathrm{dB}$, the RF amplifier must thus be able to transmit a peak power of at least $25.4\,\mathrm{dBm}$. To increase both the reliability and the output power, a thick gate-oxide transistor, with an oxide thickness of $6.5\,\mathrm{nm}$ and a zero current drain-bulk breakdown voltage of $9\,\mathrm{V}$, was selected for the nMOS switch in the Class E output stage. This transistor is actually a $0.35\,\mu\mathrm{m}$ transistor, with a minimal gate length of $0.34\,\mu\mathrm{m}$. The L-match impedance transformation network at the output is designed to transform the $50\,\Omega$ antenna impedance to $R_m = 12\,\Omega$. Together with the thick gate-oxide transistor, this will meet the output power requirement.

Two driver stages are added in order to achieve sufficient gain and to lower the input capacitance seen by the upconverter. Since the main task of the driver stages is to deliver gain, the corresponding transistors are minimum gate-length $0.18\,\mu\mathrm{m}$ transistors that benefit from the higher f_T.

The amplifier is made fully differential to increase the output power, to reduce the required on-chip decoupling capacitance and to reduce the influence of the parasitic ground inductance. A dedicated ground is provided for the differential RF output, to control the inductance of the RF signal path. A single-ended version of the complete RF amplifier circuit is shown in figure 6.14.

As stated in section 3.3.7, the Class E operating regime is not the optimum regarding efficiency and output power because of the losses in the circuit. To investigate this effect, both the values of L_1 and C_1 are varied, the circuit is simulated with a spice-like simulator and the output power and efficiency are calculated. The result of this optimization process is shown in figure 6.15.

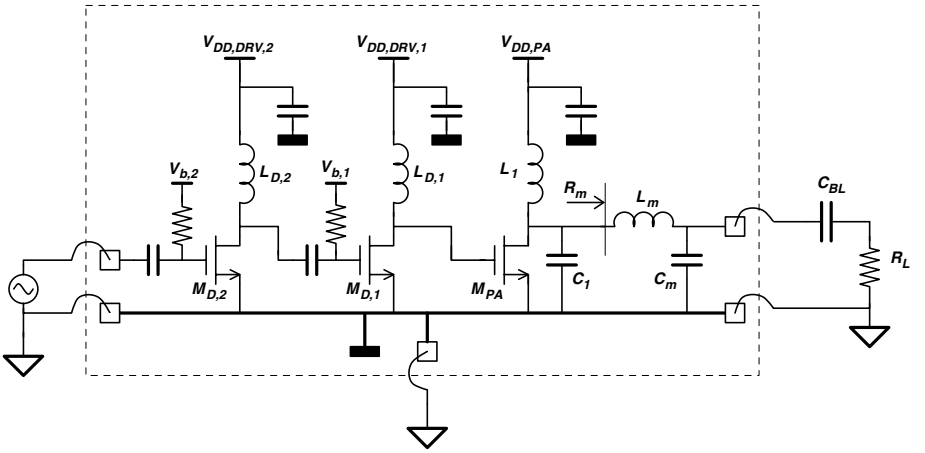
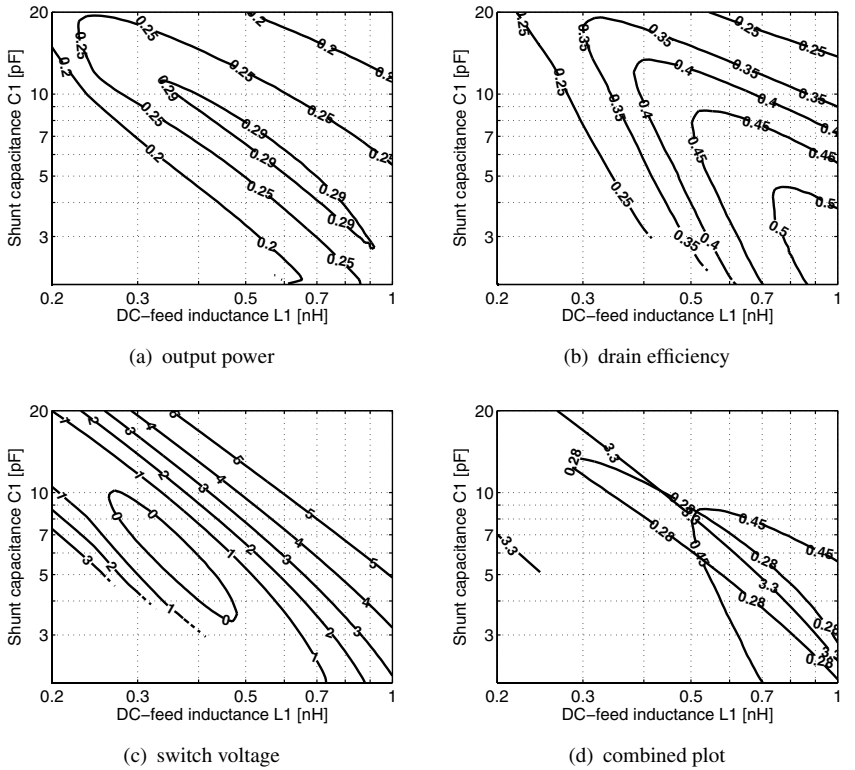


Figure 6.14. Single-ended version of the RF Class BE amplifier.

Figures 6.15(a) and 6.15(b) show the output power and efficiency. In each point, the supply voltage is adjusted to keep the maximum drain voltage below 7 V. Figure 6.15(c) depicts the switch-in voltage, which is defined as the drain voltage when the switch closes and the drain current starts to flow. As such, Class E operation is defined by the contour line of 0 V switch-in voltage. Safe operation is guaranteed as long as the switch-in voltage remains below 3.3 V, which is the nominal supply voltage of the thick gate-oxide transistor. Clearly, the Class E contour is located away from both the maximum output power and maximum efficiency. Figure 6.15(d) combines the three previous plots.

From these figures, L_1 is chosen to be $0.5\,\mathrm{nH}$ and $C_1 = 6\,\mathrm{pF}$. Slightly better performance could be achieved for a higher DC-feed inductance and lower shunt capacitance. However, one should provide some margin for the parasitic capacitance of the inductor and metal interconnections.

The contours of figure 6.15 were calculated for several transistor widths. A width of $3000\,\mu\mathrm{m}$ results in a maximal overall efficiency and was therefore chosen in this design. The resulting on-resistance is $0.5\,\Omega$ at a gate voltage of 3.3 V. The corresponding input capacitance of the thick gate-oxide transistor, taking into account the Miller effect, is $11.5\,\mathrm{pF}$. The simulations predict a maximum differential output power of $560\,\mathrm{mW}$ or $27.5\,\mathrm{dBm}$ with a drain efficiency of 44%. The gate-drain capacitance of this transistor equals $3.3\,\mathrm{pF}$ and thus it can be expected that this will cause too much AM-PM distortion. Therefore, predistortion will be needed to alleviate this and will be demonstrated during the discussion of the measurement results.



(a) output power

(b) drain efficiency

(c) switch voltage

(d) combined plot

Figure 6.15. Contour plots of the single-ended Class E output stage.

To drive the $11.5\,\mathrm{pF}$ capacitance of the nMOS switch, the driver stages consume a total DC power of $140\,\mathrm{mW}$. The RF input power of the first driver stage is only $-3\,\mathrm{dBm}$. The upconverter can thus directly be connected to the power amplifier. The first driver stage has an nMOS with a gate width of $M_{D,1} = 800\,\mu\mathrm{m}$, a gate length of $0.18\,\mu\mathrm{m}$, a DC-feed inductance of $L_{D,1} = 0.5\,\mathrm{nH}$ and a supply voltage of $1.56\,\mathrm{V}$. The second driver stage has an nMOS with a gate width of $M_{D,2} = 100\,\mu\mathrm{m}$, a gate length of $0.18\,\mu\mathrm{m}$, a DC-feed inductance of $L_{D,2} = 2.9\,\mathrm{nH}$ and a supply voltage of is equal to $1.4\,\mathrm{V}$. The supply voltages of the driver stages are chosen low enough to avoid hot carrier generation.

The actual RF amplifier is fully differential, as shown in figure 6.16. In this figure, the decoupling capacitances have been omitted. The DC-feed inductors of the Class E stage and the driver stages are merged together to form differential inductors with a center tap. The main reason is to save silicon area. The differential inductors are modeled and simulated as full four port networks, since the currents in the inductor are not fully differential due to the switching nature of the Class E amplifier.

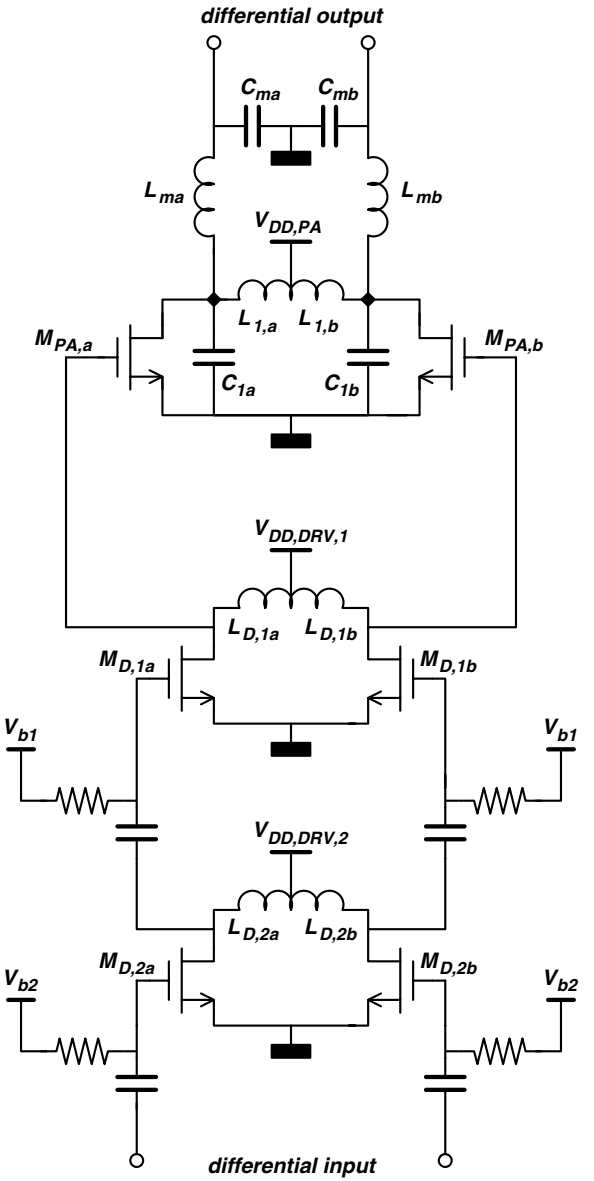


Figure 6.16. Differential structure of the Class BE amplifier. The decoupling has been omitted.

Table 6.1. Component values of the RF power amplifier

parameter	value
gate width of M_{PA}	$3000\,\mu$m
gate length of M_{PA}	$0.34\,\mu$m
C_1	6 pF
$L_{1,a} + L_{1,b}$	differential coil of 1 nH
L_m	1.7 nH
C_m	3.2 pF
gate width of $M_{D,1}$	$800\,\mu$m
gate length of $M_{D,1}$	$0.18\,\mu$m
$L_{D,1a} + L_{D,1b}$	differential coil of 1 nH
gate width of $M_{D,2}$	$100\,\mu$m
gate length of $M_{D,2}$	$0.18\,\mu$m
$L_{D,2a} + L_{D,2b}$	differential coil of 5.8 nH

All the inductors of the RF amplifier are integrated on-chip. The differential topology and the integrated impedance transformation network will result in a solution that is less sensitive to the used package. Furthermore, since no off-chip striplines or inductors are required, the total board area is reduced. Inductors L_1, L_m and $L_{D,1}$ have a quality factor Q_L of about 11.

Figure 6.17 shows the drain voltage and drain current of the Class E stage. There is an overlap between voltage and current, typically for a Class BE amplifier. However, when current starts to flow, the drain voltage is already below 3 V.

To conclude, table 6.1 summarizes the component values of the RF power amplifier.

6.4.2 Design of the Linear Amplitude Modulator

Figure 6.18 shows the simplified circuit of the amplitude modulator. For the design of this block, the RF power amplifier is first simulated for different values of $V_{DD,PA}$, the supply voltage of the Class E amplifier. From these transient RF simulations, it was found that the equivalent DC load resistance of the RF amplifier is $9\,\Omega$. R_{PA} is related to the efficiency of the switching amplifier. In an ideal case, the efficiency of the RF PA is independent of the supply voltage $V_{DD,PA}$ and therefore R_{PA} should remain constant. In a real implementation however, the equivalent DC load resistance is not constant and this variation will cause AM-AM distortion of the RF output signal. To reduce this distortion, the amplitude modulator is designed to have a very low output

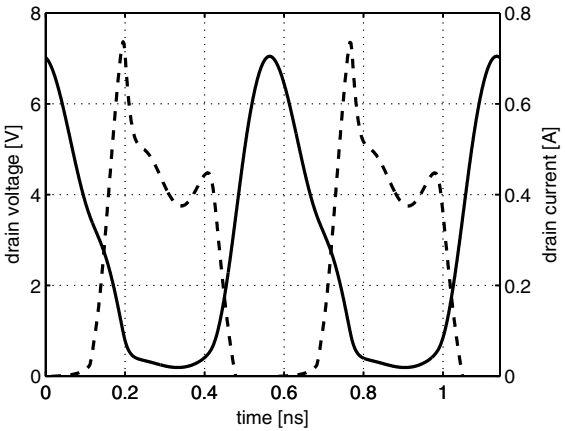


Figure 6.17. Drain voltage (solid line) and drain current (dashed line) of the Class BE output stage.

resistance, which is controlled by the loopgain and the output conductance of the large pMOS. It was found in section 5.2.4.2 that the DC output impedance is equal to

$$Z_{out,AM} \approx \frac{1}{A_0 H_0 g_m} \tag{6.6}$$

The output voltage of the amplitude modulator almost goes rail to rail, i.e. from zero to $V_{DD,AM}$. Therefore, the feedback block $H(j\omega)$ should act as an attenuator to limit the dynamic range of the feedback voltage, and a DC shift is necessary to bring the feedback voltage within the voltage range of the input differential pair. The output of the opamp, and thus the voltage at the gate of the pMOST M_{AM}, must also have a large dynamic range. To shut down the power amplifier, the opamp output must be made equal to $V_{DD,AM}$. At full output power, the output of the opamp must go as low as possible to reduce the on-resistance of M_{AM}, and thus to reduce the drain-source voltage drop of M_{AM}.

To avoid distortion due to envelope filtering, the opamp is designed to have a large bandwidth, rather than a large gain. Therefore, to have a low output impedance, the pMOST M_{PA} should have a large g_m which reflects itself in the large gate width of $8000\,\mu\mathrm{m}$. The gate length is $0.34\,\mu\mathrm{m}$ and the gate capacitance of that transistor is $15\,\mathrm{pF}$. Together with the output impedance of the OTA, this creates a dominant pole at $8.8\,\mathrm{MHz}$, which is high enough for the envelope signal. The non-dominant pole at the output node is created by $R_{PA} = 9\,\Omega$ and $C_{dec,2} = 80\,\mathrm{pF}$, resulting in $222\,\mathrm{MHz}$. The OTA has a voltage gain of $28\,\mathrm{dB}$ and a GBW of $260\,\mathrm{MHz}$. The pMOST has a transconductance

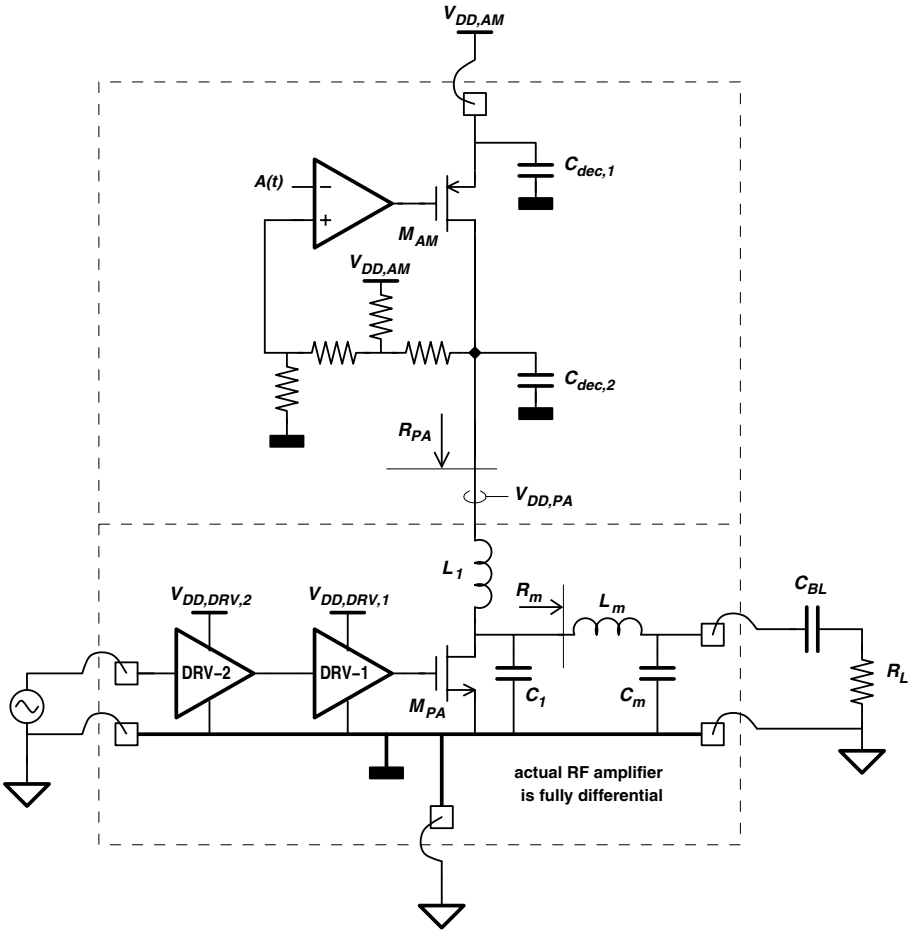


Figure 6.18. Circuit implementation of the fully integrated linear amplitude modulator.

g_m of $440\,\mathrm{ms}$ and together with $R_{PA} = 9\,\Omega$, this creates a gain of 4 or $12\,\mathrm{dB}$. A resistive feedback with an attenuation of $-18\,\mathrm{dB}$ is chosen to have a closed loop transfer function with a flat phase response. All this results in a loopgain of $22\,\mathrm{dB}$ and a phase margin of 78 degree, which is sufficient to safely close the loop. From the above figures, the closed loop output impedance of the amplitude modulator can easily be calculated and equals $0.72\,\Omega$. The bandwidth of this modulator is also wide enough to avoid degradation of the RF spectrum.

The supply voltage of the amplitude modulator is $3.3\,\mathrm{V}$ and has a decoupling capacitance of $C_{dec,1} = 110\,\mathrm{pF}$. When EDGE signals are transmitted at

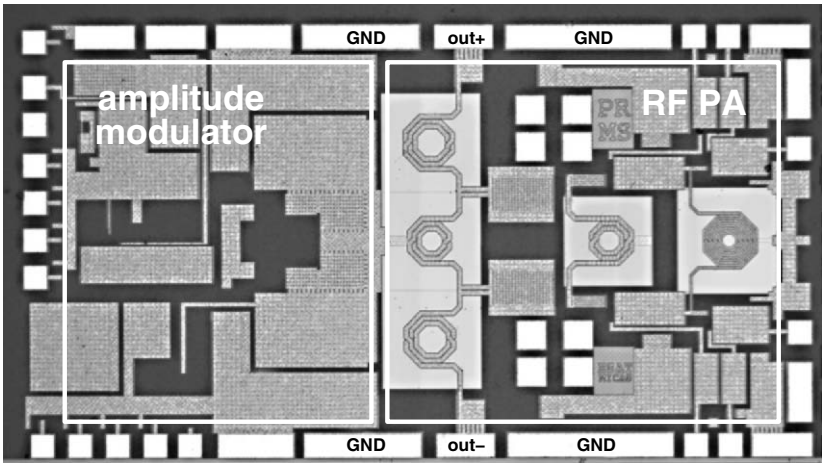


Figure 6.19. Photograph of the linearized EDGE PA in $0.18\,\mu\mathrm{m}$ CMOS.

maximum output power, the actual supply voltage of the RF PA, $V_{DD,PA}$ in figure 6.18, has a peak value of $2.9\,\mathrm{V}$, an average value of $1.9\,\mathrm{V}$ and a minimal value of $0.4\,\mathrm{V}$.

6.4.3 Layout Aspects

The amplitude modulator is integrated on the same chip as the RF PA. Figure 6.19 shows a photograph of the fully integrated linearized amplifier. The total chip area including bonding pads is $1.8\,\mathrm{mm}$ by $3.6\,\mathrm{mm}$. Wide bonding strips are used for the ground and supply connections. Sufficient decoupling is integrated on-chip to ensure stability and to provide a low impedance to the AC supply currents of the RF power amplifier.

The $0.18\,\mu\mathrm{m}$ CMOS process has six metal layers with a thick top metal layer. The inductors of the Class E stage are composed of $18\,\mu\mathrm{m}$ wide traces of the top five metal layers in parallel. The lowest metal layer is not used in order to decrease the parasitic capacitance to ground.

6.5 Measurements

6.5.1 Measurement Setup

The $0.18\,\mu\mathrm{m}$ CMOS power amplifier is mounted on a Al_2O_3 ceramic substrate and packaged in a $CuBe$ box. This allows to firmly connect the SMA connectors that carry the RF signals. A photograph of this setup is shown in figure 6.20.

To perform the measurements, both the low frequency amplitude and phase modulated RF signal need to be applied. The EDGE measurements require

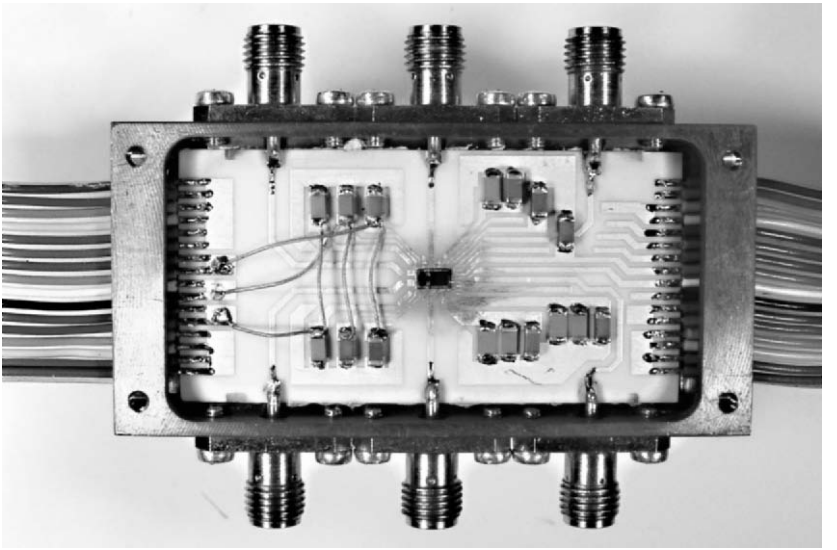


Figure 6.20. Photograph of the chip, mounted on a Al_2O_3 ceramic substrate and placed in a $CuBe$ box.

that the EDGE baseband $I(t)$ and $Q(t)$ signals are converted into an $A(t)$ and $P(t)$ format, as shown in figure 6.7. In a real implementation, this conversion is done in a cordic or in the DSP. For the measurements, the generation and the conversion are done in MATLAB.

Figure 6.21 depicts the setup and used equipment for the measurements. The complex baseband signal, represented by $I(t)$ and $Q(t)$, is created and converted to $A(t)$ and $P(t)$ in MATLAB. For a correct quadrature upconversion, $P(t)$ is represented by an in-phase and quadrature component $P_i(t)$ and $P_q(t)$. The three baseband signals are transmitted to the memory of an arbitrary waveform generator with three channels (Tektronix AWG-430). This ensures a correct time alignment of all three signals. The envelope signal is directly applied at the input of the linear amplitude modulator. The phase signal is fed to an external upconverter (Rohde & Schwarz SMIQ-06B). The output is a constant envelope phase modulated signal at $1.75\,\mathrm{GHz}$. This signal is made differential by a 0/180 degree power splitter/combiner (not shown here) and applied to the differential RF amplifier.

The differential output of the RF amplifier is first converted to a single-ended signal by the same 0/180 degree power splitter/combiner (not shown here). The signal is then fed through a 10dB RF attenuator (not shown here) before being applied to a signal analyzer (Rohde & Schwarz FSIQ-26) or an RF peak and average power meter (Rohde & Schwarz NRVS). The signal analyzer

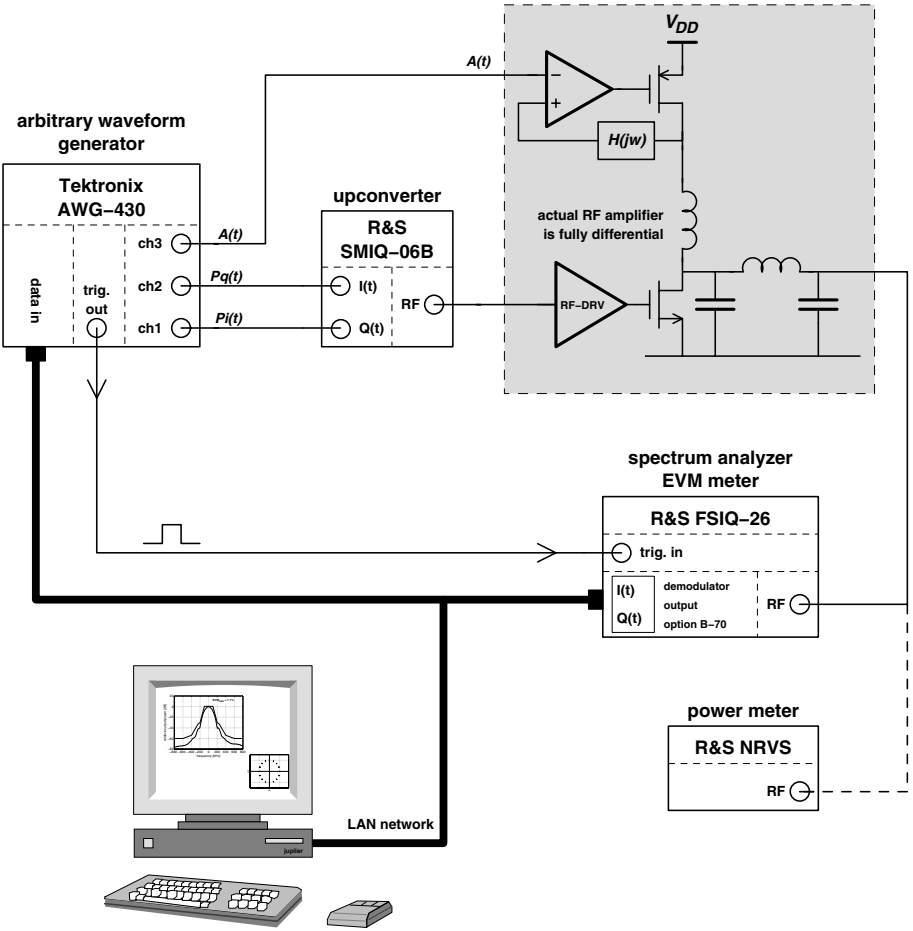


Figure 6.21. Measurement setup.

allows to calculate the spectrum of the amplified signal, to calculate the Error Vector Magnitude and to demodulate the signal and store the $I(t)$ and $Q(t)$ in the internal memory that can be read out by a computer.

6.5.2 Constant Envelope Measurements

To measure the constant envelope performance, a DC voltage is used for the envelope signal $A(t)$. The measured constant envelope maximum RF output power is $27\,\mathrm{dBm}$. For DCS-1800 Class 2, the required constant envelope output power for GMSK modulation is $24\,\mathrm{dBm}$. Therefore, the presented solution can also be used as a class 2 GSM power amplifier.

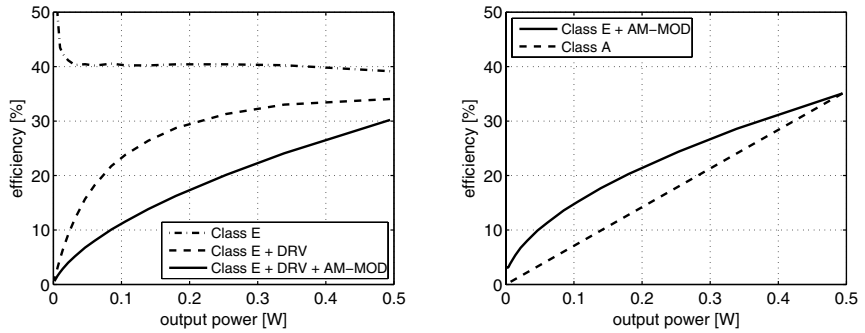


Figure 6.22. Measured efficiency versus output power, for constant envelope signals. In (b), this efficiency is compared to a Class A amplifier that would achieve the same peak envelope output power and efficiency.

The input power is $-3\,\mathrm{dBm}$, resulting in a power gain of $30\,\mathrm{dB}$ and it enables a direct connection of the upconverter to the linearized amplifier. Figure 6.22 shows the measured efficiency of the amplifier for various levels of P_{OUT}. This graph is obtained by applying several DC voltages at the input of the amplitude modulator. The different efficiencies are defined in section 5.2.3. The class E amplifier itself ($\eta_{d,E}$) maintains its high drain efficiency of $40\,\%$ over the entire power range. When the power consumption of the RF driver stages is taken into account ($\eta_{oa,E}$), the efficiency drops to $34\,\%$. Due to the voltage drop across the amplitude modulator ($\eta_{oa,PMA}$), the overall system efficiency becomes $30\,\%$. In figure 6.22(b), the efficiency of the Class E and linear amplitude modulator is compared to the efficiency of a Class A amplifier that would achieve the same output power and efficiency. Note however that the Class A amplifier can not be used for linear amplification at maximum output power due to compression. The polar modulated Class E solution on the other hand is able to transmit linearly, all the way up to the peak envelope output power of the RF amplifier since the linearity requirement is shifted to the amplitude modulator.

6.5.3 AM-AM and AM-PM Distortion Measurement

To measure the AM-AM and AM-PM distortion, a linear decreasing voltage with a period of $10\,\mu\mathrm{s}$ is used for the envelope signal. Once the voltage starts to decrease, a trigger pulse is send to the spectrum analyzer. The latter uses this trigger pulse to store the downconverted data in an internal memory, which

Table 6.2. Measured performance and process specifications.

Parameter	
technology	$0.18\mu m$ CMOS
supply voltage	1.8V and 3.3V
area	1.8x3.6mm^2
frequency	1.75GHz
Parameter	**Measured**
peak envelope output power	27dBm
input power	-3dBm
overall efficiency	34%
DC power consumption	
at peak envelope output power	
RF PA	1250mW @ 3.3V
RF Driver	186mW @ 1.8V
amplitude modulator	33mW @ 3.3V

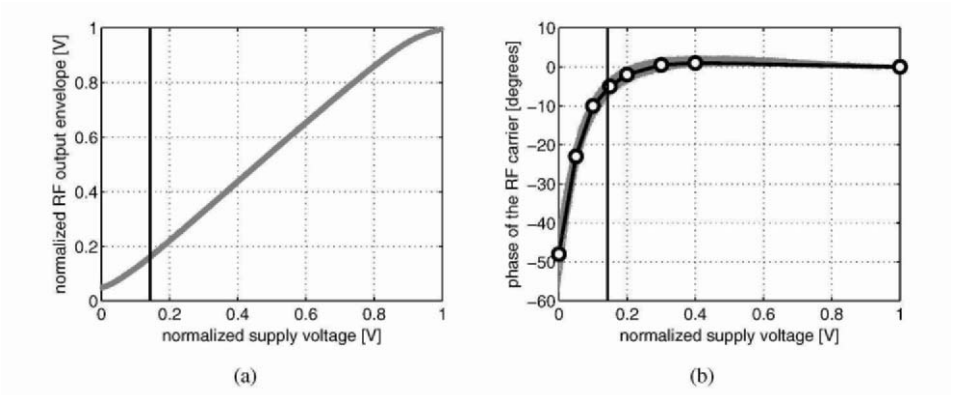


Figure 6.23. Measured AM-AM and AM-PM distortion.

can be accessed through the LAN bus. From this downconverted data, the amplitude and phase distortion can be obtained.

Figure 6.23 shows the measured AM-AM and AM-PM distortion of the amplifier. As expected, the distortion becomes severe when the supply voltage goes to zero. The dynamic range of the EDGE signal is indicated as well. However, one should realize that this dynamic range occurs at full output power. In other words, when power control is applied the average output power is reduced and the dynamic range moves to the left, introducing more AM-AM and AM-PM. Also note that the AM-AM shows some compression for a high envelope value. This is due to the saturation or compression in the opamp.

The AM-PM distortion has been measured for different samples and temperatures ranging from $20°C$ to $70°C$. All this data is plotted on figure 6.23, but only little variation can be noticed. Therefore, it is justified to apply this data for predistortion. The influence of the AM-AM distortion is less severe compared to the AM-PM and therefore only AM-PM predistortion is applied. To make a digital predistortion feasible, the measured AM-PM distortion has been averaged and is modeled by a piece-wise linear approximation which is also indicated in figure 6.23(b) by the solid line. This predistortion will be used in the EDGE measurements, discussed next.

6.5.4 EDGE Measurements

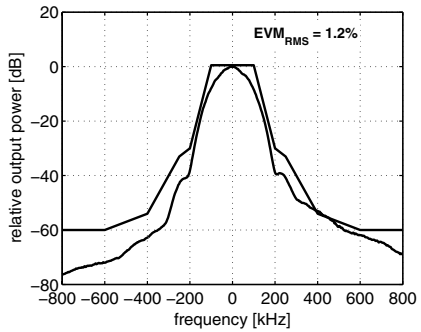
In EDGE, the peak to average power ratio is $3.4\,\mathrm{dB}$. Since the RF amplifier has a peak envelope output power of $27\,\mathrm{dBm}$, the maximum average output power, at least in theory, is about $23.6\,\mathrm{dBm}$ or $230\,\mathrm{mW}$. However, some clipping can be tolerated and the maximum average output power can therefore be higher than this number.

Figure 6.24 shows several EDGE output spectra at $214\,\mathrm{mW}$ or $23.3\,\mathrm{dBm}$ average output power. It is interesting to note how AM-PM distortion together with a delay between the amplitude path and the phase path can generate an asymmetric output spectrum. In figure 6.24(a), no predistortion or delay compensation is applied. It can be seen that the spectral mask margin is very small at $400\,\mathrm{kHz}$ offset. Furthermore, the output spectrum exhibits a large amount of asymmetry. When the 8 piece-wise linear predistortion is applied (figure 6.24(b)), the mask is easily met and the asymmetry has almost disappeared. However, the predistortion has little influence on the EVM. If the delay between amplitude and phase is completely compensated by a delay of $52\,\mathrm{ns}$, the spectrum is again symmetrical. Applying both AM-PM predistortion and delay compensation (figure 6.24(d)) results in a large margin at both $400\,\mathrm{kHz}$ and $600\,\mathrm{kHz}$ offset. In figure 6.24(a), the asymmetry of the spectrum raises the right side of the spectrum, whereas in figure 6.11, the left side of the spectrum was raised. As such, one can conclude that in the measurement setup, the phase signal is delayed compared to the envelope signal. Inserting a delay of $52\,\mathrm{ns}$ in the phase path to obtain a symmetrical spectrum, confirms this assumption.

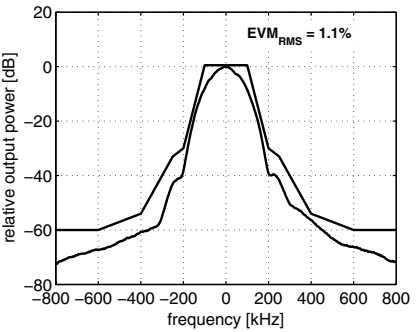
Figure 6.25 shows a wideband measurement of the modulated output spectrum at maximum output power, up to $6\,\mathrm{GHz}$ and measured with a $100\,\mathrm{kHz}$ resolution bandwidth. The amplifier noise inside the DCS-1800 frequency band is low enough, and the second and third harmonics of the RF carrier are visible at $3.6\,\mathrm{GHz}$ and $5.25\,\mathrm{GHz}$, though their power level is already low.

6.5.4.1 EDGE performance versus output power

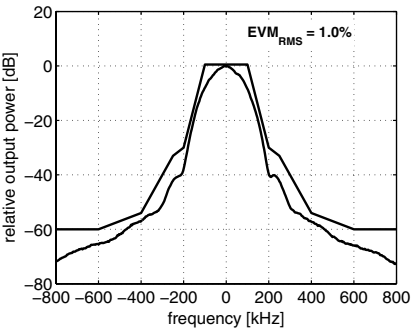
Figures 6.26(a) and 6.26(b) show the value of the output spectrum at $400\,\mathrm{kHz}$ and $600\,\mathrm{kHz}$ offset versus average output power. Both the upper and lower



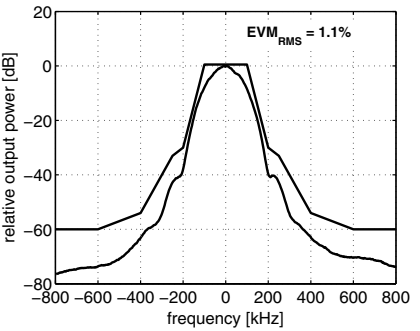
(a) no predistortion, no delay compensation



(b) AM-PM predistortion, no delay compensation



(c) delay compensation, no predistortion



(d) AM-PM predistortion and delay compensation

Figure 6.24. Measured EDGE output spectra at $214\,\mathrm{mW}$ or $23.3\,\mathrm{dBm}$.

spectral values are given to indicate the presence of an asymmetry of the output spectrum. In figure 6.26(a), the spectral mask at $+400\,\mathrm{kHz}$ offset is not met if no predistortion or delay compensation is applied. The spectral mask at $-400\,\mathrm{kHz}$ offset is much lower which corresponds to a large asymmetry of the output spectrum. When the delay compensation of $52\,\mathrm{ns}$ is applied, both curves fall together, indicating a symmetrical output spectrum. With the predistortion, the spectral mask margin can be further increased as well as the power range over which the mask is met. The spectral mask margin is high enough to allow some degradation in the upconverter.

For a low average output power, the spectral mask margin is smaller due to the increased AM-AM and AM-PM. At maximum output power, the spectral mask margin is reduced because of additional AM-AM distortion in the ampli-

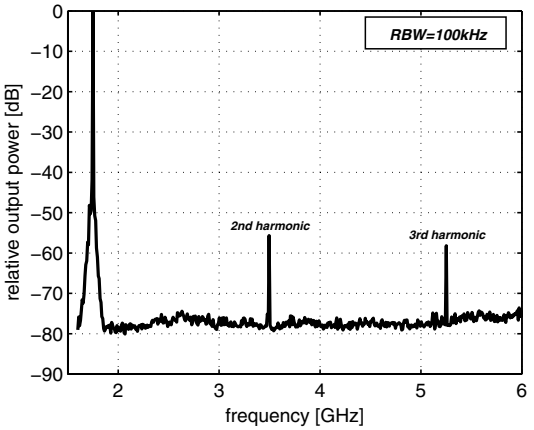


Figure 6.25. Wideband measurement of the output spectrum at maximum output power.

tude modulator, as can be seen on figure 6.26. Note that the maximum output power is $240\,\mathrm{mW}$ or $23.8\,\mathrm{dBm}$, which is slightly higher than the theoretical value of $230\,\mathrm{mW}$ or $23.6\,\mathrm{dBm}$

To achieve the spectral mask specifications over a wider power range, a more aggressive AM-AM and AM-PM predistortion should be applied, based on a look-up table that can be updated periodically. Figure 6.26(c) shows the measured EVM_{RMS} versus average output power. The EVM_{RMS} is well below the required $9\,\%$ specification and is mainly reduced by the delay compensation.

6.5.4.2 EDGE performance versus frequency

Figure 6.27 shows the performance of the linearized amplifier versus frequency. The output power and spectral mask specifications are met for both the DCS-1800 band ranging from $1.71\,\mathrm{GHz}$ to $1.785\,\mathrm{GHz}$ and the PCS-1900 band from $1.85\,\mathrm{GHz}$ to $1.91\,\mathrm{GHz}$.

6.5.4.3 EDGE performance versus load mismatch

The load resistance of $50\,\Omega$, used in these measurements, is actually the impedance of the antenna at the transmit frequency. Due to environmental changes, this impedance might change, which is characterized by the $VSWR$ that corresponds to such a mismatch. The GSM-EDGE Standard requires that the linearity requirements are met for a $VSWR$ mismatch of 3.

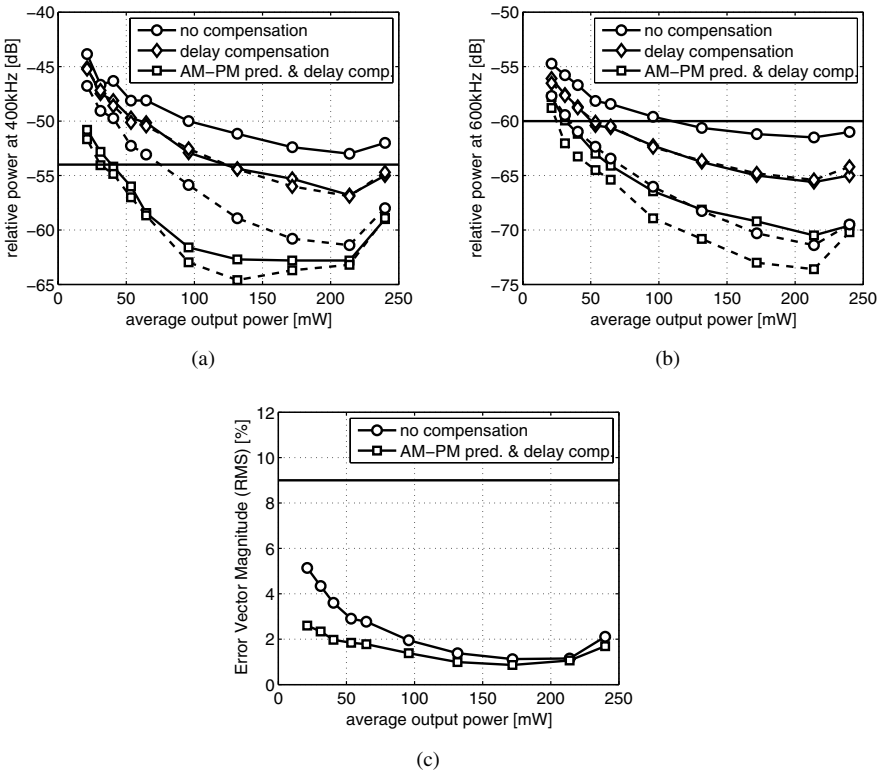


Figure 6.26. Measurement of the relative output power at $400\,\mathrm{kHz}$ and $600\,\mathrm{kHz}$ frequency offset and Error Vector Magnitude (RMS) versus average output power. In (a) and (b), the solid line indicates the measurement at a positive frequency offset and the dashed line is the measurement at the negative frequency offset.

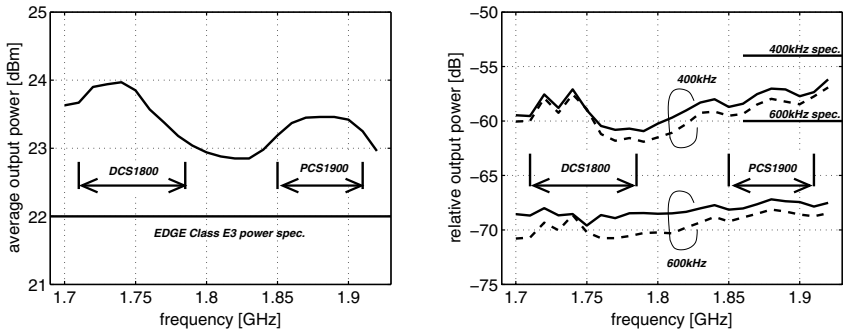
The $VSWR$ can be written as

$$VSWR = \frac{1 + |\Gamma|}{1 - |\Gamma|} \tag{6.7}$$

with Γ the coefficient of reflection, defined as

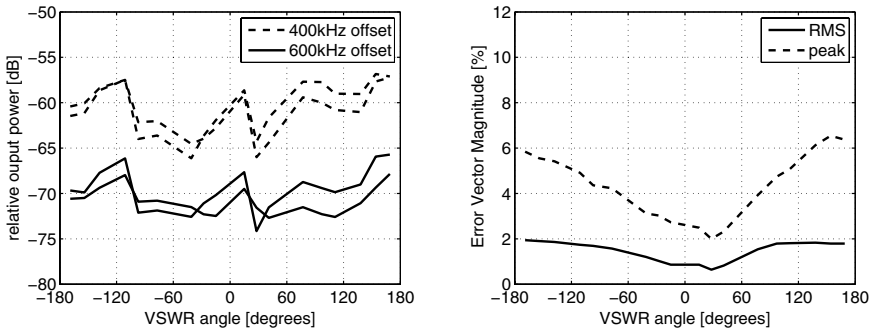
$$\Gamma = \frac{Z_L - Z_C}{Z_L + Z_C} \tag{6.8}$$

in which Z_C is the characteristic impedance of the system and Z_L is the load impedance, normally equal to $50\,\Omega$. Since Z_L and Γ are complex numbers, a $VSWR$ of 3 can be achieved with several Γ angles. This is commonly denoted as the $VSWR$ angle, though it is actually the angle of Γ.



(a) (b)

Figure 6.27. Output power and relative power at 400 kHz and 600 kHz frequency offset versus carrier frequency.



(a) (b)

Figure 6.28. Measured EVM_{RMS} and relative power at 400 kHz and 600 kHz for a $VSWR$ of 3.

The measured relative power at 400 kHz and 600 kHz and the measured EVM_{RMS} for a $VSWR$ of 3 are shown in figure 6.28. Both the spectral mask and EVM requirements are met over the entire range of $VSWR$ angles.

6.5.4.4 EDGE performance versus envelope bandwidth

To demonstrate how delay compensation can reduce the required envelope bandwidth, figure 6.29 shows the measured output spectrum at ± 400 kHz and ± 600 kHz offset for different values of the envelope bandwidth. From the simulated results of figure 6.13, it was found that the bandwidth should be at least 1.5 MHz to meet the spectral mask requirements. This is in relative good agreement with the measurement of figure 6.29. If delay compensation is ap-

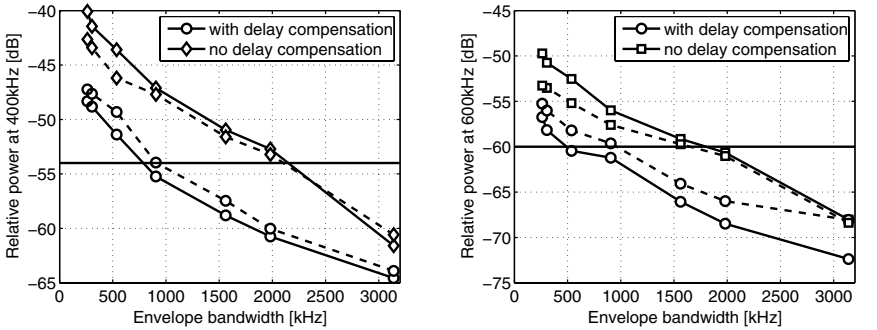


Figure 6.29. Measured relative power at $400\,\mathrm{kHz}$ and $600\,\mathrm{kHz}$ frequency offset versus envelope bandwidth, with and without delay compensation.

plied, the envelope bandwidth can be reduced below $1\,\mathrm{MHz}$. This is higher than the $600\,\mathrm{kHz}$, predicted by figure 6.13, but that figure did not include other distortion mechanisms. Altogether, the measurement of figure 6.29 clearly demonstrates how delay compensation can reduce the bandwidth of the amplitude modulator.

6.5.5 16-QAM Modulation and Two-Tone Test

In figure 6.30(a), the demodulated measured constellation diagram for an output power of $23.8\,\mathrm{dBm}$ is shown. The RMS value of the Error Vector Magnitude equals $1.7\,\%$. Figure 6.30(b) shows the demodulated measured constellation diagram if a 16-QAM constellation is transmitted, using the EDGE baseband filter. In 16-QAM, the envelope signal goes through the origin of the complex plane which was not the case in EDGE due to the $3\pi/8$ symbol rotation. Figure 6.30(b) thus demonstrates that with this architecture, it is possible to transmit an amplitude modulated signal that goes to zero. The PAPR of the 16-QAM signal is equal to $5.9\,\mathrm{dB}$ and the measured RMS value of the Error Vector Magnitude equals $3\,\%$. In figure 6.30(b), the output power is $22\,\mathrm{dBm}$ and the corresponding overall efficiency is $18\,\%$.

To conclude the measurements, figure 6.31 shows the result of a two-tone test for several values of average output power. No predistortion or delay correction was applied in these measurements. A two-tone signal has an envelope that goes down to zero and is very sensitive to phase distortion, since the phase signal is a square wave. At maximum output power, AM-AM distortion of the amplitude modulator will dominate and the output spectrum is symmetrical. For medium power levels, an asymmetry can be observed in the spectrum, due to the envelope filtering, AM-PM and differential delay. For lower output

Table 6.3. Measured performance and ETSI specifications.

Parameter	Measured	EDGE E3 Specs.
average output power	23.8dBm	22dBm
overall efficiency	22%	
modulation spectrum		
400kHz offset, 30kHz RBW	-59dB	-54dB
600kHz offset, 30kHz RBW	-70dB	-60dB
Error Vector Magnitude		
RMS EVM	1.69%	9%
Peak EVM	5.87%	30%

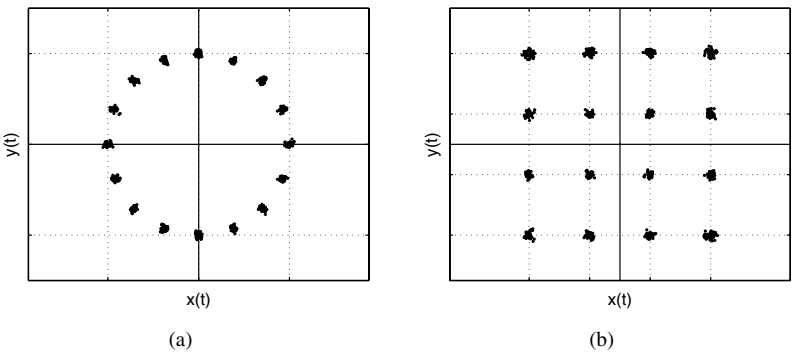


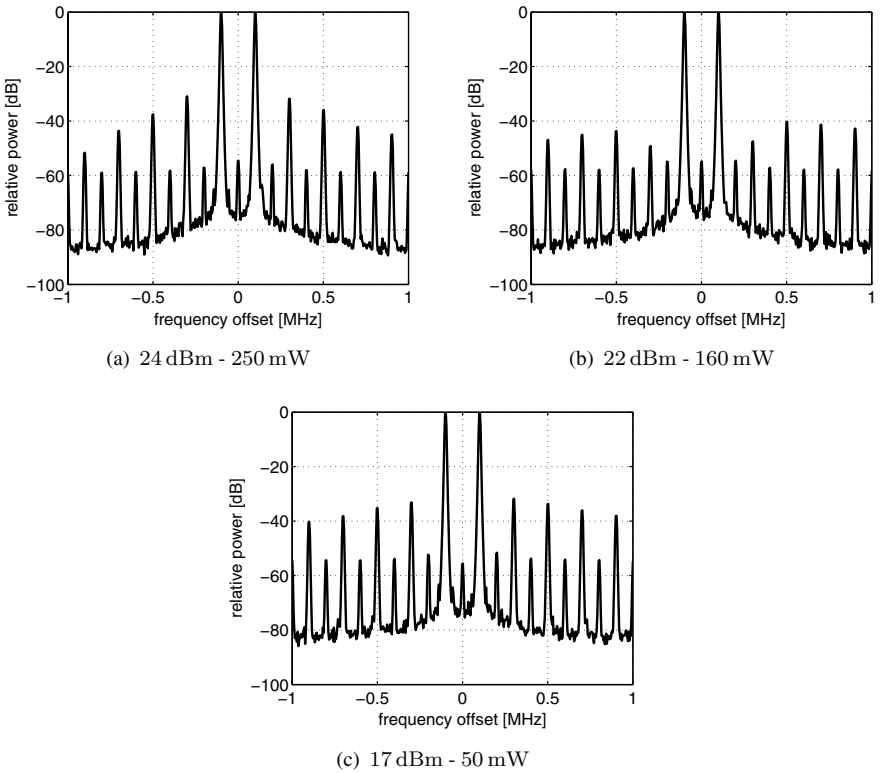
Figure 6.30. Demodulated and normalized measured constellation diagram for (a) rotated 8PSK and (b) 16-QAM modulation.

power, the AM-AM distortion is dominant and the output spectrum is symmetrical again.

This measurement also reveals that classical IM3 and IP3 simulations and measurements are not sufficient to characterize the performance of a transmitter for digital communication systems. In figure 6.31, the IM3 increases for both a high and a low output power, which would actually result in two IP3 points.

6.6 Architectural Improvements

The GSM-EDGE Standard requires that the average modulated output power can be regulated down to $0\,\mathrm{dBm}$ in $2\,\mathrm{dB}$ steps. Due to the feedforward from the driver stages, the minimal output power of this solution, while still meeting



(a) 24 dBm - 250 mW

(b) 22 dBm - 160 mW

(c) 17 dBm - 50 mW

Figure 6.31. Measured output spectrum for a two-tone signal with a 200 kHz tone spacing.

the spectral mask requirements, is $40\,\mathrm{mW}$ or $16\,\mathrm{dBm}$. To reduce this number, the supply voltage of the driver stages has to be reduced or modulated as well. In fact, one might as well turn off the supply of the Class E amplifier, apply the polar modulation on the driver stage and rely entirely on the feedforward of the Class E stage.

If the supply voltage of both the Class E and driver stages is modulated, the linearity deteriorates. An example of such a system is given in [Sand03]. The architecture uses both AM-AM and AM-PM predistortion to achieve the required linearity over the entire output power range.

Another solution is to use a low power Class AB amplifier, in parallel with the Class E amplifier, which takes over at low output power levels. The Class AB can be placed in parallel [Ding05] or the power combining architecture of chapter 4 can be used to combine the two amplifiers.

Finally, a cascode solution [Yoo00],[Sowl02] will have a less linear on-resistance, but the feedforward and hence the AM-PM distortion can be reduced considerably. This would also enable a wider power range.

6.7 Comparison with Other EDGE Solutions

To summarize, this EDGE power amplifier achieves an average modulated output power of $23.8\,\mathrm{dBm}$ with an average PAE of $22\,\%$. At maximum output power, the EVM_{RMS} is $1.7\,\%$ and the relative output power at $400\,\mathrm{kHz}$ frequency offset is $-59\,\mathrm{dB}$ and $-70\,\mathrm{dB}$ at $600\,\mathrm{kHz}$ frequency offset. Furthermore, the EVM_{RMS} and spectral mask specifications are met over a wide power range.

Comparison with other work is difficult since —to the author's knowledge— this is the only integrated CMOS power amplifier capable of amplifying EDGE signals. Most commercial available products use expensive GaAs or SiGe linear amplifiers.

- Skyworks has a polar loop transmitter for GSM-EDGE that consists of a 3-chip solution: a transceiver, a PA controller and a PA module [Sowl04]. The PA controller is implemented in a $0.35\,\mu\mathrm{m}$ BiCMOS technology. The technology of the PA itself is not mentioned in the paper. The EVM_{RMS} of $4.5\,\%$ is higher compared to the EVM_{RMS} of this work ($1.7\,\%$). The relative power at $400\,\mathrm{kHz}$ offset is $-60\,\mathrm{dB}$ which is comparable to this work. Both the EVM_{RMS} and spectral mask figures are only given for a fixed output power and no indication is given on how to implement power control and how this will influence the EVM_{RMS} and spectral mask. The higher output power of $26.5\,\mathrm{dBm}$ and the PAE of $35\,\%$ are understandable as an external, high efficiency power amplifier is used.

- Ericsson [Pehl04] and Tropian [Sand03] have an open-loop topology that is similar to this work. They both use a single-chip CMOS controller and, again, have an external RF power amplifier in a dedicated technology. These papers do not mention any figures regarding output power and efficiency.

- The CMOS power amplifier presented in [Fall01] has EDGE capabilities. The amplifier operates in Class AB, close to Class B. The relative spectrum at $400\,\mathrm{kHz}$ is only $-32\,\mathrm{dB}$ at a peak envelope output power of $28\,\mathrm{dBm}$, hence the EDGE specification of $-54\,\mathrm{dB}$ is not met. For EDGE, the average output power will be about $3.4\,\mathrm{dB}$ lower than the peak envelope output power although this average output power is not published. Also the corresponding overall modulated efficiency and EVM_{RMS} are not published.

This brief comparison demonstrates that the achieved linearity of the presented, fully integrated CMOS solution is comparable with other, more expensive, multi-chip solutions and this work also outperforms an integrated CMOS Class AB solution.

6.8 Conclusion

In this chapter, the design of a linearized polar modulated CMOS power amplifier in a $0.18\,\mu\mathrm{m}$ CMOS technology was discussed. The amplifier targets and meets the requirements of the GSM-EDGE cellular phone system.

First, in section 6.2, the EDGE system was discussed. EDGE uses existing GSM equipment and licenses and achieves a datarate that is three times as high, close to the datarate of UMTS. The investigation of the EDGE signal and the EDGE linearity requirements deliver the basic design requirements for the power amplifier.

In section 6.3 some architectural issues were discussed and the selection of the linear amplitude modulator was motivated. The high integration level is the main issue in this motivation. In the same section, the distortion mechanisms of the polar modulated power amplifier were investigated with respect to the EDGE signal and system requirements. In this regard, the spectral mask requirement is harder to meet compared to the error vector magnitude requirements.

Detailed circuit level implementation issues were given in section 6.4. First the design of the Class BE RF power amplifier was investigated. The optimization was clarified by contour plots, depicting the output power and efficiency tradeoff. Also the design of the linear amplitude modulator was covered, as well as some layout issues of the $0.18\,\mu\mathrm{m}$ CMOS technology.

The measurement results of the fully integrated power amplifier were covered in section 6.5. First, the measurement setup was discussed, followed by the constant envelope power measurement and AM-PM measurement. The RF amplifier achieves a peak envelope output power of $27\,\mathrm{dBm}$ with a drain efficiency of $40\,\%$ and an overall efficiency of $34\,\%$. The spectral mask and EVM measurements under EDGE modulation were also extensively demonstrated. The interaction between AM-PM distortion and a delay in the phase path creates an asymmetry in the output spectrum, by which the spectral mask is violated. AM-PM predistortion and delay compensation is suggested to overcome most of the distortion. This allows to meet the EDGE specifications for an average output power from $23.8\,\mathrm{dBm}$ down to $16\,\mathrm{dBm}$. The EDGE specifications were also measured and met over the DCS-1800 and PCS-1900 frequency bands, and for a $VSWR$ of 3. In the previous chapter, delay compensation of the phase signal was suggested to reduce the distortion of low-pass envelope

filtering. This was verified by measurements, and it was shown that this technique indeed allows to reduce the bandwidth of the amplitude modulator.

To conclude, this chapter has proven the feasibility of a single-chip linearized power amplifier that meets the EDGE cellular system requirements in a submicron CMOS technology.

Chapter 7

A CMOS POWER AMPLIFIER FOR BLUETOOTH

7.1 Introduction

This chapter will discuss the design and CMOS implementation of a fully integrated CMOS RF power amplifier that uses the lattice-type LC balun, introduced in chapter 4 [Reyn05c].

The amplifier is designed in a $0.13\,\mu\mathrm{m}$ CMOS technology and requires no expensive off-chip components. The amplifier operates at $2.45\,\mathrm{GHz}$ and meets the Bluetooth specifications.

The Bluetooth system is briefly discussed in section 7.2. From this, some general guidelines are obtained for the design of the RF amplifier. The power amplifier uses the lattice-type LC-balun of section 4.3 to achieve sufficient output power. The design of this network is discussed in section 7.3, as well as the layout implementation issues. The measurement results are covered in section 7.5 and final conclusions are given in section 7.7.

7.2 The Bluetooth System

Bluetooth is an industrial specification for wireless personal area networks (PANs). Bluetooth provides a way to connect and exchange information between devices like personal digital assistants (PDAs), mobile phones, laptops, PCs, printers and digital cameras via a secure, low-cost, globally available, short range radio frequency. Bluetooth is a radio standard primarily designed for low power consumption, with a short range (power class dependent: 10 centimeters, 10 meters, 100 meters or up to 400 meters) and with a low-cost transceiver microchip in each device. Bluetooth works in the $2.4\,\mathrm{GHz}$ ISM (Industrial Scientific Medicine) band that ranges from $2.4\,\mathrm{GHz}$ to $2.4835\,\mathrm{GHz}$ and has 78 channels of $1\,\mathrm{MHz}$ wide. The ISM band is a free band and is used by other equipment, like WLAN and microwave ovens. Therefore, Bluetooth

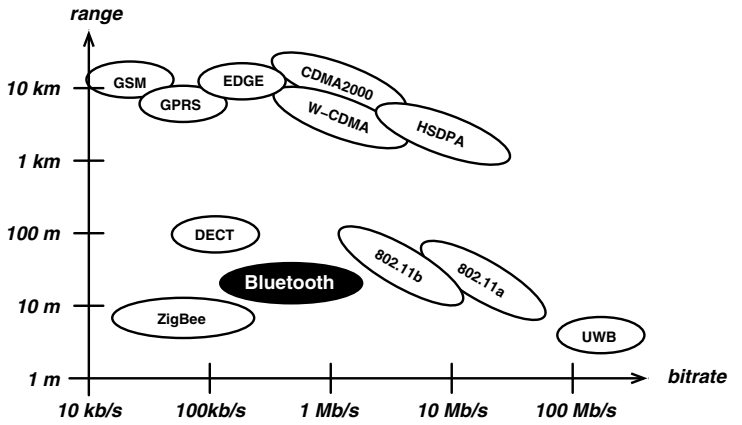


Figure 7.1. Position of the Bluetooth system between other wireless standards.

uses frequency hopping to maintain the communication link. Finally, figure 7.1 depicts the position of Bluetooth between other popular wireless standards.

7.2.1 Modulation

Bluetooth uses Gaussian Frequency Shift Keying (GFSK)[1] [SIG04]. A binary one is represented by a positive frequency deviation and a binary zero is represented by a negative frequency deviation. GFSK results in a constant envelope signal, and thus a switching amplifier can be used. A GFSK modulated RF signal can be written as

$$v(t) = A \cdot \cos\left(\omega_c t + \theta(t)\right) \tag{7.1}$$

The modulation is present in $\theta(t)$ as

$$\theta(t) = \sum_i a_i \pi h \int_{-\infty}^{t-iT} h(u) * rect\left(\frac{u}{T}\right) du \tag{7.2}$$

in this equation

- a_i are the modulating bits, either +1 or -1.
- h is the modulating index, whose value should be between 0.28 and 0.35.
- $h(t)$ is the impulse response of the baseband Gaussian filter, by which the digital data is filtered to reduce the bandwidth of the RF signal.

[1]The newest version of Bluetooth, Enhanced Data Rate (EDR), is not covered here

- $rect\left(\frac{u}{T}\right)$ is the rectangular pulse associated with each data bit.

For Bluetooth, the BT product equals 0.5 and the raw maximum datarate equals $1\,\mathrm{Mbps}$ [SIG04]. The Bluetooth system is very similar to GSM, which uses GMSK, i.e. $h = 0.5$ and $BT = 0.3$.

7.2.2 Power Amplifier Requirements

Bluetooth enables devices to talk to each other when they come in range, even if they are not in the same room, as long as they are within up to 100 meters of each other, dependent on the power class of the product. Products are available in one of three power classes:

1 Class 1: $100\,\mathrm{mW}$ or $20\,\mathrm{dBm}$, it has the longest range up to 100 meters.

2 Class 2: $2.5\,\mathrm{mW}$ or $4\,\mathrm{dBm}$, it allows a quoted transmission distance of 10 meters, and is the most common power level.

3 Class 3: $1\,\mathrm{mW}$ or $0\,\mathrm{dBm}$, it allows transmission of 10 cm, with a maximum of 1 meter. This power class is rather rare.

Bluetooth also requires that the output power of a Class 1 device can be regulated in order to reduce the power consumption of the transceiver and thus to increase the battery lifetime. For a Class 1 device, the output power has to be regulated down to $4\,\mathrm{dBm}$ or less with a step size between $2\,\mathrm{dB}$ and $8\,\mathrm{dB}$.

7.2.3 Spectral Purity and Spurious Emissions

Within the ISM band, the transmitter has to meet both a spectral mask and adjacent channel power (ACP) requirements. The spectrum is measured with a resolution bandwidth of $100\,\mathrm{kHz}$.

At $\pm 500\,\mathrm{kHz}$ frequency offset, the modulating spectrum should be $20\,\mathrm{dB}$ below the spectrum at zero frequency offset. In the Bluetooth specification, this is denoted as $-20\,\mathrm{dBc}$, but is it actually $20\,\mathrm{dB}$ below the power in a $100\,\mathrm{kHz}$ frequency band around the carrier frequency.

Besides the spectral mask, the integrated power over a $1\,\mathrm{MHz}$ band should be low enough in the neighboring or adjacent channels. This is not a spectral mask requirement, but an adjacent channel power requirement. For a channel spacing of two, the power transmitted in a $1\,\mathrm{MHz}$ band should be lower than $-20\,\mathrm{dBm}$ or $10\,\mu\mathrm{W}$. For a channel spacing of more than three, that power should be lower than $-40\,\mathrm{dBm}$ or $0.1\,\mu\mathrm{W}$.

Outside the ISM band, the transmitter has to meet following requirements:

frequency band	transmitted power in a 100 kHz band
30 MHz - 1 GHz	$-36\,\mathrm{dBm}$
1 GHz - 12.75 GHz	$-30\,\mathrm{dBm}$
1.8 GHz - 1.9 GHz	$-47\,\mathrm{dBm}$
5.15 MHz - 5.30 GHz	$-47\,\mathrm{dBm}$

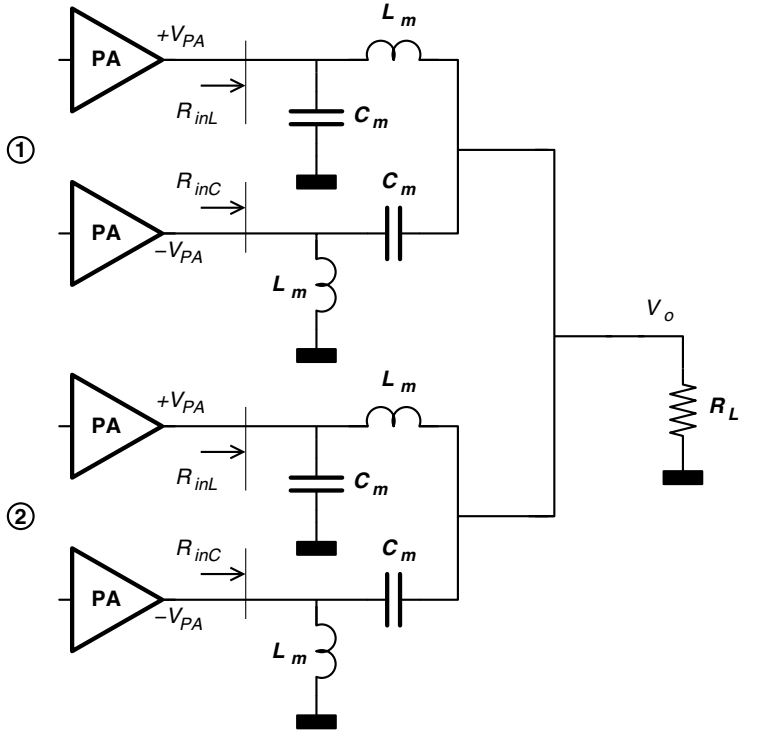


Figure 7.2. Architecture of the RF amplifier and lattice-type LC balun.

7.3 Circuit Implementation

This design uses the lattice-type LC balun of chapter 4 to achieve sufficient output power in a $0.13\,\mu\mathrm{m}$ CMOS technology. It was the aim to demonstrate that with native $0.13\,\mu\mathrm{m}$ transistors and the power combining network, sufficient output power can be achieved to meet the Class 1 Bluetooth requirements. Therefore, no thick gate oxide transistors or device stacking techniques were used.

The structure, introduced in section 4.3 and shown here in figure 7.2, consists of two L-C sections, driven by a differential voltage $(+V_{PA}, -V_{PA})$. Impedance transformation can be obtained by a proper selection of L_m and C_m. As this circuit allows multiple power amplifiers to be connected to one load, the input impedance R_{in} of each segment can be made larger, while still achieving a high output power. This reduces the losses of each section and enables to achieve sufficient output power at a low supply voltage. The single-ended output avoids the need of an external balun or RF transformer and inherently allows to efficiently control the output power.

To achieve the $20\,\mathrm{dBm}$ output power specification at $2.45\,\mathrm{GHz}$, two differential amplifiers are placed in parallel. To provide some margin, each amplifier is designed to deliver $60\,\mathrm{mW}$ of output power, resulting in a theoretical total output power of $240\,\mathrm{mW}$ or $23.8\,\mathrm{dBm}$. At zero drain current, the nMOS transistor can withstand a drain voltage of $4\,\mathrm{V}$. Assuming a Class E waveform with a peak drain voltage of $3.5\,\mathrm{V}$, the equivalent load impedance required to have an output power of $60\,\mathrm{mW}$ is $Z_{in} = 9.6\,\Omega$. In section 4.3, it was found that

$$R_{inL} = R_{inC} = R_{in} = \frac{R_L}{2} \cdot \left(\frac{B}{R_L}\right)^2 \tag{7.3}$$

From which, B can easily be obtained

$$B = \sqrt{2 \cdot N \cdot R_L \cdot Z_{in}} \approx 44\Omega$$

At a frequency of $2.45\,\mathrm{GHz}$, the corresponding values of L_m and C_m can be calculated as

$$L_m = \frac{B}{\omega} = 2.86\mathrm{nH}$$

$$C_m = \frac{1}{B\omega} = 1.48\mathrm{pF}$$

and these values are both easily implemented in CMOS.

The optimum of the nMOS switches in the Class E power amplifier, is a width of $1000\,\mu\mathrm{m}$ with a gate length of $0.13\,\mu\mathrm{m}$. Making the switch larger would result in a minor increase of η_d, but it would also increase the driver power $P_{DC,DRV}$ and hence the overall efficiency η_{oa} would reduce.

Figure 7.3 shows the complete circuit of the PA. The beauty of the lattice-type LC balun is that the inductor L_m and capacitor C_m, both connected to ground and needed to create a pure resistive input impedance, are absorbed into the Class E power amplifier. As such, less components are needed. The consequence is that the components of the two branches of the differential Class E amplifier are no longer identical, i.e. $C_{1a} > C_{1b}$ and $L_{1a} > L_{1b}$. The DC-feed inductors of the two differential branches, L_{1a} and L_{1b}, can still be merged together but it then requires a coil with a *non-centered center-tap*. This was not possible in the $0.13\,\mu\mathrm{m}$ technology and therefore the DC-feed inductors L_{1a} and L_{1b} are not merged together.

The outputs of the four RF amplifiers thus need to be combined which results in long interconnect lines that introduce a lot of parasitics. These interconnect parasitics are also included in the design and are modeled as lumped striplines. DC-blocking capacitors are included at the output of each amplifier to avoid a DC current through the load. The parasitics of the interconnection will have an influence on the LC balun and therefore, the values of the inductors and capacitors slightly change in order to maintain Class BE operation. The final values are summarized in following table:

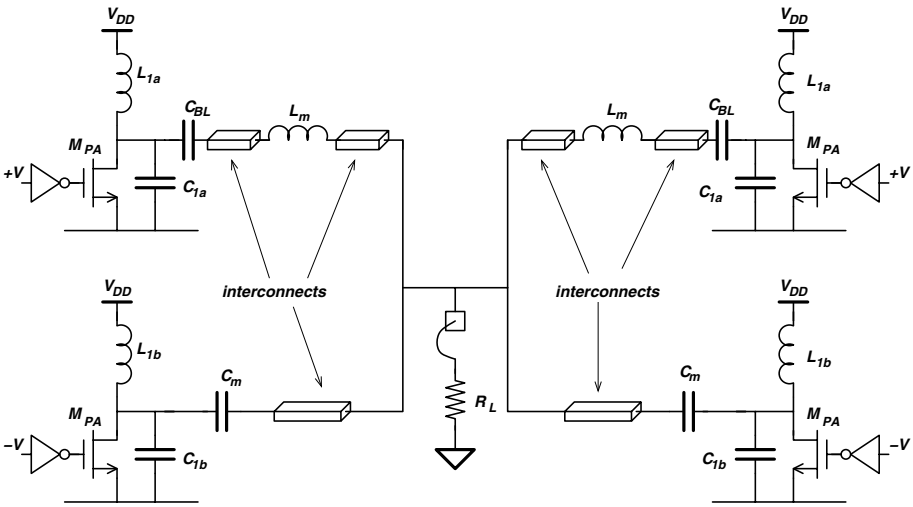


Figure 7.3. Circuit implementation of the Bluetooth PA.

component	value	component	value
L_{1a}	1.3 nH	L_{1b}	0.9 nH
C_{1a}	1.2 pF	C_{1b}	1.1 pF
L_m	3.1 nH	C_m	0.8 pF
C_{BL}	10 pF		

The RF driver stages are simple digital inverters. The main reason for this choice was the smaller required silicon area. The drawback is the increased DC power consumption, since the gate capacitance of the output transistor is not tuned by an inductor. The input power is $-6\,\mathrm{dBm}$ which enables a direct connection of the upconversion mixers to the first RF driver stage. All the inductors and capacitors of figure 7.3 are integrated on-chip and no external matching or tuning is necessary.

7.4 Layout Aspects

Figure 7.4 shows a photograph of the fully integrated RF amplifier. The size is $2.74\,\mathrm{mm}$ by $2.00\,\mathrm{mm}$. At $2.45\,\mathrm{GHz}$, the value of L_m is still relatively large. Therefore, this topology is even better suited at higher frequencies, since that would reduce the values of L_m and L_1. This would also enable to place more than four amplifiers in parallel, and to use tuned driver stages without consuming excessive silicon area.

Clearly, to connect the outputs, long interconnections are necessary. The electrical model of these lines is shown in figure 7.5. Underneath the interconnect lines lies a patterned metal ground plane to accurately model the parasitic

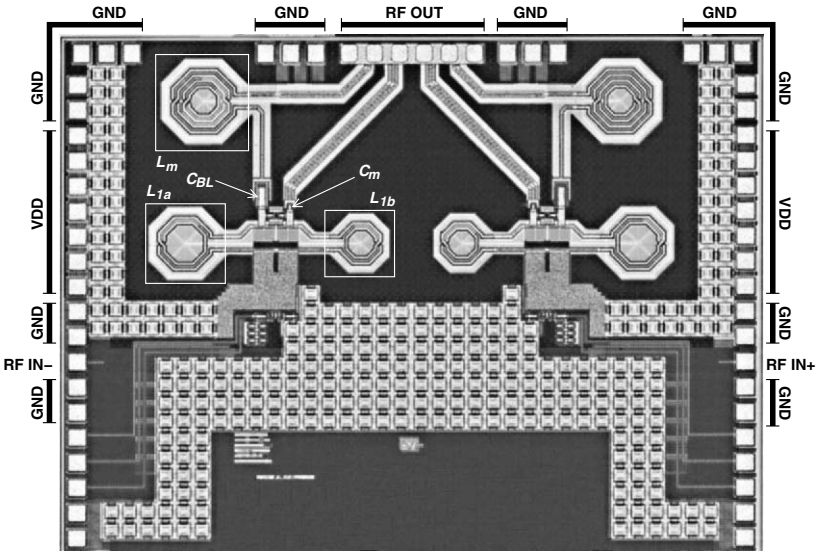


Figure 7.4. Photograph of the fully integrated RF amplifier.

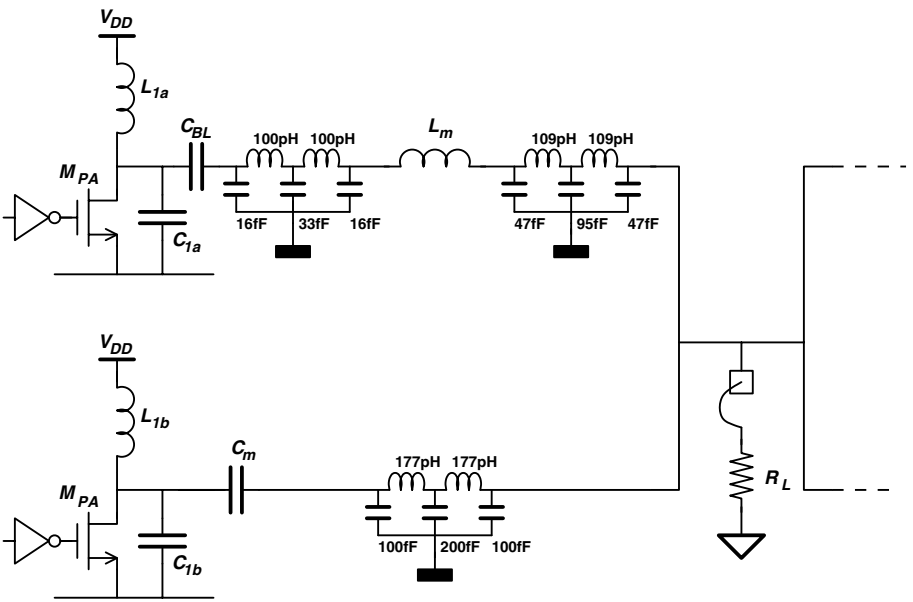


Figure 7.5. One section of the Bluetooth PA, including the models of the parasitic interconnect lines.

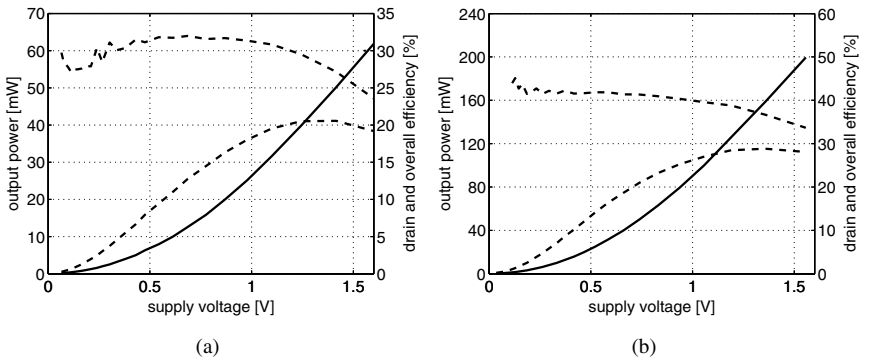


Figure 7.6. Measured output power (solid line), drain efficiency (upper dashed line) and overall efficiency (lower dashed line) for (a) one power amplifier and (b) two amplifiers in parallel.

capacitance, to avoid capacitive signal injection into the substrate and to short-circuit the substrate losses.

A total value of two times $206\,\mathrm{pF}$ is implemented to bypass the supply voltage of the Class E stage. High density Metal-insulator-Metal (MiM) capacitance with a density of $2\,\mathrm{fF}/\mu\mathrm{m}^2$ is used for this task.

7.5 Measurements

7.5.1 Output Power and Efficiency

Figure 7.6 shows the measured output power and efficiency at $2.45\,\mathrm{GHz}$ versus supply voltage, for both one and two differential amplifiers in parallel. For two amplifiers, a maximum output power of $200\,\mathrm{mW}$ or $23\,\mathrm{dBm}$ can be achieved at an overall efficiency of $28\,\%$. The corresponding drain efficiency is $34\,\%$ and the driver stages consume $118\,\mathrm{mW}$. This dissipation could be reduced if tuned driver stages were used, but it would drastically increase the chip area. Notice the reduction of the overall efficiency η_{oa} for lower output power, which is due to the constant dissipation in the switching driver stages.

Figure 7.6 also shows the measurement when only one segment of the PA is activated. If N is reduced from two to one, the output power will drop by $6\,\mathrm{dB}$ if the supply voltage is kept constant. In this case, the peak output power is $60\,\mathrm{mW}$ or $17.8\,\mathrm{dBm}$ which is indeed approximately $6\,\mathrm{dB}$ below the peak output power of two amplifiers in parallel. For one amplifier, the maximum overall efficiency is $21\,\%$ at $50\,\mathrm{mW}$ and the corresponding drain efficiency is $27\,\%$. The power dissipation in the driver stage is divided by two, resulting in a consumption of $59\,\mathrm{mW}$. The values of C_1 and L_1 were optimized for

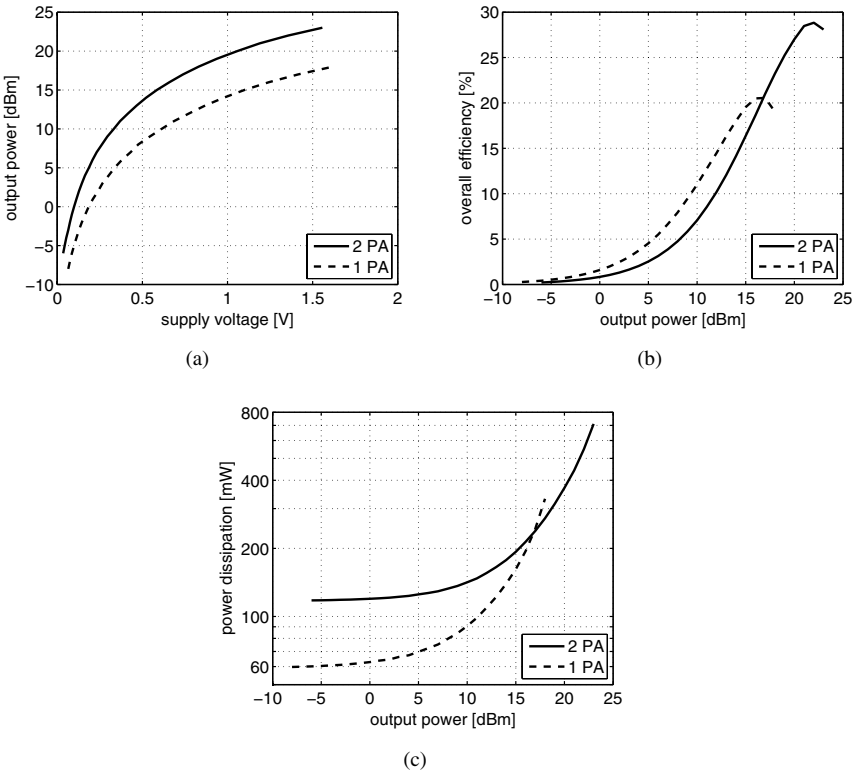


Figure 7.7. Output power, efficiency and power dissipation of one power amplifier and two amplifiers in parallel.

maximum efficiency at maximum output power. Therefore, the efficiency is maximum when both amplifiers are working in parallel.

When N is increased from 1 to 2, the transformed load impedance decreases from $19.2\,\Omega$ to $9.6\,\Omega$. For the same supply voltage the difference in output power between one and two amplifiers is close to the theoretical value of $6\,\mathrm{dB}$, as shown in figure 7.7(a). When only one is PA used, the dissipation of the driver stage is divided by two, which results in an increase of the overall efficiency of one PA $\eta_{oa,1PA}$ at lower power levels. This mechanism is clearly indicated on figure 7.7(b). At a power level of $17\,\mathrm{dBm}$ and below, it is beneficial to use only one PA. To further demonstrate the benefit to switch off one PA at lower power levels, figure 7.7(c) shows the measured power dissipation of the entire amplifier, including the driver stages.

In figure 7.8, the output power, drain efficiency and overall efficiency versus frequency is given for both one and two amplifiers working at a supply

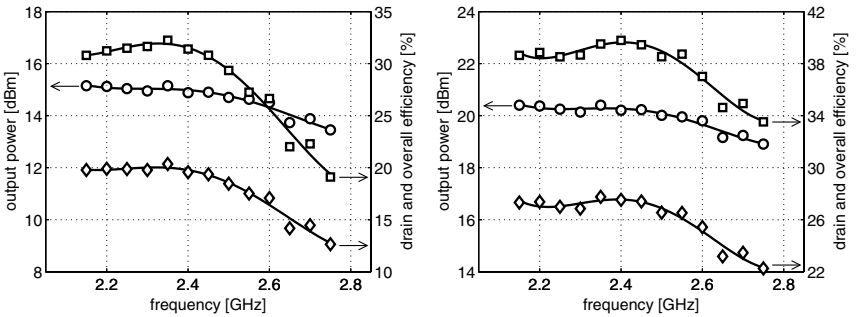


Figure 7.8. Measured output power ($\circ$), drain efficiency ($\square$) and overall efficiency ($\diamond$) versus frequency for (a) one amplifier and (b) two amplifiers in parallel. The supply voltage is 1.1 V.

Table 7.1. Measured constant envelope performance summary.

parameter	measured performance	
	2-PA	1-PA
output power	200 mW - 23 dBm	50 mW - 17 dBm
drain efficiency	42 %	32 %
overall efficiency	29 %	21 %
input power	−6 dBm	−6 dBm
driver stage power consumption	118 mW	59 mW

Table 7.2. Measured Bluetooth performance summary.

parameter	measured performance	Bluetooth specification
spectral mask	−22.1 dBc and −21.25 dBc	−20 dBc
ACP at channel offset 2	−46 dBm	−20 dBm
ACP at channel offset 3	−47 dBm	−40 dBm

voltage of 1.1 V. Notice the difference of about 6 dB in output power over the entire frequency range. Table 7.1 summarizes the power and efficiency measurements.

7.5.2 Bluetooth Measurements

Figure 7.9 shows the measured Bluetooth output spectrum at maximum output power and an input power of −6 dBm. At ±500 kHz frequency offset, the

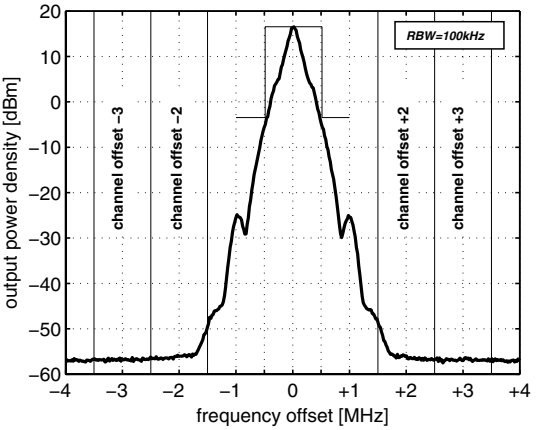


Figure 7.9. Measured Bluetooth output spectrum at maximum output power.

transmitted power is $-22.1\,\mathrm{dBc}$ and $-21.25\,\mathrm{dBc}$, which is compliant with the $-20\,\mathrm{dBc}$ specifications. The adjacent channel power is obtained by integrating the power received in a $1\,\mathrm{MHz}$ band. The adjacent channel power is $-46\,\mathrm{dBm}$ for a channel offset of two and $-47\,\mathrm{dBm}$ for a channel offset of three, surpassing the Bluetooth specifications of $-20\,\mathrm{dBm}$ and $-40\,\mathrm{dBm}$ adjacent channel power. Table 7.2 summarizes the Bluetooth measurements.

7.6 Comparison with Other Work

The comparison between this Bluetooth power amplifier and other solutions is shown in table 7.3. The table contains four published Bluetooth CMOS power amplifiers and three commercially available Bluetooth amplifiers; the T7024 from Atmel [Atm], the RF2172 from RF-MD [RFMb] and the CGB240 from TriQuint [Tri].

The presented work is the only Bluetooth power amplifier that is fully integrated and has a single-ended output. It also works at a lower supply voltage and is able to efficiently control its output power. One can also notice that the use of external striplines and external matching networks results in an increased efficiency compared to the two fully integrated designs. The efficiencies of the CMOS solutions in [Vath01], [Sowl02] and [Ho03] are comparable or even better than the more expensive technology solutions. In other words, CMOS transistors can do the job.

The power combining technique of this work outperforms the SiGe and GaAs solutions at lower power levels. This is shown in table 7.4. In theory, this

Table 7.3. Comparison with other Bluetooth power amplifiers.

Reference	Technology	P_{out} [dBm]	P_{in} [dBm]	G_P [dB]	η_d [%]	PAE [%]	V_{DD} [V]	fully integrated	differential	single-ended output
[Vath01]	0.25μm CMOS	24.0	+2	22	–	48	2.5			✔
[Sowl02]	0.18μm CMOS	23.0	-7	30	–	45	2.4			✔
[Mert02]	0.25μm CMOS	21.4	-8.6	30	38	26	2.6	✔	✔	
[Ho03]	0.35μm CMOS	23.0	–	–	–	37	1.5		✔	
[Atm]	SiGe	23.0	0	23	–	35	3			✔
[RFMb]	GaAs HBT	23.5	0	23	–	41	3.6			✔
[Tri]	InGaP HBT	23.0	+3	20	–	50	3.2			✔
this work	0.13μm CMOS	23.0	-6	29	35	29	1.5	✔	✔	✔

Table 7.4. Comparison with other Bluetooth power amplifiers: efficiency at lower output power.

		P_{out} [dBm]							
Reference	Technology	5	7	10	12	15	17	20	23
[Atm]	SiGe	4%	5%	7%	9%	13%	20%	25%	35%
[RFMb]	GaAs HBT	3%	3%	6%	9%	14%	17%	25%	41%
[Tri]	InGaP HBT	–	10%	–	20%	–	32%	–	50%
this work	0.13μm CMOS	5%	7%	11%	14%	19%	21%	27%	29%

implementation allows a two level power control, but clearly the architecture can be extended for multiple discrete power levels.

7.7 Conclusion

The main goal of this design was to demonstrate the feasability of a fully integrated Bluetooth power amplifier in a $0.13\,\mu\mathrm{m}$ technology, with a single-ended output and a capability of efficiently controlling the RF output power. If one wants to use the power combining technique to modulate the amplitude of the output RF signal, the design can be extended with the techniques discussed in section 5.5.

In section 7.2 the Bluetooth system was briefly discussed and the power amplifier requirements were given. The design of the power amplifier in a $0.13\,\mu\mathrm{m}$ CMOS technology was discussed in section 7.3. The implementation requires no expensive off-chip components and has a single-ended output. The measurement results, given in section 7.5, clearly indicate the efficiency improvement at lower output power levels. An output power of $23\,\mathrm{dBm}$ is achieved by combining two differential amplifiers in parallel, which meets the Bluetooth requirements. The spectral mask and adjacent channel power requirements are met as well. In comparison with other power amplifiers, this amplifier is the only fully integrated CMOS differential power amplifier with a single-ended output. It achieves $23\,\mathrm{dBm}$ at the lowest power supply. Compared to more expensive technologies, the power combining structure achieves a higher efficiency at lower output power. The amplifier could be further improved by using tuned driver stages.

Before concluding this chapter, table 7.5 gives a broad overview of published CMOS RF power amplifiers for wireless communications. The two implementations of this research work, the EDGE PA of chapter 6 and the Bluetooth PA of this chapter are shown at the end of the table. The Bluetooth power amplifier excells in the high integration level, single-ended output and

Table 7.5. Performance overview of published saturated CMOS power amplifiers.

Reference	Class	Technology [μm]	f [GHz]	P_{out} [dBm]	P_{out} [mW]	P_{in} [dBm]	G_P [dB]	η_d [%]	PAE [%]	V_{DD} [V]	fully integrated	differential	single-ended output
[Su97]	D	0.8	0.8	30	1000	+5	25	62	42	2.5			✔
[Tsai98]	E	0.35	1.9	30	1000	+10	20	–	48	2.0		✔	
[Yoo00]	E	0.25	0.9	29.5	900	–	–	46	41	1.8			✔
[Mert00]	E	0.35	0.7	30.0	1000	+12	18	–	60	2.3		✔	
[Kuo01]	F	0.2	0.9	31.8	1500	–	–	–	43	3.0			✔
[Shir01]	F	0.25	1.4	24.8	304	–	–	–	49	1.5			✔
[Fall01]	AB	0.35	1.8	30.4	1100	–	–	–	49	3.4			✔
[Vath01]	AB	0.25	2.4	24.0	251	+2	22	–	48	2.5			✔
[Aoki01]	B	0.35	2.4	33.4	2200	24.9	8.5	36	31	2.0	✔	✔	✔
[Zhan02]	AB	0.18	5.2	17.0	50	+3	14	–	25	1.8	✔		✔
[Sowl02]	AB	0.18	2.4	23.0	200	-7	30	–	45	2.4			✔
[Mert02]	CE	0.25	2.45	21.4	138	-8.6	30	38	25.8	2.6	✔	✔	
[Zarg02]	A	0.25	5.0	22.0	159	–	–	–	–	3.3		✔	
[Ho03]	E	0.35	2.4	23.0	200	–	–	–	37	1.5		✔	
[Aoki03]	B	0.18	1.9	34.5	2800	+17.5	17	–	50	1.8	✔	✔	✔
[Hame03]	F	0.18	8.0	22.0	158	–	–	62	–	1.0	✔		✔
[Behz03]	A	0.18	5.25	23.0	200	–	–	–	–	3.3		✔	
[Komi04]	AB	0.18	24	14.5	28.2	+12.3	2.2	11	6.5	2.8	✔		✔
[Ramo04]	E	0.35	0.85	30.0	1000	+11	19	60	60	2.3		✔	
EDGE PA (Ch. 6)	BE	0.18	1.75	27.0	500	-3	30	40	34	3.1	✔	✔	
BT PA (Ch. 7)	BE	0.13	2.45	23.0	200	-6	29	35	29	1.5	✔	✔	✔

on-chip power combining. The EDGE power amplifier is remarkable for its single chip solution, high integration level and linear behavior up to $27\,\mathrm{dBm}$.

Chapter 8

CONCLUSIONS

Wireless networks, whether used for mobile telephony or computer networks, will continue to grow. New services are introduced on existing mobile networks and new wireless systems are deployed. This evolution requires new equipment that is able to transmit and receive higher datarates. On the other hand, the cost of the mobile user devices should be kept as low a possible to ensure that new services and systems become a success and are competitive on the world market.

The power amplifier, whose task it is to amplify the radio signal as efficiently and accurately as possible, has to follow this evolution as well. Meeting the output power and linearity specifications is just not good enough for mobile communications. The power amplifier should also be as efficient as possible and produced at a low cost. This was the main motivation of this research: CMOS integration of power amplifiers for wireless and mobile communication systems.

8.1 Main Contributions and Achievements

The main theoretical contributions of this work are:

- A fast design methodology for switching amplifiers has been developed. The technique allows to design a Class E amplifier with inclusion of all losses. It also enables to investigate a deviation from Class E. A mixture between Class B and Class E, denoted as Class BE, was proposed to combine both efficiency and maximum output power in CMOS.

- The tradeoffs of the Class E and Class BE design in CMOS have been thoroughly discussed. The fast design tool allows to visualize the design space. With this visualization, the impact of CMOS scaling, the consequences of a

higher operating frequency and the influence of device stacking have been investigated.

- The benefits of power combining in CMOS have been demonstrated. The proposed lattice-type LC balun can be merged with the Class BE amplifier and as such the number of on-chip passives is reduced. This architecture results in an improved efficiency of the power amplifier and enables a discrete form of power control. This idea was demonstrated by the design and measurement of a Bluetooth CMOS power amplifier with power control.

- A detailed study of the distortion mechanisms in a polar modulated power amplifier was presented. This study has led to the development of a fully integrated polar modulated CMOS power amplifier for GSM-EDGE.

As a result of the theoretical work, two CMOS power amplifiers were successfully designed, fabricated and measured. The detailed design procedures along with the measurement results have been reported. A summary of these implementations is given below.

- An integrated CMOS power amplifier for GSM-EDGE at $1.75\,\mathrm{GHz}$ is integrated in a $0.18\,\mu\mathrm{m}$ CMOS technology and meets the stringent GSM-EDGE specifications. To efficiently amplify the non-constant envelope EDGE signal, a polar modulation architecture was developed. The amplifier achieves a peak output power of $27\,\mathrm{dBm}$ with an overall efficiency of $34\,\%$. When transmitting EDGE signals, the amplifier achieves an overall efficiency of $22\,\%$ at an output power of $23.8\,\mathrm{dBm}$ or $240\,\mathrm{mW}$. The EVM_{RMS} and spectral mask requirements are met over a power range of $7.8\,\mathrm{dB}$. The amplifier is fully integrated and requires no off-chip RF components, apart from a balun if a single-ended operation is required. Furthermore, the amplifier is less sensitive to the packaging due to the on-chip impedance matching network. The RF amplifier has an input power of only $-3\,\mathrm{dBm}$, which poses no difficulties to the upconversion mixers and thus allows a further degree of integration. The amplitude modulator is integrated on the same chip and allows the amplification of amplitude modulated signals.

- A fully integrated CMOS power amplifier for Bluetooth at $2.45\,\mathrm{GHz}$ is fully integrated in a $0.13\,\mu\mathrm{m}$ CMOS technology. This solution requires no off-chip RF components, is fully differential and has a single-ended output. The power amplifier achieves an output power of $23\,\mathrm{dBm}$ with an overall efficiency of $29\,\%$ at $1.5\,\mathrm{V}$ and $-6\,\mathrm{dBm}$ input power. A lattice-type LC power combining network is integrated on the CMOS chip and allows the parallel connection of four amplifiers. It also enables to efficiently control the transmitted output power. This power amplifier demonstrates the fea-

sibility of on-chip power combining and power control in CMOS, and it clears the road towards full digital linearization schemes.

8.2 Epilogue

Today, power amplification has become a cocktail of old vacuum tube techniques and modern digital signal processing. After many years of research, it is still an exciting area with many different roads and tradeoffs and endless discussions. CMOS has proven its capabilities in efficiently amplifying radio signals, but still a whole area of architectural challenges lies ahead.

Before drawing the final line underneath this work, I sincerely hope that the presented research contains an added value to both the novel and experienced designer, and that it may be a source from which new techniques in power amplification are explored.

List of Abbreviations and Symbols

Abbreviations

AC	Alternating Current
ACP	Adjacent Channel Power
ACPR	Adjacent Channel Power Ratio
AM	Amplitude Modulation
BALUN	Balanced to Unbalanced
BJT	Bipolar Junction Transistor
BPSK	Binary Phase Shift Keying
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CMOS	Complementary Metal Oxide Semiconductor
DC	Direct Current
DCS	Digital Cellular System
DRV	Driver
DSP	Digital Signal Processing (Processor)
EDGE	Enhanced Data rate for GSM Evolution
FDMA	Frequency Division Multiple Access
FM	Frequency Modulation
GaAs	Gallium Arsenide
GFSK	Gaussian Frequency Shift Keying
GSM	Global System for Mobile communication, Groupe Spéciale Mobile
HBT	Heterojunction Bipolar Transistor
IM3	Third Order Intermodulation Distortion
IMD	Intermodulation Distortion
InP	Indium Phosphide
ISM	Industrial Scientific and Medical
LF-PA	Low Frequency Power Amplifier
MIM	Metal Insulator Metal (capacitor)
MOS	Metal Oxide Semiconductor
MOST	Metal Oxide Semiconductor Transistor
nMOS	n-channel MOS transistor
NPO	Negative-Positive-Zero (capacitor)
OPAMP	Operational Amplifier

OTA	Operational Transconductance Amplifier
PA	Power Amplifier
PCS	Personal Communications Service
PDF	Probability Density Function
PLL	Phase Locked Loop
PM	Phase Modulation
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RBW	Resolution Bandwidth
RF	Radio Frequency
RF-PA	Radio Frequency Power Amplifier
RMS	Root Mean Square
SIA	Semiconductor Industry Association
SiGe	Silicon Germanium
SMA	Sub-Miniature version A (connector)
SSB	Single Sideband
TDDB	Time Dependent Dielectric Breakdown
TDMA	Time Division Multiple Access
UMTS	Universal Mobile Telecommunications System
VSWR	Voltage Standing Wave Ratio
WCDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network

Symbols

$A(t)$	envelope signal, amplitude signal	[−] or [V]
A_{max}	maximum value of the envelope or amplitude signal	[−] or [V]
A_{rms}	RMS value of the envelope or amplitude signal	[−] or [V]
c	speed of light in vacuum	[m/s]
C	capacitor	[F]
C_1	shunt capacitance of the Class E amplifier	[F]
C_{BL}	DC-blocking capacitance	[F]
C_d	total drain capacitance	[F]
C_{db}	drain-bulk capacitance	[F]
CF	crest factor	[−] or [dB]
C_g	total gate capacitance	[F]
C_{gd}	gate-drain capacitance	[F]
C_{gs}	gate-source capacitance	[F]
C_P	power capability	[−]
E	power enhancement ratio	[−]
EVM	magnitude of the error vector	[−] or [%]
EVM_{RMS}	RMS value of the EVM	[−] or [%]
$e^{j\theta(t)}$	complex phase signal	[−] or [V]
$e^{jP(t)}$	complex phase signal	[−] or [V]
f	frequency	[Hz]
f_c	carrier frequency	[Hz]
g_m	transistor transconductance	[S]
$g(t)$	complex envelope signal, complex baseband signal	[−] or [V]

G_P	power gain	[dB]
$i_{DS}(t)$	drain-source current	[A]
IL	insertion loss	[−]
IL_{dB}	insertion loss	[dB]
$IP3$	Third Order Intercept Point	[dBm]
I_Q	quiescent current	[A]
$I(t)$	in-phase signal	[−] or [V]
L	inductor	[H]
L_1	DC-feed inductance of the Class E amplifier	[H]
L_g	transistor gate length	[μm]
L_x	excess inductance of the Class E amplifier	[H]
N	number of parallel PA stages	[−]
$P(t)$	phase signal	[rad]
PAE	power added efficiency	[−] or [%]
$PAPR$	peak to average power ratio	[−] or [dB]
PEP	peak envelope output power	[W] or [dBm]
P_{in}	RF input power	[W] or [dBm]
$P_{o,0}$	RF output power, for one PA directly connected to $50\,\Omega$	[W] or [dBm]
$P_{o,max}$	maximum RF output power that occurs at $v_{DS,max}$	[W] or [dBm]
$P_{o,tot}$	total RF output power, including harmonic power	[W] or [dBm]
P_{o,f_0}	total RF output power at the fundamental frequency f_0	[W] or [dBm]
P_o	total RF output power at the fundamental frequency	[W] or [dBm]
P_{DC}	DC power consumption	[W]
$P_{DC,DRV}$	DC power consumption of the driver stage(s)	[W]
$P_{DC,PA}$	DC power consumption of the power amplifier stage	[W]
P_{diss}	dissipated power	[W]
Q	quality factor	[−]
$Q(t)$	quadrature signal	[−] or [V]
R_L	load resistor	[Ω]
r	impedance transformation ratio	[−]
r_{on}	resistance of the nMOS switch	[Ω]
R	resistor	[Ω]
$T(j\omega)$	loopgain	[−] or [dB]
T_d	time delay (of a filter)	[s]
T_g	group delay (of a filter)	[s]
V_{DD}	supply voltage	[V]
$V_{DD,AM}$	supply voltage of the amplitude modulator	[V]
$V_{DD,DRV}$	supply voltage of the driver stage	[V]
$V_{DD,nom}$	nominal supply voltage of the technology	[V]
$V_{DD,PA}$	supply voltage of the power amplifier	[V]
$v(t)$	modulated RF signal	[−] or [V]
$v_{DS}(t)$	drain-source voltage	[V]
$v_{DS,max}$	maximum or peak drain-source voltage	[V]
$v_P(t)$	RF phase signal	[−] or [V]
V_{th}	threshold voltage	[V]
$V_{th,n}$	threshold voltage of the nMOS transistor	[V]
V_o	amplitude of the sinusoidal output voltage	[V]
$V_{o,rms}$	RMS value of the sinusoidal output voltage	[V]
W_g	transistor gate width	[μm]

$x(t)$	in-phase signal	$[-]$ or $[\mathrm{V}]$
$y(t)$	quadrature signal	$[-]$ or $[\mathrm{V}]$
Z_C	characteristic impedance	$[\Omega]$
Z_L	load impedance	$[\Omega]$
α	inductor loss factor	$[\Omega/\mathrm{nH}]$
Γ	reflection coefficient	
η_{conv}	conversion efficiency	$[-]$ or $[\%]$
η_d	efficiency, drain efficiency, collector efficiency	$[-]$ or $[\%]$
η_{oa}	overall efficiency	$[-]$ or $[\%]$
$\theta(t)$	phase signal	rad
$\theta(\omega)$	phase response of a filter	rad
λ	wavelength	$[\mathrm{m}]$
ω	angular frequency	$[\mathrm{rad/s}]$
ω_0	resonance frequency	$[\mathrm{rad/s}]$
ω_{3dB}	$-3\,\mathrm{dB}$ frequency of a filter	$[\mathrm{rad/s}]$
ω_c	switching or carrier frequency	$[\mathrm{rad/s}]$

References

[Abid04] A. A. Abidi, "RF CMOS Comes of Age", *IEEE Journal of Solid-State Circuits*, vol. 39, no. 4, pp. 549–561, April 2004.

[Alin99] P. Alinikula, K. Choi, and S. I. Long, "Design of Class E Power Amplifier with Nonlinear Parasitic Output Capacitance", *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 46, no. 2, pp. 114–118, February 1999.

[Andr99] P. Andreani, L. Sundstrom, N. Karlsson, and M. Svensson, "A Chip for Linearization of RF Power Amplifiers Using Digital Predistortion with a Bit-Parallel Complex Multiplier", *Journal of Analog Integrated Circuits and Signal Processing*, vol. 22, pp. 25–30, 1999.

[Anne01] Anne-Johan Annema, Govert Geelen, and Peter de Jong, "5.5 V I/O in a 2.5 V 0.25μm CMOS technology", *IEEE Journal of Solid-State Circuits*, vol. 36, no. 3, pp. 528–538, March 2001.

[Aoki01] I. Aoki, S. D. Kee, D. Rutledge, and A. Hajimiri, "A 2.4-GHz, 2.2-W, 2-V Fully-Integrated CMOS Circular-Geometry Active-Transformer Power Amplifier", in *Proceedings of the Custom Integrated Circuits Conference*, May 2001.

[Aoki02a] I. Aoki, S. D. Kee, D. B. Rutledge, and A. Hajimiri, "Distributed Active Transformer—A New Power-Combining and Impedance-Transformation Technique", *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 1, January 2002.

[Aoki02b] I. Aoki, S. D. Kee, D. B. Rutledge, and A. Hajimiri, "Fully Integrated CMOS Power Amplifier Design Using the Distributed Active Transformer Architecture", *IEEE Journal of Solid-State Circuits*, vol. 37, no. 3, pp. 371–383, March 2002.

[Aoki03] I. Aoki, S. Kee, D. Rutledge, and A. Hajimiri, "A Fully-Integrated 1.8-V, 2.8-W, 1.9-GHz, CMOS Power Amplifier", in *IEEE Transactions on Microwave Theory and Techniques*, pp. 199–202, 2003.

[Ath] Atheros Communications, *http://www.atheros.com/*.

[Atm] Atmel, T-7024 Datasheet, *http://www.atmel.com/*.

[Avra89] C. P. Avratoglou, N. C. Voulgaris, and F. I. Ioannidou, "Analysis and Design of a Generalized Class E Tuned Power Amplifier", *IEEE Transactions on Circuits and Systems*, vol. 36, no. 8, pp. 1068–1079, August 1989.

[Axi] Axiom Microdevices, *http://www.axiom-micro.com/*.

[Baka02] W. Bakalski, W. Simbürger, H. Knapp, H Wohlmuth, and A. L. Scholtz, "Lumped and Distributed Lattice-type LC-Baluns", in *Proceedings of the International Microwave Symposium*, volume 1, pp. 209–212, June 2002.

[Behz03] A. R. Behzad et al., "A 5-GHz Direct-Conversion CMOS Transceiver Utilizing Automatic Frequency Control for the IEEE 802.11a Wireless LAN Standard", *IEEE Journal of Solid-State Circuits*, vol. 38, no. 12, pp. 2209–2220, December 2003.

[Boum03] S. Boumaiza and F. M. Ghanouchi, "Thermal Memory Effects Modeling and Compensation in RF Power Amplifiers and Predistortion Linearizers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 12, pp. 2427–2433, December 2003.

[Bro] Broadcom, AirForce One Transceiver, *http://www.broadcom.com/*.

[BT32] C. Lorenz AG Berlin-Tempelhof, "Schaltungsanordnung zum Übergang von einer symmetrischen, elektrischen Anordnung zu einer unsymmetrischen, insbesondere bei Hochfrequenzanwendungen", German Patent no. 603816, April 1932.

[Catr04] M. Catrysse, *Wireless Power and Data Transmission for Implantable and Wearable Monitoring Systems*, PhD thesis, K.U.Leuven, Belgium, December 2004.

[Ceyl04] N. Ceylan, J. E. Mueller, and R. Weigel, "A New Addressing Method for Look-up Table Based Digital Predistortion Linearizers", in *Proceedings of the International Symposium on Systems, Signals and Electronics*, August 2004.

[Chir35] H. Chireix, "High power outphasing modulation", *IRE*, vol. 23, no. 11, pp. 1370–1392, 1935.

[Couc97] L. W. Couch II, *Digital And Analog Communication Systems*, Prentice Hall, 1997.

[Cox74] D. C. Cox, "Linear Amplification with Nonlinear Components", *IEEE Transactions on Communications*, vol. COM-23, no. 11, pp. 1942–1945, December 1974.

[Crip99] S. C. Cripps, *RF Power Amplifiers for Wireless Communication*, Artech House, 1999.

[Deut04] Deutsche Bank, "GSM White paper: Brilliant Past, Bright Future", February 2004, http://www.gsmworld.com/news/press_2004/press04_11.shtml.

[Ding05] Y. Ding and R. Harjani, "A High-Efficiency CMOS +22-dBm Linear Power Amplifier", *IEEE Journal of Solid-State Circuits*, vol. 40, no. 9, pp. 1895–1900, September 2005.

[Doan04] C. H. Doan, S. Emami, A. M. Niknejad, and R. W. Brodersen, "Design of CMOS for 60GHz applications", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 440–441, February 2004.

[Dohe36] W. H. Doherty, "A New High Efficiency Power Amplifier for Modulated Waves", *Proceedings IRE*, vol. 24, no. 9, pp. 1163–1182, September 1936.

[Dupu02] T. J. Dupuis and S. A. Paul, "RF Power Amplifier Circuitry and Method for Amplifying Signals", United Stated Patent 6,462,620 to Silicon Laboratories, Inc., Patent and Trademark Office, Washington, D.C., October 2002.

[Elli04] M. R. Elliott, T. Montalvo, B. P. Jeffries, F. Murden, J. Strange, A. Hill, S. Nandipaku, and J. Harrebek, "A Polar Modulator Transmitter for GSM/EDGE", *IEEE Journal of Solid-State Circuits*, vol. 39, no. 12, pp. 2190–2199, December 2004.

[ETSI01] ETSI, "ETSI TS 145 005 v4.6.0 (Radio transmission and reception)", November 2001.

[Fall01] C. Fallesen and P. Asbeck, "A 1W $0.35\mu m$ CMOS Power Amplifier for GSM-1800 with 45% PAE", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 158–159, San Francisco, USA, February 2001.

[Fort01] F. Fortes and M.J. do Rosário, "A Second Harmonic Class-F Power Amplifier in Standard CMOS Technology", *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 6, pp. 1216–1220, June 2001.

[Fran00] G. Francois, "Bouwblokken in de Elektronica - II", Dutch coursetext, 1999-2000.

[Gupt01] R. Gupta, B. M. Ballweber, and D. J. Allstot, "Design and Optimization of CMOS RF Power Amplifiers", *IEEE Journal of Solid-State Circuits*, vol. 36, no. 2, pp. 166–174, February 2001.

[Hadj04] A. Hadjichristos, J. Walukas, N. Klemmer, W. Suter, S. Justice, S. Uppathil, and G. Scott, "A Highly Integrated Quad Band Low EVM Polar Modulation Transmitter for GSM/EDGE Applications", in *Proceedings of the Custom Integrated Circuits Conference*, May 2004.

[Hame03] S. Hamedi-Hagh and C. A. T. Salama, "A 1V, 8GHz CMOS Integrated Phase Shifted Transmitter for Wideband and Varying Envelope Communication Systems", in *Proceedings of the Custom Integrated Circuits Conference*, September 2003.

[Hani99] G. Hanington, P. Chen, P. M. Asbeck, and L. E. Larson, "High-Efficiency Power Amplifier Using Dynamic Power-Supply Voltage for CDMA Applications", *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 8, pp. 1471–1476, August 1999.

[HH04] S. Hamedi-Hagh and C. A. T. Salama, "CMOS Wireless Phase-Shifted Transmitter", *IEEE Journal of Solid-State Circuits*, vol. 39, no. 8, pp. 1241–1252, August 2004.

[Ho03] K.-W. Ho and H. C. Luong, "A 1-V CMOS Power Amplifier for Bluetooth Applications", *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 50, no. 8, pp. 445–449, August 2003.

[Inge97] M. Ingels and M. Steyaert, "Design Strategies and Decoupling Techniques for reducing the effects of Electrical Interference in Mixed-Mode ICs", *IEEE Journal of Solid-State Circuits*, vol. 32, no. 7, pp. 1136–1141, July 1997.

[Jaco] H. Jacobsson, *Low-COST CMOS for Microwave Radio Communication*, Ericsson.

[Kahn52] L. R. Kahn, "Single Sideband Transmission by Envelope Elimination and Restoration", *Proceedings IRE*, vol. 40, no. 7, pp. 803–806, July 1952.

[Kazi84] M. K. Kazimierczuk, "Collector Amplitude Modulation of the Class E Tuned Power Amplifier", *IEEE Transactions on Circuits and Systems*, vol. 31, no. 6, pp. 543–549, June 1984.

[Kazi86] M. K. Kazimierczuk, "Class E Tuned Power Amplifier with Nonsinusoidal Output Voltage", *IEEE Journal of Solid-State Circuits*, vol. sc-12, no. 4, pp. 575–581, August 1986.

[Kazi87a] M. Kazimierczuk, "Exact Analysis of Class E Tuned Power Amplifier with Only One Inductor and One Capacitor in Load Network", *IEEE Journal of Solid-State Circuits*, vol. sc-18, no. 2, pp. 214–221, February 1987.

[Kazi87b] M. K. Kazimierczuk and K. Puczko, "Exact Analysis of Class E Tuned Power Amplifier at any Q and Switch Duty Cycle", *IEEE Transactions on Circuits and Systems*, vol. cas-34, no. 2, pp. 149–159, February 1987.

[Kazi89] M. K. Kazimierczuk and W. A. Tabisz, "Class C-E High-Efficiency Tuned Power Amplifier", *IEEE Transactions on Circuits and Systems*, vol. 36, no. 3, pp. 421–428, March 1989.

[Keni00] P. B. Kenington, *High-Linearity RF Amplifier Design*, Artech House, 2000.

[Komi04] A. Komijani and A. Hajimiri, "A 24GHz, +14.5dBm Fully-Integrated Power Amplifier in 0.18μm CMOS", in *Proceedings of the Custom Integrated Circuits Conference*, May 2004.

[Krau80] H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, Wiley, 1980.

[Kuo01] T. C. Kuo and B. B. Lusignan, "A 1.5W Class-F Power Amplifier in 0.2 μm CMOS Technology", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 154–155, San Francisco, USA, February 2001.

[Kusu02] S. Kusunoki, K. Yamamoto, T. Hatsugai, H. Nagaoka, K. Tagami, N. Tominaga, K. Osawa, K. Tanabe, S. Sakurai, and T. Iida, "Power-Amplifier Module With Digital Adaptive Predistortion for Cellular Phones", *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 12, pp. 2979–2986, December 2002.

[Li94] C. H. Li and Y. O. Yam, "Maximum frequency and optimum performance of class E power amplifiers", *IEE Proceedings on Circuits, Devices and Systems*, vol. 141, no. 3, pp. 174–184, June 1994.

[McCu03] E. McCune and W. Sander, "EDGE Transmitter Alternative Using Nonlinear Polar Modulation", in *Proceedings International Symposium on Circuits and Systems*, pp. III–594–III–597, May 2003.

[Mert00] K. Mertens, M. Steyaert, and B. Nauwelaers, "A 700-MHz, 1-W Fully Integrated Differential Class E Power Amplifier in CMOS", in *Proceedings of the European Solid-State Circuits Conference*, September 2000.

[Mert01] K. Mertens, P. Reynaert, and M. Steyaert, "Performance Study of CMOS Power Amplifiers", in *Proceedings of the European Solid-State Circuits Conference*, September 2001.

[Mert02] K. Mertens and M. Steyaert, "A Fully Integrated Class 1 Bluetooth 0.25μm CMOS PA", in *Proceedings of the European Solid-State Circuits Conference*, September 2002.

[Mert05] K. Mertens, *Implementation of Fully Integrated RF-CMOS Power Amplifiers*, PhD thesis, K.U.Leuven, Belgium, October 2005.

[Moor65] G. E. Moore, "Cramming more components onto integrated circuits", *Electronics*, vol. 38, no. 8, April 1965.

[Otis05] B. Otis, Y. H. Chee, and J. Rabaey, "A 400 μW-RX, 1.6mW-TX Super-Regenerative Transceiver for Wireless Sensor Networks", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 396–397, February 2005.

[Paul03] S. A. Paul and T. J. Dupuis, "Power Amplifier Circuitry and Method Using An Inductance Coupled To Power Amplifier Switching Devices", United Stated Patent 6,549,071 to Silicon Laboratories, Inc., Patent and Trademark Office, Washington, D.C., April 2003.

[Pehl04] D. R. Pehlke, A. Hadjichristos, and S. Justice, "High Performance Open-Loop AM Modulator Designed for Power Control of E-GPRS Polar Modulated Power Amplifier", in *Proceedings of the Custom Integrated Circuits Conference*, May 2004.

[Pies01] T. Piessens and M. Steyaert, "SOPA: A high-efficiency line driver in 0.35μm CMOS using a self-oscillating power amplifier", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 306–307, February 2001.

[Raab75] F. H. Raab, "High Efficiency Amplification Techniques", *IEEE Transactions on Circuits and Systems*, vol. 7, pp. 3–11, April 1975.

[Raab77] F. H. Raab, "Idealized Operation of the Class E Tuned Power Amplifier", *IEEE Transactions on Circuits and Systems*, vol. cas-24, no. 12, pp. 725–735, December 1977.

[Raab78] F. H. Raab and N. O Sokal, "Transistor Power Losses in the Class E Tuned Power Amplifier", *IEEE Journal of Solid-State Circuits*, vol. sc-13, no. 6, pp. 912–914, December 1978.

[Raab96] F. H. Raab, "Intermodulation Distortion in Kahn-Technique Transmitters", *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, no. 12, pp. 2273–2278, December 1996.

[Raab97] F. H. Raab, "Class-F Power Amplifiers with Maximally Flat Waveforms", *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, no. 11, pp. 2007–2012, November 1997.

[Raab01] F. H. Raab, "Maximum Efficiency and Output Power of Class F Power Amplifiers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 6, pp. 1162–1166, June 2001.

[Raab02] F. H. Raab, P. Asbeck, S. Cripps, P. B. Kenington, Z. B. Popović, N. Pothecary, J. F. Sevic, and N. O. Sokal, "Power Amplifiers and Transmitters for RF and Microwave", *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 814 – 826, March 2002.

[Ramo04] J. Ramos, K. Francken, G. Gielen, and M. Steyaert, "Knowledge- and Optimization-Based Design of RF Power Amplifiers", in *Proceedings International Symposium on Circuits and Systems*, pp. 23–26, Vancouver, Canada, May 2004.

[Ramo05] J. Ramos, *CMOS Operational and RF Power Amplifiers for Mobile Communications*, PhD thesis, K.U.Leuven, Belgium, March 2005.

[Reyn02] P. Reynaert, K. Mertens, and M. Steyaert, "Optimizing the Dimensions of Driver and Power Transistor in Switching CMOS RF Power Amplifiers", *Analog Integrated Circuits and Signal Processing*, vol. 32, pp. 177–182, August 2002.

[Reyn03a] P. Reynaert, K. Mertens, and M. Steyaert, "A State-Space Behavioural Model for CMOS Class-E Power Amplifiers", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, no. 2, pp. 132–138, February 2003.

[Reyn03b] P. Reynaert and M. Steyaert, "Towards Fully Integrated CMOS RF Transceivers", in *Proceedings of the International Symposium on VLSI-TSA*, October 2003.

[Reyn04a] P. Reynaert and M. Steyaert, "Distortion in Supply Voltage Modulated CMOS RF Power Amplifiers", in *Proceedings of the International Symposium on Systems, Signals and Electronics*, August 2004.

[Reyn04b] P. Reynaert and M. Steyaert, "Fully Integrated CMOS RF Transceivers and Power Amplifiers", in *Proceedings of the Asia-Pacific Microwave Conference*, December 2004.

[Reyn05a] P. Reynaert and M. Steyaert, "A 1.75 GHz GSM-EDGE Polar Modulated CMOS RF Power Amplifier", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 312–313, February 2005.

[Reyn05b] P. Reynaert and M. Steyaert, "A 1.75 GHz Polar Modulated CMOS RF Power Amplifier for GSM-EDGE", *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2598–2608, December 2005.

[Reyn05c] P. Reynaert and M. Steyaert, "A Fully Integrated CMOS RF Power Amplifier with Parallel Power Combining and Power Control", in *Proceedings of the Asian Solid-State Circuits Conference*, pp. 137–140, November 2005.

[RFMa] RF Micro Devices, *http://www.rfmd.com/*.

[RFMb] RF Micro Devices, RF-2172 Datasheet, *http://www.rfmd.com/*.

[rGPP01] 3rd Generation Partnership Project, "3GPP TS 45.004 v5.0.0 (Modulation)", November 2001.

[Sahu04] B. Sahu and G. A. Rincón-Mora, "A High-Efficiency Linear RF Power Amplifier With a Power-Tracking Dynamically Adaptive Buck-Boost Supply", *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 1, pp. 112–120, January 2004.

[Sand03] W. B. Sander, S. V. Schell, and B. L. Sander, "Polar Modulator for Multi-mode Cell Phones", in *Proceedings of the Custom Integrated Circuits Conference*, September 2003.

[Schl04] N. Schlumpf and M. Declercq, "A Fast Modulator for Dynamic Supply Linear RF Power Amplifier", *IEEE Journal of Solid-State Circuits*, vol. 39, no. 7, pp. 1015–1025, July 2004.

[Schu98] K. Van Schuylenbergh, *Optimisation of Inductive Powering of Small Biotelemetry Implants*, PhD thesis, K.U.Leuven, Belgium, May 1998.

[Sern05] Bert Serneels, Tim Piessens, Michiel Steyaert, and Wim Dehaene, "A high-voltage output driver in a standard 2.5V 0.25μm CMOS technology", *IEEE Journal of Solid-State Circuits*, vol. 40, no. 3, pp. 576–583, March 2005.

[Shir01] A. Shirvani, D. K. Su, and B. A Wooley, "A CMOS RF Power Amplifier with Parallel Amplification for Efficient Power Control", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 156–157, San Francisco, USA, February 2001.

[SIA05] SIA, "Semiconductor Industry Association, 2005 Annual Report", 2005.

[SIG04] Bluetooth SIG, "Specification of the Bluetooth System, v2.0 + EDR", November 2004.

[Sila] Silicon Laboratories, Si4300 Datasheet, *http://www.silabs.com/*.

[Silb] Silicon Laboratories, Si4905 Datasheet, *http://www.silabs.com/*.

[Smit90] G. H. Smith and R. E. Zulinski, "An Exact Analysis of Class E Amplifiers with Finite DC-Feed Inductance at Any Output Q", *IEEE Transactions on Circuits and Systems*, vol. 37, no. 4, pp. 530–534, April 1990.

[Snid67] D. M. Snider, "A Theoretical Analysis and Experimental Confirmation of the Optimally Loaded and Overdriven RF Power Amplifier", *IEEE Transactions on Electron Devices*, vol. ED-14, no. 12, pp. 851–857, December 1967.

[Soka75] N. O. Sokal and A. D. Sokal, "Class E - A New Class of High-Efficiency Tuned Single-Ended Switching Power Amplifiers", *IEEE Journal of Solid-State Circuits*, vol. sc-10, no. 3, pp. 168–176, June 1975.

[Soka81] N. O. Sokal, "Class E High-Efficiency Switching-Mode Tuned Power Amplifier with Only One Inductor and One Capacitor in Load Network - Approximate Analysis", *IEEE Journal of Solid-State Circuits*, vol. sc-16, no. 4, pp. 380–384, August 1981.

[Sowl97] T. Sowlati and Y. M. Greshishchev amd C. A. T. Salama, "Phase-Correcting Feedback System for Class E Power Amplifier", *IEEE Journal of Solid-State Circuits*, vol. 32, no. 4, pp. 544–550, April 1997.

[Sowl02] T. Sowlati and D. Leenaerts, "A 2.4GHz $0.18\mu m$ CMOS Self-Biased Cascode Power Amplifier with 23dBm Output Power", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 294–295, San Francisco, USA, February 2002.

[Sowl04] T. Sowlati, D. Rozenblit, R. Pullela, M. Damgaard, E. McCarthy, D. Koh, D. Ripley, F. Balteanu, and I. Gheorghe, "Quad-Band GSM/GPRS/EDGE Polar Loop Transmitter", *IEEE Journal of Solid-State Circuits*, vol. 39, no. 12, pp. 2179–2189, December 2004.

[Stas05] R. B. Staszewski, J. Wallberg, S. Rezeq, C-M. Hung, O. Eliezer, S. Vemulapalli, C. Fernando, K. Maggio, R. Staszewski, N. Barton, M-C. Lee, P. Cruise, M. Entezari, K. Muhammad, and D. Leipold, "All-digital PLL and GSM/EDGE Transmitter in 90nm CMOS", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 316–317, February 2005.

[Stey98] M. Steyaert, M. Borremans, J. Janssens, B. De Muer, and N. Itoh, "A Single-Chip CMOS Transceiver for DCS-1800 Wireless Communications", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 48–49, San Francisco, USA, February 1998.

[Stey00] M. Steyaert, J. Janssens, B. De Muer, M. Borremans, and N. Itoh, "A 2-V CMOS Cellular Transceiver Front-end", in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 142–143, San Francisco, USA, February 2000.

[STM] ST Microelectronics, *http://www.st.com/*.

[Su97] D. Su and W. McFarland, "A 2.5V, 1-W Monolithic CMOS RF Power Amplifier", in *Proceedings of the Custom Integrated Circuits Conference*, pp. 189–192, 1997.

[Su98] D. K. Su and W. J. McFarland, "An IC for Linearizing RF Power Amplifiers Using Envelope Elimination and Restoration", *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 2252–2258, December 1998.

[Teik04] I. Teikari, J. Vankka, and K. Halonen, "Digitally Controlled RF Predistortion Based On Vector Modulator", in *Proceedings of the International Symposium on Systems, Signals and Electronics*, August 2004.

[Tri] TriQuint, CGB-240 Datasheet, *http://www.triquint.com/*.

[Tsai98] K.-C. Tsai and P. R. Gray, "A 1.9-GHz 1-W CMOS Class E Power Amplifier for Wireless Communications", in *Proceedings of the European Solid-State Circuits Conference*, pp. 76–79, September 1998.

[Vath01] V. R. Vathulya, T. Sowlati, and D. Leenaerts, "Class 1 Bluetooth Power Amplifier with 24 dBm Output Power and 48% PAE at 2.4 GHz in $0.25\mu m$ CMOS", in *Proceedings of the European Solid-State Circuits Conference*, pp. 84–87, September 2001.

[Wagh04] P. Wagh, P. Midya, P. Rakers, J. Caldwell, and T. Schooler, "An All-digital Universal RF Transmitter", in *Proceedings of the Custom Integrated Circuits Conference*, May 2004.

[Wang05] F. Wang, D. Kimball, J. Popp, A. Yang, D. Y. C. Lie, P. Asbeck, and L. Larson, "Wideband Envelope Elimination and Restoration Power Amplifier with High Efficiency Wideband Envelope Amplifier for WLAN 802.11g Applications", in *Proceedings of the International Microwave Symposium*, June 2005.

[Wir05] Wireless Web, http://wireless.iop.org/articles/news/6/10/13/1, *EDGE networks continue to boom*, 2005.

[Yang99] K. Yang, G. I. Haddad, and J. R. East, "High-Efficiency Class-A Power Amplifiers with a Dual-Bias-Control Scheme", *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 8, pp. 1426–1432, August 1999.

[Yoo00] C. Yoo and Q. Huang, "A common-gate switched, 0.9W class-E power amplifier with 41% PAE in 0.25μm CMOS", in *Symposium on VLSI Circuits Digest of Technical Papers*, pp. 56–57, June 2000.

[Yoo01] C. Yoo and Q. Huang, "A Common-Gate Switched 0.9-W Class-E Power Amplifier with 41% PAE in $0.25 - \mu m$ CMOS", *IEEE Journal of Solid-State Circuits*, vol. 36, no. 3, pp. 823–830, May 2001.

[Yue98] C. P. Yue and S. S. Wong, "On-Chip Spiral Inductors with Patterned Ground Shields for Si-Based RF IC's", *IEEE Journal of Solid-State Circuits*, vol. 33, no. 5, pp. 743–752, May 1998.

[Zarg02] M. Zargari et al., "A 5-GHz CMOS Transceiver for IEEE 802.11a Wireless LAN Systems", *IEEE Journal of Solid-State Circuits*, vol. 37, no. 12, pp. 1688–1694, December 2002.

[Zarg04] M. Zargari et al., "A Single-Chip Dual-Band Tri-Mode CMOS Transceiver for IEEE 802.11a/b/g Wireless LAN", *IEEE Journal of Solid-State Circuits*, vol. 39, no. 12, pp. 2239–2249, December 2004.

[Zhan02] W. Zhang, E.-S. Khoo, and T. Tear, "A Low Voltage Fully-Integrated 0.18μm CMOS Power Amplifier for 5-GHz WLAN", in *Proceedings of the European Solid-State Circuits Conference*, pp. 215–218, September 2002.

[Zhan03] P. Zhang et al., "A 5-GHz Direct-Conversion CMOS Transceiver", *IEEE Journal of Solid-State Circuits*, vol. 38, no. 12, pp. 2232–2238, December 2003.

[Zuli86] R. E. Zulinski and J. W. Steadman, "Idealized Operation of Class E Frequency Multipliers", *IEEE Transactions on Circuits and Systems*, vol. 33, no. 12, pp. 1209–1218, December 1986.

[Zuli87] R. E. Zulinski and J. W. Steadman, "Class E Power Amplifiers and Frequency Multipliers with Finite DC-Feed Inductance", *IEEE Transactions on Circuits and Systems*, vol. 34, no. 9, pp. 1074–1087, September 1987.

Index