Highly Efficient and Highly Linear RF Power Amplifiers

Weng Chuen Edmund Neo

Highly Efficient and Highly Linear RF Power Amplifiers

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. ir. K. C. A. M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen

op maandag 8 maart 2010 om 15:00 uur

door

Weng Chuen Edmund NEO

elektrotechnisch ingenieur,
geboren te Singapore

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. Ing. J. N. Burghartz

Copromotor:
Dr. ing. L. C. N. de Vreede

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
Prof. dr. Ing. J. N. Burghartz, Technische Universiteit Delft, promotor
Dr. ing. L. C. N de Vreede, Technische Universiteit Delft, copromotor
Prof. dr. H. F. F Jos, Chalmers University of Technology (Sweden)
Prof. dr. J. Long, Technische Universiteit Delft
Prof. dr. L. E. Larson, University of California -San Diego (USA)
Prof. dr. P. A. M. Baltus, Technische Universiteit Eindhoven
J. Gajadharsing, NXP Semiconductors, adviseur
Prof. dr. P. M. Sarro, Technische Universiteit Delft, reservelid

Weng Chuen Edmund Neo,
Highly Efficient and Highly linear RF Power Amplifiers,
Ph.D. Thesis Delft University of Technology,
with summary in Dutch.

Keywords: amplifier design, high-efficient, digital predistortion, mixed-signal, base-station, handsets, radio frequency (RF).

ISBN: 978-90-8559-934-0

Printed in The Netherlands.

"To mama
1952-2009"

Contents

Chapter 1

Introduction

The wireless communication industry has grown enormously since its humble beginnings in the early 1950's. At the start of this new century, it was estimated to be a multi-billion dollar industry and rapidly growing. The early driver for this growth is mans' needs for communication. Man being the innovative creature he is, has constantly pushed for and developed new means to communicate, from paper (the postal system) to electric current in copper cables (telephones, fax machines and later computers) and now electromagnetic waves in free-space (wireless phones). These are all brilliant innovations driven by mans' need for more effective communication. In today's world, we can all communicate with each other instantaneously across large physical distances (spanning the whole globe), anytime and anywhere in the civilized world.

The information age, which first started with radio and TV broadcast and really took off when the internet arrived, gave man for the first time in history wide access to a vast amount of information. However in the early years of the internet, and still very much valid today, the access to the internet is still limited to fix immobile physical locations. Today, there is an increasing demand to be able to access the internet anywhere, and this is becoming a strong driver behind the growth of the wireless communication industry.

Yet another area, which is earmarked to influence the wireless communication business, is machine to machine communication (loosely referred to M2M communication). M2M enabled hardware is basically a group of devices/sensors capable of responding to external request (most likely from another machine) for data it holds and are capable of transmitting these data autonomously using wireless technology. An application of such a technology could e.g. be in facility management where the status of the portable fire extinguishers are monitored wirelessly throughout a campus. It eliminates the need for manual verification of the pressure gauge of if the extinguisher is discharged.

Table 1.1 gives a generic summary of the different forms of communication just described and the typical applications that drives it. This summary is often called the "3G and Beyond", "4G vision". It is amazing to witness how far wireless communication has evolved since its humble beginnings. In the following section, we will digress

Table 1.1: Support for real time and non-real time service between humans, human-machine and machine-machine

	Human	Machine
Human	VOIP (videophone conference), interactive Games, chatting (visual mail, audio mail, text mail)	video broadcasting, video supervision, human navigation, Internet browsing, music download
Machine	remote control, recording to storage device: voice, video etc	location information services, distribution systems, data transfer, consumer electronics device maintenance

source: NTTDoCoMo

through the evolution of the wireless standards (0G to 3G) and show its increasing complexity.

1.1 From 0G to 3G and Beyond

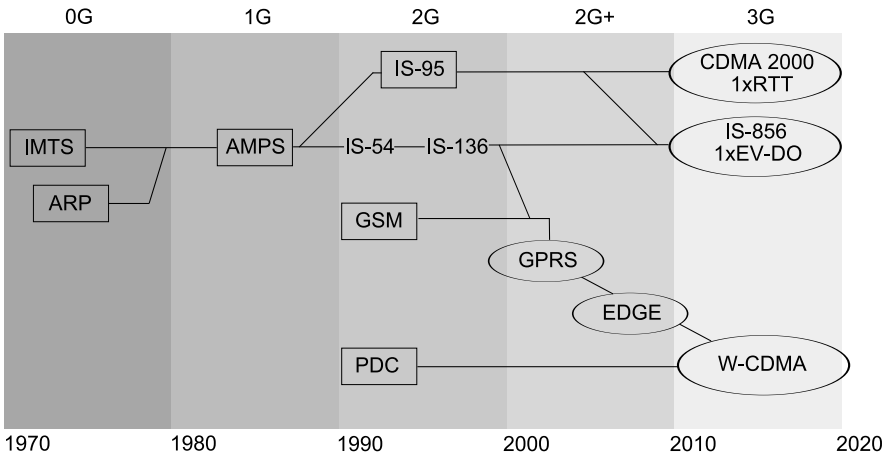


Figure 1.1: Evolution of standards from 0G to 3G

Figure 1.1 shows the development of the major standards as they evolved over the years. Every conceptual shift in wireless technology has been characterized or named as a generational change, thus the naming convention 0G, 1G, 2G, 2G+ and 3G.

The first commercial mobile communication system was perhaps the Improved Mobile Telephone Service (IMTS) that began in 1969 (0G). At that time, the concept of cellular communication (dividing physical space into hexagonal cells, resulting

in frequency reuse) was not yet developed (see Figure 1.2. Furthermore only analog modulation was used and no air access techniques were employed. The standard supported only 13 full-duplex channels and radio coverage was up to a radius of 100 km. This was at that time sufficient, as there were not many users subscribed for the service. In 1971 (0.5G), cellular communication and air access techniques were introduced into the commercial system, beginning with AutoRadioPuhelin (ARP), which was commonly known as the car radio phone system. Cells of 30 km radius were used but no hands-off system was yet in place, thus when a user crosses the boundary between two cells, the connection was lost. This system uses Frequency Division Multiple Access (FDMA) techniques and could support up to 80 fully duplexed channels.

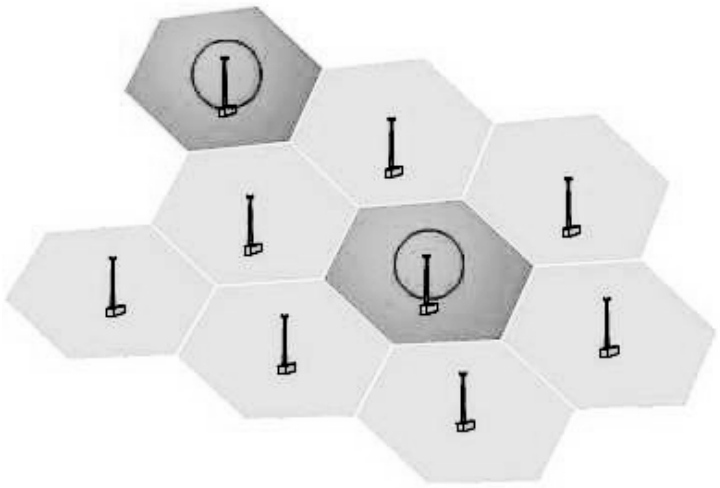


Figure 1.2: Physical area where wireless communication is supported is divided into hexagonal cells. Base-station in adjacent cells transmits on a different frequency preventing interference.

Then in the early 1980s (1G), the Advanced Mobile Phone System (AMPS) was introduced in the United States. AMPS was originally developed by Bell Labs. It uses analog modulation, but it boosts higher capacity through FDMA techniques, additional computing power, and frequency reuse through the cellular system.

By this time (1990s), the digital revolution has taken up so much momentum that it is rapidly adopted by the wireless communication industry as a means to support the increasing number of wireless users. IS-54 (DAMPS) which is a 2G system is the successor of the AMPS system. It uses the existing AMPS channel with increased capacity by splitting each channel further into three time slots (time division multiplexing technique) and digitally compressing the voice data, yielding three times the capacity in a single cell. This marked the start of the digital age in the business of wireless communication. In the 1980s, work had already begun in Europe on a superior 2G system known as the Global System for Mobile Communications (GSM). It was

introduced to the market in 1991 and quickly became very popular. It is now widely used in over 160 countries, and it is estimated that 82% of the global market applies to this standard today. GSM is a constant envelope modulation scheme (GMSK) and uses Time Division Multiple Access (TDMA) as air access. It features a range of digital techniques, which lead to significant improvement in connectivity and voice quality, as well as the introduction of new digital services like Short Message Service (SMS).

Another 2G system, which played an important role in future standards to come, was the CDMA One (also known as IS-95) pioneered by QUALCOMM. IS-95 uses Code Division Multiple Access (CDMA). This is a digital technique that transmits a stream of bits and allows several users to share the same frequency simultaneously. This implies that this system has a higher capacity then TDMA-based systems such as GSM.

Increasing demand for information orientated features on our handsets, such as Wireless Application Protocol (WAP) access, Multimedia Messaging Service (MMS), streaming video, and for internet email and world wide web (www) access led to development of new standards capable of supporting these applications. General Packet Radio Service (GPRS) introduced in 2000 is a Mobile Data Service compatible with GSM systems. GSM systems in combination with GPRS is often known as 2.5G, which is in fact a transition between 2G and 3G. Shortly after, Enhanced Data rates for GSM Evolution (EDGE) a 2.75G system compatible with GSM, was introduced which doubles the bandwidth capacity of GPRS systems, allowing for many high-speed data applications such as streaming video to be supported.

In 1999, the International Telecommunications Union (ITU) called for proposals for 3G protocols in an effort to ensure the smooth transition from 2G to 3G. These proposals were termed International Mobile Telecommunication 2000 (IMT-2000) and one of the adopted proposals which is becoming widely implemented is the Universal Telecommunications Service (UMTS). According to the ITU, 3G technologies should enable network operators to offer users a wider range of more advanced services (wide-area wireless voice telephony and broadband wireless data) while achieving greater network capacity through improved spectral efficiency. Wideband-CDMA (WCDMA) is an approved 3G protocol/technology used within the UMTS system which enables the network operators to achieve these stated goals. It works on the principles of Code Division Multiplexing for air access. This direct spreading technology, spreads its transmission over a 5 MHz bandwidth and can carry both voice and data simultaneously. It features a peak data rate of 384 kbps and has support for a wide array of new and emerging multimedia services. NTTDoCoMo (subsidiary of Japan's largest telephone operator NTT) launched the first WCDMA service in 2001 and now has over a million subscribers.

At the time of this thesis work, there was no formal definition of "3G and beyond", 4G requirements. But the trend of the wireless communication industry is moving towards full integration of voice, data and video streaming. To achieve this, system architects have predicted that bandwidth of more than 100Mbit/sec are required. This is a great challenge for designers of the essential components of the wireless communication system.

1.2 RF Front-end

The core element to the wireless communication system is perhaps the RF front-end. The RF front-end consist of electronic circuitry and an antenna. Its function is to convert electrical signals to electromagnetic waves, which is subsequently transmitted (transmitter) and received (receiver) through the air.

Early implementation of the transmitter consisted of an oscillator which emits continuously a train of undamped waves. By interrupting these undamped waves into long and short pulses, the information to be transmitted/received was included. In early years, the active element in the oscillator was a vacuum tube and today in commercial RF transmitters for wireless communication, the vacuum tube has been replaced by the solid state transistor.

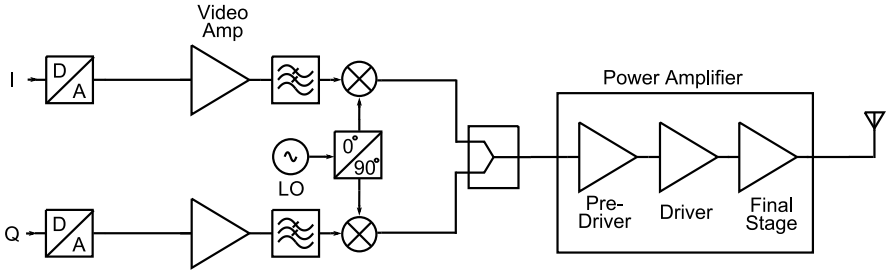


Figure 1.3: Direct Conversion Transmitter Architecture.

Modern RF transmitters can be viewed as composition of several sub-circuits (see Figure 1.3), namely power amplifier, base band circuitry, oscillators, mixers, filters and circulators. These blocks are often implemented in different technologies (different sub-vendors), and often the main goal is to optimize on cost, part count and size. These goals are considered to be the driving forces for high levels of integration.

Other important design considerations are to:

- improve power efficiency to increase talk time,
- improve interference performance,
- improving sensitivity (receiver) so as to obtain more range.

Two of the biggest challenges associated with the new generation (3G and 4G) that RF hardware designers face are: (1) the coexistence of different standards in the handset or base station, and the need to switch or adapt the hardware between the standards; (2) extremely large bandwidth requirements which are needed to support the high data rates supported in these new protocols. Furthermore, these challenges should be met without compromising too much on cost, power efficiency, part count and size. Up to date, a huge amount of work has been done for the various sub-circuit blocks. Novel circuit implementation techniques together with innovation in semiconductor technology have accomplished improved performances, for LNAs [1], mixers [2, 3], filters [4] and power amplifiers.

In this thesis, the work is focused primarily on the power amplifier block in the transmitter. In view of this, we will present novel design techniques and findings that enable the power amplifier to meet the requirements for 3G and 4G systems.

1.3 The RF Power Amplifier

The Power Amplifier is a very crucial component in the transmitter as it is the last active link between the small signal part of the components and the antenna. The role of the PA is to bring the transmitted power up to the level required in the specifications for base stations or handsets of the various communication links. Typical Power Amplifiers consist of three gain stages (Figure 1.3), namely the pre-driver, driver and the final-stage. Ideally, in an application where linear amplification is required, each of the gain stages in the PA should be able to provide amplification without any distortion and with 100% efficiency. However, with real devices this is not possible as the characteristics of the devices change with drive power and self heating. Furthermore real devices have losses and are non-linear, subjects that will be discussed in Chapter 2.

Among the three gain stages in the PA, the final-stage is always the most costly in terms of device size and current consumption. Furthermore, 3G protocols which boast increase user capacity and high data rates (e.g. W-CDMA) often employ spread spectrum techniques. This results in signals having noise-like behavior with occasional large peaks in power (high crest factor also known as high Peak-to-Average Power Ratio). In base-stations these peak power levels are often > 50 dBm. Therefore designers of the final-stage PA in base station have to oversize their PA (typically 10dB greater than the average power), this often results in low efficiency (as further explained in Chapter 2).

PA designers targeting handsets, on the other hand do not have to deal with such high RF output power. Furthermore the crest factor for the signals at the handset is lower than in the base-stations, and linearity specifications are more relaxed. However, the market now demands that the phone be able to support multiple protocols in multiple bands. Current implementations to support this, is to replicate the transceiver chain in every band. However this takes up precious real estate on the PCB. The challenge facing the PA designer for handsets is to make the PA tunable or adaptable by software control.

1.4 Research Objectives

The focus of this thesis work is on the final-stage power amplifier. The objective is to provide innovative design solutions which can be adopted by the industry in the future. The following objectives are kept in mind throughout the thesis work.

Design solution for the **handset PAs** should meet the following objectives:

- high power efficiency;
- linearity;
- compactness;
- multi-band capabilities;

where multi-band capabilities refer to using the same PA in various frequency bands so that multiple protocols can be supported within the same handset.

In contrast, design solution for the **base-station PAs** should meet the objectives:

- high RF output power;

- high efficiency despite high crest factor of the signals;
- linearity;
- low component count and reliability.

All concepts developed in this thesis havebeen demonstrated through hardware prototypes. It was our aim that these demonstrators show a better performance than current PA implementations deployed in the field.

1.5 Outline

Figure 1.4 gives a flow chart outline of the thesis and represents the path that the author has taken to reach the objectives listed in Section 1.4.

Chapter 1 gives an introduction to the wireless communication industry from its beginning to the state where it is now. Furthermore, it outlines the future of wireless communication and its accompanying challenges. This is followed by a short introduction of the RF front-end hardware, with emphasis given to the transmitter. The power amplifier (which is part of the transmitter chain) is the subject of research in this thesis work and is described in more detail, followed by stating the list of objectives in this work.

Chapter 2 provides the concepts required to understand the power amplifier. In the first part, we explain the different classes of power amplifier operation and show how the classes of operation affect its efficiency. Subsequently, the concept of linearity/distortion in the power amplifier is introduced, and the various sources of distortion are discussed. This is followed by a discussion on how linearity and efficiency are related and how designers make a balance of these two important parameters in a real power amplifier design.

In Chapter 3, based upon the understanding of the distortion sources within the power amplifier, we introduce the Derivative Superposition (DS) technique to reduce distortion in the Class-B/-AB amplifiers. Furthermore, we investigate harmonic impedance control both at the source and load side and its effect on the overall linearity on the traditional power amplifiers.

Chapter 4 looks into predistortion as a proposed system level approach in solving the linearity problems of the power amplifier without trading in any efficiency. A digital memoryless predistortion algorithm is developed within the scope of this work, which is suitable for linearizing classical power amplifiers. This algorithm is later adapted and used in more advanced amplifier concepts as described in Chapter 5. Time dependent non-linear effects (also known as memory effects) are also addressed, and a compensation technique is introduced. The work up to this point has been focused on improving the linearity of the power amplifier. At this juncture we make a switch to investigate techniques which will improve the average efficiency of the power amplifier, which is highly desirable for complex modulated signals. In Chapter 5, we look into the general concept behind high efficiency power amplifier architectures and review the major important high efficient architectures in use today.

Chapter 6 introduces a novel approach for handset power amplifiers to simultaneously allow for multi-band tuning and high efficiency performance based on an adaptive matching network. And lastly in Chapter 7, the concept of the N-way Do-

herty Power Amplifier is introduced, and it is experimentally shown why this is a good candidate for 3G base station power amplifiers.

Lastly, this thesis work is concluded in Chapter 8 with recommendations for future research directions.

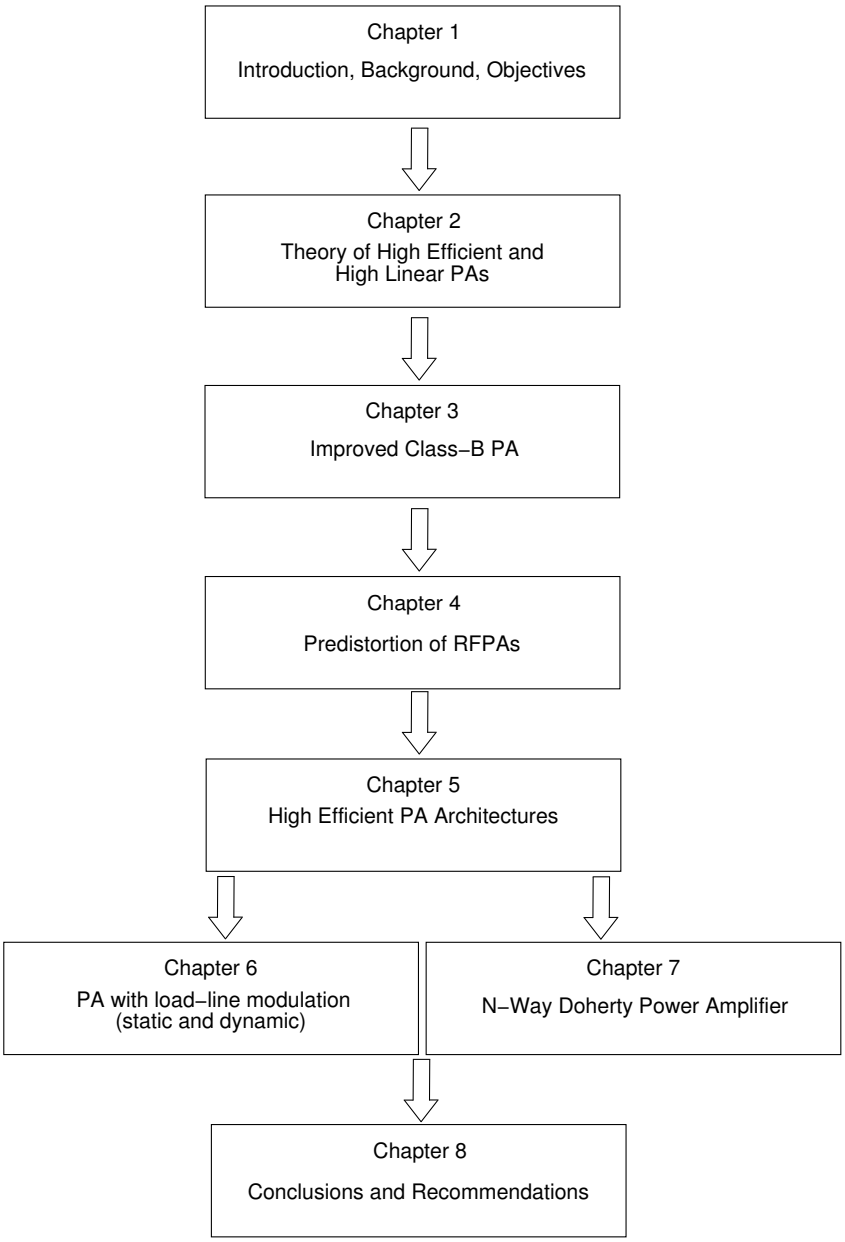


Figure 1.4: Organization of the thesis

Chapter 2

Theory of Highly Efficient and Linear Power Amplifiers

This chapter covers the theory on RF power amplifiers (RFPA), which is required to understand the conditions when the amplifier is operating linearly and efficiently. This serves as a foundation for the content of the later chapters. Although this has already been presented in many ways in literature [5–7], it is still useful to review the theory in light of the interpretation given in this thesis work.

Section 2.1 discusses the well known Class-A, Class-AB, Class-B and Class-C power amplifier, explaining its linearity and efficiency performance in the ideal case. In Section 2.2, non idealities associated with real transistors and their effect on linearity, bandwidth and efficiency are discussed. Lastly, Section 2.3 discusses the new generation wireless protocols (3G,4G) and the challenges it presents to designers of RF power amplifiers.

2.1 Ideal Class-A to C Power Amplifier

To explain the traditional classes of the RFPA (A, AB and C), the following "ideal" device is used in a circuit setup as shown in Figure 2.1. In this thesis work, the definition of an ideal transistor refers to a voltage controlled source with DC characteristics such that the device turns on at 0 volts and ramps up linearly till the maximum current. For the purpose of verifying the concepts in this thesis work using a simulator, a simple transistor model based on a Symbolic Defined Device (SDD in Agilent's ADS) will be used in the schematic of Figure 2.1, with maximum output power of the amplifier normalized to 1-Watt. This simple model is able to take into account the effect of saturation, as shown in equation (2.1),

$$I_d = \begin{cases} 0 & \text{if } V_g \le 0, \\ (k \cdot V_g) - \underbrace{10^{-12} \cdot \exp\left(\frac{-V_d}{10^{-3}}\right)}_{\text{models voltage saturation}} & \text{if } V_g > 0. \end{cases} \tag{2.1}$$

where I_d is the current at the output of the transistor, V_g is the voltage across the gate of the input of the transistor, k is a constant multiple, and V_d is the voltage across the output of the transistor.

In Figure 2.1, the DC characteristic plot shows what was just described but for the normalized case.

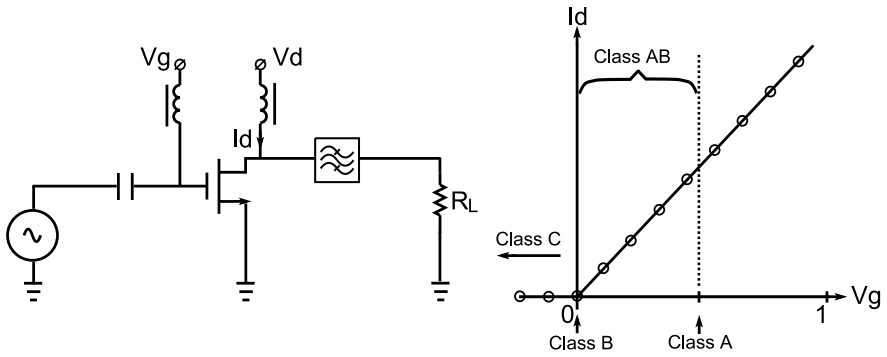


Figure 2.1: Generic RF power amplifier with ideal transistor

The classes of operation are determined by the gate bias V_g. For Class-A, the gate bias is chosen such, that even under maximum sinusoidal excursion around the bias at the input, the output current would remain sinusoidal (a linear transfer from the voltage to current). This would imply that for Class-A operation the designer has to set the bias at half the maximum DC current in order that the device can deliver its maximum RF power under maximum drive. Moving down from the Class-A bias point towards zero, there are a whole range of bias values which corresponds to the Class-AB mode of operation. For the special case where the bias V_g is zero, the amplifier is said to be operating in Class-B. If the bias is set further down into the negative region, the operation becomes Class-C.

Figure 2.2 shows the various input voltage and corresponding current waveform at the drain of the device, for the different classes of operation, generated using our simple model as described above in ADS. An important thing to note from this figure is the concept of the conduction angle α. α is related to the time in a RF cycle where the device is actually conducting current. As the bias is moved from Class-A to Class-B, the conduction angle moves from a maximum of 180 degrees (always conducting) towards 90 degrees (only half the time), and this has an effect on the efficiency of the power amplifier as illustrated in the following equations.

$$I_d = \begin{cases} I_q + A_{out} \cdot \cos(\theta) & \text{for } \theta \leq \alpha \text{ and } \theta \geq 2\pi - \alpha, \\ 0 & \text{for } \theta \text{ otherwise.} \end{cases} \tag{2.2}$$

where

$$\cos\alpha = -\frac{I_q}{A_{out}}$$
$$A_{out} = kA_{in}$$

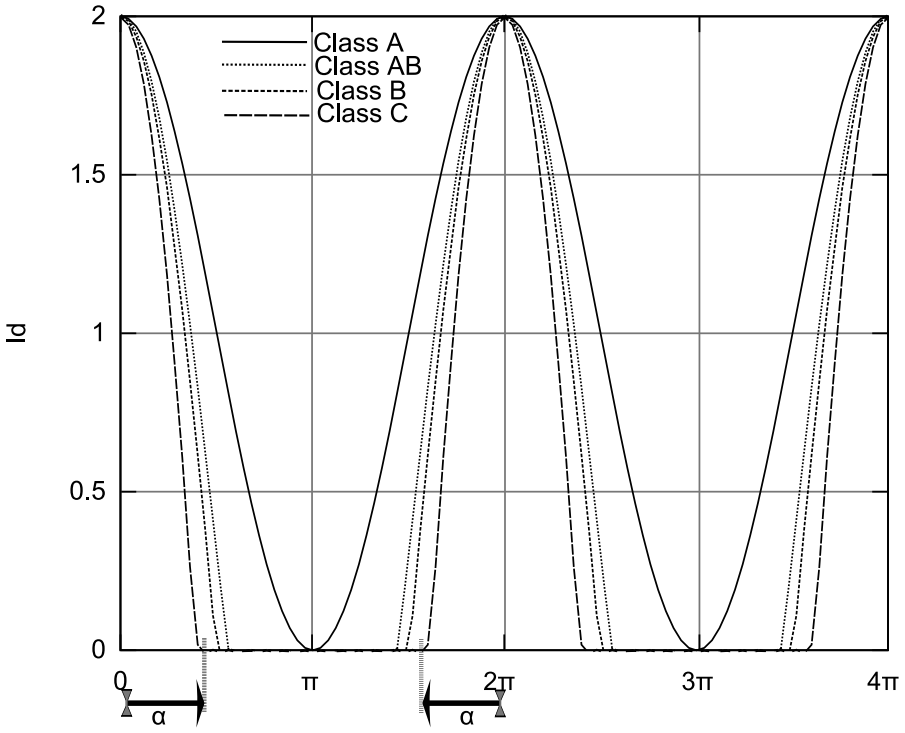


Figure 2.2: Output current waveforms for the various Classes

Note, that θ corresponds to one period of the RF sinusoid input waveform. I_q refers to the quiescent current. A_{out} is the magnitude of the sinusoidal peak current corresponding to a sinusoidal peak voltage excitation of magnitude A_{in}, under a linear gain of k.

Applying Fourier Series analysis to equation 2.2, the following expression for the current at the fundamental and DC are obtained as follows [5]:

$$\begin{aligned} i_{rf} &= \frac{1}{\pi}\int_0^{2\pi} i_{DC}\cos(\theta)\,\mathrm{d}\theta \\ &= \frac{2}{\pi}\left(I_q\sin\alpha + \frac{A_{out}}{4}\sin 2\alpha + \frac{A_{out}\,\alpha}{2}\right) \end{aligned} \tag{2.3}$$

$$\begin{aligned} I_{DC} &= \frac{1}{2\pi}\int_0^{2\pi} i_{DC}\,\mathrm{d}\theta \\ &= \frac{1}{\pi}\left(I_q\,\alpha + A_{out}\,\sin\alpha\right) \end{aligned} \tag{2.4}$$

The maximum efficiency achieved for any given class of operation occurs when

the voltage across the device is in voltage saturation. This implies that the maximum efficiency is only determined by the ratio of the currents.

$$\eta = \frac{i_{rf}}{I_{DC}} \tag{2.5}$$

$$= \frac{\sin 2\alpha - 2\alpha}{4\left(\alpha \cos\alpha - \sin\alpha\right)} \tag{2.6}$$

Substituting $\alpha = 180°$ (which corresponds to the Class-A operation) into equation 2.5, results in $\eta = 50\%$ and $\alpha = 90°$ (Class-B) results in $\eta = 78.5\%$. As the conduction angle decreases beyond 90° (corresponding to Class-C operation), the theoretical efficiency approaches 100% as shown in equation 2.7.

$$\eta = \lim_{\alpha \to 0°} \frac{\sin 2\alpha - 2\alpha}{4\left(\alpha \cos\alpha - \sin\alpha\right)} \tag{2.7}$$

$$= 100\%$$

Figure 2.3 gives a summary on how the maximum efficiency varies with conduction angles α.

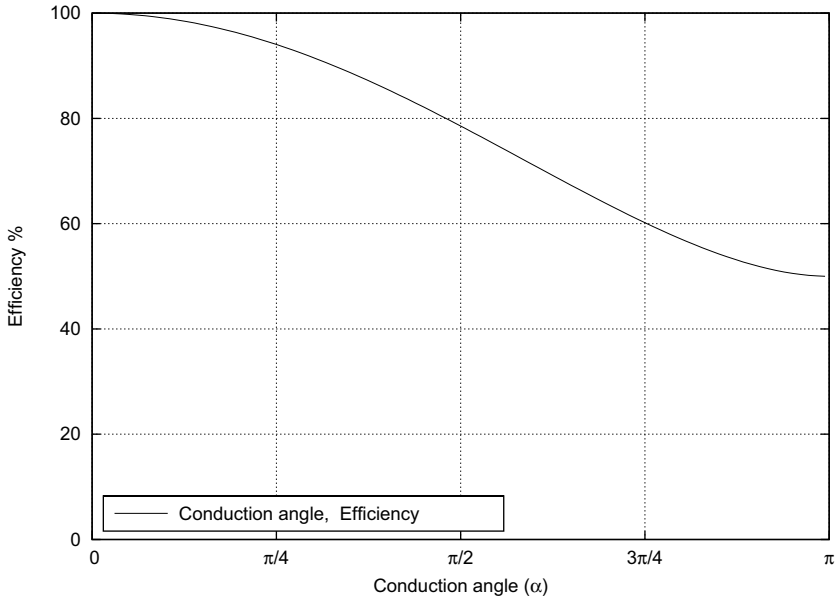
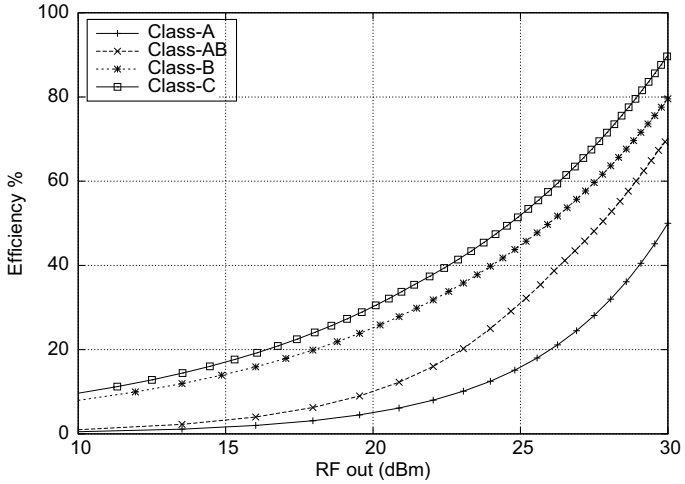


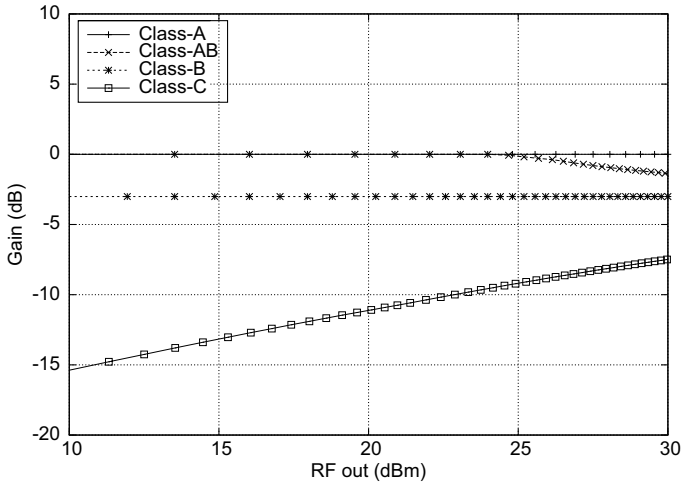
Figure 2.3: Efficiency versus conduction angle

Although moving from Class-A to Class-C mode of operation results in an increase in efficiency, the negative impact is that more input drive is required in Class-C to reach the same maximum power in Class-A. Figure 2.4 shows how both efficiency/gain versus output power changes as the mode of operation is changed from Class-A to

Class-C. Note that the gain here refers to the voltage gain, as the power gain of an ideal FET device is infinite. Real devices have losses at the input and this makes the definition of the power gain valid. The trend shown here with voltage gain will be the same as the power gain, when real devices are used.



(a) Efficiency versus output power.



(b) Voltage gain versus output power. Gain is normalized to the Class A case.

Figure 2.4: Simulated Voltage gain and Efficiency versus output power of a 1-Watt Power Amplifier under the different operating classes.

Besides Class-A and Class-B operation, none of the other classes exhibit linear gain. Although the gain in Class-B is constant, its value is theoretically 3dB lower

than what is achieved in Class-A. In Class-AB, the gain starts off constant and similar to Class-A, however at a certain input drive, the output RF current starts to clip and the conduction angle α starts being a function of the input drive. In Class-C, the conduction angle α is always a function of the input drive, resulting in a gain that changes as a function of power.

2.2 Distortion, Bandwidth and Efficiency issues in RF Power Amplifiers

Real transistors unlike the ideal model used in Section 2.1 are unfortunately far more complex. Besides the basic voltage-controlled current source function, these real transistors have parasitics that leads to reduction in bandwidth, in-band distortion, and deviation from the ideal efficiency computed in the previous section.

2.2.1 Bandwidth

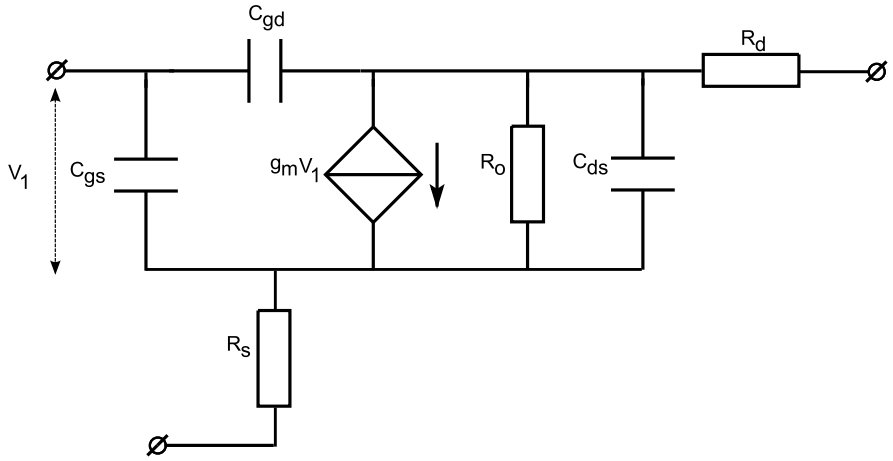


Figure 2.5: Typical equivalent FET model.

Figure 2.5 shows a typical model use in analyzing FET devices [8]. At the output we typically have a resistance R_o and capacitance C_{ds} in parallel with the voltage-controlled current source. The resistances R_d and R_s are metal resistances connecting the transistor to the outside world. To get maximum power out of the transistor, an impedance needs to be offered that causes voltage saturation in the transistor and also resonate out the imaginary components (caused mainly by C_{ds}).

The real part of this impedance can be estimated based on the power rating of the transistor,

$$R_{match} = \frac{V_d^2}{2 \cdot Pout_{max}} \quad , \tag{2.8}$$

which in larger transistors can result in small values of R_{match}. Note also that the parasitic capacitances also scale with size of the transistor. In most RF power amplifier applications for wireless communications, the system impedances are fixed at 50Ω, this implies that a lossless impedance transformation is required (same applies for the input). The need for impedance transformation and the presence of these frequency dependence parasitics in the transistor causes bandwidth limitation in any power amplifier design. Well known design techniques with regards to wide-band matching can be found in [7], but this often results in a trade-off with the physical size of the power-amplifier hardware.

2.2.2 Efficiency

The deviation from the ideal efficiency for the various classes computed in Section 2.1 is attributed to two loss mechanism for FET transistors. The first is due to, what is commonly referred to as on-resistance. This resistance is frequency independent and its impact on the efficiency can be estimated by the following equation (using Class-B as an example):

$$\eta = 78.5\% \frac{1}{1 + 2 \cdot \frac{R_o}{R_L}} \quad , \tag{2.9}$$

where R_L refers to the impedance which the device sees at its output.

The second loss mechanism is frequency dependent and relates to the losses in the output capacitance (C_{ds}) where the resistive part is a combination of drain resistance and substrate resistance. The resulting degradation in efficiency due to this effect is estimated as follow:

$$\eta = 78.5\% \frac{1}{1 + \omega^2 \cdot C_{ds}^2 \cdot R_p \cdot R_L} \quad , \tag{2.10}$$

where R_p represents the losses in that output capacitance C_{ds}.

It is the goal of device technologist to minimize these losses [9], bringing the efficiency of these transistor closer to its ideal performance.

2.2.3 Distortion

The voltage-controlled current source in real devices follows a so called “square-law” relationship (in reality the order is much higher than two) instead of the piecewise linear we have assumed for the ideal device (Figure 2.1). Furthermore, the parasitic capacitances in the transistor are also functions of drive power. These non-linear component in the transistors add up in a vectorial fashion, and this interaction can be analyzed with the Volterra Series analysis.

This technique allows one to derive close form expressions showing how the non-linear components interact with each other to generate the final distortion products [10]. However, the disadvantage of this technique is that, it is only valid in the range where the coefficients are extracted. Furthermore, to use the analysis at the point where the power amplifier is driven close to its compression, a large number of terms are needed to be extracted and included in the analysis, which makes the

resulting equations cumbersome. To make these close form expressions interpretable, assumptions are usually required to simplify these expressions. As a result, this analysis only gives an approximate estimate of the distortion terms, which at best is within 3dB of the actual values.

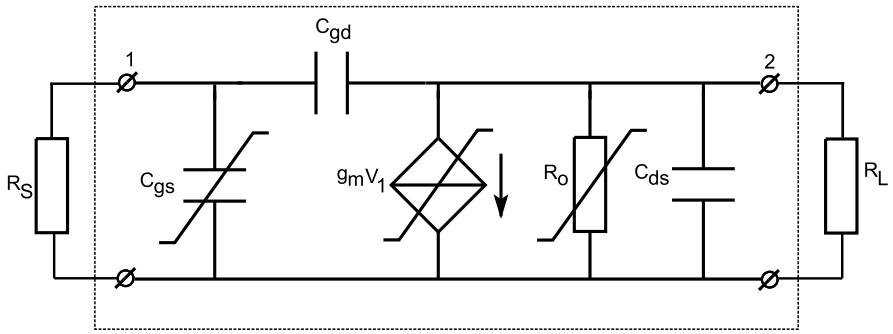


Figure 2.6: Typical equivalent FET model used in Volterra Series analysis.

Figure 2.6 shows a typical non-linear model used in analyzing FET devices. The third order distortion is directly related to equation 2.11. This is derived using the Volterra series technique explained in [10]. Note that the sources of non-linearity are here restricted to the transconductance g_m, input capacitance C_{gs} and the output resistance R_o,

$$H_{3,2}(s_1, s_2, s_3) = \frac{i_{NL3,g_o}\alpha + i_{NL3,g_m}\alpha - i_{NL3,C_{gs}}\gamma}{\delta\alpha - \gamma\beta} \quad , \tag{2.11}$$

where

$$\begin{aligned}
\alpha &= \frac{1}{R_S} + (s_1 + s_2 + s_3)(C_{gs} + C_{gd}) \\
\beta &= -(s_1 + s_2 + s_3)\, C_{gd} \\
\gamma &= (s_1 + s_2 + s_3)\, C_{gd} - g_m \\
\delta &= -[(s_1 + s_2 + s_3)\, C_{gd} + \left(\frac{1}{R_o} + \frac{1}{R_L}\right)]
\end{aligned}$$

Here s_1, s_2 and s_3 refer to the frequency components in a third order system. To obtain the expression for the IM3 products, typically we set $s_1 = j\omega_1$, $s_2 = j\omega_2$ and $s_3 = j\omega_1$. Note that i_{NL3,g_m} and $i_{NL3,C_{gs}}$ are the third order non-linear current generators, whose values are related to the Taylor coefficients of the non-linear current and charge functions, as well as the composed third order transfer functions.

$$
\begin{aligned}
i_{NL3,g_m} = {} & K_{3,g_m} H_{1,1}(s_1) H_{1,1}(s_2) H_{1,1}(s_3) \\
& + \frac{2}{3} K_{2,g_m} [H_{1,1}(s_1) H_{2,1}(s_2, s_3) \\
& \qquad + H_{1,1}(s_2) H_{2,1}(s_1, s_3) \\
& \qquad + H_{1,1}(s_3) H_{2,1}(s_1, s_2)]
\end{aligned} \tag{2.12}
$$

$$
\begin{aligned}
i_{NL3,g_o} = {} & K_{3,g_o} H_{1,2}(s_1) H_{1,2}(s_2) H_{1,2}(s_3) \\
& + \frac{2}{3} K_{2,g_o} [H_{1,2}(s_1) H_{2,2}(s_2, s_3) \\
& \qquad + H_{1,2}(s_2) H_{2,2}(s_1, s_3) \\
& \qquad + H_{1,2}(s_3) H_{2,2}(s_1, s_2)]
\end{aligned} \tag{2.13}
$$

$$
\begin{aligned}
i_{NL3,C_{gs}} = {} & (s_1 + s_2 + s_3)\, K_{3,C_{gs}} H_{1,1}(s_1) H_{1,1}(s_2) H_{1,1}(s_3) \\
& + \frac{2}{3} K_{2,C_{gs}} (s_1 + s_2 + s_3) [H_{1,1}(s_1) H_{2,1}(s_2, s_3) \\
& \qquad + H_{1,1}(s_2) H_{2,1}(s_1, s_3) \\
& \qquad + H_{1,1}(s_3) H_{2,1}(s_1, s_2)]
\end{aligned} \tag{2.14}
$$

In this notation, the first subscript of H refers to the order, and the second subscript refers to the node (see Figure 2.6). Thus, $H_{2,1}$ and $H_{2,2}$ are second-order transfer functions of the voltage at the input of the circuit to the nodes 1 and 2 respectively. Likewise $H_{1,1}$ and $H_{1,2}$ are the voltage transfer functions that only depend on the linear network response. Equation 2.11 shows that the third order intermodulation (which falls in band), is a result of a complex interaction of third-order non-linear currents from the various distortion components, as well as indirect third-order terms that arise from the interaction of fundamental components with second order frequency products. This later contribution makes the circuit conditions that control the base band and second harmonic voltages very important to the overall linearity performance of the design.

2.3 Wireless Communication Standards and the RF Power Amplifier

Modern wireless standards (2G and 3G) employ digital modulation and multiple air access technique to increase the capacity of the system (increase number of users). Understanding the wireless communication standard being employed is an important task for the power amplifier designer as it often states what the transmitted signal will "looks like" (i.e. in terms of CF and bandwidth). Furthermore, it states the specification on the accepted Signal-Noise-Ratio (SNR), the Bit-Error-Rate (BER) and

extraneous emissions tolerated. These criteria should all be met and where possible with the highest efficiency possible.

GSM the most common wireless standard for 2G uses GMSK (Gaussian Minimum Shift Keying) modulation and TDMA/FDD (Frequency Division Duplexing). GMSK is a continuous-phase frequency-shift keying modulation scheme (i.e. constant envelope) similar to MSK but with the digital data stream first being preprocessed by pulling it through a Gaussian filter. This results in a constant envelope RF signal which allow power amplifier designers for GSM application to dimension the transistor such that it always is in saturation, resulting in high efficiency. Furthermore, it also allows for very relaxed linearity specifications as amplitude distortion is of no concern. Note also that in GSM systems, the power amplifier has to amplify signals in burst (TDMA), thus the thermal settling time of the transistor becomes important. Commonly used technique to get around this is to use a Class-A or AB biasing where the quiescent current is significantly high enough to mitigate the fluctuation in temperature due to the RF power.

For 3G systems, WCDMA is one of the main standards. WCDMA employs QPSK as modulation and CDMA as air access technique. QPSK in itself is a constant envelope signal, but when used in a CDMA environment, where the information is spread over the whole spectrum and all channels are just transmitted on the same frequency, the signal becomes noise like and have occasionally high peaks. The crest factor of a WCDMA signal is approximately 10-15 dB depending on the number of channels. This would imply that the devices used in the power amplifier has to be "over dimensioned" to handle the occasional peak powers, while the amplifier remains most of the time at its average power, 10dB lower than the maximum dimensioned power.

Figure 2.7 gives the plot of the Probability Distribution Function (PDF) of a WCDMA signal versus normalized power levels and the efficiency performance of an ideal transistor in Class-B operation. Note that for most of the time, the signal is being amplified at a low efficiency and only occasionally at the peaks the amplifier is operated at its maximum efficiency. This implies that on average, the power amplifier is operated at a much lower efficiency than its peak efficiency capability. Given the CW efficiency characteristic of the amplifier and the PDF of the signal, the average efficiency can be computed using Equation 2.15, and is found for WCDMA to be 22% for an ideal device when operated in Class-B.

$$\langle \eta \rangle = \frac{\langle p_{out} \rangle}{\langle P_{DC} \rangle} = \frac{\int p_{out} \cdot p(p_{out})\, \mathrm{d}p_{out}}{\int \dfrac{p_{out} \cdot p(p_{out})}{\eta(p_{out})}\, \mathrm{d}p_{out}} \tag{2.15}$$

With a real transistor, it could be even required to operate the transistor further at back-off due to stringent linearity requirements. This together with the losses in the device described in Section 2.2.2 will cause the efficiency value to be lower than the ideal case, and in practise this will range from 5 to 10%.

This low average efficiency for signals with high CF even for an ideal device, is the motivation at this time of writing for research within the RF power amplifier community, into high efficient power amplifier classes [11], high efficient power amplifier architectures [12, 13] and crest factor reduction techniques [14].

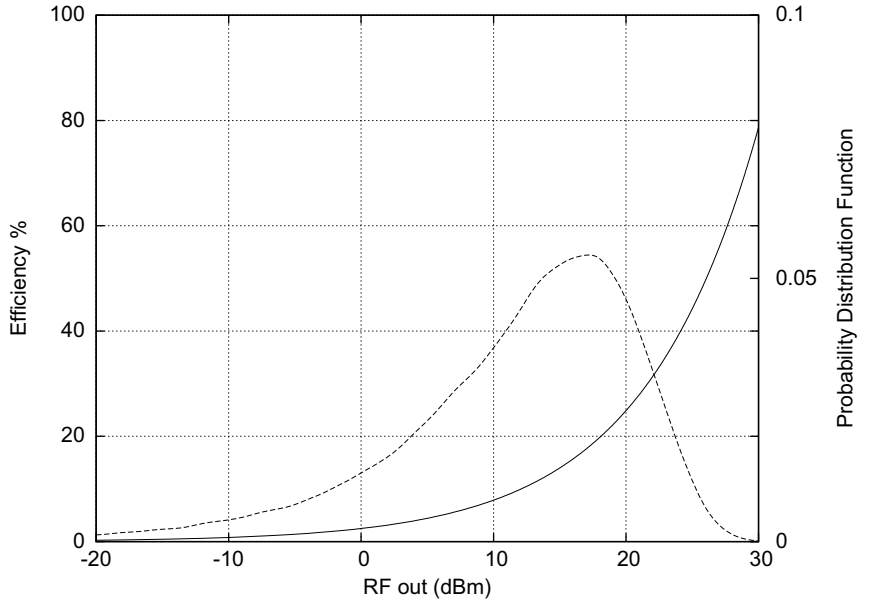


Figure 2.7: Efficiency versus power back-off and PDF of a WCDMA signal

In this thesis work, we will focus on improving the well known Class-B/-AB amplifier, and look into high efficient power amplifier architectures which are suitable for signals with high CF.

Chapter 3

Improved Class-B, Class-AB Power Amplifiers

This chapter gives an overview on the progression in class-AB oriented power amplifier stages over the last few years. This progression has yielded a dramatic improvement in linearity & efficiency, which can be contributed to technology, as well as, to improved design techniques. Following a historical perspective, we will first discuss in Section 3.1 how the transfer function of the active device itself can be made linear using the Derivative Superposition technique (DS) [15]. This technique, which uses multiple devices in parallel with small bias adjustments to shape the overall transconductance of the composed device, proves to be very effective in the far output power back-off condition. In order to improve the device linearity closer to compression, as next step the always present secondary mixing phenomena over the input non-linearities of the active device input are utilized to compensate direct third-order intermodulation distortion products. This so called "input out-of-band" matching strategy was originally developed for bipolar devices [16] but proves to be also effective for FET devices at moderate power back-off conditions. To improve linearity both in the far as well medium range output power back-off conditions, in this work, a combination of derivative superposition with input-out-of-band matching was applied to LDMOS base station technology. Finally, to improve linearity close to compression, simultaneous input and output out-of-band matching has been utilized in Section 3.3 [17]. This technique yields rather dramatic improvements in the classical linearity – efficiency trade-off at the expense of some decreased linearity in the far power back-off conditions. We conclude with a consideration on the progress made in LDMOS technology within the constrains of (modified) "Class-B" operation.

3.1 Derivative Superposition

The technique of derivative Superposition (DS) aims to improve amplifier linearity by shaping the overall transconductance of an amplifier stage. This is implemented by using several FET devices with appropriate gate widths in parallel, while biasing

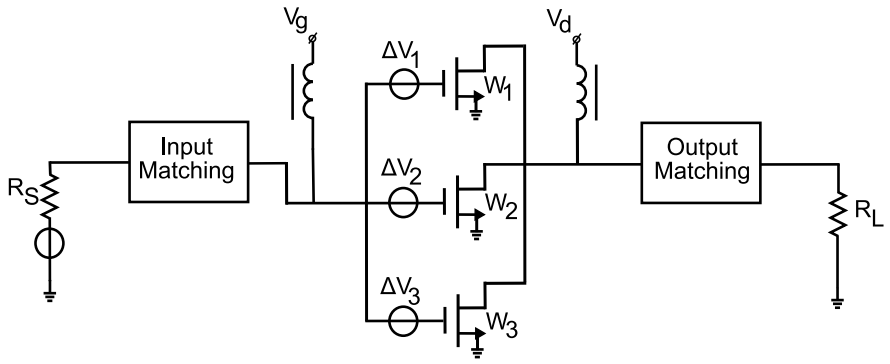


Figure 3.1: Generic setup for Derivative Superposition using FET devices in Class-AB operation.

each of them with a specific gate-offset voltage (Figure 3.1). Using this technique, it is theoretically possible to shape the transconductance (g_m) such that the higher odd-order derivatives g_{m3} and g_{m5} are significantly reduced, or even approach zero. This technique was first demonstrated by Webster et al. [15] showing 8dB reduction of third-order intermodulation distortion 10dB from the P1dB point using HEMT transistors from NEC (NE33284). Subsequently in [16], the DS approach has been used on the Philips (NXP) GEN 2 LDMOS technology to obtain better linearity performance in back-off power levels. Although very convincing results have been achieved using this method (see Figure 3.2), very extensive prototyping was used in most of all the above referenced works. To better understand the interaction of the non-linear sources of distortion in the transistor when DS is applied, we have developed in the following section, analytical equations for this linearization method using Volterra Series techniques.

3.1.1 The Volterra Series approach to Derivative Superposition

Using Volterra series calculation techniques, one can obtain analytical expressions for the intermodulation terms, based on the nonlinear description of the active device. Although labor intensive, this method provides clear insight in the dominant device contributions to intermodulation distorion and indicates how DS can be utilized to reduce this intermodulation.

Figure 3.3 gives the large signal transistor model that was employed for analysis in this chapter. The model shown is a simplified model of a LDMOS transistor. Although a more complex model can be used, this would yield unmanageable equations for the resulting intermodulation terms. For this reason, in our assumptions the gate-source capacitance C_{gs} and transconductance g_m are modeled here as one-dimensional nonlinearities (i.e. they only depend on the gate-voltage V_{gs}). This simplification has be verified through careful inspection that for LDMOS devices, the dominant nonlinearities are mainly represented by C_{gs} and g_m [18].

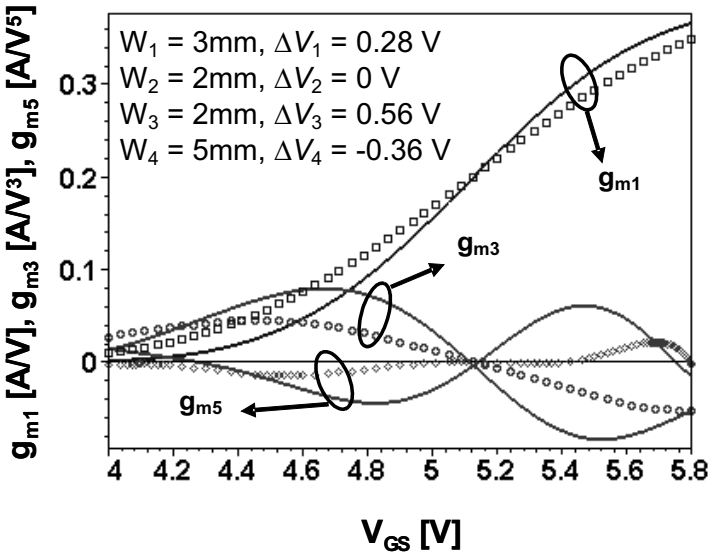


Figure 3.2: Measured odd-order Taylor coefficients g_{m1}, g_{m3} and g_{m5} of a 12mm gate width GEN 2 Philips LDMOS device (solid lines) and experimental DS optimized odd-order Taylor coefficients using four GEN 2 devices with the same total gate width (dotted lines) (V_{DS} = 26V). One can observe that the g_{m3} and g_{m5} are significantly reduced arround the IM3 sweet spot (g_{m3} = g_{m5} = 0) **Courtesy of Dr. Mark van der Heijden**

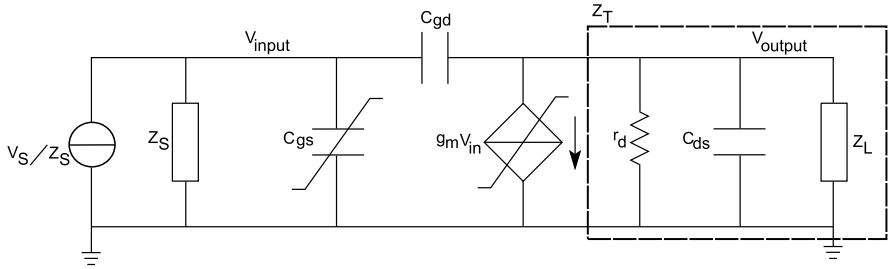


Figure 3.3: Schematic used for the Voltera-Series evaluation of the Derivative Superposition method.

In order to predict the IM3 of the device over a considerable input-power range, a fifth-order Volterra-series expression has been developed for a two-tone drive signal. In our calculations, conventional class-AB operation was assumed, using short circuit conditions for the base-band as well the 2nd harmonic frequencies. Note that these short circuit conditions prevents the additional generation of IM3 components through secondary mixing over the out-of-band impedances (as seen in Chapter 2). In addition, it strongly simplifies the resulting equations, something that is highly desirable when dealing with Volterra series. Under these conditions, the IM3 output voltage at the output of our LDMOS device in class-B operation can be formulated as,

$$v_{out(\omega_{IM3})} = \frac{\left(g_m - j\omega_{IM3}C_{gd}\right) i_{C_{gs}} - \left(Y_S + j\omega_{IM3}\left(C_{gd}+C_{gs}\right)\right) i_{gm}}{j\omega_{IM3}C_{gd}\left(g_m + Y_S + Y_L + j\omega_{IM3}C_{gs}\right) + Y_L\left(Y_S + j\omega_{IM3}C_{gs}\right)}, \tag{3.1}$$

where

$$i_{C_{gs}} = j\omega_{IM3}\{\frac{3}{4}C_{gs3}v^2_{in(\omega_2)}v_{in(-\omega_1)} + \frac{25}{8}C_{gs5}v^3_{in(\omega_2)}v_{in(-\omega_2)}v_{in(-\omega_1)}\}, \tag{3.2}$$

$$i_{g_m} = \frac{3}{4}g_{m3}v^2_{in(\omega_2)}v_{in(-\omega_1)} + \frac{25}{8}g_{m5}v^3_{in(\omega_2)}v_{in(-\omega_2)}v_{in(-\omega_1)}, \tag{3.3}$$

$$v_{in(\omega)} = \frac{Y_S v_S\left(j\omega\, C_{gd} + Y_L\right)}{j\omega\, C_{gd}\left(g_m + Y_S + Y_L + j\omega\, C_{gs}\right) + Y_L\left(Y_S + j\omega\, C_{gs}\right)}\ . \tag{3.4}$$

Note, that g_{m3} and g_{m5} are the third and fifth-order Taylor coefficients of I_{ds} versus V_{gs} respectively. Similarly, C_{gs3} and C_{gs5} are third and fifth-order Taylor coefficients of C_{gs} versus V_{gs}, respectively. While, Y_S and Y_L are the source and load admittances at the fundamental frequency.

The above expressions can be further simplified by neglecting the feedback capacitance C_{gd}. This proves to be a valid assumption for modern LDMOS devices due to the presence of the grounded shield between the gate and the drain [19], which strongly reduces the feedback capacitance. Furthermore, the two-tone difference frequency (baseband or IF frequency) can be considered to be small compared to the design frequency (RF frequency). Consequently, by assuming ($\omega = \omega_1 = \omega_2$), the following simplified IM3 expression can be obtained:

$$\begin{aligned} v_{out(\omega_{IM3})} = \frac{-v_{in_\omega}|v_{in_\omega}|^2}{Y_L\left(Y_S + j\omega\, C_{gs}\right)}\{ &\left(Y_S + j\omega\, C_{gs}\right)\left(\frac{3}{4}g_{m3} + \frac{25}{8}g_{m5}|v_{in_\omega}|^2\right) \\ &- j\omega\, g_m\left(\frac{3}{4}C_{gs3} + \frac{25}{8}C_{gs5}|v_{in_\omega}|^2\right)\} \end{aligned} \tag{3.5}$$

From this expression, an important conclusion can be drawn on the use of DS on LDMOS devices. Namely, the use of DS to shape the transconductance such that its derivatives (g_{m3} and g_{m5}) approaches zero, is insufficient when C_{gs} is non-linear in the area of operation. In addition, at higher power levels, other components such as C_{ds} becomes non-linear and will start contributing to the distortion. These facts limit the

use of classical DS techniques for achieving improved linearity close to compression of the active device.

To improve the linearity also at higher power levels, we will now consider the use of baseband and harmonic source tuning to influence the IMD behavior of a power amplifier. This was first demonstrated in the work of De Carvalho [20]. In these works the short circuited conditions for baseband and second harmonics are omitted, yielding secondary mixing of the double and base-band frequency components with fundamental frequency. This results in the generation of additional (secondary) IM3 components with different phases and magnitudes. Since these IM3 components can have opposite signs compared to the (direct) IM3 components which result from third-order nonlinearities, cancellation phenomena can occur, which can be used to our advantage.

These phenomena have been extensively explored in the past for bipolar devices [16]. However, FET devices show a more complicated behavior for their transconductance and input capacitance, leaving still room for research at this point. To fill this gap and to improve the linearity of LDMOS FET devices closer to their 1dB compression point, we now extend our Volterra analysis of Section 3.1.1 by allowing a non-short circuit condition for the second harmonic at the source side, leaving ($Y_{S_{2\omega}}$) as a variable. Besides using the out-of-band termination technique, derivative superposition is applied to improve linearity in the far back-off region. This is important to communication signals with very high peak-to-average ratios.

In the following sub-sections, Volterra series analysis is used to evaluate the input out-of-band matching conditions for improved linearity and efficiency. We will verify these results experimentally in this chapter.

3.1.2 Finding the optimum 2nd harmonic source impedance

To keep our equations manageable we ignore again the feedback capacitance and considering here only terms that contribute at higher powers (i.e fifth-order terms), the resulting IM3 voltage at the output can be written as,

$$v_{out(\omega_{IM3})} \cong -Z_T\{ \overbrace{g_m v_{in(\omega_{IM3})}}^{\text{direct IM3 from input}} + \overbrace{\frac{3}{8} g_{m3} v_{in(-2\omega_1)} v_{in(2\omega_2)} v_{in(\omega_1)}}^{\text{direct IM3}} + \underbrace{\frac{25}{8} g_{m5} v_{in(\omega_2)}^2 v_{in(-\omega_1)} |v_{in(\omega_2)}|^2\}}_{\text{fifth order originated IM3}} \tag{3.6}$$

where

$$v_{in(\omega_{IM3})} = -\frac{j\,\omega_{IM3}}{Y_{S(\omega_{IM3})} + j\,\omega_{IM3}C_{gs}}\{\frac{25}{8}C_{gs5}v_{in(\omega_2)}^3 v_{in(-\omega_2)}v_{in(-\omega_1)}\} \tag{3.7}$$

$$v_{in(2\omega)} = \frac{j\,\omega C_{gs2}v_{in(\omega)}^2}{Y_{S(2\omega)} + j\,\omega C_{gs}} \tag{3.8}$$

$$Z_T = \frac{1}{Z_{L(\omega_{IM3})}} + \frac{1}{r_d} + j\,\omega C_{ds} \tag{3.9}$$

Of which Z_T is the impedance seen at the IM3 frequency by the Drain current source (approximately equal to the impedance at the fundamental frequencies) (Figure 3.3). Equation (3.6) and (3.8) show that by adjusting the second harmonic source-termination $Y_{S_{2\omega}}$, the phase of the secondary-mixing products at the IM3 frequency can be controlled to cancel the direct IM3 term generated by third-order non-linearities.

3.2 Measurement verification

In this measurement verification we use for the non-optimized LDMOS device a Philips (NXP) GEN 2 device with the gate dimensions $3\times0.8\times100\mu\,m$. the derivative superposition optimization is done in the same LDMOS technology "on wafer" as described in [19]. The on-wafer linearity measurements were carried out using an in house developed active load/source pull setup [21] that provides independent control of the fundamental, baseband and second-harmonic impedances at the source and the load side of the DUT. This setup allows a straight forward evaluation of the influence of $Z_{S(2\omega)}$ on the overall linearity of the PA stage.

Figure 3.4 gives the results of a two-tone IM3 measurement. Both the lower and upper IM3 exhibits a similar trend and are within 0.5dB of each other. All out-of-band impedances were terminated with a short circuit except when explicitly stated in the figure. The transducer power gains of all 4 tests were similar (22dB) and the efficiency performances were also comparable. Using only derivative superposition, an improvement in IM3 can only be achieved up to 10dB back-off from compression. If only out-of-band optimization is used with a focus to achieve linearity at high powers, linearity at lower power levels is degraded. However, using the combination of both DS and $Z_{S(2\omega)}$ optimization, the advantages of both techniques line up, providing a linearity improvement from lower power levels up to 4dB back-off from the 1 dB compression point.

To further verify the effectiveness of this proposed technique, the PAs were driven at various power levels closer to compression with a realistic test signal (CDMA IS-95). A transducer power gain of 22dB was achieved with all 4 setups and similar efficiency performance. The ACPR was measured at 750kHz offset at both the adjacent channels (Note that the ACPR specs for base-stations is 45 dBc). Figure 3.5 gives the average ACPR of the 2 adjacent channels versus output channel power. Note that for a modulated signal as used in our experiment, the output channel power represents the average power level, but the instantaneous power varies over a large power range, as

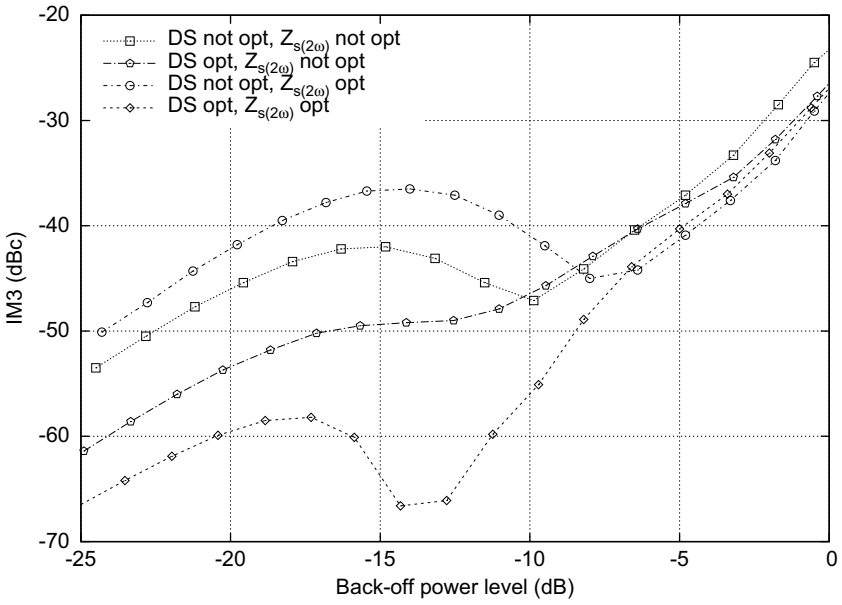


Figure 3.4: Measured IM3 versus Output power back-off from P1dB. Where P1dB=33dBm, f_o = 1.96GHz and Δf = 100KHz

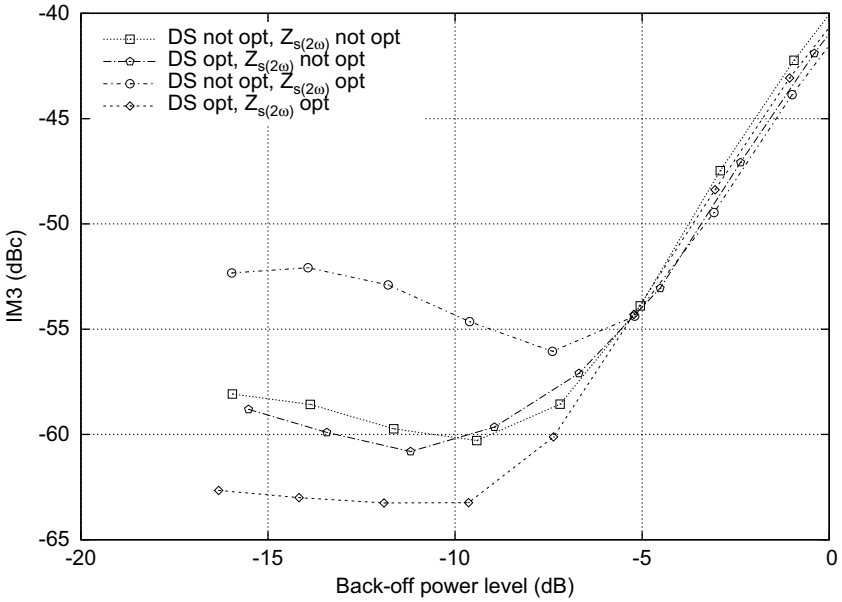


Figure 3.5: Measured ACPR versus Output power back-off from P1dB. Signal used is Forward Channel IS95. Peak Power=33.5dBm. Crest Factor=8.5dB

explained in Chapter 2. This is the reason why the two-tone intermodulation data versus power does not follow exactly the characteristic of ACPR versus power.

Using only DS only (⬠ plot, Figure 3.5), the linearity improvement is marginal over the conventional LDMOS operation (□ plot) and only effective up to 10dB back-off. With $Z_{S(2\omega)}$ tuning (∘ plot), some improvement in linearity at power levels close to P1dB compression point were achieved but at the expense of linearity at lower powers. Finally, with the combination of DS and $Z_{S(2\omega)}$ tuning (⋄ plot), significant improvements in linearity is possible up to 4dB back-off from the compression point.

3.3 Input and output harmonic impedance control

In the previous section, linearity improvement at higher power levels was achieved by tuning only the second harmonic impedance at the source ($Y_{S_{2\omega}}$). This basic concept of harmonic impedance tuning was further extended in [17, 22] by including also the tuning of the second harmonic at the load side $Y_{L_{2\omega}}$. The combination of both input and output second harmonic tuning gives a greater flexibility to control the IM3 cancellation effects and can provide better results closer to compression.

In [17] the first combined load and source pull experiments were performed on Gen 4 LDMOS devices from NXP. The device used has a gate width of 2mm. They were biased in Class B operation and the fundamental load and source match were provided for maximum RF output power and efficiency. Linearity (two-tone) measurements were performed at a center frequency of 2GHz and a tone spacing of 100KHz. Figure 3.6 gives the results of the experiment. Note that in Gen 4 LDMOS devices, various technology improvement has been made [9] resulting in improved intrinsic linearity performance. The linearity performance of the Gen 4 device is comparable to that of Gen 2 with DS applied.

The □ trace gives the reference case where no optimization for the out-of-band harmonic impedances was used (i.e. all out of band impedances were set to zero). The ⬠ trace shows the case where only $Y_{S_{2\omega}}$ was optimized with the second harmonic termination at the output set to a short (similar as in Section 3.2). For these conditions the optimum value of $Y_{S_{2\omega}}$ works out to be zero (open circuit condition), yielding an extended linear power range (up going IM3 slope occurs at the higher power level) at the price of some linearity degradation at lower power. It can be noted that the IM3 sweet spot moves closer to the P1dB compression point, while the linearity is slightly better everywhere in the high power region. These results are consistent with the results in Section 3.2. Lastly, the ∘ trace shows the case where both $Y_{S_{2\omega}}$ and $Y_{L_{2\omega}}$ are simultaneously optimized. This optimization yields an IM3 improvement even closer to the P1dB compression point. The corresponding second harmonic admittances that correspond to this optimum are:

$$Y_{S_{2\omega}} = Y_{L_{2\omega}} = 0 \tag{3.10}$$

Under these 2nd conditions, the IM3 sweet spot is 7 dB closer to the P1dB compression point (in comparison to the reference case). This drastic change results in a significantly better linearity for the high-power region. The P1 dB compression point

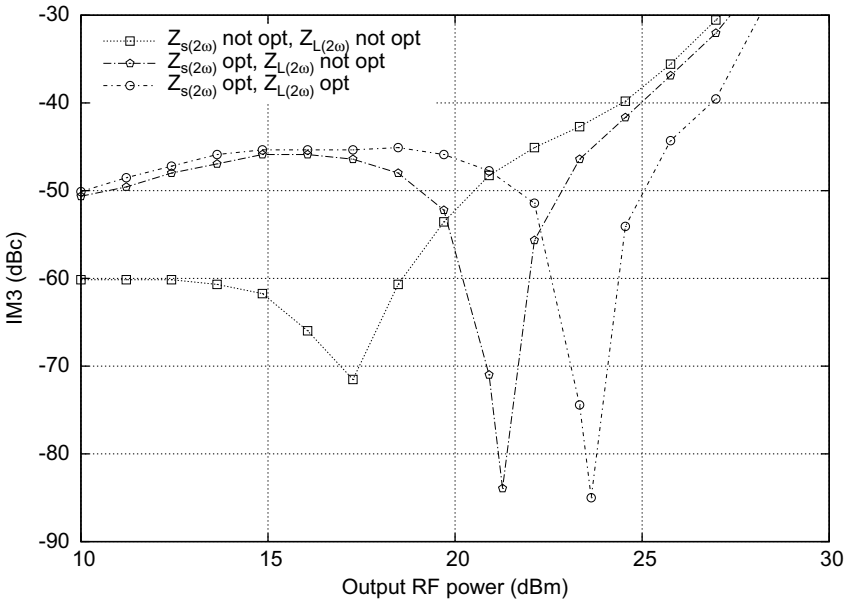


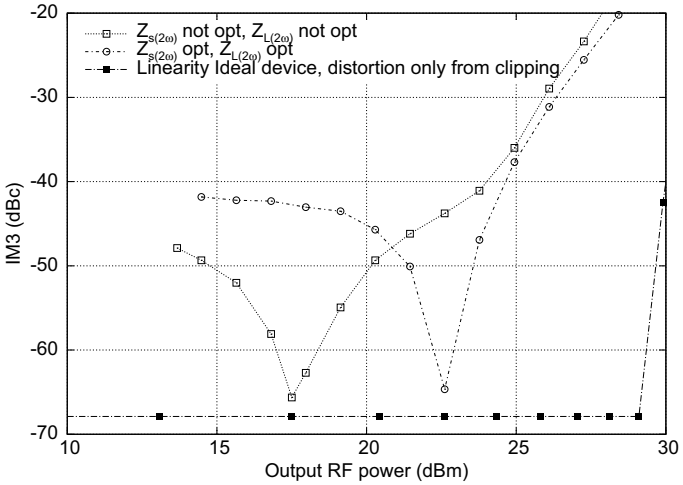
Figure 3.6: Measurements from [17], demonstrating linearity improvements for Gen 4 when using out-of-band tuning at both input, as well as, output of the Gen 4 LDMOS device. (f_o = 2GHz, Δf = 100KHz). Courtesy of Mr. Dave Hartskeerl and Dr. Marco Spirito.

itself is hardly affected by these 2nd harmonic manipulations, since it is predominantly set by the fundamental matching conditions.

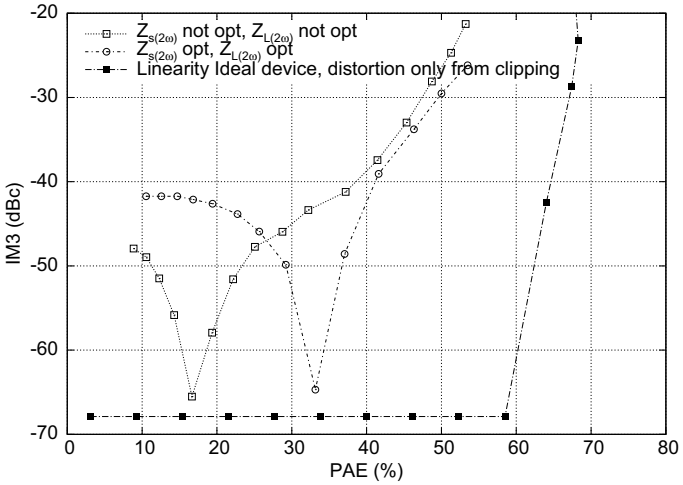
The positive effect of improved linearity close to the P1dB point is two fold. First the power utilization factor (PUF) is maximized and second, higher efficiency can be achieved while still meeting the linearity specifications, yielding significant improvements when using complex modulated signals such as IS-95 or WCDMA.

3.4 LDMOS Technology progress

With the optimum matching strategy known, it is interesting to evaluate the progress in LDMOS technology. In [22], the NXP LDMOS technology progress was evaluated up to Gen 6. In [17] also the same optimization procedure was utilized for the Gen 6 LDMOS technology generation. Linearity measurements were made using the same center frequency and tone spacing as before. The second harmonic conditions resulting in optimum linearity at high power were found to be close to those for the Gen 4 device evaluated in [17]. Also a similar IM3 trend was observed for the Gen 6 device as for the previously evaluated Gen 4 device (see Figure 3.7). Subsequent measurement with an IS-95 forward link signal (crest factor of approx 6.5dB) resulted in a drain efficiency of 44% at an ACLR1 level of -45dBc (see Figure 3.8), just 5% shy of the efficiency of an ideal device (normalized to the same maximum power) at the -45dBc



(a) Two-tone measurements (f_o = 2GHz, Δf = 100KHz). Linearity at output of a two-tone signal versus average output power.



(b) Two-tone measurements (f_o = 2GHz, Δf = 100KHz). Linearity at output of a two-tone signal versus efficiency.

Figure 3.7: Measurements from [22], demonstrating linearity improvements for Gen 6 when using out-of-band tuning at both input, as well as, output of the Gen 6 LDMOS device. Efficiency, linearity computation of the ideal device is based on an ideal Class-B amplifier which has peak CW efficiency of 78% and peak power normalized to that of the devices used in the simulation (32.5dBm). Distortion components for the ideal device are generated when the signal clips. Courtesy of Dr. Marco Spirito.

ACLR1 level. The source of distortion for the ideal device is from signal clipping.

This impressive efficiency performance was reported as a linearity-efficiency record for "Class-B" operated amplifiers using LDMOS technology in [22]. In this work we have extended these experiments to include the latest progress in LDMOS technology. In addition we have also added the theoretical limit for an ideal device (peak power is normalized to that of the devices used in the measurements) that is only limited by clipping of the modulated signal.

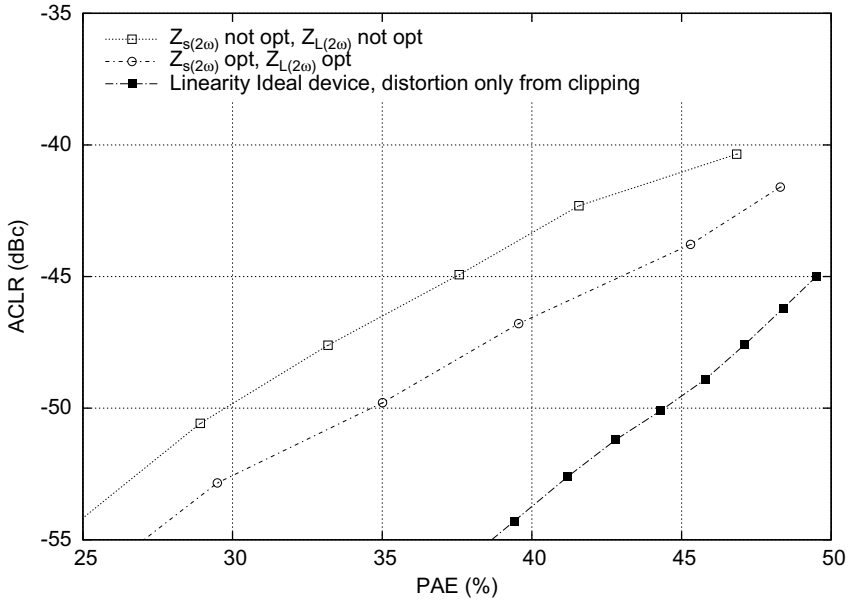


Figure 3.8: IS-95 forward link measurements from [22] at f_o = 2GHz, demonstrating linearity improvements for Gen 6 when using out-of-band tuning at both input, as well as, output of the Gen 6 LDMOS device. Efficiency, linearity computation of the ideal device is based on an ideal Class-B amplifier which has peak CW efficiency of 78% and peak power normalized to that of the devices used in the simulation (32.5dBm). Distortion components for the ideal device are generated when the signal clips. Courtesy of Dr. Marco Spirito.

3.5 Conclusion

In this part of the thesis work, two techniques have been utilized to improve the linearity and efficiency performance of "class-AB" amplifier stages, namely derivative superposition and the out-of-band matching technique. Both techniques, as well as their combined use, rely on basic understanding of the distortion mechanism present in the single stage PA cell and prove to be useful in minimizing the intermodulation distortion. In summary,

1. **Derivative Superposition**
   Originally introduced to shape only the transconductance through superposi-

tion of individual device characteristic, in this work this technique was analyzed with close form analytical equations and shown that the non-linear input capacitance C_{gs} also needs to be considered while applying DS, to obtain an overall improvement in linearity at low to medium power levels (relative to the 1dB compression point).

2. **Derivative Superposition with second harmonic control**
   Since derivative superposition proved to be inadequate to provide better linearity close to compression, as next step, 2nd harmonic input control was utilized to move the IM3 sweet spot closer to compression. It was found that using this second harmonic source impedance, a significant linearity improvement at medium to high power levels could be established. Experiments have shown that the combination of DS and optimum 2nd harmonic input matching can provide improved linearity for all power levels.
3. **Derivative Superposition with 2nd harmonic input and output control**
   To push further the linearity vs. efficiency trade-off and obtain linearity improvements at even higher power levels, 2nd harmonic tuning at both source and load can be utilized. By independent control of the input and output 2nd harmonic impedances, using a custom active load-pull setup, linearity improvements at high power levels were demonstrated up to the P1dB compression point, with only a very limited linearity degradation at low powers levels. For modern complex modulated signals (e.g. IS-95) this later technique using the latest LDMOS technology provides a linearity-efficiency trade-off which becomes closer to the theoretical limit given by an ideal signal clipping limiting device.

By using DS and no longer demanding second harmonic shorts, "Class-B" operation (which originally assumes 2nd harmonic shorts) has been given a new future, since it provides low-cost, low-complexity, high-linearity and high-efficiency single-stage amplifier performance. With these latest results the author is of opinion that the modified Class-B PA cell is now very close to its theoretical limit. Marginal improvements can still be made by looking at techniques to reduce losses in both the active device (processing technology) or the peripheral matching network. However, to realize a further significant break through in efficiency performance, it is necessary to look at high-efficiency PA architectures (system solutions), a subject which we will consider in the next chapters.

In view of this, when aiming for highly efficient power amplifier architectures, such as Doherty and LINC, also linear operation can be achieved, when the power amplifier cells are based on ideal devices. However in practical implementation of these high efficient architectures, the different power amplifier cells interact with each other; consequently even if the PA cells are "linear enough", the resulting interaction would more often results in a unacceptable overall linearity performance. In these configurations, a circuit solution is often not practically feasible due to the complexity of the problem statement or the required complexity of this linearization. As such a system solution such as predistortion (digital for added flexibility/reconfigurability) is proposed to be used in conjunction with these high efficient PA architectures.

The following chapter looks into predistortion as a linearization technique, and for this purpose a predistortion (PD) algorithm was implemented, for both the memoryless and the case with memory. We will use the PD algorithm in our later Doherty power amplifier experiments and the adaptive PA for handsets.

Chapter 4

Predistortion of Power Amplifiers

As industry embraces more complex PA architectures such as Doherty and Envelope Tracking to improve for efficiency, simple circuit solutions to achieve high linearity are no longer become feasible and digital predistortion becomes mandatory. This has yielded the situation that at the time of writing, Digital Predistortion (DPD) has become an industrial standard for tackling nonlinear distortion in higher complexity power amplifiers. This trend is accelerated by the the fall in cost of digital electronics, making it very attractive to incorporate a digital predistorter within the transceiver system/ICs.

In this thesis work, the focus has been on the development of custom predistortion algorithms to meet the requirements of highly efficient but rather complex PA concepts, which will be covered in later chapters. In this chapter, the basic principles of predistortion for a single Class-B PA stage will be first introduced and further investigated. Section 4.1 considers the simplest case, namely the predistortion of a memoryless non-linear system. In Section 4.2, the case of a non-linear PA with memory effects is discussed and the required extensions in the compensation algorithm are described. The effectiveness of the algorithms are verified with both simulations and measurements.

4.1 Linearizing the memoryless Power Amplifier

To aid understanding of predistortion, it is helpful to use a system theory approach. A system is linear if the output signal is always a linear scaled representation of the input signal. This implies that the Gain (G) is constant and independent of the actual input signal level. A system is considered to be non-linear if its transfer (e.g. Gain) depends on the input signal (x) as illustrated in Figure 4.1.

The nonlinearity of a system can be modeled in various ways. The easiest would be the use of a polynomial approximation as in equation (4.1). Note that such a power series approximation is only valid in a very limited frequency range and assumes the

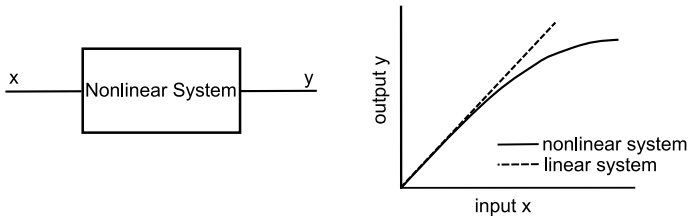


Figure 4.1: A non-linear system with its in and output relation

absence of memory effects like charge storing.

$$y = a_1\,x + a_2\,x^2 + a_3\,x^3 + a_4\,x^4 + a_5\,x^5 \tag{4.1}$$

Furthermore, a_1 to a_5 are the complex power series coefficients, which can be simplified to real in cases where no phase information is required. a_1 is the desired linear transfer and the rest of the other coefficients causes the deviation from the ideal constant gain. In the following analysis, relatively low power levels are considered (with respect to the peak power that the device or system can deliver), since this would result in $x \ll 1$ and as such only the lower order terms play a significant role, and terms raised to a power above three can be omitted.

We will now assume that the input is excited with a two-tone signal both of equal magnitude and here conveniently set to unity:

$$x = \cos(\omega_1 t) + \cos(\omega_2 t) \tag{4.2}$$

Substituting (4.2) representing a two-tone signal into equation (4.1), and collecting the terms which are associated with the fundamental and IM3 frequencies, the following expression can be obtained.

$$y = (a_1 + \frac{9}{4}a_3)[\cos(\omega_1 t) + \cos(\omega_2 t)] + \frac{3}{4}a_3[\cos(2\omega_2 t - \omega_1 t) + \cos(2\omega_1 t - \omega_2 t)] + \cdots \tag{4.3}$$

Note that the term $\frac{9}{4}a_3$ represents gain compression or expansion of the fundamental tones as function of the input variable x and $\frac{3}{4}a_3$ is the coefficient related to the third order IM3. When using the nonlinear system of Figure 4.1 as part of a larger configuration which needs to behave linear for its input – output relation, one has to compensate for the nonlinear part of the system. For this purpose another non-linear system (the predistorter) can be placed in cascade as shown in Figure 4.2.

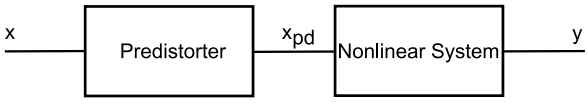


Figure 4.2: Cascaded non-linear system with predistorter

For the transfer of the predistorter block we assume, the following:

$$x_{pd} = b_1\, x + b_3\, x^3 \tag{4.4}$$

Note that this predistorter has besides its linear transfer only a third order non-linearity. The overall transfer function of the predistorter and non-linear system to be linearized can now be written as:

$$\begin{aligned} y = a_1 b_1 x + a_2 b_1^2 x^2 + \left(a_1 b_3 + a_3 b_1^3\right) x^3 \\ + 2\, a_2 b_1 b_3 x^4 + 3\, a_3 b_1^2 b_3 x^5 + a_2 b_3^2 x^6 \\ + 3\, a_3 b_1 b_3^2 x^7 + a_3 {b_3}^3 x^9 \end{aligned} \tag{4.5}$$

By setting $b_1 = 1$ and $b_3 = -\frac{a_3}{a_1}$ the third order coefficient of the overall transfer would be set to zero. It is important to note that although the original non-linear system has only second and third order nonlinearities; due to the use of predistorter higher order non-linearities are introduced/generated. Note that these higher odd order non-linearities (fifth, seventh and ninth) directly contribute to the distortion products at the IM3 frequencies. From this result it can be concluded that a third-order predistorter used to compensate a third-order non-linear system will never be able to completely eliminate all third-order distortion, furthermore higher-order distortion (which is typically lower in power) are added into the process. This phenomenon can be also explained by the following.

The following mathematical equations represent the transfers of the block as functions.

$$y = k\, \mathbb{F}(x_{pd}) \tag{4.6}$$

$$x_{pd} = \mathbb{G}(x) \tag{4.7}$$

Where k is a constant. The output as a result of the cascade is:

$$y = k\, \mathbb{F}(\mathbb{G}(x)) \tag{4.8}$$

To obtain a linear output, the transfer of the predistorter should be $\mathbb{G} = \mathbb{F}^{-1}$. In the example that was used before, $\mathbb{F}$ was a third order polynomial. However, it can be shown that the inverse of a third-order polynomial will be a polynomial of an infinite order. This implies that in order to completely cancel out the distortion of a nonlinear system, a predistorter with an infinite amount of higher-order non-linearity is required. In practice such a requirement translates to an infinite bandwidth for the predistorter to handle al the distortion products. Therefore at this juncture, it can be concluded that there is no practical predistorter possible that can completely reduce the distortion of the overall system to zero.

4.1.1 Practical implementation of a memoryless predistorter

We will now discuss a practical implementation of a memoryless predistorter. To best illustrate the algorithm for memoryless correction, accompanying simulation results

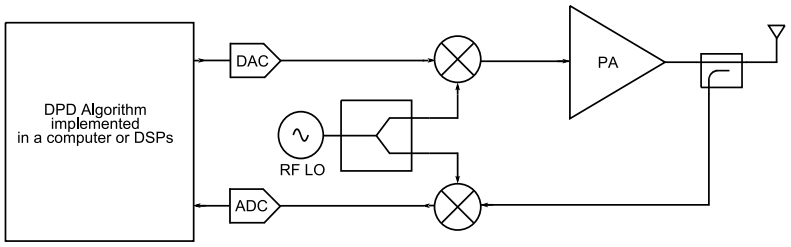


Figure 4.3: General hardware setup for digital predistortion

are provided. A realistic Class-B PA design using accurate transistor models of Motorola devices (LDMOS LV1 technology) and realistic models of passive components for matching and bias are used. This PA model is provided in the design examples of ADS, and has been optimized for applications in the 870MHz range.

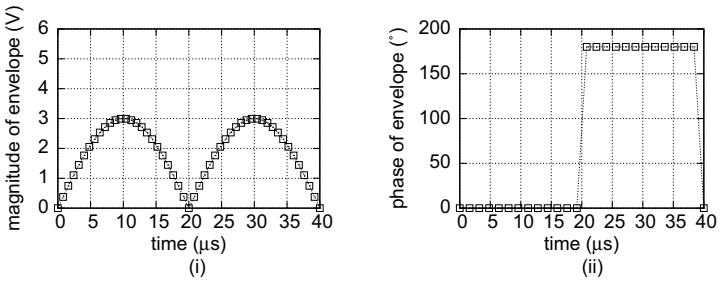
The first step in the predistortion, involves characterizing the non-linear behavior of the PA (determining $\mathbb{F}$ in equation (4.6)). In a memoryless system all the information for $\mathbb{F}$ can be found in its AM-AM and AM-PM characteristics. The AM-AM and AM-PM data can be collected in two ways. The first method is stepping an input CW signal for its power level and measure the resulting magnitude of the output power and its corresponding phase. The second method, which is the one used in this thesis, uses a two-tone test signal at the input of the PA. The instantaneous peak power of the two-tone should correspond to the deliverable peak power of the PA. Under this input drive condition, the PA is excited with instantaneous power over its complete range. The two-tone signal is created by first generating a base-band sine wave and mixing this directly with a RF LO signal to the frequency of interest, as demonstrated by the following equation:

$$\begin{aligned} x = & \overbrace{cos(\omega_{bb} \cdot t)}^{\text{base-band signal}} \cdot \overbrace{cos(\omega_{rf} \cdot t)}^{\text{RF LO}} \\ = & \underbrace{\frac{1}{2} cos(\omega_{rf} \cdot t + \omega_{bb} \cdot t)}_{\text{high tone}} + \underbrace{\frac{1}{2} cos(\omega_{rf} \cdot t - \omega_{bb} \cdot t)}_{\text{low tone}} \end{aligned} \tag{4.9}$$

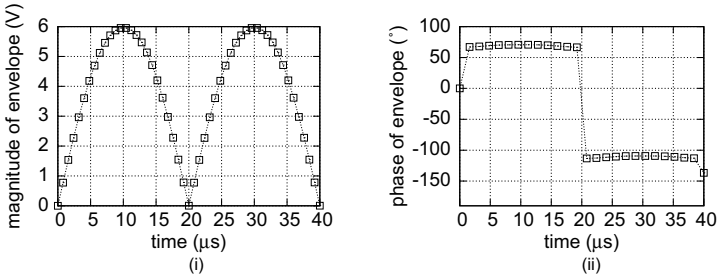
Where $\omega_{bb} = 2\pi f_{bb}$ refers to the base-band frequency and $\omega_{rf} = 2\pi f_{rf}$ refers to the RF frequency.

Figure 4.3 shows a block diagram schematic which is commonly used for digital predistortion (DPD). The base-band signals are generated in the digital domain (in this thesis work, it is done in MATLAB), and up-converted to RF through a mixer. At the output of the PA, the signal is down converted and subsequently digitally sampled to obtain the base band envelope and phase information extracted (Figure 4.4). This down conversion from RF to IQ is an available feature provided in the Agilent PSA.

Figure 4.5 shows the extracted AM-AM and AM-PM using a base-band sine wave of 25 KHz. Note that since we are working predominantly with base-band envelope data in DPD systems, it is useful to analyze data in the voltage domain.



(a) Input base-band signal



(b) Output down-converted base-band signal

Figure 4.4: Simulations are done in ADS based on Envelope Simulation using a Motorola LDMOS LV1 based PA (also provided in the ADS examples).

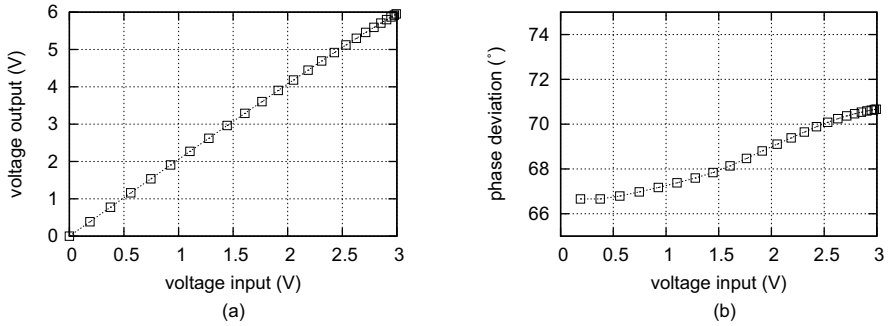


Figure 4.5: Simulated input-output AM-AM and AM-PM voltage transfer functions of a Motorola LDMOS LV1 based PA.

The plots compare the input base-band voltage to the output base-band voltage and its phase deviation (complex envelope). For this particular PA, the AM-PM is a much more severe source of distortion than the AM-AM (Gain compression). The output magnitude compresses only minimally in the high voltage region, compared to the phase deviation which starts already at low input voltages.

The next step in the process is to arrange the collected data into a look-up-table (LUT), with the output voltage (y) as the index to the LUT. In this way the index points to the relevant input voltage x and its corresponding phase to obtain the desired y. Figure 4.6 shows the graphical representation of this LUT.

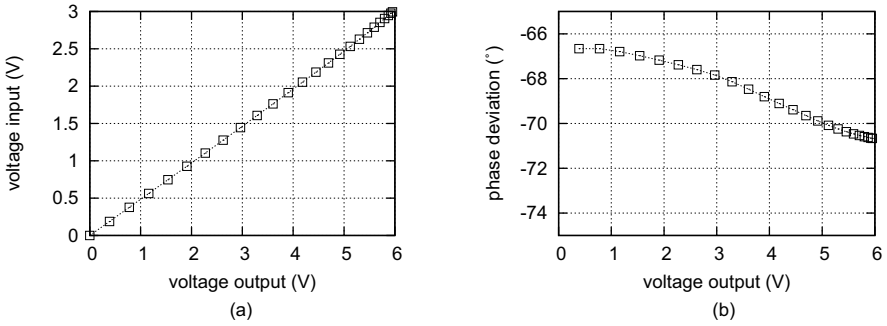


Figure 4.6: Graphical representation of a voltage based Look-Up-Table (LUT) for pre-distortion of a Motorola LDMOS LV1 based PA.

Thus, to obtain a base-band sine wave (clean two-tone) of peak envelope voltage of 6 volts at the output of the PA. The values of the base-band sine wave are used as index to the LUT, and the resulting predistorted input signal to drive the PA is shown in polar form in Figure 4.7. Note that these plots represents the base-band samples in time and the time per unit should correspond to the sampling rate in an appropriate way since this sets the frequency spacing of the two-tone signal at the RF frequency. As expected, not much predistortion is required for the amplitude of the envelope but the phase of the base envelope requires introducing phase offsets. And the maximum offset corresponds to the position where the envelope has its peak value.

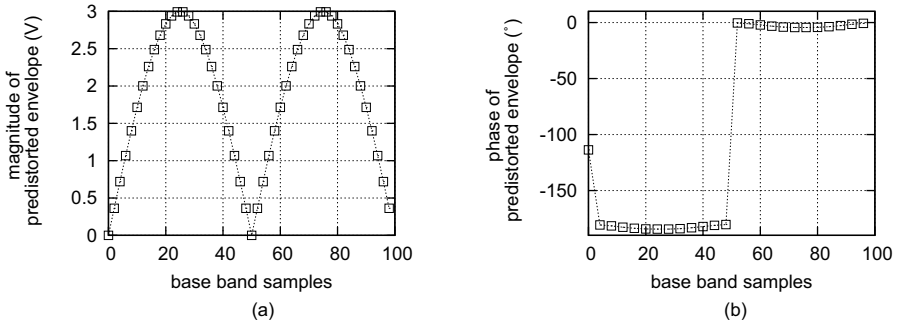


Figure 4.7: Predistorted base-band signal used to correct the distortion of the PA.

To obtain a "clean" two-tone signal with a tone spacing 50KHz at the output of the PA. The predistorted signal is up converted to a RF center frequency of 870MHz

and used as the new input to the PA. Figure 4.8 shows the output power spectrum of the PA with and without the predistortion. The IM3 and IM5 side bands are suppressed by approximately 25 dB and 22 dB respectively. It is important to note that this is at the maximum output power level that this PA can deliver for a two-tone signal.

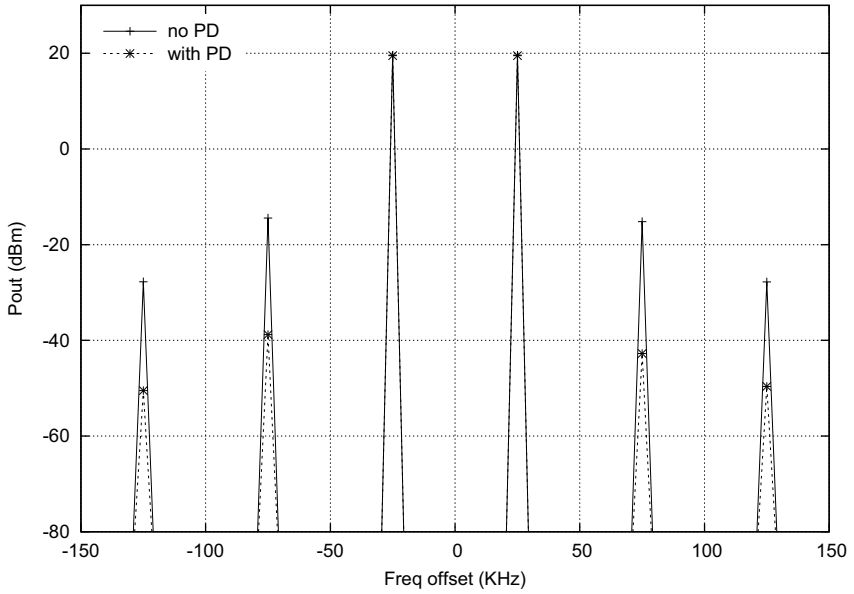


Figure 4.8: Simulated two-tone output spectrum at maximum output power for the Motorola LDMOS LV1 based PA with and without memoryless predistortion. The tone spacing and center frequency of the two-tone signal are 50KHz and 870MHz.

Assuming now that the non-linear properties of the PA remains approximately the same when a two-tone with a wider tone spacing is used, the same signal shown in Figure 4.7 can be used with a higher sampling rate and introduced as the new input to the PA (two-tone of 5MHz tone spacing). Figure 4.9 shows the new results of these simulations with this larger tone spacing. Note that now the suppression of IM3 and IM5 is no longer symmetrical (high and low are not equally suppressed). Even more the IM3 and IM5 reduction due to the predistortion is also not as effective as for the narrow tone case handled previously.

To further understand why memoryless predistortion is not effective for a large two-tone spacing, it is helpful to examine the envelope waveform of the down converted signal. Figure 4.10 shows the simulated envelope waveform at the output of the PA. In comparison to the envelope waveform of a two-tone with narrow tone spacing, the waveforms here exhibit a distinct asymmetry. If a base-band input and output mapping was constructed from this data, there would be for every output y two different values of x. This has serious implications for the memoryless predistorter, regardless of how the LUT is constructed (using CW AM-AM and AM-PM measurements or with a slow two-tone signal as described in the previous section),

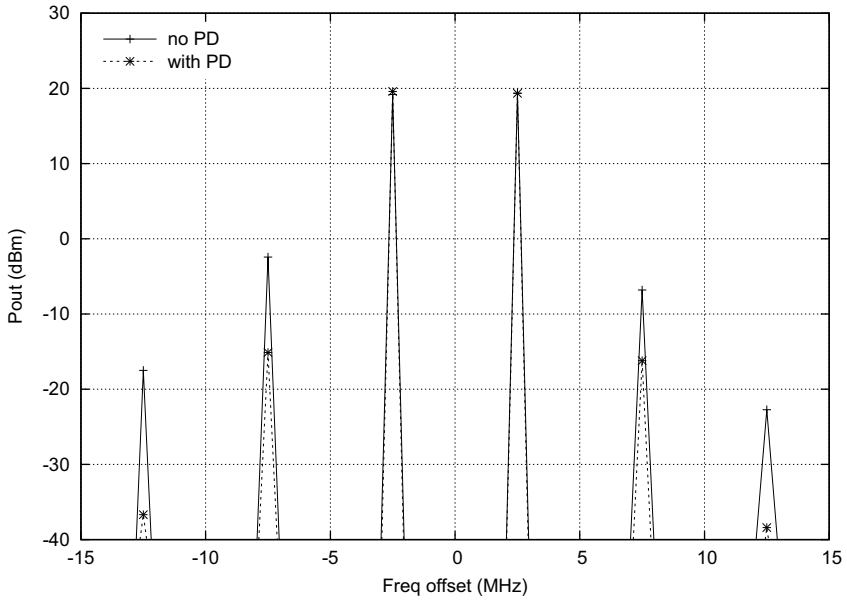


Figure 4.9: Simulation result of a memoryless predistortion of a two-tone signal. Tone spacing=5MHz. Center frequency=870MHz. Peak power of two-tone reaches peak power of PA

because memoryless predistortion works on the assumption that the system to be linearize has a unique input-output relationship. This sort of distortion is often referred to slow-wave effects or otherwise known as memory effects and would be extensively discussed in the next section.

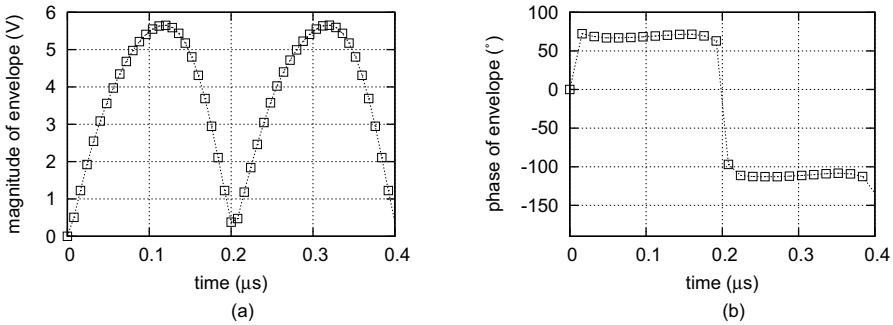


Figure 4.10: Waveform envelope at output of PA. Tone spacing=5MHz. Center frequency=870MHz. The peak power of two-tone equals the peak power of PA. No pre-distortion is employed in this experiment.

From these series of experiments, it has been shown that for a modulated signal

that has a bandwidth that is narrow enough such that it falls in the range where the PA is memoryless, then the above mentioned predistortion algorithm is effective in improving the linearity performance. Measurement verification of these experiments are shown in Chapter 6, where this predistortion concept was used to improved the linearity of a class AB PA under static load-line modulation.

4.2 PA with memory effects

In a system with memory, the output signal at time t is not only a function of the input signal applied to the amplifier at that time t but also from the forgoing history of the input signal. This is mathematically represented by the well known convolution of the input signal with the impulse response of the system as shown in Equation (4.10).

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)\mathrm{d}\tau \tag{4.10}$$

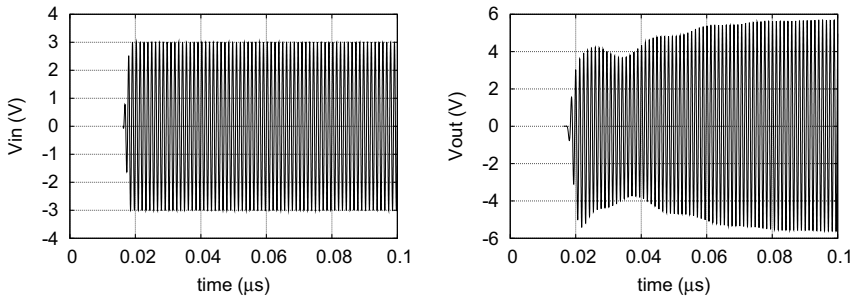
However, in the case of a non-linear system with memory effects, the situation becomes more complicated since the impulse response is now also a function of the instantaneous power. A mathematical formulation of such a system can be expressed using the Volterra/Wiener equations, which are the basis of the well known Volterra Series calculation techniques. Equation (4.11) shows the third-order Volterra kernel, representing the third-order linearity with memory effects. This equation shows that the output is a of three dimensional convolution of the input signal with the impulse response, which is again a function of three time constants.

$$y(t) = \iiint_{-\infty}^{\infty} h(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)\mathrm{d}\tau_1\mathrm{d}\tau_2\mathrm{d}\tau_3 \tag{4.11}$$

Although Volterra series analysis is probably the most appropriate tool to analyze and model non-linear systems with memory. It has been shown in Chapter 2 that when higher order nonlinearities above the third are taken into account, the complexity of the expression increases dramatically. Furthermore, if impedances at the IF and higher harmonics are taken into account, the equations become soon uninterpretable. The other even more important reason to avoid Volterra series for this application, lies in the difficulty of finding an analytical inverse function required for the pre-distortion action. As such, a logic compromise is to seek where ease of implementation is traded off against accuracy of representation of the system.

Memory effects can come from distortion products at higher harmonic frequencies mixing back to the in-channel, but more often the main contributors come from effects at the IF frequency, such as self heating (thermal memory effects) and the effects of the bias-circuitry. In a RF power amplifier, the only memory effects that are detrimental to in channel spectral regrowth are the ones that affect the envelope (amplitude and phase), as was seen in Figure 4.10.

Thus to characterize a non-linear PA with memory, it is helpful to introduce the concept of an envelope step/pulse response. Figure 4.11(a) shows in simulation, the pulsed RF waveform at the input of the Motorola LDMOS LV1 PA, and Figure 4.11(b)



(a) Input RF Signal (b) Output RF Signal

Figure 4.11: Transient Simulations done in ADS to demonstrate the transient response of the Motorola LDMOS LV1 PA.

shows the result of the transient simulation at the output of the PA. Subsequently, if one were to down-convert the signal at the output to base-band (magnitude and phase) and differentiate it once, this will result in what we refer in this thesis work to the envelope pulse response.

Fortunately, the simulator ADS allows us to obtain this information directly using Envelope Simulation. An envelope pulse response describes how the output (magnitude and phase) changes as a consequence of exciting the non-linear system with a very narrow modulated pulse, equivalent to an impulse in the envelope sense. Because this system is both non-linear and has memory, this envelope pulse response will be a function of the magnitude of the pulse itself.

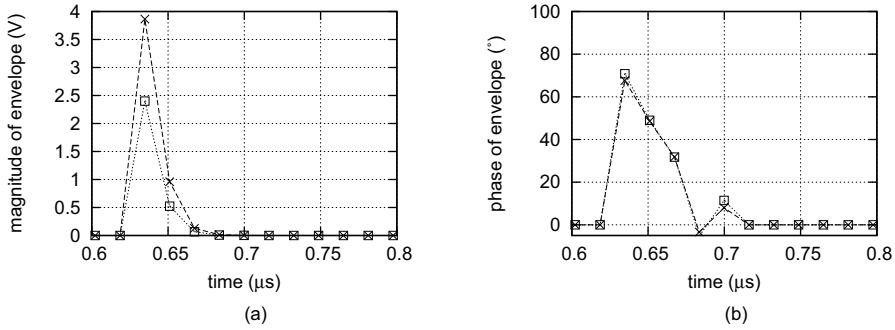


Figure 4.12: Envelope waveform at output of PA in response to a impulse at the base-band at the input.

Figure 4.12 shows the simulated envelope pulse responses of the same amplifier using Envelope Simulation. The envelope pulse response at the output of the amplifier results from an RF modulated pulse stimulus at maximum input power and at 6 dB back-off provided at the input of the amplifier. Note that if the PA was memoryless, the output envelope would also have been a pulse. i.e. no transient decay would occur

as shown in Figure 4.12. Furthermore this "smearing" out of the energy takes place over approximately 50 ns, which is a considerable fraction of the envelope of a two-tone signal with 5 MHz bandwidth. This explains the asymmetry in the envelope of the output two-tone waveforms (Figure 4.10). In the following section, the technique used for memory effect compensation in this thesis work will be discussed.

4.3 Linearizing Power Amplifiers with memory effects

A mathematical way to formulate the concept of memory in a non-linear system that was discussed in the previous section is the use of memory polynomials, as shown in equation (4.12). This is basically the familiar power series representation introduced in Section 4.1, extended in time (delay taps), but note that v_{out} and v_{in} are complex base-band variables. For this particular application of linearizing a non-linear system with memory only in the band of interest, the RF carrier can be ignored.

$$v_{out}(n) = \sum_{q=0}^{Q} \sum_{k=1}^{K} a_{k\,q} v_{in}(n-q)|v_{in}(n-q)|^{2(k-1)} \tag{4.12}$$

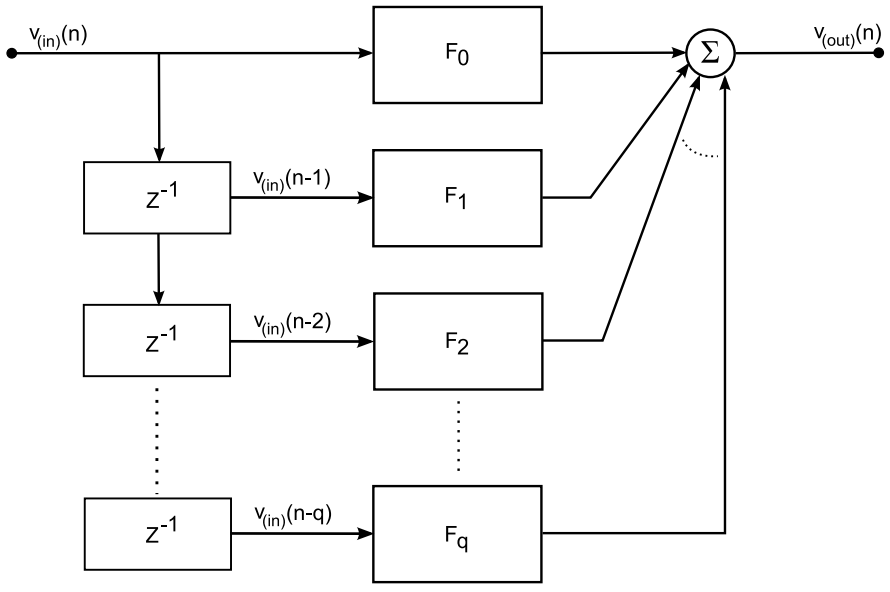


Figure 4.13: Block diagram demonstrating memory polynomials to model a non-linear PA with memory effects

Figure 4.13 shows figuratively the above equation. Note that the present output $v_{out}(n)$ is a linear combination of the non-linear transfer of the current input $v_{in}(n)$ and the other previous inputs $v_{in}(n-q)$. Though this mathematical formulation is

straight forward, it does not accurately model the non-linear system. This analysis technique fails to take into account cross term non-linearities, which is the effect where the non-linear transfer F_0 being also a function of $v_{out}(n-1)$ and other larger delay taps. This inadequacy also extends to the other transfer functions F_1 to F_q. But as was previously discussed, this compromise in accuracy of representation is justified by the ease of implementing the predistorter.

From the above figure and making the assumption that time delay greater than 3 time steps away no longer has any influence on the output, equation (4.12) can be simplified to the following.

$$\begin{aligned} v_{out}(n) &= F_0\left(v_{in}\left(n\right)\right) + F_1\left(v_{in}\left(n-1\right)\right) + F_2\left(v_{in}\left(n-2\right)\right) + F_3\left(v_{in}\left(n-3\right)\right) + \cdots \\ &\approx F_0\left(v_{in}\left(n\right)\right) + F_1\left(v_{in}\left(n-1\right)\right) + F_2\left(v_{in}\left(n-2\right)\right) + F_3\left(v_{in}\left(n-3\right)\right) \end{aligned} \tag{4.13}$$

Note that in this thesis, the second, third and fourth term in equation (4.13) are referred to residual memory. To determine the input v_{in} at time n necessary to obtain a v_{out} at time n equal to y, equation (4.13) can be rearranged to give the following.

$$v_{in}(n) = F_0^{-1}\left(y\right) - F_0^{-1}\left(F_1\left(v_{in}\left(n-1\right)\right) + F_2\left(v_{in}\left(n-2\right)\right) + F_3\left(v_{in}\left(n-3\right)\right)\right) \tag{4.14}$$

This implies that to compute the required v_{in} at time n, the vector sum at the output due to the previous inputs must first be computed (residual memory). Then a subsequent computation F_0^{-1} of the residual memory to determine the input at the current time n is required to cancel this memory residue. With the memory residue canceled, $F_0^{-1}(y)$ at the input would give y at the output resulting in the final predistorted v_{in} which is the vector sum of these two components. The above is the operation principle of a predistorter with memory compensation.

4.3.1 Practical implementation

A practical implementation of the predistorter with memory compensation, is nothing more than taking the information of the desired output signal and applying equation (4.14) to the complete output signal in order to compute the required predistorted input. In order to do this, it is clear that F_0 to F_q first have to be characterized, where q is equal to the number of delay taps that is required. For this action, it is useful to look at the envelope pulse response described in section 4.2. Figure 4.12 shows the envelope pulse response at the output of the non-linear PA with memory, for two different input voltages. The peak value at the magnitude plot and its corresponding phase $v_{out}(n)$ is related to the input $v_{in}(n)$ through F_0, and one time step away (how this time step is chosen will be explained in the following paragraph,) the output magnitude and phase of $v_{out}(n+1)$ is related to the input $v_{in}(n)$ through F_1. The same conclusions can be drawn for the rest of the transfer functions F_2 to F_q. If a

full range of input voltage pulses are used covering all possible input levels, then all the required transfer functions can be completely characterized.

To illustrate the concept, the above mentioned algorithm has been applied to the PA demonstrator that is used in Section 4.1. This PA was shown to be affected by memory effects when the envelope frequency is in the MHz range. A sweep for the spacing of the two-tone tones was performed to study the arising asymmetry of the IM3 components. Note that this is one of the traditional proposed methods [23] to determine memory effects. Following this test its is claimed that the more asymmetrical the IM3 components, the more difficult it is to compensate the amplifier for memory effects. It was found that the PA exhibits the biggest IM3 asymmetry at a tone spacing of 1.2 MHz.

Therefore, the goal is now to show that this predistortion algorithm is capable of compensating the non-linearity of a two-tone signal with tone spacing of 1.2 MHz. This implies that the base-band sine wave has a frequency of 600 KHz. Using 100 discrete samples to represent the sine wave of 600 KHz, will result in each sample having a time step of approximately 16 nsec. From figure 4.12, this time step falls into the category where the PA is still affected by residual memory, and the number of delay taps to accurately represent the residual memory is approximately five. The characterization procedure is done with a series of input modulated pulse voltages ranging from the minimum to maximum allowed input voltage, and the output envelope magnitude and phase information is tabulated to represent F_0 to F_5. Figure 4.14 shows the magnitude and phase of F_0 to F_2.

Applying the algorithm based on equation (4.14) on the PA under test already results in a significant suppression of the IMD products, however not as good as might be expected. This can be attributed to the earlier explained phenomenon that was excluded in our analysis. Namely, the fact that the transfer function itself, e.g. F_k where k can be zero or any of the tap values, is also a function of the previous inputs $v_{in}(n-k-m)$ where m is greater than zero. This implies that to get a better cancellation the transfer functions F_0 to F_5 have to be slightly modified. In this thesis, an iterative approach is used to implement this correction where after every iteration the look-up-tables are modified (by modifying the complex coefficients of the polynomial of every F_n) and the iteration stops when the desired specifications of IMD suppression is reached.

Figure 4.15 shows the simulation result of the two-tone signal before and after predistortion where three iterations of the algorithm was employed. The IMD components are suppressed by more than 60 dBc. Figure 4.16 shows the envelope waveform at the output of the PA before and after applying the predistortion with memory compensation. Note that the asymmetry in the magnitude and phase is completely removed after applying the predistortion algorithm.

4.3.2 Hardware Experimental Verification

With a successful verification of the predistorter with memory compensation in a simulation environment, the final step is to verify it with hardware. This brings in two extra elements that were not accounted for in simulations and which will affect the performance of the predistorter. The first being the presence of noise in the

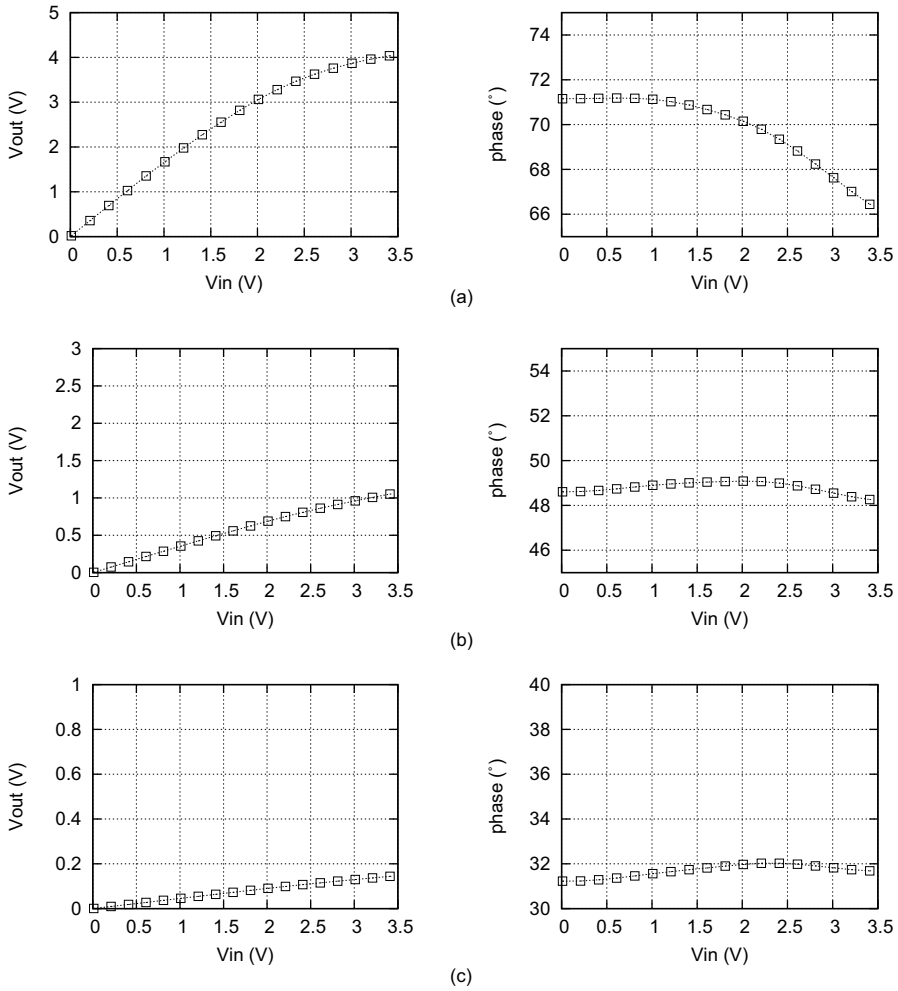


Figure 4.14: (a) Magnitude and phase of F_0. (b) Magnitude and phase of F_1. (c) Magnitude and phase of F_2. Values are based on simulated envelope pulse response data.

calibration routine (when building the LUT) and the second, thermal memory effects which were not included in the models used in the simulation.

Figure 4.17 shows the block diagram schematic of our custom amplifier test bench for performing digital predistortion. Its structure is based on the heterodyne transceiver architecture. The complex base-band signals are generated using MATLAB and uploaded as a digital 400 MHz IF signal to the Agilent N6030A Arbitrary Waveform Generators (AWG). This signal is further up-converted in the analogue domain to a center frequency of 2.14 GHz using a mixer with the LO port connected to a ESG providing a CW signal at 1.74 GHz. A high-pass filter suppresses the lower-

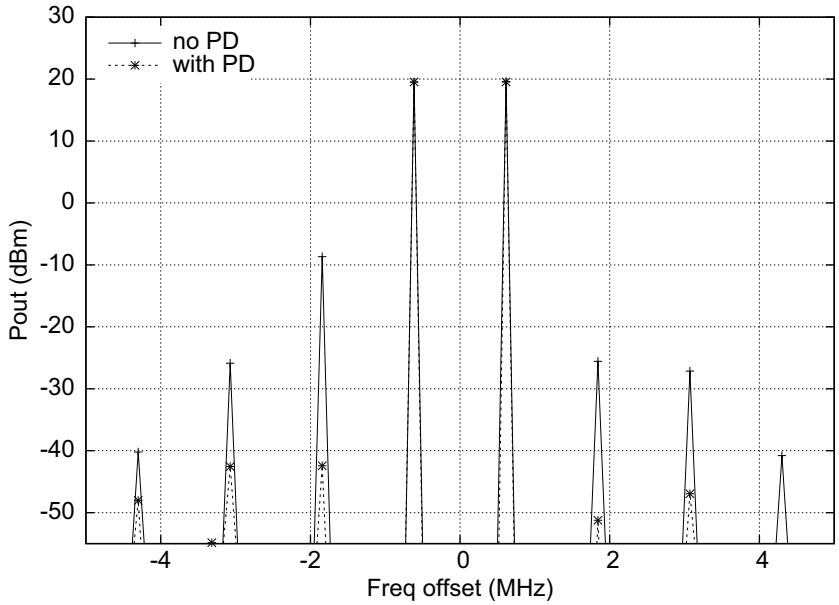


Figure 4.15: Simulation result of a predistortion with memory compensation of a two-tone input signal. Tone spacing=1.2 MHz. Center frequency=870MHz. Peak power of two-tone reaches peak power of PA. Power Amplifier under consideration is a Class-AB amplifier based on the Motorola LDMOS LV1.

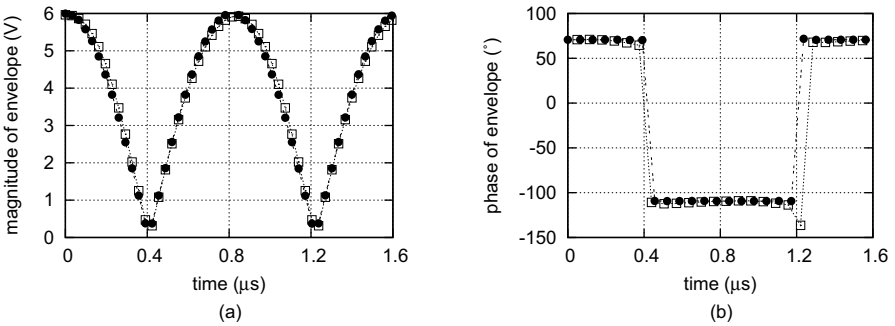


Figure 4.16: Envelope of the two-tone waveform at output of PA. Tone spacing=1.2 MHz. Center frequency=870MHz. The peak power of two-tone equals the peak power of PA. Before and after applying predistortion with memory compensation.

sideband and leakage of the oscillator. After filtering, the drive signals are amplified to reach the required input power levels for our PA under test.

At the output of the DUT, the maximum power is relatively high, thus to protect the instruments and to ensure detectability (without saturating the instruments), a two couplers configuration is used to split off a fraction of the output power, while

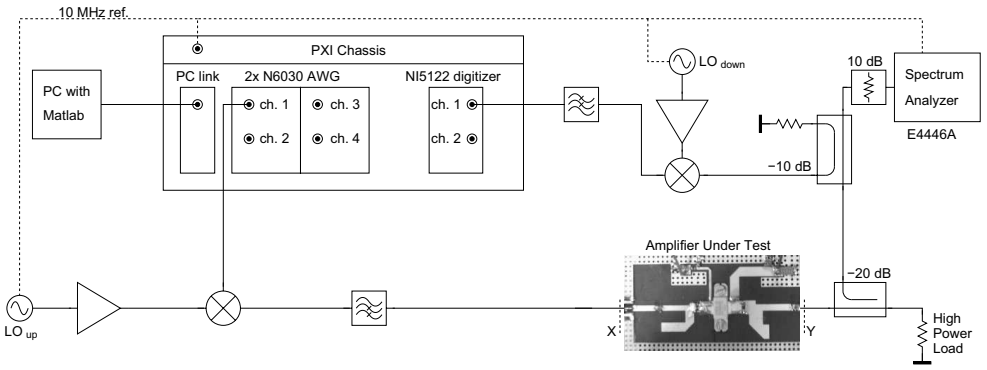


Figure 4.17: Hardware setup for digital predistortion with memory compensation

the remaining power is delivered to a dummy load. The detection signals are fed to a spectrum analyzer for the measurement of the (channel) powers and spectral purity, as well as to a down-converting mixer. This high-linearity down-converting mixer is driven by a 2115 MHz LO signal, which is frequency locked to the LO signal used for the up-conversion. Consequently, the output signal of the amplifier under test is down-converted in a linear fashion to an IF signal of 25 MHz. By using high speed digitizer cards from National Instruments (NI-PXI5122), this signal is acquired and through software operations converted to a complex baseband representation used for collecting the data to build the LUTs or for further iterations in the memory compensation algorithm.

At the input, a simple calibration routine is required to relate the base-band voltages at the AWG to actual input powers at plane X. This is done by generating a DC voltage at the AWG and stepping them from a minimum to maximum output voltage and measuring with a power meter the power at plane X. At the output, the power loss in the path between plane Y and the spectrum analyzer has to be calibrated to ensure accurate measurement of the output powers.

The PA hardware under test (see figure 4.18) is a 30 watt GaN (HEMT) Class B PA designed in house with the active device provided by CREE. Figure 4.19 shows the measured CW gain and efficiency of this PA. A peak power of 45 dBm is obtained when the device is driven 5 dB into compression and this corresponds to a peak efficiency of 68%.

The first step in the procedure involves "checking the amount of memory" in the PA. This would enable the operator of the DPD to determine the number of delay taps required. In a simulator environment it is relatively easy to perform a RF pulse modulated signal response. In fact, this is a matter of just having a pulse width, which is equivalent to one sample time in the base-band and then mixing it with the RF signal. At the output, in simulations the sampling rate can be made much higher than the actual base-band sample rate making it easy to accurately detect the response. Unfortunately in the hardware domain, the signal down conversion to IF is done in the analogue domain and the effective sampling rate will be equivalent to that of the base-band generated signal.

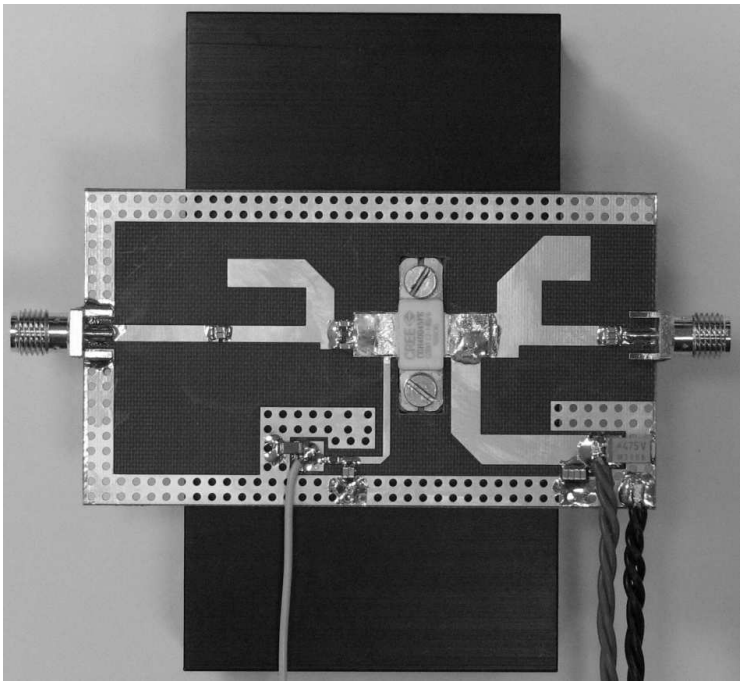


Figure 4.18: A 30 watt GaN (HEMT) Class B power amplifier stage used in the predistortion experiment including memory compensation.

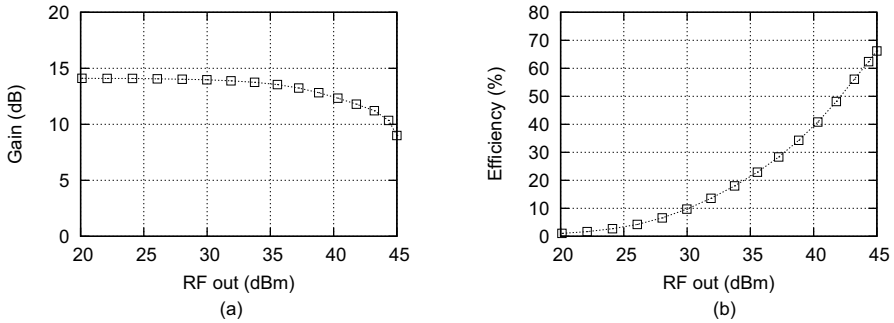


Figure 4.19: Measured CW (a) gain and (b) efficiency of the 30 watt GaN Class B PA stage at frequency of 2.14 GHz.

To get around this hardware restricted sampling speed speed, a modulated step response is used instead of the impulse response. To obtain the envelope information to build the LUT, the modulated pulse response is differentiated once. Figure 4.20 shows the measured magnitude and phase of the modulated step response and the subsequent derived magnitude and phase of the modulated pulse response, for two input voltage levels. Note that the time between each sample is 10 nsec, which is

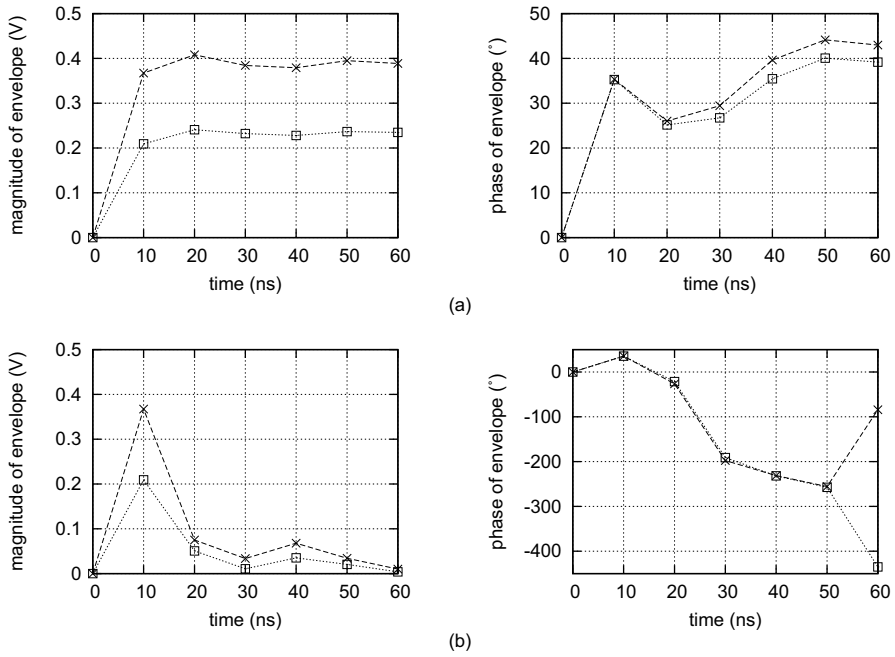


Figure 4.20: (a) Measured base-band magnitude and phase of a modulated step response at the output of the PA. (b) Derived based-band magnitude and phase of a modulated pulse response at the output of the PA.

sufficient for the predistortion of a WCDMA signal. From these data, it is clear that five delay taps are required, thus having a total of 6 LUTs (F_0 to F_5).

A WCDMA signal, test model 1 with 64 dedicated physical channels (DPCHs) is applied at the input of the PA under test. The signal has a 5 MHz bandwidth and a crest factor of 11.5 dB. The PA is driven such that the peak instantaneous power at the output reaches approximately the peak output power of the PA which is 45 dBm. The resulting measured average output power is 34.4 dBm which results in a peak power of 45.9 dBm. The predistortion with memory compensation algorithm is automated in MATLAB and set to iterate till the adjacent channel leakage ratio (ACLR) specification, ACLR1=-45dBc and ACLR2=-50dBc at 5 and 10 MHz offset are met.

Figure 4.21 shows the power spectrum at the output of the PA before and after applying PD with memory compensation, and Table 4.1 summarizes the other measured parameters. Note that there is still significant asymmetry in the ACLR2, which could in principle be fixed if more iterations are applied. However the routine in this experiment is chosen such that it stops when all the specifications are met.

Subsequently, a power sweep is made where the predistortion routine is applied at various average output power (for the same WCDMA signal). Figure 4.22 shows the measured drain efficiency of the GaN class B PA. From which it can be concluded that

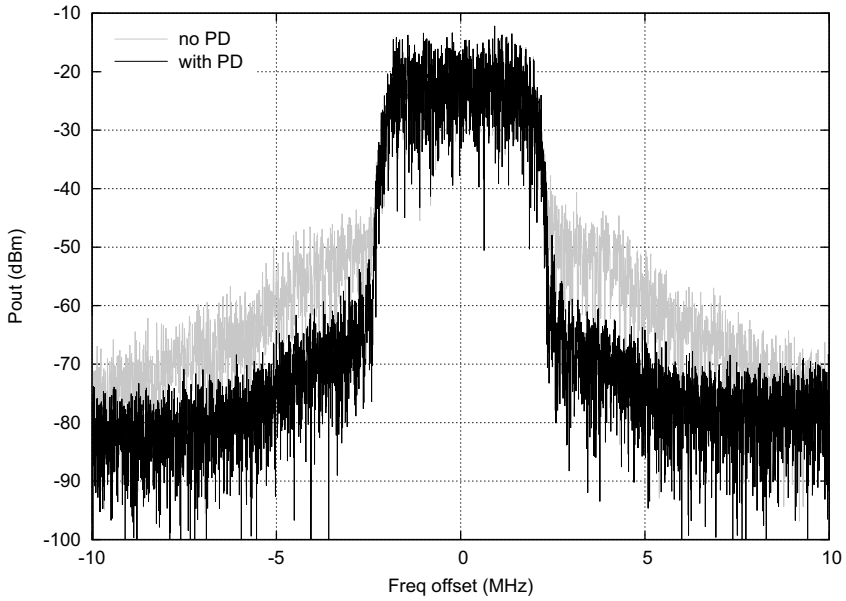


Figure 4.21: Power spectrum of a WCDMA signal at the output of the PA under test, before and after PD with memory compensation. Center freq.=2.14 GHz. Pout average=34.4dBm.

Table 4.1: Measured results of GaN Class B PA driven with test model 1 WCDMA signal, with and without predistortion.

	no DPD	with DPD
Inband Output Power (dBm)	34.4	34.5
Drain Efficiency (%)	24.2	24.8
ACLR1 low (dBc)	-33.8	-49.1
ACLR1 high (dBc)	-32.1	-49.4
ACLR2 low (dBc)	-51.6	-59.4
ACLR2 high (dBc)	-52.0	-53.6

the use of predistortion does not in anyway affect the efficiency of the PA. Note that the predistortion can not correct for hard clipping effects, so beyond 25% efficiency it is not possible to obtain sufficient correction to meet the ACLR specifications. Up to that point where clipping occurs the ACLR measurements shown (Figure 4.23) that the implemented predistortion routine is capable of linearizing the PA over a wide

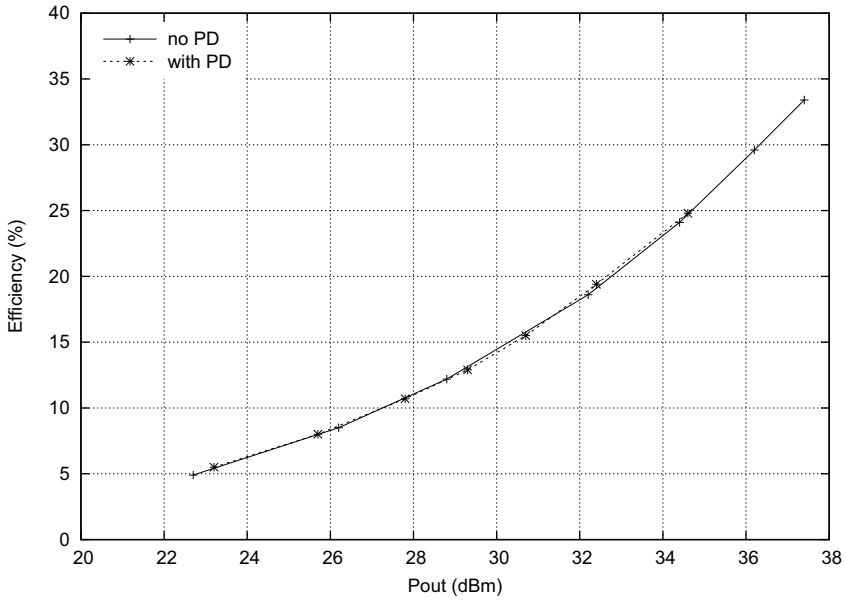


Figure 4.22: Measured drain efficiency of the GaN class B PA as a function of swept average output power of the WCDMA test model 1 signal. Center freq.=2.14 GHz.

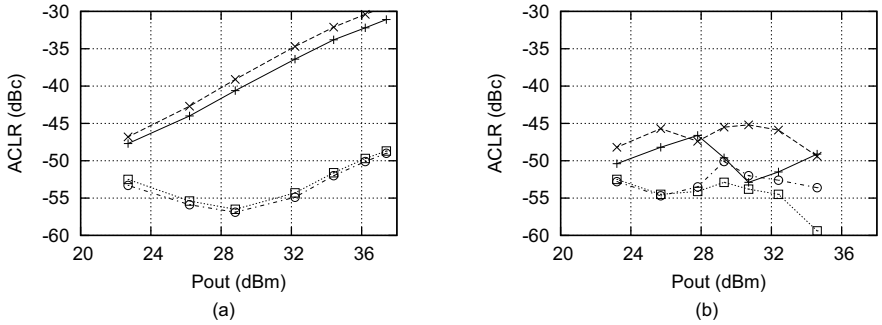


Figure 4.23: Measured ACLR of the GaN class-B PA driven by WCDMA signal test model 1 at various P_{avg} for (a) no predistortion (b) with predistortion. × ACLR1 low + ACLR1 high □ ACLR2 low ∘ ACLR2 high.

range of power levels with signals actually used in the UMTS band.

4.4 Conclusions

The goal of this chapter was to look into PD algorithms that can be used for the classical PAs, but also for highly efficient PA architectures that are the focus in the remaining chapters of this thesis. The two main categories where predistorters can

be classified in are:

1. **Memoryless predistorters.**
   When a non-linear PA is memoryless, the output signal is only a function of the currently applied input signal. This makes predistortion straightforward. Namely the predistorter function is simply the inverse of the non-linear transfer function of the non-linear PA itself. For most PAs, the memoryless situation is valid when the bandwidth of the modulated signal is small. This is equivalent to having a very slowly changing envelope (magnitude and phase) of the modulated signal with time. A simple memoryless predistortion algorithm was implemented that first requires the AM-AM and AM-PM characteristic of the non-linear PA to be characterized, this algorithm uses a LUT approach to implement the inverse function of the non-linear transfer of the PA. Simulation results show this straight forward implementation is effective when used with a realistic PA design. In the next Chapter, this memoryless predistortion algorithm will be utilized to predistort an adaptive PA prototype for future handset implementations.
2. **Predistorters with memory effect compensation.**
   When the bandwidth of the modulated signal becomes high (i.e. rapidly changing envelope of the modulated signal), the bias circuitry of the PA is often not fast enough to track the changes in the envelope and this shows up as distortion of the envelope (skewing of the envelope waveform). This effect reduces the effectiveness of the memoryless predistorter since the output is no longer only a function of the presented input at the same time slot but also from the history of this input signal. In this thesis work, a predistorter with memory compensation was developed which is based on memory polynomials. This concept has proven to be effective in simulations and further verified on a UMTS GaN class-B power amplifier. Using this predistortion algorithm with memory compensation, the PA can be used up to its peak output power (without signal clipping) and still meeting the stringent linearity specifications for UMTS base-stations.

With the linearization techniques developed in this chapter, the focus from here on (chapter 6 and 7) will be on highly efficient PA concepts, which require digital predistortion techniques to meet their spectral specifications. It will be this combination that set the trend for future amplifier implementations for use in third and fourth-generation communication applications.

Chapter 5

High Efficient Power Amplifier Architectures

In Chapter 3, it has been shown how improvements can be made in the design of Class-AB power amplifiers such that they can reach average efficiency numbers in the order of 40%, when operated with complex modulated signals(such as IS95). The next milestone set by industry, is amplifier efficiencies greater than 50% for 3G signals like WCDMA, WiMax, LTE etc. Note that due to the significantly higher peak-to-average power ratios (>8dB) of these signals, this is a very challenging task, even when assuming that the peak efficiency of the active devices can reach 100%. Note that such a number is for example theoretically feasible for ideal Class-F operation, however, in power back-off this efficiency number would typically degrade as,

$$\eta = \sqrt{p_{out}}\,\eta_{max} \tag{5.1}$$

In which p_{out} is the normalized output power varying between 0 and 1, with 1 being the maximum (saturated) output power. Note that this relation implies that for operation in 6dB back-off of the saturated output power (usually the 1dB compressed point), the efficiency drops to only half its maximum value.

One approach to overcome this limitation, is to use pulsewidth modulation (PWM) techniques at RF frequency to drive switch-mode power amplifiers (Class-D). This has the potential of having 100% efficiency even when amplifying signals with significant peak-to-average power ratios. This PWM RF amplifier architecture and technique will be briefly discussed in section 5.1. At this time of writing, the most common approach used in the industry to overcome this on first sight rather fundamental limitation, is to use the "classical" amplfier core (e.g. an class-B or class-F operated device) in a larger system where this efficiency-output-power roll-off behavior is modified. These so called high-efficiency power amplifier architectures, manipulate their efficiency-output power behavior in such a way that a significantly lower efficiency degradation versus back-off power is achieved. In section 5.2, we explain the basic principles how these amplifier architectures manage to keep their amplifying devices at their peak efficiency, even when operated in power back-off. In section 5.3 and

section 5.4, we introduce the most commonly used high efficiency RF power amplifier architectures of today.

5.1 Pulse width Modulation for RF Power Amplifiers

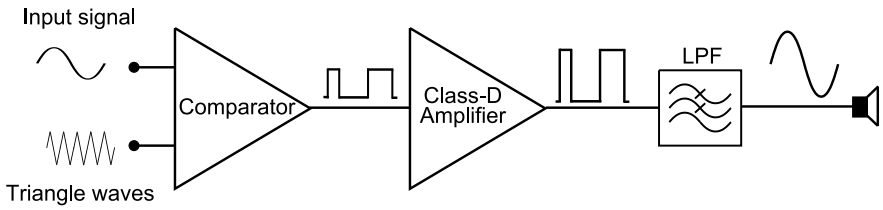


Figure 5.1: Typical Class-D audio amplifier

The concept of driving a switching class amplifier with pulsed width modulated (PWM) signals to boost efficiency performance is not new, and it is first used in audio amplifiers. Figure 5.1 shows a typical audio amplifier architecture that uses this setup. The input signal (in this particular case, a sinusoid) is converted to a sequence of pulses, where the pulse width and density of the pulses in time is a function of the amplitude and frequency of the audio signal being amplified. The resulting frequency of the pulses is usually greater than 5 times the highest frequency of interest of the input audio signal. These train of pulses are used to drive a switched-mode type power amplifier (typical Class-D), which is 100% efficient in theory. To retrieve the amplified version of the original audio signal, a low pass filter is required to remove the higher frequency components. This particular architecture of pulse width modulator and switched-mode power amplifier is also referred to as PWM amplifiers.

With current advancement in transistor technology, transistors can now switch reasonably at RF frequencies. This would raise the expectation that the PWM amplifier concept is also feasible at RF frequencies. However, to achieve good efficiency number the switching needs to be close to perfect something that is extremely difficult to achieve in practice due to the unavoidable device parasitics.

The operation principle of the PWM is as follows [24]. The average frequency of the pulses is centered at the center of the frequency band. By varying the position of the pulses with regards to their average position, the phase information of the base-band signal is coded. And similarly,the amplitude information of the base-band signal is coded by varying the widths of the pulses.

Figure 5.2 shows a very simplified block diagram schematic of a RF PWM amplifier. The challenges of this amplifier architecture is both in the implementation of the switched-mode amplifier [25] and the encoding of the base-band signals into pulse trains [26–28] which can be efficiently used with the switched-mode amplifier. The RF PWM amplifier though having the potential of having very high efficiency performance, suffers from reduced dynamic range, as the quantization technique employed will inevitably raise the noise floor. Secondly, the spectrum at the output of the

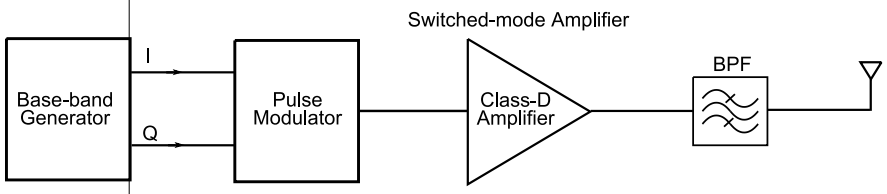
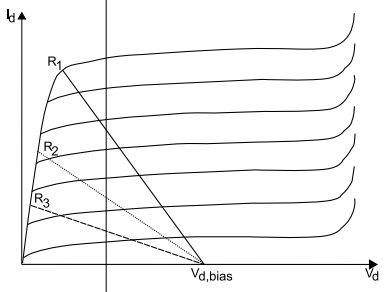


Figure 5.2: Simplified block diagram schematic of a RF PWM amplifier

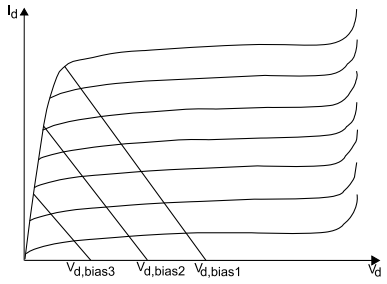
switched-mode amplifier has very strong harmonic contents, this puts very stringent requirements on the band-pass filter. Any loss in the band-pass filter will result in reduction of the overall efficiency of the PWM amplifier. Furthermore, the amount of linearity correctability of the PWM amplifier system with standard DPD algorithms (currently employed in the industry) remains doubtful. For these reasons, it is the author's opinion that the PWM amplifier will not be a good candidate for short to midterm time frame solutions to improve the efficiency of the RF power amplifiers required by 3G/4G applications. As such, these topics are not further investigated in this thesis, and the reader is referred to the cited references for details.

5.2 Basis of High-Efficiency Power Amplifier Architectures

Almost all the current high-efficiency RF power amplifier architectures are based around a classical amplifier core and employ either (1) load line modulation or (2) supply voltage modulation at the drain/collector to control the output power. Both methods have only one aim, namely to keep the output transistor at its maximum voltage swing, which results in the highest efficiency.



(a) Load presented to the device is changing as function of power, keeping the output stage at its maximum voltage swing.



(b) Bias at the drain is changing as function of output power, keeping the output stage at is maximum voltage swing.

Figure 5.3: Illustrates the principles behind the high-efficiency architectures which relies on supply tracking and load-line tracking.

For the first case, where the load-line is varied, the load which the transistor sees when delivering maximum power is:

$$R_{L,opt} = \frac{V_{d,bias}^2}{2 \cdot Pout_{max}} \tag{5.2}$$

In normal Class-B operation, when the power amplifier is backed-off 3dB from its maximum output power ($Pout_{max}/2$), given that the load remains the same, the voltage at the drain will decrease by a factor of the square root of 2 ($\frac{V_{d,bias}}{\sqrt{2}}$). To get the voltage at the drain/collector again back to its maximum voltage swing, which implies high efficiency, the impedance offered to the transistor needs to be doubled ($2 \cdot R_{L,opt}$). This is illustrated in Figure 5.3(a), which shows a typical plot of drain current versus drain voltage for various gate biases. R_1 corresponds to the load-line where the transistor is matched for maximum power out. When the transistor is backed-off in power, the load-line is changed to R_2 and in further back-off its changed to R_3 (where $R_3 > R_2 > R_1$). In theory, this principle is able to offer the peak efficiency performance (78.5% in the case for Class-B) at all output power levels. The limitation of this approach is the feasibility to create a load that approximates infinity when the desired output power level approaches zero.

The second approach to obtaining high-efficiency in back-off powers is to change the drain/collector bias voltage as a function of output power. This also results in keeping the output stage at its maximum voltage swing for all output power-levels. This is illustrated in Figure 5.3(b). Here the load offered to the transistor remains constant, but when the power is backed-off the drain bias is also decreased ($V_{d,bias1} > V_{d,bias2} > V_{d,bias3}$). In practice this will keep the efficiency of the output stage for a particular class of operation at its theoretical maximum.

5.3 Bias Supply Variation Based Power Amplifier Architectures

The Kahn power amplifier architecture, commonly also known as Envelope Elimination and Restoration (EER) is one technique that implements the bias supply variation concept to achieve high efficiency. Figure 5.4 shows a block diagram schematic of a typical Kahn architecture implementation.

The Kahn high-efficient architecture in its pure form consist of a high efficiency RF power amplifier cell, which linearity is of little concern (e.g. a switching mode type amplifier like Class-D, E or F). The bias supply of this RF amplifier stage is connected to the output of a video-bandwidth amplifier. The incoming RF signal is amplitude clipped (phase information is retained) and this clipped signal is use to drive the input of the RF power amplifier. Since the driving signal has a constant envelope, the RF power amplifier can be always operated in voltage saturation to maintain high efficiency at all times. To restore the desired envelope amplitude modulation, a part of the incoming RF signal is used to extract the envelope signal, which is fed to a high voltage video-bandwidth amplifier in the supply path of the RF-stage. Doing so will bring the original envelope information of the input signal on the supply voltage of

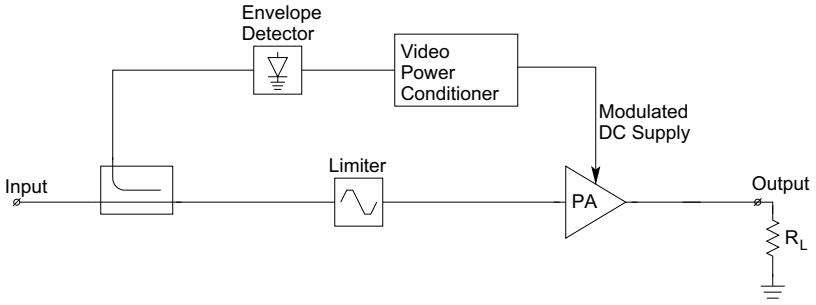


Figure 5.4: Typical EER architecture

the "saturated" RF stage. Consequently, the output signal of the RF-stage will now contain both the original phase, as well as, the "restored" amplitude modulation of the input signal, resulting in linear amplification from input-to-output of the complete system.

Actual amplifiers designed based on the Kahn architecture [29] has shown efficiency improvements over the classical "linear" Class-B amplifier.

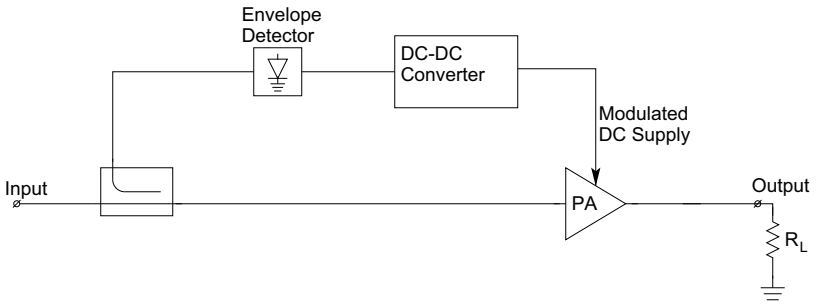


Figure 5.5: Typical ET architecture

The other high-efficient architecture based on bias supply variation is that of the Envelope Tracking Power Amplifier. The envelope tracking (ET) architecture is almost identical to that of the Kahn technique (see Figure 5.5). However, unlike the Kahn architecture, ET requires the use of a linear output stage for the RF power amplifier (e.g Class-AB is commonly used). This is because in ET, the RF path contains both amplitude and phase information (note that the Limiter is absent in the ET block schematic). This linear amplifier is operated efficiently by varying its drain bias as a function of the envelope power. This is done by detecting the envelope and using this information to control a DC-DC converter (typically high efficient buck-boost converters are used).

There are two bias supply tracking schemes in use today. In the first scheme, the bias supply tracks the long-term average output power of the power amplifier [30]. This is commonly referred to Average ET (AET), and it is found in handset power amplifiers, where due to power control schemes (CDMA systems), the average output

power can vary significantly as a function of proximity from the base-station.

In the second tracking scheme, the bias supply tracks the instantaneous output RF envelope. This is commonly referred to as Wide Bandwidth ET (WBET). Just as the Kahn technique, this later technique has the highest potential for enhanced efficiency when amplifying signals with a high CF. However, the difficulty for the WBET, as well for the Kahn technique is the design of the energy efficient wide-band DC-DC converter, which is basically a buck-boost converter combined with a high band-width voltage amplifier able to drive a very low impedance (output of the transistor). A task that becomes increasingly more difficult at higher bandwidths and power levels. Moreover, RF power transistors commercially available today are often optimized for performance at a single bias drain voltage, when varying the drain voltage such that it approaches zero for low output power levels, causes the transistor to be operated far from its peak performance, resulting in higher losses and lower efficiency. In spite of these challenges, successful demonstrators have been demonstrated with above 50% overall efficiency (power amplifier and DC-DC converter) for W-CDMA signals [31].

5.4 Load-line Variation Based Power Amplifier Architectures

As mentioned in the introduction of this Chapter, one way to keep the voltage swing of the amplifying output stage at is maximum is to adjust its output loading. To achieve this in practice, various amplifier concepts have been introduced over time. One of them is out-phasing, a technique that makes use of two parallel signal paths, whose input signals are adjusted in phase to achieve the desired signal vector summation to construct the output signal. This technique, which was originally introduced for improved linearity using an isolating output signal combiner, is commonly referred to as, "Linear Amplification using Nonlinear Components (LINC)". When aiming also for high efficiency, load modulation of the amplifying stages is mandatory. Therefore, the isolating signal combiner needs to be replaced by a loss-less non-isolating version. Doing so, yields the well known Chireix amplifier configuration [32]. Its first implementation at microwave frequencies was reported in the 1970s [33].

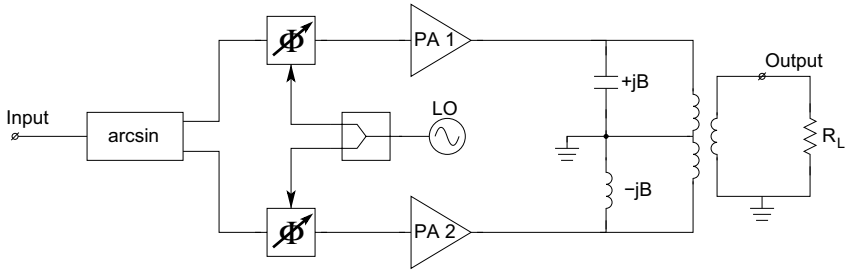


Figure 5.6: Typical Chireix/LINC architecture

Figure 5.6 shows a typical block diagram schematic of a Chireix/LINC high-efficiency power amplifier. The two power amplifier in the block diagram are classical

implemented (Class-B/-AB) and designed to always operate in voltage saturation. The incoming base-band signal is split into two separate constant envelope phase modulated signals, which are opposite in phase with respect to each other (therefore also the name outphasing amplifier). In modern day implementations, this is done completely in the digital domain with DSPs. The output of these two voltage saturated amplifiers are combined in the load, and this results in a linear amplification of the input signal. The output impedance seen by each of these amplifiers are complex and varies with the outputphasing phase. Here we see the effect of the load-line variation, which results in high efficiency as explained in Section 5.2. However, in this case, the reactive part of the load is not desired, as it will lead to degradation in efficiency. This is circumvented by placing shunt susceptance elements in the combiner to tune out the undesired reactive components. But because these reactive components vary with the outphasing phase, it is impossible to tune it out completely over the whole dynamic range of the signal. Therefore a careful design choice has to be made, which often corresponds to maximizing efficiency at the back-off power which corresponds to the CF of the signal used.

Figure 5.6 shows a typical block diagram schematic of a Chireix/LINC high efficiency power amplifier. The two power amplifiers in the block diagram are typically Class-B stages, which are preferably operated in voltage saturation. To achieve the proper signal summation, the incoming signal is split into two separate constant envelope signals with phase modulation. The offset phase between the channels depends on the desired output power level. Note that in modern implementations of outphasing amplifiers [34, 35] this is done completely in the digital domain. Due to the outphasing the output impedance seen by each of the two amplifiers is complex and varies with the offset phase. The variation in the real part gives us the desired load-line modulation, which results in high efficiency operation as explained in Section 5.2. However, the changing reactive part of the load is not desired, as it will lead to degradation in efficiency. This is circumvented by placing shunt susceptance elements in the combiner to tune out the undesired reactive components at a particular outphasing angle. However, since these reactive components varies with the outphasing angle, it is impossible to tune out the reactance completely over the whole dynamic range of the signal. Therefore a careful design choice has to be made, which often corresponds to maximizing the efficiency at a back-off power that corresponds to the CF of the signal under consideration [36].

Another high-efficiency architecture that relies on load-line variation is the Doherty amplifier. The Doherty amplifier has a long history as its operational principle was first demonstrated in 1936 by W.H. Doherty [37].

Figure 5.7 shows a block schematic of the most common implementation of the Doherty architecture. Typically it consists of two “classical” amplifier cells, namely: the main amplifier, which is mostly operated in Class-AB and the peaking amplifier, which is typically operated in Class-C. The output powers of these two amplifier cells are “added” through a quarter-wavelength line power combining network. In this configuration, the main amplifier acts like a normal Class-AB cell until it goes into voltage saturation at a specified back-off level from the maximum power out of the total Doherty architecture. Above this power level, the peak amplifier is also activated by the higher amplitude of the input signal and starts to drive current into the power

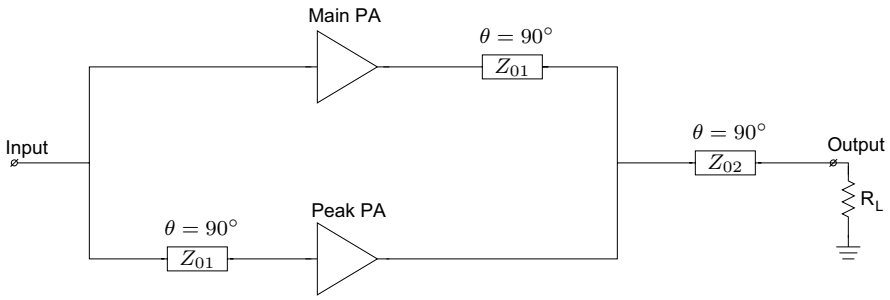


Figure 5.7: Typical Doherty architecture

combining network. This power injection will cause an effective load-line modification for the main amplifier. As result, the main amplifier remains in voltage saturation, while it is able to source more power to the load over the range that the peaking amplifier is activated. The overall effect is that the efficiency of the total Doherty amplifier peaks when the main amplifier reaches voltage saturation (when the peak amplifier is still off) and at the point where both the main and the peak amplifier reach voltage saturation (maximum output power).

In this thesis work the Doherty amplifier architecture will further analyzed and improved in Chapter 7.

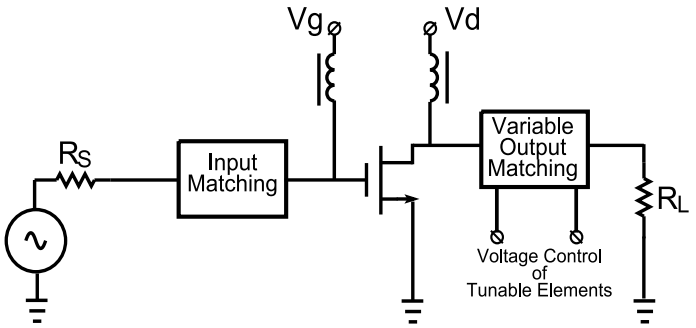


Figure 5.8: Typical Direct Load-line modulation architecture

On first sight the simplest dynamic load-line based high-efficient amplifier architecture is that of the direct load-line modulation architecture as shown in Figure 5.8. Here the basic concept is to use a tunable matching network to adjust the loading conditions of the amplifier cell. Note that such a tunable network, which operation is typically based on tunable elements like varactors, can change the load offered to the output stage transistor, as a function of (1) the long term average power (similar to the case of ET), which is very useful for handset applications in CDMA power control loops; (2) the envelope power,resulting in voltage saturation and thus high efficiency over a large dynamic range.

These tunable network elements need to be controlled externally by a signal processor. Although simple in concept, practical implementations proves to be very chal-

lenging. This due to very high requirements on the tunable components, which need to be lossless, add no distortion to the system, and should be able to be switched/varied with the speed of the RF envelope. In [38] a dynamic load-line amplifier with tunable matching networks has been implemented and demonstrated for signals with CF approximately 10dB, a two-fold improvement in efficiency over a conventional Class-B amplifier.

In summary, although many high efficiency amplifier concepts have been proposed, at the time of writing. Each concept seems to have its own unique advantages and disadvantages, making one concept more, or in some cases less applicable to requirements of a specific application. In short, handset applications seem to benefit most of the direct Load-line modulation concept. Since this concept can in principle also provide the selection of different frequency bands, antenna mismatch correction, etc., while the power levels involved remain limited, something that significantly relaxes the requirements of the tunable components. Base station applications, however, seems to benefit most of ET (high RF bandwidth), Chireix and Doherty (high modulation bandwidth) based amplifier concepts. Here the very high power levels often complicate the use of tunable network elements as varactors. In the following Chapters, we aim to evaluate and develop load modulation based amplifier concepts for improved efficiency in power back-off, in addition, we also give attention to the RF reconfigurabiltiy when dealing with amplifiers for handset applications.

Chapter 6

PA with adaptive matching

Today, cellular phones supporting multiple communication standards such as GSM, EDGE and WCDMDA are already commercially available. To enable this multi-band, multi-mode operation, current handset implementations are using parallel line-ups for the transmit and receive paths in combination with antenna duplexer and various switches as shown in Figure 6.1 a to meet the specific requirements of each standard.

However, with the increasing user demand to add more and more functionality like GPS, WiFi, WiMax, TV on Mobile etc. manufacturers of handsets are confronted with increasing costs, complexity and form factor, while customers require cost and size reductions. With the architecture of Figure 6.1 a, the hardware complexity increases linearly with the number of frequency bands that need to be supported. A proposed architecture that will support the multi-band multi-mode operation and does not introduce significant increase in cost and area is shown in Figure 6.1 b. Although conceptually simple, practical design considerations place severe constraints on the technology implementation to realize the required adaptive circuit blocks. For the receive path most of the circuit functions have already been demonstrated [39–41]. Major bottlenecks remain however in the implementation of the tunable filters and adaptive power amplifiers [42, 43].

Besides the size reduction of the multi-band multi-mode Power Amplifier, there is also the desire to improve efficiency and amplifier operation under antenna mismatch conditions. Current amplifier implementations already include power control loops to reduce output power when operating in close proximity of a base-station. In this mode of operation the transmit signal is backed-off from its maximum, yielding a significantly decreased efficiency (see also Chapter 2). Increasing the loading of the output stage can reduce this loss in efficiency under power backed-off operation, something that in principle can be combined with the required tunability in output matching when aiming for multi-band operation.

To address these upcoming demands on the handset amplifier, this thesis worjk has focused on implementing an adaptive handset PA using tunable matching networks. This PA should have also improved efficiency at all output power levels, to combine the needs of multi-mode and multi-band operation. Key elements in this amplifier design are the adaptive matching networks, which are discussed in Section 6.1.

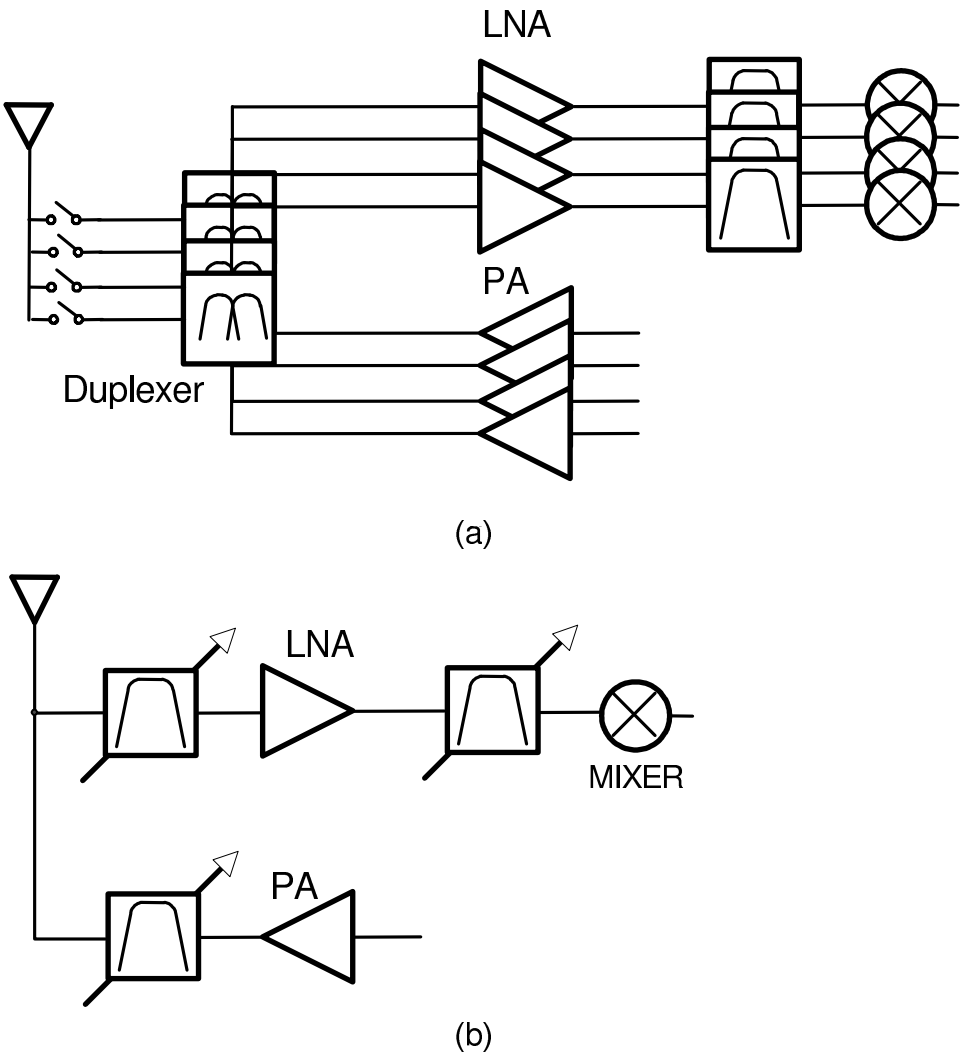


Figure 6.1: (a) Traditional parallel path multi-band multi-mode approach based on static RF circuit blocks. (b) Simplified next-generation transceiver concept using adaptive RF function blocks.

Various tunable matching network topologies are considered and evaluated for their performance. Section 6.2 shows how these matching networks can provide improved efficiency in power back-off operation and facilitate frequency band switching.

6.1 Reconfigurable/Adaptive matching networks

In order to effectively change the loading impedance of the output stage as a function of the output power level and frequency of operation, tunable matching networks are required. In this section, the basic requirements that these networks must satisfy to obtain the desired PA functionality in terms of power control range and frequency tuning are discussed. Following that, the design of the matching network within the technology constraints is given and the resulting implementation is verified by measurement.

6.1.1 Design Goal of Adaptive Matching Networks

For the implementation of the tunable elements, DIMES silicon-on-glass Schottky diodes are used [44–48]. To achieve high linearity, these uniformly doped devices are placed in anti-series configuration [49] with a high impedance connection to establish the biasing of the center node. Compared to other RF tunable components, these elements provide superior linearity, reliability, tuning speed and quality factor ($Q >$ 100 @ 2 GHz).

In the time frame of this work the varactor devices in the silicon-on-glass process are characterized by the following design constraints:

- effective capacitance tuning range of the varactor stack. $\frac{C_{max}}{C_{min}} < 2.5$.
- control voltages < 18 V (to avoid diode breakdown).
- no forward biasing of any of the diodes by the RF signal.

Although in principle, control voltages up to 18 V are required to obtain the full capacitance tuning range for this technology version, it is noted that the varactor diodes are always operated in reverse bias and therefore require practically no DC current. Consequently, the generation of these control voltages by DC-to-DC up-conversion is relatively straightforward without any severe requirement on the conversion efficiency. The actual value of the control voltages depends on the intended capacitance control range and the RF voltage swing present in the matching network. By using a low impedance matching network solution, one can improve the power handling of the tunable matching networks. With the constraints above, the following specifications for the reconfigurable matching network are set:

- tunable impedance transformation ratio 10.
- continuously tunable with high speed (20 MHz bandwidth).
- operating frequencies of 900, 1800, 1900, and 2100 MHz (covering the most important bands for handsets).
- losses of less than 1 dB for the whole input impedance tuning range.
- number of independent control voltages ≤ 2.
- high linearity (≥ 55 dBc at 1 W output power level).
- output power range from 0.1 to 1W, while assuming a 3V supply voltage.
- improved amplifier efficiency for power back-off conditions.

In order to meet these requirements, the following subsection investigate the most promising network topology, and motivates the final choice made in this work.

6.1.2 Choice of Adaptive Matching Network Topology

Since low-loss matching network implementations with only two control voltages are prefered, the number of reactive elements in the following analysis is limited to a maximum of four. The networks considered in this analysis are given in Figure 6.2. Variable inductors are composed of an inductor/varactor combination with a net positive reactance.

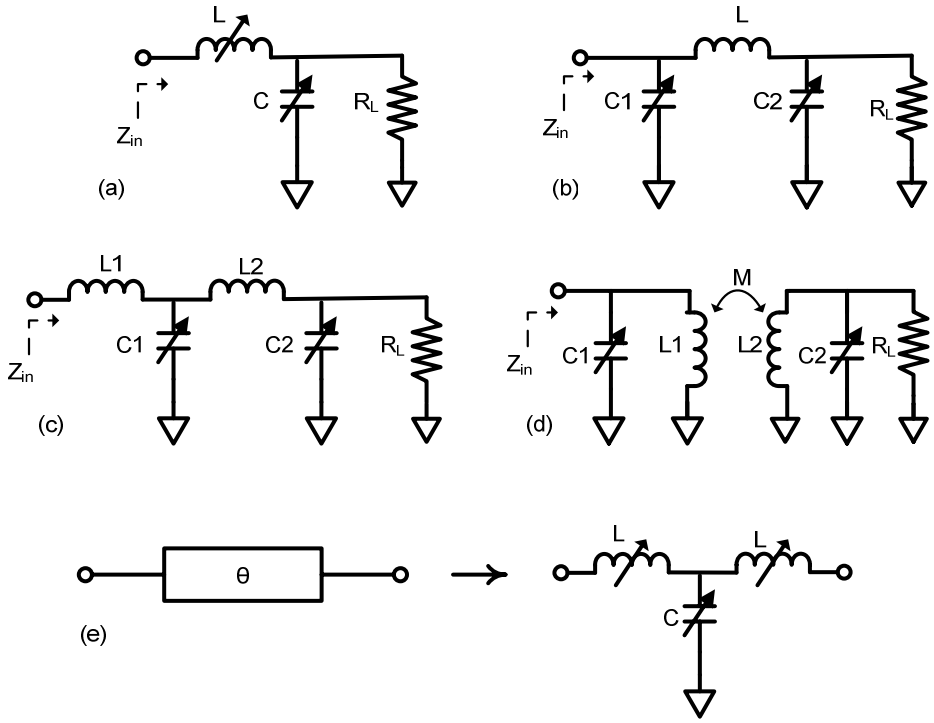


Figure 6.2: Network topologies considered for the implementation of the adaptive matching network. (a) L-type network. (b) π-type network. (c) Two section L-type network. (d) Transformer coupled network. (e) Constant delay network.

Of these networks, the L-type network is the simplest. In principle this configuration can provide impedance transformation up to the loading impedance of the network (50Ω). Unfortunately, to obtain an acceptable impedance transformation range, it requires a very large tuning range for its component values. A π-type matching network has one more element than a L-type network, which in principle allows a load termination to be transformed to any position within the Smith-chart. When using two L-type matching networks in cascade, we obtain the so called two-stage ladder matching network. As with the π-type matching network, this topology allows an impedance transformation that covers the entire Smith chart. The advantage of the ladder network over the π-type network is that, for higher impedance transformation ratios, the Q of the two-stage network is significantly lower. Note that high Q condi-

tions in the matching network give rise to increased losses and higher voltage swings, limiting the power handling capabilities of the resulting network.

When implementing a transformer or coupled line-based matching network, one suffers from the trade-off between the impedance tuning range and coupling factor of the coils. As a result, high Q conditions are required to obtain the intended tuning range, yielding higher losses. All these matching networks yield large S_{21} phase variations (10 degrees over a 5 to 50 Ω impedance range) when changing the impedance transformation by adjusting the network elements. When changed dynamically [38, 50], this phase variation appears as an AM-to-PM distortion and degrades the linearity of a dynamic load modulated amplifier when no compensating measures like digital predistortion are taken. If one wants to avoid the use of predistortion in this application, one can consider network solutions that provide no phase variation when changing the impedance ratio. A transmission $\frac{\lambda}{4}$ transformer is such a network (see Figure 6.2 e, which, for a limited bandwidth, can be approximated by its lumped equivalent. Consequently, by changing both the L and C simultaneously, one can vary the characteristic impedance while keeping the phase delay constant. This approach requires tunable inductors, which in principle can be implemented using a combination of an inductor and a capacitor. However, in practice this will result in increased losses making this topology less favorable compared to the other suggested matching network solutions.

The results of the above analysis are summarized in Table 6.1. Based on this table, the discussion above and the fact that the priority here is for static adaptive matching, the two-stage ladder network is the most appropriate topology. In the following section, the implementation details of this matching network and its measurement performance is given.

6.1.3 Implementation and Measurements of the adaptive matching network

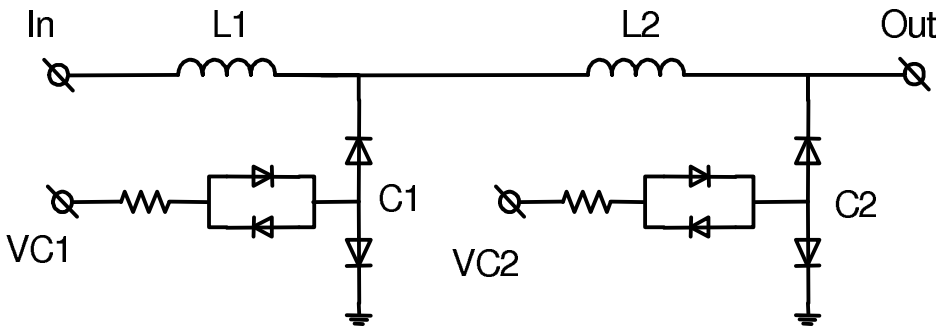


Figure 6.3: Schematic diagram of the two-stage ladder network.

The schematic diagram of the two-stage ladder network is shown in Figure 6.3. The tunable capacitors are composed of an anti-series configuration of two varactors,

Table 6.1: Evaluation of matching networks.

Types of matching network	Impedance ratio >10	Q at high transfer ratio	Complexity	Compactness
L	No	–	Low	Good
π	Yes	Low	Low	Good
Two stage ladder	Yes	Lowest	Low	Good
Coupled line	Yes	Very High	High	Moderate
Constant Phase Shift	Yes	Low	Very High	Bad

which form the so-called varactor stack (VS). A high value resistor and two diodes in anti-parallel configuration have been used to realize sufficiently high impedance for the varactor stack center tap to avoid linearity degradation for narrow tone spacing [47]. Each varactor stack is independently controlled for its effective capacitance through its center-tap voltage. The inductors can be realized either by coplanar waveguide [48] or bond-wires; all reported matching networks are implemented using silicon-on-glass technology [46].

In the experiment, bond-wires are used for the implementation of the inductors to facilitate tuning of the integrated matching network and simplify also for the connection to the active device. The values of the varactors and the inductances of the bondwires were chosen to achieve a large impedance control range at 1900 MHz and to obtain the lowest loss condition for the highest output power. The final bond-wire dimensions were verified using a three-dimensional (3-D) field simulator. The layout and components values are given in Figure 6.4 and Table 6.2, respectively.

The S-parameters of the matching network were measured using a HP8510C network analyzer. Figure 6.5 shows the measured S_{11} of the integrated matching networks at 900, 1800, 1900, and 2100 MHz; this measurement illustrates the band-switching and impedance transformation capabilities of the implemented adaptive matching network loaded with 50Ω at its output. The data of Fig 6.5 is based on the on-wafer measured S parameters of the device shown in Fig 6.4. With this, the S_{11} data at the reference plane of the output of the active device is obtained. From Figure 6.5, note that the coverage in the Smith chart is very large for the higher

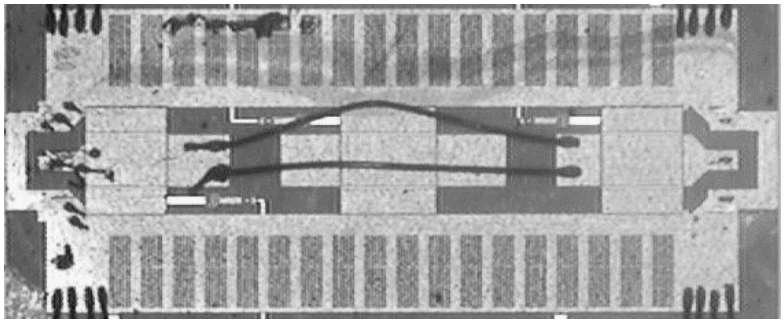


Figure 6.4: Micro photograph of matching network using bondwires to realize the inductor.

Table 6.2: Element values of output matching network

Element	Values
L_1	Length≈ $600\mu m$, radius wire≈ $15\mu m$, Height≈ $350\mu m$, $L \approx 0.3nH$
C_{jo_1}	36 pF(zero bias)
L_2	Length≈ $1870\mu m$, radius wire≈ $15\mu m$, Height≈ $800\mu m$, $L \approx 1.2nH$
C_{jo_2}	36 pF (zero bias)

frequency bands, providing a resistive impedance transformation range much larger than 10. Although the control range for the 900 MHz band is very limited, it is still possible to offer a matching condition to the active device for the highest intended output power (28 dBm). Note that the 28 dBm optimum loading condition indicated on the Smith Chart assumes a lossless matching network. To compensate for the losses of the actual network, a slightly lower impedance has to be offered in practice.

One of the most important aspects of tuner design is to maximize the power gain ($\frac{P_{out}}{P_{in}}$) or minimize the losses of the structure. The definition of this gain is:

$$G_p(dB) = 10\,log_{10}\left(\frac{P_{out}}{P_{in}}\right) = 10\,log_{10}\left(\frac{|S_{21}|^2}{1-|\Gamma|^2}\right) \tag{6.1}$$

Note that G_p is a more appropriate measure than G_{max}, since G_{max} assumes

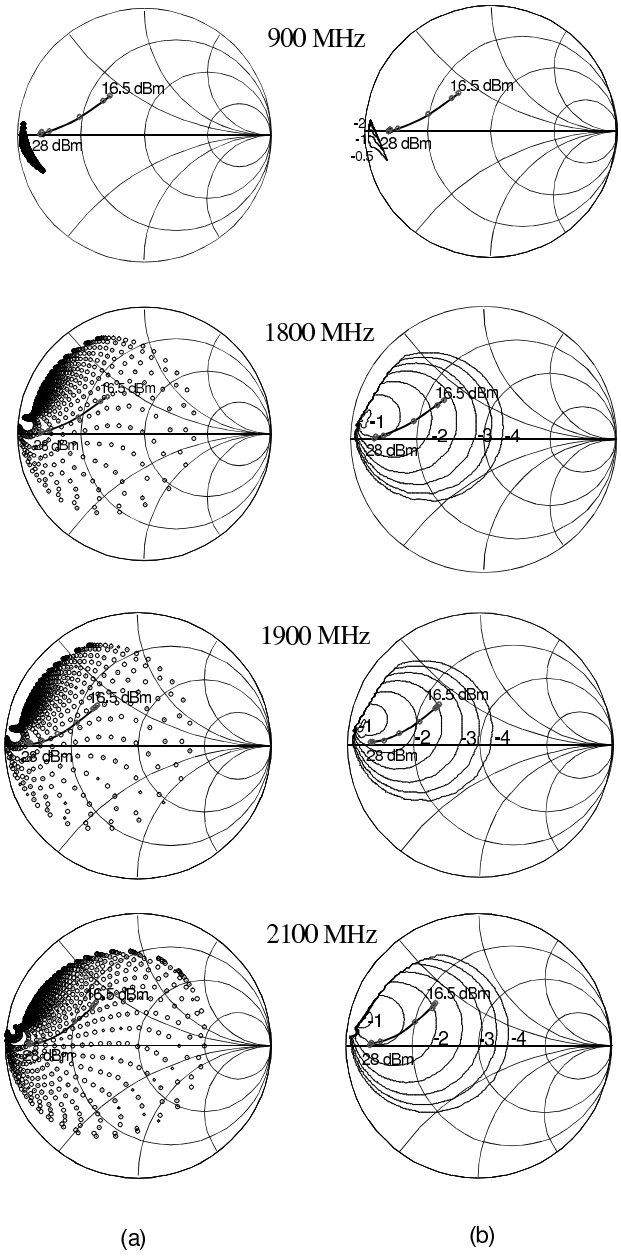


Figure 6.5: (a) Measured s11 and (b) contours of constant Gp of the adaptive output matching network at 900, 1800, 1900 and 2100 MHz, the simulated optimum loading trajectories for the highest efficiency are indicated. Vcontrol < 18V. Required control voltage for the optimum load trajectory < 10V.

conjugate matching conditions at both input and output. (G_{max} was approximately −0.3 dB for all tuning values). Figure 6.5 also plots the measured loss contours up to −4 dB. The losses of these matching networks are caused by the limited Q of the inductors and the parasitic resistance of the varactor stacks. The Q of the matching network was optimized to be low for the high power condition in order to avoid forward biasing or voltage breakdown conditions of the varactors. As a consequence, the highest matching conditions are found for the lowest impedance ratio transformations, yielding higher losses in these situations.

In Figure 6.5, the simulated optimum load trajectories as a function of output power, at the device reference plane, for maximum efficiency of the QUBIC device (which will be used in the full PA design) are superimposed on the Smith Chart. The expected gain and efficiency for the optimum load trajectory at 1800 MHz, assuming a lossless matching network is shown in Figure 6.6. Note the significantly higher efficiencies achievable over a large power control range with the configurable load. The slight drop in efficiency at very high power back-off conditions can be partly explained by the remaining constant (RF power independent) part of the DC power consumption in Class-AB operation. Also, the impact of device parasitics (the output resistance) becomes increasing dominant in these large back-off conditions.

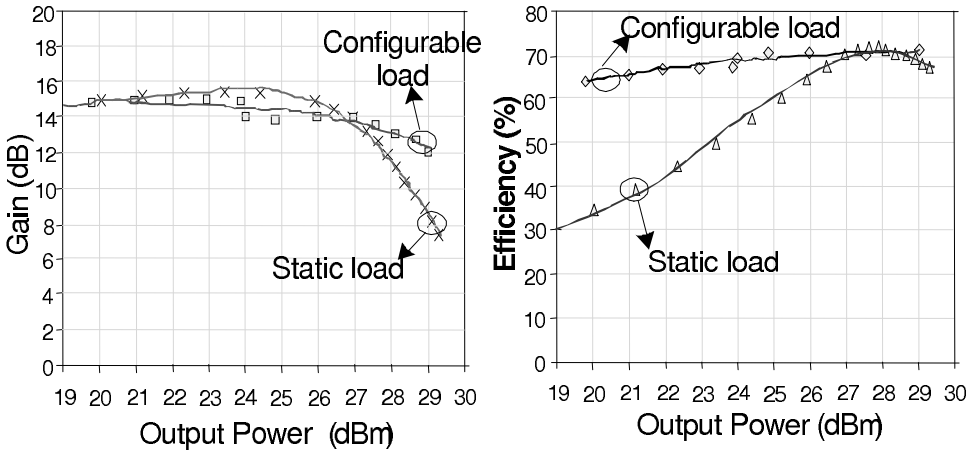


Figure 6.6: Simulated (a) gain and (b) efficiency of the QUBIC device for the optimum loading trajectory of 1800 MHz.

The large-signal performance of the varactor-based matching networks has been reported in [48] and the IM3 was better than 50 dBc (f_c = 2GHz Δf = 20MHz) with an output power control range of 17 to 27 dBm. Note that this performance meets the constraints of most communication standards for handsets. No forward biasing occurs for either of the diodes for RF powers up to 1 W. Simulations to investigate the effect of noise on the control voltages of the varactor show that a 10-mV sinusoid perturbation of 100 kHz on the control lines gives rise to spectral regrowth of -66 dBc relative to the carrier, which is sufficiently low for current communication standards.

These results demonstrate that the reconfigurable matching networks can provide the necessary impedance transformation when inserted into the PA module, with almost no degradation of the linearity due to modulation of the varactors or voltage fluctuation on the control lines.

6.2 Design of a Power and Frequency Adaptive PA

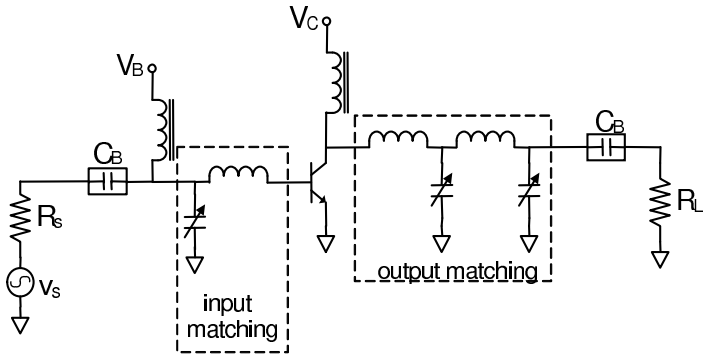


Figure 6.7: Schematic of the adaptive power amplifier.

Figure 6.7 shows the schematic of the planned test circuit which consists of an active device (2-W SiGe HBT transistor with an f_t of 50 GHz [51]) between two tunable integrated input and output matching networks. This configuration enables the verification of the use of an adaptive matching network to perform static load-line modulation so as to improve the collector efficiency, and also to achieve frequency-band switching.

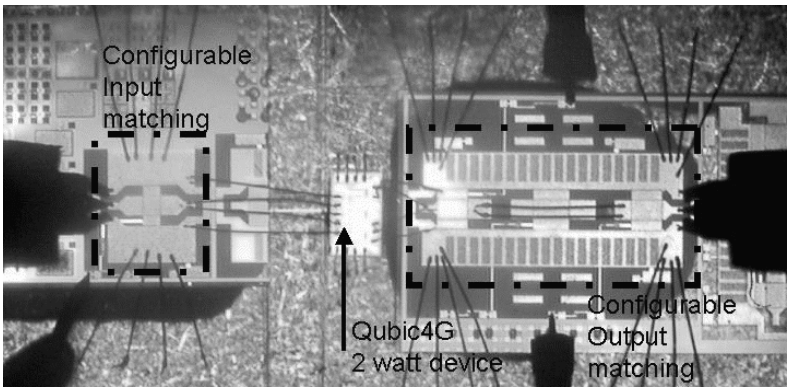


Figure 6.8: Power amplifier with reconfigurable input and output matching network.

At the input, a tunable L-type matching configuration, consisting of a shunt varactor and a series bond wire inductor, is used as pre-match for the active device.

Figure 6.8 shows a micro photograph of the final implemented PA module. The total chip area (active device + integrated matching networks) is 8 mm^2.

From Section 6.1.3, it was shown that the output matching network exhibits variable losses as a function of impedance. To estimate the degradation in the power efficiency due to losses in the output matching network, simulation is done with a ideal Class-B device along with the measured S-parameter values of the output matching network. Figure 6.9 a plots the simulated efficiency for different output loading conditions as a function of output power for a lossless matching network. In Figure 6.9 b, this simulation has been repeated, but now using the measured losses of the adaptive output-matching network. As a consequence of the losses in the output matching network, the efficiency decrease is approximately 12% (from 80% to 68%) at high output power levels (30 dBm) and 30% (from 76% to 46%) at 10 dB back-off (20 dBm). Despite this, the efficiency is still much better at the 10 dB back-off point than the same amplifier implemented with a fixed output matching network.

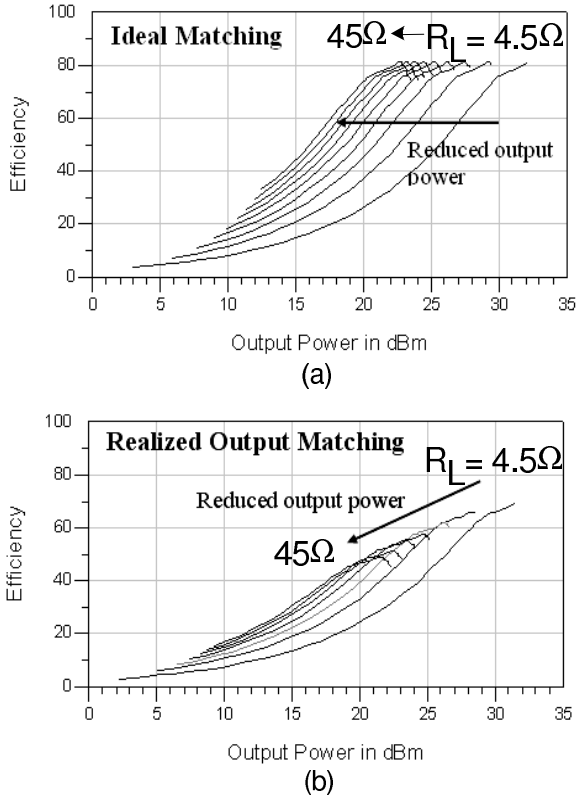


Figure 6.9: Simulated power efficiency of an ideal Class-B amplifier with (a) ideal matching network and (b) using the measured S-parameters of the fabricated matching network.

Although the optimum loading condition for the active device and input drive

power level can be found through simulation, it is preferred to use an experimental approach to find the optimum control voltage for the varactors. Note that the best loading conditions are those that yield the highest efficiency for a given output power level. In the following experiments, pulsed conditions for the biasing of the active device are used to avoid thermal problems. The schematic of the setup is given in Figure 6.10.

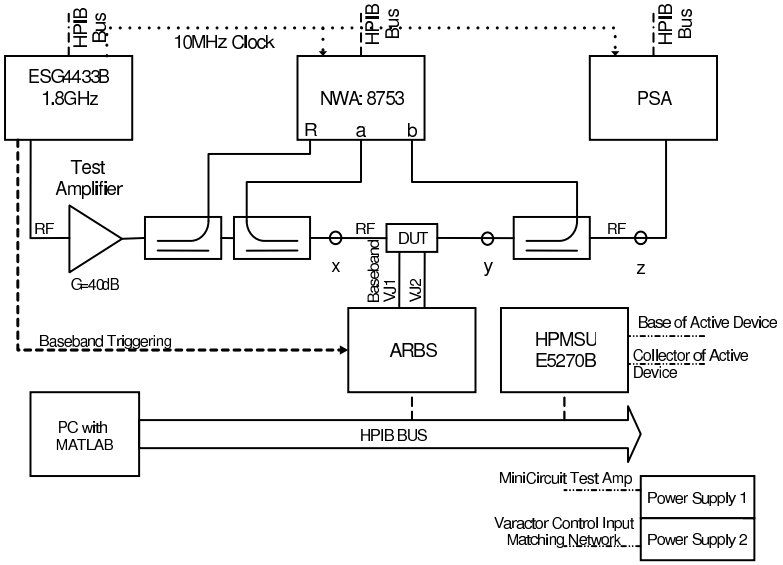


Figure 6.10: Schematic of the test-bench for the adaptive PA.

Figure 6.11 plots the measured gain versus output power for the 900, 1800, 1900, and 2100 MHz bands when the PA matching networks are fixed to the maximum output power settings at each frequency.

Figure 6.12 plots the measured efficiency of the adaptive amplifier for the 900, 1800, 1900, and 2100 MHz bands. Two cases are described. First, the amplifier matching networks are fixed to their maximum output power settings at each frequency band (this result in the classical Class-AB efficiency roll-off with back-off power). Second, for the 1800, 1900, and 2100 MHz bands, the efficiency is optimized for each P_{out} level, by providing the optimum combination of V_{C1}, V_{C2} and P_{in} (Note that for the 900 MHz band, this freedom is lost, because of the limited impedance control (Figure 6.5).

The improvement in efficiency over the classical Class-AB amplifier is clear and is typically more then 15%, resulting in an effective doubling of the efficiency when compared to the PA in power back-off using a fixed load, with minimal differences in the gain of the PA. As a result of this optimization, 30%–55% efficiency is achieved over a 10 dB range for the 1800, 1900, and 2100 MHz bands. Note that at 10 dB back-off, this efficiency exceeds the theoretical upper limit for Class-B operation (24.8%). When considering the PAE, the gain should be taken into account, so in this

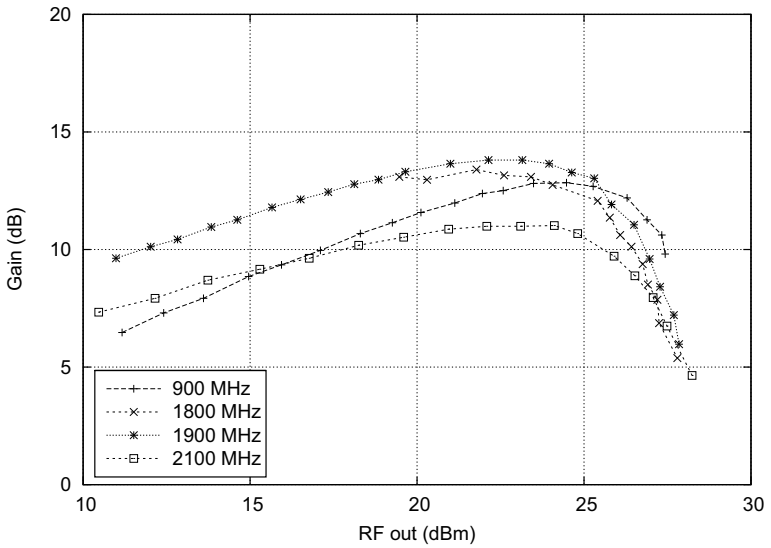


Figure 6.11: Measured gain versus output power at 900, 1800, 1900, and 2100 MHz when the PA matching networks are fixed to their maximum output power settings for each band.

case a drop in gain will affect its PAE, however, also for this parameter significant improvements are found in power back-off using the static load-line adjustments.

The varactor-based matching networks have been independently characterized for their linearity using a two-tone signal, yielding IM3 levels below -50 dBc for output powers up to 27 dBm output power (f_c = 20GHz, Δf = 20MHz) [48]. Consequently, the intermodulation products at the output of the adaptive PA is caused by the AM-AM and AM-PM distortion of the active device itself, in combination with secondary mixing products due to the non optimized second harmonic impedances at the input and output of the active device. Although the amplifier circuit itself was not optimized for linearity, two-tone testing (f_c = 900MHz, Δf = 6.67KHz) was performed on the complete amplifier close to its maximum power ratings (Figure 6.13), which is a worst-case operating condition for linearity and is thus a good gauge of the overall linearity performance.

Note that the power of the fundamental tones is 22 dBm, yielding 25 dBm of average output power, which represents the maximum power level for unclipped two-tone operation (3 dB headroom, since P_{peak} = 28dBm from Figure 6.12). As expected, the linearity for this non optimized amplifier is not impressive (IM3= −28dBc) owing to the poor AM-AM and AM-PM characteristics of the device itself. However, using memoryless predistortion which was developed in Chapter 5, the IM3 level is improved to 35dBc, which is a typical level accepted by handset applications. Note that in handsets the memoryless predistorter is the de facto implementation, as the use of DPDs with memory compensation capabilities is way too costly, furthermore adjacent channel leakage specifications are more relaxed in handsets.

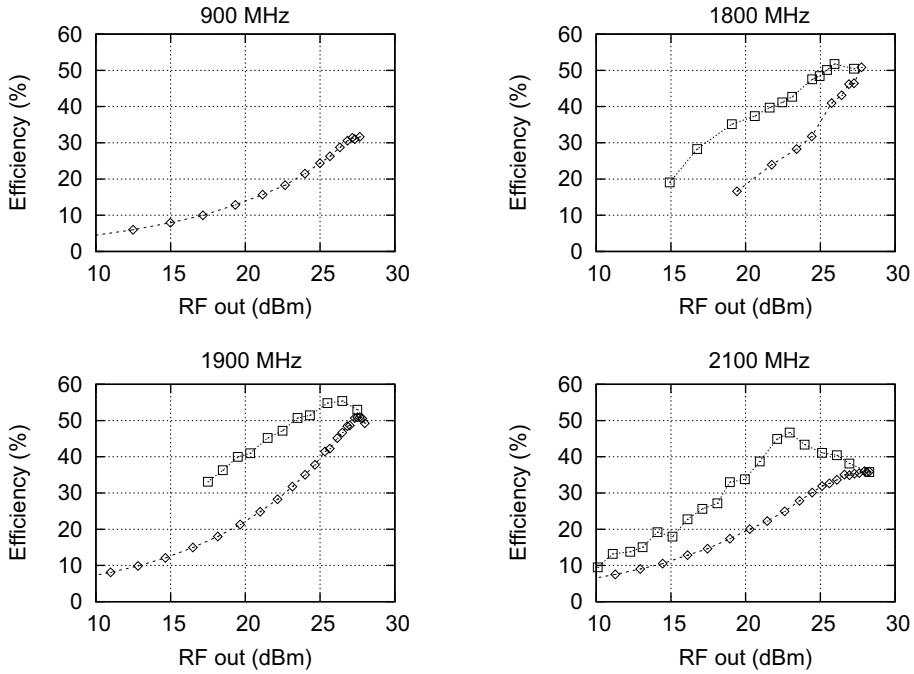


Figure 6.12: Measured collector efficiency versus output power at 900, 1800, 1900 and 2100 MHz. First the adaptive matching networks are fixed to the optimum maximum output power condition. Second the output loading is optimized for efficiency at each power level. □ adaptive matching. ◇ no adaptive match.

6.3 Dynamic Load-line PA

The adaptive matching network shown in Figure 6.4 realized with tunable varactors is capable of impedance adaptation at very high speeds of up to 20 MHz, therefore this makes it the ideal candidate to track the output stage loading conditions as a function of the envelope of the modulated signal (the envelope signal of a 3G signal is typically within this 20 MHz bandwidths). This approach, known as dynamic load-line modulation [38], enables the PA to operate at its maximum efficiency at all times, yielding considerable efficiency improvements especially for modulated signals with high crest-factors.

Though the available adaptive matching network from Section 6.2 is not designed for dynamic load line tracking, it is still possible to demonstrate the possible improvements in efficiency using this existing hardware; but some problems with linearity due to the phase distortion properties of this network will be encountered, as will be made clear in the following subsections. Section 6.3.1 starts by giving an analysis of the dynamic load-line PA. In this analysis, the load the device should see as a function of the envelope of the transmit signal is derived. Following that, Section 6.3.1 shows how complications arise, in the sense of added distortion, when a practical adap-

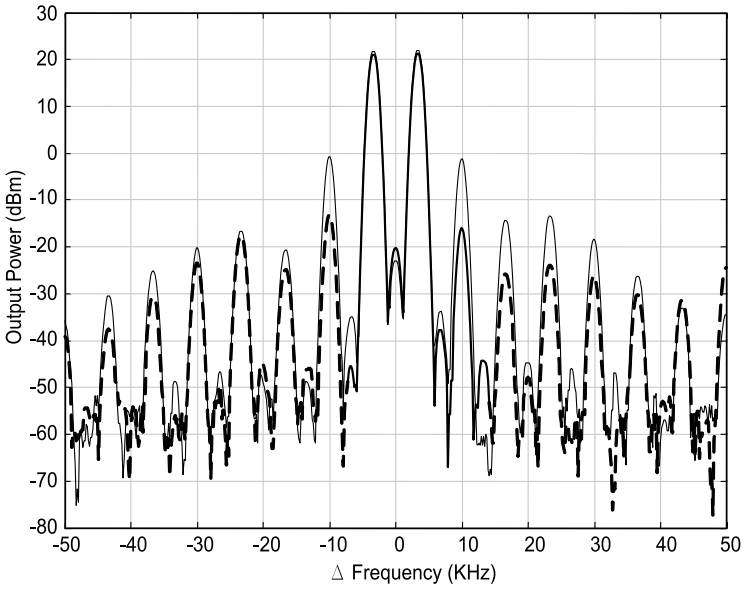


Figure 6.13: Measured two-tone linearity of the complete amplifier close to its maximum output power (f_c=900MHz, Δf=6.67KHz) solid line: without predistortion, broken line: with memoryless predistortion.

tive matching network is used. Lastly in Section 6.3.3, simulations and measurement results demonstrate the efficiency improvement by this dynamic load-line PA concept.

6.3.1 Dynamic Load-line Analysis

A RF modulated signal can be described by a time-varying envelope signal (commonly referred to as a baseband signal) which modulates a RF carrier. Equation 6.2 shows a signal at the input of the PA, where $\epsilon_{in}(t)$ and $\Phi_{in}(t)$ is respectively the magnitude and phase of the complex base-band signal at the input.

$$v_{in}(t) = \epsilon_{in}(t)\ \cos(\omega_{RF}t + \Phi_{in}(t)) \tag{6.2}$$

In the case of a linear amplification, the signal at the output should be a constant multiple of the input, thus:

$$v_{out}(t) = \epsilon_{out}(t)\ \cos(\omega_{RF}t + \Phi_{out}(t)) \tag{6.3}$$

Where $\frac{\epsilon_{out}(t)}{\epsilon_{in}(t)}$ and $\Phi_{out}(t) - \Phi_{in}(t)$ are constant values for all values of t such that no amplitude or phase distortion occurs.

As was discussed in chapter 2, the device operates under its maximum efficiency when the voltage swing at the collector or drain is saturated to the supply voltage. Therefore it is no surprise that the desired signal (for high efficiency operation) should

be a constant envelope signal, while phase modulation is permitted, as shown in the following:

$$v_{dev}(t) = V_{cc}, \cos(\omega_{RF} t + \Phi_{dev}(t)) \tag{6.4}$$

Where $v_{dev}(t)$ is the voltage at the output of the active device, and V_{cc} is the DC bias voltage provided to the output of the device. Assuming that the adaptive matching network is lossless and that the PA behaves quasi-static within one RF cycle (this assumption is valid when the RF frequency is much higher than the envelope or base-band frequency); the following equation can be written:

$$\int_0^{T_{rf}} \frac{V_{cc}^2 \cos(\omega_{RF} t + \Phi_{out}(t))^2}{R_{var}} \mathrm{d}t = \int_0^{T_{rf}} \frac{\epsilon_{out}^2 \cos(\omega_{RF} t + \Phi_{out}(t))^2}{R_L} \mathrm{d}t \tag{6.5}$$

Which states that for a lossless adaptive matching network, the power delivered by the active device to the input of the load varying network (R_{var}) is equal to the power delivered at the output into the fixed load connected at the output of the adaptive matching network (typically the antenna). Note that the integration is done over one RF period (T_{rf}).

Rearranging this equation (6.5) in terms of R_{var}, the following equation is obtained, which gives the loading conditions which the adaptive matching network needs to offer to the output stage device as a function of the envelope signal.

$$R_{var} = \frac{V_{cc}^2 R_L}{\epsilon_{out}^2} \tag{6.6}$$

This concludes that the impedance offered to the active device should change inversely with the square of the envelope. From the equation (6.4) and (6.6), the current sourced by the active device can be computed as follows:

$$i_{dev} = \frac{\epsilon_{out}(t)^2 \cos(\omega_{rf} t + \Phi_{dev}(t))}{V_{cc} R_L} \tag{6.7}$$

Assuming that the active device is a voltage controlled current source, this would imply that the driving signal at the input of the active device is a constant scalar multiple of it. Figure 6.14 shows the generic implementation of a dynamic load-line PA based on the above developed equations.

The envelope squarer at the input (in the ideal case) only squares the envelope of the RF signal as this information is not only required by the controller of the adaptive matching network, but also by the active device itself. Under these conditions, the voltage at the output of the active device is a constant envelope making it highly efficient. While the resulting signal at the output of the lossless adaptive matching network offered to R_L is a linear amplification of the input signal.

6.3.2 Limitations of the adaptive matching network

In a classical PA, the active device is the only source of non-linearity in the PA. However, in the dynamic load-line PA, not only does the active device generates non-

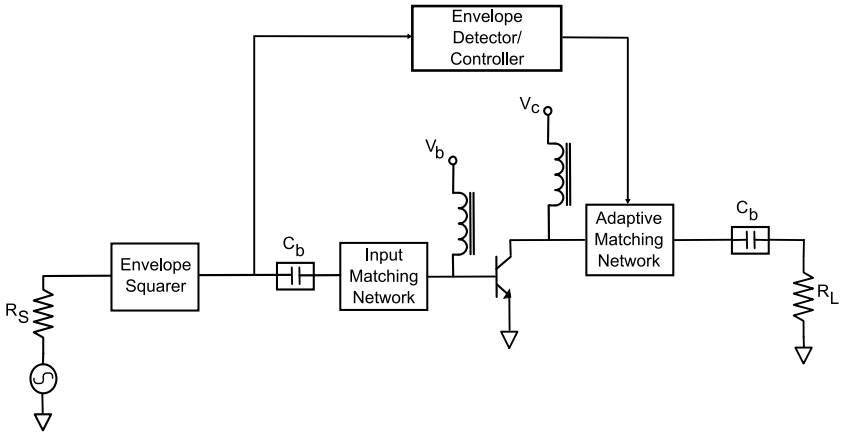


Figure 6.14: Generic implementation schematic of a dynamic load-line PA.

linearity, but distortion is also caused by the time varying nature of the matching network.

Equation (6.6) shows that if the envelope of the signal at a certain point in time tends to zero, then R_{var} which is the load which the adaptive matching network has to offer to the device tends to infinity. However, all practical adaptive matching networks have a finite impedance transformation ratio. In the following discussion, this limitation is tested in simulation (ADS) using the generic setup shown in Figure 6.14 and the equations developed in the previous section. The active device is implemented using an ideal voltage current transfer similar to what was done in Chapter 2. The biasing and current is set such that the active device delivers a maximum of 1 watt to the load. The adaptive matching network is described using a three port SDD block, which allows the impedance it offers to the active device to be controlled by a third port. A two-tone signal with a 1 MHz tone-spacing is used in this test since its envelope goes from zero to the maximum that the device can handle, thus offering a perfect test vehicle.

Figure 6.15 a shows the impedance offered to the active device which varies from 4.5Ω to 450Ω. This is the reciprocal of the envelope square of the base-band signal with its maximum value truncated at 450Ω. Figure 6.15 b shows the envelope of the signal at the output of the device, which is expected from theory to have a constant amplitude. The periodical dip to zero corresponds to the point where R_{var} stays constant and as the envelope of the actual signal goes to zero. The same experiment is again repeated but this time, the impedance transformation ratio is limited even more by setting the maximum impedance the network can provide to 45Ω.

Figure 6.16 shows the simulated output power spectrum for both cases. Limiting the tuning range by a factor 10 in the second situation increases the spectral regrowth by approximately 25 dBc. The simulated efficiency performance for both cases is 78.5%, which is identical to the maximum efficiency for a Class-B amplifier at full power.

The signal distortion is caused due to the fact that the active device is still driven

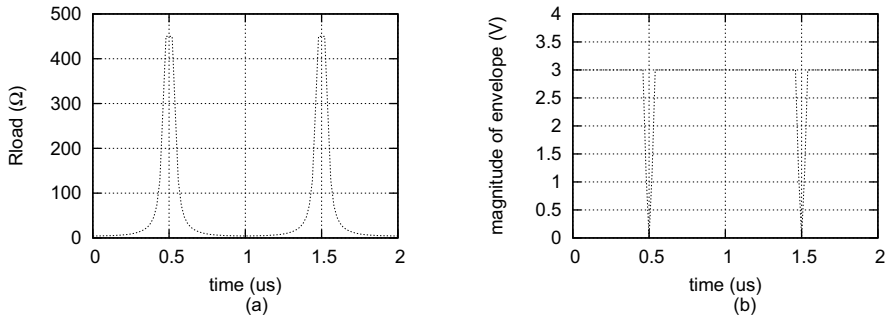


Figure 6.15: (a) The impedance that the device sees as a function of time. (b) Envelope voltage at the output of the active device.

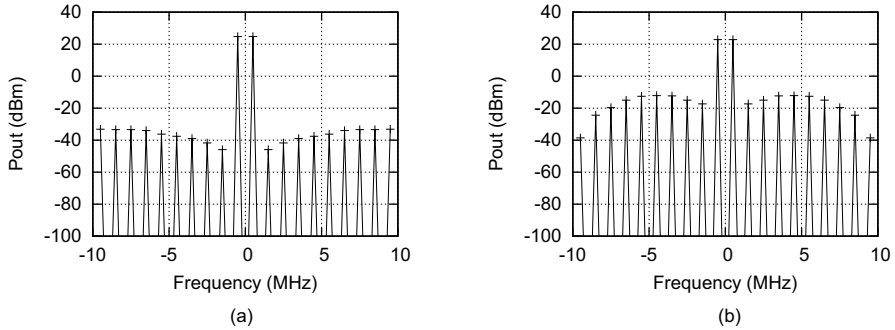


Figure 6.16: (a) Output power spectrum (Pout) when tuning range of load is from 4.5 to 450Ω (b) Output power spectrum (Pout) when tuning range of load is from 4.5 to 45Ω

by originally intended predistorted signal (from the envelope squarer), when R_{var} is limited to a constant upper limit. To solve this problem, the dynamic load line PA has to switch to classical PA opearation (constant output loading) when R_{var} reaches its upper limit. Incorporation of this mode into the test above, results in 84 dBc suppression of IMD products and 1% reduction in efficiency.

In addition to the limited tuning range, phase variation from the input to output of the matching network, as a function of the impedance transformation ratio will introduce phase distortion that also shows up as intermodulation distortion. To investigate this effect, the ideal matching network that was implemented using SDD blocks above is replaced by a two stage L matching network, just like the one in the actual hardware, but now with component values that provide a R_{var} range of 4.5 to 45Ω instead.

Figure 6.17 shows the phase of S_{21} of the adaptive matching network as a function of R_{var}. Note that there is a change of over more than 10 degrees over the entire impedance range. Figure 6.17 b shows the resulting output power spectrum due to

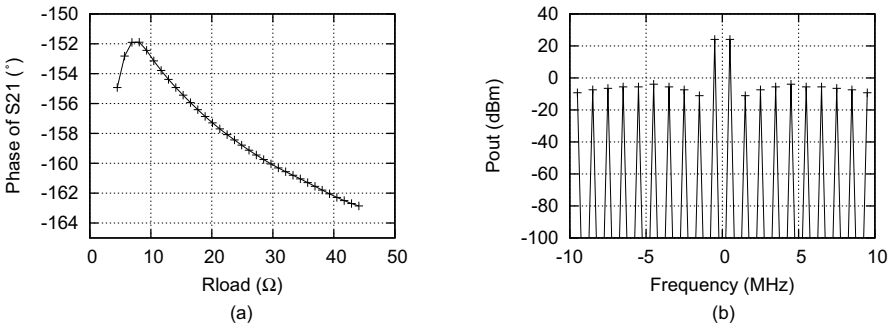


Figure 6.17: (a) Phase distortion as a function of the provided load impedance to the active device. (b) Output power spectrum (Pout) of an ideal device with the implemented adaptive matching network (using ideal components) under dynamic load line operation.

the phase distortion. This effect can be compensated easily by adjusting the phase of the complex base-band signal at the input as was done in the predistortion procedure described in chapter 5.

Lastly, the adaptive matching network block in Figure 6.14 is replaced by actual measured s-parameter data of the implemented adaptive matching network (Figure 6.4), using the import data set feature from ADS. Any real implemented matching network will have, besides its finite impedance transformation ratio, also losses. Figure 6.18 shows how these losses vary as function of time when tracking a complex base-band sine wave of 25 KHz.

This fluctuation in loss as function of the base-band signal will show up as distortion at the output and slightly decreases the efficiency. Also this problem can be solved by modifying the baseband content of the input signal, by using a slightly higher input drive at the times where the losses are higher. The envelope detector will take this predistorted envelope information and change the adaptive matching network correspondingly. Note that this approach will not eliminate the losses, but will correct for the nonlinear distortion caused by the loss fluctuation with the baseband signal.

Figure 6.19 a shows the resulting power-output spectrum of this simulation setup without any correction of; (1) finite load impedance transformation ratio, (2) phase variations with changing impedance transformation of the matching network and (3) fluctuation of the losses with the changing impedance transformation. Note that although the adaptive matching network is based on hardware measurements, the active device is still ideal. And Figure 6.19 b shows the results after performing the correction for all the above mentioned "defects".

The resulting simulated IMD components are below -50 dBc, and the collector efficiency is computed to be 65% for this two-tone signal. In comparison a conventional ideal Class-B PA with no-losses in its matching networks, will exhibit a collector efficiency of 54% for a two-tone signal when driven to its peak power capabilities. Consequently, with this result it has been shown that the dynamic load-line concept,

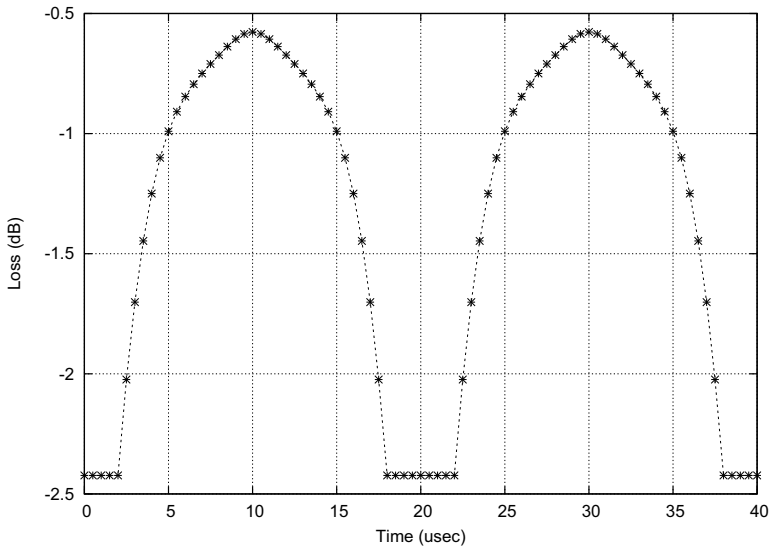


Figure 6.18: Loss of the actual implemented adaptive matching network as a function of time for a base band signal of 25 KHz.

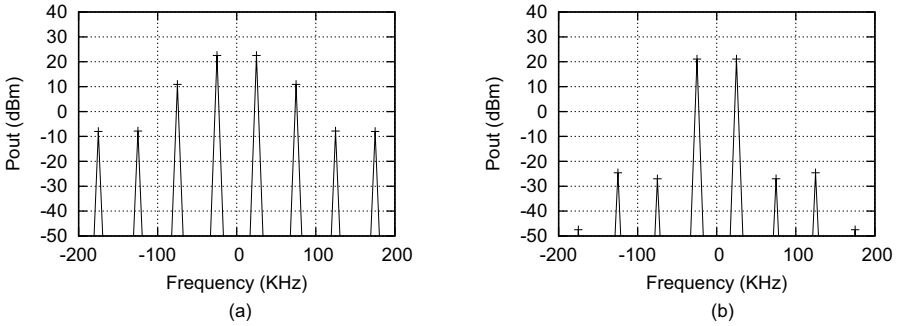


Figure 6.19: (a) Output power spectrum (Pout) of an ideal device with the actual implemented adaptive matching network under dynamic load-line operation. (b)Output power spectrum (Pout) of an ideal device with actual implemented adaptive matching network under dynamic load-line operation after predistortion.

even when including the losses of a real adaptive matching network, can outperform ideal Class-B operation, making it an interesting candidate for the high efficient amplification of communication signals with a high crest factor.

6.3.3 Verification of the dynamic Load-Line PA concept

In the above analysis, all simulations were based on the assumption of an ideal active device. In this section, this ideal model is replaced by a MEXTRAM model of the

QUBIC4G device actually used in the hardware demonstrator. Unlike operating with an ideal device, the optimum load (R_{var}) is now no longer purely ohmic, but should have also a reactive part to cancel output capacitance and other parasitics of the active device. Furthermore, as was mentioned in Chapter 2, these output parasitics are a function of output power, increasing the complexity of the dynamic load-line operation. To determine the optimum output impedance at the various output powers, a load pull simulation of the device biased in Class-B was performed at these levels. Using this data as starting point, the optimum impedance transformation was determined and the complete dynamic load-line PA (just as in the hardware available as shown in Figure 6.8) was simulated using a two-tone signal (Δf = $50KHz$, f_c = $1.8GHz$) as input.

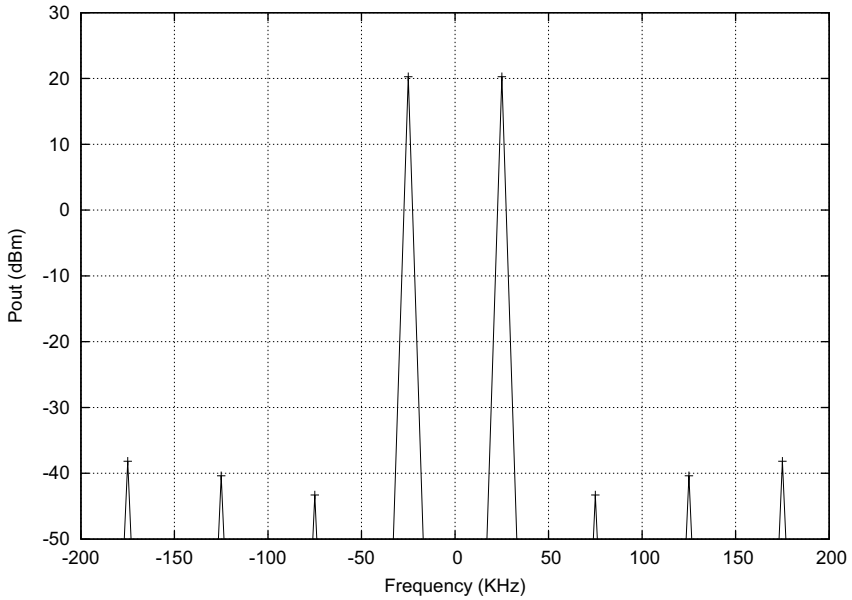


Figure 6.20: Simulated output power spectrum (Pout) of a two-tone signal (Δf = $50KHz$, f_c = $1.8GHz$) for the dynamic load line PA with QUBiC device and measured s-parameter data for the adaptive matching network. The peak power of the envelope is set equal to the maximum output power available under CW condition. The memoryless predistortion compensates for the imperfections of the adaptive matching network such as: dynamic changes in phase, losses and the limited impedance transformation ratio.

Figure 6.20 shows the simulated output power spectrum (Pout) of the dynamic load line PA with memoryless predistortion applied to correct for the active device non-linearities, and all above mentioned limitations of the adaptive matching network. A simulated collector efficiency of 56.2% was obtained with all intermodulation products below 60 dBc. When comparing this result with conventional Class B PA operation for this QUBiC device using a fixed load for maximum output power, the simulated efficiency for this Class-B operation is 50% and an IM3 level of -15 dBc

when no predistortion is applied.

Finally, to verify the dynamic load-line concept in hardware, the setup shown in Figure 6.10 was used. Due to logistic reasons, a slightly different matching network sample was used for the following tests rather then the ones used to generate the results in Section 6.2. The high-voltage arbitrary waveform generators that were used previously to provide the DC control voltages to the varactors of the integrated matching network, are now programmed to deliver the dynamically varying control signals needed to achieve the desired impedance variation with the envelope of the signal to be offered to the output of the active device.

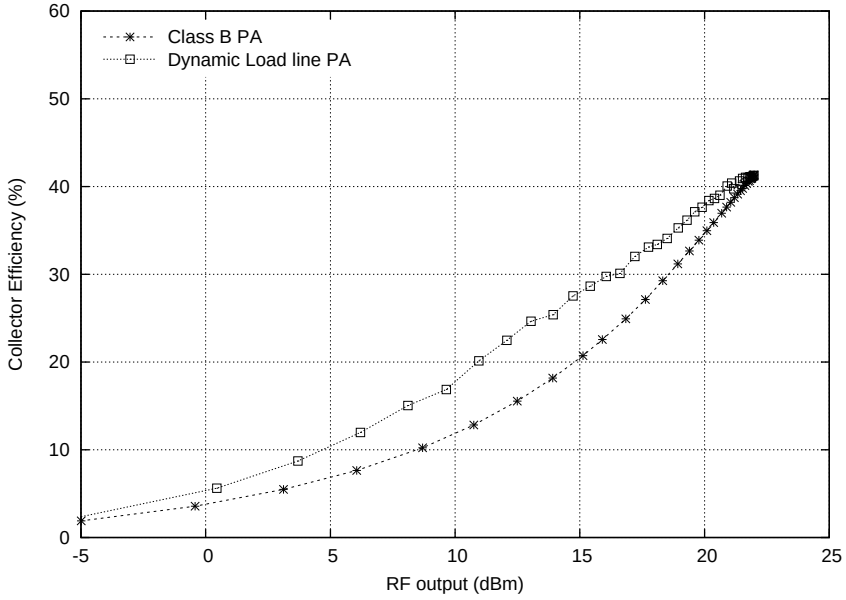


Figure 6.21: Collector efficiency versus RF output power of the dynamic load-line amplifier when the load is adaptive for optimum efficiency at each output power level and when for the same amplifier the load is fixed to the maximum power-out condition (Class-B operation)

Figure 6.21 shows the CW-collector efficiency measured at 1800 MHz for the case where the load is fixed at maximum RF output power and the case where the impedance is adapted with RF output power. Note that the maximum output power reached with this sample is 22 dBm and the maximum collector efficiency is 41%. This is clearly not a optimum sample as its performance is far from what was achieved in Section 6.2, but it is sufficient for the purpose of verifying the dynamic load-line concept. The following figure shows the instantaneous efficiency for both the dynamic load-line PA and the Class-B PA when driven by a two-tone signal with 50 KHz tone spacing. Note that for the Class-B PA operation, it is driven into compression.

Table 6.3 shows the measured collector efficiency for the two-tone signal in both cases. A 3% efficiency improvement was obtained for the two-tone signal, something that could have been estimated based on the CW efficiency profile shown in

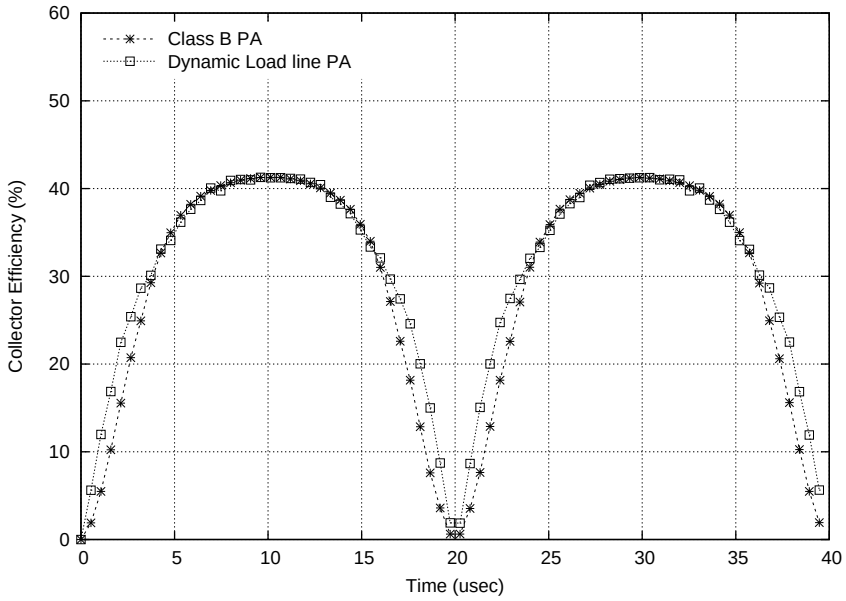


Figure 6.22: Comparing the instantaneous collector efficiency of a two-tone signal for the full dynamic load-line PA and a Class-B PA.

Figure 6.21. Linearity results are not presented, because even though all the corrections for the limitations of the adaptive matching network was made and memoryless predistortion applied, the IM3 was not better than -35 dBc. This limited linearity improvement can be attributed to memory effects of the PA which were not addressed by these corrections. Nevertheless, the concept of the dynamic load line concept was shown to be working and is capable of improving efficiency performance when the PA driven with modulated signals, especially the ones with high crest factor.

Table 6.3: Measured performance of the dynamic load line PA

	Dynamic Load Line PA	Class B PA
Average Pout	18.52 dBm	18.57 dBm
Collector Efficiency	29.4%	26.4%

6.4 Conclusion

To the author's knowledge, for the first time, a multi-band multi-mode power amplifier, which is continuously tunable in operating frequency and output power level, has been demonstrated. The demonstrator PA has a peak output power between 27 and 28 dBm (0.5 to 0.7 watts) at the 900, 1800, 1900, and 2100 MHz bands with very reasonable efficiencies. The capability to adapt the loading of the active device as function of desired output-power results in high-efficiency performance under static power back-off, e.g. in power control loop applications. The integrated varactor-based matching networks are low-loss and provide low distortion due to the use of the anti-series silicon-on-glass technology. The required control voltages in this amplifier for the intended load trajectory were restricted to 10 V. Since all varactors are reversed biased, the control of the adaptive network requires almost no DC power. The total chip area including integrated matching networks is restricted to 8 mm^2, and can be made even more compact.

Another strength of the varactor-based matching network is that it in principle can handle high tuning speeds (up to 20 MHz). Therefore, it is capable to track the output stage loading conditions with the envelope of the modulated signal. This approach, known as dynamic load-line modulation, enables the PA to operate at its maximum efficiency at each point in time, yielding considerable efficiency improvements also for modulated signals with a high crest factor. Using the design which was meant for the multi-band multi-mode power amplifier, it was shown in simulations and measurement that the efficiency for signals with a non-constant envelope can be improved. Although the principles are demonstrated, still increased efforts must be made to solve the higher distortion due to the continuous adaptation of the matching network. Contributors to the overall distortion in a dynamic load-line PA were identified to be:

1. Finite impedance transformation ratio of the adaptive matching network,
2. Impedance dependent phase shift from input to output of the adaptive matching network,
3. Impedance dependent losses of the adaptive matching network,
4. Distortion from the active device.

Although simulation results of this concept show high efficiency together with good linearity after applying memoryless predistortion, in practise improvement in efficiency was marginal and memoryless predistortion was not effective. This is attributed to the presence of memory effects that were not included in the simulation. At this moment it is not customary to apply memory compensation algorithms in handsets, due to cost and power consumption considerations. This issue is not further investigated within these thesis. However, that leaves us with the question about the cause of the memory effects in our practical demonstrator. After close inspection it proved that one of the dominant contributors to memory effects was the requirement for high impedance in the center node connection of the anti-series configured (uniformly doped) varactors diodes. This impedance is needed to guarantee a linear operation of the varactors also for band modulated signals [52]. It works out that this requirement is somewhat conflicting with the desire to modulate the varactors with the envelope of the modulated signal. This gives rise to increased distortion in

practical implementations. In the initial simulations this phenomena does not appear due to the fact that *s*-parameters where used to model the matching network. After recognizing the problem, new developments were started to find a new varactor topology that is suited for fast modulation while not introducing memory effects. This new activity already has created many exciting results, which will find soon their application in a new generation of dynamic load-line amplifiers [36]. It is expected that for these new implementations no memory correction will be required, so the use of memoryless predistortion as introduced in this chapter will be sufficient. The combination of these techniques can provide compact, low-cost, band-switching, highly-efficient power amplifiers for future mobile phones.

In the next chapter, the focus is on power amplifiers for base station applications, which should provide even higher levels of output power and efficiency, while band-switching is less a concern. These slightly different requirements, which make the hardware demands for base station applications much more severe, demand a different approach than the use of dynamic matching networks. Consequently, new design techniques and signal processing concepts will be developed to meet also here the challenges set by the latest generation of wireless communication standards.

Chapter 7

Doherty Power Amplifier for base stations

In Chapter 1, the important considerations in the development of base-station PAs were discussed. The increasing information bandwidth and amount of users per channel, led to the use of coding and modulation schemes, which require RF signals with significantly large Crest Factors (CF), also commonly refered to large peak-to-average power ratios. The Class-B/AB amplifier, which was and still is in many places the de-facto base-station PA, is no longer very attractive, simply because when the Class-B/AB amplifier is driven with RF signals with high CF, it exhibits typically a relatively low average efficiency.

It was also discussed in chapter 5 that to avoid this efficiency degradation, many high efficient PA architectures have been proposed over the past 5 years. Despite of the large variety in amplifier concepts, only the Envelope Tracking [31, 38] and Doherty amplifier [53, 54] concepts seem to be actively pursued by the industry to become commercial products for the base-station market. In these approaches, the Envelope Tracking amplifier, in its most promising implementation, makes use of an efficient DC-to-DC converter in combination with a video amplifier to modulate the supply voltage. This increases cost and gives rise to some additional power losses that slightly degrade the overall efficiency of the amplifier. As such, although the supply tracking PA is promising, commercialize implementation is still many years away. The Doherty amplifier, which does not require additional "peripheral" hardware, seem to have been adopted by the industry as the way forward in boosting efficiency performance at this present time, especially when signals with high CF are involved. Base station manufacturers such as Nokia (Europe) and Huawei (China) are at this moment already selling "green" base-stations based on the two-way symmetrical Doherty power amplifier. Unfortunately, practical implementations of three-way and higher branch Doherty PAs, which enables even higher efficiency performance possible for signals with CF>10 dB, are still rare and not fully convincing in their efficiency and linearity performance (even in the research labs). The main reason for this, is the rather difficult design and the delicate fine tuning of the power distribution/combining networks and the sensitive biasing of the main and peaking amplifier

stages.

In this Chapter the Doherty Power Amplifier (DPA) for base-station is investigated. The goal is to study the theoretical limitations of the three-way DPA, and to discover design techniques that will allow the design of a three-way Doherty to come close to the theoretically allowed efficiency performance. Section 7.1 starts with the well known two-way DPA, a novel approach to analyzing the two-way Doherty is shown, and note that this technique is easily scalable to the analysis of a three-way Doherty. Concepts presented here are further supported by actual designs of a two-way symmetrical and asymmetrical Doherty based on LDMOS technology from NXP. In section 7.3, the step to the three-way DPA is made, a detailed theoretical analysis is given, followed by verification of the new concepts. Lastly, with the insights obtained in section 7.5 and 7.6, the development of two novel modified three-way Doherty design is given.

7.1 Two-way Doherty Analysis

The analysis of a two-way Doherty amplifier is well documented in literature [5]. In all of these analysis, they assume a fixed network topology, which is designed for its optimum element values in order to obtain the desired high-efficiency power back-off point(s). In this section, a more rigorous approach is taken, where no assumption of the output power combining network is made besides the fact that it is lossless and reciprocal. Figure 7.1 shows the basis of the proposed analysis, where the output power combining network is represented by its black-box parameters (which is to be determined), and at one end it is driven by voltage controlled current sources, representing the active devices, and at the other end the resistive load is connected.

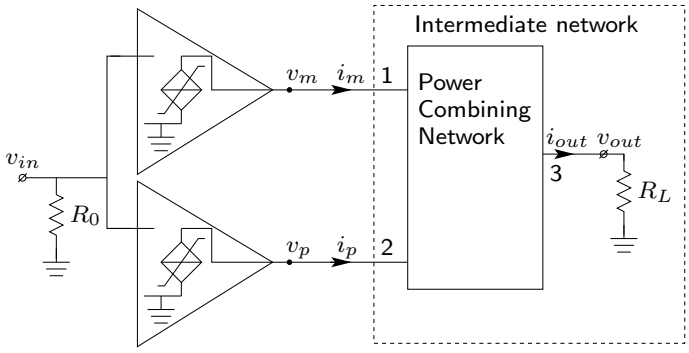


Figure 7.1: General two-way DPA with ideal voltage-controlled current sources and a lossless power-combining network.

In the first part of the analysis, the load R_L is incorporated into the DPA output power combining network. This is to reduce the complexity from a three-port to a two-port problem. This "intermediate" network though still reciprocal is no longer lossless since it now includes the dissipating element R_L.

Next the voltages and currents at the output of the main and peak devices (Figure 7.1) are expressed as a function of the black-box parameters of the intermediate network for two specific drive conditions, namely, the point where the DPA amplifier delivers its maximum (full) power (F), as well as the point where the main device reaches its high-efficiency condition at the back-off point (B) (see Figure 7.2 (d)); thus, at the power back-off point (B):

$$v_{m,B} = i_{m,B}\, z_{11} + i_{p,B}\, z_{12} \tag{7.1}$$

$$v_{p,B} = i_{m,B}\, z_{21} + i_{p,B}\, z_{22} \tag{7.2}$$

And at full power (F),

$$v_{m,F} = i_{m,F}\, z_{11} + i_{p,F}\, z_{12} \tag{7.3}$$

$$v_{p,F} = i_{m,F}\, z_{21} + i_{p,F}\, z_{22} \tag{7.4}$$

Where $v_{m,F}$, $i_{m,F}$ and $i_{m,B}$, are the voltages and currents of the main device at the full output power and the back-off point, respectively, while $v_{p,F}$, $i_{p,F}$ and $v_{p,B}$, $i_{p,B}$ are the corresponding voltages and currents for the peaking device (see Figure 7.2). Solving equation (7.1)–(7.4) simultaneously for the z-parameters yields

$$\begin{aligned}
z_{11} &= \frac{-v_{m,F} i_{p,B} + v_{m,B} i_{p,F}}{-i_{m,F} i_{p,B} + i_{m,B} i_{p,F}} \\
z_{12} &= -\frac{v_{m,B} i_{m,F} - v_{m,F} i_{m,B}}{-i_{m,F} i_{p,B} + i_{m,B} i_{p,F}} \\
z_{21} &= -\frac{v_{p,F} i_{p,B} - v_{p,B} i_{p,F}}{-i_{m,F} i_{p,B} + i_{m,B} i_{p,F}} \\
z_{22} &= \frac{v_{p,B} i_{m,F} + v_{p,F} i_{m,B}}{-i_{m,F} i_{p,B} + i_{m,B} i_{p,F}}
\end{aligned} \tag{7.5}$$

The goal at this point is to determine all the voltage and current variables such that the black-box parameters of the intermediate network can be subsequently computed. To do this, the voltages and currents that are already "fixed" by the efficiency requirements are first written down. Note that high efficiency requires that the voltage amplitude at the output of the active devices be maximized with respect to their dc supply voltage when an output current is present. For the two-way DPA, the output voltage amplitude of the main device should be maximum at full output power (F), as well as at a specific back-off power level (B) [see Figure 7.2 d]. Consequently, the amplitude of $v_{m,F}$, $v_{m,B}$, and $v_{p,F}$ must be set to their maximum value, which in Class-B operation equals their dc supply voltage, thus, $v_{m,F} = v_{m,B} = v_{p,F} = V_{DC}$. Furthermore, to guarantee maximum efficiency in power back-off, the peaking amplifier is off until the backed-off point is reached, yielding $i_{p,B} = 0$. This leaves $i_{m,F}$, $i_{p,F}$, $i_{m,B}$, and $v_{p,B}$ as the remaining unknowns to solve for.

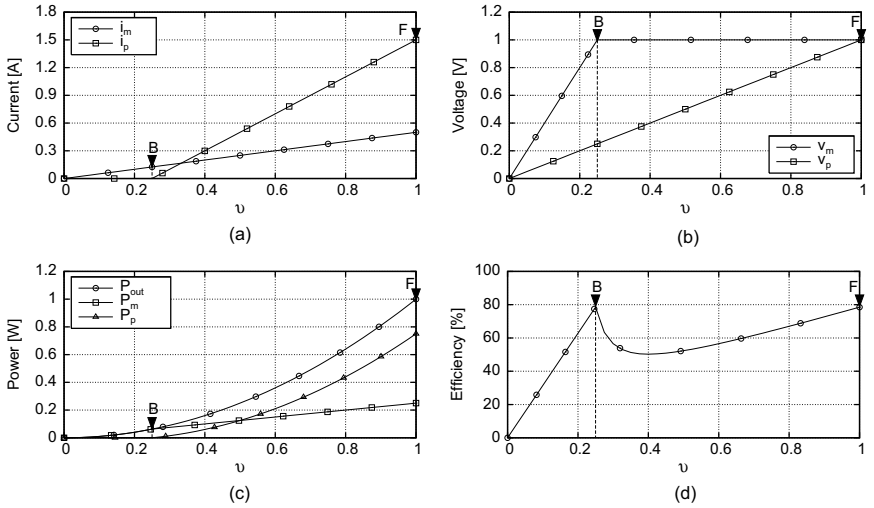


Figure 7.2: ADS simulation results based on the schematic of Figure 7.1 as verification of the two-way DPA analysis. The power-combining network is calculated for $k = 0.25$. The devices operate in Class-B (shorted higher harmonics at their output) and the design is normalized to 1-W output power and 1-V RF input voltage. (a) Output current for main and peak devices versus normalized RF input voltage. (b) Resulting voltage amplitude at the output of the main and peaking devices as a function of the normalized RF input voltage. (c) Power contributions and resulting total power of the main and peaking devices. (d) Resulting efficiency as a function of the normalized RF input voltage.

To address the above, the maximum output power delivered by the DPA, assuming a lossless power-combining network is considered:

$$p_{out,max} = \frac{1}{2}\,\mathbb{R}e\left(v_{m,F}i^*_{m,F} + v_{p,F}i^*_{p,F}\right) \tag{7.6}$$

Note that, in this equation, high efficiency is only achieved when the current and voltage related to a device (e.g. $v_{m,F}$ and $i_{m,F}$) are in phase, yielding maximum (real) power delivered to the load. Similarly, the following expression for the output power at the back-off point can be written:

$$p_{out,B} = \frac{1}{2}\,\mathbb{R}e\left(v_{m,B}i^*_{m,B} + v_{p,B}i^*_{p,B}\right) = \frac{1}{2}\,\mathbb{R}e\left(v_{m,B}i^*_{m,B}\right) \tag{7.7}$$

Next the power back-off/full-power ratio k^2 can be written as:

$$k^2 = \frac{p_{out,B}}{p_{out,F}} = \frac{\mathbb{R}e\left(v_{m,B}i^*_{m,B}\right)}{\mathbb{R}e\left(v_{m,F}i^*_{m,F} + v_{p,F}i^*_{p,F}\right)} \tag{7.8}$$

When the DPA is backed-off from full power to lower output levels, it is essential that the output power change linearly with the input power or with the square of the

input voltage. Consequently, the following condition enforces the output current of the active devices to their input-driving voltage. Here, the most logical choice from an implementation point-of-view is to assume a linear relation between output current and input voltage when the device is turned on. Thus, this follows as:

$$i_{m,B} = k\, i_{m,F} \tag{7.9}$$

The final condition to solve for the black-box parameters of the intermediate two-port network is found by making use of the reciprocal property $z_{12} = z_{21}$ yielding:

$$v_{m,B}\, i_{m,F} - v_{m,F} i_{m,B} = v_{p,F} i_{p,B} - v_{p,B} i_{p,F} \tag{7.10}$$

Using equation (7.6)–(7.10), the variables $i_{m,F}$, $i_{p,F}$, $i_{m,B}$ and $v_{p,B}$ are uniquely determined and the z-parameters of the intermediate network network are given by equation (7.5). The related s-parameters of the two–port network are:

$$s_{11} = -\frac{k^2 - 2}{k^2 + 2} \tag{7.11}$$

$$s_{12} = s_{21} = -\frac{j2k}{k^2 + 2} \tag{7.12}$$

$$s_{22} = -\frac{k^2}{k^2 + 2} \tag{7.13}$$

Note that to keep the above equations to a manageable size, the phase angle between the power waves of port 1 and port 2 have been set to 90° .

Having solved the two-port parameters of the intermediate network, the last remaining step in determining the power-combining network for the two-way DPA is to reintroduce the black-box parameters that involve the third port connected to the load, namely, s_{13}, s_{31}, s_{23}, s_{32} and s_{33}. Note that the original two-port -parameters do not change when going to a three-port network, provided that the loading impedance is identical to the normalization impedance used for the s-matrix. The remaining s-parameters involving the third port can be found by imposing the well-known lossless three-port s-parameter conditions using ($N = 3$) using:

$$\sum_{n=1}^{N} |s_{np}|^2 = 1 \qquad \forall\, p \tag{7.14}$$

$$\sum_{n=1}^{N} s_{np}\, s_{nq}^{*} = 0 \qquad \forall\, p \neq q \tag{7.15}$$

The expansion to the three-port case results in six equations. Since $s_{13} = s_{31}$ and $s_{23} = s_{32}$ due to the reciprocity, there remain six unknowns (since each s-parameter is, in general, complex) and six independent equations. Consequently, under the condition that the RF signal in port 2 lags with respect to the signal in port 1 by 90 degree, a unique analytical solution can be found for the three-port s-matrix representing the power-combining network of the two-way DPA. The resulting s-parameter solution is a function of k and α , where α is the phase of s_{23} representing the electrical angle of the signals between ports 2 and 3

$$
\begin{aligned}
s_{13} &= s_{31} \\
&= \frac{2\,k\tan(\alpha) - j\,2\,k}{\sqrt{\frac{1}{1+\tan(\alpha)^2}}(k^2\tan(\alpha)^2 + k^2 + 2 + 2\,\tan(\alpha)^2)}
\end{aligned} \tag{7.16}
$$

$$
\begin{aligned}
s_{23} &= s_{32} \\
&= \frac{2\sqrt{\frac{1}{1+\tan(\alpha)^2}}(1 + j\tan(\alpha))}{k^2+2}
\end{aligned} \tag{7.17}
$$

$$
s_{33} = \frac{k^2(\tan(\alpha)^2 - 1) + j\,2\,k^2\tan(\alpha)}{k^2\tan(\alpha)^2 + k^2 + 2 + 2\,\tan(\alpha)^2} \tag{7.18}
$$

The following equations give the required current characteristics of the active devices as a function of the normalized input drive υ, which represents the ratio of the actual drive voltage to the maximum drive voltage and is, therefore, dimensionless:

$$
i_m = 2\,k\,p_{out,max}\frac{\upsilon}{V_{dc}} \tag{7.19}
$$

$$
i_p = \begin{cases} \frac{2\,p_{out,max}}{V_{dc}}(\upsilon - k) & \text{if } \upsilon \geq k, \\ 0 & \text{if } \upsilon < k. \end{cases} \tag{7.20}
$$

With the above, the theoretical analysis of the two-way DPA is fully complete.

7.1.1 Two-way Doherty Analysis Verification

For the verification of the previous analysis, the schematic of Figure 7.1 is simulated using Agilent's Advanced Design System (ADS). For the power-combining network, the calculated three-port s-parameters with k =0.25 and α =0 are used. Note that this relates to an asymmetrical DPA design [53] with an efficiency peak at 12-dB back-off of the full power. The main and peak devices are modeled with symbolically defined devices (SDDs) just as was also done in Chapter 2, which implement the current transfer functions given in equations (7.19) and (7.20). The supply voltage of the active devices is set to 1 V while the maximum current of the main and peaking devices is chosen such that the maximum output power equals 1W. The ADS simulation results are given in Figure 7.2. In this figure, the amplitude of the current of the main and peaking devices is plotted versus input drive voltage. Note that the peaking device is only activated beyond the power back-off point B. The resulting RF amplitude for the RF output voltage of the main and peaking devices is given in Figure 7.2 b. Note that at the high efficiency points (B) and (F), the voltage amplitude in the main device indeed reaches the level of the supply voltage. Moreover, for the main amplifier device, the voltage amplitude remains constant and equal to the supply voltage between these points, while for the peaking device, there is a linear increase in voltage amplitude reaching V_{DC} at maximum output power.

Figure 7.2 c shows the power delivered by the main and peaking devices, as well as the total power. In Figure 7.2 d, the resulting efficiency is plotted versus input voltage. As expected, the efficiency peaks at full power (F) and at the chosen back-off point (B).

One can conclude from this result that the analysis method proposed indeed results in the desired DPA behavior. However, unlike previous design methods, no presumptions were made on the network topology. An additional advantage is that, one can also consider alternative driving conditions for the main and peak amplifiers. The following section shows how to relate the above analysis to the well known micro-strip line implementation of a two-way symmetrical and asymmetrical Doherty, and discusses an efficient design strategy for the proper design of the two-way Doherty power amplifier.

7.2 Design of a two-way asymmetrical DPA

In the following design verification, a two-way DPA targeted for WiMAX base-station will be designed and made. The general specifications are that it should be able to deliver a peak power of at least 51 dBm, and have at least 13dB gain. Note that WiMAX uses OFDM techniques to stack numerous users in a limited frequency band, as such, like WCDMA the resulting RF signals have high crest-factors, typically in the range of 10-13dB. To ensure that the DPA can still deliver good efficiency performance when driven by signals with such high CF, the efficiency peaks of the DPA should be engineered, such that the back-off peak lies in the same magnitude of the CF of the signal.

Substituting the relevant voltage and current into equation (7.8) and simplifying, results in the following:

$$k = \frac{i_{m,F}}{i_{m,F} + i_{p,F}} \tag{7.21}$$

This indicates that the back-off ratio k is related to the maximum current which the device can deliver, which is indirectly related to the size of the device. Thus, assuming a linear drive profile (as indicated in (7.19)) and that the size of the transistor used for the peak is m times greater than that of the main (i.e. $i_{p,F} = m\, i_{m,F}$). The following equation is obtained which relates the device sizes used to that of the power back-off ratio for which the DPA has its efficiency peak.

$$k = \frac{1}{1 + m} \tag{7.22}$$

Where m is the ratio of the power rating of the transistor in the peak with respect to the transistor used in the main.

From the selection of transistors available, the BLF6G27-45 LDMOS transistor from NXP is chosen, which has a power rating of 45W for its 1dB compression point. The main amplifier will be designed with one of these transistors and the peak amplifier will employ two of these transistors in parallel, giving a total of 135 Watts (51.3dBm) output power in the Doherty configuration. This sets the value of m to 2

and $k = \frac{1}{3}$ accordingly. This results in a potential asymmetrical Doherty, which has its second efficiency peak at 9.5 dB back-off, which is in the order of the crest factor of the signals used in WiMAX.

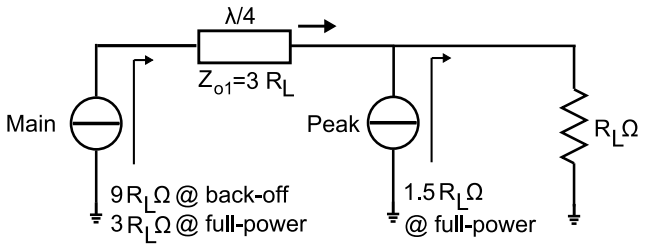


Figure 7.3: Simplest transmission line implementation of a two-way Doherty output power combiner.

For the frequency and power range, typically used for base-station applications, microstrip is the preferred choice for the implementation of the matching networks. Using the black-box parameter analysis above, the simplest solution for the Doherty power combining network occurs when we substitute $\alpha = 0$. This results in the black-box s-parameters reducing to the following:

$$s = \frac{1}{k^2+2}\begin{bmatrix} -k^2+2 & -j\,2k & -j\,2k \\ -j\,2k & -k^2 & 2 \\ -j\,2k & 2 & -k^2 \end{bmatrix} \tag{7.23}$$

Converting this three-port s-parameter to its z-parameter representation results in the following:

$$z = \begin{bmatrix} 0 & -j\,\frac{Z_o}{k} & -j\,\frac{Z_o}{k} \\ -j\,\frac{Z_o}{k} & 0 & 0 \\ -j\,\frac{Z_o}{k} & 0 & 0 \end{bmatrix} \tag{7.24}$$

Where Z_o is the normalizing impedance of the s-parameter black-box.

Synthesizing this z-parameter black-box results in a quarter-wave length transmission line connected between the main and peak amplifier (see Figure 7.3). The following equations give the relationship between the terminating impedance (in most case being 50 Ω), and the impedance which the main and peak amplifier will see during the Doherty operation.

$$Z_{o1} = \frac{1}{k}\,Z_o = \frac{1}{k}\,R_L \tag{7.25}$$

Note that R_L is in this case the terminating impedance which was referred to in the s-parameter black-box analysis.

Under full power drive:

$$Z_{m,F} = Z_{o1} = \frac{1}{k} R_L \tag{7.26}$$

$$Z_{p,F} = R_L \frac{1}{1-k} \tag{7.27}$$

At desired back-off drive (9.5dB in this case):

$$Z_{m,B} = \frac{Z_{o1}^2}{R_L} = \frac{1}{k^2} R_L \tag{7.28}$$

Substituting $k = \frac{1}{3}$ and $R_L = 50\Omega$ into the above equations results in the characteristic impedance of the line Z_o being 150Ω, and the impedance which the main and peak amplifier have to see at full power as 150Ω and 75Ω respectively. Knowing the peak power capability of the transistor ($P_{max,main}$,$P_{max,peak}$) and the drain supply voltage the devices are biased at, the impedance which the devices need to see for maximum power transfer is approximated as (not taking into consideration package parasitics):

$$R_{opt} = \frac{V_{DC}^2}{2\,P_{max}} \tag{7.29}$$

For the device that is used here, it works out that the impedance that this device needs to see is approximately 8.7Ω. Therefore to use this device in the Doherty combiner as it is now, an impedance transformation of greater than 15 is required for the Main amplifier, which would require a large amount of board space to make this impedance transformation wideband.

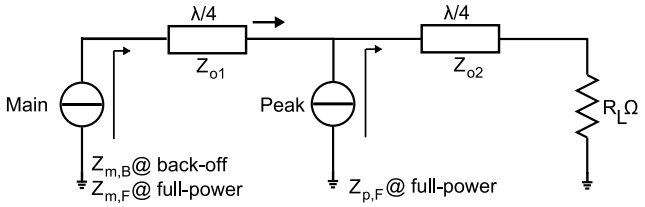


Figure 7.4: Two quarter-wave length line transmission line implementation of a two-way Doherty output power combiner. This implementation has the flexibility of choosing the characteristic impedance of the lines and the impedances offered to the Main and Peak amplifier.

To get around the need for such a large impedance transformation ratio, an extra quarter-wave length line with characteristic impedance Z_{o2} (see Figure 7.4) is added between the output of the Peak amplifier and the load. This same topology can also be arrived at from the black-box parameter analysis by renormalizing the s-parameter in the two-port stage and setting $\alpha = -90°$ when introducing the third port. This analysis is given in Appendix A.

This topology result in the following equations for the impedances offered to the Main and Peak amplifier cells as a function of the characteristic impedances of the lines (in which we replace k as a function of m (7.22)).

Under full power drive:

$$Z_{m,F} = \frac{Z_{o1}^2 R_L}{Z_{o2}^2(1+m)} \tag{7.30}$$

$$Z_{p,F} = \frac{Z_{o2}^2}{R_L}\frac{1+m}{m} \tag{7.31}$$

and at the desired back-off drive (9.5dB in this case):

$$Z_{m,B} = \frac{Z_{o1}^2 R_L}{Z_{o2}^2} = (1+m)Z_{m,F} \tag{7.32}$$

The choice of $Z_{m,F}$ and $Z_{p,F}$ will set the characteristic impedance, Z_{o1} and Z_{o2} of the quarter-wave length lines as given by the following equations:

$$Z_{o2} = \sqrt{Z_{p,F} R_L \frac{m}{1+m}} \tag{7.33}$$

$$Z_{o1} = \sqrt{Z_{m,F} Z_{p,F}\, m} \tag{7.34}$$

The above equations for $Z_{m,F}$, $Z_{p,F}$, Z_{o1} and Z_{02} allow the designer to make trade-offs between the impedance which we need to match the device to (impedance transformation ratio) and the thickness of the quarter-wave length lines in the Doherty power combining network. In this design, $Z_{m,F}$ is chosen equal to 50Ω. The motivation for this is to allow the independent characterization of the Main PA in a 50Ω system, which is the default industrial practise at the time of writing.

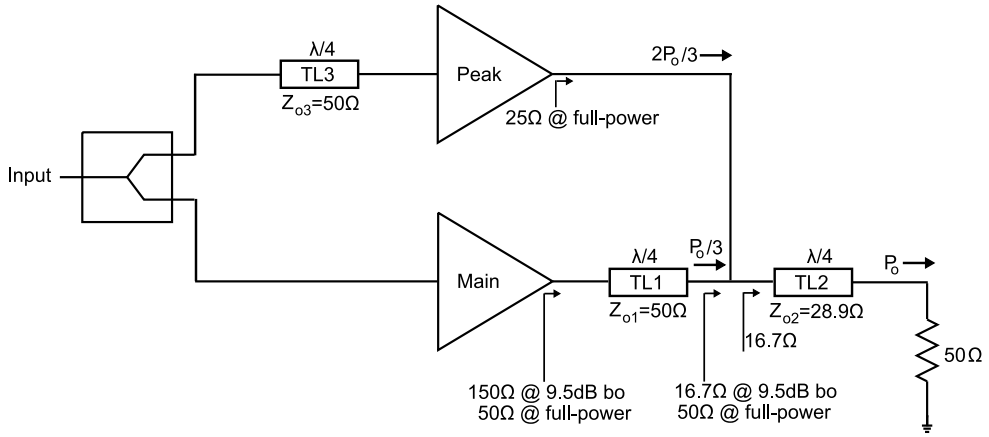


Figure 7.5: Asymmetrical DPA Wimax.

The final terminating impedances for the main and peak and their resulting characteristic impedance of the $\frac{\lambda}{4}$ lines for the WiMAX asymmetrical DPA design is

shown in Figure 7.5. With this, the Doherty power combining network for this design is properly defined and the attention is now turned to designing the Main and Peak amplifier cells.

The design of the Peak amplifier cell is relatively straight forward. In this case, it has to be matched to an impedance of 25Ω for maximum efficiency and power out. The active device is biased in Class-C and a load pull measurement/simulation is done to find the optimum impedance which the device has to see for maximum efficiency and still meeting the output power requirement. At the input, a conjugate match is provided. Note that in practise, the actual gate bias has to be fine tuned at a later phase but this would not affect the matching condition significantly. The final step requires the use of a transmission line transformers to match to the optimum impedance of 25Ω at the output and 50Ω at the input.

The design of the Main amplifier cell is slightly more involved as it requires a good power and efficiency match at two power levels. In this particular case, the active device has to be matched to 50Ω when delivering maximum output power and 150Ω when the power is backed of by $\frac{1}{3}$. This can be done by first performing a load-pull measurement/simulation at the maximum output power (Z_{opt1}) and repeat this for an approximately three times lower input drive such that a optimum impedance (Z_{opt2}) can be found that maximizes efficiency when the output power is backed off by 3 times. Note that in general $\mathbb{R}e(Z_{opt1}) \neq \mathbb{R}e(Z_{opt2})$ due to the package parasitics.

The next step involves transferring the loading conditions offered by the Doherty power combiner to the main amplifier at full power (50Ω) and back-off power (150Ω) to the optimum loading conditions found for the main device (Z_{opt1} and Z_{opt2} respectively) using the same matching network. Clearly, in this case a single transmission line matching network will not be sufficient, as this will only allow Z_{opt1} to be transformed to 50Ω or Z_{opt2} to be transformed to 150Ω but not simultaneously. A simple solution for this is to use two transmission line in series for matching (literally a step, a thick line close to the device to a narrower line, see Figure 7.6, where the main is at the top of the figure). This will give four independent parameters (two different characteristic impedances and two different line lengths) to tune in the matching to match two complex loads. At the input it is less complex as only a conjugate match is required to 50Ω and this is done with a single piece of transmission line.

The following step in the asymmetrical Doherty design is the power splitter at the input. The power splitter should provide an asymmetrical split such that the branch connected to the peak will have two times as much power as the branch connected to the main. To do this, a rat race type splitter is employed with slight modifications to the characteristic impedance to achieve this asymmetrical power split (See Figure 7.6).

The last step in this asymmetrical Doherty design is to put these separate blocks together, and make some fine adjustments. The first consideration/adjustment would be to ensure that at power levels where only the main is on, the peak amplifier should not load the Doherty power combining network (adsorb power). To ensure this, a piece of transmission line with the characteristic impedance equal to the terminating impedance of 25Ω can be added to rotate s_{22} of the peak amplifier cell to a high ohmic value. In this particular design this was not necessary. Secondly, when both Main and Peak are delivering power, these powers should reach the load in phase for maximum output power, this might require the use of phase offset lines, such as TL3

in Figure 7.5. And lastly, to further fine tune when the peak amplifier turns on, the gate bias of the peak device can be modified, and this will affect the efficiency versus back-off power behavior.

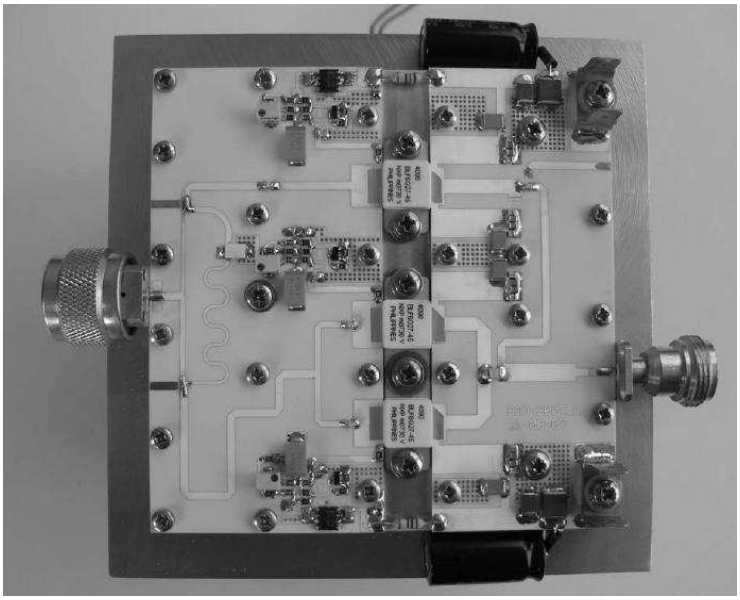


Figure 7.6: WiMAX Asymmetrical Doherty Base Station PA for 2.5 to 2.6 GHz. Max. output power of 51dBm.

Figure 7.6 shows the final implemented WiMAX asymmetrical Doherty for 2.5GHz to 2.6GHz band, implemented with the above procedure on Rogers 4530B PCB. This PA is verified using a standard Base station Amplifier test bench, with CW power sweeps and WiMAX signals.

Figure 7.7 shows the measured drain efficiency of this WiMAX DPA using a CW signal at 2.555GHz. It reaches a peak drain efficiency of 52% at full power (51 dBm) and at 10 dB back-off the drain efficiency is 34%. Compared with a ideal class-B amplifier normalized to the peak efficiency of this Doherty, the asymmetrical Doherty offers 15% efficiency increase at 10 dB power back-off, which is rather significant. Note that higher back-off efficiencies are possible but this have been traded off for a flatter gain versus output power characteristic (a cheaper commercial predistorter can be used), and also for better Gain flatness across the whole band (WiMAX standards specifies a less than 0.5 dB gain fluctuation across the band). The gain characteristics are shown in Figure 7.8.

The DPA is now driven with actual WiMAX signals (OFDMA 64QAM, 10MHz, 5ms frame length, TDD, CF =9.5 dB) such that the actual peak power reaches the maximum CW power that the DPA can deliver. The measured results are given in Table 7.1.

The measured performance shows a 36% measured efficiency for the WiMAX signal. Note that, when signals of comparable CF is used with an ideal Class-B

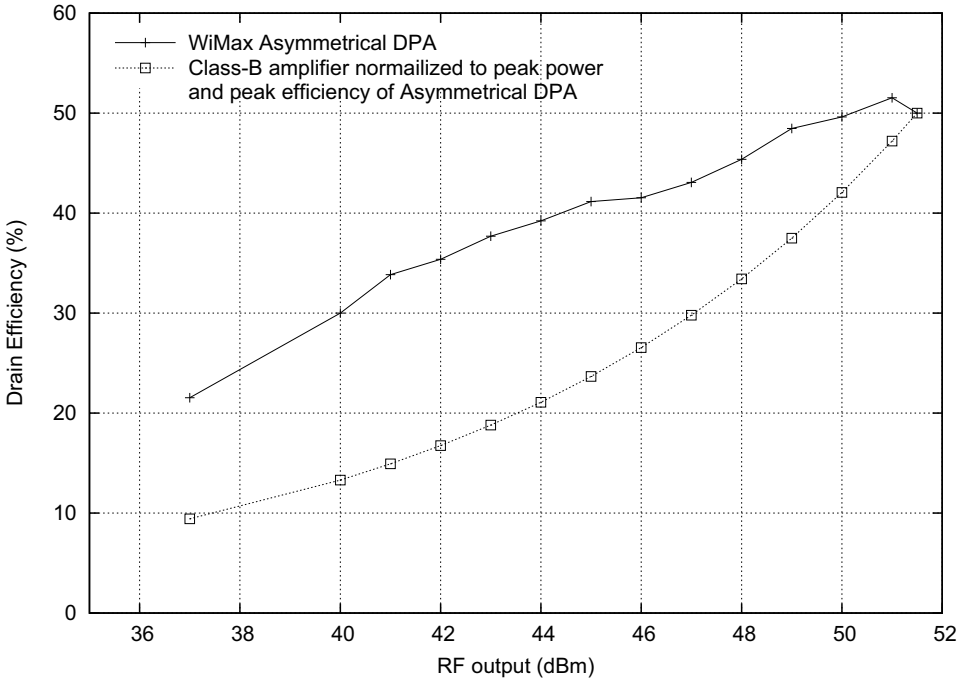


Figure 7.7: Measured drain efficiency of the WiMAX Asymmetrical Doherty Base Station PA with CW at 2.555GHz, comparison with an ideal Class-B amplifier with peak efficiency and peak power normalized to the maximum efficiency and maximum power of this Doherty amplifier.

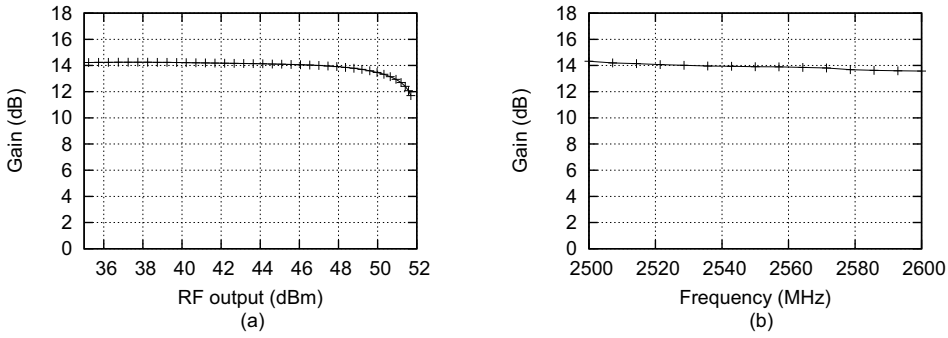


Figure 7.8: (a) Gain of the WiMAX asymmetrical DPA as a function of RF output power at CW of 2.555GHz. (b) Gain of the WiMAX asymmetrical DPA as a function of frequency at input RF power of 33 dBm.

amplifier with peak efficiency at 78%, the calculated efficiency is 28%. Thus, it can be concluded that a asymmetrical Doherty is already a good candidate to serve the

Table 7.1: Measured performance of the WiMAX asymmetrical Doherty with WiMAX signals (Center Freq=2.555 GHZ $\Delta Freq$ =10 MHz CF≈9.5 dB)

Parameters	Performance
Peak Power	52 dBm
Average Power	42 dBm
Drain Efficiency	36%
ACP @ 10 MHz	-30 dBc
EVM	-24 dB
Gain	14 dB

base-station market for 3G and beyond applications. In the coming sections, the focus is turned to the three-way Doherty, which can provide even better efficiencies for signals with this or higher CF.

7.3 Three-way Doherty Analysis

The analysis that was performed on the two-way DPA in section 7.1 can be extended to a three or even N-way DPA. As the whole analysis is based on working with black-box parameters, going to a N-way DPA simply implies dealing with larger matrices and solving a larger system of equations. To demonstrate the scalability of this approach, the following will show the analysis of a three-way DPA using the concepts described in section 7.1.

The first step in the analysis is again the reduction of the number of network ports by "absorbing" the output load as part of the network. In Figure 7.9, this reduces a four-port network problem to a three-port problem. Next the RF output voltages are expressed as currents of the main, peaking 1 and peaking 2 devices, through the three-port z-parameters, at each of the desired efficiency peaks (see Figure 7.10), namely at full power (F), first back-off point (B_1) and second back-off point (B_2);

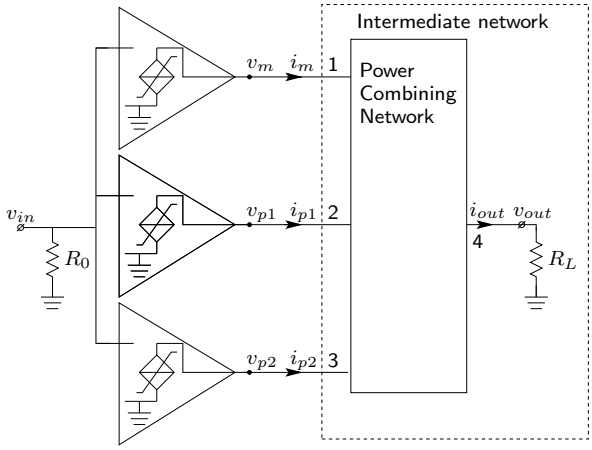


Figure 7.9: General three-way DPA with ideal voltage-controlled current sources and lossless power-combining network.

arriving thus at the following expression:

$$\begin{pmatrix} \begin{bmatrix} \upsilon_{m,F} \\ \upsilon_{p1,F} \\ \upsilon_{p2,F} \end{bmatrix} \\ \begin{bmatrix} \upsilon_{m,B_1} \\ \upsilon_{p1,B_1} \\ \upsilon_{p2,B_1} \end{bmatrix} \\ \begin{bmatrix} \upsilon_{m,B_2} \\ \upsilon_{p1,B_2} \\ \upsilon_{p2,B_2} \end{bmatrix} \end{pmatrix} = \begin{pmatrix} [z] & [0] & [0] \\ [0] & [z] & [0] \\ [0] & [0] & [z] \end{pmatrix} \begin{pmatrix} \begin{bmatrix} i_{m,F} \\ i_{p1,F} \\ i_{p2,F} \end{bmatrix} \\ \begin{bmatrix} i_{m,B_1} \\ i_{p1,B_1} \\ i_{p2,B_1} \end{bmatrix} \\ \begin{bmatrix} i_{m,B_2} \\ i_{p1,B_2} \\ i_{p2,B_2} \end{bmatrix} \end{pmatrix} \tag{7.35}$$

Where $[z]$ represents the three-port z-parameters of the intermediate network.

The subscripts $_m$, $_{p1}$ and $_{p2}$ refer to the main, peaking 1, and peaking 2 device, respectively, while the second subscripts $_F$, $_{B_1}$ and $_{B_2}$ refer to the voltage and current variables when the DPA is delivering full power, when the DPA is backed-off by a factor of k_1^2, and when the DPA is backed-off by a factor k_2^2, respectively.

Using equation (7.35), the z-parameters can be expressed in terms of the voltages and currents, which, due to the size of these expressions, are intentionally not included in this thesis.

Just like in the two-way DPA, the next step is to determine all the voltage and current variables, to uniquely determine the z-parameters of the intermediate three-port network. To maximize the efficiency at full power and the two other back-off points (B_1) and (B_2), the RF voltage amplitude at the output of the relevant devices (main, peaking 1 and peaking 2) has to be maximized at these points. In addition, one should ensure that, for high efficiency, the devices are only activated ($i \neq 0$) beyond the proper back-off point. Consequently, assuming class-B operation, the following variables are already fixed due to these considerations:

$$|v_{m,F}| = |v_{m,B_1}| = |v_{m,B_2}| = V_{dc} \tag{7.36}$$
$$|v_{p1,F}| = |v_{p1,B_1}| = V_{dc} \tag{7.37}$$
$$|v_{p2,F}| = V_{dc} \tag{7.38}$$
$$i_{p1,B_2} = 0 \tag{7.39}$$
$$i_{p2,B_1} = i_{p2,B_2} = 0 \tag{7.40}$$

The following shows the nine unknown variables left to solve, for the three-way Doherty power amplifier:

$$i_{m,F},\ i_{m,B_1},\ i_{m,B_2},\ i_{p1,B_1},\ i_{p2,F},\ v_{p1,B_2},\ v_{p2,B_1},\ v_{p2,B_2} \tag{7.41}$$

Again, the output power of the three-way DPA, assuming a lossless power-combining network, can be expressed as a function of the voltage and current variables.

$$p_{out} = \frac{1}{2}\,\mathbb{R}e\left(v_m i_m^* + v_{p1} i_{p1}^* + v_{p2} i_{p2}^*\right) \tag{7.42}$$

Here too, high efficiency is only achieved when the current and voltage related to a specific device (e.g., $v_{m,F}$ and $i_{m,F}$) are in phase, yielding maximum (real) power delivered to the load. Next the power back-off/full-power ratios are specified as:

$$\begin{aligned} k_1^2 &= \frac{p_{out,B_1}}{p_{out,F}} \\ &= \frac{\mathbb{R}e\left(v_{m,B_1} i_{m,B_1}^* + v_{p1,B_1} i_{p1,B_1}^*\right)}{\mathbb{R}e\left(v_{m,F} i_{m,F}^* + v_{p1,F} i_{p1,F}^* + v_{p2,F} i_{p2,F}^*\right)} \end{aligned} \tag{7.43}$$

$$\begin{aligned} k_2^2 &= \frac{p_{out,B_2}}{p_{out,F}} \\ &= \frac{\mathbb{R}e\left(v_{m,B_2} i_{m,B_2}^*\right)}{\mathbb{R}e\left(v_{m,F} i_{m,F}^* + v_{p1,F} i_{p1,F}^* + v_{p2,F} i_{p2,F}^*\right)} \end{aligned} \tag{7.44}$$

Furthermore, assuming a linear dependency of the current source on the RF input voltage when the active device is turned on, the following current relationships can be written:

$$i_{m,B_1} = i_{m,F} \cdot k_1 \tag{7.45}$$
$$i_{m,B_2} = i_{m,F} \cdot k_2 \tag{7.46}$$
$$i_{p1,B_1} = i_{p1,F}\left(\frac{k_1 - k_2}{1 - k_2}\right) \tag{7.47}$$

At this stage, similar to the two-way DPA, the reciprocal properties of the intermediate network ($z_{12} = z_{21},\ z_{23} = z_{32},\ z_{13} = z_{31}$) are used to complete the set of equations to solve for the 9 unknowns.

Using equation (7.42)–(7.47), the 9 unknown variables listed in (7.41) can be solved for specified back-off levels k_1, k_2 and phase relations between the main and peaking devices. The z-parameters of the intermediate network is now fully determined. Contrary to the two-way DPA analysis, here we present the intermediate network in z-parameters (instead of s) since the resulting equations are more compact and interpretable.

$$z_{11} = \frac{V_{dc}^2}{2 \cdot p_{out,max} \cdot k_2^2} \tag{7.48}$$

$$z_{12} = z_{21} = \frac{j\,V_{dc}^2}{2 \cdot p_{out,max} \cdot k_1 \cdot k_2} \tag{7.49}$$

$$z_{23} = z_{32} = \frac{jV_{dc}^2}{2 \cdot p_{out,max} \cdot k_1 \cdot (k_2 - 1)} \tag{7.50}$$

$$z_{13} = z_{31} = z_{22} = z_{33} = 0 \tag{7.51}$$

The current source relations are also determined by the above choices and are as follows:

$$i_m = 2 \cdot k_2 \cdot p_{out,max} \frac{v}{V_{dc}} \tag{7.52}$$

$$i_{p1} = \begin{cases} \frac{2 \cdot p_{out,max}}{V_{dc}} k_1 (v - k_2) & \text{if } v \geq k_2, \\ 0 & \text{if } v < k_2. \end{cases} \tag{7.53}$$

$$i_{p2} = \begin{cases} \frac{2 \cdot p_{out,max}}{V_{dc}} (1 - k_2)(v - k_1) & \text{if } v \geq k_1, \\ 0 & \text{if } v < k_1. \end{cases} \tag{7.54}$$

The last step is to reintroduce the fourth port (to which the resistive load is connected) back into the analysis. Taking the reciprocity condition into account, the unknowns at this stage are s_{14}, s_{24}, s_{34} and s_{44} (the s-parameters are complex in general, each with a magnitude and phase). Expanding equation (7.14) and (7.15) for the lossless four-port situation, ten equations are obtained (4 from equation (7.14) and 6 from (7.15)). Since there are more equations than unknowns, this makes the system overdetermined. Analytical analysis at this point is cumbersome and tedious as there will probably be many solutions to this system of equation. As such we have resorted to the use of numerical techniques to solve for the s-parameters of the last port.

For k_1 =0.5 and k_2 =0.25, and assuming the current conditions of (7.52)-(7.54) with V_{dc} =1V and $p_{out,max}$ =1 W, the four-port s-matrix normalized to 1Ω is found to be as follows:

$$s = \begin{bmatrix} 0.864 & -j0.195 & 0.260 & 0.372 - j0.092 \\ -j0.195 & 0.561 & -j0.585 & 0.133 + j0.536 \\ 0.260 & -j0.585 & -0.220 & -0.714 + j0.177 \\ 0.372 - j0.092 & 0.133 + j0.536 & -0.714 + j0.177 & -0.074 + j0.039 \end{bmatrix} \tag{7.55}$$

7.3.1 Verification of the three-way Doherty Power Amplifier

For the verification of the above three-way analysis, the schematic of Figure 7.9 with the s-matrix above for the power-combining network are used (where k_1 =0.5, k_2 =0.25). Again, the main and peaking devices are modeled with SDDs, using the current dependency given in (7.52)-(7.54). The supply voltage of the active devices is set to 1 V, while the maximum current of the main and peaking devices is chosen such that the total maximum output power equals 1W. The simulation results from ADS are given in Figure 7.10.

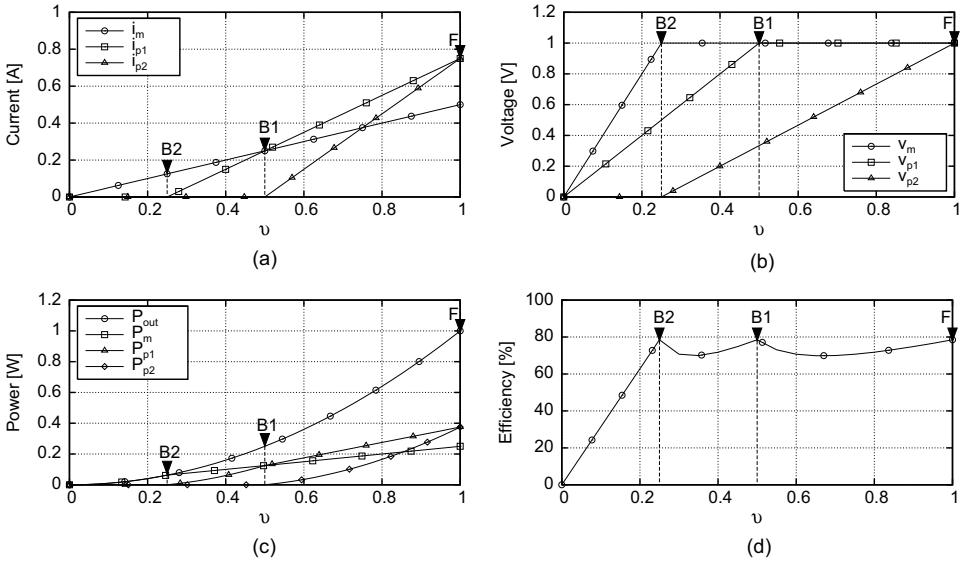


Figure 7.10: ADS simulation results based on the schematic of Figure 7.9 as validation of the three-way DPA analysis. The power-combining network is calculated for k_1 =0.5 and k_2 =0.25. All devices use shorts for the higher harmonics (class B). The design is normalized to 1-W output power and 1-V RF input voltage. (a) Output current for main and peaking devices versus normalized RF input voltage. (b) Resulting RF voltage at the output of the main and peaking devices as a function of the normalized RF input voltage. (c) Power contributions and resulting total power of the main and peaking devices. (d) Resulting efficiency as a function of the normalized RF input voltage.

In Figure 7.10 a, the amplitude of the main and peaking device current are plotted versus input drive voltage. Figure 7.10 b shows the resulting RF voltage amplitude for the output of the main and peaking devices. Figure 7.10 c gives the power delivered by the main and peaking devices, as well as the total power. In Figure 7.10d, the resulting efficiency is given versus input voltage.

7.3.2 Analyzing the existing three-way Doherty solution

As was earlier mentioned, the advantage of this analysis strategy, is that the designer is not committed to a "fixed" power combining network. The designer is free to experiment with the current drive profiles, the phase relationships between them, and even in the optimization process to solve for the s-parameter of the load port, to arrive at an output power-combining network that meets the application needs. As such, it is an interesting exercise in this section to apply this analysis technique to the well known classical power-combining network topology for the three-way DPA, and see if this uncovers new insights for this well known solution.

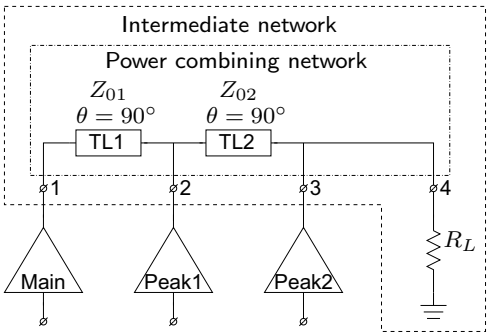


Figure 7.11: Principle schematic three-way DPA setup using the classical two-stage quarter-wave length transmission line implementation. The network parameters for k_1 =0.5 and k_2 =0.25, assuming 1-V supply voltage and 1-W output power are, Z_{o1} =4, Z_{o2} =1 and R_L =0.5.

Figure 7.11 shows the well-known two-stage quarter-wavelength transmission line network solution [54] with parameters Z_{o1} and Z_{02} representing the characteristic impedances of the $\frac{\lambda}{4}$ lines and with load R_L. The values of these parameters can be calculated using the back-off ratios k_1 and k_2 [54]:

$$Z_{o2} = \frac{R_L}{k_1} \tag{7.56}$$

$$Z_{o1} = \frac{R_L}{k_1 k_2} \tag{7.57}$$

By reincorporating the load into an intermediate power-combining network, a three-port z-matrix is obtained. Furthermore, by inspection and by making use of the network reciprocity, the following expression is obtained:

$$z = \begin{bmatrix} z_{11} & z_{12} & z_{13} \\ z_{12} & 0 & 0 \\ z_{13} & 0 & z_{33} \end{bmatrix} \tag{7.58}$$

Where

$$z_{11} = \left(\frac{Z_{o1}}{Z_{o2}}\right)^2 R_L \tag{7.59}$$

$$z_{12} = -jZ_{o1} \tag{7.60}$$

$$z_{13} = -\frac{Z_{o2}}{Z_{o1}R_L} \tag{7.61}$$

$$z_{33} = R_L \tag{7.62}$$

To maximize the efficiency under full power and the two back-off conditions, the voltage and current variables are chosen as in the prior analysis. In this classical DPA implementation, the phase relations of the main and peaking amplifiers are determined by the network topology and expressed relative to the main amplifier, yielding:

$$\upsilon_{m,F} = \upsilon_{m,P1} = \upsilon_{m,B2} = V_{dc} \tag{7.63}$$

$$\upsilon_{p1,F} = \upsilon_{p1,B1} = -jV_{dc} \tag{7.64}$$

$$\upsilon_{p2,F} = -V_{dc} \tag{7.65}$$

$$i_{p1,B2} = 0 \tag{7.66}$$

$$i_{p2,B2} = i_{p2,B1} = 0 \tag{7.67}$$

At this juncture, substituting the above variables together with the z-parameters of the network into equation (7.35), and solving for the following nine remaining unknowns results in:

$$i_{m,F} = i_{m,B1} = -j\frac{V_{dc}}{z_{12}} \tag{7.68}$$

$$i_{m,B2} = \frac{V_{dc}}{z_{11}} \tag{7.69}$$

$$i_{p1,F} = \frac{V_{dc}}{z_{12}}\left[\left(1 - \frac{z_{13}}{z_{33}}\right) + j\left(\frac{z_{11}}{z_{12}} + \frac{z_{13}^2}{z_{33}z_{21}}\right)\right] \tag{7.70}$$

$$i_{p1,B1} = \frac{V_{dc}}{z_{12}}\left(1 + j\frac{z_{11}}{z_{12}}\right) \tag{7.71}$$

$$i_{p2,F} = \frac{V_{dc}}{z_{33}}\left(1 - j\frac{z_{13}}{z_{12}}\right) \tag{7.72}$$

$$\upsilon_{p1,B2} = V_{dc}\frac{z_{12}}{z_{11}} \tag{7.73}$$

$$\upsilon_{p2,B1} = -j\,V_{dc}\frac{z_{13}}{z_{12}} \tag{7.74}$$

$$\upsilon_{p2,B2} = V_{dc}\frac{z_{13}}{z_{11}} \tag{7.75}$$

Equation (7.68) shows that $i_{m,F} = i_{m,B_1}$; this implies that the current provided by the main device should be equal in magnitude when the three-way DPA is delivering

full power, as well as when the DPA is at its first back-off level. This, in itself, is a remarkable conclusion, which, to our knowledge, has never before been emphasized in literature. Note that this conclusion has a significant impact on physical implementations of the three-way DPA since it requires the main device to ramp up linearly with the input RF voltage up to the first back-off point, then saturate, and remain saturated and constant up to full power. This undoubtedly causes linearity problems in practical three-way DPA implementations.

$$i_m = \begin{cases} \frac{2\,p_{out,max}}{V_{dc}} k_2\, v & \text{if } v \le k_1, \\ \frac{2\,p_{out,max}}{V_{dc}} k_2\, k_1 & \text{if } v > k_1. \end{cases} \tag{7.76}$$

$$i_{p1} = \begin{cases} \frac{2\,p_{out,max}}{V_{dc}} k_1 (v - k_2) & \text{if } v \ge k_2, \\ 0 & \text{if } v < k_2. \end{cases} \tag{7.77}$$

$$i_{p2} = \begin{cases} \frac{2\,p_{out,max}}{V_{dc}} (v - k_1) & \text{if } v \ge k_1, \\ 0 & \text{if } v < k_1. \end{cases} \tag{7.78}$$

To verify the above, an ADS simulation was performed using the classical three-way configuration of Figure 7.11 for k_1 =0.5 and k_2 =0.25, yielding R_L =0.5, Z_{o1} =4, and Z_{o2} =1 for the network parameters. To implement the desired dependency of the current sources, SDDs are again use to represent the main and peaking devices, with characteristic shown in equation (7.76)-(7.78). Again, the amplifier is dimensioned for 1 W output power and the RF input voltage is normalized to the supply voltage of 1 V.

The ADS harmonic balance simulation results are given in Figure 7.12. Note the saturation of the current of the main device after the first back-off point is reached. Fulfilling this condition provides a similar efficiency performance as found in the general three-way analysis. When the saturation condition for the current of the main device is violated (e.g., see dashed lines in Figure 7.12 a and 7.12 d), both the efficiency and linearity significantly degrade in power back-off. This poses a practical implementation issue, which is further discussed in the following section.

7.4 Practical issues in implementing the three-way DPA

In Sections 7.1 and 7.3, the Doherty amplifiers were analyzed for their requirements in terms of power-combining network and current profile versus input voltage. As was demonstrated in section 7.2, in the case of the two-way DPA implementations; the desired current profile for the main and peaking amplifier are approximated by using class-AB biasing of the main device and class-C biasing for the peaking device, in conjunction with the asymmetrical power splitter. Although these bias conditions yield only an approximation of the desired current profile, reported implementations show an improved efficiency at an acceptable linearity level compared to single-stage amplifiers.

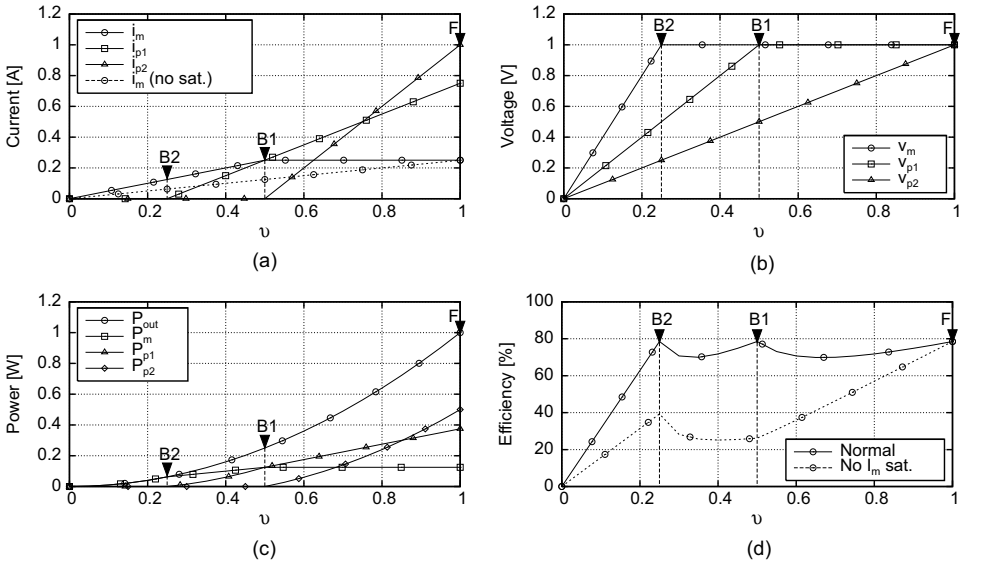


Figure 7.12: ADS simulation results based on the classical three-way DPA. The power-combining network is calculated for k_1 =0.5 and k_2 =0.25. All devices use shorted higher harmonics resulting in a class-B-like operation. The main device saturates at the first back-off point. The design is normalized to 1-W output power and 1-V RF input voltage. (a) Output current for main and peaking devices versus normalized RF input voltage. (b) Resulting voltage amplitude at the output of the main and peaking devices as a function of the normalized RF input voltage. (c) Power contributions and resulting total power of the main and peaking devices. (d) Resulting efficiency as a function of the normalized RF input voltage.

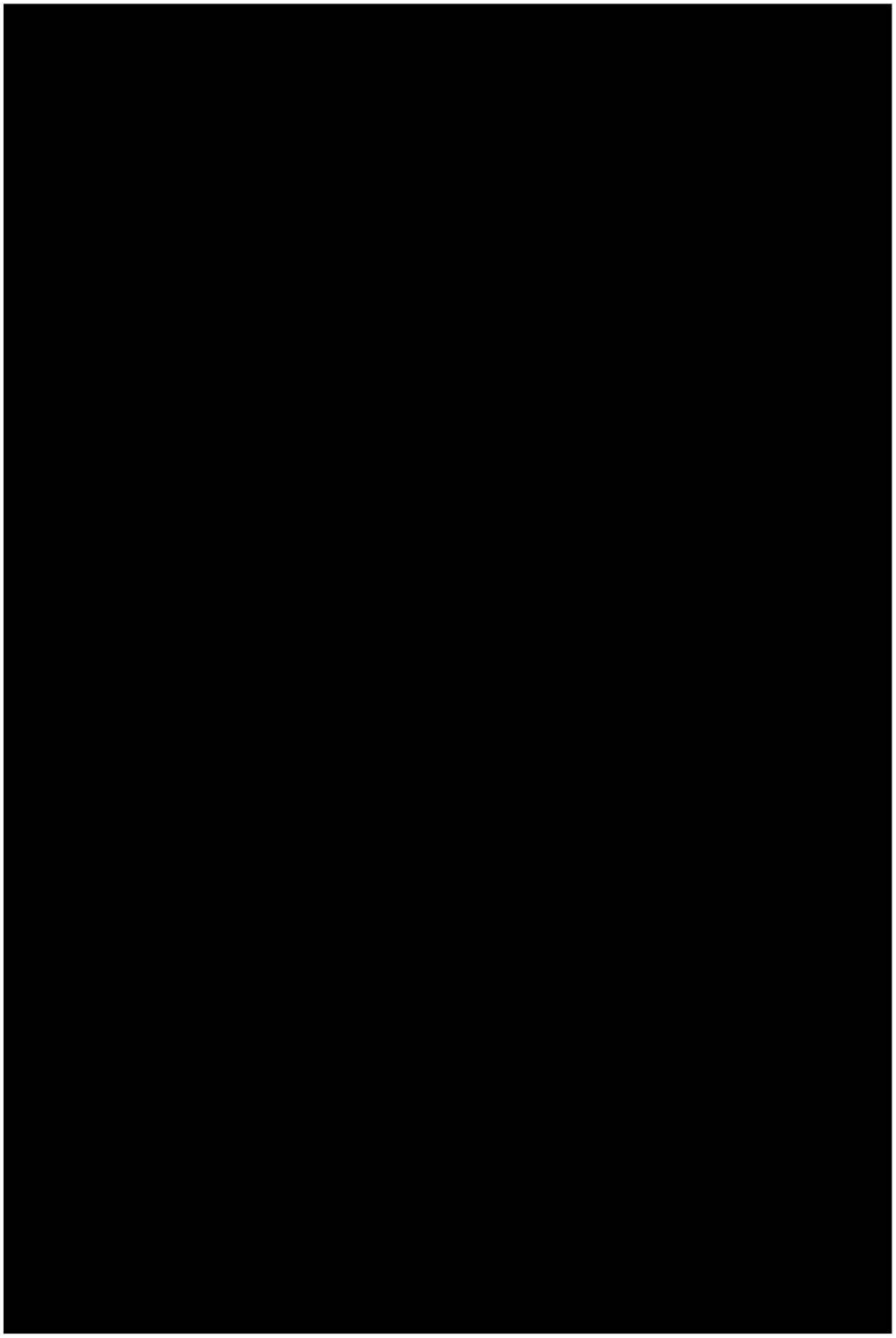
A more detailed inspection of these results demonstrates that most DPA implementations compromise the bias conditions of the main and peaking devices to improve linearity at the cost of efficiency. This is one of the main reasons why reported efficiency numbers for DPA are significantly lower than their theoretical optimum performance.

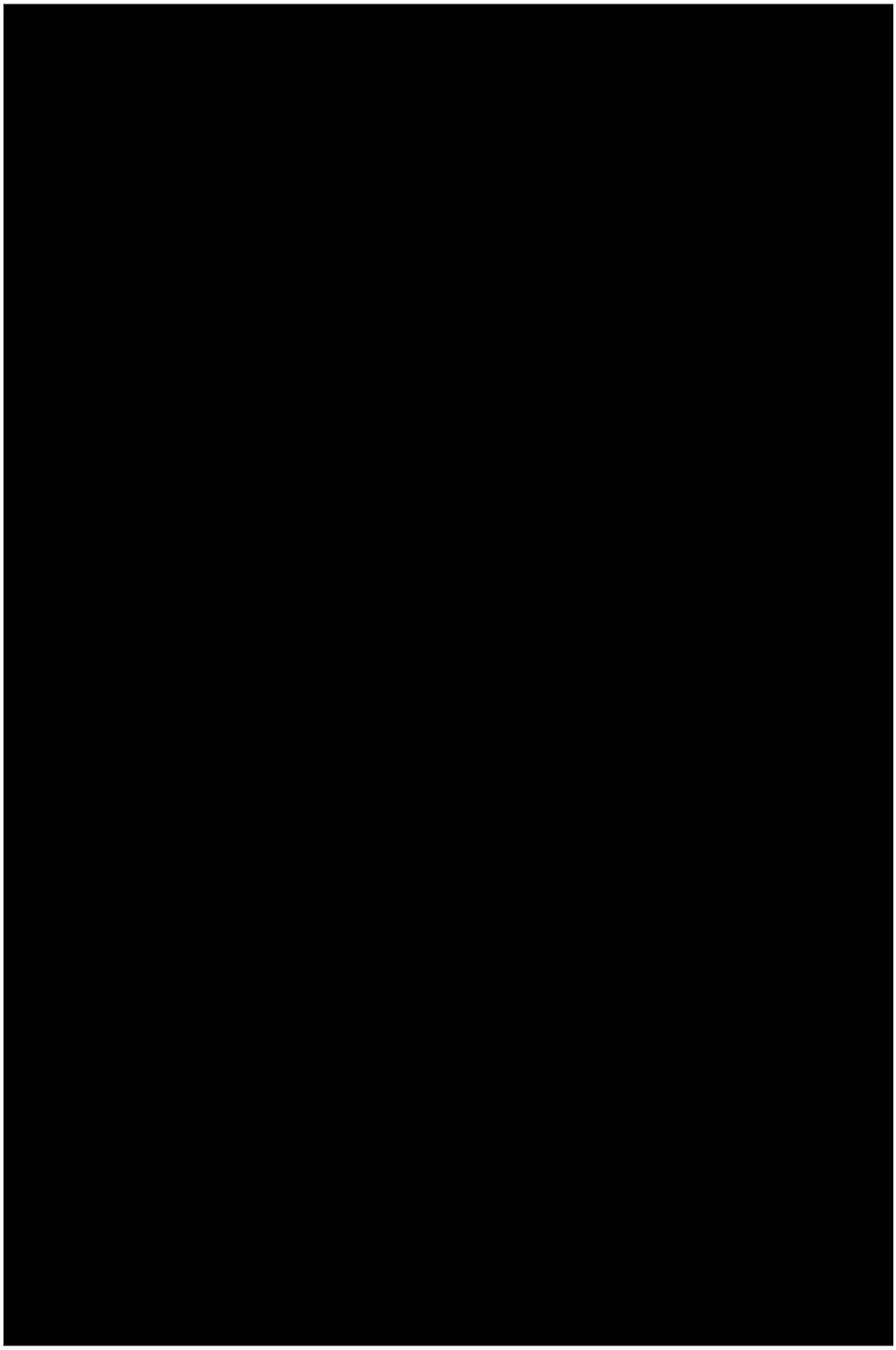
For three-way DPA implementations, the situation is even worse since now three active devices must exhibit a current versus RF input voltage that closely matches the desired current profile. An additional complication is found for the classical three-way DPA approach. Here the current of the main device should ramp up linearly with the RF input voltage and saturate at the first back-off point to a constant amplitude value. It is obvious that such a requirement is not easy to fulfill using conventional DPA techniques in a controllable and reproducible manner.

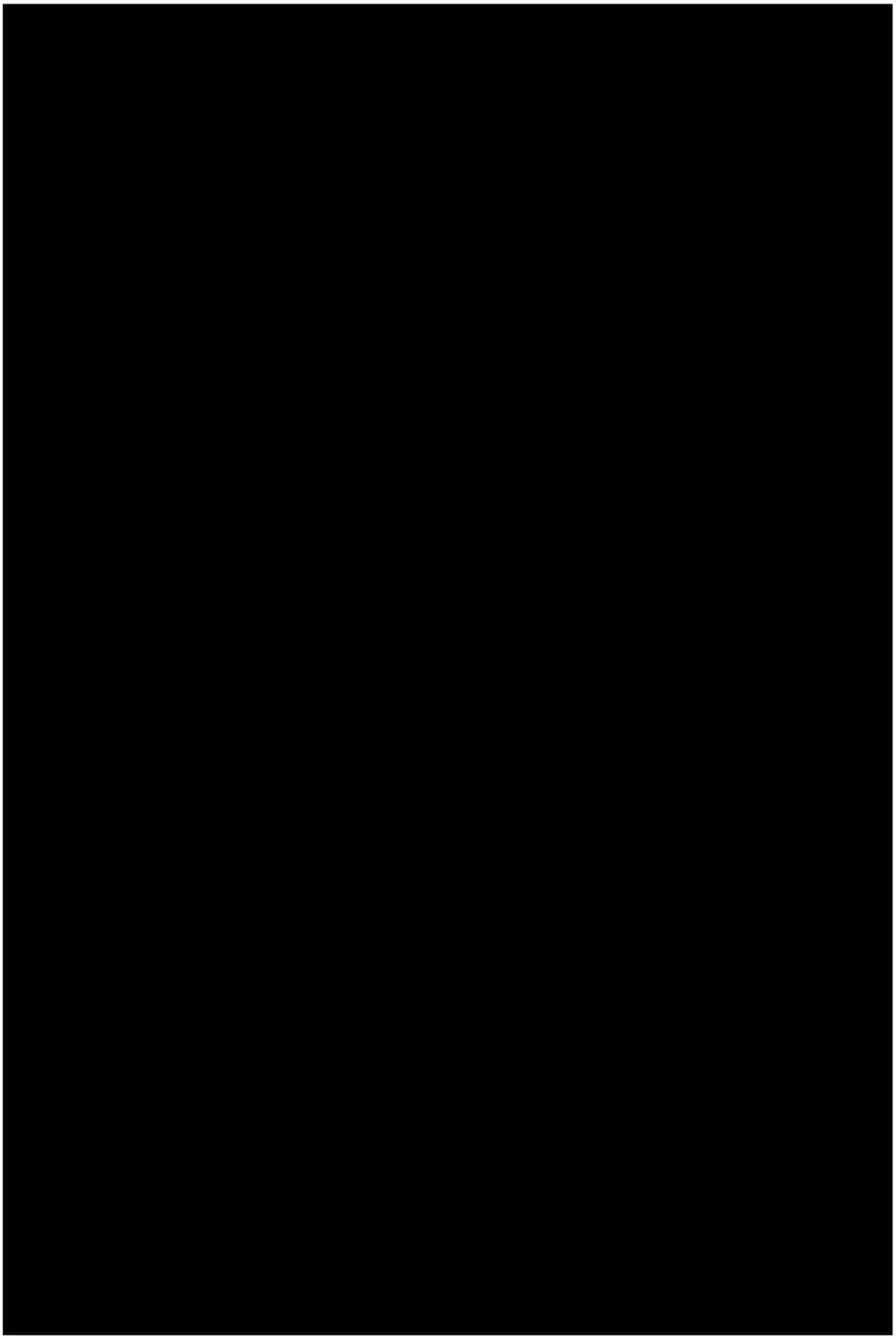
In the following sections, two radical design concepts of a three-way DPA will be presented together with their simulation and measured verification; these concepts allow the designer to overcome the mentioned implementation difficulties and allow for true three-way DPA operation. The design concept in Section 7.5 is based on having the linear drive profiles shown in Figure 7.10 a and synthesizing a new power-

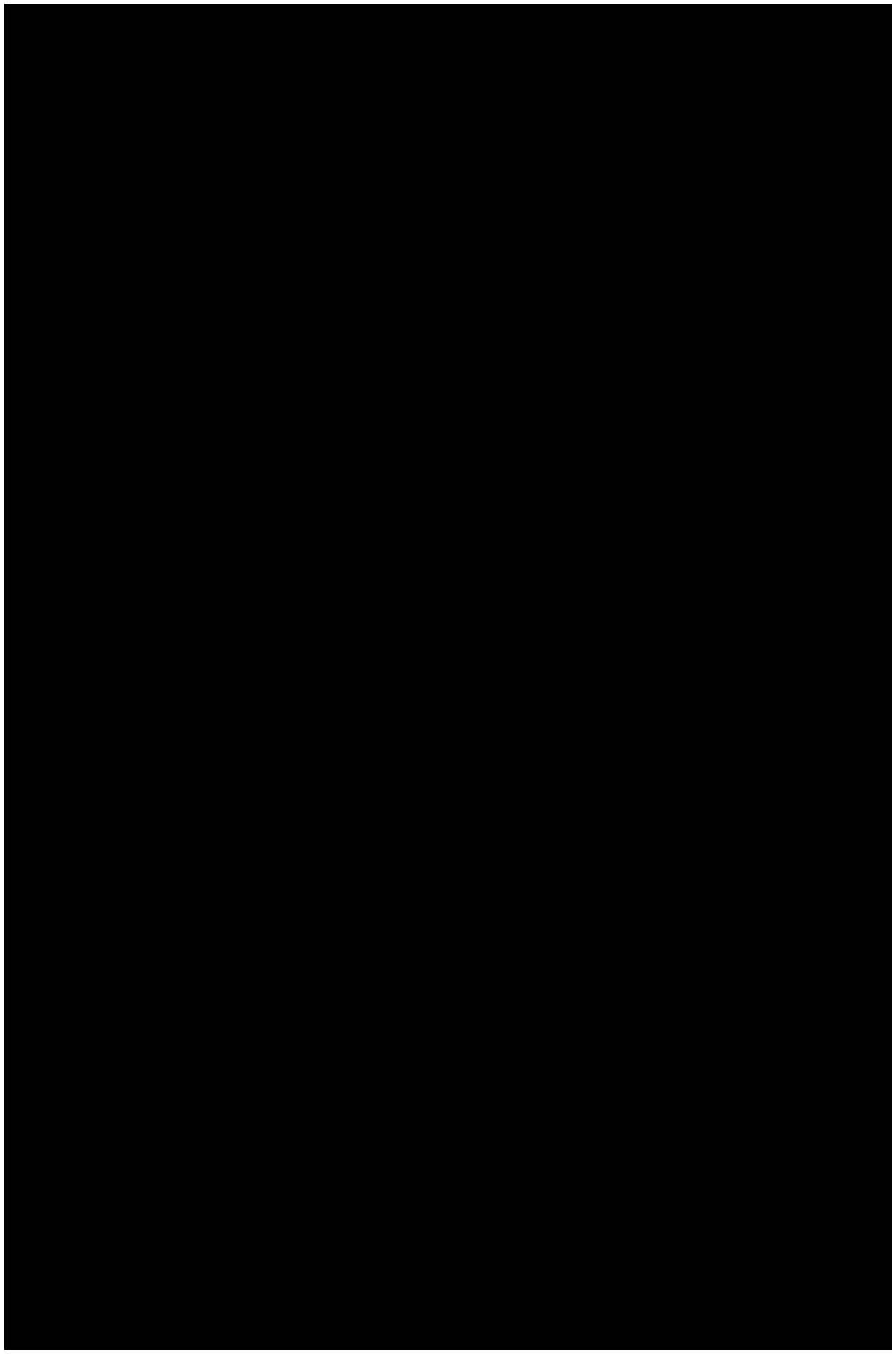
combining network from the described black-box parameter approach. In the second design concept (Section 7.6) a mixed-signal approach to individually control the inputs of the main and peaking devices is introduced. This approach facilitates arbitrary input drive conditions and "free-to-choose" biasing condition (e.g., class-B for both the main and peaking devices) since the intended current dependencies versus input power are no longer obtained through the manipulation of dc bias points of the output stages.

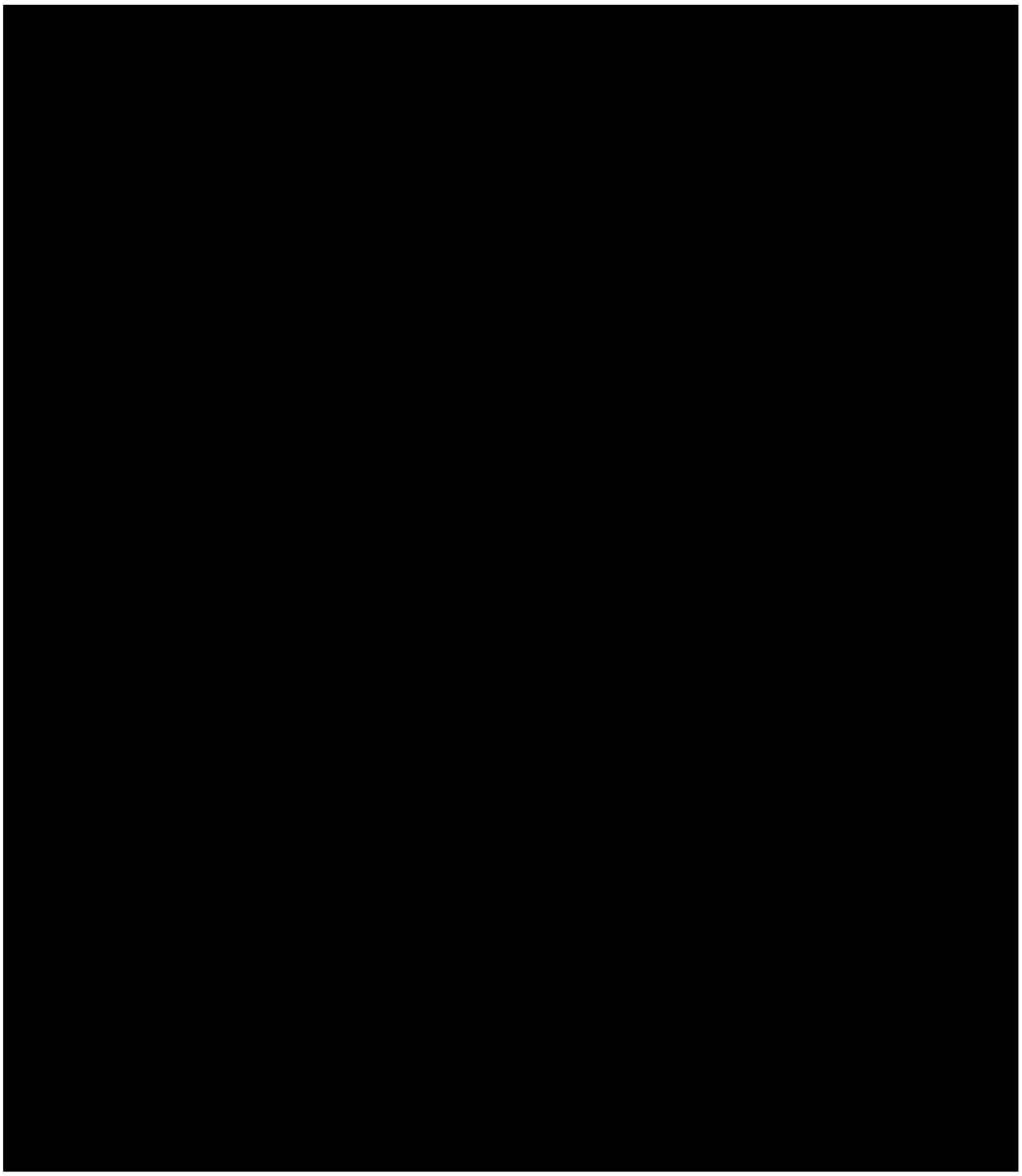
7.5 The "Linear" Three-Way Doherty





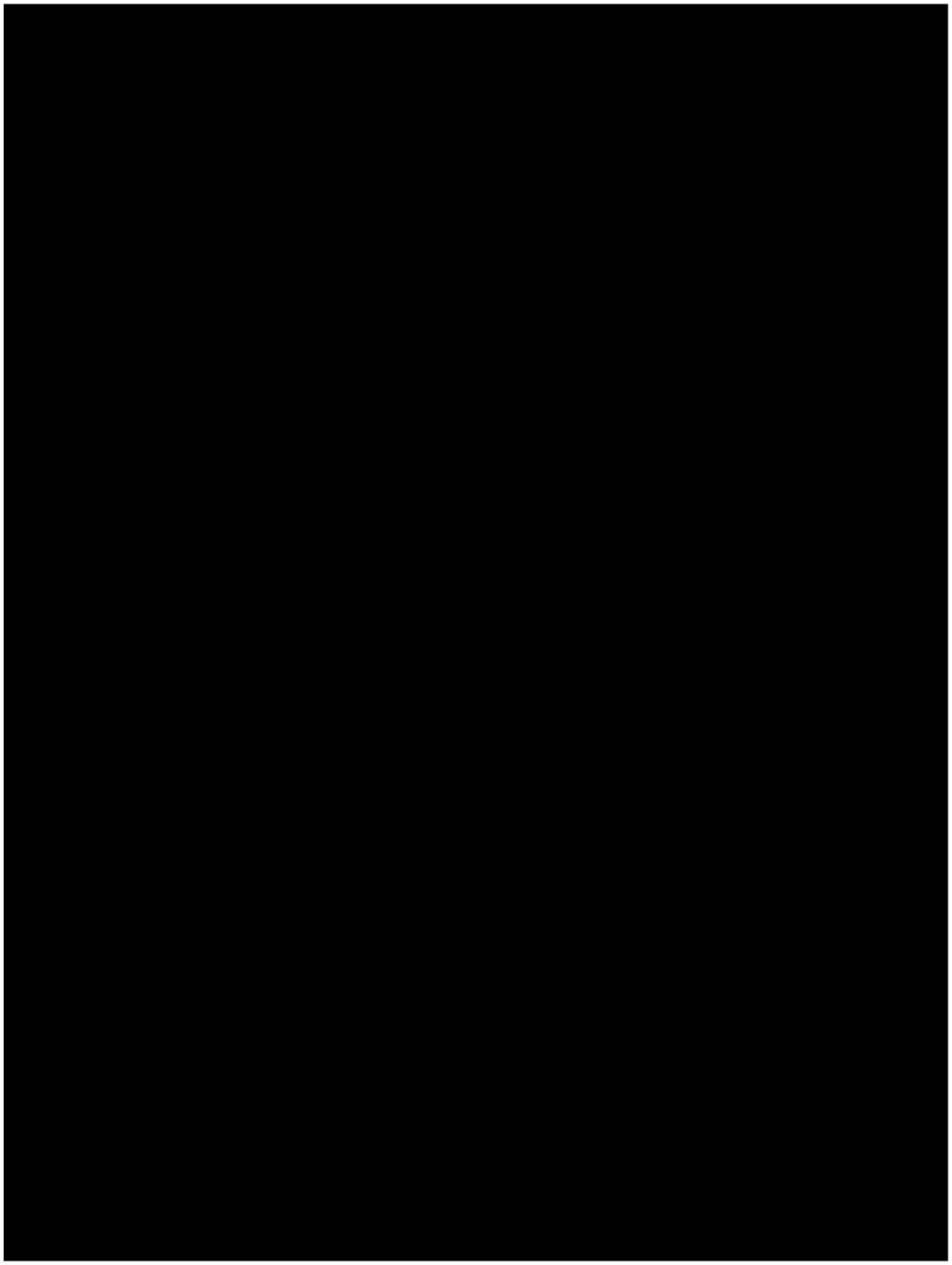






7.6 The Mixed-Signal Three-Way Doherty

As was clear from Section 7.5, a full analogue solution to the three-way Doherty is a very complex design task. A lot of fine tuning is required for the bias voltages of the devices, the lengths of the offset transmission lines and the input power splitter to get good performance out of the Doherty amplifier. To overcome these difficulties, a "mixed-signal approach" to the Doherty amplifier is introduced in this section. This approach can be considered as having a digital power splitter at the input, and this facilitates the independent optimization of the drive condition for Main, Peak 1 and



Peak 2 resulting in maximum efficiency. Furthermore, no more offset transmission lines are required at the input to obtain the correct power combining at the output, as this can be easily taken care of in the digital domain.

To highlight the advantages of the mixed-signal approach, a three-way DPA using the traditional power combining network (see Section 7.3.2) will be designed. Note that this choice allows the verification of the conclusions made in Section 7.3.2 with regards to the required current drive conditions for optimum three-way DPA behavior, as the mixed signal concept provides for the design of any arbitrary drive profile. Furthermore, this verification exercise will give a clear comparison with existing DPA solutions. Note that this work is done in close collaboration with another PhD project (performed by Marco Pelk) which investigates innovative RF measurement setup and the work has been published in [56].

7.6.1 Design of the Mixed-Signal three-way Doherty

The mixed-signal Doherty that will be prototyped here is targeted for UMTS base-stations at a center frequency of 2.14 GHz (same frequency as the "Linear" three-way Doherty). The general design goal is to aim for a 100 Watt (peak power) amplifier with gains above 10 dB and as high efficiency as possible when driven with a WCDMA signal.

At present, LDMOS is still the preferred technology for base station power amplifiers due to its high reproducibility, good power handling, high-linearity and gain, but most importantly, due to its low cost. However, over time it is expected that wide-bandgap materials like GaN will challenge the position of silicon LDMOS technology, since GaN can offer higher power densities and higher gain at microwave frequencies. To investigate these prospects, this Doherty prototype will be designed with GaN devices and this will give a good indication of the potential of GaN technology for three-way Doherty power amplifiers with improved efficiency in far back-off operation to better serve signals with very high crest factors.

Just like in two-way Asymmetrical Doherty and the "linear" three-way Doherty, the first step in the design procedure is to determine the relationship between the device sizes and the back-off powers B_1 and B_2 where the efficiency peaks of the three-way Doherty occurs. Equation (7.92) which can be derived from the required drive profiles (see equation (7.76)-(7.78)) and the power relationships in back-off, gives this relationships.

$$k_1 = \frac{1 + m_1}{1 + m_1 + m_2}, \qquad k_2 = \frac{1}{1 + m_1} \tag{7.92}$$

If the back-off points B_1 and B_2 are desired to be at -6 dB ($k_1 = 0.5$) and -12 dB ($k_2 = 0.25$) respectively, the resulting required device size ratio will be $1{:}m_1{:}m_2$=1:3:4. Unfortunately, in practical situations these device sizes are not always available. Also from the GaN HEMT devices donated by CREE for this work, there were only two power ratings available, namely: 15 W and 45 W. Consequently, the configuration that approximates the intended device ratios best is 1:3:3. Substituting this into equation (7.92) results in high-efficiency operation points being located at -4.9 dB and

-12 dB power back-off respectively. Note that just like in the all the other Doherty concepts, this device size selection will fix also the load modulation ratios of the main (Υ_m) and peak1 device (Υ_{p1}). The following equations give these relationship.

$$\Upsilon_m = \frac{(1+m_1)^2}{1+m_1+m_2} \tag{7.93}$$

$$\Upsilon_{p1} = \frac{m_1(1+m_1+m_2)}{m_1^2+m_1-m_2} \tag{7.94}$$

This yields $\Upsilon_m = 2.3$ and $\Upsilon_{p1} = 2.3$, which are very favorable numbers compared to the situation of the "linear" three-way Doherty ($\Upsilon_m = 3$ and $\Upsilon_{p1} = 4$) solution discussed in Section 7.5. The next step is to find the proper trade-offs between the characteristic impedance and the loading impedances (impedance which the active device has to be matched to) required by the amplifier cells. For this purpose, the following equations which relates the characteristic impedance of the lines in the power combining network and the loading impedance of the Main, Peak 1 and Peak 2 are also derived for the classical three-way DPA power combiner.

$$Z_{01} = \sqrt{Z_{m,F} \cdot Z_{p1,F} \cdot m_1} \tag{7.95}$$

$$Z_{02} = \sqrt{\frac{Z_{p2,F} \cdot Z_{p1,F} \cdot m_1 \cdot m_2}{(1+m_1)^2}} \tag{7.96}$$

$$Z_{03} = \sqrt{\frac{Z_{p2,F} \cdot R_L \cdot m_2}{1+m_1+m_2}} \tag{7.97}$$

Where $Z_{m,F}$, $Z_{p1,F}$ and $Z_{p2,F}$ are the impedances which we need to match the active device to for maximum power out, and they are given by the following equations.

$$Z_{m,F} = \frac{Z_{03}^2 \cdot Z_{01}^2 \cdot (1+m_1+m_2)}{Z_{02}^2 \cdot R_L \cdot (1+m_1)^2} \tag{7.98}$$

$$Z_{p1,F} = \frac{Z_{02}^2 \cdot R_L \cdot (1+m_1)^2}{Z_{01}^2 \cdot m_1 \cdot (1+m_1+m_2)} \tag{7.99}$$

$$Z_{p2,F} = \frac{Z_{01}^2 \cdot (1+m_1+m_2)}{R_L \cdot m_2} \tag{7.100}$$

The following give the equations of the impedances that are offered to the Main and Peak 1 amplifier in back-off, as a result of the load-modulation effect of the network.
At back-off 1:

$$Z_{m,B1} = \frac{Z_{03}^2 \cdot Z_{01}^2 \cdot (1+m_1+m_2)}{Z_{02}^2 \cdot R_L \cdot (1+m_1)^2} \tag{7.101}$$

$$Z_{p1,B1} = \frac{Z_{02}^2 \cdot R_L \cdot (1+m_1)^2}{Z_{01}^2[(1+m_1)^2 - (1+m_1+m_2)]} \tag{7.102}$$

At back-off 2:

$$Z_{m,B2} = \frac{Z_{03}^2 \cdot Z_{01}^2}{Z_{02}^2 \cdot R_L} \tag{7.103}$$

As was also discussed in the other Doherty designs, Section 7.5 and 7.2, the characteristic impedance of the lines can be influenced by selecting convenient loading conditions of the amplifier cells at the full power point ($Z_{m,F}$,$Z_{p1,F}$ and $Z_{p2,F}$). Selecting too high impedance levels for the amplifier cells will complicate the output matching due to the large impedance transformation ratios required. Choosing too small values results in unpractical widths for the quarter-wave lines of the power-combining network. In this design, the loading conditions shown in Figure 7.18 were used for the amplifier cells, providing the best trade-off in amplifier cell matching and impedance levels of the quarter-wave lines, namely: Z_{01} = 32.7 Ω, Z_{02} = 26.5 Ω and Z_{03} = 61.2 Ω. Note that these values can be easily implemented in the used substrate technology (Taconic, TLT 0.8 mm).

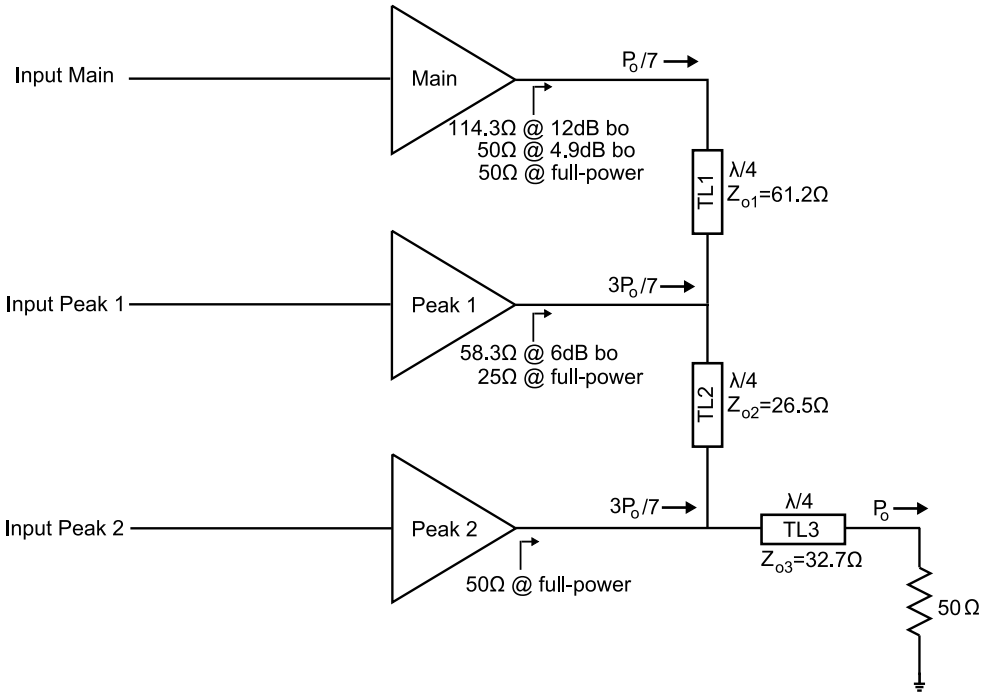


Figure 7.18: Schematic of the transmission line implementation of the "mixed-signal" three-way Doherty.

The remaining step is to design the matching network for the Main, Peak 1 and Peak 2, which should provide matching for optimum efficiency at all the loading conditions specified for the different power back-off levels, something that has been already discussed extensively in Section 7.5 and 7.2.

Figure 7.19 shows the final "mixed-signal" GaN three-way Doherty prototyped, designed with the above procedure.

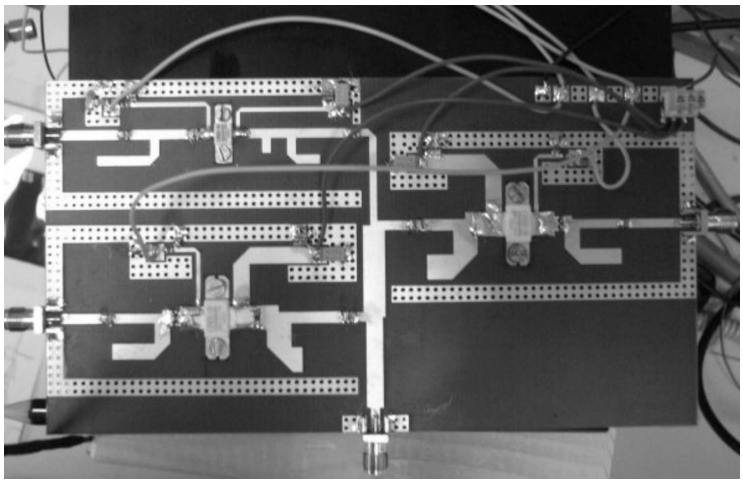


Figure 7.19: Hardware prototype of the "mixed-signal" GaN three-way Doherty.

7.6.2 Simulated Performance of the Mixed-Signal three-way Doherty

After completion of the hardware design, it is still required to find the optimum input drive conditions for the mixed-signal three-way Doherty to achieve maximum efficiency when driven with WCDMA signals. To address this problem, the single-tone DPA efficiency as function of power in the back-off region will be optimized. Note that this is a very effective approach, since once the probability distribution function (PDF) of a complex modulated signal is known, the average drain efficiency $\langle \eta_d \rangle$ can be calculated based on the PDF and the DPA single-tone efficiency versus power back-off [31], using equation (2.15) from Chapter 2.

In order to optimize the input drive signals of the DPA, also the phase relations (φ_{peak1} and φ_{peak2}) between the amplifier cells need to be determined (this is done with the offset transmission lines at the input for the "linear" three-way Doherty). In practice, the following three phase dependencies are considered:

1. Phases optimized at the full power condition.
2. Optimization at the second backoff point.
3. Phase optimization at all power levels.

Figure 7.20 shows the related simulated drain efficiencies. Also shown is the power probability distribution function (PDF) for the W-CDMA signal without de-cresting that will be used later in the efficiency considerations.

By comparing the results of Figure 7.20 with DPA results found in literature [57, 58], it can be concluded that most conventional DPA implementations with a passive input power splitter use phase relations which are optimized for high-efficiency at the maximum output power condition.

In order to make a careful selection of the optimum phase relations for driving the DPA, equation (2.15) is used with the efficiency curves of Figure 7.20 to calculate the W-CDMA efficiency performance for all three cases. The results of this calculation

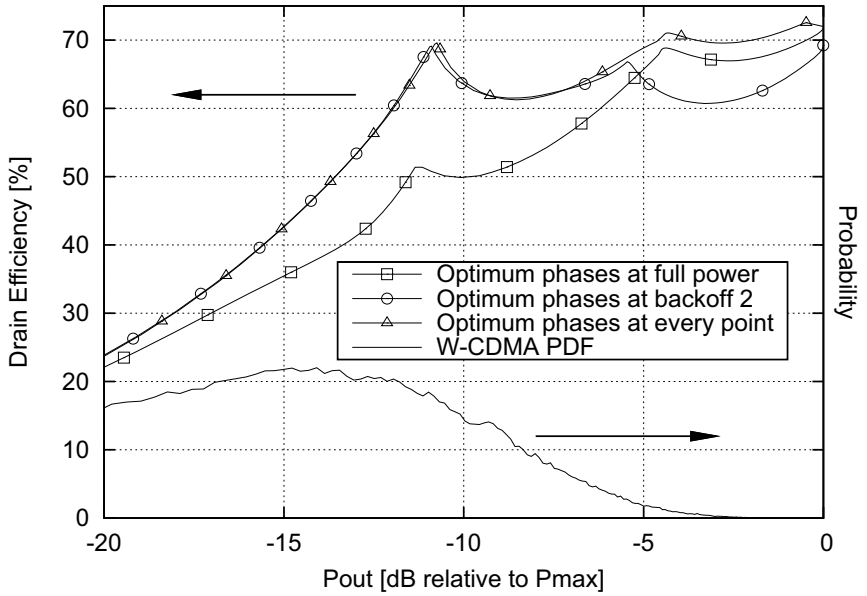


Figure 7.20: Simulated drain efficiency using three different phase optimization methods: (1) The relative phases for the peak1 and peak2 amplifiers optimized at full output power. (2) The relative phases for the peak1 and peak2 amplifiers optimized at the second back-off point. (3) The relative phases for the peak1 and peak2 amplifiers optimized at each power level. Also shown is the probability distribution function for a W-CDMA signal.

are presented in Table 7.3.

Table 7.3: Calculated W-CDMA Drain Efficiency Based on Single Tone Simulation Data for Three Different Cases

	Full Power	Back-off 2	Everywhere
η_d	54.4%	60.2%	61.2 %

Clearly, the situation where the phase angles are optimized at each power level yields the highest efficiency. A 1% reduction in efficiency is found for the case the efficiency at back-off point 2 is optimized. Worse performance is found for the situation where the phase DPA phase relations are only optimized at maximum output power.

It should be mentioned that for single-tone operation, phase optimization at each output power level is relatively easy to implement. However, when using complex modulated signals together with phase relations that change continuously as a function

of instantaneous power, this concept proves to be very difficult to implement. Thus to avoid too much computational overhead in signal processing for only 1% efficiency improvement, it is decided to restrict to the use of fixed but optimized phase relations for signals fed to the amplifier cells.

7.6.3 Mixed-signal three-way Doherty Experimental Verification

As mentioned above, this sort of three-input one-output PA concept is to the author's knowledge new, and a custom test setup has to be built and used to verify the above conclusions. The description and feature of this prototype setup can be found in the work of M.Pelk [56].

Single tone Measurement of the "mixed-signal" three-way Doherty

From the simulations in Section 7.6.2, a good estimate is achieved for the power levels of the input signals for the main, peak 1 and peak 2 amplifiers, as well for the relative phases to reach maximum PAE at the 12 dB back-off point (38 dBm). Using these initial values as starting point in the MATLAB measurement optimization, the optimum input powers and phase relations for maximum PAE are found. The same procedure is repeated at 4.9 dB back-off (45.1 dBm) and at full power (50 dBm). At each point optimizing the three input powers P_{main}, P_{peak1}, P_{peak2} and the two relative phases φ_{peak1}, φ_{peak2}. Linear interpolation is used to obtain values between these points. Note that this approach is almost equivalent to the use of optimized drive conditions at each power level, as discussed in the simulation section. The resulting PAE as function of normalized output power is plotted in Figure 7.21, indicating the upper bound for PAE that can be achieved with the "mixed-signal" GaN three-way Doherty. Note the rather exceptional PAE performance as function of back-off power.

As a second experiment, the more simplistic drive condition with constant phase relations between the amplifier cells is used. Doing so avoids the complication of continuously varying phases in combination with complex modulated signals as discussed in the simulation section. Consequently, the phase relations are kept constant, but their relative angles are optimized for maximum PAE at the 2nd power back-off point. The resulting PAE as function of back-off power is also plotted in Figure 7.21. Table 7.4 gives the resulting phases of φ_{peak1} and φ_{peak2} at the three high-efficiency points for both drive profiles.

It can be observed from this table that the phases are relatively close to the theoretical values of $-90°$ and $-180°$ as one would expect for a three-way Doherty amplifier.

It is clear that the curve with continuously optimized phase relations (Figure 7.21) indeed gives a better performance over the use of fixed phase relations at higher power levels. However, as discussed in the simulation section, this difference is not significant enough to justify a much more complicated signal processing when the DPA is driven with a complex modulated signal. Another reason why the fixed-phase drive condition is preferred, is its more well-behaved gain-flatness (Figure 7.22), yielding a less critical predistortion.

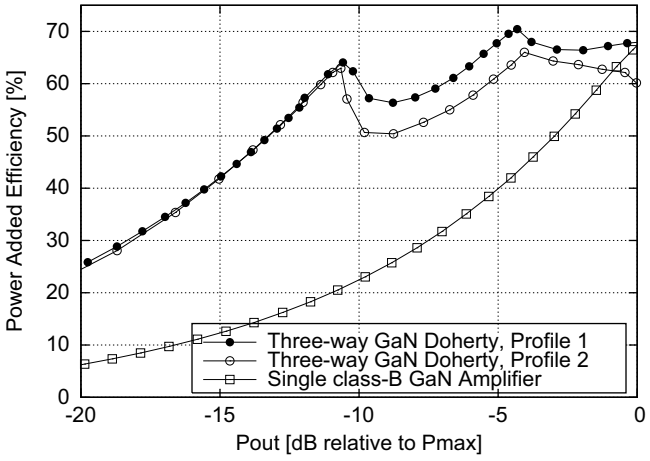


Figure 7.21: Measured single tone PAE versus output power backoff for different drive profiles. For the GaN three-way DPA, P_{max} = 50 dBm and for the single class-B amplifier P_{max} = 44.5 dBm. (in cooperation with M.Pelk [56])

Table 7.4: Relative Phase for Peak1 and Peak2

	Profile 1		Profile 2	
	φ_{peak1}	φ_{peak2}	φ_{peak1}	φ_{peak2}
Full Power	−81.7°	−191.2°	−99.8°	−198.2°
Backoff 1	−97.1°	−198°	−99.8°	−198.2°
Backoff 2	−99.8°	−198.2°	−99.8°	−198.2°

Figure 7.21 also shows the normalized measured PAE of a 45 W class-B GaN amplifier that was presented in Chapter 4. Note that this amplifier utilizes an identical device as in the peak 1 amplifier. It is interesting to see that at maximum output powers both the DPA, as well as, the class-B amplifier using the same device technology reach a maximum PAE of almost 70%, confirming the close to ideal operation of the DPA design at full power. Note, that the PAE of class-B GaN amplifier decreases proportional with the square of the back-off power, whereas the GaN three-way DPA demonstrates very high efficiency throughout the entire back-off range of 12 dB. At the 12 dB back-off point, the GaN three-way DPA provides a three time higher PAE than the class-B amplifier for CW signals, indicating the very high efficiency potential of the three-way DPA for complex modulated signals with a high peak-to-average power ratio.

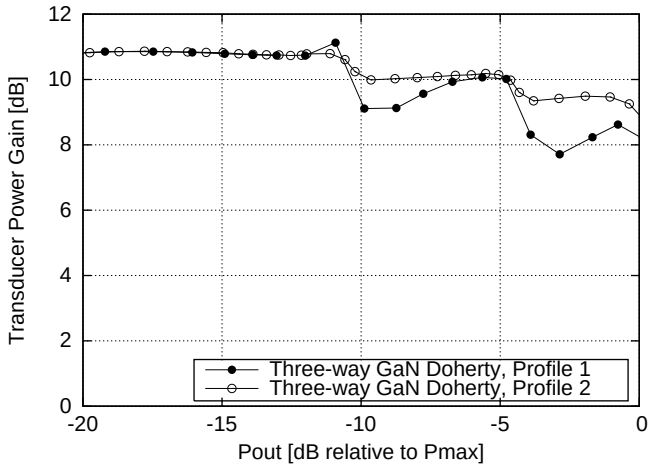


Figure 7.22: Measured gain versus output power backoff for the two different drive profiles. Drive profile 1 is with continuously varying phase and drive profile 2 is with fixed phase relations. As can be seen, the gain flatness is much better for the second drive profile. Also here, $P_{max} = 50$ dBm.

Measured W-CDMA Performance

One of the most important specifications for UMTS base station amplifiers is the Adjacent Channel Leakage Ratio (ACLR), since this gives a direct indication on the achieved spectral purity. To test amplifiers for this property, Test Model 1 [59] specifies a realistic W-CDMA traffic scenario with 16, 32 or 64 dedicated physical channels (DPCHs). Currently, to improve base-station efficiencies, people decrest W-CDMA signals [60] to reduce their peak-to-average power ratio, enabling amplifiers to operate at higher efficiencies. This decresting of the WCDMA-signal, typically results in a crest factor reduction of several dB without any significant degradation of the EVM and ACLR. In this mixed-signal three-way Doherty amplifier however, decresting is no longer needed, since its PAE vs. back-off power dependence is optimized to handle modulated signals with crest factors in the order of 10 to 12 dB.

For the characterization of our 3-way DPA a W-CDMA signal with 64 DPCHs was created using Agilents Advanced Design System (ADS) having a crest factor of 11.5 dB. Next, this signal is inputted in MATLAB, which returns the 3 input signals needed to drive the 3-way DPA. For this purpose, the optimum PAE single-tone drive conditions found in the previous section were used. Figure 7.23 gives the resulting AM-AM and AM-PM characteristics of the GaN three-way DPA when driven by the W-CDMA test signal. The left plots show the initial performance when no predistortion is applied. In these graphs, the "haze" found around the AM-AM and AM-PM characteristics indicate the presences of memory effects, which are caused by thermal effects, trapping in the devices and imperfect DC biasing of the devices. To suppress the influence of these effects, the DPD with memory compensation algorithm

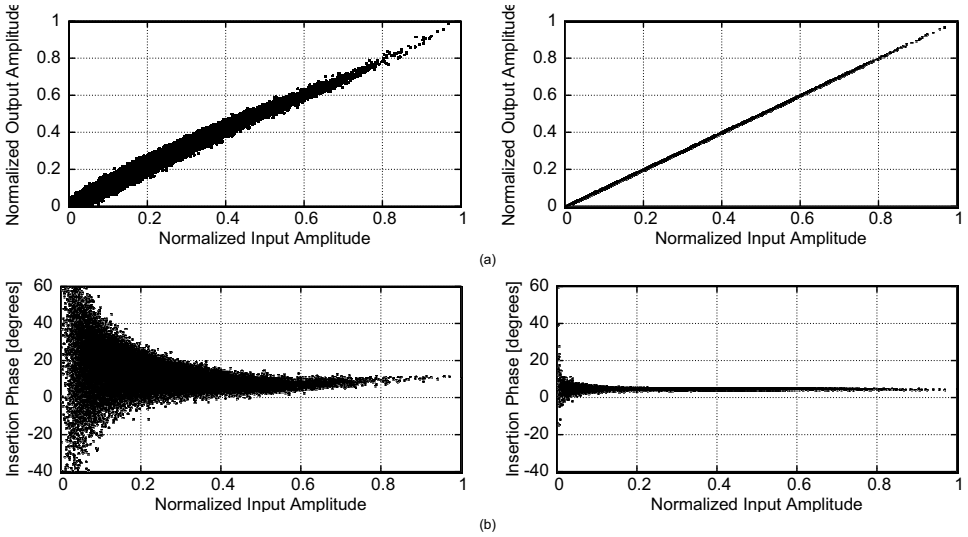


Figure 7.23: Measured AM-AM distortion (a) and AM-PM distortion (b) for the GaN three-way DPA. At the left the performance without predistortion can be seen and the right plots show the performance after predistortion with memory effect compensation.

developed in house and described in Chapter 4 is used. The system to be linearized is treated to be a single input and a single output system. The splitting of the signal to the Main, Peak 1 and Peak 2 is in the software part as a software splitter block. The plots on the right of Figure 7.23 show the resulting AM-AM and AM-PM distortion when this algorithm is applied. Comparing the later results with the non pre-distorted results, a clear reduction in haze can be observed, while the resulting characteristic approaches closely the behavior of an ideal linear system, indicating the correct operation of the algorithm.

Consequently, Figure 7.24 shows the spectrum of the signal before and after pre-distortion. The measured channel power over a bandwidth of 5 MHz is 38.5 dBm (approx. 7 watts) and the measured peak power is 50 dBm (100 watts). Note that with pre-distortion indeed a very significant improvement in linearity is achieved, resulting in a measured ACLR of better than -50 dBc for both ACLR 1 and ACLR 2. The related measured PAE for this power level is 53% with a crest factor of 11.5 dB. To the author's best knowledge, this is the highest PAE reported in literature for a base-station amplifier operating with a W-CDMA signal under these conditions.

In addition to the above, Figure 7.25 shows the measured PAE of the GaN three-way DPA and class-B GaN amplifier driven with the W-CDMA signal as function of back-off power. The back-off power level here is taken relative to the average W-CDMA output power. Note that although the class-B and mixed-signal three-way Doherty have a similar peak-efficiency for their single-tone operation (Figure 7.21), W-CDMA operation results in a more than 2 times higher PAE for the three-way Doherty, while still meeting all linearity specifications. The slow reduction of PAE

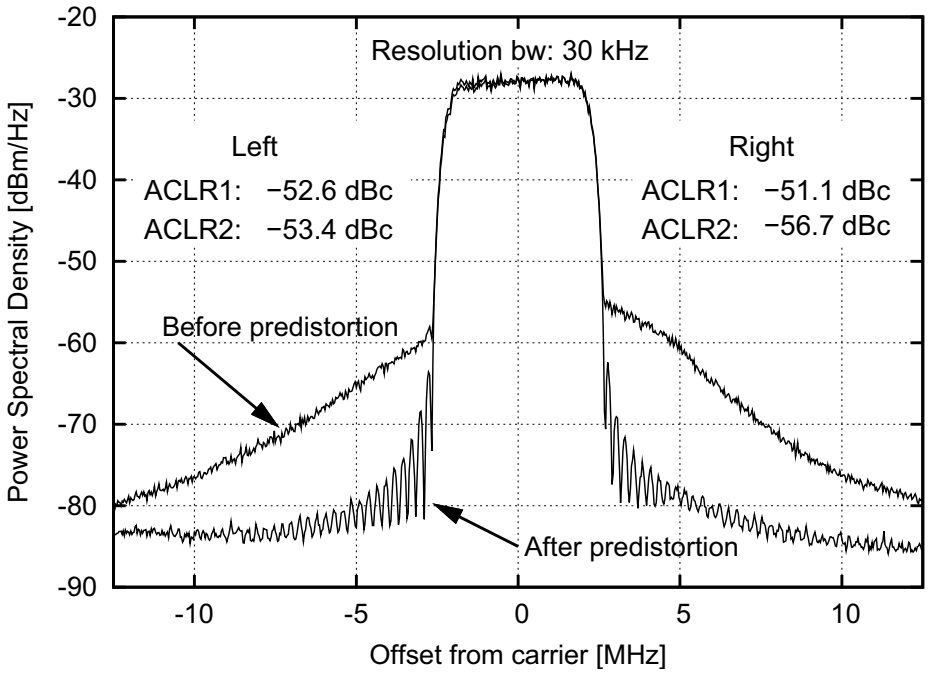


Figure 7.24: Measured spectrum for a W-CDMA signal before and after predistortion. The ACLR values shown are after predistortion. The average channel output power is 38.5 dBm and the PAE is 53%.

for the 3-way DPA versus increased power back-off, indicates once more the enormous performance potential of GaN technology for base-station applications, when signals with an even higher crest factor come in use.

7.7 Conclusions

In this chapter, the Doherty amplifier concept targeted for base-station application has been thoroughly investigated. 3G and 4G wireless standards have, as a result for supporting higher data through-put, RF signals with crest factors in the order of 10dB and higher. It has been shown analytically in this work that the Doherty concept can provide high-efficiency even when the output power is backed-off over a wide range. It requires the use of multiple devices in parallel, with their outputs combined in a "correct" way such that efficiency peaking occurs at multiple power levels. A novel way to analyze this Doherty combining network has been introduced in this thesis work. This gives the user the flexibility to "invent" a new power combining topology, if the ones available in literature do not meet the needs.

This technique was used to analyze the power combining network for a two-way asymmetrical Doherty, and it was found that the topology for the power combining network widely available in literature is the simplest strip-line implementation avail-

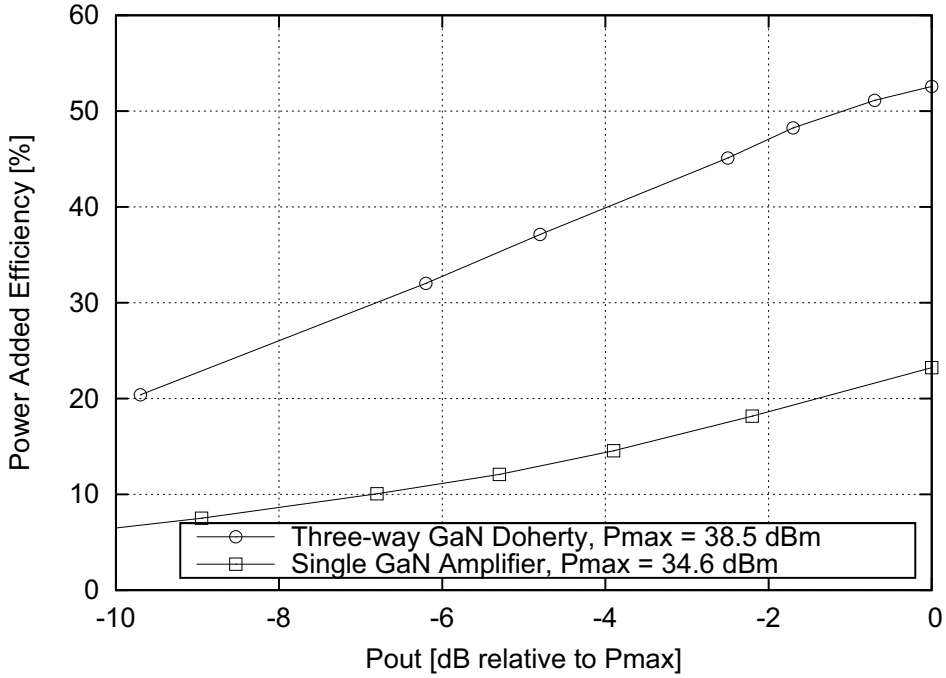


Figure 7.25: Measured spectrum for a W-CDMA signal before and after predistortion. The ACLR values shown are after predistortion. The average channel output power is 38.5 dBm and the PAE is 53%.

able. A set of design procedures with complete design equations are given in this chapter for the design of an asymmetrical Doherty amplifier. A demonstrator was built for WiMAX base-station application at the 2.5 to 2.6 GHz frequency band, and measurements demonstrate a significant improvement in efficiency over a class-B PA when driven by WiMAX test signals with a CF of 9.5 dB.

The two-way asymmetrical Doherty theoretically suffers from a significant dip in efficiency between its back-off peak efficiency point and the peak efficiency at full power (from 78% to 50% see Figure 7.2), thus to gain efficiency improvement for signals with comparable CF, this work recommends the use of the three-way Doherty concept.

The technique that was employed to analyze the power combining network for the two-way Doherty is extended to handle the three-way Doherty concept. New insights are obtained for the classical three-way Doherty power combining network, indicating the need for the main PA to have both current and voltage saturation. This further explains the enormous effort required to get a good performance for this classical Doherty design, in terms of efficiency and linearity performance. However, with the provided black-box parameter analysis of the power combining network, a new solution is also found, which avoids the current saturation of the main amplifier cell; and allows for the current profile to be accurately engineered just by the use of

asymmetrical splitters and the operating bias of the active devices.

In this work, we have labelled this new concept, the "Linear" three-way Doherty. This concept was verified in simulation, with a design aiming for UMTS application at 2.14 GHz. Simulation results show the linear ramp current profiles for the Main, Peak 1 and Peak 2, and the peak efficiencies at the targeted backed-off powers. Based on the design procedure outlined in the work for the "Linear" three-way Doherty, a Master of Science project performed by Ji Zhao and supervised by J. Gajadharsing at NXP was started to prototype the "linear" three-way hardware design [55]. Measurement results on the first prototype boards show improved performance over the asymmetrical Doherty amplifier prototype, but further design iterations are required to fully benefit from the performance potential.

To allow one to engineer any drive profile needed for the Main, Peak 1 and Peak 2, we have introduced the "mixed-signal" three-way Doherty. This concept combines digital software control with the Doherty hardware. This is done by removing the analogue symmetrical splitter and replacing it with a digital splitter. But, because any arbitrary drive profile can be made, it is now feasible to test this new concept with the well known three-way Doherty power combining network from literature.

To also investigate the potential of GaN as the technology of choice in future base-station implementations, a prototype mixed-signal three-way Doherty with GaN HEMT devices from CREE was built. To facilitate its calibrated testing, a custom 3-way test-bench has been developed, which can provide the required input signals, as well as the measurement of the resulting input/output port powers [56]. Using this setup, the CW-performance of the 3-way GaN DPA has been characterized and optimized using software control and yielding a measured performance of: 68% PAE at 50 dBm (full power), 70.4% at 45 dBm (first back-off point) and 64% at 38 dBm (second back-off point), while the measured transducer power gain was greater than 10 dB at all times.

To demonstrate that this exceptional high-efficiency performance can be effectively utilized for practical base-station operation, our GaN three-way DPA was driven by a W-CDMA signal with a crest factor of 11.5 dB. Using a dedicated, memory-effect compensating pre-distortion algorithm, the resulting measured PAE for this signal was 53% at an average power of 38.5 dBm, while meeting all linearity specifications.

To the author's best knowledge this is the highest PAE performance ever reported up to date, for any power amplifier operating with a W-CDMA signal without using crest factor reduction techniques.

To support this conclusion, various recent amplifier designs were considered, which also provide excellent W-CDMA performance (Table 7.5, column A). Note, that a straightforward comparison is troubled by the fact that W-CDMA signals with different crest factors have been used. However, to obtain a good estimate of the performance of these amplifiers for the unaltered W-CDMA signal (crest factor 11.5dB) as used in our experiments, their estimated efficiency for this signal is computed as well. The computations are based on provided data for their single-tone Drain Efficiency as function of back-off power, using the procedure demonstrated in Chapter 2. This estimate will be accurate within a few percent. The results are given in Table 7.5, column B.

The above demonstrates the performance potential of GaN technology for Doherty

Table 7.5: W-CDMA Performance of Recent Amplifier Designs

Design	W-CDMA crest factor used in ref.	A Measured drain efficiency with crest factor used in ref.	B Estimated drain efficiency with crest factor of 11.5 dB
Steinbeiser et al. [58]	6.5 dB	57 %	45.2 %
Yamamoto et al. [61]	9 dB	50 %	44.9 %
Ui et al. [62]	7.8 dB	42 %	36.2 %
Kimball et al. [31]	7.67 dB	53.4 %	45.4 %*
This work	**11.5 dB**	**55 %**	**55 %**

*Assuming crest factor independent efficiency for the envelope amplifier.

amplifiers, making the mixed-signal GaN three-way DPA an interesting candidate for base-station power amplifiers in the future.

Chapter 8

Conclusions and Recommendations

The objectives of this thesis work is to investigate innovative design solutions for RF power amplifiers, which provide efficient and linear amplifier operation. Besides these two basic requirements, future handset transmit amplifiers need to provide multi-band operation, which requires adjustments in the broadcast frequency. For base-station amplifiers emphasis is put on linear and efficient operation at very high output power levels (>100-Watts), while operating with signals that are characterized by a high peak-to-average power ratios.

To reach these research objectives, we first considered the traditional Class-B/-AB amplifier, and searched for ways to make this operation mode more linear without sacrificing efficiency. In view of this, to achieve a better linearity in power back-off operation, Chapter 3 proposes the use of Derivative Superposition. This technique, which manipulates the transconductance of the amplifying device by putting more devices with different gate bias offset voltages in parallel, proves to be most effective for low output powers up to 10dB of the 1dB compression point. However, when the transistors are operated closer to their compression (e.g. 5 to 7dB from the P1dB point) the classical derivative superposition technique loses its effectiveness. By extending this linearization technique and include also the second harmonic source and load terminations, improved linearity at much higher power levels (close to the P1dB compression point) were achieved. This "modified" Class-B operation resulted in a linearity-efficiency performance, with the latest generation of LDMOS devices, that approximates the theoretical limit of an ideal Class-B operated device, which is only limited by signal clipping. This latter technique is of great importance when aiming at low-cost high-bandwidth linear amplifier circuits, which does not require digital predistortion.

At the time that this research was performed, Digital Predistortion(DPD) was already gaining popularity as a system approach to correct for the distortion generated by the power amplifier. Due to its flexibility and "ease" of implementation, Digital Predistortion seems to be the linearization concept of choice when dealing with advanced high-efficiency power amplifier architectures. Although in theory these

high-efficiency power amplifier architectures are linear systems, in practice it proves that they are much more susceptible to nonlinear signal distortion than traditional or "modified" Class-B/-AB operated devices, due to the complex signal interaction within these architectures.

Therefore, before stepping into high-efficiency power amplifier architectures, we have considered and developed an effective digital predistortion approach that is suitable to be used in conjunction with the high-efficiency amplifier architectures that we discuss later in the thesis. Consequently, for this purpose in Chapter 4 we give predistortion solutions for the case the amplifier can be considered to have no memory and the more general case, where the memory effects of the amplifier can be no longer ignored.

In Chapter 5, we give an overview of the most common high-efficiency amplifier architectures reported in literature to date, listing the advantages and disadvantages of each architecture.

In Chapter 6, we discuss a high-efficiency amplifier architecture suitable for future handset applications. This direct load-line adaptation amplifier has been developed and implemented using the DIMES in-house "ultra-low distortion varactor technology". This concept can provide improved amplifier efficiency under "static" power back-off conditions. Note, that this feature is very useful for communication systems that make use of control loops for the transmitted output power (e.g. (W)CDMA systems). In addition to this, it was shown that the realized adaptive load amplifier is also capable of tracking the instantaneous envelope of a modulated RF signal. This latest feature provides improved efficiency for the amplification of signals with a high peak-to-average power ratio. Equally important is the fact that the tunable matching networks used to implement the load variable amplifier architecture provides the ability to adjust the operating frequency and serve more communication frequency bands with a single amplifier. In support of this, a demonstrator realized using these integrated varactor based matching networks, managed to cover all frequency bands from 900 MHz to 2100 MHz, while maintaining sufficient gain, output power and efficiency.

In Chapter 7, focus is placed on high-efficient power amplifier architectures suitable for base-station applications. Here emphasis is on the Doherty power amplifier, which at the time of research investigations, was already being actively pursued in the industry to meet the high-efficiency demands in 3G/4G applications. In this thesis work, we have contributed to a deeper understanding of the N-way Doherty architecture and contributed the novel linear three-way Doherty design procedure. The design concepts developed are translated to functional hardware demonstrators, which yielded record high efficiency performance. Based on these results these new developments have reached already early adaptation in industry and will most likely be utilized in the next generation of highly efficient base-stations for 3G/4G communication systems.

In the next few sections, we will like to succinctly raise a few issues that were not part of the thesis work, but which we see as potential to impact the RF power amplifier community in the future.

8.1 High-Efficient Amplifier Classes of Operation

In this work, the "Class-AB" amplifier formed the backbone of the improved efficiency architectures investigated in Chapter 3 (modified Class-B), Chapter 6 (adaptive matching) and Chapter 7 (Doherty amplifier). The reason for this choice is that the Class-B is simple, effective and can be easily adapted in more complex high efficiency amplifier architectures.

However, to achieve peak efficiency operation beyond 78.5%, in various research projects focus is currently placed on the so called switched-mode classes, which in principle can provide peak efficiencies beyond 78.5%. The basic concept behind these high-efficiency classes is to shape the voltage or/and current waveforms at the drain of the device such that the voltage/current waveforms becomes "squarish". This effectively minimizes the overlap between voltage and current at the device terminals, resulting in lower power losses and thus increasing efficiency. In view of this, currently the following amplifier classes have received an increased interest to enable the future generation of high efficiency RF amplifier implementations:

Class-F [11], this amplifier class uses short-circuited conditions for the even harmonics and open conditions for the odd harmonics. When controlling the active device terminations up to the third harmonic, theoretical 90.7% efficiency can be achieved. This theoretical efficiency approaches 100% for maximum output power, when an infinite amount of harmonics are controlled. Note that Class-F is still a "linear" operating amplifier and allows operation in power back-off. For this reason it is in principle also suited for application in Doherty and LINC amplifiers [63]. Practical implementations of Class-F amplifiers however are constrained in their effective control of the higher harmonics, which becomes very problematic when aiming for larger bandwidths.

Class-D also utilizes square-waveforms at the output of the transistor, but achieves this by using two or more transistor as switches. Due to this "ideal" switching its theoretical efficiency is 100%, but power back-off operation is now only possible by modulating the supply voltage or using a kind of pulse modulation of the output. This amplifier concept is extensively used in audio amplifiers and DC-DC converters. However at RF frequencies, practical implementations are scarce, as the transistor will suffer at these frequencies from limitations in switching speed (deviating from the ideal square waveforms), and losses due to the output capacitance. Improvement in device technology has elevated some of these limitations to a certain extend, and to date Class-D demonstrators have been reported up to 2GHz [64–66].

Class-E power amplifiers [67] use a single transistor operated as a switch. But the concept here is not to create square voltage waveforms at the drain of the transistor, but rather to ensure that the drain voltage drops to zero just before the drain current comes on. Figure 8.1 shows a generic Class-E voltage and current waveform at the output of the transistor, notice that the overlap in voltage and current is kept to a bare minimum. The advantage of Class-E, is that it makes use of the parasitic drain-source capacitance as part of its matching, resulting in "skewed" voltage waveforms. As such Class-E compared to Class-D, seems much more practical at RF frequencies [68]. Since it also does not require true "square like" waveforms at the output of the transistor, the switching speed is more relaxed than in Class-D. For these reasons

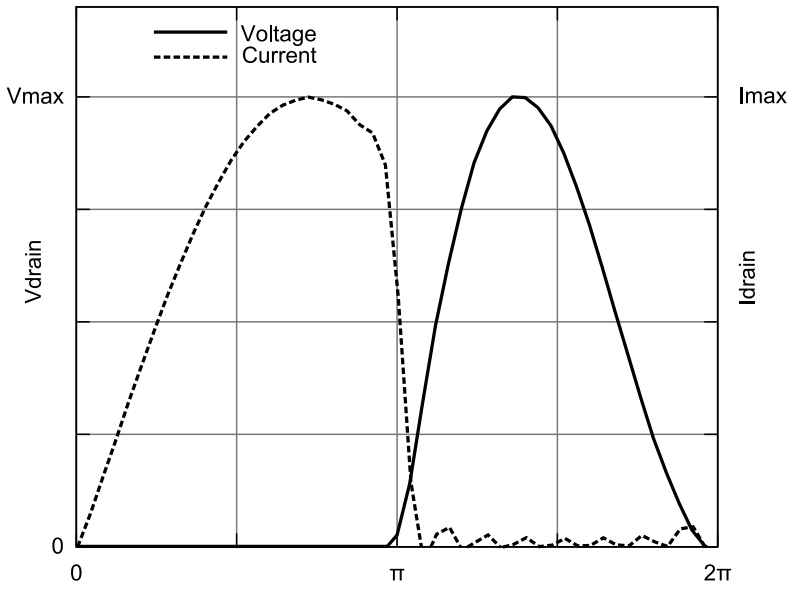


Figure 8.1: Typical Class-E voltage and current waveforms at the output of the transistor.

Class-E has already been effectively used by various research groups in Envelope Elimination and Restoration amplifier architectures [69, 70], as well as, in Envelope Tracking amplifiers [71]. Both concepts make use of supply voltage modulation to control their output power and enhance their efficiency.

Due to their advantages for improved peak efficiency the high efficiency amplifier classes of above are intensively investigated in various research projects, however, currently industrial RF implementations are still very rare. This is partly due to limitations in gain and therefore also in PAE, the additional cost of linearization, and the in practice often very limited improvement over a well designed class-B amplifier stage. It is therefore the author's personal opinion that the benefits in peak efficiency are only worthwhile when these switching classes are operated in combination with high efficiency amplifier architectures, like LINC, Doherty, EER, Envelope tracking etc. Such that their improved peak efficiency is beneficial in achieving a higher average efficiency for modern complex modulated signals, while their shortcomings in linearity are eased by digital predistortion techniques.

8.2 Technologies for Base-Station RF Power Amplifiers

Currently, LDMOS is the standard technology of choice within the RF industry for base-station power amplifiers. It has maintained its lead over other technologies through constant improvement. In [9], device specialist at NXP have shown how power density, linearity and power efficiency of their LDMOS devices have improved

over various technology generations (Figure 8.2). As such, LDMOS has the important advantage that, by being around for a long time, it has become a mature low-cost process technology with high yield and high reliability.

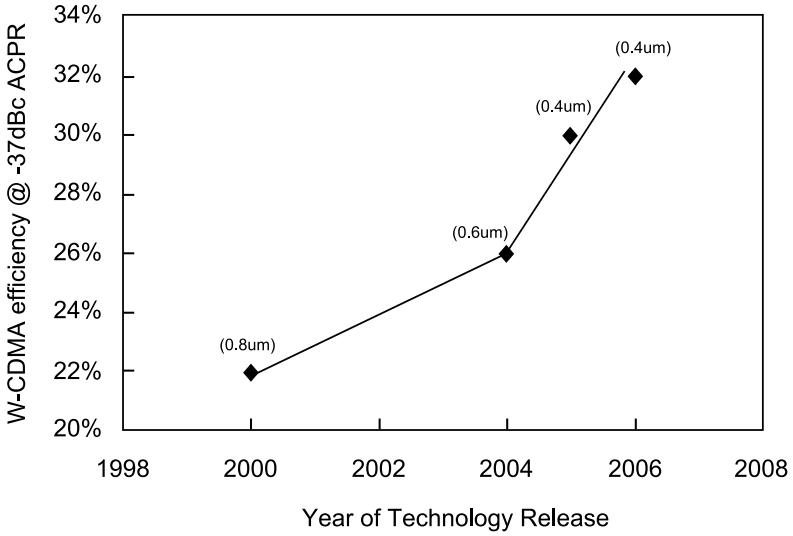


Figure 8.2: Realized W-CDMA efficiency at 2.14 GHz versus time for 100W NXP LDMOS-transistors. The used gate-length is shown in brackets. Courtesy of Fred van Rijs and Steven Theeuwen from NXP.

GaN is considered as an up-and-coming technology and poses the general threat of displacing LDMOS as the technology of choice for base-station power amplifiers. GaN power devices can provide higher power densities than LDMOS devices, while higher higher breakdown voltage allows the use of much higher drain voltages. Both properties result in a much more relaxed impedance levels for the matching of the device input and output when comparing to LDMOS devices with a similar output power levels. The higher impedance levels allow designs that are more wideband or aim for higher operating frequencies. In addition, the effective output resistance of GaN devices is much higher than that of what is typically found for LDMOS devices, making GaN an excellent choice for high efficiency amplifier architectures that use load modulation like e.g. Doherty or LINC.

To utilize this latest feature, in this thesis work, GaN devices from Cree have been used in a "mixed-signal" three-way Doherty (100 Watt amplifier). This experiment resulted in a record high power average efficiency for non de-crested (11.5dB peak-to-average power ratio) WCDMA signal. Another example where GaN was successfully used is in [31], where the authors demonstrated a GaN Envelope Tracking amplifier with efficiencies above 50% operating with WCDMA signals with a crest factor of 7.67dB.

Despite having better performance capabilities than LDMOS, it will still take sometime before GaN will achieve a majority share in the base-station market. This is because LDMOS is a much more mature technology, and therefore exceeds by

far GaN in reliability and cost per watt. It are these two criteria that carry the most weight, when deciding for a particular technology. In spite of these economical arguments, research to investigate the intrinsic advantages of GaN (such as its high break-down voltage), is extremely useful, since it will help us to find high performance amplifier solutions for the future which are more environmental friendly.

8.3 The Future – "Mixed-Signal", "Software Radio"

One of the buzzwords in the wireless communication community today is "Software-defined Radio". The concept of "Software-defined Radio" as visualized by the author in [72] consist of a transmitter that is capable of characterizing the propagation path, adapt the signal to best suit the propagation environment, electronically steer the transmit beam in the most optimal direction, set the optimal power level of the signal to be transmitted before the actual transmission takes place. Likewise, the receiver should be able to detect the frequency of the broad casted signal of interest. For this purpose it needs to be able to sense the transmitted energy in multiple communication channels, while the modulation schemes used for these channels, must be detected on the fly, resulting in a reconfiguration on demand of the demodulator of the receiver.

It is clear from this description that such an implementation will require tremendous amount of computing power for the signal processing and control (FPGAs and dedicated DSPs are usually employed). On the RF/analogue end, the challenge is to move the digital interface as close to the antenna as possible. For the transmitter, the ideal situation would be to just have the power amplifier stage as the only remaining analogue part, which is directly driven by an extremely fast (D/A) converter. However, this approach proves to be extremely difficult to realize in practice, since this would require a D/A converter that is able to operate with a high dynamic range at frequencies above 1GHz. It is obvious that such a brute force approach would consume a tremendous amount of DC power. As such, the proposed "Software Radio" architecture these days uses a digital IF interface, where channel selection and filtering can be done in the digital domain. Note that in spite the progressing digitalization of the radio, critical parts like the PA stage and antenna remain analogue in nature, and for this reason are restricted in bandwidth. To easily change operating frequency and make band-switching feasible, low-loss tunable matching elements are required [73–79].

Consequently, although true "Software-defined Radio" still seems to be a distant goal, digitally enhanced RF circuits are already appearing. E.g. in [80, 81], a LINC high-efficiency architecture is demonstrated that seems to be suitable to be employed within the "Software Radio" architecture. Also in this thesis, digitally enhanced RF circuits have been demonstrated; e.g. digital predistortion for correcting the non-linearities of the amplifier core, band-switching handset amplifiers that utilize software control for pre-distortion and setting the value of the tunable elements to achieve the desired frequency band of operation. One of the most outspoken examples of digitally enhanced RF functionality in this work is the "Mixed-Signal" Doherty approach as presented in Chapter 7. Here software control is used to optimize the phase-offsets and

drive profiles of each individual device for achieving the very best Doherty operation. This was up to date, always done in an analogue fashion by the use of offset lines and analogue signal splitters, by pushing these analogue functions to the digital domain, directly a record breaking efficiency performance was achieved. As a matter of fact, this result triggered new research activities within our group to integrate these to digital functions into a dedicated driver chip for compact highly efficiency Doherty amplifiers.

It is clear from the examples given in this thesis work, that following a "Mixed-Signal"/software approach brings performance advantages, since it relaxes design specifications on the analogue/RF part, as "errors" can be corrected easily corrected within the software. It is therefore our expectation that the future of RF power amplifiers will be in the "Mixed-Signal" domain, where full use of the advancements in digital signal processing will result in power amplifier that are more efficient, more linear, and reconfigurable. Consequently, the day that a typical RF amplifier will have a digital I-Q baseband inputs is closer than ever before ···

Appendix A

Black-box Parameter derivation of the Two-Way Doherty Output Power Combiner

This appendix describes how we can use the black-box parameter approach described in Chapter 7, to derive a two-way Doherty output power combiner, that allows flexibility in choosing the characteristic impedance of the lines and the impedances offered to the Main and the Peak amplifier.

From Figure 7.1, we can relate the z-parameter of the intermediate power combining network (with the load incorporated) to the voltages and currents of the Main and Peak amplifier. Solving the voltages and currents against the boundary conditions (equations (7.6)– (7.10)), and setting the phase between the Main and the Peak to 90°, results in the following compact z-parameter description of the intermediate power combiner.

$$z = \begin{bmatrix} \frac{Z_{o,int}}{k^2} & -j\,\frac{Z_{o,int}}{k} \\ -j\,\frac{Z_{o,int}}{k} & 0 \end{bmatrix} \tag{A.1}$$

Where $Z_{o,int}$ refers to the intermediate load impedance that is embedded within the power combining network, described by the black-box parameter. And k refers to the back-off ratio of the two-way Doherty amplifier.

Transforming the z-parameters to s-parameters with a system impedance $Z_{o,sys}$ results in the following:

$$s = \frac{1}{\zeta}\begin{bmatrix} Z_{o,int}\cdot Z_{o,sys} - (k\cdot Z_{o,sys})^2 + Z_{o,int}^2 & -2\cdot j\cdot Z_{o,int}\cdot Z_{o,sys}\cdot k \\ -2\cdot j\cdot Z_{o,int}\cdot Z_{o,sys}\cdot k & -Z_{o,int}\cdot Z_{o,sys} - (k\cdot Z_{o,sys})^2 + Z_{o,int}^2 \end{bmatrix} \tag{A.2}$$

Where

$$\zeta = Z_{o,int} \cdot Z_{o,sys} + (k \cdot Z_{o,sys})^2 + Z_{o,int}^2$$

To derive the full three-port s-parameter that describes the power combining network, we introduce the s-parameters of the third port.

$$\begin{aligned} s_{13} &= |s_{13}| \cdot \exp(j \cdot \beta) \\ s_{23} &= |s_{23}| \cdot \exp(j \cdot \alpha) \\ s_{33} &= |s_{33}| \cdot \exp(j \cdot \sigma) \end{aligned} \tag{A.3}$$

Where β, α and σ are phases to be determined.

Using the lossless three-port s-parameter conditions (equation (7.14)) we can write down the following equation where only the magnitude of s_{13} is unknown.

$$|s_{11}|^2 + |s_{12}|^2 + |s_{13}|^2 = 1 \tag{A.4}$$

Rearranging and substituting the known s-parameters.

$$\begin{aligned} |s_{13}|^2 &= 1 - |s_{11}|^2 - |s_{12}|^2 \\ |s_{13}| &= \frac{2 \cdot k \cdot Z_{o,sys} \sqrt{Z_{o,int} \cdot Z_{o,sys}}}{Z_{o,int} \cdot Z_{o,sys} + (k \cdot Z_{o,sys})^2 + Z_{o,int}^2} \end{aligned} \tag{A.5}$$

The negative solution is ignored as we are referring here to magnitudes.

Similarly to solve for the magnitude of s_{23} we can write the following equation:

$$|s_{12}| + |s_{22}| + |s_{23}| = 1 \tag{A.6}$$

$$|s_{23}| = \frac{2 \cdot Z_{o,int} \sqrt{Z_{o,int} \cdot Z_{o,sys}}}{Z_{o,int} \cdot Z_{o,sys} + (k \cdot Z_{o,sys})^2 + Z_{o,int}^2} \tag{A.7}$$

Knowing the magnitude of s_{23}, we can solve for the magnitude of s_{33} using:

$$|s_{13}| + |s_{23}| + |s_{33}| = 1 \tag{A.8}$$

$$|s_{33}| = \left| \frac{-Z_{o,int} \cdot Z_{o,sys} + (k \cdot Z_{o,sys})^2 + Z_{o,int}^2}{Z_{o,int} \cdot Z_{o,sys} + (k \cdot Z_{o,sys})^2 + Z_{o,int}^2} \right| \tag{A.9}$$

At this point, the magnitudes of the s-parameters of the third port is uniquely known. To solve for their phases β, α and σ, the last three equations of the lossless three-port s-parameter are used.

$$\bar{s_{11}} \cdot s_{12} + \bar{s_{12}} \cdot s_{22} + \bar{s_{13}} \cdot s_{23} = 0 \tag{A.10}$$

$$\bar{s_{13}} \cdot s_{11} + \bar{s_{23}} \cdot s_{12} + \bar{s_{33}} \cdot s_{13} = 0 \tag{A.11}$$

$$\bar{s_{12}} \cdot s_{13} + \bar{s_{22}} \cdot s_{23} + \bar{s_{23}} \cdot s_{33} = 0 \tag{A.12}$$

Substituting the known values into the above equations, results in a system of equations where there is no unique solution. The phases are related to each other through the following equations:

$$\beta = \alpha - \frac{\pi}{2} \tag{A.13}$$

$$\sigma = \pi + 2 \cdot \beta \tag{A.14}$$

When the phase α, between the Peak amplifier and the output load $Z_{o,sys}$ is set to $-\frac{\pi}{2}$. This results in $\beta = -\pi$ and $\sigma = -\pi$. With this, the s-parameter description of the output power combining network of the two-way Doherty is complete. Transforming this into its y-parameter equivalent for synthesis reasons results in:

$$y = \begin{bmatrix} 0 & j\frac{k}{Z_{o,int}} & 0 \\ j\frac{k}{Z_{o,int}} & 0 & j\frac{1}{\sqrt{Z_{o,int} \cdot Z_{o,sys}}} \\ 0 & j\frac{1}{\sqrt{Z_{o,int} \cdot Z_{o,sys}}} & 0 \end{bmatrix} \tag{A.15}$$

Which results in the network of Figure A.1.

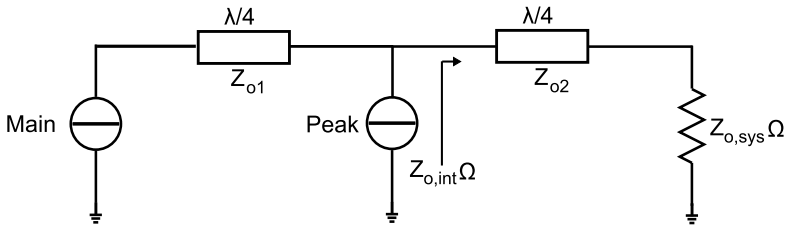


Figure A.1: Resulting transmission line network from synthesizing the three-port y-parameters given in equation (A.15)

The characteristic of the quarter-wave transmission lines are as follows:

$$Z_{o1} = \frac{Z_{o,int}}{k} \tag{A.16}$$

$$Z_{o2} = \sqrt{Z_{o,int} \cdot Z_{o,sys}} \tag{A.17}$$

Bibliography

[1] M. van der Heijden, L. de Vreede, and J. Burghartz, "On the design of unilateral dual-loop feedback low-noise amplifiers with simultaneous noise, impedance, and IIP3 match," *Solid-State Circuits, IEEE Journal of*, vol. 39, no. 10, pp. 1727–1736, Oct. 2004.

[2] S.-T. Lim and J. Long, "A Low-Voltage Broadband Feedforward-Linearized BJT Mixer," *Solid-State Circuits, IEEE Journal of*, vol. 41, no. 9, pp. 2177–2187, Sept. 2006.

[3] L. Sheng and L. Larson, "An Si-SiGe BiCMOS direct-conversion mixer with second-order and third-order nonlinearity cancellation for WCDMA applications," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 51, no. 11, pp. 2211–2220, Nov. 2003.

[4] S. Li, J. Huang, Q. Meng, L. Sun, Q. Zhang, F. Li, A. He, X. Zhang, C. Li, H. Li, and Y. He, "A 12-Pole Narrowband Highly Selective High-Temperature Superconducting Filter for the Application in the Third-Generation Wireless Communications," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 55, no. 4, pp. 754–759, April 2007.

[5] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*. Norwood, MA: Artech House, 1999.

[6] P. B. Kennington, *High-Linearity RF Amplifier Design*. Norwood, MA: Artech House, 2000.

[7] G. Gonzalez, *Microwave Transistor Amplifiers Analysis and Design*. Upper Saddle River: Prentice Hall, 1984.

[8] D. A. Neamen, *Semiconductor Devices and Physics, Basic Principles*, 2nd ed. New York: McGraw-Hill, 1997.

[9] F. van Rijs and S. Theeuwen, "Efficiency improvement of LDMOS transistors for base stations: towards the theoretical limit," *Electron Devices Meeting, 2006. IEDM '06. International*, pp. 1–4, Dec. 2006.

[10] P. Wamback and W. Sansen, *Distortion Analysis of Analog Integrated Circuits*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1998.

[11] F. Raab, "Class-F power amplifiers with maximally flat waveforms," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 45, no. 11, pp. 2007–2012, Nov 1997.

[12] ——, "Efficiency of Outphasing RF Power-Amplifier Systems," *Communications, IEEE Transactions on [legacy, pre - 1988]*, vol. 33, no. 10, pp. 1094–1099, Oct 1985.

[13] ——, "Efficiency of Doherty RF Power-Amplifier Systems," *Broadcasting, IEEE Transactions on*, vol. BC-33, no. 3, pp. 77–83, Sept. 1987.

[14] S. Marsili, "Technique for Peak to Average Power Ratio Reduction suited for M-carrier WCDMA Base Station Transmitters," *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, pp. 3522–3525, May 2007.

[15] D. R. Webster, J. Scott, and D. G. Haigh, "Control of circuit distortion by the derivative superposition method," *IEEE Microw. Guided Wave Lett.*, vol. 6, no. 3, pp. 123–125, Mar. 1996.

[16] M. van der Heijden, "RF Amplifier Design Techniques for Linearity and Dynamic Range," Ph.D. dissertation, Delft University of Technology, Delft, The Netherlands, 2005.

[17] D. Hartskeerl, I. Volokhine, and M. Spirito, "On the optimum 2^{nd} harmonic source and load impedances for the efficiency-linearity trade-off of RF LDMOS power amplifiers," *Radio Frequency integrated Circuits (RFIC) Symposium, 2005. Digest of Papers. 2005 IEEE*, pp. 447–450, 12-14 June 2005.

[18] J. Vuolevi, J. Aikio, and T. Rahkonen, "Extraction of a polynomial LDMOS model for distortion simulations using small-signal S-parameter measurements," *Microwave Symposium Digest, 2002 IEEE MTT-S International*, vol. 3, pp. 2157–2160, 2002.

[19] J. Gajadharsing, "Low distortion RF-LDMOS Power Transistor for Wireless Communications Base Stations Application," *Microwave Symposium Digest, 2003 IEEE MTT-S International*, vol. 3, pp. 1563–1566, 2003.

[20] N. de Carvalho and J. Pedro, "Large signal IMD sweet spots in microwave power amplifiers," *Microwave Symposium Digest, 1999 IEEE MTT-S International*, vol. 2, pp. 517–520 vol.2, 1999.

[21] M. Spirito, "Enhanced Techniques for the Design and Characterization of RF Power Amplifiers," Ph.D. dissertation, Delft University of Technology, Delft, The Netherlands, 2005.

[22] M. Spirito, M. J. Pelk, F. van Rijs, S. J. C. H. Theeuwen, D. Hartskeerl, and L. C. N. de Vreede, "Active Harmonic Load-Pull for On-Wafer Out-of-Band Device Linearity Optimization," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 54, no. 12, pp. 4225–4236, Dec. 2006.

[23] J. Vuolevi and T. Rahkonen, *Distortion in RF Power Amplifiers*. Artech House, 2003.

[24] F. Raab, "Radio Frequency Pulsewidth Modulation," *Communications, IEEE Transactions on*, vol. 21, no. 8, pp. 958–966, Aug 1973.

[25] A. Grebennikov and N. O. Sokal, *Switchmode RF Power Amplifiers*. Burlington, MA: Newnes, 2007.

[26] A. Jayaraman, P. Asbeck, K. Nary, S. Beccue, and K.-C. Wang, "Bandpass delta-sigma modulator with 800 Mhz center frequency," Oct 1997, pp. 95–98.

[27] J. Keyzer, R. Uang, Y. Sugiyama, M. Iwamoto, I. Galton, and P. Asbeck, "Generation of RF pulsewidth modulated microwave signals using delta-sigma modulation," vol. 1, 2002, pp. 397–400.

[28] S. Rosnell and J. Varis, "Bandpass pulse-width modulation (BP-PWM)," June 2005, pp. 4 pp.–.

[29] F. Raab, B. Sigmon, R. Myers, and R. Jackson, "L-band transmitter using Kahn EER technique," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 46, no. 12, pp. 2220–2225, Dec 1998.

[30] J. Staudinger, B. Gilsdorf, D. Newman, G. Norris, G. Sadowniczak, R. Sherman, and T. Quach, "High efficiency cdma rf power amplifier using dynamic envelope tracking technique," *Microwave Symposium Digest., 2000 IEEE MTT-S International*, vol. 2, pp. 873–876 vol.2, 2000.

[31] D. F. Kimball, J. Jeong, C. Hsia, P. Draxler, S. Lanfranco, W. Nagy, K. Linthicum, L. E. Larson, and P. M. Asbeck, "High-Efficiency Envelope-Tracking W-CDMA Base-Station Amplifier Using GaN HFETs," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 11, pp. 3848–3856, Nov. 2006.

[32] H. Chireix, "High Power Outphasing Modulation," *Proceedings of the IRE*, vol. 23, no. 11, pp. 1370–1392, Nov. 1935.

[33] D. Cox and R. Leck, "A VHF Implementation of a LINC amplifier," *Communications, IEEE Transactions on*, vol. 24, no. 9, pp. 1018–1022, Sep 1976.

[34] W. Gerhard and R. Knoechel, "LINC digital component separator for single and multicarrier W-CDMA signals," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 1, pp. 274–282, Jan. 2005.

[35] M. Helaoui, S. Boumaiza, A. Ghazel, and F. Ghannouchi, "Digital compensation of branches imbalance effects in LINC transmitters," Dec. 2004, pp. 688–691.

[36] J. Qureshi, R. Lui, A. de Graaw, M. P. van der Heijden, J. Gajadharsing, and L. C. de Vreede, "A Highly Efficient Chireix Amplifier using Adaptive Power Combining," *Microwave Symposium Digest, 2008 IEEE MTT-S International*, 2008.

[37] W. Doherty, "A New High Efficiency Power Amplifier for Modulated Waves," *Proceedings of the IRE*, vol. 24, no. 9, pp. 1163–1182, Sept. 1936.

[38] F. Raab, "High-efficiency linear amplification by dynamic load modulation," *Microwave Symposium Digest, 2003 IEEE MTT-S International*, vol. 3, pp. 1717–1720 vol.3, June 2003.

[39] J. Ryynanen, K. Kivekas, J. Jussila, A. Parssinen, and K. Halonen, "A dual-band RF front-end for WCDMA and GSM applications," *Solid-State Circuits, IEEE Journal of*, vol. 36, no. 8, pp. 1198–1204, Aug 2001.

[40] D. Wang, X. Wang, A. Eshraghi, D. Chang, and P. Bacon, "A fully integrated GSM/DCS/PCS Rx VCO with fast switching auto-band selection," *Radio and Wireless Conference, 2002. RAWCON 2002. IEEE*, pp. 209–212, Aug. 2002.

[41] A. Tasic, W. Serdijn, and J. Long, "Multi-standard/multi-band adaptive voltage-controlled oscillator," *Radio Frequency Integrated Circuits (RFIC) Symposium, 2004. Digest of Papers. 2004 IEEE*, pp. 135–138, June 2004.

[42] D. Qiao, R. Molfino, S. Lardizabal, B. Pillans, P. Asbeck, and G. Jerinic, "An intelligently controlled RF power amplifier with a reconfigurable MEMS-varactor tuner," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 3, pp. 1089–1095, March 2005.

[43] H. T. Jeong, H. S. Lee, I. S. Chang, and C. D. Kim, "Efficiency enhancement method for high-power amplifiers using a dynamic load adaptation technique," *Microwave Symposium Digest, 2005 IEEE MTT-S International*, pp. 4 pp.–, June 2005.

[44] K. Buisman, L. C. N. de Vreede, L. E. Larson, M. Spirito, A. Akhnoukh, Y. Lin, X. Liu, and L. K. Nanver, "A Monolithic Low-Distortion Low-Loss Silicon-on-Glass Varactor-Tuned Filter with Optimized Biasing," *Microwave and Wireless Components Letters, IEEE*, vol. 17, no. 1, pp. 58–60, Jan. 2007.

[45] K. Buisman, C. Huang, A. Akhnoukh, M. Marchetti, L. de Vreede, L. Larson, and L. Nanver, "Varactor Topologies for RF Adaptivity with Improved Power Handling and Linearity," *Microwave Symposium, 2007. IEEE/MTT-S International*, pp. 319–322, June 2007.

[46] K. Buisman, L. Nanver, T. Scholtes, H. Schellevis, and L. de Vreede, "High-performance varactor diodes integrated in a silicon-on-glass technology," *Solid-State Device Research Conference, 2005. ESSDERC 2005. Proceedings of 35th European*, pp. 117–120, Sept. 2005.

[47] K. Buisman, L. de Vreede, L. Larson, M. Spirito, A. Akhnoukh, T. Scholtes, and L. Nanver, "Distortion-free varactor diode topologies for RF adaptivity," *Microwave Symposium Digest, 2005 IEEE MTT-S International*, pp. 4 pp.–, June 2005.

[48] K. Buisman, L. de Vreede, L. Larson, M. Spirito, A. Akhnoukh, Y. Lin, X. Liu, and L. Nanver, "Low-distortion, low-loss varactor-based adaptive matching networks, implemented in a silicon-on-glass technology," *Radio Frequency integrated Circuits (RFIC) Symposium, 2005. Digest of Papers. 2005 IEEE*, pp. 389–392, June 2005.

[49] R. Meyer and M. Stephens, "Distortion in variable-capacitance diodes," *Solid-State Circuits, IEEE Journal of*, vol. 10, no. 1, pp. 47–54, Feb 1975.

[50] G. Leuzzi and C. Micheli, "Variable-Load Constant-Efficiency Power Amplifier for Mobile Communications Applications," *European Microwave Conference, 2003. 33rd*, pp. 375–377, Oct. 2003.

[51] P. Deixler, R. Colclaser, D. Bower, N. Bell, W. De Boer, D. Szmyd, S. Bardy, W. Wilbanks, P. Barre, M. v Houdt, J. Paasschens, H. Veenstra, E. v d Heijden, J. Donkers, and J. Slotboom, "Qubic4g: a f_T/f_{max} = 70/100 GHz 0.25μm low power SiGe-BiCMOS production technology with high quality passives for 12.5 Gb/s optical networking and emerging wireless applications up to 20 Ghz," *Bipolar/BiCMOS Circuits and Technology Meeting, 2002. Proceedings of the 2002*, pp. 201–204, 2002.

[52] C. Huang, L. de Vreede, F. Sarubbi, M. Popadic, K. Buisman, J. Qureshi, M. Marchetti, A. Akhnoukh, T. Scholtes, L. Larson, and L. Nanver, "Enabling Low-Distortion Varactors for Adaptive Transmitters," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 56, no. 5, pp. 1149–1163, May 2008.

[53] M. Iwamoto, A. Williams, P. Chen, A. G. Metzger, L. E. Larson, and P. M. Asbeck, "An extended Doherty amplifier with high efficiency over a wide power range," *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 12, pp. 2472–2479, Dec. 2001.

[54] N. Srirattana, A. Raghavan, D. Heo, P. Allen, and J. Laskar, "Analysis and Design of a High-Efficiency Multistage Doherty Power Amplifier for Wireless Communications," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 3, pp. 852–860, Mar. 2005.

[55] J. Zhao, "High Efficiency three-way Doherty Amplifier with a Passive Input Splitter," Master's thesis, Technische Universiteit Delft, Department of Microelectronics, Delft, The Netherlands, 2008.

[56] M. Pelk, W. Neo, J. Gajadharsing, R. Pengelly, and L. de Vreede, "A High-Efficiency 100-W GaN Three-Way Doherty Amplifier for Base-Station Applications," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 56, no. 7, pp. 1582–1591, July 2008.

[57] J. Gajadharsing, O. Bosma, and P. van Westen, "Analysis and design of a 200 W LDMOS based Doherty amplifier for 3 G base stations," *Microwave Symposium Digest, 2004 IEEE MTT-S International*, vol. 2, pp. 529–532 Vol.2, June 2004.

[58] C. Steinbeiser, T. Landon, C. Suckling, J. Nelson, J. Delaney, J. Hitt, L. Witkowski, G. Burgin, R. Hajji, and O. Krutko, "250 W HVHBT Doherty with 57% WCDMA Efficiency Linearized to -55 dBc for 2c11 6.5 dB PAR," *Solid-State Circuits, IEEE Journal of*, vol. 43, no. 10, pp. 2218–2228, Oct. 2008.

[59] *3G TS 25.141 Base station conformance testing (FDD)*, 3rd Generation Partnership Project; Technical Specification Group Radio Access Networks Technical Specification, Rev. V3.1.0, 2000.

[60] L. Larson and D. K. P. Asbeck, "Challenges and Opportunities for Compound Semiconductor Devices in Next Generation Wireless Base Station Power Amplifiers," in *Proc. IEEE Compound Semiconductor Integrated Circuit Symposium (CSIC'05)*, Nov. 2005.

[61] M. Iwamoto, A. Williams, P.-F. Chen, A. Metzger, C. Wang, L. Larson, and P. Asbeck, "An extended Doherty amplifier with high efficiency over a wide power range," *Microwave Symposium Digest, 2001 IEEE MTT-S International*, vol. 2, pp. 931–934 vol.2, 2001.

[62] N. Ui, H. Sano, and S. Sano, "A 80W 2-stage GaN HEMT Doherty Amplifier with 50dbc ACLR, 42% Efficiency 32dB Gain with DPD for WCDMA Base station," *Microwave Symposium, 2007. IEEE/MTT-S International*, pp. 1259–1262, June 2007.

[63] J. Kim, B. Kim, and Y. Y. Woo, "Advanced Design of Linear Doherty Amplifier for High Efficiency using Saturation Amplifier," June 2007, pp. 1573–1576.

[64] A. Long, J. Yao, and S. Long, "A 13 W current mode class D high efficiency 1 GHz power amplifier," vol. 1, Aug. 2002, pp. I–33–6 vol.1.

[65] H. Kobayashi, J. Hinrichs, and P. Asbeck, "Current mode class-D power amplifiers for high efficiency RF applications," vol. 2, 2001, pp. 939–942 vol.2.

[66] A. Al Tanany, A. Sayed, and G. Boeck, "A 2.14 GHz 50 Watt 60% Power Added Efficiency GaN Current Mode Class D Power Amplifier," Oct. 2008, pp. 432–435.

[67] N. Sokal and A. Sokal, "Class e-a new class of high-efficiency tuned single-ended switching power amplifiers," *Solid-State Circuits, IEEE Journal of*, vol. 10, no. 3, pp. 168–176, Jun 1975.

[68] S. Pajic, N. Wang, P. Watson, T. Quach, and Z. Popovic, "X-band two-stage high-efficiency switched-mode power amplifiers," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 9, pp. 2899–2907, Sept. 2005.

[69] A. Diet, C. Berland, M. Villegas, and G. Baudoin, "EER architecture specifications for OFDM transmitter using a class E amplifier," *Microwave and Wireless Components Letters, IEEE*, vol. 14, no. 8, pp. 389–391, Aug. 2004.

[70] D. Y. Lie, J. Popp, F. Wang, D. Kimball, and L. Larson, "Linearization of Highly-Efficient Monolithic Class E SiGe Power Amplifiers with Envelope-Tracking (ET) and Envelope-Elimination-and-Restoration (EER) at 900MHz," Nov. 2007, pp. 1–4.

[71] J. Popp, D. Lie, F. Wang, D. Kimball, and L. Larson, "Fully-integrated highly-efficient RF Class E SiGe power amplifier with an envelope-tracking technique for EDGE applications," Jan. 2006, pp. 231–234.

[72] J. Mitola, "The software radio architecture," *Communications Magazine, IEEE*, vol. 33, no. 5, pp. 26–38, May 1995.

[73] K. Buisman, L. de Vreede, L. Larson, M. Spirito, A. Akhnoukh, T. Scholtes, and L. Nanver, "Distortion-free varactor diode topologies for RF adaptivity," *Microwave Symposium Digest, 2005 IEEE MTT-S International*, pp. 4 pp.–, 12-17 June 2005.

[74] R. Malmqvist, A. Ouacha, and R. Erickson, "Multi-Band and Reconfigurable Front-Ends for Flexible and Multi-Functional RF Systems," Dec. 2007, pp. 1–4.

[75] N. Haridas, A. Erdogan, T. Arslan, A. Walton, S. Smith, T. Stevenson, C. Dunare, A. Gundlach, J. Terry, P. Argyrakis, K. Tierney, A. Ross, and T. O'Hara, "Reconfigurable MEMS Antennas," June 2008, pp. 147–154.

[76] R. Whatley, Z. Zhou, and K. Melde, "Reconfigurable RF impedance tuner for match control in broadband wireless devices," *Antennas and Propagation, IEEE Transactions on*, vol. 54, no. 2, pp. 470–478, Feb. 2006.

[77] L. Rabieirad and S. Mohammadi, "A reconfigurable MEMS-less CMOS tuner for software defined radio," June 2008, pp. 779–782.

[78] D. Kelly, C. Brindle, C. Kemerling, and M. Stuber, "The state-of-the-art of silicon-on-sapphire CMOS RF switches," Oct.-2 Nov. 2005, pp. 4 pp.–.

[79] A. Botula, A. Joseph, J. Slinkman, R. Wolf, Z.-X. He, D. Ioannou, L. Wagner, M. Gordon, M. Abou-Khalil, R. Phelps, M. Gautsch, W. Abadeer, D. Harmon, M. Levy, J. Benoit, and J. Dunn, "A Thin-Film SOI 180nm CMOS RF Switch Technology," Jan. 2009, pp. 1–4.

[80] S. Moloudi, K. Takinami, M. Youssef, M. Mikhemar, and A. Abidi, "An Outphasing Power Amplifier for a Software-Defined Radio Transmitter," Feb. 2008, pp. 568–636.

[81] Y. Y. Woo, J. Yi, Y. Yang, and B. Kim, "SDR transmitter based on LINC amplifier with bias control," vol. 3, June 2003, pp. 1703–1706 vol.3.

Summary

This thesis introduces several concepts with their experimental verification for the design of efficient and linear RF power amplifiers.

In **Chapter 1**, a brief historical overview of the wireless communication market is given together with the author's vision on the future of wireless communications with subsequent related challenges for RF power amplifier designers. The main targets for both handset power amplifiers and base-station power amplifiers are that they become more efficient, as well as, more linear. Besides that, for handset power amplifiers, there is additional requirement for multi-band operation with emphasis on the compactness of their implementation. For base-station power amplifiers the main emphasis is on increasing the power handling capability of the amplifier and its ever increasing efficiency requirement for signals that have high peak-to-average power ratios.

Chapter 2 introduces the basic concepts of linearity and efficiency associated with the power amplifier. The traditional amplifier classes of operation are discussed with regards to their efficiency performance, assuming an ideal active device. Following this discussion, the various dominant sources of distortion are identified for a practical device, and how this distortion increases when device operation is pushed towards higher efficiency.

To operate at higher efficiencies and still reach an acceptable linearity performance, **Chapter 3** introduces the use of Derivative Superposition as a technique to improve device linearity. Doing so, the resulting amplifier stage can be operated at higher power levels, while still meeting its linearity requirements, resulting in an improved efficiency performance. However, at very high power levels close to compression, Derivative Superposition is no longer effective. Consequently, to overcome this limitation, in this chapter, harmonic manipulation is introduced to shift the linearity "IM3 sweet-spot" of the device closer to compression. This yields improved linearity close to the compression point and allows therefore device operation at a reduced power back-off levels yielding improved efficiency. Applying this technique on a NXP Gen 6 LDMOS device results in an efficiency of 44% at an ACLR1 level of -45dBc for a IS-95 forward link signal.

Relying entirely on analogue circuit techniques to meet the high linearity requirements of present/future wireless communication standards, while simultaneously aiming for average efficiencies in excess of 50% for complex modulated signals, is close to impossible. As such, the trend in the industry is to "solve" the linearity issues of the power amplifier in the digital domain, a technique commonly referred to as digital predistortion. **Chapter 4** introduces the theoretical concepts involved in digital pre-

distortion and discusses the current techniques for predistortion for power amplifiers with and without memory effects.

Digital predistortion allows the linearity specifications of the RF power amplifier to be "relaxed", as the overall transmitter linearity can be improved digitally. As such, this opens up the possibility to use more aggressive power amplifier architectures such as LINC and Doherty, which although in theory are both linear and efficient, but in their practical implementation suffer from severe distortion. **Chapter 5** gives an overview of the common high-efficient power amplifier concepts reported in literature, and discusses their advantages and disadvantages.

Chapter 6 presents a multi-band, multi-mode class-AB power amplifier targeted for handset applications, which utilizes continuously tunable input and output matching networks integrated in a low-loss silicon-on-glass technology. The tunable matching networks make use of very high-Q varactor diodes in a low distortion anti-series configuration to achieve the desired source and load impedance tunability. A QUBIC4G high voltage breakdown transistor is used as active device. The realized adaptive amplifier provides 13 dB gain, 27-28 dBm output power at the 900, 1800, 1900 and 2100 MHz bands. For communication bands above 1 GHz optimum load adaptation is facilitated, resulting in efficiencies between 30%-55% over a 10 dB output power control range.

Chapter 7 presents a mixed-signal approach for the design and testing of high-performance N-way Doherty amplifiers. In support of this N-way power-combining networks are analyzed for their optimum design, by examining the relationship between the input drive conditions of the active devices and their efficient output power combining. This analysis makes no prior assumption on the network topology and facilitates free-to-choose levels for the high-efficiency power back-off points. By comparing the results of this analysis with prior work, it is shown that very specific drive conditions apply to traditional three-way Doherty amplifier implementations to obtain simultaneously high-efficiency and high-linearity operation. To demonstrate this approach, a three-way Doherty 100-W GaN base-station power amplifier at 2.14 GHz is designed and built. Mixed-signal techniques are utilized for uncompromised control of the amplifier stages to optimize efficiency, as well as, linearity. The combination of the above techniques resulted in an unprecedented high efficiency over a 12 dB power back-off range, facilitating a record high PAE for a W-CDMA test signal with high crest factor, while meeting all the spectral requirements for UMTS base stations.

Chapter 8 presents the conclusions of this thesis work and provides suggestions for further research and development in the field of RF power amplifiers for wireless communication applications.

Samenvatting

Dit proefschrift introduceert verschillende concepten voor efficiënte en lineaire RF vermogensversterkers. Deze concepten worden allemaal experimenteel geverifiëerd.

In **Hoofdstuk 1** wordt een kort historisch overzicht van de draadloze communicatie gegeven alsmede de visie van de auteur op de toekomst van draadloze communicatie en de daarmee samenhangende uitdagingen voor de ontwerpers van de RF vermogensversterkers. Het blijkt dat verbeterde efficiëntie en lineariteit de belangrijkste trends zijn voor zowel de vermogensversterkers van mobiele telefoons als die van basisstations. Verder is er voor de vermogensversterkers van de mobiele telefoon, de aanvullende eis voor multi-band multi-mode functionaliteit waarbij de nadruk ligt op de compactheid van de uiteindelijke versterker implementatie. Voor basisstations vermogensversterkers ligt de nadruk op het vergroten van het maximale vermogen dat de versterker kan leveren met behoud van een goede efficiëntie voor signalen die een groot verschil kennen in hun piek en gemiddelde vermogen.

Hoofdstuk 2 introduceert de basisconcepten voor lineariteit en efficiëntie m.b.t. de RF vermogensversterker. De traditionele versterkerklassen worden besproken samen met hun efficiëntie prestaties. In deze discussies wordt eerst een ideale transistor verondersteld waarna voor een echte transistor de dominante bronnen van distortie worden geïdentificeerd. Er wordt aangetoond dat de signaalvervorming toeneemt wanneer de transistor wordt gebruikt bij een hoger uitgangsvermogen om een beter rendement te verkrijgen.

Om een hogere efficiëntie te bereiken bij een aanvaardbare lineariteit, introduceert **Hoofdstuk 3** het gebruik van de "Derivative Superposition" techniek om de lineariteit te verbeteren. Door het benutten van deze techniek blijft de versterking lineair op een hoger vermogensniveau wat resulteert in een beter rendement. Op zeer hoge vermogensniveaus, dicht bij het P1dB compressiepunt, is de Derivative Superposition techiek niet meer effectief. Om deze beperking te overwinnen wordt harmonische manipulatie geïntroduceerd om het optimale IM3 punt dichter bij het P1dB compressiepunt te plaatsen. Hierdoor verbetert de lineariteit bij het P1dB compressiepunt waardoor men minder vermogensreductie hoeft te gebruiken voor de transistor. Dit leidt tot een beter rendement. Het toepassen van deze techniek op een NXP Gen 6 LDMOS transistor resulteert in een efficiëntie van 44% met een ACLR1 niveau van -45dBc voor een IS-95 verbindingssignaal.

Als men streeft naar een gemiddelde efficiëntie van meer dan 50% voor complexe gemoduleerde signalen dan is het gebruik van uitsluitend analoge circuittechnieken, om aan de lineariteit eisen van de huidige/toekomstige draadloze communicatiesys-

temen te voldoen, niet toereikend. Daarom is de algemene indruk dat digitale compensatietechnieken moet worden gebruikt om de lineariteits problemen van de vermogensversterker op te lossen. **Hoofdstuk 4** introduceert de theoretische concepten van digitale precompensatie en bespreekt de gangbare technieken voor precompensatie voor vermogensversterkers met en zonder geheugeneffecten.

Door het gebruik van digitale compensatietechnieken wordt de lineariteit specificatie van de RF vermogensversterker minder streng. Dit komt omdat nu de digitale precompensatie de lineariteit van de totale zender verbetert. Dit biedt de mogelijkheid om een meer vooruitstrevende versterker architectuur zoals LINC en Doherty te gebruiken, die weliswaar in theorie zowel lineair als efficiënt zijn maar in het praktijk aanleiding geven tot sterke signaalvervorming. **Hoofdstuk 5** geeft een overzicht van de meest gangbare hoog-efficiënte RF vermogensversterker architecturen die in literatuur worden vermeld en bespreekt hun voor- en nadelen.

Hoofdstuk 6 introduceert een multi-band, multi-mode klasse-AB vermogensversterker voor mobiele telefoon toepassingen. Dit versterker concept maakt gebruik van continu verstembare in- en uitgangs impedantie aanpasnetwerken, die zijn geïntegreerd in een silicium-op-glas technologie met zeer lage verliezen. De afstembare aanpasnetwerken maken gebruik van varactordiodes in een anti-series configuratie die vrijwel geen signaalvervorming geeft. Een QUBIC4G transistor met verbeterde "breakdown" eigenschappen wordt gebruikt als versterkend element. De gerealiseerde multi-band, multi-mode versterker levert 13 dB gain bij 27-28 dBm uitgangsvermogen in de frequentiebanden van 900, 1800, 1900 en 2100 MHz. Voor de communicatiebanden boven de 1 GHz kan voor ieder afzonderlijk vermogensniveau, d.m.v. het afstembare uitgangsnetwerk, de optimale impedantie worden aangeboden aan de transistor wat resulteert in 30%-55% efficiëntie over een 10 dB vermogensbereik.

Hoofdstuk 7 introduceert een mixed-signaal benadering voor het ontwerp en het testen van "N-way" Doherty vermogensversterkers. Dit hoofdstuk begint met een algemene analyse voor het ontwerp van "N-way" vermogencombinerende netwerken. De analyse is gebaseerd op het verband tussen de optimale aansturing van de actieve elementen en het combineren van hun uitgangsvermogens d.m.v. een verliesvrij netwerk. De analyse maakt geen enkele aanname wat betreft de netwerktopologie en laat de ontwerper vrij om de "back-off" punten, waar de versterker zijn hoge efficiëntie bereikt, te kiezen. Door deze analyse met eerder werk te vergelijken wordt aangetoond dat er zeer specifieke aanstuurvoorwaarden zijn om de traditionele "3-way" Doherty versterker goed te laten werken met een hoog rendement en lineariteit. Om dit aan te tonen wordt een "3-way" Doherty van GaN met 100-Watt uitgangsvermogen voor base-stations bij 2.14 GHz ontworpen en gebouwd. De mixed-signaal techniek is gebruikt om de individuele versterkertrappen optimaal aan te sturen om zowel een goed rendement als lineariteit te bereiken. De combinatie van de bovengenoemde technieken resulteerde in een tot nu toe niet geevenaarde efficiency over een 12 dB vermogens bereik t.o.v. het P1dB compressie punt. Dit resulteerde in een nieuw record voor de PAE van een W-CDMA testsignaal met hoge "crest-factor" terwijl voldaan wordt aan alle spectrale eisen van de UMTS specificatie.

Hoofdstuk 8 geeft een overzicht van de belangrijkste conclusies van dit proefschrift en geeft nieuwe suggesties voor verder onderzoek en ontwikkelingen op het gebied van de RF vermogensversterkers voor draadloze communicatie toepassingen.

List of Publications

Journal Papers

W. C. E. Neo, Y. Lin, X. Liu, L. C. N. de Vreede, L. E. Larson, M. Spirito, M. J. Pelk, K. Buisman, A. Akhnoukh, A. de Grauw, and L. K. Nanver, "Adaptive Multi-Band Multi-Mode Power Amplifier Using Integrated Varactor-Based Tunable Matching Networks," *IEEE J. Solid-State Circuits*, vol. 41, no. 9, pp. 2166–2176, Sept. 2006.

W. C. E. Neo, J. H. Qureshi, M. J. Pelk, J. R. Gajadharsing, and L. C. N. de Vreede, "A Mixed-Signal Approach Towards Linear and Efficient N-Way Doherty Amplifiers," *IEEE Trans. Microwave Theory Tech.*, vol. 55, no. 5, pp. 866–879, May. 2007.

Conference Papers

W. C. E. Neo, M. P. van der Heijden, L. C. N. de Vreede, M. Spirito, V. Cuoco and F. van. Rijs, "A technique to linearize LDMOS power amplifiers based on derivative superposition and out-of-band impedance optimization," in *Eur. Microwave Conf.*, vol. 3, Oct. 2004, pp. 1173–1175.

W. C. E. Neo, X. Liu, Y. Lin, L. C. N. de Vreede, L. E. Larson, M. Spirito, A. Akhnoukh, A. de Grauw, and L. K. Nanver, "Improved hybrid SiGe HBT class-AB power amplifier efficiency using varactor-based tunable matching networks," in *Proc. IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM2005)*, Oct. 2005, pp. 108–111.

Co-Authorship

V. Cuoco, **W. C. E. Neo**, L. C. N. de Vreede, H. C. de Graaff, L. K. Nanver, K. Buisman, H. C. Wu, H. F. F. Jos, and J. N. Burghartz, "A new extraction technique for the series resistances of semiconductor devices based on the intrinsic properties of bias-dependent y-parameters," in *Proc. IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM2004)*, Sep. 2004, pp. 148–151.

V. Cuoco, **W. C. E. Neo**, M. Spirito, O. Yanson, N. Nenadovic, L. C. N. de Vreede, H. F. F. Jos, and J. N. Burghartz, "The electro-thermal Smoothie database model for LDMOS devices," in *Proc. ESSDERC 2004*, Sept. 2004, pp. 457–460.

V. Cuoco, O. Yanson, P. Hammes, M. Spirito, L. C. N. de Vreede, A. van Steenwijk, M. Versleijen, **W. C. E. Neo**, H. F. F. Jos, and J. N. Burghartz, "Large signal verification of the circuit-oriented smoothie database model for LDMOS devices," in *Eur. Microwave Conf.*, vol. 1, Oct. 2004, pp. 217–220.

M. J. Pelk, **W. C. E. Neo**, J. R. Gajadharsing, R. S. Pengelly, and L. C. N. de Vreede, "A High-Efficiency 100-W GaN Three-Way Doherty Amplifier for Base-Station Application," *IEEE Trans. Microwave Theory Tech.*, vol. 56, no. 7, pp. 1582–1591, Jul. 2008.

M. Marchetti, M. J. Pelk, K. Buisman, **W. C. E. Neo**, M. Spirito, and L. C. N. de Vreede, "Active Harmonic Load-Pull with Realistic Wideband Communications Signals," *IEEE Trans. Microwave Theory Tech.*, vol. 56, no. 12, pp. 2979–2988, Dec. 2008.

J. H. Qureshi, M. J. Pelk, M. Marchetti, **W. C. E. Neo**, J. R. Gajadharsing, M. P. van der Heijden, and L. C. N. de Vreede, "A 90-W Peak Power GaN Outphasing Amplifier With Optimum Input Signal Conditioning," *IEEE Trans. Microwave Theory Tech.*, vol. 57, no. 8, pp. 1925–1935, Aug. 2009.

Acknowledgments

The PhD experience in Delft will become one of the most memorable periods of my life, and I am very fortunate that this opportunity was open up to me. This journey began when Prof. Kees Beenakker came to Singapore to recruit students for his now famous Master of Science program in Microelectronics at the Delft University of Technology. Thank you, Prof. Kees Beenakker for selecting me for this program.

After my Master of Science, I was very fortunate to be offered a PhD position by Dr Leo de Vreede in the group of then Prof. Joachim Burghartz (my promotor). I am very grateful to both of them for having the confidence that I can undergo the grueling 4 years of a PhD student. Dr. Leo de Vreede who then became my supervisor (co-promotor) for my PhD studies, did not only play a big role in my PhD thesis work, he also acted as role model for a good engineer, scientist and teacher, which I strive to become. Leo is someone who will never say "no"when approached for discussions, I benefited a lot from these discussions, which led eventually to the many new concepts that are presented in this thesis work. Thank you Leo for the great time I had during the PhD, it was not an easy time, but a very memorable period. In this PhD work, I have received excellent support from NXP. Not only was this project financially supported by them, I also received a lot of technical input, as well as, market input from them. Here, I would like to acknowledge in special John Gajadharsing and Rik Jos for their strong support over the last few years and their interest in this thesis work.

Contrary, to the popular belief that a PhD study is very lonely work where one spends long hours alone in the laboratory or behind the computer, my PhD experience is completely the opposite, as I have spend a big proportion of my time collaborating with other people. In the beginning of my PhD, I benefited from working closely with Mark van der Heijden on improving the linearity of LDMOS devices. Also, Marco Spirito provided me with his support when characterizing devices on his "famous" Load-pull setup. Not to forget Vittorio Cuoco, who taught me about modeling and characterization of the transistors. Subsequently, I got to work with Marco Pelk, who was the guy we all went to, if we had technical problems, no matter software or hardware. Together with Marco Pelk, we put many hours into the mixed-signal based test bench, which also has DPD capabilities. I will never forget the "hardship" we faced in that laboratory, in getting everything to work together, as well as the many dinners we had at Burger King during our evening break. Marco, I am indebted to you for all the help and want to thank you for your friendship, a great proportion of this work would never see light without your collaboration. I also like to acknowledge

my office room-mate and fellow PhD, Koen Buisman. Besides providing me with hours of interesting discussions, Koen also designed those very linear RF tunable varactors, which was crucial for the adaptive PA in my thesis work. Here, I will also like to acknowledge and thank Prof. Lis Nanver and the DIMES technology team for making those varactors. Also a special word of thanks has to be given to Atef Aknouk, who helped me in many ways to ensure that everything was operating smoothly in the measurement laboratory.

Coming from a foreign country and living in The Netherlands can be a very challenging task. Fortunately for me, I had the support from our ever helpful secretaries. Here, I will like to acknowledge; Birgit Rademakers, Pien Rijnink, Marysia Lagendijk-Korzeniewski who definitely went beyond the scope of their job to help me with issues concerning housing, health insurance, taxes etc, etc. Last but not least, I also want to acknowledge my other fellow PhDs, Huang Cong, Jawad Qureshi and Mauro Marchetti. The company of you guys definitely made the PhD experience much more "tolerable".

A great part of the thesis write up was done, after I left the university while working full-time in NXP. This would not have been possible without the huge support and understanding of; Mark Murphy, Rik Jos, Korné Vennema and John Gajadharsing. I also like to acknowledge Sjoerd van Nederveen, Hans Kartman, Harrie Rahangmetan, and Ton Schouten, my colleagues at NXP, for helping me make that transition from university to company life, and their support and understanding while I was writing my PhD thesis.

During my PhD studies, I was very fortunate to have the support of my family and friends. I will like to acknowledge my family for all their love and encouragement. Their encouragements have kept me going through the tough times. At this juncture, I will like to acknowledge and thank my mentor Ruth Winden. I have benefited a lot from her advice, guidance and encouragement through these years. I also like to thank my partner Shannon Koh for her love and patience, writing this thesis did not put me in the best of moods.

For my final acknowledgement, I like to dedicate this PhD thesis to my late mother Mak Yoke, who passed away during the writing of this thesis.

About the Author

W. C. Edmund Neo was born in Singapore in 1977. He received the B.Eng. degree in electrical engineering with First Class honors from the National University of Singapore in 2002. He subsequently pursued his master degree in electrical engineering in Delft, and in 2004 he was awarded the M.Sc. degree in electrical engineering (cum laude) from the Delft University of Technology, Delft, The Netherlands. From 2004 to 2008 he worked towards his Ph.D. degree in electrical engineering at the Delft University of Technology in the area of novel circuit design techniques for high-efficiency power amplifiers. Since mid 2008, he joined NXP Semiconductors in Nijmegen where he holds the responsibility as application engineer in the RF power transistor product line.