Broadband Doherty Power Amplifiers with Enhanced Linearity for Emerging Radio Transmitters

by

Hamed Golestaneh

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2016

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Hamed Golestaneh

Statement of Contributions

The work presented in Chapter 6 of this thesis was the result of collaboration with Kasyap Patel, who validated the proposed current contour based design methodology by designing a proof-of-concept circuit prototype (Section 6.5).

Abstract

The ever-increasing demand for utilizing wireless spectra has led to development of spectrally efficient radio systems. While these systems offer much higher data throughput, they employ more sophisticated modulation schemes, which result in wideband signals with high peak-to-average power ratios. These signal characteristics significantly complicate the design of RF transmitters, particularly power amplifiers, in terms of power efficiency and linearity requirements. Furthermore, upcoming wireless standards, such as long term evolution advanced (LTE-A) require adoption of carrier aggregation which incorporates multiple component carriers to yield aggregated channels of larger bandwidth (up to 100 MHz). On the other hand, the emerging systems are expected to support legacy standards with minimum area, cost, and power overhead, and thus call for highly-efficient linear broadband power amplifiers capable of efficiently amplifying concurrent modulated signals located over a broad carrier frequency range.

This thesis focuses on Doherty power amplifiers (DPAs) with extended high-efficiency range, enhanced bandwidth and improved linearity as a solution for high-efficiency multi-band multi-standard transmitters. It addresses three major concerns associated with DPAs, namely, back-off efficiency, bandwidth, and linearity. The Thesis begins with a detailed theoretical analysis of two-way and three-way Doherty configurations from which the governing equations are derived. This is followed by a comprehensive study of bandwidth limitation in DPA variants.

As the first contribution, it is shown that the two existing three-way Doherty structures, i.e., conventional and modified DPAs have inherently broadband characteristics and thus are promising solutions for multi-standard base station transmitters. As a proof of concept, a 30-W three-way modified Doherty amplifier was designed and implemented using packaged GaN transistors over 0.73–0.98 GHz. The prototype was successfully linearized under modulated signals with up to 20 MHz modulation bandwidth.

To further improve the linearizability of the DPAs under wideband and multi-band modulated signals, this thesis investigates major sources of static and dynamic nonlinearity in two-way DPAs both at device and circuit levels and explores circuit techniques to mitigate them. Furthermore, the challenges of applying the Doherty technique for concurrent transmission of multiple modulated signals are tackled.

The most significant contribution of this thesis is to develop a novel waveform engineering approach to designing ultrawideband DPAs. This approach completely reformulates the DPA's output combiner conditions in order to accommodate complex-valued load modulation. Moreover, it relaxes the harmonic termination requirements of the DPAs

to further enlarge the Doherty design space, thereby enhancing the bandwidth. A 50-W waveform-engineered two-way DPA prototype was designed for 1.5–2.5 GHz range and was successfully linearized under intra- and inter-band carrier-aggregated signals with up to 600 MHz carrier spacing.

Lastly, an input matching network design methodology is proposed for broadband DPAs. This methodology uses the novel concept of "current contours" to minimize the bandwidth, efficiency and linearity degradation of DPAs caused by device input non-idealities.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. Slim Boumaiza for his priceless help, support, and guidance throughout my PhD. I sincerely thank him for teaching me various aspects of independent research. I would also like to thank my committee members, Prof. de Vreede, Prof. Long, Prof. Safavi, and Prof. Nieva for reviewing my thesis and providing insightful feedback.

Over the past four years, I have been blessed with wonderful friends in the EmRG group who always provided me with great encouragement and cheer. It is my great pleasure to take this opportunity to also thank all my friends in Canada. A special thanks to Sajjad Mosharraf and Kamal Rahimi for their immense support and companionship.

Last but not least, I would like to thank my parents for their endless love, support, and prayers throughout the years. There are no words that can express my gratitude and appreciation for all you have done for me.

Dedication

To my loving family

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2W-DPA	Two-Way Doherty power amplifier
3G	Third generation
3W-CDPA	Conventional three-Way Doherty power amplifier
3W-DPA	Three-Way Doherty power amplifier
3W-MDPA	Modified three-Way Doherty power amplifier
4G	Fourth generation
5G	Fifth generation
AC	Alternating current
ACLR	Adjacent channel leakage ratio
ADS	Advanced design system
AM/AM	Amplitude to amplitude modulation
AM/PM	Amplitude to phase modulation
CMOS	Complementary metal-oxide-semiconductor
CW	Continuous wave
DC	Direct current
DE	Drain efficiency
DLM	Dynamic load modulation

DPA	Doherty power amplifier
DPD	Digital pre-distortion
ET	Envelope tracking
EVM	Error vector magnitude
FBW	fractional bandwidth
FET	Field effect transistor
GaAs	Gallium arsenide
GaN	Gallium nitride
HBT	Heterojunction bipolar transistor
HEMT	High electron mobility transistor
IMN	Input Matching Network
ITR	Impedance transformation ratio
LDMOS	Laterally diffused metal oxide semiconductor
LTE	Long term evolution
MEMS	Micro-electro-mechanical systems
MMIC	Monolithic microwave integrated circuit
OFDM	Orthogonal frequency division multiplexing
PA	Power amplifier
PAE	Power added efficiency
PAPR	Peak-to-average power ratio
PDF	Probability density function
RF	Radio frequency
SOI	Silicon on insulator

SRFT	Simplified real frequency technique
VCCS	Voltage-controlled current source
VCVS	Voltage-controlled voltage source
WCDMA	Wideband code division multiple access
WeDPA	Waveform-engineered Doherty power amplifier
WiMAX	Worldwide Interoperability for Microwave Access

Chapter 1

Introduction

1.1   Motivation

The growing demand for multimedia services via wireless communications with higher throughput necessitates modern wireless networks to support spectrally-efficient modulation schemes and wideband signals. For instance, emerging 4G standards employ multi-carrier multiplexing techniques such as orthogonal frequency division multiplexing (OFDM) with overlapping carriers to optimally use the spectrum. This, in turn, has led to large variation in the envelope of the signal waveform, namely, characterized as peak-to-average power ratio (PAPR). This signal characteristic significantly increases the design complexity of the underlying transmitters, particularly the constituent radio frequency power amplifiers (RF PAs) to be deployed in modern wireless infrastructure. On the one hand, the PA is required to linearly amplify the input signal so as to maintain signal quality and minimize out-of-band emissions in adjacent channels, and thus it must be backed off from the nominal output power. On the other hand, statistics suggest that the signal is mostly transmitted at much lower power levels (see Figure 1.1), meaning that the PA operates at its inefficient region most of the time. Figure 1.2 illustrates the drain efficiency of an ideal class-B PA, as a function of the power back-off. It shows that, for a signal with PAPR = 10 dB, for example, the instantaneous drain efficiency is most often around 25% (53.5% lower than the peak efficiency), which substantially degrades the average efficiency of the PA.

Furthermore, upcoming networks are expected to handle legacy communication standards. This calls for multi-standard radio systems that are capable of operating over a broad range of carrier frequencies and of concurrently processing various signals encoded

(a) Envelope waveform indicating a high PAPR signal.

(b) Probability density function (PDF) of the signal.
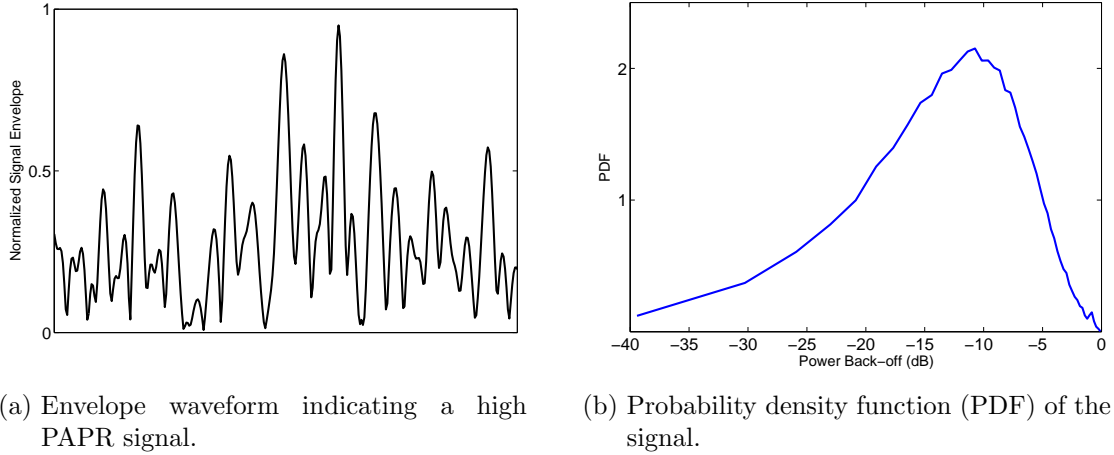
Figure 1.1: Characteristics of a 10 MHz long term evolution (LTE) signal.
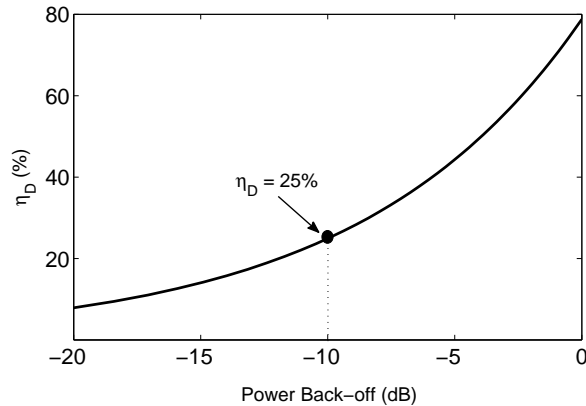
Figure 1.2: Drain efficiency of an ideal class-B power amplifier vs. power back-off

according to dissimilar standards. On top of that, communication signals required by the 3GPP specifications for Long Term Evolution Advanced (LTE-A) will have to utilize carrier aggregation to meet the need for transmission bandwidth. Carrier aggregation refers to the use of more than one spectrum portion to deploy multiple component carriers in order to obtain aggregated channels of wider transmission bandwidth (up to 100 MHz). Figure 1.3 illustrates the three carrier aggregation scenarios envisaged: (i) intra-band contiguous carrier aggregation (contiguous component carriers aggregated in the same operating band), (ii) intra-band non-contiguous carrier aggregation (non-contiguous carriers
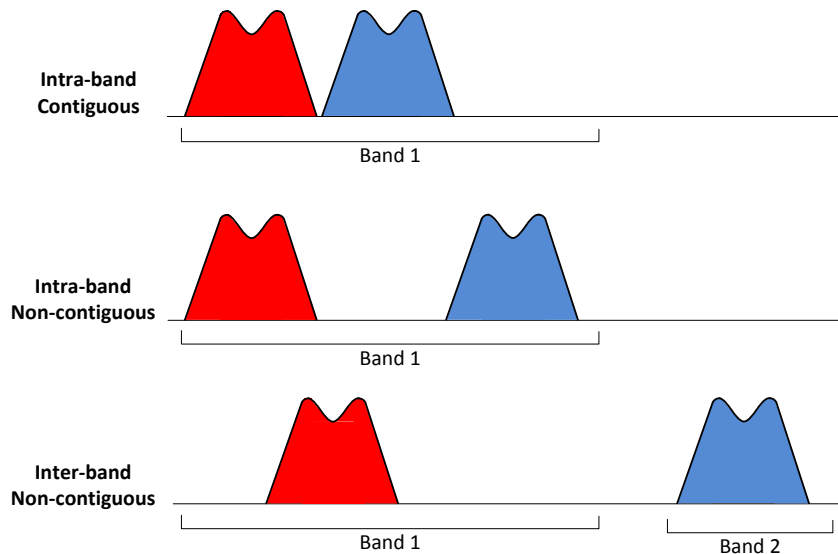
Figure 1.3: Three options for carrier aggregation.

aggregated in the same operating band) and (iii) inter-band carrier aggregation (carrier aggregation of component carriers in different operating bands which may be contiguous or non-contiguous within each band). These scenarios impose additional challenge for designing high efficiency linear PAs. In fact, a single RF PA is expected to efficiently amplify multi-band high-PAPR signals concurrently. In addition, the same RF PA must comply with the linearity requirements imposed by wireless standards.

In conclusion, despite all the advantages provided by the the modern wireless standards, more challenges are also introduced in designing RF PAs, in terms of linearity, power efficiency, and fractional bandwidth (FBW), which are yet to be addressed.

1.2 Thesis Objectives

To tackle the challenges discussed in Section 1.1, various efficiency enhancement techniques were proposed in the literature, among which Doherty power amplifier (DPA) has gained continued interest in recent years. Although primarily employed for narrowband amplification, the DPA has also shown potential for broadband transmitters. Nevertheless, few publications have tackled concurrent amplification of multi-band signals. This limits the

application of such wideband or multi-band DPAs for carrier aggregated signals, such as those used in LTE-A.

This thesis is thus focused on DPAs with extended FBW and enhanced linearity for base-station applications, including both theoretical analysis and implementation aspects. The objectives of the thesis are enumerated below:

1. Study of bandwidth limitation in two- and three-way DPAs. The analysis is initially based on an ideal transistor model and is later extended to include the transistor non-idealities.

2. Developing a three-way DPA (3W-DPA) with extended FBW ($\sim$30%), suitable for efficient amplification of signals with PAPRs up to 12 dB. Two distinct architectures, i.e., conventional and modified DPAs are analyzed in terms of FBW capability. An approach is also devised to compensate for the device parasitics. Finally, the theoretical analyses are validated by developing a proof-of-concept prototype.

3. Analysis of static and dynamic sources of distortion in broadband DPAs and investigation of techniques to minimize/mitigate them. This study provides guidelines for designing broadband DPAs with enhanced linearity and linearizability when driven with carrier-aggregated signals.

4. Developing a waveform engineering approach to enhance the bandwidth of the output combiner in 2W-DPAs far beyond the conventional bandwidth extension techniques proposed in the literature (i.e. more than 50%) and with minimum linearity compromise. This approach combined with guidelines from 3, enables efficient amplification of inter-band carrier-aggregated signals with arbitrary carrier spacings across the bandwidth.

5. Developing a systematic approach to designing input matching networks for DPAs based on the linearity analysis conducted in 3. While the input matching design problem has not been addressed in literature, the proposed design technique enables the DPA to reach the full potential FBW of the output combiner obtained in item 4, with minimum compromise in linearity/efficiency.

The power, bandwidth, and linearity requirements of the DPAs in this thesis for CA signal transmission are as follows:

- Average output power > 1 W under modulated signals with PAPR of 6–12 dB

- FBW $> 50\%$

- Adjacent channel leakage ratio (ACLR) < -45 dB

- Error vector magnitude (EVM) $< 5\%$.

Note that the linearity specifications must be met after applying an existing linearizer [e.g., digital pre-distortion (DPD)] to the DPA.

1.3   Thesis Outline

This thesis is organized as follows: Chapter 2 overviews some of the existing efficiency enhancement techniques, in which state-of-the-art DPAs are particularly discussed. Bandwidth analysis of DPAs and asymmetrically-biased DPAs are presented in the light of recent publications.

Chapter 3 presents novel analysis and design methodologies of three-way Doherty architectures devised to improve efficiency over a broad bandwidth and back-off range. The efficiency-bandwidth capability of these topologies is also compared. Based on the presented analysis, a proof-of-concept prototype was designed and fabricated with excellent efficiency improvement across 29% fractional bandwidth over a wide back-off range (9–10 dB).

Chapter 4 investigates static and dynamic sources of distortion in DPAs. The static nonlinearities on device and circuit level are detailed and how different sources contribute to the overall nonlinearity. Moreover, sources of memory effects are discussed along with techniques to minimize/mitigate them. The ultimate goal of the chapter is to provide design guidelines to ensure linearizability of a DPA under different carrier-aggregation scenarios.

Chapter 5 proposes a novel approach for designing ultrawideband output combiners for DPAs using the concept of waveform engineering. Theoretical analysis as well as practical considerations are detailed. Based on the proposed approach, a broadband linearity-enhanced 2W-DPA is designed and fabricated with 50% fractional bandwidth. The overall

DPA's bandwidth was limited due to the challenges of designing broadband input matching networks.

Chapter 6 presents a novel methodology to design input matching networks for broadband DPAs, taking advantage of the linearity study conducted in Chapter 4. The proposed technique is validated through simulation and measurement results.

Finally, Chapter 7 concludes the thesis by summarizing the contributions of this work, as well as suggesting potential research directions for future.

Chapter 2

High Efficiency Multi-Standard RF Transmitters

2.1 Introduction

This chapter begins with describing an ideal transistor model used throughout this proposal for RF amplifier analysis. This is followed by introducing typical high power field-effect transistor (FET) technologies for base stations and their large-signal models. The rest of the chapter expounds on prevalent techniques for improving the efficiency of RF transmitters and their suitability for multi-standard radio. The main focus of this chapter is on Doherty technique and its usability for broadband/multiband operation in the light of recent publications. It also presents an analysis of the conventional two-way DPA to develop the governing equations. These equations provide a better understanding of the load modulation mechanism which is required to analyze more advanced DPA architectures, such as asymmetrically biased DPAs where the drains of the main and peaking transistors are biased differently. Finally, bandwidth analysis of the two-way DPAs are presented in light of recent publications.

2.2 Ideal FET Model

FET devices, the currently dominant transistors in base station PAs, in the most ideal case consist of a linear voltage-controlled current source (VCCS) when operating in saturation region in common source configuration, as shown in Figure 2.1. This means that the
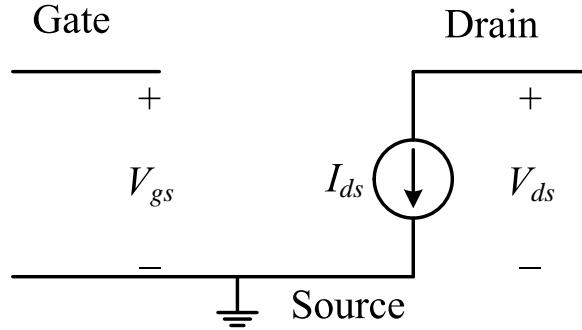
Figure 2.1: Ideal FET model.
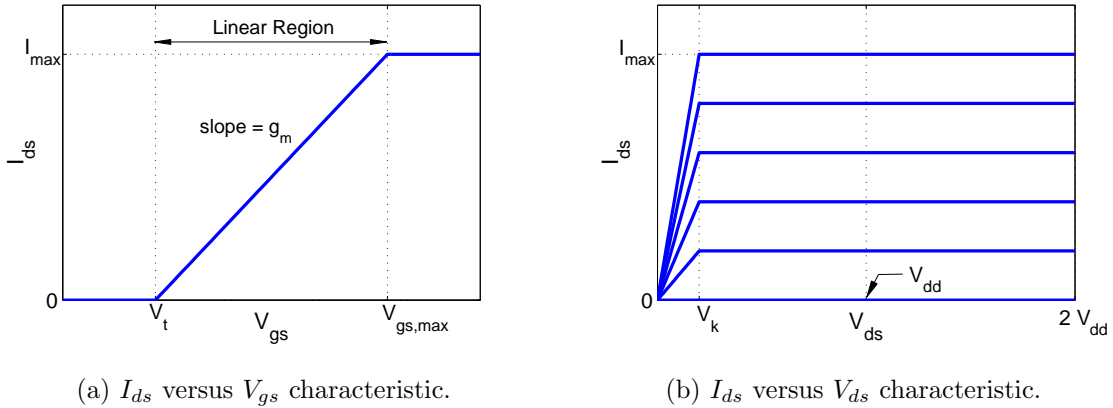
(a) I_{ds} versus V_{gs} characteristic.

(b) I_{ds} versus V_{ds} characteristic.

Figure 2.2: DC characteristics of an ideal transistor.

input voltage swing, V_{gs}, is unilaterally converted to the drain (output) current, I_{ds}, by a constant bias-dependent transconductance, g_m. The output current is in turn transformed into the output voltage through a finite load resistance. Note that unlike a small-signal amplifier, the transistor may move to the cut-off or knee region back and forth along a loadline, depending on the chosen class of operation. Figures 2.2(a) and (b) illustrate the profile of the instantaneous I_{ds} as a function of V_{gs} and V_{ds}, respectively. As can be seen, linear amplification occurs when $V_t < V_{gs} < V_{gs,max}$ and $V_k < V_{ds} < 2V_{dd}$, where V_t and V_k represent the threshold and knee voltage, respectively, and V_{dd} is the drain supply voltage.

As can be seen, the model includes no parasitic elements, and thus the input and output impedances are infinity. Furthermore, the ideal model does not consider the feedback or Miller effect, g_m nonlinearity as a function of V_{gs}, channel length modulation, and transistor

8
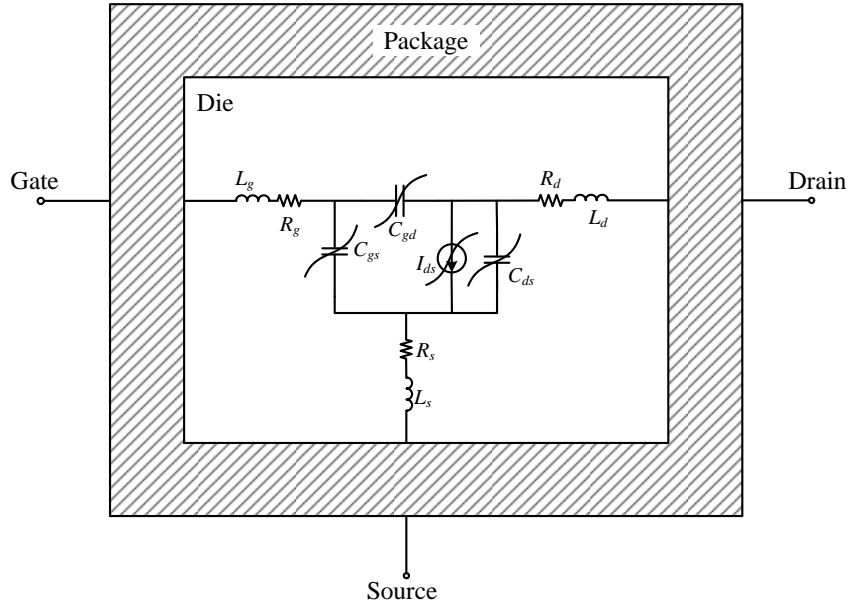
Figure 2.3: A generic transistor large-signal model.

breakdown effects. However, this model can be used to illustrate the harmonic components of the current at the biasing point, the nonlinearities associated with the class of operation (such as the soft turn-on effect of a class-C-biased transistor) and the knee effect.

2.3   Device Technologies

Even though the ideal FET model is often very helpful for theoretical analysis purposes, it has to be replaced by an accurate large-signal model for the device, in a practical design. The base-station power amplifier market is dominated by two major technologies, i.e., laterally diffused metal oxide semiconductor (LDMOS) and GaN high electron mobility transistor (HEMT) devices. A generic large-signal bare die model that applies to both technologies is depicted in Figure 2.3. As can be seen, other than the nonlinear drain current source, $I_{ds}(V_{gs}, V_{ds})$, the model includes some parasitics, which are often divided into intrinsic and extrinsic elements. Intrinsic elements, including C_{gs}, C_{gd}, and C_{ds} are associated with the device active channel, where the fundamental transistor action occurs,

and are nonlinear functions of "intrinsic" gate/drain terminal voltage in general. The extrinsic components, on the other hand, are bias-independent passives that connect the active element to the outside world, including manifolds, bond-pads, bondwires, etc. [2]. Commercial transistors are often embedded in a package, as shown in Figure 2.3 for the sake of protection and better thermal dissipation. The package introduces additional parasitics, usually modelled by EM simulations. Sometimes the package includes other integrated passives, known as pre-matching circuitry, to facilitate impedance matching over a certain frequency range.

Figure 2.4 illustrates the drain current of typical LDMOS and GaN transistors, as compared with ideal characteristics shown in Figure 2.2. It is observed that both devices deviate from the ideal case, due to the gradual turn-on (expansion) and gradual saturation (compression). Figures 2.4(a) and (b) also provide a better understanding of the nonlinear I-V characteristics of the transistors by illustrating the small-signal g_m profile as a function of V_{gs}. More discussions on the nonlinear profile of the intrinsics elements and their impacts on the PA performance will be presented in Chapter 4.

2.4 Efficiency Enhancement

The devices introduced in Section 2.3 serve as the core of the RF PA. As illustrated in Chapter 1, the average efficiency of a linear power amplifier (operating in class B or AB) when employed in modern radio transmitters is very low due to the high PAPR of the signals being used in those systems. Various techniques has thus been proposed in the literature to enhance the efficiency of the amplifier at power levels backed off from the peak value.

The basic idea of improving the back-off efficiency arises from the fact that the efficiency of a class-B biased PA is linearly proportional to the ratio of the drain voltage swing of the transistor to the drain supply voltage. Assuming the ideal transistor model described in Section 2.2 and a load impedance, R_L, the amplitude of the output voltage swing follows Ohm's law, i.e.,

$$V_1 = I_1 \cdot R_L \tag{2.1}$$

where I_1 and V_1 denote the fundamental drain current and voltage components, respectively, and R_L is the load resistance (at fundamental frequency). Generally, R_L is selected to achieve the optimum performance at the peak power, thus is referred to as the optimum load impedance of the transistor, R_{opt}. In other words, setting $R_L = R_{opt}$ results in full rail-to-rail voltage swing at the nominal peak current of the device, meaning that the
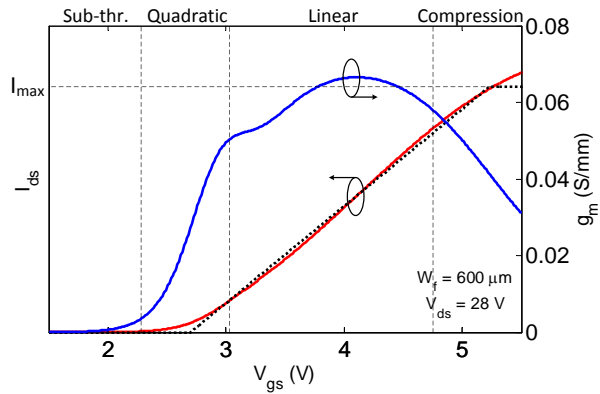
(a) I_{ds} and g_m vs. V_{gs} characteristics of an LDMOSFET.
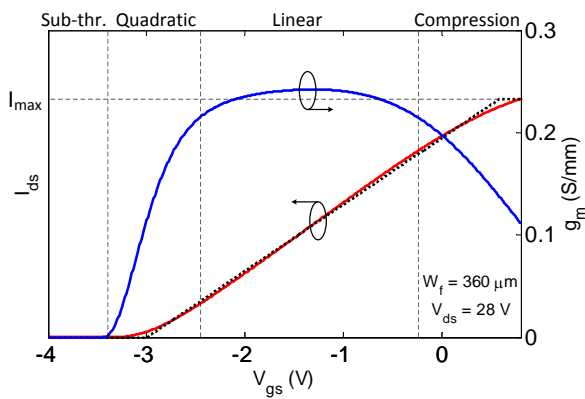
(b) I_{ds} and g_m vs. V_{gs} characteristics of a GaN HEMT.
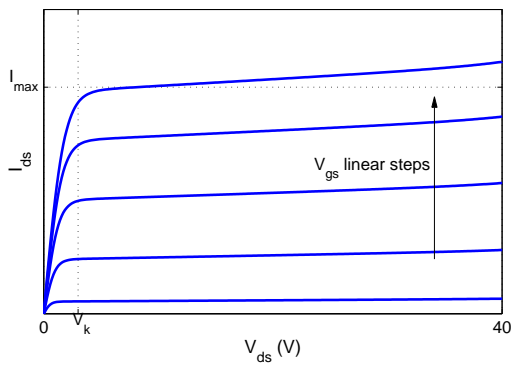
(c) DC-IV characteristics of LDMOSFET.
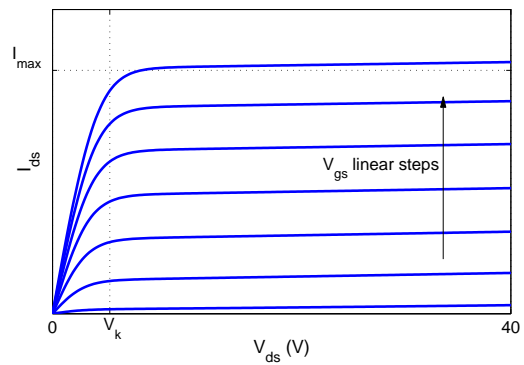
(d) DC-IV characteristics of GaN HEMT.

Figure 2.4: Drain current characteristics of a typical LDMOS and GaN transistor.

maximum efficiency coincides with the peak output power. R_{opt} for a particular device is given by,

$$R_{opt} = \frac{(V_{dd} - V_k)^2}{2P_{out,max}}. \tag{2.2}$$

Eq. (2.2) is the origin of various techniques devised to maximize efficiency at a desired power level. These techniques are broadly divided into two main categories. In the first one, the load resistance R_L is preserved as the optimum load impedance over the entire power range. In fact, the transistor drain supply voltage is varied as a function of P_{out}, i.e., proportional to the signal envelope, according to (2.2). These techniques are known as drain supply modulation techniques. Unfortunately, the efficiency improvement is significantly reduced at high back-off levels due to the nonzero knee voltage of the device. Alternatively, one could modulate the effective load impedance of the transistor (also called the load impedance "seen" by the transistor) to maintain the maximum voltage swing and efficiency over a given power range. The latter group are called the load modulation techniques, the most prominent of which is the Doherty architecture. The two groups of techniques will be reviewed in this section.

2.4.1   Drain Modulation Technique

As mentioned before, in this technique the drain supply voltage is adjusted as a function of input power to reduce the DC power consumption at back-off levels. There are various architectures based on drain modulation technique. Envelope elimination and restoration (EER) transmitter combines high-efficiency but nonlinear RF PAs with a linear envelope amplifier to realize a highly efficient linear transmitter [3]. In this transmitter, the envelope is eliminated from the RF signal fed into the PA, and then restored by drain modulation of the PA. Therefore, the linearity of the systems is directly dependent on the linear operation of the envelope amplifier and the phase alignment between the two paths.

Envelope tracking (ET) system is one of the most widely-used transmitters based on the drain modulation technique. Unlike the EER system, ET employs a linear RF PA, whose drain supply voltage is modulated by the signal envelope to maintain high efficiency in back-off levels, as illustrated in Figure 2.5. Ideally, except for the operation in knee region, the linearity of the transmitter is determined only by the linear PA, which relaxes the design of envelope modulator as compared to the EER system. However, similar to an EER system, the bandwidth and efficiency of the envelope amplifier is critical.

Recent publications on ET, have mainly focused on linearity and efficiency improvement of the envelope modulator for multi-standard transmitters. As the drain supply decreases
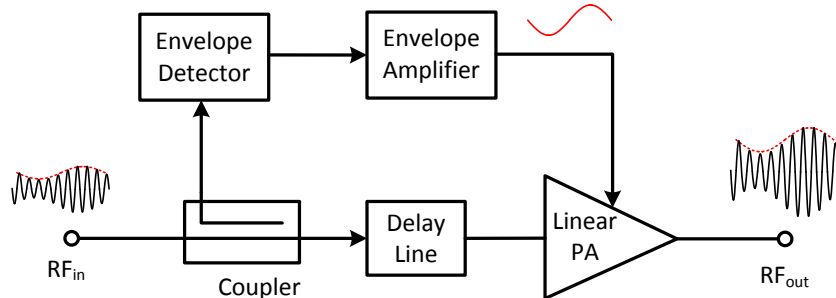
Figure 2.5: Simplified envelope tracking block diagram.

at low power levels, the device is more likely to intrude into the knee region which results in significant nonlinearity and memory effects when the system is driven with wideband signals. For instance, authors in [4, 5] suggested shaping of the original envelope to avoid the knee region at the expense of decreased power efficiency. Even though ET technique has been recently considered for base-station transmitters [6, 7], the application of ET systems are mainly limited to handset applications.

2.4.2   Load Modulation Technique

In the load modulation technique, load impedance of the device is varied as a function of power to enhance the back-off efficiency. Dynamic load modulation (DLM) technique [8] uses electronically tunable output matching network to dynamically modulate the impedance according to the signal envelope, and thus requires load-pull measurements/simulations to determine the optimum impedance trajectory vs. power, and varactors to realize the tunable matching networks. Thus far, variants of DLM technique has been proposed in the literature. For instance, the PA has been implemented in class-E [8, 9] or class-J [10] configurations, and has been successfully designed for broadband operation [11]. However, the linearity of the transmitter is a major concern and the technique has not been successfully deployed to amplify modulated signals yet.

Doherty power amplifier (DPA) is one of the most widely-adopted transmitters based on load modulation technique, for both handset and base station applications. A detailed overview of the Doherty technique will be presented in Section 2.5.

13
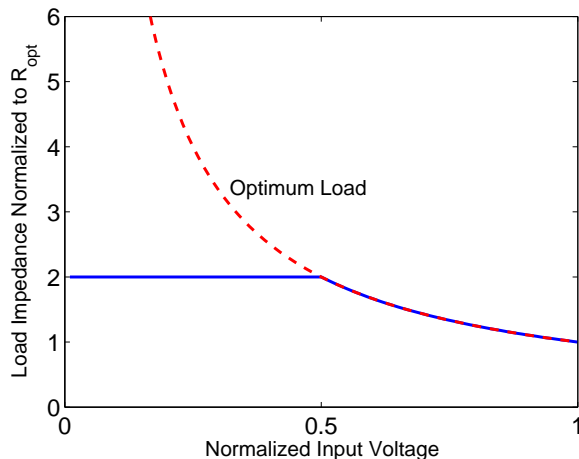
Figure 2.6: Optimal load impedance for ideal load modulation as well as the practical implementation

2.5 Doherty Amplifier Technique

2.5.1 Basic Principle

As mentioned before, the load impedance, R_L, in a class-B is typically selected to simultaneously achieve the peak output power and peak efficiency. Alternatively, R_L can be increased such that the peak efficiency occurs at a power backed off from the peak value. This would, however, lead to a decreased output power compared to the nominal peak value, since the output current swing would be less than the peak value. Any further increase in input voltage from this point on, would push the transistor into the knee region and cause nonlinearity problems. To maintain the linearity, one can gradually decrease R_L, inversely proportional to the input voltage (or drain current), so that the maximum swing is maintained without entering into the knee region. For example, Figure 2.6 shows the required impedance profile of the transistor in the most ideal case in order to achieve the maximum efficiency over the entire input voltage range. As can be seen, this impedance approaches infinity as the input drive becomes too small, which is not realizable. In practice, one may fix the impedance below a certain threshold voltage instead of further increasing it. For instance, Figure 2.6 illustrates a case where the load impedance below $v_{in} = 0.5$ is kept constant at $2R_{opt}$. Obviously, it will cause the efficiency to drop for the input drives below this threshold level. This is the core of the DPA idea [12].

An interesting approach to modulate the load impedance of a transistor is to use

14

(a) Initial realization of load modulation with two VCCSs. (b) modified circuit with the main VCCS replaced by a VCVS.

Figure 2.7: Conceptual representation of load modulation through Doherty technique.

an "auxiliary" or "peaking" transistor as illustrated in Figure 2.7(a). In this case, the impedance seen by the "main" VCCS can be written as,

$$Z_M = \frac{V_m}{I_m} = R_L \left(1 + \frac{I_p}{I_m} \right). \tag{2.3}$$

This equation shows that the load impedance of the main VCCS is controlled by the current supplied by the peaking VCCS. If the peaking VCCS is turned off, $I_p = 0$; then $Z_M = R_L$. As I_p increases in phase with I_m, Z_M also increases, according to (2.3). For instance, if $I_p = I_m$ the effective impedance seen by the main transistor will be twice the load impedance R_L. Nevertheless, this approach does not work properly for our purpose, since the impedance of the main transistor increases as a result of the current injected by the peaking transistor. One could combine I_p in antiphase with I_m to reduce Z_M as I_p grows, but that would lead to power and efficiency loss in the output combiner.

A modification to the previous circuit is depicted in Figure 2.7(b), in which the main amplifier is implemented as a voltage-controlled voltage source (VCVS). The advantage of this topology is that while the voltage across the load (or equivalently linearity) is enforced by the main amplifier, the current supplied by the peaking VCCS controls the admittance seen by main VCVS (or equivalently efficiency) as,

$$Y_M = \frac{I_m}{V_m} = \frac{1}{R_L} - \frac{I_p}{V_m} \tag{2.4}$$

Since all the transistors behave as a VCCS by nature, one can simply add an impedance inverter (such as a $\lambda/4$ transmission line) to the output of a transistor to realize the voltage source, V_m, as shown in Figure 2.8. This architecture is known as the DPA in its classical form [12]. The main device is usually biased in class B or deep AB, whereas the peaking one is biased in class C, as it must be operating only after the main device reaches its maximum voltage swing at the desired power back-off.

15
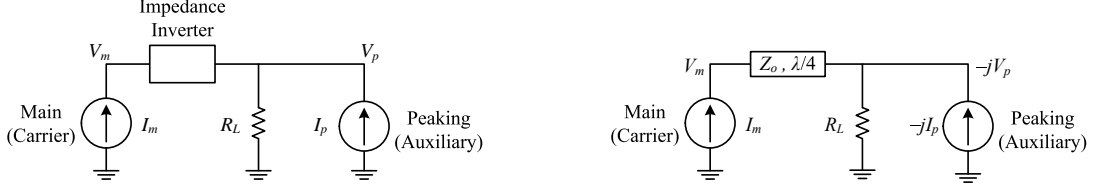
(a) Doherty amplifier representation with ideal impedance inverter.

(b) Doherty amplifier representation with $\lambda/4$ impedance inverter.

Figure 2.8: Final schematic representation of the DPA.

2.5.2 Circuit Analysis

For the simplified schematic shown in Figure 2.8, the governing equations of the DPA can be derived in terms of fundamental currents of the main and peaking devices, I_m and I_p, and the optimum impedance of the main device at peak power, R_{opt} [13]. This analysis helps to understand the efficiency enhancement of the DPA structure in a qualitative manner.

In general, the main and peaking currents are nonlinear functions of both input and output signal amplitudes. To simplify our theoretical analysis, however, it is assumed that the fundamental current profile for the main and peaking transistors are expressed by,

$$I_m(v_{in}) = \begin{cases} v_{in} \cdot I_M & 0 < v_{in} < 1 \\ I_M & v_{in} > 1 \\ 0 & \text{elsewhere} \end{cases} \tag{2.5}$$

$$I_p(v_{in}) = \begin{cases} 2(v_{in} - 0.5)I_P & k < v_{in} < 1 \\ I_P & v_{in} > 1 \\ 0 & \text{elsewhere} \end{cases} \tag{2.6}$$

where v_{in} is the normalized instantaneous input voltage, and $v_{in} = k$ is the breakpoint at which the peaking current starts to flow ($k = 1/2$ for the classical DPA). It is assumed that the currents are combined in phase at the output combiner, meaning that the voltage and current of the peaking device must lag by $90\,^\circ$. To analyze the circuit, the ABCD-parameters of the impedance inverter shall be used to relate the current and voltage components at its input and output terminals, i.e.,

$$\begin{pmatrix} V_2 \\ I_2 \end{pmatrix} = \begin{pmatrix} 0 & -jZ_o \\ -j/Z_o & 0 \end{pmatrix} \begin{pmatrix} V_1 \\ I_1 \end{pmatrix} \tag{2.7}$$

16

where ports 1 and 2 represent the output of main and peaking devices, respectively.

The following relations can help understand the circuit function:

$$Y_M = \frac{1}{R_L} - \frac{I_p}{V_p} \qquad\qquad\qquad \text{from (2.4)} \qquad\qquad\qquad (2.8)$$

$$V_p = Z_o I_m \qquad\qquad\qquad\qquad \text{from (2.7)} \qquad\qquad\qquad (2.9)$$

$$Z_M = \frac{V_m}{I_m} = Z_o^2 Y_m = Z_o \left(\frac{Z_o}{R_L} - \frac{I_p}{I_m} \right) \qquad \text{from (2.8) and (2.9).} \qquad (2.10)$$

Eq. (2.10) reveals that the impedance seen by the main transistor, Z_M, is "modulated" as I_p starts to increase. This provides the desired load modulation of the main transistor to prevent its saturation, yet maintain high efficiency at power back-off. The circuit parameters, Z_o and R_L, can be determined in terms of R_{opt}, which in turn depends on the peak output power, considering (2.9) and (2.10), (i) at low power levels ($v_{in} < \frac{1}{2}$), and (ii) at peak power ($v_{in} = 1$). In the low power region, $I_p = 0$; thus,

$$Z_M^L = Z_o^2 / R_L = 2R_{opt} \qquad \text{(low power)} \qquad\qquad (2.11)$$

where the "L" superscript denotes the quantity evaluated in the low power region. At the peak power, on the other hand, both transistors should see their optimum impedances, i.e., $Z_M^F = R_{opt}$, where "F" superscript represents the impedance of the main transistor at full power. Assuming that both the transistors share the same drain supply voltage, then $V_m^F = V_p^F = V_{dd}$ at the full power. Hence, one can deduce the circuit parameters, R_L and Z_o as

$$Z_o = \frac{V_{dd}}{I_M} = R_{opt} \qquad\qquad\qquad \text{from (2.9)} \qquad\qquad\qquad (2.12)$$

$$R_L = \frac{R_{opt}}{2} \qquad\qquad\qquad\qquad \text{from (2.11).} \qquad\qquad\qquad (2.13)$$

Moreover, evaluating (2.10) at peak power yields,

$$Z_o \left(\frac{Z_o}{R_L} - \frac{I_P}{I_M} \right) = R_{opt} \qquad\qquad\qquad\qquad\qquad (2.14)$$

which can be manipulated to give the ratio between the transistors' peak current, i.e.,

$$I_P = I_M. \qquad\qquad\qquad\qquad\qquad (2.15)$$

It can be shown that (2.12)–(2.15) are sufficient to fully describe the DPA's behavior. The fundamental current and voltage profiles of the classical DPA is depicted in Figure 2.9.
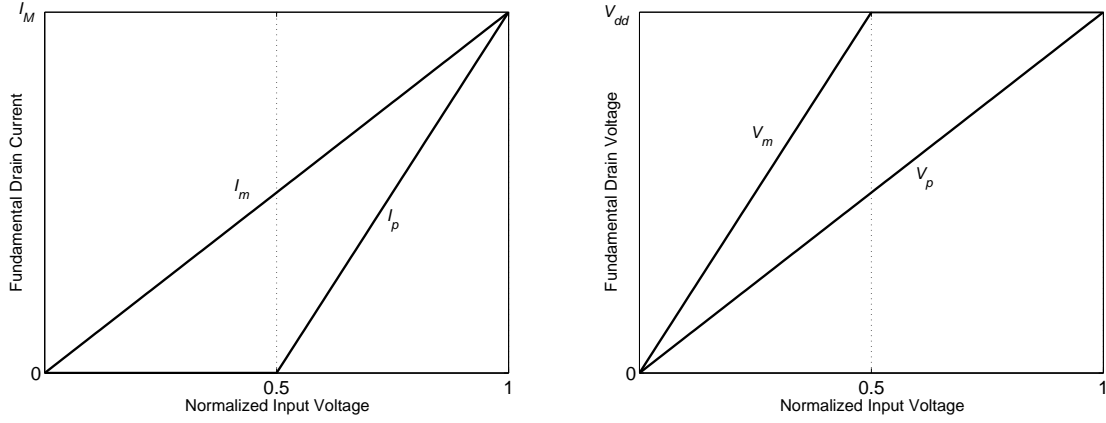
(a) Fundamental component of drain current vs. normalized input voltage.

(b) Fundamental component of drain voltage vs. normalized input voltage.

Figure 2.9: Fundamental current and voltage profiles of transistors in the classical DPA.

Based on the current profiles, one may notice an immediate challenge in realization of a DPA. Since the peaking transistor starts operating at half the maximum input voltage drive, it will not be able to provide the same peak current as the main device (I_M) if identical devices are used. One solution could be using a larger peaking device [ideally twice as large according to Figure 2.9(a)]. This would, however, lead to a decreased power utilization factor (PUF), since the power delivered by the peaking device is at best half its power capability. Alternatively, one could identical devices and unequally split the input power instead to compensate for the peaking gain shortage [14]. Ideally, the power must be divided with a ratio of 2:1 between the two cells in favor of the peaking amplifier for proper Doherty operation. This will however lead to an overall gain loss of at least $20\log(2/3) = 3.5$ dB which may not be acceptable for many cases. It should also be noted the synthesis problem of the peaking devices' current exacerbates when using a real transistor, due to the soft turn-on effect and inherently lower g_m of the class-C operation as compared to class B.

The current and voltage profiles presented can be used to calculate the impedance seen

(a) Main and peaking impedance functions vs. normalized input voltage.

(b) Main and peaking admittance functions vs. normalized input voltage.
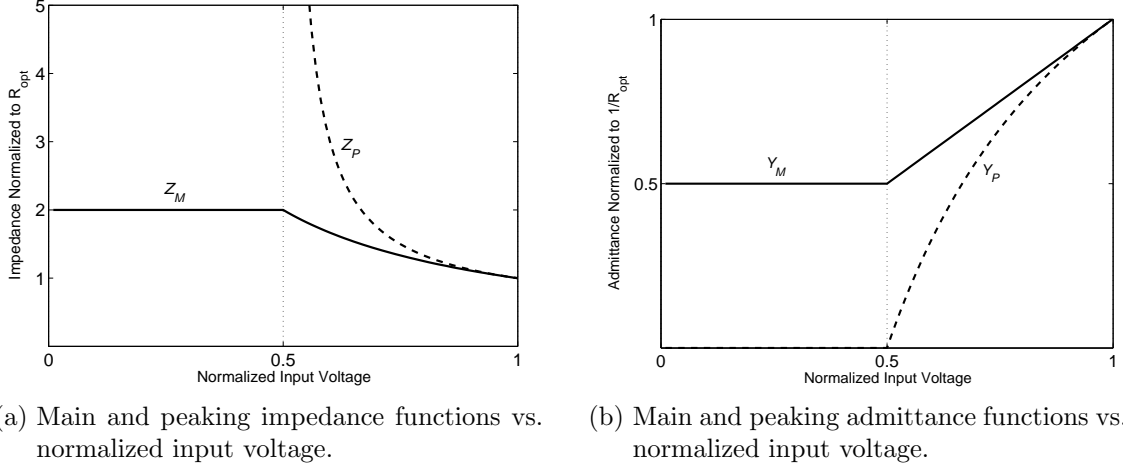
Figure 2.10: Impedance and admittance profiles of the main and peaking transistors in a DPA.

by each transistor as a function of v_{in}, i.e.,

$$Z_M(v_{in}) = \begin{cases} 2R_{opt} & 0 < v_{in} < 0.5 \\ R_{opt}/v_{in} & 0.5 < v_{in} < 1 \end{cases} \tag{2.16}$$

$$Z_P(v_{in}) = \begin{cases} \infty & 0 < v_{in} < 0.5 \\ \frac{1}{2} \cdot \frac{R_{opt}}{v_{in}-0.5} & 0.5 < v_{in} < 1 \end{cases}. \tag{2.17}$$

which are plotted in Figure 2.10, along with the admittance functions. It can be seen that the main impedance exactly follows the required profile shown in Figure 2.6 after the breakpoint. This confirms while the the efficiency of the main stage is maintained after the breakpoint, linearity is also preserved over the entire input drive range. Note that the admittance functions can be interpreted in the same manner and moreover, they are easier to verify in practice, thus the admittance will only be plotted for the next Doherty architectures discussed in this chapter.

The output power delivered by the main and peaking amplifiers is given by,

$$P_{out,M} = \frac{1}{2}\Re\{\mathbf{V_m} \cdot \mathbf{I_m}^*\} \tag{2.18a}$$

$$P_{out,P} = \frac{1}{2}\Re\{\mathbf{V_p} \cdot \mathbf{I_p}^*\}. \tag{2.18b}$$
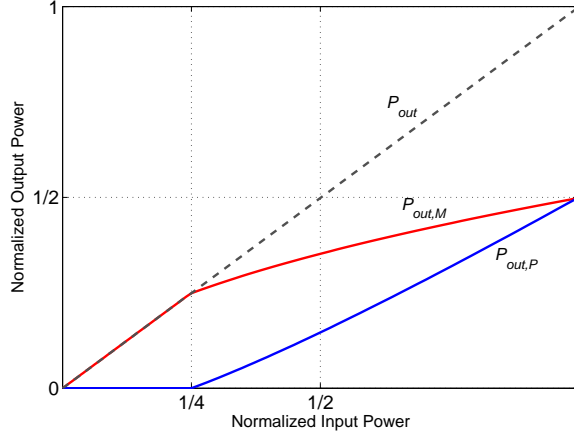
Figure 2.11: Output power profiles of each cell as well as the total power of the DPA.

The total output power can then be calculated by,

$$\begin{aligned} P_{out} &= P_{out,M} + P_{out,P} \\ &= V_L^2/R_L. \end{aligned} \tag{2.19}$$

where $V_L = V_p$ is the fundamental voltage across the load, R_L. The plot of $P_{out,M}$, $P_{out,P}$, and P_{out} is shown in Figure 2.11. It would be interesting to consider the overall output power at the two important levels, (a) the amplitude of the input voltage is equal to the breakpoint voltage ($|v_{in}| = 0.5$,) and (b) the maximum input voltage ($|v_{in}| = 1$). At power level in (a), the peaking transistor is still off and so it provides no power, and the power delivered by the main amplifier is half of its peak value, which is

$$P_{out} = P_{out,M} = \frac{1}{2}V_{dd} \cdot \frac{I_M}{2} \qquad \text{(at breakpoint)}. \tag{2.20}$$

At power level in (b), the main and peaking devices contribute equally to the output power, since,

$$\begin{aligned} P_{out} &= \frac{1}{2}(V_{dd} \cdot I_M + V_{dd} \cdot I_P) \\ &= V_{dd} \cdot I_M \qquad \text{(at peak power)}. \end{aligned} \tag{2.21}$$

It is observed that the output power at the breakpoint is one-fourth (or 6 dB lower than) the peak power. This is expected since the input voltage amplitude is one-half of the peak value and the amplifier is linear (see Figure 2.11), which means that the input power

back-off is equal to the output power back-off. In fact, even though the main and peaking output powers are nonlinear functions of the input drive, their combination is a linear one. This is because the voltage across the load is dictated by the main current, I_m according to (2.9). In other words, linearity of the DPA depends entirely on the main device, provided that the main transistor does not enter into the knee region by the function of the peaking device.

Assuming ideal class-B waveforms, the DC power consumed by the main and peaking transistor is given by,

$$P_{dc,M} = V_{dd} \cdot I_{dc,M}$$

$$= V_{dd} \cdot \frac{2I_m}{\pi} \tag{2.22a}$$

$$P_{dc,P} = V_{dd} \cdot I_{dc,P}$$

$$= V_{dd} \cdot \frac{2I_p}{\pi}. \tag{2.22b}$$

The overall power dissipated by the DPA is obviously equal to $P_{dc,M} + P_{dc,P}$. Hence, one can readily calculate the drain efficiency of the main and peaking devices as well as the overall efficiency of the DPA using (2.18), (2.19), and (2.22) as,

$$\eta_M = \frac{Pout_M}{P_{dc,M}} \times 100\% \tag{2.23a}$$

$$\eta_P = \frac{Pout_P}{P_{dc,P}} \times 100\% \tag{2.23b}$$

$$\eta_{DPA} = \frac{Pout_M + Pout_P}{P_{dc,M} + P_{dc,P}} \times 100\%. \tag{2.23c}$$

The resulting efficiency plots are illustrated in Figure 2.12, as a function of the input voltage amplitude and output power back-off. The main efficiency is exactly as was aimed for, i.e., it peaks at the breakpoint (6-dB back-off) and is maintained thereafter. The overall drain efficiency, however, slightly decreases past this point, due to the low efficiency of the peaking transistor.

It is helpful, at this point, to assess the power gain of a DPA and compare it with that of a class-B PA. Since the DPA is ideally a linear amplifier, the gain is calculated only in the low power region, for the sake of simplicity. It is also assumed for both PAs that the input side is conjugately matched to the signal source impedance, i.e., $R_{in} = Z_S^*$. Assuming equal power is fed into the both Doherty and class-B amplifiers for a fair comparison, the

(a) Efficiency vs. normalized input voltage amplitude.

(b) Efficiency vs. output power back-off.
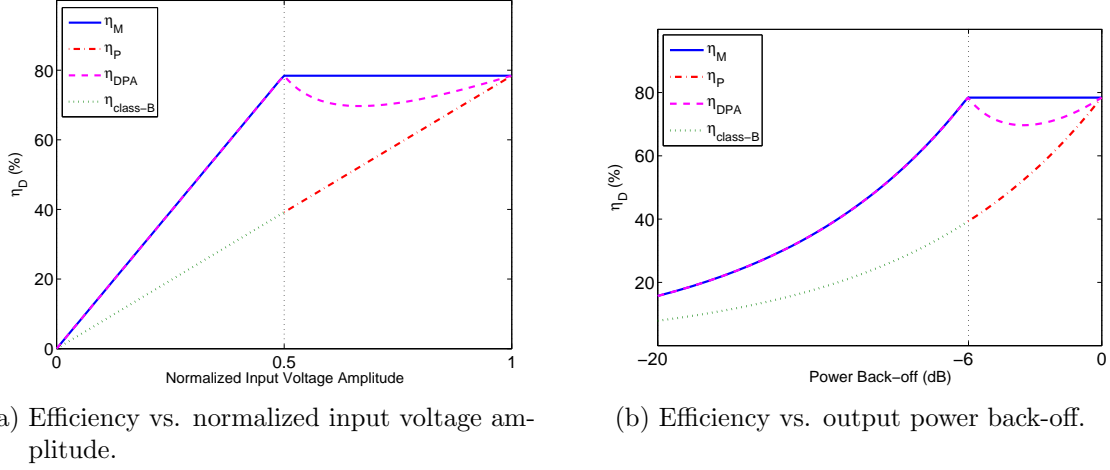
Figure 2.12: The efficiency profile of the classical DPA.

ratio of the gains will be equal to the ratio of the output powers, which is expressed as

$$\frac{G_p^D}{G_p^B} = \frac{P_{out}^D}{P_{out}^B} = \frac{\frac{1}{2}(G_m V_{im})^2 \cdot 2R_{opt}}{\frac{1}{2}(G_m V_S)^2 \cdot R_{opt}}. \tag{2.24}$$

where V_{im} and V_S denote the voltage swings across the gate of the transistor in the Doherty and main topology, respectively, and G_p. Note that $V_{im} = V_S/\sqrt{2}$ assuming equal input power division for the DPA. Hence, it can be inferred that

$$\boxed{G_p^D = G_p^B}. \tag{2.25}$$

Therefore, employing the Doherty technique to enhance the back-off efficiency does not affect the gain of the main (class B) amplifier.

2.5.3 Practical Design Challenges

Since there is no ideal current source in real world, any nonideality associated with the active device operation has to be somehow taken into account. In this section, the most prominent sources of nonideality, namely the soft-turn on effect, the devices' output parasitics, and harmonic terminations will be discussed in brief. More detailed discussion about transistor nonidealities which affect the linearity/efficency will be presented in Chapter 4.

Soft Turn-on Effect

As mentioned before, the current, voltage and efficiency plots presented in Section 2.5.2 are based on the assumption that the transistors behave as ideal current source, whose fundamental current component follows a class-B curve. In reality, however, the peaking amplifier is usually implemented using a class-C biased transistor, thus suffers from soft turn-on effect, and nonlinear gain characteristics, since the conduction angle of a class-C transistor is a function of the drive voltage. Furthermore, the transconductance of a transistor when biased in class C is lower than that of the one biased in class B which necessitates a peaking device even more than twice as large as the main one, formerly inferred from the theory.

Figure 2.13 shows the fundamental current, voltage, admittance, and efficiency profiles of the two transistors when the peaking device is biased in class C, compared against those of the ideal case (assuming a class B profile). As can be seen, due to the soft turn-on effect, there are multiple choices for the gate bias voltage of the peaking transistor:

(i) If $V_{Gp} = -0.5$, then the I_p profile is poorly approximated, leading to linearity degradation over the upper 6-dB back-off (BO) range as the main device enters the knee region. Nevertheless, perfect efficiency, ideally 78.5% at 6-dB BO and 83.5% at peak power is achieved. Moreover, peaking device periphery must be 2.55 times larger than the main one.

(ii) $V_{Gp} \lesssim -0.5$, then a more accurate approximation of I_p profile is obtained (as the voltage drive exceeds ~ 0.6), hence satisfactory linearity results. The drain efficiency, however, reduces by $\sim 5\%$ at 6-dB BO, due to the added power consumption of the peaking device. The size of the peaking device in this case is set 2.2 times greater than the main one.

where V_{Gp} is the DC gate voltage of the peaking transistor normalized to the maximum input voltage drive. This argument demonstrates that there is a trade-off between efficiency and linearity of a DPA by selecting the gate bias voltage of the peaking transistor, although linearity is usually the priority in any analog amplifier design (including PAs). Note that all the values mentioned here are based on a linear transistor model (constant g_m in the saturation region) and will obviously change when using a real transistor.

Device Output Parasitics

Figure 2.14(a) shows the DPA when current sources are replaced by transistors. As mentioned before, a phase compensation line has to be used (at the input of the peaking
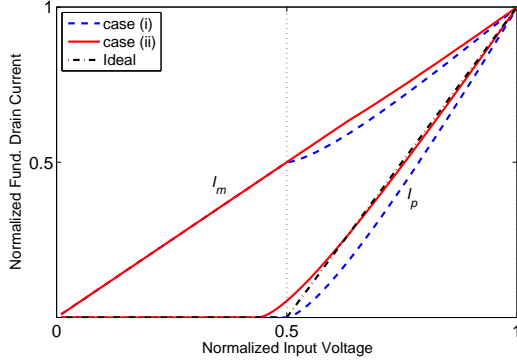
(a) Current profile of a DPA considering the turn-off effect.
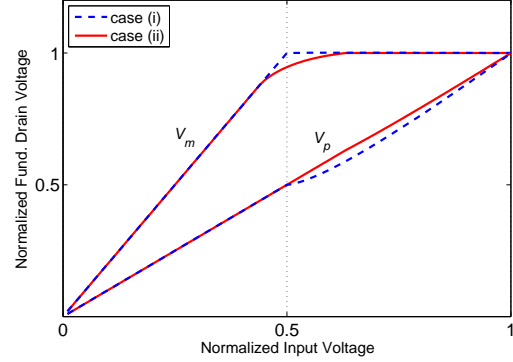
(b) Voltage profile of a DPA considering the turn-off effect.
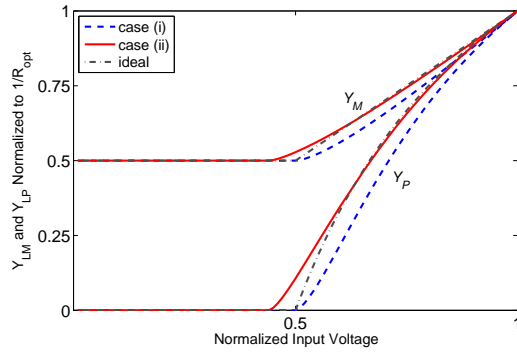
(c) Admittance profile of a DPA considering the turn-off effect.
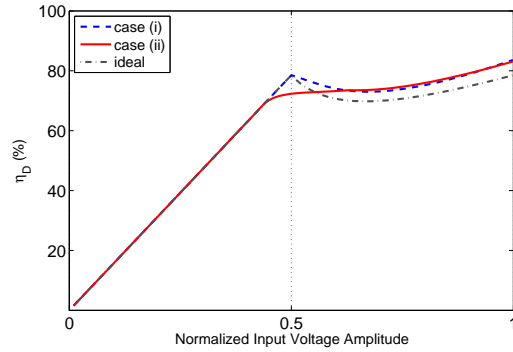
(d) Efficiency profile of a DPA considering the turn-off effect.

Figure 2.13: DPA characteristic profiles considering the soft turn-on effect of the peaking device for two class-C biasing cases.

transistor in this case) to ensure an in-phase signal combination at the output. In reality, a transistor is not just a VCCS; on the contrary, there are lots of intrinsic and extrinsic elements inside a packaged transistor, known as parasitics, that have to be taken into account in the design. These parasitics can be quite significant in the case of high power transistors embedded in large packages and/or at high operating frequencies, and may considerably change the magnitude and phase of the output impedance of the main and peaking devices. As a result, to insure a proper load modulation across the bandwidth, the combination of transistor parasitics and matching networks must behave as an impedance inverter. In terms of phase shift, for instance, the main and peaking VCCS's must stand $\pm 90°$ and

0° or 180°

Main CS

IMN + Intrinsic + ⋯

Intrinsics + Extrinsics + Pre-match + Package + ⋯

OMN

Power Divider

λ/4

IMN + Intrinsic + ⋯

Peaking CS

Intrinsics + Extrinsics + Pre-match + Package + ⋯

OMN

0° or 180°

A

λ/4

Output Matching

50 Ω

(a)

180° ± 90°

θ_M

Main CS

IMN + Intrinsic + ⋯

Intrinsics + Extrinsics + ⋯

OMN

Power Divider

Phase Comp.

θ_P

IMN + Intrinsic + ⋯

Peaking CS

Intrinsics + Extrinsics + ⋯

OMN

0° or 180°

Output Matching

50 Ω

(b)

Figure 2.14: Generic DPA block diagram with actual transistors.

0/180° away from the output tee-junction, respectively, as illustrated in Figure 2.14(b).

Another challenge arises when the output capacitance of the transistor is nonlinear. Since the output impedance of the transistor will be a function of the drive level, the load modulation may be greatly affected at high power level. If the output capacitance is compensated at high power levels, the impedance transformation and thus the efficiency enhancement will degrade in low power region. The problem worsens if the capacitance value is large, which is the case for LDMOS transistor. This will further complicate the compensation process over the bandwidth.

Harmonic Terminations

The theoretical analysis assume all of the harmonics to be ideally short circuited at the intrinsic drain of both transistors. This may not be always satisfied when packaged devices are to be used or a wideband performance is targeted. Most often, a short-circuit shunt stub, either separately or as the biasing feed, is used as a harmonic termination for power amplifier applications. While short circuit at the second harmonic and all other even harmonics, the stub is open circuit at the fundamental and odd harmonics. The impedance at the third harmonic must hence be controlled by the output matching network. Irrespective of the transistor output parasitics, the second and third harmonic impedances as well as the fundamental one may be difficult to control simultaneously over the entire bandwidth. This particularly occurs when dealing with high power transistors having a low R_{opt}. In fact, the magnitude of the second harmonic impedance must be low compared to the fundamental impedance, $2R_{opt}$ and R_{opt} in the low and peak power levels, respectively, to realize the required sinusoidal drain voltage waveform.

2.5.4 DPA for extended PAPR

The DPA theory presented in Section 2.5.1 can be extended for an arbitrary breakpoint voltage, k, equivalent to a back-off region different from 6 dB [15]. This is especially of interest for emerging wireless transmitters with increased PAPR signals (larger than 6 dB). Lowering the breakpoint ($k < 0.5$) enables a DPA to amplify the modulated signal more efficiently. Note that this does not impact the main current, but only changes the required current profile of the peaking device and thus the total output power. Hence, this architecture is sometimes referred to as the asymmetrical DPA. The general expressions for the peaking current in this case is:

$$I_p(v_{in}) = \begin{cases} \frac{v_{in}-k}{1-k}I_P & k < v_{in} < 1 \\ I_P & v_{in} > 1 \\ 0 & \text{elsewhere} \end{cases} \quad (2.26)$$

Similar to the procedure presented in Section 2.5.1, we have

$$Z_M^L = \frac{V_{dd}}{k \cdot I_M} = R_{opt}/k \qquad \text{similar to } (2.11) \qquad (2.27)$$

$$Z_o = R_{opt} \qquad \text{similar to } (2.12) \qquad (2.28)$$

$$R_L = k \cdot R_{opt} \qquad \text{similar to } (2.13). \qquad (2.29)$$

To deduce the ratio of the transistors' peak currents, (2.10) can be evaluated at $v_{in} = 1$ to give

$$\frac{I_P}{I_M} = \frac{1}{k} - 1. \tag{2.30}$$

It can be inferred that the peaking current at full power must be increased with respect to the main current as k decreases; hence the resulting amplifier is referred to as the asymmetrical DPA. Figure 2.15(a) and (b) show the fundamental voltage and current profiles for the special case of $k = 1/3$. This also means that the peaking amplifier contributes more than the main one to the overall output power. The profile of output power of the main and peaking amplifier as well as the total output power versus input power is illustrated in Figure 2.15(c) for $k = 1/3$. It can be readily inferred that the efficiency peak in this case is equal to $20 \log(1/k) = 9.5$ dB.

The admittance seen by the main and peaking transistors are also expressed as

$$Y_M(v_{in}) = \begin{cases} \frac{k}{R_{opt}} & 0 < v_{in} < k \\ \frac{v_{in}}{R_{opt}} & k < v_{in} < 1 \end{cases} \tag{2.31}$$

$$Y_P(v_{in}) = \begin{cases} 0 & 0 < v_{in} < k \\ \frac{k}{R_{opt}}(v_{in} - k) & k < v_{in} < 1 \end{cases}. \tag{2.32}$$

The admittance functions defined in (2.31) and (2.32) for $k = 1/3$ are plotted in Figure 2.15(e). It can be seen that unlike the classical case, the optimum impedance of the two transistors at peak power are no longer equal. A major disadvantage of this technique is the deep decrease of the efficiency between the two efficiency peaks, as shown in Figure 2.15(d), which leads to a considerable drop of the average efficiency for modulated signal excitation.

2.5.5 DPA with Asymmetrical Drain Voltages

A variant of the 2W DPA is using different bias voltages for the main and peaking transistors. This architecture has been proposed in the literature for a special case [16, 17]. A generalization of the idea is presented in this section which defines a large design space that provides more design flexibility. Recalling from Section 2.5.2, the load modulation of the main impedance is governed by four key parameters, I_m, I_p, Z_o, and R_L, and is described
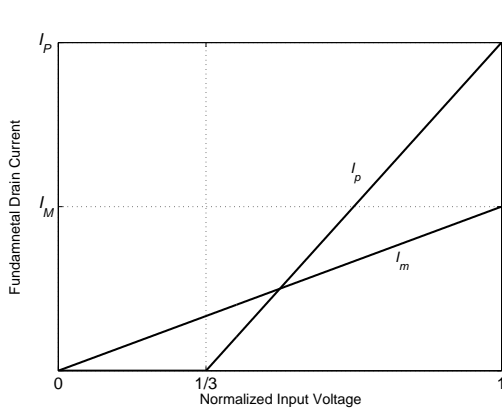
(a) Current profile of an asymmetrical DPA when $k = 1/3$.
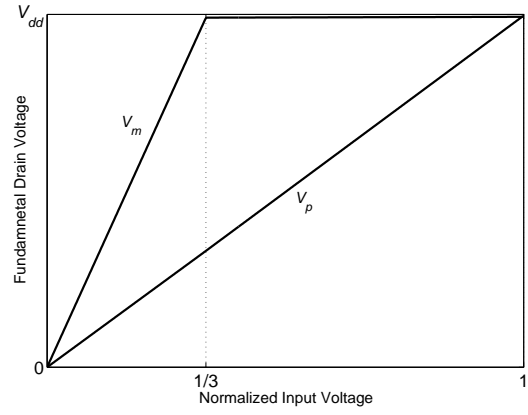
(b) Voltage profile of an asymmetrical DPA when $k = 1/3$.
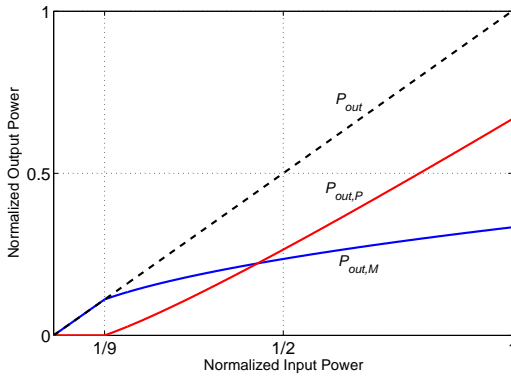
(c) Output power profile of an asymmetrical DPA when $k = 1/3$.
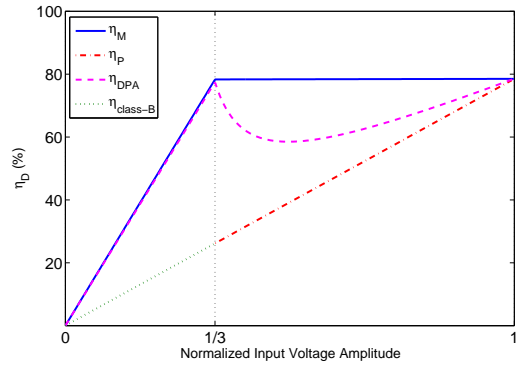
(d) Efficiency profile of an asymmetrical DPA when $k = 1/3$.
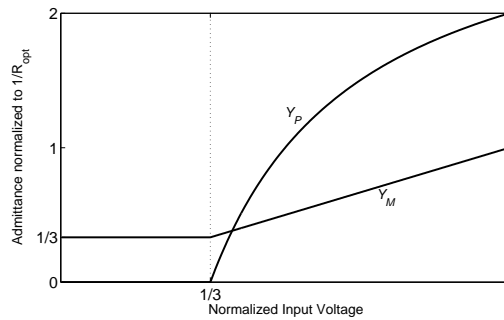
(e) Admittance profiles in an asymmetrical DPA with $k = 1/3$.

Figure 2.15: Characteristic profiles of an asymmetrical DPA.

by

$$Z_M = Z_o \left(\frac{Z_o}{R_L} - \frac{I_p}{I_m} \right). \tag{2.33}$$

It is intuitively expected that by choosing one of these parameters (e.g. the current profiles) the value of remaining parameters (Z_T and R_L) are dictated such that the load modulation conditions are satisfied. Consider a general case in which at peak power $|I_M/I_P| = \sigma$ and $|V_P/V_M| = \rho$ where $V_M \approx V_{dd}$. Thus, (2.33) can be evaluated in the low- and high-power regions for 6-dB BO efficiency peak as

$$Z_M^L = Z_o^2/R_L = 2R_{opt} \tag{2.34}$$

$$Z_M^F = Z_o^2/R_L - \frac{Z_o}{\sigma} = R_{opt}. \tag{2.35}$$

Equations (2.34) and (2.35) can be simultaneously solved for Z_o and R_L to give

$$Z_o = \sigma R_{opt} \tag{2.36}$$

$$R_L = \frac{\sigma^2 R_{opt}}{2}. \tag{2.37}$$

Furthermore, one can find the necessary relation between ρ and σ as

$$V_P = Z_o I_M$$
$$\Rightarrow \rho = \sigma. \tag{2.38}$$

In fact, σ or ρ defines a design space with different current and voltage profiles but identical performance (in terms of power, efficiency, etc.) at the design frequency. In other words, σ or ρ can be considered as a degree of freedom for DPA design depending on the required current/voltage profile requirement or frequency response. Note that if $\sigma = \rho = 1$, then $I_M = I_P$ and $V_M = V_P = V_{dd}$ which is the classical symmetrically-biased case.

As mentioned in Section 2.5.2, one major challenge in the DPA realization is the need for a peaking device at least twice larger than the main one for proper load modulation, which can greatly increase the cost since the peaking transistor is under-utilized. Asymmetrical drain biasing technique can be employed to resolve this limitation. By selecting $\sigma = \rho = 2$, for instance, the fundamental current of the peaking device is half that of the main one; thus the DPA can be realized using identical devices. Consequently, the peaking device is fully utilized and a 100% PUF is achieved. This obviously comes at the expense of an increase in the required drain-source breakdown voltage of the peaking transistor, in addition to the added complexity since two separate voltage supplies are required. The
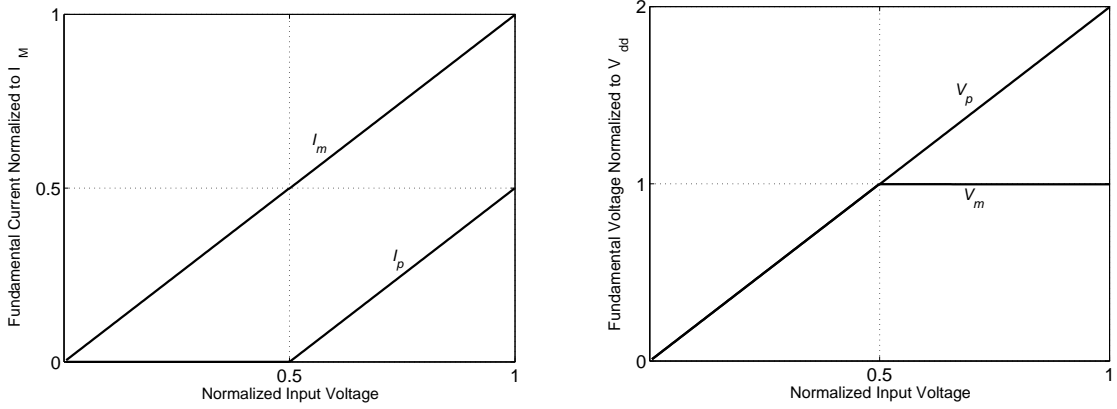
(a) Current profile of a DPA with asymmetrical drain biasing when $\sigma = \rho = 2$.

(b) Voltage profile of a DPA with asymmetrical drain biasing when $\sigma = \rho = 2$.

Figure 2.16: Current and voltage profiles of an asymmetrically-biased DPA.

technique is thus mainly useful for high breakdown voltage device technologies, such as gallium-nitride (GaN). The main transistor has to be down-biased when the technique is applied to LDMOS transistors, which in turn decreases the PUF. The current and voltage profiles for this case ($\sigma = \rho = 2$) are depicted in Figure 2.16.

It should be noted that the asymmetrical drain biasing idea can be extended for a back-off efficiency peak higher than 6 dB. Following a similar analysis conducted in Section 2.5.4, one may conclude that

$$Z_o = \sigma R_{opt} \tag{2.39}$$

$$R_L = k\sigma^2 R_{opt} \tag{2.40}$$

$$\rho = \sigma. \tag{2.41}$$

As an example, to design a DPA with $\gamma = 3$ (corresponding to 9.5-dB BO efficiency peak) using identical devices for carrier and peaking amplifiers, $\sigma = 3$. Hence, $Z_o = R_L = 3R_{opt}$, and $\rho = 3$. This, however, may not be a viable solution, because as the BO efficiency enhancement range enlarges, a higher supply voltage is required for the peaking transistor, which necessitates a higher breakdown voltage. In addition, asymmetrically biasing the two transistors will introduce additional memory effects when the DPA is driven by wideband/multi-band modulated signals, as will be discussed in Chapter 4.

2.5.6 Bandwidth Analysis of Two-way DPAs

Thus far, the operation of the 3W-DPA has been analyzed at the center frequency. As the carrier frequency shifts away from the design frequency, both the efficiency and linearity are expected to degrade. The sources of bandwidth limitation in a DPA can be categorized into two general groups:

- Choice of Z_o and R_L and the corresponding current profiles, I_m and I_p which is inherent to the given topology.

- The DPA realization technique which involves frequency dependence of the phase compensation and offset lines, quarter-wave transformer, matching networks, pre-matching circuits and device parasitics.

In the literature, the quarter-wave impedance inverter is mostly blamed for the narrow bandwidth of the DPA. Nevertheless, as will be shown later, the frequency dependence of the $\lambda/4$ inverter only manifests itself at large bandwidths well beyond those reported in the literature.

It should be noted, the circuit parameters, not only the impedance inverter, are generally tuned to satisfy the load modulation conditions at a given frequency. These conditions are rapidly violated as the frequency varies. For instance, neither the main nor the peaking amplifier acts as a VCCS, due to the transistor parasitics, as well as the pre-matching and matching networks. So the peaking amplifier will not present a high impedance at low power levels and it cannot properly modulate the impedance of the main amplifier at high power levels either.

To analyze the bandwidth of the DPA, we start by using ideal current sources instead of actual transistors. This is equivalent to exploring the first source of bandwidth limitation mentioned above. It is also assumed that the load R_L is frequency independent. Figure 2.17 illustrates the drain efficiency and output power of a classical DPA versus output power back-off and frequency. As can be seen, the back-off power and efficiency degrade with frequency. For instance, drain efficiency drops by 15% at 40% bandwidth. It is however observed that peak power and efficiency remain constant versus frequency. This is attributed to the fact that the characteristic impedance of the impedance inverter is equal to its termination impedance at peak power, i.e. R_{opt}. Figure 2.18 shows the efficiency profiles for an asymmetrical DPA designed for 9.5-dB efficiency peak. It can be observed that if higher back-off efficiency enhancement is intended, the bandwidth is further confined. This phenomenon can be explained by considering the impedance transformation ratio (ITR) of the $\lambda/4$ impedance inverter. In general, the impedance inverter

31
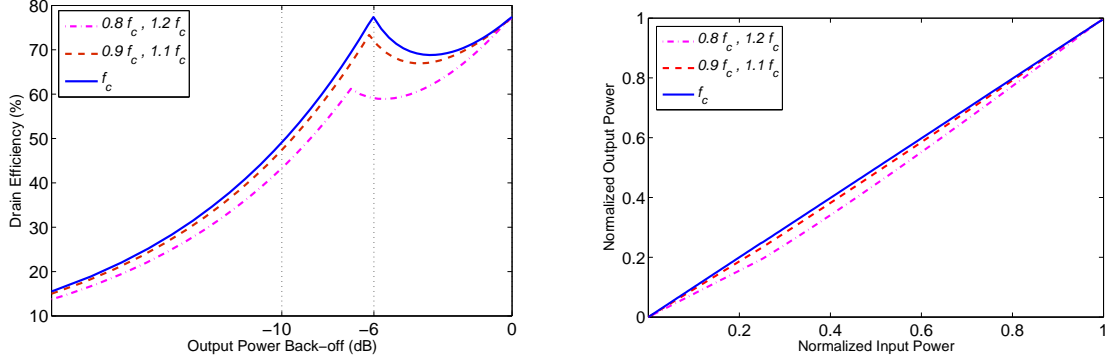
(a) Drain efficiency profile of the classical DPA vs. frequency.

(b) Output power profile of the classical DPA vs. frequency.

Figure 2.17: Frequency characteristics of the classical DPA.
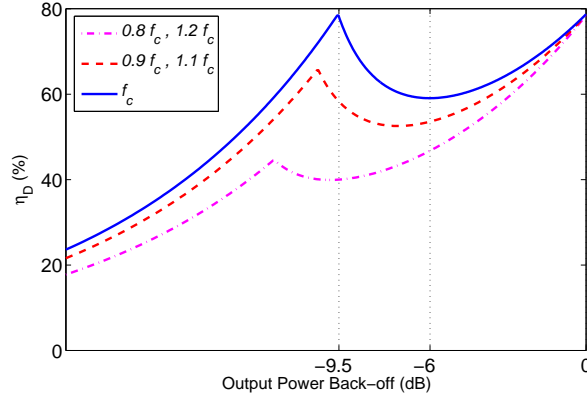
Figure 2.18: Drain efficiency of an asymmetric DPA vs. bandwidth.

transforms an impedance of $R_L = kR_{opt}$ to $Z_M = R_{opt}/k$. Evidently, for higher k values, the Q-factor of the inverter increases, hence the efficiency-bandwidth decreases in back-off. This is however not the case at the peak power, as the transmission line is still terminated with its characteristic impedance, and thus frequency independent. As discussed before, in modern transmitters, the PA is driven by modulated signals with high PAPR, and thus the average efficiency is dominated by the efficiency in back-off power levels rather than the peak power.

One idea to improve the back-off efficiency vs. bandwidth is to make the quarter-wave line terminate with its characteristic impedance at 6-dB BO rather than the peak power.

This can be achieved by using the asymmetric drain biasing technique [16]. For instance, if $\sigma = 2$, then $Z_o = R_L = 2R_{opt}$. As a result, the back-off efficiency peak will not vary with frequency which is a major advantage over the classical DPA, in addition to the fact that optimal load modulation can be achieved with identical devices, as mentioned before. The efficiency plots for this case is depicted in Figure 2.19. The second efficiency peak occurs at different BO levels for different frequencies, since the peak output power decreases with frequency. Efficiency profiles for various σ values are also plotted in Figure 2.19. It is observed that the efficiency peak both at back-off and full power degrades versus frequency; however, frequency dependence of efficiency is also reduced on average. This offers a better compromise between efficiency versus bandwidth and drain supply voltage.

2.5.7 DPAs in Literature

The recent research initiated in the area of Doherty amplifier has aimed at developing techniques to amplify: 1) signals with low to moderate PAPR (up to 7 dB) scattered over broad range of carrier frequencies [16,18–23] and 2) single-band signals with high PAPR (up to 12 dB) [15,24–27], Authors in [16,18–21] conducted comprehensive studies to identify the theoretical and practical sources of bandwidth limitations, commonly observed in different implementations of the classical two-way Doherty power amplifier (2W-DPA), as briefly discussed in Section 2.5.6. These studies triggered attempts to devise solutions to extend the bandwidth of 2W-DPAs. For example, authors in [18,20] transformed the quarter-wave impedance inverter into a quasi- or complete- lumped equivalent circuit. This allowed absorption of the two transistors' output capacitance and the bond-wires' inductance and consequently alleviated the bandwidth limitation. This technique has been applied to both base-station and mobile implementations. In spite of the excellent back-off efficiency obtained in [18] over the frequency range of 1.7–2.3 GHz (30% fractional bandwidth), this solution required a mixed-signal setup to ensure proper operation.

In [16], authors suggested a novel topology that mitigated the bandwidth limitations of a classical 2W-DPA. The novel approach required two different drain supply voltages for the main and peaking transistors (asymmetrical drain biasing technique presented in Section 2.5.5). A minimum 6-dB back-off efficiency of 52% was maintained over a fractional bandwidth of 35%. This significant bandwidth of 2W-DPAs was achieved at the expense of an increase in the required breakdown voltage of the peaking transistor. More recently, modified output combiners were proposed in [28–30] to improve the FBW of a DPA. For instance, in [28], a $\lambda/2$ line was added to the output of the peaking transistor which compensates for the frequency dispersion of the quarter-wave impedance inverter, and consequently, 52% FBW was reported with average drain efficiency in excess of 38% at
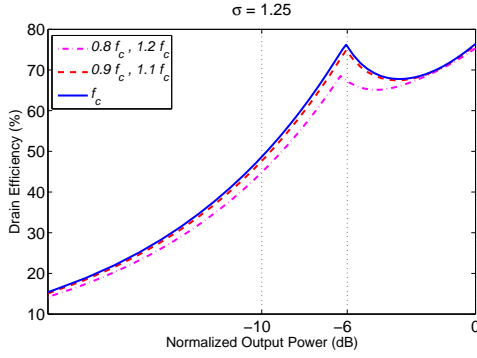
(a) Drain efficiency vs. bandwidth profile for $\sigma = 1.25$.
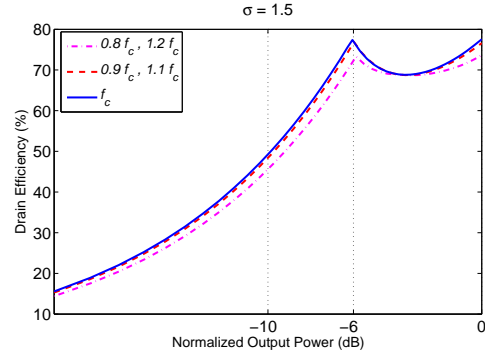
(b) Drain efficiency vs. bandwidth profile for $\sigma = 1.5$.
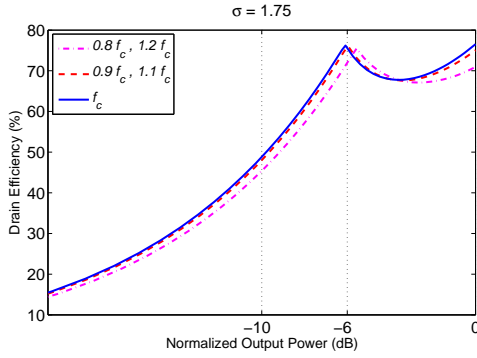
(c) Drain efficiency vs. bandwidth profile for $\sigma = 1.75$.
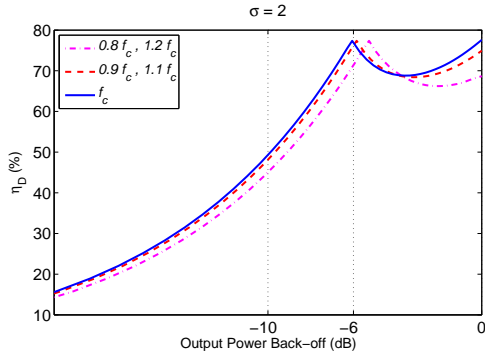
(d) Drain efficiency vs. bandwidth profile for $\sigma = 2$.
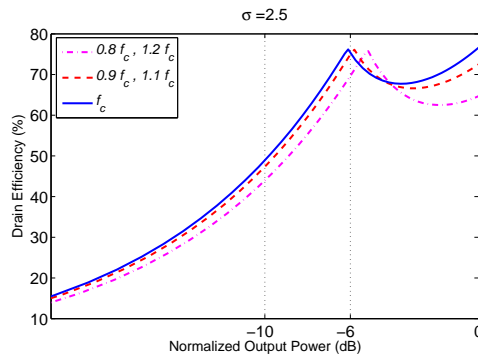
(e) Drain efficiency vs. bandwidth profile for $\sigma = 2.5$.

Figure 2.19: Efficiency-bandwidth characteristics for different σ values

115-W average output power for DVB-T broadcast applications. Although the above references were successful in extending the FBW of a DPA, they did not address the case where the DPA is driven with wideband/multiband CA signals; hence, the modulation bandwidth of the applied single-band signals were limited to 20 MHz or less. Additionally, a few attempts have been targeting the development of multi-band Doherty amplifier, by replacing the individual Doherty components (input/output matching networks, $\lambda/4$ line, etc.) with their dual-band equivalent components [31–33]. This approach is viable only for a limited range of f_2/f_1 and results in infeasible line widths otherwise. Furthermore, the efficiency enhancement at the upper band is barely sensible due to the significant impacts of device parasitics.

Various approaches have also been reported to improve the average efficiency of single-band DPAs when driven with high PAPR signals [15, 24–27]. The theory of the classical 2W-DPA has been generalized in [15] and [24] to yield what is called an asymmetrical 2W-DPA. This latter required larger peaking transistors to extend the high-efficiency power range. Nevertheless, the bandwidth of the asymmetrical 2W-DPA was further confined by the increased impedance transformation ratio of the impedance inverter (as shown in Section 2.5.6). Moreover, the noticeably increased peaking transistor size caused further complexity and narrower bandwidth for the input/output matching network.

Alternatively, multi-way architecture was reported in [27] and [25] to improve the average efficiency of a DPA under single-band and high PAPR signals. A three-way Doherty power amplifier (3W-DPA) exhibits multiple efficiency peaks at different back-off levels, and thus maintains a higher efficiency as compared to its two-way counterparts throughout the back-off region. However, in its original topology, also referred to as the conventional topology, the efficiency improvement is only achieved under proper control of the main transistor drain current to maintain proper load modulation in the high power range [25]. The additional need for controlling the main transistor hinders the analog realization of this type of 3W-DPA and only a mixed-signal version has been reported to date. To address the complexity of a classical 3W-DPA, a novel output combining network has been proposed in [34], known as modified topology. A gate-adaptation technique was added to this architecture to ensure the proper load modulation versus power [26]. The gate-adaptation feature imposed the addition of two envelope tracking modules that complicated the design.

An alternate approach has been suggested in [35, 36] for both base-station and handset applications by combining ET and Doherty technique to further improve the efficiency at high back-off levels. In this technique the drain supply voltage of the main device is modulated based on the envelope signal before the breakpoint, when the peaking transistor is turned on. Beyond the breakpoint, the drain supply voltage remains constant and efficiency enhancement is obtained through load modulation. Even though this approach

shows decent back-off efficiency enhancement, it has a narrow bandwidth and thus is not suitable for multi-standard or multi-band transmission.

Recently, reconfigurable Doherty topologies have been explored for multi-band/extended back-off efficiency enhancement. In [37], electronically tunable devices (MEMS switches) were used to achieve efficiency enhancement at multiple radio frequencies and back-off levels. However, due to the low power handling of tunable devices the application of these DPAs was limited to low power levels. The authors of [17,38] introduced an elegant approach to maximize the average efficiency for high PAPR signals over a broad range of frequencies by reconfiguring the drain supply voltage of the main transistor according to the input signal PAPR. However, the approach suffers from a low PUF when reconfigured for high PAPR values. [38] proposed a mixed-technology DPA in which the main and peaking cells are realized with LDMOS and GaN devices, respectively, thereby improving the PUF of the main device. It should be noted that none of the reconfigurable Doherty topologies are suitable for concurrent transmission of multi-band signals, which highlights the need for realizing a fully analog broadband/multiband DPAs for true multi-band, multi-standard transmission.

Table 2.1 summarizes state-of-the-art publications on Doherty amplifier technique. The bandwidth of 2W-DPAs (with 6-dB back-off range) has been successfully extended to ∼30%, while 3W-DPAs were only implemented for narrowband scenarios using mixed signal setup. The goal of this thesis is thus to enhance the bandwidth of 3W-DPAs to ∼30% and also develop a strategy to extend the bandwidth of 2W-DPAs to 50% or more, without compromising the linearity and linearizability of the DPA as required for concurrent amplification of multi-band signals.

Table 2.1: Doherty power amplifier benchmarks

Year	Technique	Frequency (GHz)	Peak P_{out} (dBm)	η (%) (BO)	Signal	PAPR (dB)	PAE_{avg}
2009 [36]	ET+DPA	2.14	43	50/58 (10/6 dB)	WCDMA (5 MHz)	8.0	50
2007 [25]	3W-DPA	2.14	41	43/50 (9/6 dB)	WCDMA (5 MHz)	10.0	40
2010 [26]	3W-DPA	2.65	50.5	51/56 (9/6 dB)	WiMAX (10 MHz)	7.8	48
2011 [39]	3W-DPA	2.14	50	60 (10 dB)	WCDMA (5 MHz)	6.5	48
2012 [32]	2W-DPA	1.8/2.4	43	60/44 (6 dB)	LTE (10 MHz)	7.0	54, 45
2013 [17]	Reconf. 2W-DPA	1.6–2.4 (40%)	42	50–62 (6 dB)	N/A	N/A	N/A
2013 [38]	Reconf. 2W-DPA	790–960 (19%)	52.5	50–62 (10 dB)	LTE (20 MHz)	10.5	44
2011 [21]	2W-DPA	2.2–2.96 (29%)	42	40–48 (6 dB)	WiMAX	10.0	25
2012 [23]	2W-DPA	1.7–2.25 (28%)	49	53–65 (6 dB)	WCDMA (5 MHz)	8.0	35
2012 [40]	2W-DPA	1.96–2.46 (23%)	42	40–43 (6 dB)	WiMAX	7.3	40
2012 [16]	2W-DPA	0.7–1.0 (35%)	50	50–68 (6 dB)	LTE (20 MHz)	10.5	42
2013 [41]	2W-DPA	1.6–2.25 (34%)	53	41–61 (6 dB)	LTE (3.8 MHz)	7.3	~47

Chapter 3

Multi-Way DPA With Extended Bandwidth

3.1   Introduction

As described in Chapter 2, the fractional bandwidth of 2W-DPAs has been successfully extended to about 30% in previous works. This chapter expounds the analysis and design of multi-way DPAs in order to maintain efficiency over a an extended power range (up to 12 dB) and extended fractional bandwidth. The architectures presented include conventional and modified multi-way DPAs. The synthesis equations are derived based on the basic analysis presented in Chapter 2. In particular, the behavior of three-way Doherty power amplifiers (3W-DPAs) is investigated as the carrier frequency shifts away from the center frequency. The proposed analysis helped identify circuit parameters that can be harnessed to maximize the bandwidth and improve back-off efficiency. A circuit prototype was designed and implemented to validate the analysis, which demonstrated excellent back-off efficiency enhancement over a wide frequency range.

3.2   Multi-Way DPA Analysis

An alternative for extended BO efficiency improvement is a "multi-way" or "N-way" DPA, sometimes called "multi-stage" DPA [25–27]. In the most general case, an N-way DPA consists of one main and $(N-1)$ peaking amplifiers, each biased at a different gate-source voltage, which results in N distinct definable efficiency peaks (see Figure 3.1 for instance).
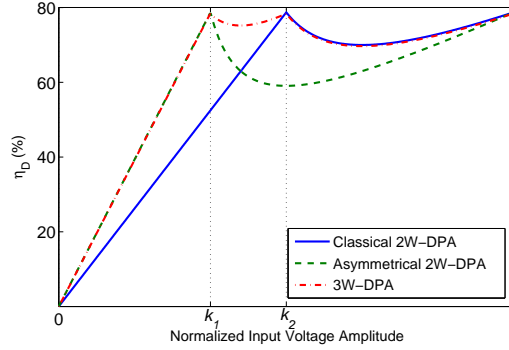
Figure 3.1: Drain efficiency comparison of three Doherty topologies.
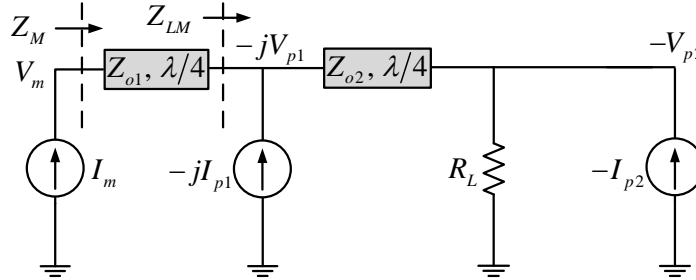
Figure 3.2: Simple schematic of a 3W-CDPA.

In practice, nevertheless, N is usually limited to 3, as adding further branches increases the complexity of the circuit. There are two multi-way DPA versions proposed in the literature, namely, the "conventional" and "modified" structures. The operation mechanism of the conventional and modified three-way DPAs are described in this section, for the purpose of simplicity. The presented analysis can be extended for an N-way DPA without loss of generality.

3.2.1 Conventional Three-Way DPA Theory

A general schematic of a three-way "conventional" DPA (3W-CDPA) architecture is depicted in Figure 3.2. The amplifier comprises a main and two peaking transistors, represented by ideal current sources, I_m, I_{p1}, I_{p2}, respectively. The idea is to employ the basic Doherty cell (the main and first peaking PA) as a new main amplifier for the second

peaking amplifier. A similar procedure to the one used in Section 2.5.2 for the 2W DPA must be followed to determine the circuit parameters for a given efficiency profile and a given output power. In other words, Z_{o1}, Z_{o2}, R_L, and the ratio between the transistors' peak current will be expressed as functions of k_1, k_2, and R_{opt}, where k_1 and k_2 represent the breakpoints at which the two efficiency peaks occur (see Figure 3.1) and R_{opt} denotes the optimum load impedance of the main transistor.

The general expressions of the current profiles are given by equations (3.1) through (3.3):

$$I_m(v_{in}) = \begin{cases} I_M \cdot v_{in}/k_2 & 0 < v_{in} < k_2 \\ I_M & v_{in} > k_2 \\ 0 & \text{elsewhere} \end{cases} \tag{3.1}$$

$$I_{p1}(v_{in}) = \begin{cases} I_{P1}\frac{v_{in}-k_1}{1-k_1} & k_1 < v_{in} < 1 \\ I_{P1} & v_{in} > 1 \\ 0 & \text{elsewhere} \end{cases} \tag{3.2}$$

$$I_{p2}(v_{in}) = \begin{cases} I_{P2}\frac{v_{in}-k_2}{1-k_2} & k_2 < v_{in} < 1 \\ I_{P2} & v_{in} > 1 \\ 0 & \text{elsewhere} \end{cases}. \tag{3.3}$$

where v_{in} is the normalized instantaneous input voltage. It should be noted that unlike the two-way DPA, the current of the main transistor must stay constant from the second breakpoint ($v_{in} = k_2$) up to the full power to enable a proper load modulation [25], as discussed later. Using (3.1)–(3.3), the equations for impedances seen by the transistors will be derived in three different power regions (i.e., low power, medium power, and high power).

In the low-power region ($v_{in} < k_1$), both peaking transistors are off, meaning that I_{p1} and I_{p2} are equal to 0 (see Figure 3.2). Hence,

$$V_m = Z_M \cdot I_m$$
$$= \left[\left(\frac{Z_{o1}}{Z_{o2}} \right)^2 \cdot R_L \right] \cdot I_m. \tag{3.4}$$

Note that at the first breakpoint ($v_{in} = k_1$), maximum efficiency should be obtained, meaning that $V_m = V_M \approx V_{dd}$, where V_{dd} represent the bias supply voltage of the main transistor. Therefore, (3.1) and (3.4) yield

$$\left(\frac{Z_{o1}}{Z_{o2}} \right)^2 \cdot R_L \cdot \left(\frac{k_1}{k_2} I_M \right) = V_M. \tag{3.5}$$

By definition, $V_M/I_M = R_{opt}$, where R_{opt} represents the optimum load impedance of the main transistor. Consequently,

$$\left(\frac{Z_{o1}}{Z_{o2}}\right)^2 R_L = \frac{k_2}{k_1} R_{opt}. \tag{3.6}$$

In the medium power region ($k_1 < v_{in} < k_2$), the first peaking transistor is turned on, so the impedance of the main transistor is modulated by I_{p1}, which can be expressed as

$$Z_M(v_{in}) = Z_{o1}^2/Z_{LM} \tag{3.7}$$

Using the ABCD-parameters of the two $\lambda/4$ transmission lines, Z_{LM} can be written as

$$\begin{aligned} Z_{LM} &= Z_L \left(1 + \frac{-jI_{p1}}{I_o}\right) \\ &= Z_L \left(1 + \frac{-jI_{p1}}{-jV_m/Z_{o1}}\right) \\ &= \frac{Z_{o2}^2}{R_L}\left(1 + \frac{Z_{o1}I_{p1}}{V_m}\right). \end{aligned} \tag{3.8}$$

Substituting (3.8) into (3.7) yields

$$\begin{aligned} Z_M &= V_m/I_m \\ &= \frac{R_L(Z_{o1}/Z_{o2})^2}{1 + Z_{o1}I_{p1}/V_m} \end{aligned} \tag{3.9}$$

which can be rearranged to give

$$Z_M = R_L \left(\frac{Z_{o1}}{Z_{o2}}\right)^2 - Z_{o1}\frac{I_{p1}}{I_m}. \tag{3.10}$$

As can be seen, (3.10) is very similar to the load modulation expression of a 2W-DPA, (2.10). It shows that the high output impedance of the main transistor in low power (first term in (3.10)) reduces to R_{opt} at $v_{in} = k_2$, owing to the current of the first peaking transistor, which prevents the main transistor saturation.

On the other hand, the voltage of the first peaking transistor can be determined using the current of the main transistor and the ABCD-parameters of the Z_{o1} transmission line (see Figure 3.2), as

$$V_{p1} = Z_{o1}I_m. \tag{3.11}$$

41

Note that the first peaking transistor reaches the maximum efficiency at $v_{in} = k_2$ meaning that $V_{p1} = V_{dd}$. Also, the main transistor provides its maximum current at this point, i.e., $I_m = I_M$. So it can be concluded that

$$Z_{o1} = R_{opt} \qquad (3.12)$$

which is exactly the same expression as in a 2W-DPA. On the other hand, writing the KCL at the output node of the second peaking transistor yields (note that $V_{p2} = k_2 \cdot V_{dd}$ at this point)

$$V_{p1}/Z_{o2} = V_{p2}/R_L$$
$$= k_2 \cdot V_{dd}/R_L \qquad (3.13)$$

or,

$$Z_{o2} = R_L/k_2. \qquad (3.14)$$

We can now readily determine Z_{o2} and R_L in terms of known parameters, k_1, k_2, and R_{opt}, using (3.14), combined with (3.6) and (3.12) as

$$Z_{o2} = k_1 R_{opt} \qquad (3.15)$$
$$R_L = k_1 k_2 R_{opt} \qquad (3.16)$$

Moreover, (3.10) evaluated at $v_{in} = k_2$ together with (3.12), can be used to determine the ratio of the peak current of the main device to that of the first peaking device, I_M/I_{P1}, as

$$\frac{I_M}{I_{P1}} = \frac{1}{k_1} - 1 \qquad (3.17)$$

In the high power region, all three transistors are operating; so the current and voltage of the main device must remain constant in this region; thus the required power is supplied only by the two peaking transistors. One may notice that the entire circuit parameters (Z_{o1}, Z_{o2}, and R_L) are already determined based on the operation mechanism of the DPA in low to medium power ranges. This is related to the fact that unlike a classical two-way DPA, in this architecture the second peaking transistor does not modulate the impedance of the main or the first peaking transistor. In fact, the voltage across the first peaking transistor is dictated by the main current source transformed into a voltage source by the Z_{o1} impedance inverter, which is already kept constant. In other words, the current from the main transistor needs to be clipped at $v_{in} = k_2$ or otherwise the current injected from the second peaking transistor will interfere with the load modulation of the main transistor which drives it excessively into the knee region, thereby causing significant linearity

42
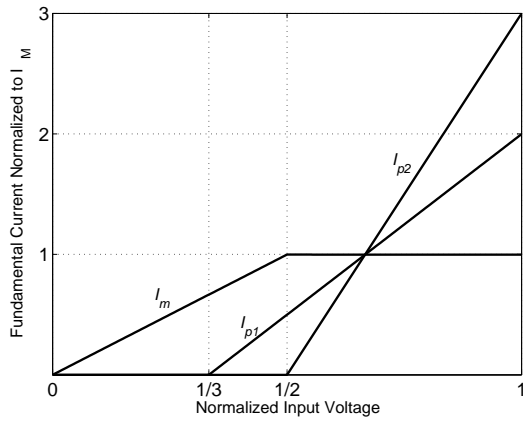
(a) Current profile of a 3W-CDPA for $k_1 = 1/3$ and $k_2 = 1/2$.
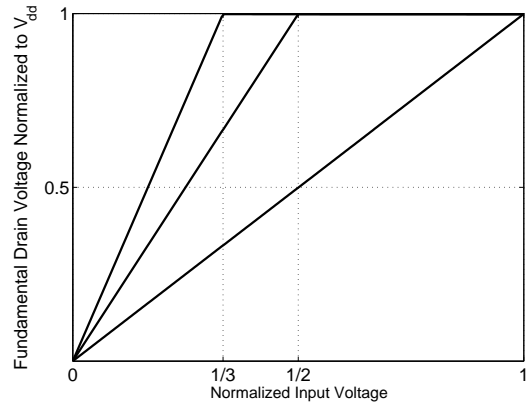
(b) Voltage profile of a 3W-CDPA for $k_1 = 1/3$ and $k_2 = 1/2$.
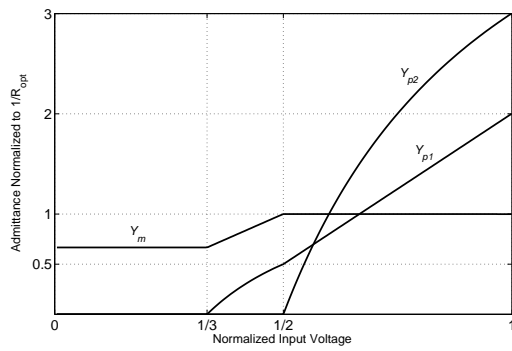
(c) Admittance profile of a 3W-CDPA for $k_1 = 1/3$ and $k_2 = 1/2$.
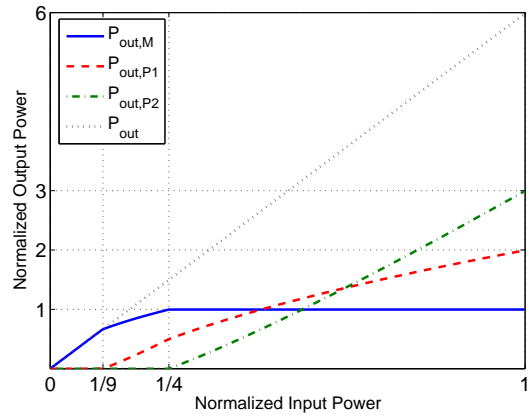
(d) Output power profile of a 3W-CDPA for $k_1 = 1/3$ and $k_2 = 1/2$.

Figure 3.3: Characteristic profiles of a 3W-CDPA.

degradation. Hence, it is inferred that load modulation concept does not exist during the high power region anymore, and the current (or power) from the second peaking transistor only maintains linearity of the 3W-CDPA. This linearity requirement defines the amount of current needed from the peaking #2 device. The KCL at the output node can be rewritten to account for the second peaking current, I_{p2}, in the high power region. So,

$$V_{p1}/Z_{o2} = V_{p2}/R_L - I_{p2} \qquad (3.18)$$

Substituting for Z_{o2} and R_L using (3.15) and (3.16), (3.18) evaluated at the peak power $(v_{in} = 1)$ results in

$$\frac{I_M}{I_{P2}} = \frac{1}{k_1}\left(\frac{1}{k_2} - 1\right) \qquad (3.19)$$

which completes our analysis of the 3W-CDPA. To sum, knowing R_{opt} which is defined by the targeted output power, together with k_1, and k_2, which are specified by the desired profile of efficiency versus power, the circuit parameters, Z_{o1} and Z_{o2}, and R_L as well as the peak current ratios of the transistors are uniquely determined. The fundamental current and voltage profiles are depicted in Figure 3.3(a) and (b), using (3.1)–(3.3), (3.17), and (3.19) for $k_1 = 1/3$ and $k_2 = 1/2$, corresponding to 9.5 dB and 6 dB power back-off, respectively. The admittance seen by the transistors can be expressed as a function of normalized input voltage as

$$Y_M(v_{in}) = \begin{cases} \frac{k_1}{k_2}G_{opt}, & 0 < v_{in} < k_1 \\[2mm] \frac{v_{in}}{k_2}G_{opt}, & k_1 < v_{in} < k_2 \\[2mm] G_{opt}, & k_2 < v_{in} < k_1 \end{cases} \qquad (3.20)$$

$$Y_{P1}(v_{in}) = \begin{cases} 0, & 0 < v_{in} < k_1 \\[2mm] \frac{k_2}{k_1}G_{opt}\left(1 - \frac{k_1}{v_{in}}\right), & k_1 < v_{in} < k_2 \\[2mm] \frac{G_{opt}}{k_1}(v_{in} - k_1), & k_2 < v_{in} < 1 \end{cases} \qquad (3.21)$$

$$Y_{P2}(v_{in}) = \begin{cases} 0, & 0 < v_{in} < k_2 \\[2mm] \frac{G_{opt}}{k_1 k_2}\left(1 - \frac{k_2}{v_{in}}\right), & k_2 < v_{in} < 1 \end{cases} \qquad (3.22)$$

where $G_{opt} = 1/R_{opt}$. The admittance functions are depicted in Figure 3.3(c), for $k_1 = 1/3$ and $k_2 = 1/2$. As can be seen, the main impedance is only modulated in the medium power region $(k_1 < v_{in} < k_2)$, and is constant elsewhere. Note that there is a fundamental

44

challenge in physical realization of the 3W-CDPA, as mentioned before, due to the required main current profile. Hence, mixed-signal approaches have been proposed in the literature as the only solution to control the voltage drive of the main stage.

It is noteworthy at this point to calculate the power distribution among the three transistors. At the peak power, the output power from each transistor is given by

$$P_{out,M} = \frac{1}{2}\Re\{\mathbf{V_m} \cdot \mathbf{I_m}^*\}$$
$$= \frac{1}{2}V_{dd} \cdot I_M \qquad (3.23)$$
$$P_{out,P1} = \frac{1}{2}\Re\{\mathbf{V_{p1}} \cdot \mathbf{I_{p1}}^*\}$$
$$= \left(\frac{1}{k_1} - 1\right)P_{out,M} \qquad (3.24)$$
$$P_{out,P2} = \frac{1}{2}\Re\{\mathbf{V_{p2}} \cdot \mathbf{I_{p2}}^*\}$$
$$= \frac{1}{k_1}\left(\frac{1}{k_2} - 1\right)P_{out,M}. \qquad (3.25)$$

For instance, for $k_1 = 1/3$ and $k_2 = 1/2$, the output power ratio from the main, and peaking #1 and #2 transistors is 1:2:3, as shown in Figure 3.3(d). In other words, even though the main amplifier has a dominant impact on the back-off efficiency, it provides only one-sixth of the total output power. This is a totally different situation than the 2W-DPA, where the main and peaking transistor contribute equally to the output power, and thus the main transistor plays far more important role than the peaking transistor, as discussed in Section 2.5. In short, in a 3W-CDPA the efficiency at back-off is related to the main device performance, while the linearity mainly depends on the performance of the peaking transistors.

3.2.2   Modified Three-Way DPA Theory

To address the complexity of a 3W-CDPA, an alternative output combining network has been proposed in [34], shown in Figure 3.4. The resultant topology is referred to as the three-way "modified" Doherty power amplifier (3W-MDPA). It can be shown that the new topology enables the main transistor to follow the normal linear trend with no need for
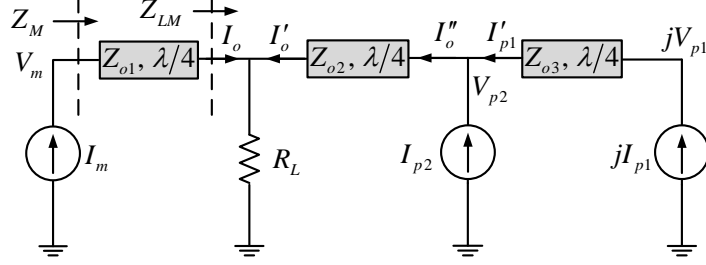
Figure 3.4: Simple schematic of a 3W-MDPA.

additional circuitry. The current profile of the transistors are expresses by

$$
I_m(v_{in}) = \begin{cases} I_M \cdot v_{in}, & 0 < v_{in} < 1 \\ I_M, & v_{in} > 1 \\ 0, & \text{elsewhere} \end{cases} \tag{3.26}
$$

$$
I_{p1}(v_{in}) = \begin{cases} I_{P1}\frac{v_{in}-k_1}{1-k_1}, & k_1 < v_{in} < 1 \\ I_{P1}, & v_{in} > 1 \\ 0, & \text{elsewhere} \end{cases} \tag{3.27}
$$

$$
I_{p2}(v_{in}) = \begin{cases} I_{P2}\frac{v_{in}-k_2}{1-k_2} & k_2 < v_{in} < 1 \\ I_{P2}, & v_{in} > 1 \\ 0, & \text{elsewhere} \end{cases}. \tag{3.28}
$$

In the low power region $(v_{in} < k_1)$, I_{p1} and I_{p2} are equal to 0 (see Figure 3.4). Hence,

$$
\frac{Z_{o1}^2}{R_L}I_m = V_m. \tag{3.29}
$$

Note that at $v_{in} = k_1$, the maximum efficiency should be attained, meaning that $V_m = V_M = V_{dd} - V_k$, where V_k and V_M denote the knee voltage and the maximum voltage swing of the main transistor, respectively. Therefore, using (3.29) and (3.26), we obtain

$$
\frac{Z_{o1}^2}{R_L}(k_1 I_M) = V_M. \tag{3.30}
$$

Note that $V_M/I_M = R_{opt}$, where R_{opt} represents the optimum load impedance of the main transistor. Thus, Z_{o3} can be expressed as

$$
Z_{o1} = \sqrt{\frac{1}{k_1} \cdot R_L \cdot R_{opt}}. \tag{3.31}
$$

46

In the medium power region ($k_1 < v_{in} < k_2$), the first peaking transistor is turned on, so the impedance of the main transistor is modulated by I_{p1}, which can be calculated as

$$Z_M(v_{in}) = Z_{o1}^2 / Z_{LM}. \tag{3.32}$$

Using the ABCD-parameters of the three $\lambda/4$ transmission lines, Z_{LM} can be written as

$$\begin{aligned}
Z_{LM} &= R_L \left(1 + \frac{I_o'}{I_o} \right) \\
&= R_L \left(1 + \frac{-(jI_{p1}) \cdot (-jZ_{o1}) \cdot (-j/Z_{o2})}{-jV_m/Z_{o3}} \right) \\
&= R_L \left(1 + \frac{Z_{o1}Z_{o3}}{Z_{o2}} \cdot \frac{I_{p1}}{V_m} \right).
\end{aligned} \tag{3.33}$$

Substituting (3.33) into (3.32) yields

$$\begin{aligned}
Z_M(v_{in}) &= V_m / I_m \\
&= \frac{Z_{o1}^2}{R_L \left(1 + \frac{Z_{o1}Z_{o3}}{Z_{o2}} \cdot \frac{I_{p1}}{V_m} \right)}
\end{aligned} \tag{3.34}$$

which can be rearranged to give

$$Z_M(v_{in}) = \frac{Z_{o1}^2}{R_L} - \frac{Z_{o1}Z_{o3}}{Z_{o2}} \cdot \frac{I_{p1}}{I_m}. \tag{3.35}$$

As can be seen, (3.35) is very similar to the load modulation expression derived in Section 2.5 for 2W-DPA. It shows that the current injected to the load by the first peaking transistor reduces the output impedance of the main one, thereby preventing its saturation.

On the other hand, the voltage of the first peaking transistor can be determined using the current of the main transistor and the ABCD-parameters of the transmission lines as

$$jV_{p1} = I_m \cdot (-jZ_{o1}) \cdot (-j/Z_{o2}) \cdot (-jZ_{o3}). \tag{3.36}$$

Note that the first peaking transistor reaches the maximum efficiency at $v_{in} = k_2$ meaning that $V_{p1} = V_{dd} - V_k = V_M$. Also, $I_m = k_2 I_M$ at this point, so we can conclude that

$$\frac{Z_{o1}Z_{o3}}{Z_{o2}} = \frac{R_{opt}}{k_2}. \tag{3.37}$$

Moreover, (3.35) evaluated at $v_{in} = k_2$ together with (3.37), can be used to determine the ratio of the peak current of the first peaking device to that of the main device, I_{P1}/I_M, as

$$I_{P1}/I_M = k_2 \left(1/k_1 - 1 \right). \tag{3.38}$$

In the high power region, all three transistors are operating. Note that the current of the second peaking transistor only modulates the load of the first peaking transistor, and thus does not affect the impedance seen by the main transistor, as long as the first peaking device does not enter into the knee region. As a result, (3.35) still holds in the high power region. In fact, the voltage across the second peaking transistor is dictated by the first peaking current source transformed into a voltage source by the $\lambda/4$ impedance inverter with the characteristic impedance of Z_{o1}. Thus, at the maximum input voltage

$$V_{P2} = V_{dd} - V_k = Z_{o3} I_{P1}. \tag{3.39}$$

In other words, Z_{o1} must be equal to the optimum load resistance of the first peaking transistor. Substituting for I_{P1} using (3.38), we can conclude that

$$Z_{o3} = \frac{R_{opt}}{k_2(1/k_1 - 1)}. \tag{3.40}$$

Note that Z_{o2} can be readily determined from (3.37), using Z_{o1} and Z_{o3} of (3.40) and (3.31), respectively.

Additionally, (3.36) can be rewritten to account for the second peaking current, I_{p2}, in the high power region. So,

$$jV_{p1} = jZ_{o3}I'_{p1} = jZ_{o1}(I''_o - I_{p2})$$
$$= (jZ_{o3}) \left[I_m(-jZ_{o1})(j/Z_{o2}) - I_{p2} \right]$$

or,

$$V_{p1} = Z_{o3} I_m \left(\frac{Z_{o1}}{Z_{o2}} - \frac{I_{p2}}{I_m} \right). \tag{3.41}$$

Evaluating (3.41) at $v_{in} = 1$ gives

$$\frac{I_{P2}}{I_M} = \frac{Z_{o1}}{Z_{o2}} - \frac{R_{opt}}{Z_{o3}} \tag{3.42}$$

which can be combined with (3.37) and (3.40) to conclude

$$I_{P2}/I_M = (1/k_1 - 1)(1 - k_2). \tag{3.43}$$
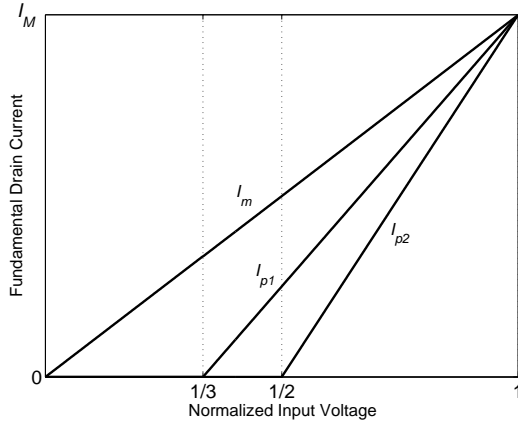
(a) Current profile of a 3W-MDMA for $k_1 = 1/3$ and $k_2 = 1/2$.
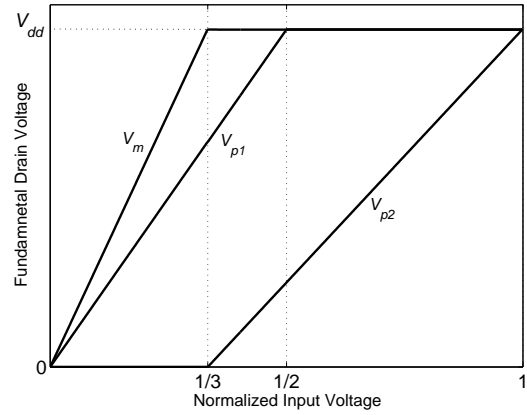
(b) Voltage profile of a 3W-MDMA for $k_1 = 1/3$ and $k_2 = 1/2$.
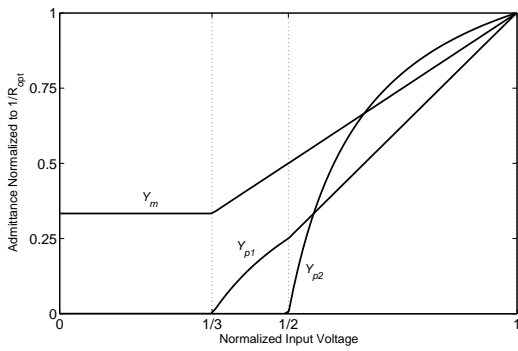
(c) Admittance profile of a 3W-MDMA for $k_1 = 1/3$ and $k_2 = 1/2$.
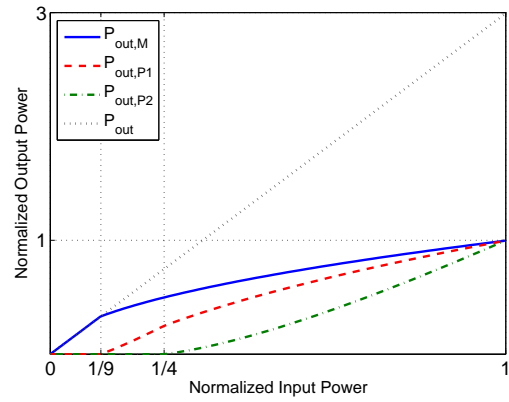
(d) Output power profile of a 3W-MDMA for $k_1 = 1/3$ and $k_2 = 1/2$.

Figure 3.5: Characteristic profiles of a 3W-MDPA.

49

The fundamental current and voltage profiles are depicted in Figure 3.5(a) and (b), using (3.26)–(3.28), (3.38), and (3.43) for $k_1 = 1/3$ and $k_2 = 1/2$. It is worth noting that the voltage swing of the second peaking transistor, V_{p2}, is 0 in the low power region, since it is proportional to I_{p1}, which is 0 for $v_{in} < k_1$. Using the fundamental current and voltage profiles, one can calculate the admittance seen by each transistor in different power regions as,

$$Y_M(v_{in}) = \begin{cases} k_1 \cdot G_{opt}, & 0 < v_{in} < k_1 \\ \\ v_{in} \cdot G_{opt}, & k_1 < v_{in} < 1 \end{cases} \tag{3.44}$$

$$Y_{P1}(v_{in}) = \begin{cases} 0, & 0 < v_{in} < k_1 \\ \\ \frac{k_1}{(1-k_1)^2} G_{opt}\left(1 - \frac{k_1}{v_{in}}\right), & k_1 < v_{in} < k_2 \\ \\ \frac{k_1}{k_2} G_{opt} \frac{v_{in}-k_1}{(1-k_1)^2}, & k_2 < v_{in} < 1 \end{cases} \tag{3.45}$$

$$Y_{P2}(v_{in}) = \begin{cases} \text{indefinite}, & 0 < v_{in} < k_1 \\ \\ 0, & k_1 < v_{in} < k_2 \\ \\ \frac{k_1}{(1-k_2)^2} G_{opt} \frac{v_{in}-k_2}{v_{in}-k_1}, & k_2 < v_{in} < 1 \end{cases} . \tag{3.46}$$

The admittance profiles of the three transistors versus the normalized input voltage are plotted in Figure 3.5(c) for the special case of $k_1 = 1/3$ and $k_2 = 1/2$. As can be seen, in this particular case, the optimum load resistances (at peak input voltage) of all the transistors are equal to R_{opt}, which may not be the case in general, according to (3.44)–(3.46).

It can be observed from Table 3.1 that for a given output power (or equivalently R_{opt}) and efficiency profile (k_1 and k_2), the three remaining equations in the table are sufficient to fully describe the 3W-DPA's behavior. Considering the four unknown circuit parameters (Z_{o1}, Z_{o2}, Z_{o3}, and R_L) it can be inferred that there is not a unique solution for this linear system of equations. Thus, we can set R_L as an independent parameter and calculate the other parameters based on that. This is a significant conclusion, since it implies that there are infinite possible sets of design parameters, all yielding the exact same performance (output power, efficiency, etc.) at the frequency of design. As will be explained later, one can harness R_L to maximize the efficiency of the 3W-DPA over the bandwidth.

Table 3.1: 3-Way conventional and modified DPA design parameters

Circuit Parameter	3W-CDPA	3W-MDPA
R_{opt}	$\frac{V_{dd}-V_k}{I_M}$	$\frac{V_{dd}-V_k}{I_M}$
Z_{o1}	R_{opt}	$\sqrt{\frac{1}{k_1}R_L R_{opt}}$
Z_{o2}	$k_1 R_{opt}$	$\frac{1}{1/k_1-1}\sqrt{\frac{1}{k_1}R_L R_{opt}}$
Z_{o3}	—	$\frac{R_{opt}}{k_2(1/k_1-1)}$
R_L	$k_1 k_2 R_{opt}$	—

The output power of each PA cell in a 3W-DPA is expressed as

$$
\begin{aligned}
P_{out,M} &= \frac{1}{2}\mathfrak{R}\{\mathbf{V_m}\cdot\mathbf{I_m}^*\} \\
&= \frac{1}{2}(V_{dd}-V_k)\cdot I_M
\end{aligned}
\tag{3.47}
$$

$$
\begin{aligned}
P_{out,P1} &= \frac{1}{2}\mathfrak{R}\{\mathbf{V_{p1}}\cdot\mathbf{I_{p1}}^*\} \\
&= k_2\left(1/k_1-1\right)P_{out,M}
\end{aligned}
\tag{3.48}
$$

$$
\begin{aligned}
P_{out,P2} &= \frac{1}{2}\mathfrak{R}\{\mathbf{V_{p2}}\cdot\mathbf{I_{p2}}^*\} \\
&= (1/k_1-1)(1-k_2)P_{out,M}.
\end{aligned}
\tag{3.49}
$$

For example, for $k_1 = 1/3$ and $k_2 = 1/2$, the output power ratios from all of the transistors are equal, similar to the 2W-DPA. This benefit results in better thermal distribution/management, particularly for high output power applications. Moreover, for $k_2 = 1/2$, i.e. the second efficiency peak at 6-dB BO which is often the case in practice, the first and second peaking transistors provide the same current and output power at peak point. The output power profiles are depicted in Figure 3.5(d).

Table 3.1 summarizes the synthesis equations for conventional and modified three-way DPAs for a given k_1 and k_2. It can be observed from Table 3.1 that for a given output power (or equivalently R_{opt}) and efficiency profile (k_1 and k_2), the three remaining equations in the table are sufficient to fully describe the 3W-DPA's behavior. Considering the

four unknown circuit parameters (Z_{o1}, Z_{o2}, Z_{o3}, and R_L) it can be inferred that there is not a unique solution for this linear system of equations. Thus, we can set R_L as an independent parameter and calculate the other parameters based on that. This is a significant conclusion, since it implies that there are infinite possible sets of design parameters, all yielding the exact same performance (output power, efficiency, etc.) at the frequency of design. As will be explained later, one can harness R_L to maximize the efficiency of the 3W-DPA over the bandwidth.

3.2.3 Bandwidth Analysis of 3W DPAs

Thus far, the operation of the 3W-DPA has been discussed at the center frequency. Both the efficiency and linearity will degrade as the carrier frequency shifts above or below the design frequency. This section studies the frequency behavior of the 3-way modified and conventional DPA architectures, respectively. For the purpose of simplicity, the ideal current sources will still be used to replace the actual transistors. Therefore, the frequency variation will be solely attributed to the output combining network.

As described in Section 3.2.2, there are four circuit elements in the architecture to be determined, which include the three inverter impedances (Z_{o1} through Z_{o3}) and the load resistance (R_L); while the number of conditions to satisfy the governing equations is three (namely at $v_{in} = k_1, k_2, 1$). Therefore, there will be a degree of freedom, say R_L, which can be judiciously selected for optimal performance versus bandwidth. In contrast, in the two-way or conventional three-way topology, all the circuit parameters are uniquely determined given the current profiles.

Simulations based on ideal class B and C modes of operation can provide insights about the bandwidth of the structure. For all cases, the optimum impedance R_{opt} is assumed to be 35 Ω, corresponding to a 10-W GaN HEMT, and k_1 and k_2 are chosen to be 1/3 and 1/2, respectively. Figures 3.6(a)-(c) illustrate the frequency behavior of the 3W-MDPA's drain efficiency, compared with the 2W-DPAs, for three different R_L values and three different input back-off (IBO) levels (9.5 dB, 6 dB, and 0 dB or full power). Note that for a fair benchmarking, the efficiency bandwidth of 3W-MDPA is compared against asymmetrical 2W-DPA at 9.5-dB IBO and classical 2W-DPA at 6-dB IBO to demonstrate the superior bandwidth of the 3W-MDPA at the respective back-off levels. As illustrated in Figure 3.6(a), at 9.5 dB IBO, the drain efficiency response reaches a plateau at R_L equal to 17 Ω and the efficiency drops as we deviate from this value. The same optimum point holds at 6-dB IBO compared to the classical 2W-DPA, as depicted in Figure 3.6(b).

Taking the 4% efficiency drop bandwidth as a reference, we observe that the fractional
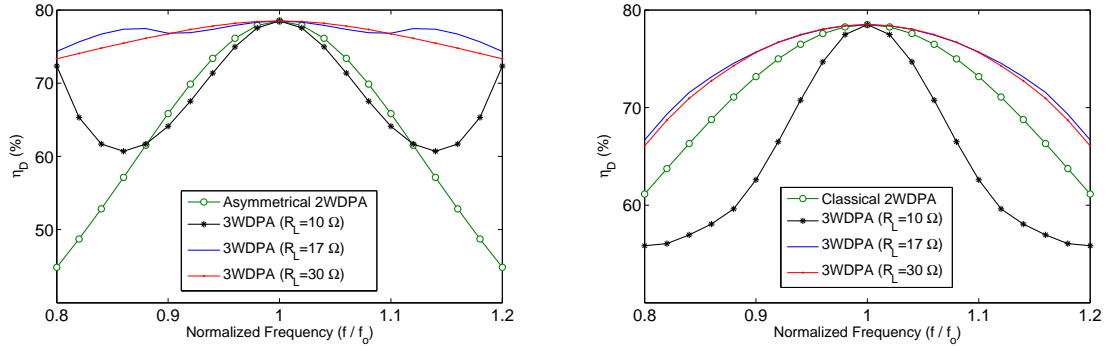
(a) Drain efficiency vs. frequency of a 3W-MDPA with $R_{opt} = 35\,\Omega$ at 9.5 dB IBO.

(b) Drain efficiency vs. frequency of a 3W-MDPA with $R_{opt} = 35\,\Omega$ at 6 dB IBO.
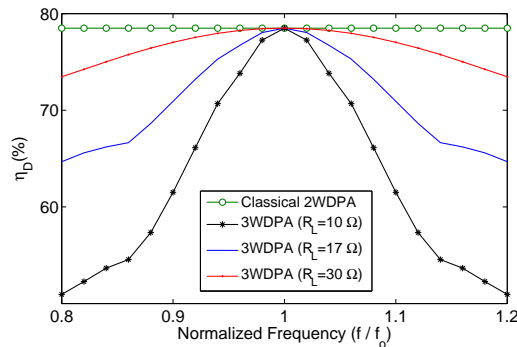
(c) Drain efficiency vs. frequency of the 3W-MDPA at 0 dB IBO (full power).

Figure 3.6: Drain efficiency bandwidth comparison of the 3W-MDPA with three different load resistances versus classical and asymmetric 2W-DPA, for $R_{opt} = 35\,\Omega$.

bandwidth of the optimal 3W-MDPA at 9.5-dB back-off is 40%, while that of the asymmetric 2W-DPA is only 10% for the same back-off level, according to Figure 3.6(a). At 6-dB IBO, the bandwidth of the 3W-MDPA is 24%, compared to 17% of the 2W-DPA, as shown in Figure 3.6(b). Finally, at full output power, the fractional bandwidth of the 3W-MDPA is 14%, whereas the ideal 2W-DPAs (both classical and asymmetrical configurations) are frequency independent. As mentioned before, since the PA is often driven by modulated signals with a typical PAPR of larger than 9 dB, optimal R_L should be selected based on the first two back-off levels, as they will have the dominant impact on the average efficiency over the bandwidth. It should be noted that for all the simulations, the standard 50 Ω

53

load is matched to R_L through a $\lambda/4$ transmission line, as it resulted in larger bandwidth compared to an ideal R_L load.

Mathematically speaking, Z_M can be expressed as a function of R_L, frequency (f), and R_{opt}, referred to as $Z_M(f, R_L, R_{opt})$. The act of providing a fixed impedance for the main transistor at 9.5-dB back-off determines the 9.5-dB efficiency bandwidth of the DPA. Therefore, a cost function should be defined to maximize the efficiency bandwidth versus R_L. The proper cost function is the standard deviation of the impedance of the main transistor versus frequency, defined as

$$
\text{stdev}[Z_M] \triangleq \left\{ \frac{1}{\Delta f} \int_{f_o - \Delta f/2}^{f_o + \Delta f/2} \left| Z_M(f, R_L, R_{opt}) \right. \right.
$$
$$
\left. \left. - Z_M(f_o, R_L, R_{opt}) \right|^2 df \right\}^{1/2} \tag{3.50}
$$

There will be a global minimum for the cost function for a given bandwidth Δf and R_{opt} (which depends on the target output power). Thus, R_L can be set to provide the maximum efficiency over the bandwidth. Figure 3.7 illustrates the standard deviation function defined in (3.50), calculated over 30% fractional bandwidth versus the load resistance, for different values of R_{opt}.

Similar analyses can be conducted for the output impedance seen by each of the two peaking transistors. Nevertheless, It was realized that the standard deviation function is much more frequency sensitive at 9.5-dB IBO. Hence, the cost function at 9.5-dB IBO [Figure 3.6(a)] could be used for the initial design of the 3W-DPA.

The efficiency versus bandwidth of the three-way modified (with optimal R_L) and the conventional architecture is compared in Figure 3.8 for $k_1 = 1/3$, $k_2 = 1/2$, and $R_{opt} = 35\Omega$. This figure clearly highlights the advantage of the conventional DPA over the modified topology, even with optimal R_L. In terms of output power, as illustrated in Figure 3.8(b), the conventional configuration is also superior since all the $\lambda/4$ lines are terminated with their characteristic impedance as is the case in the conventional 2W-DPA architecture, and thus are frequency-independent. A solution will be proposed in Chapter 4 for realizing the required current profile of the main transistor (saturation at k_2).
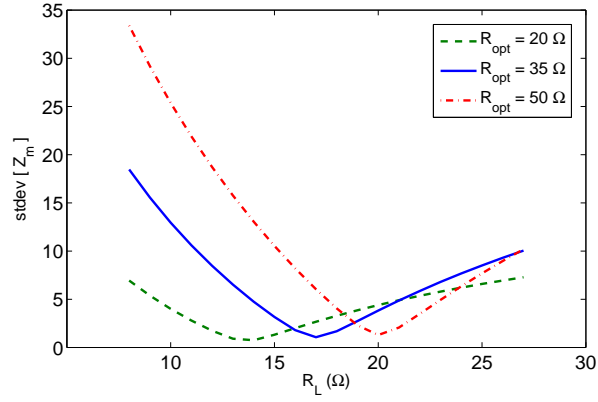
Figure 3.7: The standard deviation of the main transistor impedance vs. the load resistance for different R_{opt} values.
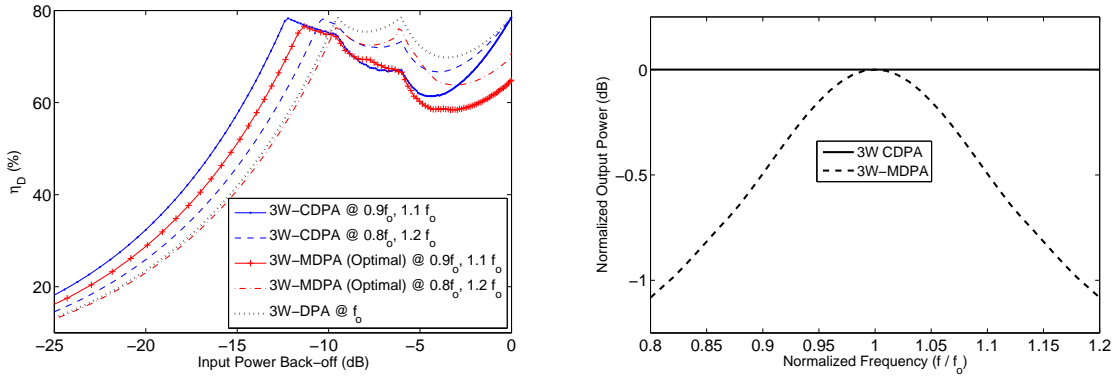
(a) Efficiency vs. bandwidth of the conventional and modified 3W-DPAs.

(b) Output power vs. bandwidth of the conventional and modified 3W-DPAs.

Figure 3.8: Comparison between frequency characteristics of the conventional and modified architectures.
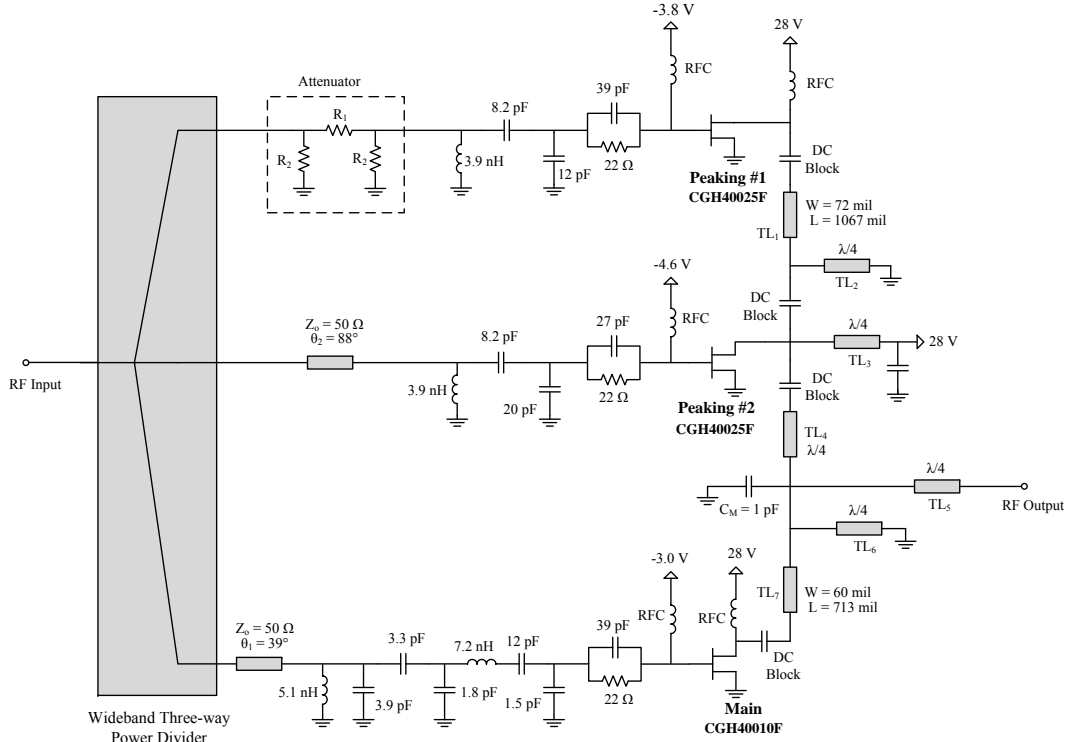
Figure 3.9: Complete schematic of the 3W-DPA.

3.3 Design of A Broadband 30-W GaN Three-Way Modified DPA

Among various DPA configurations discussed thus far, 3W-MDPA is chosen as a proof of concept. Therefore, this section briefly describes the design and validation of a broadband 3W-MDPA prototype More detailed explanations can be found in [42]. The 3W-MDPA was designed with $k_1 = 1/3$ and $k_2 = 1/2$, peak output power of 30 W, and a targeted frequency band of 0.73 to 0.98 GHz. To attain this power level, a 10 W, and two 25 W GaN packaged transistors from Cree® were used for the main, peaking #1, and peaking #2 PAs, respectively. Larger devices were exploited for the two peaking PAs due to the larger transconductance required for class C biased transistors to insure proper load modulation. An oversized device was used for the first peaking PA (25 W instead of 15 W) due to the device availability limitation. In this section, the strategies to design each part of the
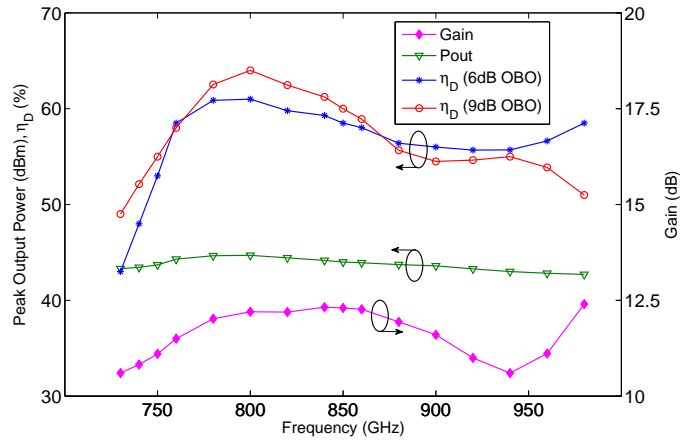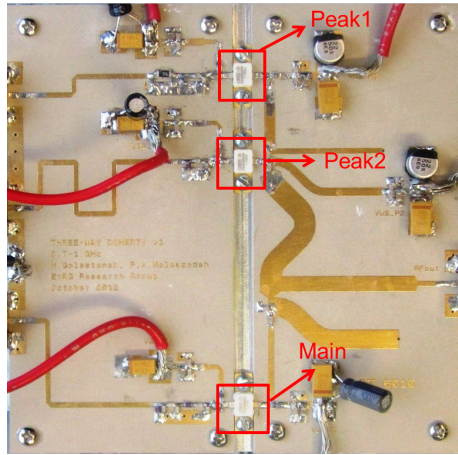
(a) A photo of the fabricated 3W-DPA.

(b) The measured CW peak output power, the associated gain, and drain efficiency at 6 and 9 dB back-off power vs. frequency.
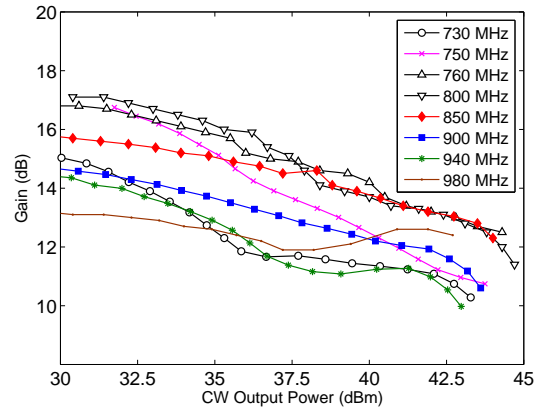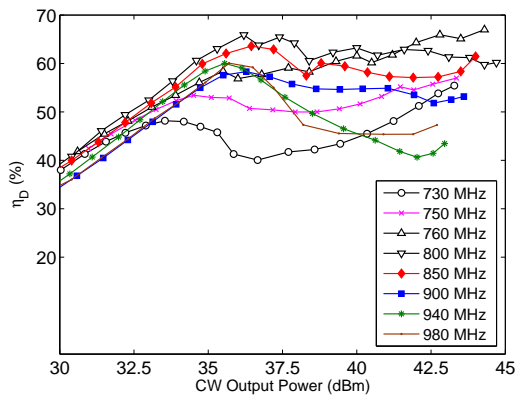
(c) Measured CW drain efficiency vs. output power and frequency.

(d) Measured CW gain vs. output power and frequency.

Figure 3.10: The fabricated 3W-DPA and its CW measurement results.
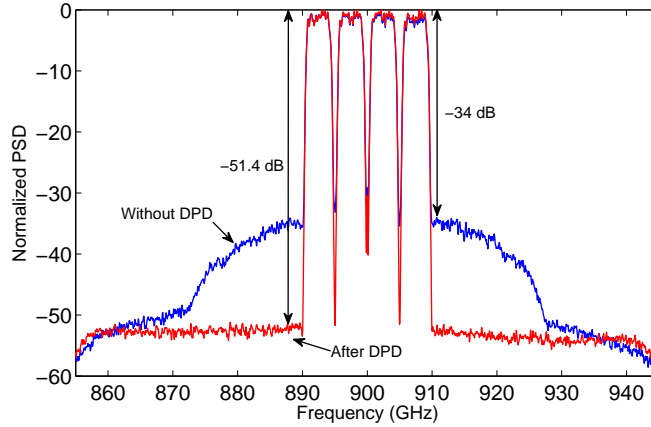
(a) The output power spectral density (PSD) with and without linearization under clipped 20 MHz WCDMA 1111 signal at 900 MHz.
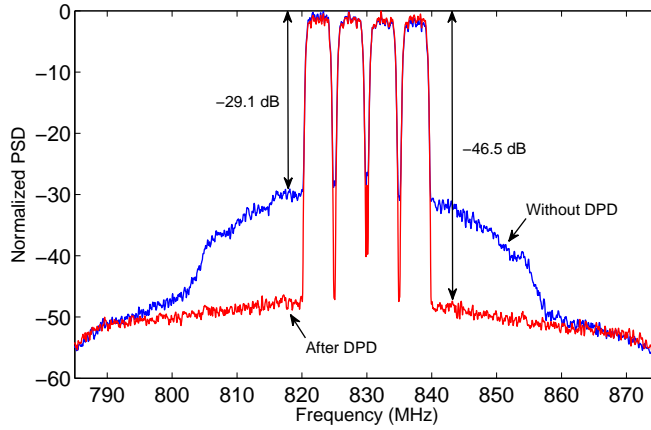
(b) The output PSD with and without linearization under unclipped 20 MHz WCDMA 1111 signal at 830 MHz.

Figure 3.11: The output PSD of the 3W-MDPA under modulated test signals.

circuit will be briefly described.

For the 10-W main transistor, R_{opt} is $35\,\Omega$. As per Figure 3.7, the optimal R_L for maximizing efficiency over the bandwidth was found to be $17\,\Omega$. Subsequently, a $\lambda/4$ line (TL$_5$ in Figure 3.9) is used to match $17\,\Omega$ to the standard $50\,\Omega$ termination. In order to compensate for a device's intrinsic and parasitic output elements, a number of solutions have been proposed in the literature. In particular, it has been suggested that the linear

(a) AM/AM and AM/PM characteristics at 900 MHz.

(b) AM/AM and AM/PM characteristics at 830 MHz.
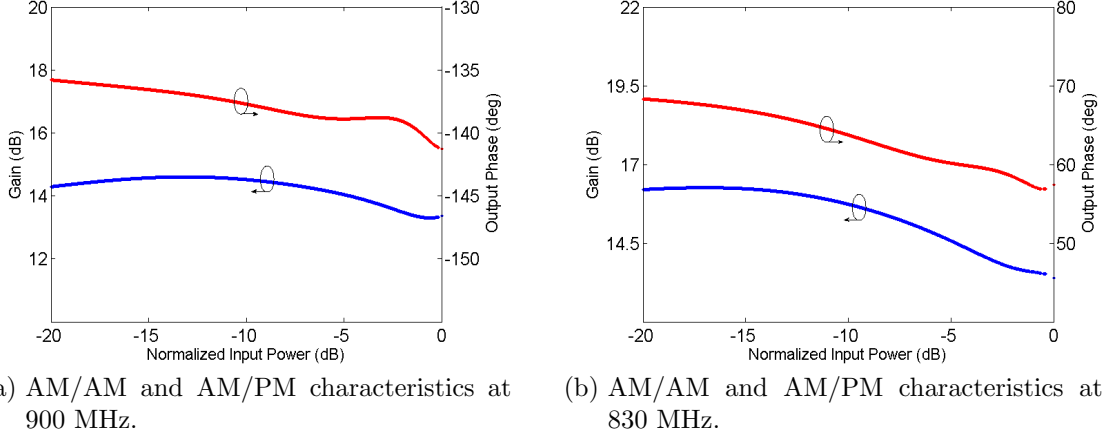
Figure 3.12: Measured AM/AM and AM/PM of the wideband 3W-DPA under modulated signal.

output capacitance of a GaN HEMT can be absorbed using the Π-equivalent circuit of a quarter-wave line (also known as quasi-lumped transmission line) [16, 18]. Alternatively, an inductance may be used to simply cancel (resonate) out the capacitance [20]. The latter may result in a narrower bandwidth due to the resultant high-Q resonance circuit. A wider bandwidth can be obtained using a negative parallel capacitance, provided it is absorbable into an equal or larger external capacitor. For this design, we have used the former (absorption method) for the Z_{o1} and Z_{o3} transmission lines to compensate for the output capacitance of the main and the first peaking transistor, and the cancellation method was applied to the second peaking transistor. It is of note that the two peaking transistors have equal output capacitances since they are the same size (25 W); thus, the positive and negative capacitances will cancel each other out, and therefore, no external inductance is required in this case.

To control the second harmonic impedance, an explicit short circuit $\lambda/4$ stub was employed for each device (TL_6, TL_2 and TL_3). The characteristic impedance of the stubs were individually optimized to achieve high-efficiency across the bandwidth. Higher order harmonics were found to be insignificant and so were not considered in the design. It is worth mentioning that the design did not require explicit matching network due to the low parasitic effects of the transistor package at the targeted frequency band.

Harmonic source-pull simulation has been used at each frequency point within the band of interest to determine the required fundamental and second harmonic impedances

for each transistor. The broadband input matching network (IMN) was then realized using the simplified real frequency technique to achieve a bandpass frequency response. The input delay-lines (θ_1 and θ_2 in Figure 3.9) were tuned to align the phase responses of the main and peaking #1 paths. As previously mentioned, the peaking #1 device was oversized. Since an external equi-splitting power divider has been used, a 2 dB broadband resistive attenuator was added to the input of the first peaking PA to adjust the transconductance of this stage. Alternatively, one could have designed a broadband three-way divider with the desired power division ratio to avoid the additional attenuator. Figure 3.9 depicts the complete circuit schematic including microstrip lines and discrete L-C elements.

Figure 3.10(a) shows a picture of the fabricated 3W-DPA. Note an external wideband power divider with even power division has been employed at the input to obtain a fully analog single-ended 3W-DPA. The peak output power and the associated gain as well as the drain efficiency (at 6 and 9-dB output back-off) of the 3W-DPA versus frequency measured under continuous wave (CW) stimulus are shown in Figure 3.10(b). At 9-dB output power back-off, a drain efficiency of higher than 49% was achieved between 730 MHz and 980 MHz, and higher than 51% in the frequency range of 750–980 MHz. Also, the peak output power varies between 42.7 dB and 44.6 dBm across the bandwidth. Figure 3.10(c) displays the measured drain efficiency as a function of output power. One can clearly observe the efficiency plateau in the back-off power range, which confirms proper Doherty operation over the entire frequency band. The peak drain efficiency was degraded compared to the simulated results, attributed to the thermal effects arising under CW stimulus. Figure 3.10(d) shows the CW gain of the amplifier at various test frequencies.

To evaluate the average efficiency as well as the linearizability of the wideband 3W-DPA, the PA was driven by a 20 MHz four-carrier wideband code division multiple access (WCDMA) signal at 900 MHz and 830 MHz, separately. While the WCDMA signal at 900 MHz was clipped to a PAPR of 7.14 dB, we applied an unclipped signal at 830 MHz, with a PAPR of 11.7 dB. To linearize the 3W-DPA, a digital pre-distortion (DPD) platform based on pruned Volterra series was used [43]. The output spectra of the 3W-DPA driven by the clipped 20 MHz WCDMA test signal at 900 MHz is depicted in Figure 3.11(a). As can be seen, ACLR was improved from –34 dB to –51.4 dB after applying the DPD, while an average power-added efficiency (PAE) of 53% and an EVM of 1.2% was achieved at the average output power of 35 dBm. Thus the PA demonstrated excellent efficiency enhancement and linearity at 900 MHz. Figure 3.11(b) shows the output spectra under the unclipped 20 MHz WCDMA signal at 830 MHz, at an average output power of 32 dBm. The ACLR of the linearized PA improved to –46.5 dB. The average PAE and EVM were measured to be 47% and 2.8%, respectively. It can be inferred that the proposed 3W-DPA would be able to provide high efficiency, for a wide range of PAPR values, over a broad range

Table 3.2: Benchmarks of DPAs with extended bandwidth/BO efficiency

					CW		Modulated	
	Freq	FBW	Gain	Peak P_{out}	η (9 dB)	η (6 dB)	PAPR	PAE_{avg}
	(GHz)	(%)	(dB)	(dBm)	(%)	(%)	(dB)	(%)
[21]	2.2–2.96	29	8	$\sim$41	26–34	40–48	10.0	$\sim$25[a]
[23]	1.7–2.25	28	N/A	$\sim$49	38–46	53–65	8.0	$\sim$35[b]
[16]	0.7–1.0	35	17	50	41–45	52–68	10.5	42
[25]	2.14	N/A	$\sim$13	41	43	50	10.0	40
[26]	2.655	N/A	9	50.5	51	56	7.8	$\sim$48[c]
This Work	0.73–0.98	29	15.5	44	49–64	42–61	11.7	47
							7.1	53

[a] Based on the reported peak power at the carrier frequency and the signal PAPR
[b] Based on ACLR of –45 dB
[c] Based on ACLR of –40 dB with linearization

of carrier frequencies. Figures 3.12(a) and (b) display the static AM/AM and AM/PM responses of the 3W-DPA. A considerable difference in the AM/AM profiles can be clearly observed when compared with the CW measurements [Figure 3.10(d)]. Particularly, at high power levels, the PA's gain has significantly decreased due to the thermal effects that occur during the CW measurements. Table 3.2 compares the performance of this 3W-DPA to state-of-the-art results from the literature. As can be seen, this work outperforms all the other works in terms of the 9 dB back-off efficiency across the bandwidth and the average efficiency under high-PAPR modulated signals.

3.4   Conclusion

In this chapter, the theoretical basis for 3W-DPAs was presented. The two currently used structures, namely conventional (3W-CDPA) and modified (3W-MDPA), were reformulated and analyzed in terms of bandwidth capability. Design synthesis equations for a given power and efficiency profile were derived using the load modulation concept. It was shown that both structures are inherently wideband compared to their 2W-DPA counterparts. However, the realization of 3W-CDPAs is not practical as they require additional

hardware to control the current profiles. The modified structure, on the other hand, provides an additional knob to maximize efficiency over the bandwidth and is more feasible for fully-analog implementation. The proposed theory was validated by designing and fabricating a fully-analog 30-W 3W-MDPA demonstrator made to operate between 730 and 980 MHz (equivalent to 29% fractional bandwidth). The linearizability of the 3W-DPA was verified using digitally modulated test signals of up to 20 MHz. Even though this prototype achieved outstanding efficiency under high PAPR signals (e.g., 47% PAE with 11.7 dB PAPR signal), it was not linearizable using conventional DPD algorithms when driven with dual-band carrier-aggregated signals. This highlights the need for linearity enhanced DPAs suitable for carrier-aggregated signal applications. This challenge will be tackled in Chapter 4.

Chapter 4

Analysis of Distortion Mechanisms in DPAs

4.1   Introduction

Doherty amplifiers are traditionally designed with exclusive focus on maximizing efficiency over a wide frequency/power range. Hence, PA designers expect that DPD will compensate for all the nonlinearities caused by the high efficiency operation. This makes sense for macro or even micro base stations, where a power overhead of 1–2 W is acceptable for the DPD engine. Nevertheless, as the bandwidth of the modulated signals increases in more recent wireless standards, DPAs exhibit more nonlinearities. These excessive nonlinearites call for more sophisticated DPD algorithms and increased power overhead. In some applications such as concurrent transmission of multi-band signals, the DPA may no longer meet the linearity requirements, even after applying using a DPD. As base stations shrink in size and micro-cells are replaces by pico- and femto-cells, the power overhead of DPD is of increasing importance. Hence, it is absolutely crucial to reduce the complexity of DPD from circuit design stage, where nonlinearities are originated. This requires deep understanding of distortion mechanisms in DPAs in order to devise circuit techniques to mitigate them.

A detailed analysis of sources of nonlinearities in 2W-DPAs is presented in this chapter. Firstly, sources of quasi-static nonlinearity are studied, both at the device and circuit levels. Then dynamic nonlinearities (also known as memory effects) are explored in light of DPA operation. The ultimate goal of this study is to understand distortion mechanisms in a DPA in order to develop a strategy to design a DPA that can be linearized using a simple DPD algorithm such as a memoryless polynomial. Special care must be taken when the DPA
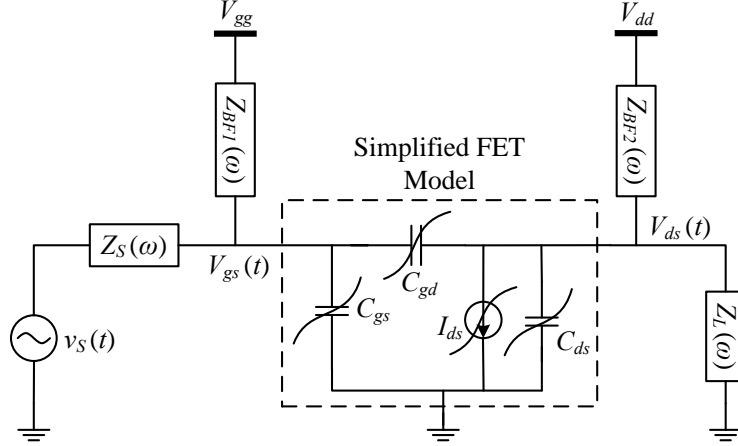
Figure 4.1: A simplified FET large-signal model.

is concurrently driven by multiple modulated signals to ensure the linearity requirements are satisfied with no significant drop in gain, output power and efficiency. The presented analysis can be extended to 3W-DPAs in future.

4.2 Quasi-Static Distortion Mechanisms in DPAs

In order to investigate the sources of distortion when a DPA is driven with CA signal stimuli, we start by analyzing the quasi-static distortion mechanisms which describe the nonlinearities under narrowband input signals, particularly single-tone (CW) excitation. The quasi-static nonlinearities are divided into device-level and circuit-level sources depending on where they originate from, and are characterized by static AM/AM and AM/PM responses.

4.2.1 Sources of Distortion at Device Level

The first step toward studying the quasi-static distortion mechanisms is to identify the sources of nonlinearity inherent to transistors, including GaN HEMT and LDMOS devices, which are the core of RF PAs.

The analysis presented is based on a simplified large-signal FET model, shown in Figure 4.1. The model includes a nonlinear current source, $I_{ds}(v_{gs}, v_{ds})$, and three nonlinear capacitors, i.e. $C_{gs}(v_{gs})$, $C_{ds}(v_{ds})$, and $C_{gd}(v_{gd})$. The die and package extrinsic elements are considered as a part of input/output matching networks. Figure 4.1 also illustrates the Thévenin equivalent model of an RF PA, where the device is terminated in $Z_S(\omega)$ and $Z_L(\omega)$, at the input and output sides, respectively. Note that the goal of this analysis is not to provide an exact analytical solution, but rather to offer insights into the distortion mechanisms by the device under large-signal regime and possible ways to mitigate them during the design. This analysis can also be considered as an initial step that needs to be further extended in the context of a Doherty amplifier.

This section particularly deals with the quasi-static nonlinearities, i.e., AM/AM and AM/PM distortions arisen under single-tone (CW) excitations, meaning that,

$$v_S(t) = V_S \cos(\omega_o t + \theta). \tag{4.1}$$

Generally speaking, the output voltage of a nonlinear circuit , can be written as

$$v(t) = \sum_{n=0}^{\infty} V_n \cos(n\omega_o t + \vartheta_n) \tag{4.2}$$

where V_n and ϑ_n are the amplitude and phase of the n-th harmonic. In the current analysis, however, unless otherwise stated, the intrinsic gate and drain terminals are assumed to be ideally short circuited at all the harmonics by proper source and load impedances. Therefore, gate and drain voltage waveforms can be expressed as the sum of quiescent (DC) and time-varying (AC) terms, i.e.,

$$V_{gs}(t) = V_{GS} + v_i(t) \tag{4.3}$$
$$V_{ds}(t) = V_{DS} + v_o(t) \tag{4.4}$$

$$v_i(t) = V_i \cos(\omega_o t + \phi_i) = \Re\left\{V_i \cdot e^{j(\omega_o t + \phi_i)}\right\} = \Re\left\{\mathbf{V_i} \cdot e^{j\omega_o t}\right\} \tag{4.5}$$
$$v_o(t) = V_o \cos(\omega_o t + \phi_o) = \Re\left\{V_o \cdot e^{j(\omega_o t + \phi_o)}\right\} = \Re\left\{\mathbf{V_o} \cdot e^{j\omega_o t}\right\} \tag{4.6}$$

where $\mathbf{V_i}$ and $\mathbf{V_o}$ are phasor representations and V_i, V_o, φ_i, and φ_o are generally functions of V_S and θ. Therefore, without loss of generality, the voltage drive waveform can be simply expressed as

$$v_S(t) = V_S \cos \omega_o t. \tag{4.7}$$

In the following sections, the nonlinearities introduced by each of the aforementioned sources will be investigated. Obviously, the overall AM/AM and AM/PM will be a nonlinear combination of these sources.

(a) Nonlinear C_{gs} profile of a 15-W GaN die.

(b) Nonlinear C_{gs} profile of a 10-W LDMOS die.
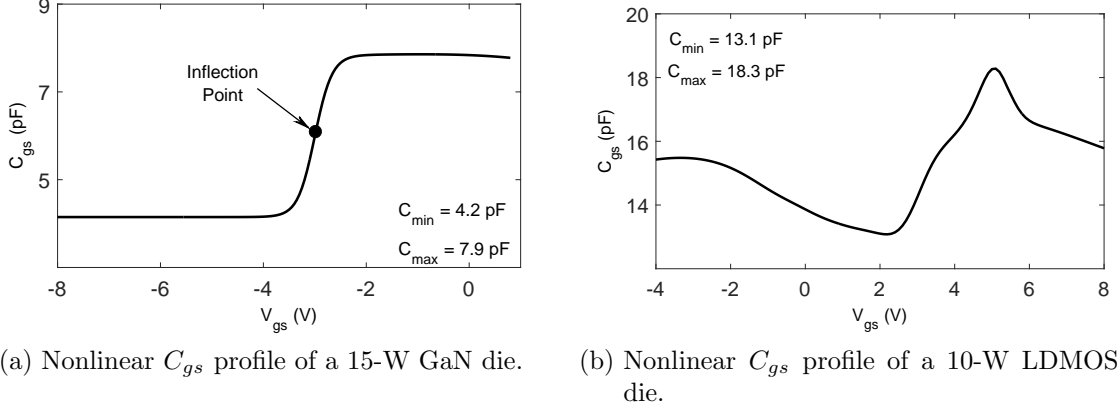
Figure 4.2: Nonlinear profile of the gate-source capacitance for GaN and LDMOS transistors.

Nonlinear C_{gs}

The profile of nonlinear gate-source capacitance of 10-W GaN and LDMOS transistors are depicted in Figure 4.2. It is typically observed that C_{gs} is a function of both v_{gs} and v_{ds}; although for the sake of simplicity here it is assumed to solely depend on V_{gs}. As can be seen, even though in both technologies C_{gs} sharply varies in the vicinity of the threshold voltage, much stronger nonlinearity is observed in case of GaN, in terms of both variation range and nonlinearity order. As will be shown, this is particularly important when designing a GaN DPA (or any III-V based PA in general), as it results in significant difference in the input impedance of class-B (main) and C (peaking) devices. For LDMOS devices, the input is usually considered to be linear for practical design.

The AC component of the gate-source nonlinear charge is defined as,

$$q_{gs}(v_i) = Q_{gs}(V_{gs}) - Q_{gs}(V_{GS}) = c_{i1}v_i + c_{i2}v_i^2 + c_{i3}v_i^3 + \cdots \tag{4.8}$$

where c_{ik} is the bias-dependent k-th order coefficient of the power series. Note that c_{i1} denotes the linear component of gate-source capacitance. Note that $q_{gs}(v_i)$ is a periodic function of time, and thus the current flowing through the capacitance can be represented by a Fourier series. Hence, the fundamental current component (at ω_o) is written as,

$$i_{gs,f_o} = \frac{dq_{gs,f_o}}{dt} = \Re\left\{\mathbf{I_{gsf_o}} \cdot e^{j\omega_o t}\right\} \tag{4.9}$$

where $\mathbf{I_{gs}}$ is the phasor of the fundamental current, which is

$$\begin{aligned}
\mathbf{I_{gs,f_o}} &= j\omega_o\mathbf{V_i}\left(c_{i1} + \frac{3}{4}c_{i3}V_i^2\right) \\
&\triangleq j\omega_o\mathbf{V_i}C_{gs,f_o}
\end{aligned} \tag{4.10}$$

C_{gs,f_o} is called the effective nonlinear capacitance at fundamental frequency, or simply the fundamental capacitance. The AM/AM and AM/PM distortions due to the nonlinear C_{gs} can be determined by writing the input KCL at ω_o and using (4.10) (refer to Appendix A for detailed derivation), which are:

$$|\mathbf{V_i}|^2 = V_S^2 \cdot \frac{G_S^2(\omega_o) + B_S^2(\omega_o)}{G_S^2(\omega_o) + [B_S(\omega_o) + \omega_o(c_{i1} + 3/4c_{i3}V_i^2)]^2} \qquad \text{(AM/AM)} \tag{4.11}$$

$$\angle\mathbf{V_i} = -\tan^{-1}\left[\frac{B_S(\omega_o) + \omega_o(c_{i1} + 3/4c_{i3}V_i^2)}{G_S(\omega_o)}\right] \qquad \text{(AM/PM)}. \tag{4.12}$$

Even though (4.11) and (4.12) are recursive and not in closed form, they clearly show the nonlinear dependence of amplitude and phase of $\mathbf{V_i}$ upon V_S. Moreover, according to Figure 4.2(a), C_{gs} of a GaN device follows a nonlinear function of the form $\tanh(\cdot)$, which has odd symmetry around the inflection point, that corresponds to a deep class-AB biasing point. This means that around class-B or deep AB biasing point, all the odd-order coefficients of (4.8) other than c_{i1} are equal to zero. Therefore, it can be inferred from (4.11) and (4.12) that no AM/AM and AM/PM is induced. In fact, C_{gs} of a GaN HEMT acts as a perfectly linear capacitor when it is biased at the inflection point. In practice, the input impedance may still be slightly nonlinear due to the Miller effect of gate-drain capacitance, as discussed later. Note that this conclusion is based on the assumption that all of the harmonics are shorted out at the intrinsic gate terminal. In fact, it has been shown in the literature that GaN devices are extremely sensitive to the second harmonic source impedance in particular, as second harmonic contents cause asymmetry in the voltage swing around the quiescent point (the inflection point in this case), and thus the capacitance nonlinearity becomes important. Table 4.1 summarizes the sign of AM/AM and AM/PM caused by nonlinear C_{gs} in GaN HEMTs, as the bias point varies from class AB to C.

Nonlinear C_{gd} and Miller Effect

The profiles of the feedback capacitance, C_{gd}, versus V_{gd} for 10-W GaN and LDMOS transistors are shown in Figure 4.3. As can be seen, in the case of LDMOS nonlinearity is only pronounced in the knee region, and also the capacitance value is much smaller than

Table 4.1: AM/AM and AM/PM distortions due to the nonlinear C_{gs}

Class	c_{i3}	AM/AM	AM/PM
AB	–	Expansion	Expansion
B	≈ 0	NO	NO
C	+	Compression	Compression
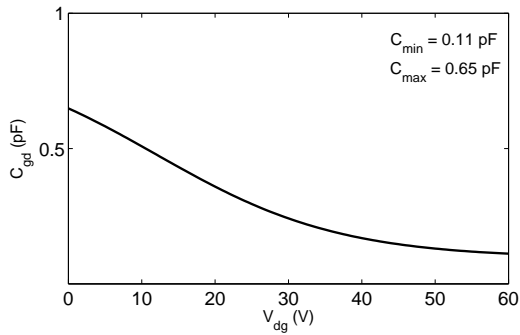
(a) Nonlinear C_{gd} profile of a 15-W GaN die.
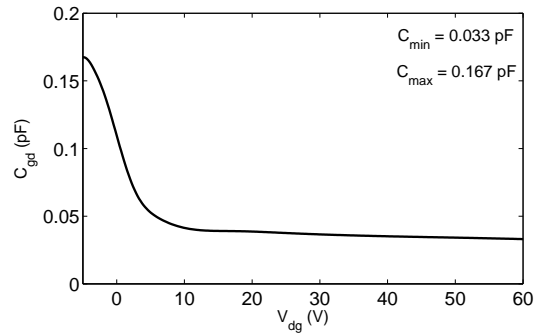
(b) Nonlinear C_{gd} profile of a 10-W LDMOS die.

Figure 4.3: Nonlinear C-V profile of the gate-drain capacitance for for GaN and LDMOS transistors.

C_{gs} and C_{ds} and thus its nonlinearity can be neglected. For GaN HEMT, nevertheless, C_{gd} capacitance shows a strong nonlinearity, which plays an important role, specially in the input AM/AM and AM/PM nonlinearity.

A similar analysis to the one presented for nonlinear C_{gs} can be conducted to evaluate the effect of feedback capacitance on AM/AM and AM/PM characteristics. If the nonlinear gate-drain charge is expressed with power series as

$$q_{gd}(v_f) = Q_{gd}(V_{gd}) - Q_{gd}(V_{GD}) = c_{f1}v_f + c_{f2}v_f^2 + c_{f3}v_f^3 + \cdots \qquad (4.13)$$

where $v_f = v_i - v_o$, then the overall fundamental input capacitance can be calculated as (see Appendix A for detailed derivation)

$$C_{in,f_o} = C_{gs,f_o} + (1 + A_v)C_{gs,f_o}$$
$$\approx \left(c_{i1} + \frac{3}{4}c_{i3}V_i^2 \right) + (1 + A_V) \left(c_{f1} + \frac{3}{4}c_{f3}V_o^2 \right) \qquad (4.14)$$

68

where $A_V = -V_o/V_i$ is the gain of the amplifier. It can be seen the effect of nonlinear C_{gd} is multiplied by a factor of $1 + A_V$, which is referred to as Miller effect. In fact, the Miller capacitance is often the dominant nonlinearity source at the input, according to the value of C_{gd} (Figure 4.3) as compared to C_{gs} (Figure 4.2) and $A_V \approx 10$.

To determine the AM/AM and AM/PM resulted from the nonlinear input capacitance, C_{in,f_o} defined in (4.14) must replace C_{gs,f_o} in (4.11) and (4.12), i.e.,

$$V_i^2 = V_S^2 \cdot \frac{G_S^2(\omega_o) + B_S^2(\omega_o)}{G_S^2(\omega_o) + [B_S(\omega_o) + \omega_o C_{in,f_o}]^2} \tag{4.15}$$

$$\phi_i = -\tan^{-1}\left[\frac{B_S(\omega_o) + \omega_o C_{in,f_o}}{G_S(\omega_o)}\right]. \tag{4.16}$$

Note that $c_{f3} > 0$ for both technologies; hence, it can be inferred that nonlinear C_{gd} introduces gain and phase compression irrespective of the biasing point. It should be noted in the current analysis it has been assumed that A_V is a positive constant. Obviously, any (amplitude and phase) nonlinearity in the gain will be mixed with the C_{gd} nonlinearity which further complicates the analysis. In addition, a phase shift other than $180°$ will change the phase of I_{gd,f_o} and generates a resistive Miller component.

Nonlinear Transconductance

Nonlinear drain current source has been known as the main nonlinearity source, as it defines the basis of the transistor operation. In this section, only the nonlinearity of the current source with respect to variation in V_{gs}, i.e., nonlinearity in the saturation region, will be studied. It means that the drain current is assumed to be solely dependent upon V_{gs}, i.e., no channel length modulation exists. The V_{ds} dependence arising from operation in triode region will be discussed in the following section.

As discussed in Chapter 2, the small-signal transconductance can deviate from the ideal characteristics quite significantly, due to the gradual turn-on and saturation characteristics. The AC component of drain-source current is defined based on power series expansion as

$$i_d(v_i) = I_{ds}(V_{gs}) - I_{DS} = g_{m1}v_i + g_{m2}v_i^2 + g_{m3}v_i^3 + \cdots \tag{4.17}$$

where $I_{DS} = I_{ds}(V_{GS})$ is the quiescent drain current. By substituting $v_i = V_i \cos(\omega_o t)$ in (4.17), one can determine the large-signal transconductance as

$$G_m = \frac{\mathbf{I_{ds,f_o}}}{\mathbf{V_i}} = g_{m1} + \frac{3}{4}g_{m3}V_i^2 + \frac{5}{8}g_{m5}V_i^4 + \cdots . \tag{4.18}$$
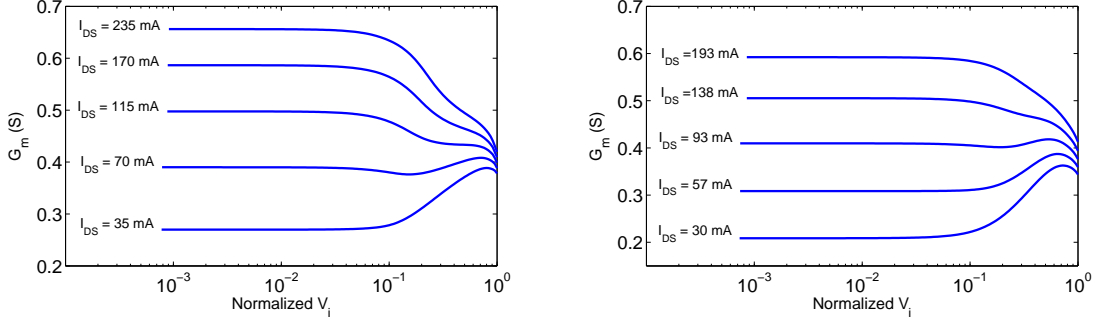
(a) Large-signal short-circuit G_m profile of a 15-W GaN die.

(b) Large-signal short-circuit G_m profile of a 10-W LDMOS die.

Figure 4.4: Large-signal transconductance profile of GaN and LDMOS transistors.

Note that for a class A/B biased device, $g_{m3} < 0$, whereas for a class-C biased device $g_{m3} > 0$. Therefore, at least 5 coefficients (up to g_{m5}) has to be considered for a class-C biassed transistor to properly model the initial gain expansion and the final gain compression versus input voltage ($g_{m5} < 0$). Figure 4.4 illustrates the large-signal G_m characteristics of 10-W GaN and LDMOS devices vs. gate voltage swing, V_i, for various biasing points. As can be seen, LDMOS device exhibits a slightly flatter (and thus more linear) characteristic as compared to GaN HEMT. Figure 4.4 Also shows that proper choice of gate bias point plays an important in the G_m nonlinearity. These plots can also be used in particular designs to cancel out the nonlinearities caused by other sources, e.g., the input capacitance.

Nonlinear Output Capacitance and Knee Effect

Nonlinear drain-source current is generally a mixed function of both V_{gs} and V_{ds}, although in most of the simplified models it is separated as

$$I_{ds}(V_{gs}, V_{ds}) = F(V_{gs}) \cdot G(V_{ds}) \tag{4.19}$$

where $F(\cdot)$ can be represented by a polynomial function, and

$$G(V_{ds}) = \tanh(\alpha V_{ds}) \cdot (1 + \lambda V_{ds}) \tag{4.20}$$

where α and λ are model parameters that determine the knee voltage and the output conductance in the saturation region, respectively. Unfortunately, this representation of drain
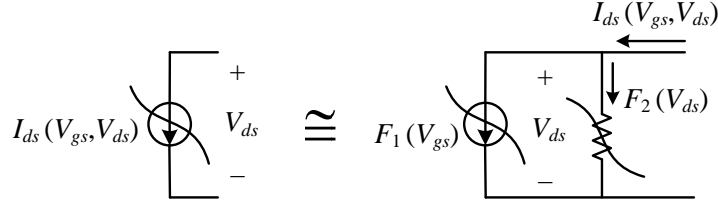
Figure 4.5: Drain current source can be approximated by a gate voltage-controlled current source in parallel with a nonlinear conductance to model the knee effect.

current significantly complicates the analysis. Alternatively, for the purpose of simplicity, one can approximate the drain current as

$$I_{ds}(V_{gs}, V_{ds}) \approx F_1(V_{gs}) + F_2(V_{ds}). \tag{4.21}$$

This is equivalent to connecting a nonlinear conductance, dF_2/dV_{ds}, in parallel with the nonlinear VCCS, $F_1(V_{gs})$, as shown in Figure 4.5. As the transistor intrudes into the knee region, the nonlinear conductance and current through it start to increase from zero, i.e.,

$$i_o = F_2(V_{ds}) - F_2(V_{DS}) = g_{o1}v_o + g_{o2}v_o^2 + g_{o3}v_o^3 + \cdots . \tag{4.22}$$

Note that $F_2(V_{DS}) \approx 0$, as the output impedance in the saturation region is assumed to be infinite. The large-signal output conductance can be determined as

$$G_o = g_{o1} + \frac{3}{4}g_{o3}V_o^2 + \cdots . \tag{4.23}$$

Nonlinearity in the output can also be caused by the nonlinear output capacitance, i.e., $C_{out} \approx C_{ds} + C_{gd}$, neglecting the feed-forward effect of C_{gd}. Note that C_{out} represents the effect of output capacitance on the AM/AM and AM/PM and not the equivalent output capacitance seen looking back into the device output terminal. The gate-drain capacitance has been discussed earlier. The only difference with the input capacitance is that Miller effect does not exist for the output assuming the voltage gain is sufficiently large. The drain-source capacitance of a GaN HEMT is considered to be linear in most of the models, while it is strongly nonlinear in case of LDMOS devices. Even though it is generally a function of both V_{ds} and V_{gs} especially in the knee region, here it is assumed to be solely dependant on V_{ds} for simplicity. As can be seen in Figure 4.6, the overall output capacitance is nonlinear for both technologies, however, since C_{ds} is the dominant portion, the GaN

(a) Nonlinear C_{out} profile of a 15-W GaN die.

(b) Nonlinear C_{out} profile of a 10-W LDMOS die.
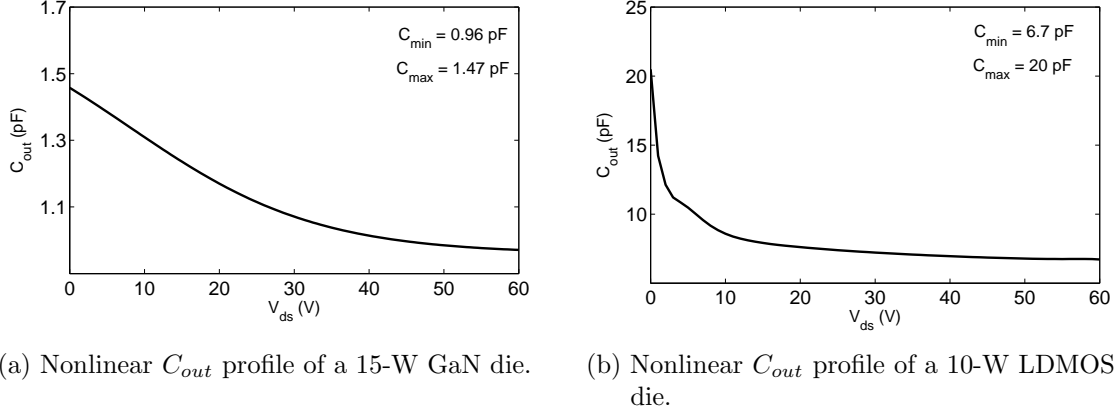
Figure 4.6: Nonlinear output capacitance profiles of GaN and LDMOS transistors.

output capacitance exhibits a mild (and often negligible) nonlinearity profile. For the same reason, C_{out} of LDMOS device can be considered a function of V_{ds} only, i.e.,

$$q_{out}(v_o) = Q_{out}(V_{ds}) - Q_{out}(V_{DS}) = c_{o1}v_o + c_{o2}v_o^2 + c_{o3}v_o^3 + \cdots . \qquad (4.24)$$

Assuming a single-tone sinusoid at the output represented by (4.4), The nonlinear fundamental output capacitance is defined as,

$$C_{out,f_o} = c_{o1} + \frac{3}{4}c_{o3}V_o^2 + \cdots . \qquad (4.25)$$

Writing KCL the output node at ω_o using one can express the nonlinear voltage transfer function as

$$\frac{\mathbf{V_o}}{\mathbf{V_i}} = -\frac{G_m}{[G_o + G_L(\omega_o)] + j[\omega_o C_{out,f_o} + B_L(\omega_o)]} \qquad (4.26)$$

where $Y_L(\omega_o) = G_L(\omega_o) + jB_L(\omega_o)$ is the load admittance seen by the transistor. Therefore, The resultant AM/AM and AM/PM are given by

$$|V_o|^2 = V_i^2 \cdot \frac{G_m^2}{[G_o + G_L]^2 + [\omega_o(c_{o1} + 3/4c_{o3}V_o^2) + B_L]^2} \qquad \text{(Output AM/AM)} \qquad (4.27)$$

$$\Delta\phi = \phi_o - \phi_i = \tan^{-1}\left[\frac{B_L + \omega_o(c_{o1} + 3/4c_{o3}V_o^2)}{G_L + \omega_o(g_{o1} + 3/4g_{o3}V_o^2)}\right] \qquad \text{(Output AM/PM)}. \qquad (4.28)$$

Hence, it can be concluded that both the knee region and output capacitance nonlinearities can contribute to output AM/AM and AM/PM distortions. It is especially interesting that

72
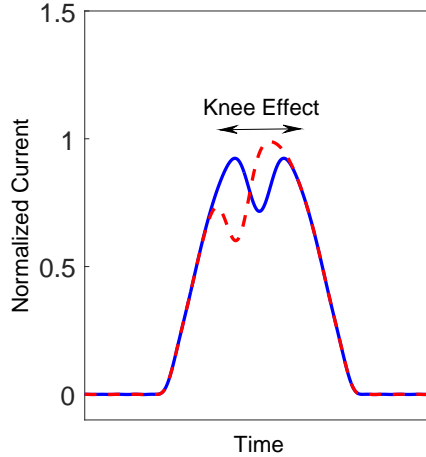
Figure 4.7: Knee region intrusion when the transistor is terminated with a pure resistive load (solid) and an inductive load (dashed).

the knee effect can introduce phase distortion even if the output capacitance is linear. This occurs when the output capacitance is not fully cancelled out by the load susceptance. Consequently, the drain current waveform will be asymmetrically distorted as the drain voltage enters the knee region, as shown in Figure 4.7. This can also occur when the harmonics are not properly short circuited. Both cases are very likely to happen in class-J [44, 45], class-F [46], and also broadband PAs where perfect compensation for output capacitance over the bandwidth is practically impossible. Furthermore, it can be inferred from Figure 4.6 that C_{out,f_o} is a monotonically increasing function of V_o, and the same holds for G_o. Hence, even though the overall AM/PM is alleviated as a result, the gain compression versus input drive is aggravated. Note that the direction of phase distortion(compression/expansion) due to the knee effect depends on the fundamental or harmonic reactance (inductive/capacitive), which affects the phase shift between the voltage and current waveforms. For instance, for the asymmetrically-distorted waveform shown in Figure 4.7, output phase will decrease with voltage drive (phase compression).

It should be noted that even though the knee effect and output capacitance nonlinearities arise close to peak power regardless of the PA configuration, in case of DPA they can appear at back-off levels as high as 6–12 dB. This may considerably affect the linearity and linearizability of a DPA, as it cannot be linearized by simply backing-off the input power.
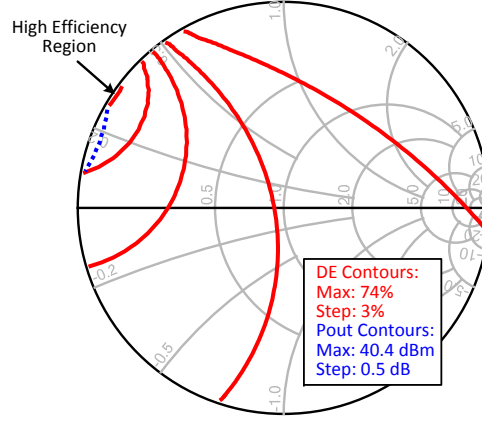
Figure 4.8: Second harmonic source-pull contours of a class-B biased GaN HEMT at 0.8 GHz.

4.2.2 Input Second Harmonic Termination

As mentioned in Section 4.2.1, GaN-based PAs exhibit strong input nonlinearity attributed to the nonlinear input capacitance. This means that the instantaneous gate-source voltage not only depends on the source impedance at the fundamental, $Z_S(f_o)$, as well as harmonics, $Z_S(2f_o), Z_S(3f_o)$, etc. Particularly, as will be shown, the second harmonic impedance can have a great impact on the performance. Note that in the analysis presented in Section 4.2.1, harmonics were assumed ideally shorted. This section studies the effects of second harmonic impedance for class-B and C biased PAs used in a DPA. In order to do that, one needs to set the fundamental source impedance first and then sweep the second harmonic impedance across the Smith chart and plot constant power and efficiency contours (also known as source-pull contours). Fortunately, simulations confirm that choice of fundamental impedance does not impact on the second harmonic source-pull contours, and only affects the gain (regardless of the gate bias voltage). The output is set to the optimum load impedance, i.e. $R_{opt} = 30\,\Omega$ for all cases.

Figure 4.8 shows the second harmonic source-pull contours of a class-B biased 15-W GaN HEMT die at 0.8 GHz and peak power. The fundamental impedance is $Z_S(f_o) = 30\,\Omega$, corresponding to 18.7 dB gain. As can be seen, efficiency, unlike output power, greatly depends on the second harmonic impedance, to the extent that peak efficiency may degrade as much as 10% by poor choice of second harmonic impedance. Efficiency and gain profiles versus output power are depicted in Figure 4.9 for a good and a poor choice of second

74

(a) Drain efficiency profiles.   (b) Gain profiles.
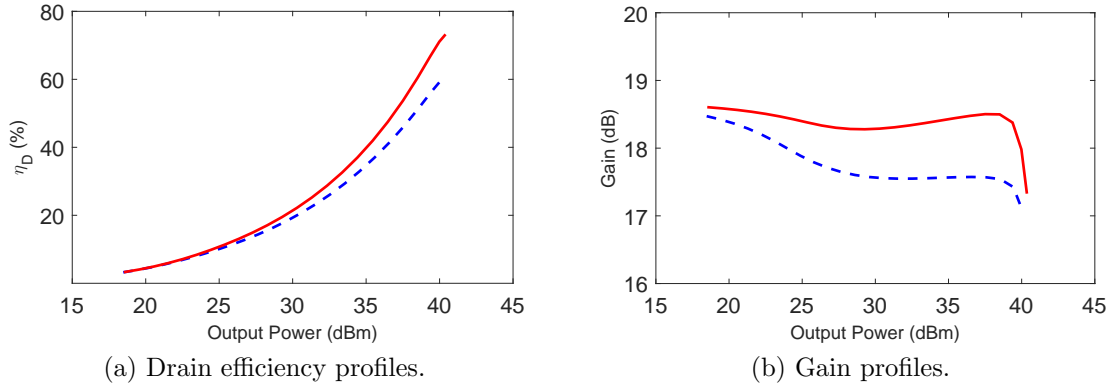
Figure 4.9: Efficiency and gain profiles of a class B GaN PA with a good (solid) and a poor (dashed) input second harmonic termination.
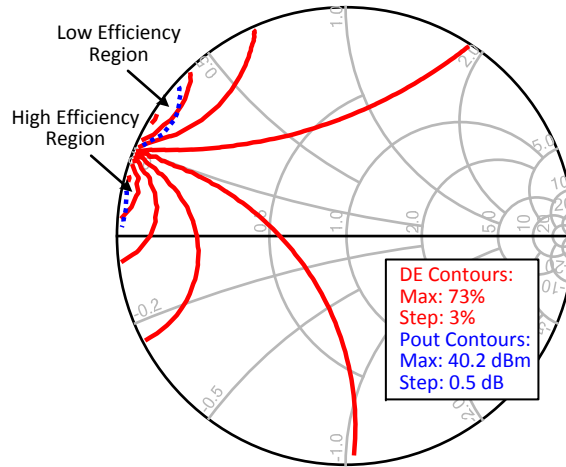
Figure 4.10: Second harmonic source-pull contours of a class-B biased GaN HEMT at 2 GHz and peak power.

harmonic impedance. It is observed that the second harmonic impedance not only affects efficiency, but it also degrades linearity by aggravating the slow compression in GaN PAs. It is worth mentioning that the small signal gain is the same for both cases since nonlinearity is negligible at low power levels.

To investigate this effect versus frequency, the same source-pull simulation was con-

(a) Gate voltage waveforms.

(b) Drain current waveforms.
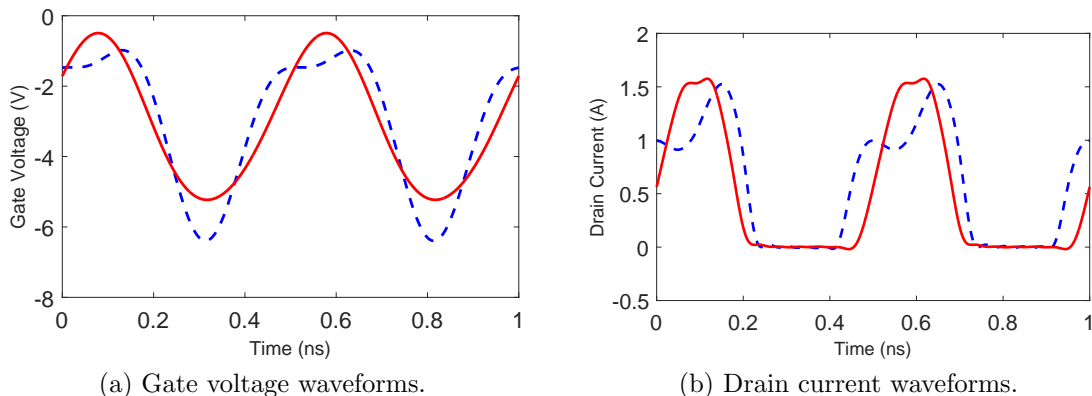
Figure 4.11: Waveforms of a class-B GaN PA with a good (solid) and a poor (dashed) input second harmonic termination.

ducted at 2 GHz, as shown in Figure 4.10. The fundamental impedance is $Z_S(f_o) = 30\,\Omega$, corresponding to 12.4 dB. Comparing with Figure 4.8, it can be seen that the second harmonic sensitivity of a class-B biased GaN PA increases with frequency. Hence, designing the input matching network for a broadband class B (or AB) PA in III-V technologies is more challenging at higher frequencies. It is of note that increasing the device size will have the same effect as operation at higher frequencies, since the admittance of the gate-source capacitance, ωC_{gs}, is linearly proportional to the device periphery as well the frequency.

Figure 4.11(a) illustrates the gate voltage waveform for the two mentioned cases. One can clearly see the unwanted second harmonic contents on the blue waveforms, attributed to the poor second harmonic termination at this frequency. Looking at Figure 4.11 more closely reveals the reason for significant degradation of efficiency when second harmonic is presented in the waveform. Comparing the red (solid) and blue (dashed) waveforms, one can see the increased conduction angle due to the asymmetrical gate voltage waveform. This fact is shown more clearly in Figure 4.11(b), the drain current waveforms for the two cases. In fact, the class of operation has shifted toward class A for the blue case, due to the poor second harmonic termination. Hence, the input second harmonic effects can be mitigated to some extent by lowering the gate bias voltage.

This analysis can be repeated for a class-C biased PA, representing the peaking amplifier in a Doherty configuration. Similar to a class-B PA, second harmonic contours are independent of the selected fundamental impedance at a given frequency. Figure 4.12 displays the source-pull contours of a 30-W GaN die at 0.8 GHz and 27 dBm input power with $Z_S(f_o) = 30 - j14\,\Omega$ and optimum load impedance, $R_{opt} = 30\,\Omega$. As can be seen,

76
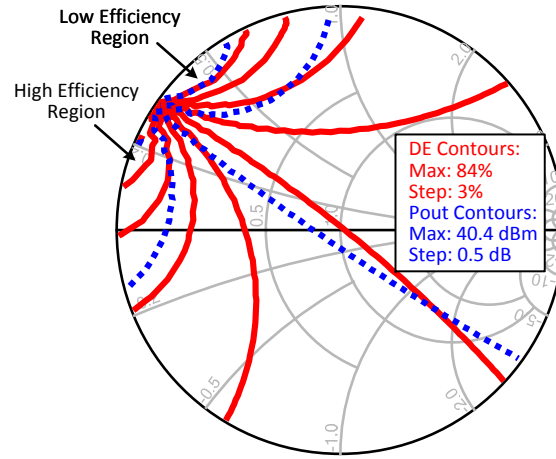
Figure 4.12: Second harmonic source-pull contours of a class-C biased GaN HEMT at 0.8 GHz and PEP.

(a) Drain efficiency profiles.

(b) Drain current profiles.
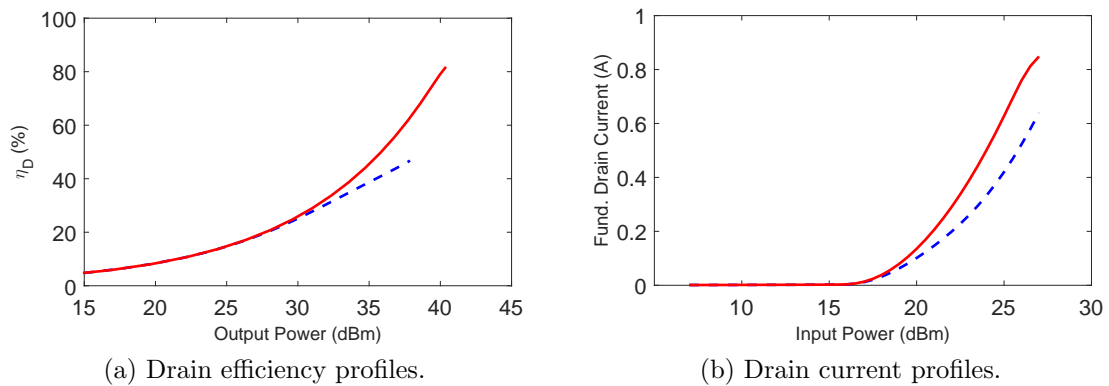
Figure 4.13: Efficiency and current profiles of a class C GaN PA with a good (solid) and a poor (dashed) input second harmonic termination.

unlike a class-B PA, not only efficiency, but also output power highly depend on the second harmonic input impedance.

Figure 4.13 shows fundamental drain current as well as efficiency versus input and output power, respectively, for a good and a poor second harmonic termination. Note that in a Doherty configuration, the fundamental drain profile is a critical parameter, as it defines the load modulation and as can be seen greatly depends on the input second
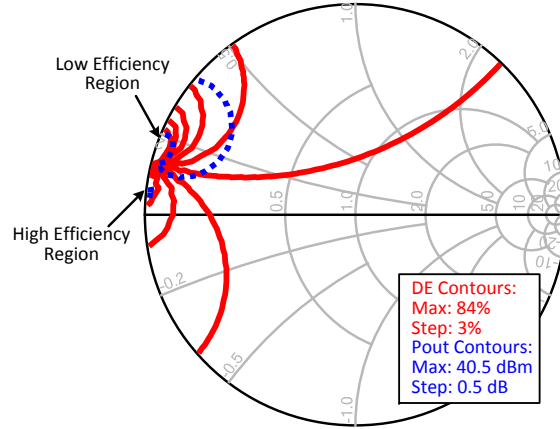
Figure 4.14: Second harmonic source-pull contours of a class-C biased GaN HEMT at 2 GHz and PEP.

harmonic impedance. This highlights the role of the input second harmonic termination for the peaking PA in the overall performance of the DPA.

Figure 4.14 depicts the source-pull contours of the same transistor at 2 GHz with $Z_S(f_o) = 12 + j6\,\Omega$. Comparing these contours with those of Figure 4.12, one can infer that unlike a class B PA, the sensitivity to the input second harmonic termination decreases with frequency.

Note that the output load was assumed constant and equal to R_{opt} throughout this analysis. Ideally, changing the output load should not affect the source-pull contours. However, due to the device nonidealities (particularly Miller capacitance), the transistor is not unilateral. Hence, the effect of the output termination may be non-negligible especially at higher frequencies (e.g., above 3 GHz). This subject is discussed in more detail in Chapter 6.

In short, the input second harmonic termination plays an important role in the performance of GaN PAs. Hence, conducting the second harmonic source-pull simulation/measurement is a crucial step in designing input matching networks, particularly in GaN DPAs. The input matching design challenges are further explored in Chapter 6.

4.2.3 Sources of Distortion at Circuit Level

As mentioned in Chapter 3, the linearity of the 2W-DPA at the center frequency is solely determined by the main amplifier at center frequency. This is because the voltage across the load is dictated by the main current flowing through the $\lambda/4$ line. The function of the peaking PA is to keep the main PA away from the knee region by providing proper current (both in magnitude and phase). This section studies the distortion mechanisms caused by nonidealities in Doherty amplifier configuration.

Peaking Device Current

As mentioned before, the peaking transistor plays a critical role in the linearity of the DPA in the medium to high power levels by modulating the effective load impedance of the main device. Otherwise, the main device will enter the knee region and gain linearity is significantly degraded as a result. There are several factors affecting the magnitude of the peaking current: gate bias selection, size, and nonlinear input capacitance of the peaking device. The mutual effects of the peaking size and gate bias was already discussed in Section 2.5.3. Due to the soft turn-on effect of the class-C biased device, a trade-off has to be made between AM/AM response (linearity) and back-off efficiency.

According to the analysis conducted in Section 4.2.1, as a result of nonlinear input capacitance, the fundamental gate voltage of the peaking transistor is a mixed function of the drive voltage and frequency, i.e., $H(\omega, V_s)$. This in turn renders the turn-on point of the peaking transistor a strong function of frequency, and thus is a major cause of nonlinearity in GaN-based broadband DPAs. Figures 4.15(a) and (b) illustrate the plots of drain efficiency and third-order intermodulation distortion (IMD3) of a GaN HEMT current source under a two-tone stimulus, as the normalized gate bias voltage of the peaking device varies between -0.45 and -0.55. As can be seen, 8–10 dB improvement of IMD3 can be gained in IMD3 by sacrificing few percent of efficiency.

Phase Alignment

Phase imbalance between the the main and peaking paths reduces the combining efficiency and introduces distortions close to peak power. From a different point of view, considerable phase difference between the two paths hinders proper load modulation of the main amplifier due to a decreased effective current from the peaking device. The admittance
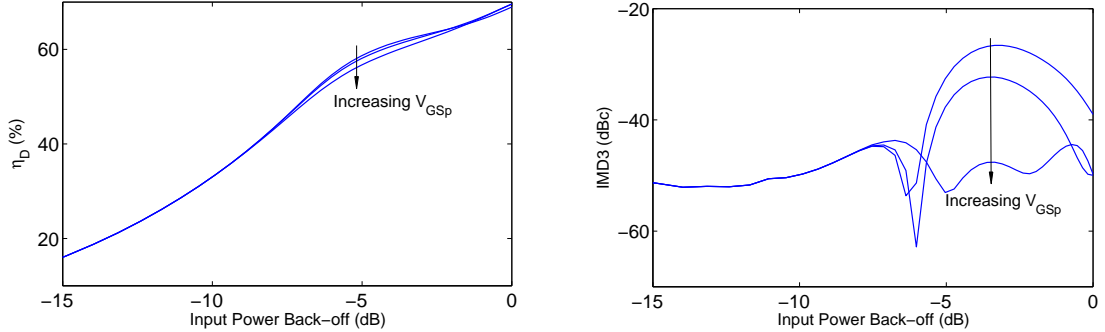
(a) Drain efficiency degradation vs. input power back-off for different peaking biasing points.

(b) IMD3 vs. input power back-off for different peaking biasing points.

Figure 4.15: Two-tone performance degradation vs. peaking gate biasing.

seen by the main transistor as a function of peaking current is given by

$$Y_M = \frac{R_L}{Z_o^2} \left[1 + Z_o \frac{I_p \angle \Delta \varphi_p}{V_m} \right] \tag{4.29}$$

where $\Delta \varphi_p$ represents the phase misalignment. It is observed that an imaginary part is introduced to Y_M (or to the impedance Z_M) due to the phase misalignment between the main and peaking currents, and this generates both AM/AM and AM/PM distortion. The major origin of this phase imbalance is the AM/PM generated by device-level sources, such as nonlinear input/output capacitance. Different phase responses attributed to different input/output capacitance of the main (biased in class-AB) and peaking (biased in class-C) amplifiers prevents in-phase combination of the output currents (or equivalently output powers) over the upper 6-dB power range. The problem is exacerbated for broadband DPAs, as the AM/PM response would also be a function of frequency, which prevents independent alignment at different frequencies.

Figure 4.16 depicts the real part of Z_M as the phase imbalance varies from 0 to 40°. Obviously, negative $\Delta \varphi_p$ values will result in similar results, but the sign of AM/PM will change. Figures 4.17(a) and (b) show the added AM/AM and AM/PM attributed to the phase imbalance. As can be seen, the real part of the main impedance is still modulated with no significant deterioration. Figure 4.18 displays the drain efficiency and IMD3 at peak envelope power (PEP) versus $\Delta \varphi_p$ under a two-tone excitation with 10-MHz spacing centered at 1 GHz. It can be seen that up to $\pm 15° \sim 20°$ phase imbalance has insignificant impact on IMD3 and efficiency.

80
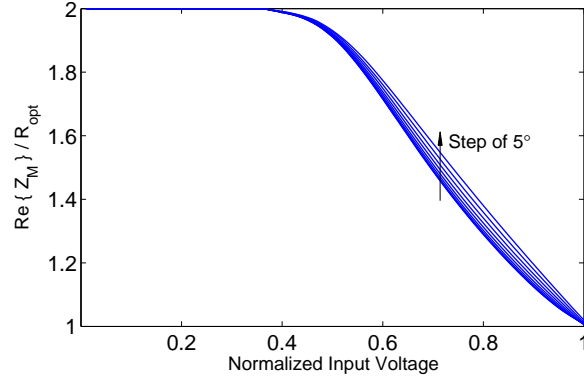
Figure 4.16: Load modulation curves for different $\Delta\varphi_p$ values.
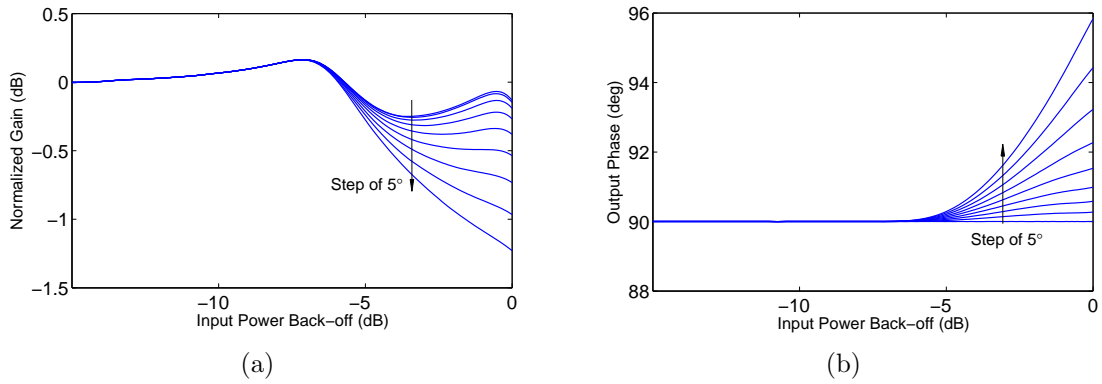
Figure 4.17: AM/AM and AM/PM caused by misalignment between the main and peaking paths.

Impedance Inverter's Dispersion

As the carrier frequency shifts from the design frequency, the $\lambda/4$ line will act as an imperfect impedance inverter, which presents a different impedance to the main transistor, causing both efficiency and linearity degradation. Note that this is completely different from the case where the two paths are not aligned (discussed earlier), as this affects the performance of the DPA not only at high power level, but also at BO power. Assuming the electrical length of the inverter is equal to $\theta = \pi/2 + \Delta\theta$ radians at a given frequency f, The voltage across the load can be described by

$$V_L = V_m \cos\theta - jZ_o I_m \sin\theta. \tag{4.30}$$

81

(a) Two-tone Drain efficiency vs. $\Delta\varphi_p$ at PEP.　　(b) IMD3 vs. $\Delta\varphi_p$ at PEP.
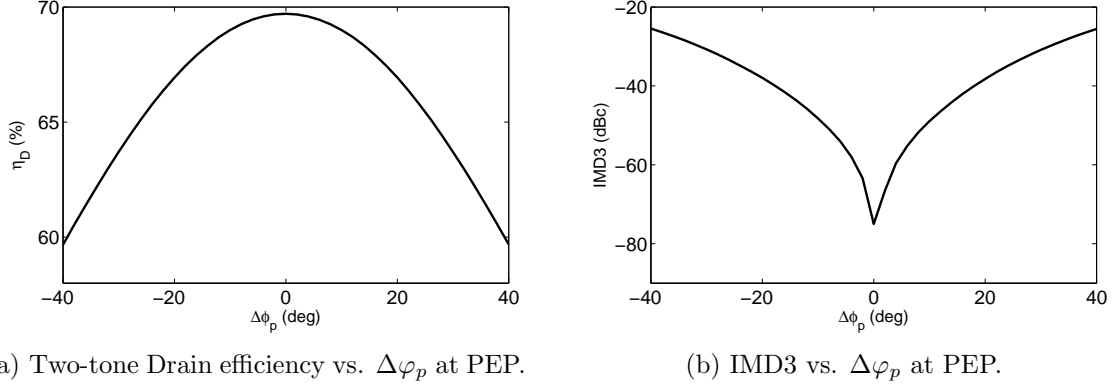
Figure 4.18: Performance degradation due to the misalignment between the main and peaking paths obtained from a two-tone simulation at PEP.

It is observed that frequency dispersion of the impedance inverter line affects both AM/AM and AM/PM of a DPA. In fact, it can be shown that if the frequency shift is small, then the low power gain of the DPA will be reduced by $1 - 3\Delta\theta^2$ and the high power phase will change by $\pm\Delta\theta$ (refer to Appendix B for detailed proof). Note that these approximations are based on the assumption that the two paths are perfectly aligned and the impedance inverter is only frequency dispersive. For instance, at 10% above/below the center frequency, a phase distortion of $\sim \pm 9°$ and gain expansion of ~ 0.3 dB are recorded. One can clearly conclude that the resultant AM/PM is much more significant. Furthermore, its direction (phase expansion vs. compression) depends on the frequency of operation with respect to the center frequency, and thus cannot be compensated for by using existing analog predistorion techniques. It also substantially complicates the DPD algorithm for wideband/ concurrent multi-band amplification, since the DPA's group delay greatly varies with frequency. Figure 4.19 shows the AM/AM and AM/PM responses for the example mentioned above, which confirms the approximation used. A solution to alleviate the frequency dispersion effect is to use the asymmetric drain biasing technique. It can be shown that the AM/PM in that case will reduce to $\Delta\theta/\sigma$. This is another advantage of the asymmetrical biasing technique.
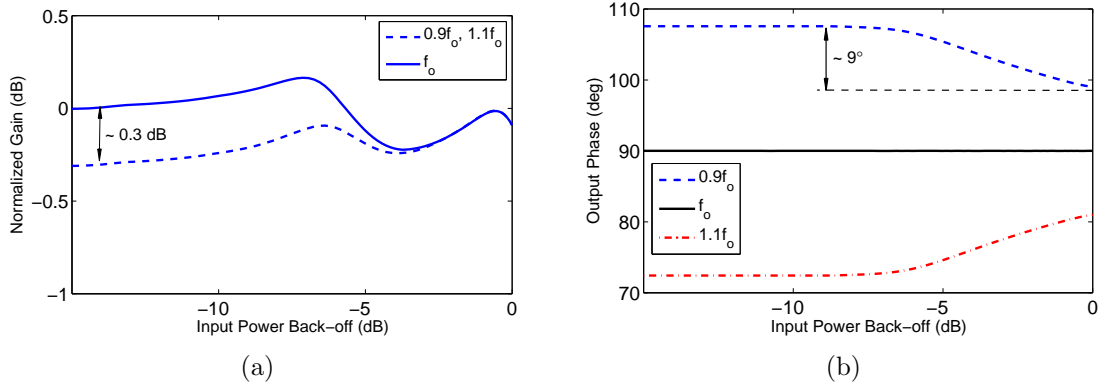
Figure 4.19: AM/AM and AM/PM caused by frequency dispersion of the quarter-wave impedance inverter.
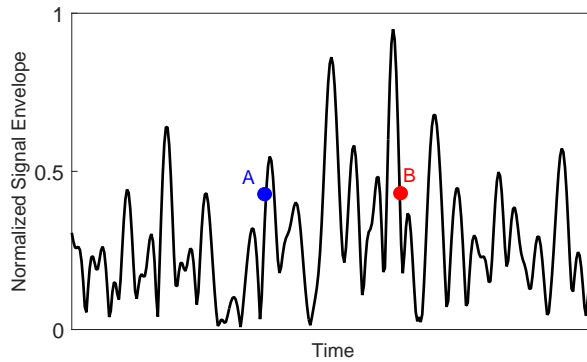
Figure 4.20: Envelope waveform of a typical modulated signal.

4.3 Dynamic Distortion Mechanisms in DPAs

The sources of nonlinearity discussed in previous sections deal with the case where the PA is driven by CW (single sinusoidal) stimuli. When the PA is excited by a modulated signal having a non-constant envelope (as is the case for modern communication signals used in 3G and 4G standards) the behavior of the PA cannot be described by static AM/AM and AM/PM characteristics. In these cases, other than the quasi-static sources of nonlinearity, dynamic sources also known as memory effects come into play as well. In other words, when a PA is driven with a modulated signal, the PA's response at any envelope instant not only depends on the envelope at that instant, but also the history of the envelope.

Figure 4.20 illustrates this effect more clearly. As can be seen, the amplitude at instants A and B is similar, however the PA's response at the points are different due to the memory effects. This is often observed as a spread around the AM/AM and AM/PM responses, as will be shown later.

Dynamic sources of distortion can be categorized into short-term and long-term memory effects, depending on their time-scale relative to the signal modulation bandwidth. Major sources of memory effects in PAs include:

- thermal and trapping effects

- magnitude/phase dispersion, and

- drain induced memory effects.

Note that these sources are common among all PA structures and are not specific to DPAs. However, special attention must be paid in case of DPAs as the main and peaking transistors operate interactively under modulated signal excitations. The three sources will be discussed further in the following sections.

4.3.1   Thermal and Trapping Effects

When a PA is driven by modulated signals, at some instants when the amplitude is large, the transistor channel heats up. This temperature change results in dynamic variation of some of the device parameters such as gain. In other words, in a short time later, the input signal amplitude reduces, whereas the channel has not cooled down as quickly, meaning that the PA's response is affected by the signal's history [2]. Thermal memory effects become significant when the thermal time constants (of the order of milliseconds to microseconds) become comparable with the timescale of the envelope and can be minimized by proper thermal management.

Charge trapping is another source of long-term memory effect which is attributed to the imperfections and defects in the device channel [2]. These imperfections result in trapping and release of carriers which effectively change the charge density in the channel and hence the gain of the PA over time. This mechanism highly depends on the device technology and fabrication process and its rate is of the order of kilohertz to megahertz. It is usually negligible in LDMOS transistors; however, III-V based devices such as GaN HEMTs exhibit significant trapping effects. In this work, the minimum time-constant associated with the trapping is assumed to be much larger than the timescale of the signal envelope, so the resultant memory effects would be negligible.

4.3.2   Magnitude/Phase Dispersion

Magnitude/phase Dispersion is a short-term memory effect and is not specific to nonlinear circuits such as PAs. Group delay dispersion in linear circuits can be seen in filters in stop band or passband edges where the phase and magnitude responses are no longer linear and constant, respectively, with frequency. Therefore, when passing a band-limited signal such as a pulse through the filter, signal components at different frequencies will experience dissimilar time delays and/or attenuation/amplification constants, hence the pulse shape will be distorted. Obviously, in linear circuits this effect will not manifest under single-tone stimuli, thus it is a dynamic phenomenon.

In PAs, input/output matching networks combined with transistor parasitics can be considered as filters, having a magnitude and phase response which can contribute to amplitude/phase dispersion. This fact highlights the importance of matching network design with a sufficiently flat gain and linear phase (constant group delay) responses. This effect in DPAs is similar to single-ended PAs, hence the IMNs of the main and peaking PAs as well as the combining network need to be carefully designed to minimize magnitude/phase dispersion.

4.3.3   Drain Induced Memory Effects

Unwanted drain modulation is one of the most common sources of short-term memory effects in PAs. The drain biasing feed is designed to provide a low-impedance path at low frequencies and a high impedance path at RF carrier frequency. The bandwidth of the low-frequency path, also know as video bandwidth, is limited and depends on the inductive and capacitive components in the biasing feed. This bandwidth dictates the maximum instantaneous bandwidth of the input modulated signal, as any signal components beyond this bandwidth will experience drain induced memory effects.

Drain induced memory effects are rooted in even-order nonlinearities of the PAs. To explain the mechanisms of this dynamic nonlinearity, consider the PA driven by a two-tone stimulus at f_1 and f_2. The inherent nonlinear drain current source introduce multiple distortion products distributed over a broad frequency range. Among those distortion products, the second-order inter-modulation term is very critical because it occurs at a low-frequency current component (i.e., $\Delta f = f_2 - f_1 < 100$ MHz in case of single-band modulated signals). If the impedance of the DC supply network is significant (Δf is greater than the video bandwidth of the supply network), a non-negligible low-frequency drain voltage is induced which dynamically changes the supply voltage of the transistor on a timescale appropriate to the signal envelope, also called unwanted drain modulation.

Note that this undesired voltage is not in phase with the signal envelope due to the typical inductive nature of the biasing feed at low frequencies. Hence, the lowered supply voltage at some instants can push the transistor into the knee region. Recalling from Section 4.2, knee region intrusion can cause significant quasi-static static nonlinearity (AM/AM and AM/PM); in this case, dynamic knee region intrusion results in dynamic nonlinearity or memory effects. This is known as the dominant source of memory effects in GaN PAs [47]. Another mechanism occurs if the output capacitance of the device is strongly nonlinear (e.g., in LDMOS transistors). Capacitance variation caused by the drain modulation (on the signal envelope timescale) causes time-varying AM/AM and AM/PM distortions, i.e. memory effects. The third mechanism is attributed to the Miller capacitance, C_{gd}. If the undesired voltage component created at $f_2 - f_1$ is fed back to the input (gate terminal), it will mix with the fundamental components at f_1 and f_2 to create components at IMD3 frequencies, $2f_1 - f_2$ and $2f_2 - f_1$. Since the gain and phase shift of the PA could be different at f_1 and f_2, the two IMD3 components are imbalanced, which is another indicator of memory effects [48].

In the case of concurrent dual-band excitation, Δf can be larger than the conventional video bandwidth of the biasing feed network ($\sim$100 MHz). Hence, if a broadband PA is of interest, maintaining a small drain node impedance from DC to the bandwidth of the PA is imperative for concurrent multi-band signal transmission. This is a challenging task, since in a broadband PA, the bandwidth is comparable with the center frequency, at which the drain node should see a high impedance. Note that even though this nonlinearity exists for all PA topologies, in case of DPAs, this issue is aggravated due to the presence of two transistors, and thus two drain biasing networks.

In a conventional DPA realization, the main and peaking drains are biased separately [see Figure 4.21(a)] to allow careful monitoring the drain current of both transistors during measurements for troubleshooting. Hence, high-Z_o quarter-wave lines (or large inductors known as RF chokes) and large capacitors are employed as the biasing feed and DC blocks, respectively, in order not to affect the RF performance. However, this large L-C combination often resonates at low frequency, resulting in a high impedance at the drain of the main and peaking transistors, as illustrated in Figure 4.22. This problem is exacerbated when designing multi-way DPAs, since multiple biasing feeds and DC-blocks can create multiple resonances at low frequencies. One remedy for the problem is to absorb the biasing feed into the output combining network and use a single biasing feed for both transistors (if they share the same drain supply voltage), as shown in Figure 4.21(b). Therefore, the value of effective L and C at low frequency reduce, which pushes the resonant frequency to higher values outside the required video bandwidth, so that a small low-frequency impedance is maintained at the drain node (Figure 4.22).

(a) Conventional biasing network.

(b) An example of biasing network to reduce the biasing feed impedance.
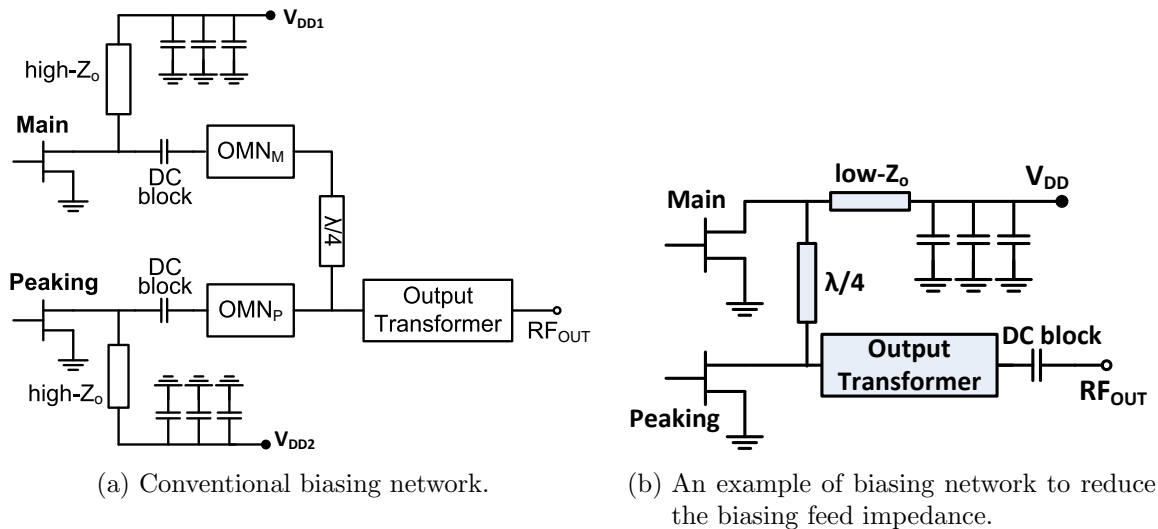
Figure 4.21: Illustration of conventional vs. recommended biasing networks for DPAs.
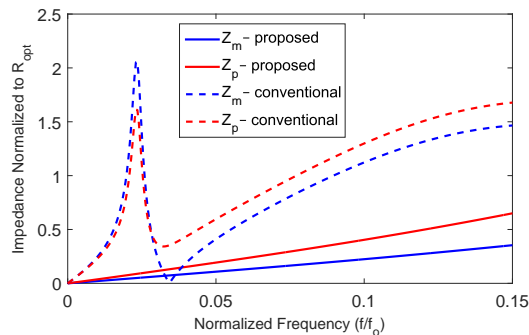
Figure 4.22: Biasing feed of a conventional vs. proposed DPA.

The previous description of drain induced memory effects is validated by modulated signal simulation of the two following cases:

- Case I – A DPA with large RF chokes as bias feeds and separate drain supplies for the two transistors [Figure 4.21(a)].

- Case II –A DPA with only one supply voltage connected to the main transistor through a low-Z quarter-wave line [Figure 4.21(b)].

The two DPAs were designed using GaN devices to operate at 2 GHz with more than 44 dBm output power. Note that in both DPAs harmonics are ideally shorted for a fair

87
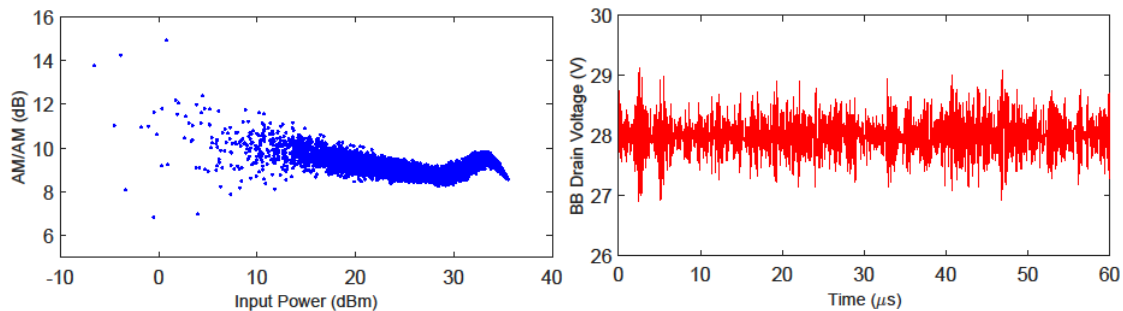
(a) Case I – Significant memory effect due to undesired drain modulation.
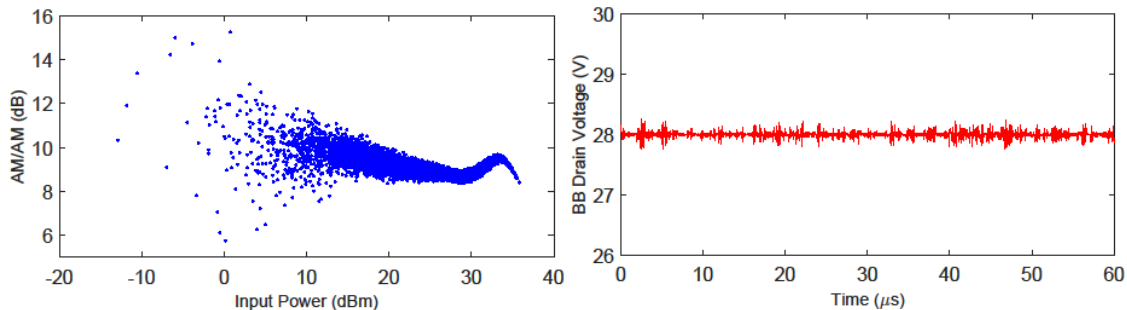
(b) Case II – Reduced drain induced memory effect by reducing the bias feed impedance.

Figure 4.23: Simulated AM/AM responses and baseband drain voltages of Case I and II.

comparison. Moreover, both DPAs share the same input matching and output combining networks (except for the biasing feed), so the static AM/AM and AM/PM responses are the same at 2 GHz. The two DPAs were simulated under a 10-MHz modulated signal using Keysight ADS Ptolemy co-simulation tool. Figure 4.23 illustrates a qualitative picture of drain induced memory effects using simulated AM/AM responses for the two cases. As can be seen, more spread (and thus more memory effect) is observed in Case I as a result of undesired drain modulation caused by the biasing feeds.

To quantify these memory effects, a 5th order memoryless polynomial DPD function is applied to linearize the DPAs. Since a memoryless DPD only corrects for static nonlinearities, the residual distortion indicates the intensity of memory effects in the PA. Figure 4.24 displays the output spectrum of the two cases, with and without memoryless DPD. It can be seen that for Case II, ACLR has been successfully reduced to –50.2 dBc which meets the spectrum mask. However, in Case I, ACLR was only improved to –42.8 dBc by using memoryless DPD. Table summarizes the simulation results for the two cases with and

88
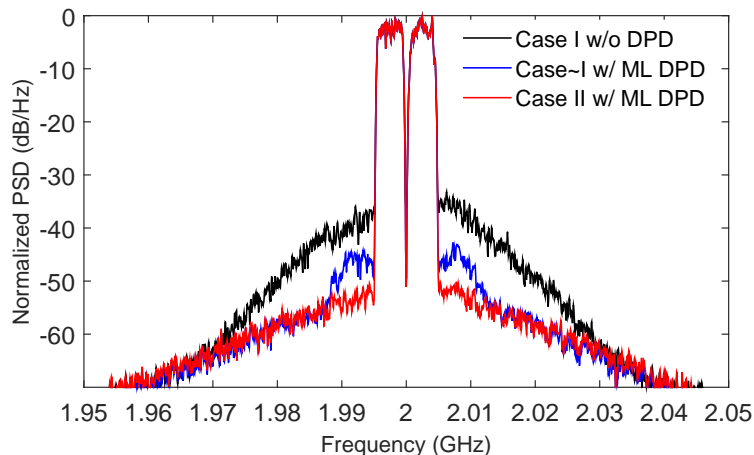
Figure 4.24: Simulated spectrum of Case I and II with and without memoryless DPD.

Table 4.2: Summary of the simulation results for the two DPA cases.

		Avg. Power (dBm)	NMSE (dB)	ACLR (dB)	EVM (%)
Case I	No DPD	37.0	–34.7	–33.2	4.1
	With DPD	37.0	—	–42.8	1.5
Case II	No DPD	37.0	–39.8	–35.1	3.6
	With DPD	37.0	—	–50.2	0.9

without memoryless DPD.

4.4 Conclusion

This chapter presented a detailed analysis of sources of distortion in broadband DPAs. The presented analysis provides a designer with insight into static and dynamic distortion mechanisms in place when the PA is driven with single-band and multi-band modulated signals. Furthermore, it helps one to reduce the nonlinearities in DPAs with minimum sacrifice in RF performance (efficiency and power vs. bandwidth). This is to ensure linearizability of the DPA with a fairly simple DPD algorithm which becomes crucial as the base stations move towards pico- and femto-cells.

Chapter 5

Broadband Waveform-Engineered DPA

5.1 Introduction

As discussed in Chapters 2 and 3, broadband two-way and multi-way DPAs have been developed to improve the back-off efficiency of PAs over a fractional bandwidth of 30%. On the other hand, the introduction of continuous modes of operation including class B/J [44] and continuous class F [46] design spaces, has defined a continuum of fundamental and harmonic impedances that employ the concept of waveform engineering to deliver the same output power, efficiency and linearity. The use of waveform engineering provides more flexibility in achieving a broadband match (50% fractional bandwidth or more) which has been confirmed with experimental results [49,50]. For instance, in [50] a broadband single-ended PA based on class B/J design space was implemented using simplified real frequency technique (SRFT) and achieved peak drain efficiency in excess of 60% from 1.9 to 2.9 GHz. Despite achieving a significant bandwidth, still the back-off efficiency was poor and not suitable for high-PAPR applications.

This chapter attempts to bridge the gap between the waveform engineering concept and class B based 2W-DPA theory. This approach offers significant flexibility for designing broadband DPAs, similar to the design flexibility provided by "continuous" modes of operation for broadband single-device PAs. It will be shown that the typical bandwidth limiting factors of conventional DPAs do not exist in a waveform-engineered DPA (WeDPA); thus, theoretically, the achievable bandwidth of a WeDPA is equal to that of a single-device PA. In fact, the frequency response of the input matching networks dictate
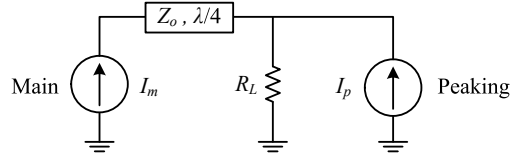
Figure 5.1: Simplified schematic of a conventional DPA.

the bandwidth of the WeDPA rather than the output combining network. The chapter begins with the theoretical foundation for WeDPAs using the concept of class B/J design space. Based on the proposed formulation, an approach to designing a broadband WeDPA is presented. As proof of the proposed theory, a broadband WeDPA prototype was designed, fabricated, and measured.

5.2   Doherty Amplifier and Waveform Engineering

The theory of conventional 2W-DPA was described in detail in Chapter 2. For ease of reference, the simplified schematic as well as the current and voltage profiles of a conventional 2W-DPA are shown again in Figures 5.1 and 5.2, respectively. As shown in Section 2.5, with proper choice of the characteristic impedance of this line and the load impedance, $R_L = R_{\mathrm{opt}}/2$, the maximum output voltage swing (and thus maximum efficiency) can be obtained over the 6-dB back-off range, while linearity is maintained.

Based on the fundamental impedance characteristics illustrated in Figure 5.2(c), it can be observed that,

1. The real part of the main impedance is modulated from $2R_{\mathrm{opt}}$ at low power levels to R_{opt} at peak power.

2. The real part of the peaking impedance varies from ∞ at low power levels to R_{opt} at peak power.

3. The imaginary parts of both impedances are zero at all power levels.

4. All of the harmonics are shorted at the intrinsic drain terminal of both devices.

Even though several variations of the conventional topology (Figure 5.1) have been proposed in the literature, they all satisfy the above conditions at the center frequency. As

91

(a) Current profiles.

(b) Voltage profiles.

(c) Impedance profiles.
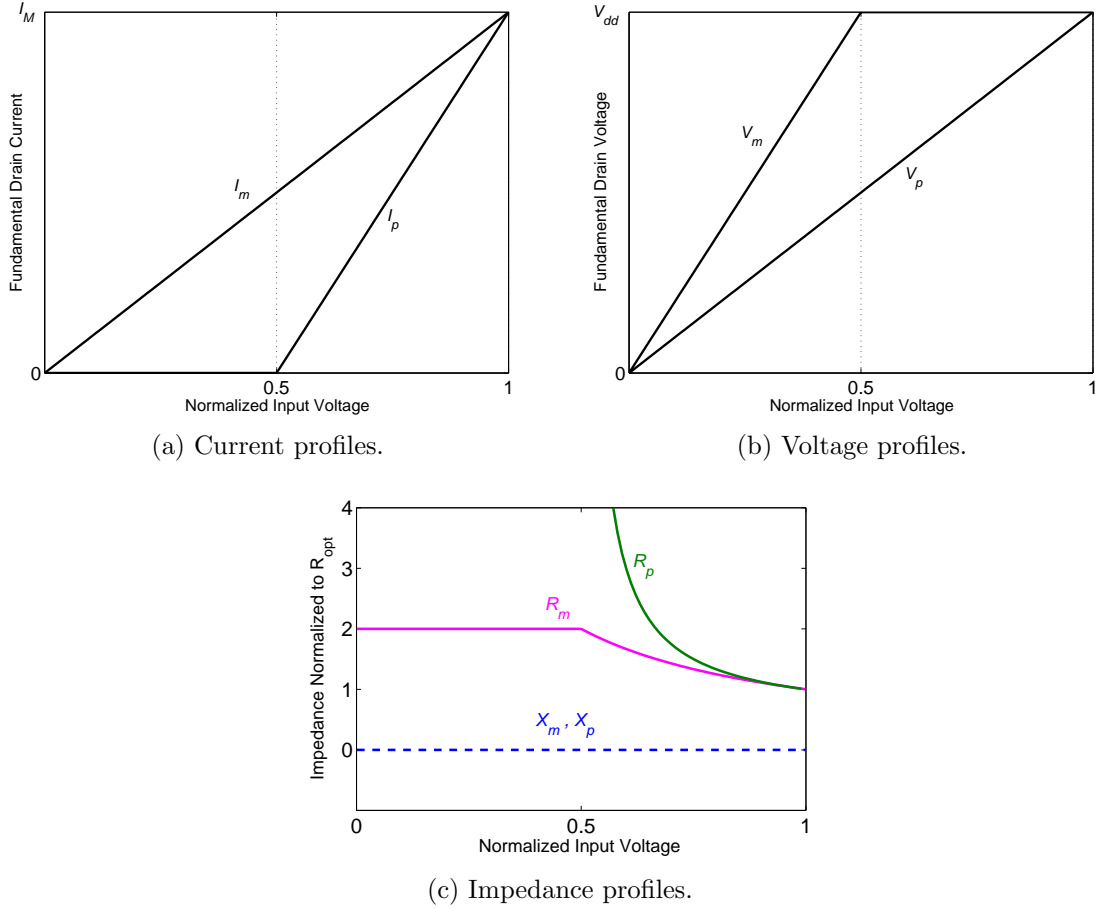
Figure 5.2: Fundamental components of drain current, voltage, and impedance of the main and peaking transistor in a conventional DPA.

frequency varies, the transmission line will no longer be a perfect impedance inverter and as a result, the imaginary parts will become non-zero. However, an extended bandwidth is obtained as long as the imaginary parts of the fundamental impedances are small compared to the real parts and if the harmonic contents of both drain voltages are negligible with reference to the fundamental components. This brings up the following question: Is it possible to extend the theory of DPAs for complex-valued fundamental/harmonic impedances instead of the real-to-real class B based load modulation? The rest of this section attempts to provide an answer for this crucial question.
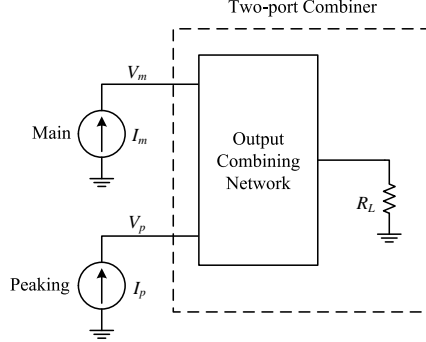
Figure 5.3: Generic illustration of a DPA with ideal current sources and a lossless combining network. Device output parasitics are embedded in the output combining network.

5.2.1   WeDPA Concept

In this subsection a novel approach to extending DPA theory using the concept of waveform engineering and black-box analysis of the output combining network, is proposed. For simplicity, the class B/J design space was used in this analysis, however, the approach can be used for other continuous modes of operation [44, 46]. Figure 5.3 is a generic illustration of a DPA that includes a three-port lossless combining network to deliver power to the load. Throughout the analysis, the 50 Ω termination is embedded into the output combining network; hence, the three-port combiner will be reduced to a two-port network as illustrated in Figure 5.3. It will be shown that this greatly simplifies the analysis. Note that the resulting two-port network, referred to as the reduced combiner hereafter, is no longer lossless, though it is still reciprocal.

In this analysis, ports 1 and 2 represent the intrinsic drains of the main and peaking transistors, and the output device parasitics are included in the output combining network. Using linear network theory, one can express the fundamental currents and voltages of the main and peaking devices in terms of the Z-parameters of the black-box two-port combiner, as

$$\begin{bmatrix} V_m \\ V_p \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \cdot \begin{bmatrix} I_m \\ I_p \end{bmatrix}. \tag{5.1}$$

Evaluating (5.1) under two specific drive conditions: B, the back-off point at which the efficiency is maximized and F, the point where the DPA delivers its full power, results in

$$V_m^B = Z_{11}I_m^B + Z_{12}I_p^B \tag{5.2}$$

$$V_p^B = Z_{21}I_m^B + Z_{22}I_p^B \tag{5.3}$$

$$V_m^F = Z_{11}I_m^F + Z_{12}I_p^F \tag{5.4}$$

$$V_p^F = Z_{21}I_m^F + Z_{22}I_p^F \tag{5.5}$$

where V_m^B, I_m^B and V_m^F, I_m^F are the voltages and currents of the main device at the back-off and peak power levels, respectively, and V_p^B, I_p^B and V_p^F, I_p^F are the voltages and currents of the peaking device. Equations (5.2)–(5.5) form a system of linear equations, including four unknowns, Z_{11}, Z_{12}, Z_{21}, and Z_{22}. By solving the equations simultaneously for the Z-parameters [25], we have

$$Z_{11} = \frac{-V_m^F I_p^B + I_p^F V_m^B}{-I_m^F I_p^B + I_m^B I_p^F} \tag{5.6}$$

$$Z_{12} = -\frac{I_m^F V_m^B - I_m^B V_m^F}{-I_m^F I_p^B + I_m^B I_p^F} \tag{5.7}$$

$$Z_{21} = -\frac{V_p^F I_p^B - I_p^F V_p^B}{-I_m^F I_p^B + I_m^B I_p^F} \tag{5.8}$$

$$Z_{22} = \frac{-I_m^F V_p^B + I_m^B V_p^F}{-I_m^F I_p^B + I_m^B I_p^F}. \tag{5.9}$$

Considering ideal current profiles for the main and peaking transistors, as is the case for a conventional DPA [see Figure 5.2(a)], we have

$$I_m^B = \frac{1}{2} \cdot I_m^F \tag{5.10}$$

$$I_p^B = 0 \qquad\qquad \text{Class C biasing,} \tag{5.11}$$

$$|I_p^F| = |I_m^F| = I_{\max} \qquad \text{Equal power contribution.} \tag{5.12}$$

Also, if Z_{opt}^B and Z_{opt}^F represent the optimum load impedances of the main transistor corresponding to class B/J operation at back-off and full power, respectively, and Z_{opt}^P denotes the optimum peaking load impedance at full power, then,

$$V_m^B = Z_{\text{opt}}^B \cdot I_m^B \tag{5.13}$$

$$V_m^F = Z_{\text{opt}}^F \cdot I_m^F \tag{5.14}$$

$$V_p^F = Z_{\text{opt}}^P \cdot I_p^F. \tag{5.15}$$

where

$$Z_{\text{opt}}^B = 2R_{\text{opt}} + jX_B \tag{5.16}$$

$$Z_{\text{opt}}^F = R_{\text{opt}} + jX_F \tag{5.17}$$

$$Z_{\text{opt}}^P = R_{\text{opt}}^P + jX_P. \tag{5.18}$$

In (5.16)–(5.18), $|X_B| \leq 2R_{\text{opt}}$, $|X_F| \leq R_{\text{opt}}$, and $|X_P| \leq R_{\text{opt}}^P$ will ensure proper operation according to class B/J theory [44]. The goal is now to determine the Z-parameters of the combining network in order to satisfy the above conditions. However, in this case, unlike that of a conventional DPA, the harmonics must not be shorted. This will be discussed in more detail later in this section. Using (5.6), (5.7), and (5.9) at a given fundamental frequency, f_o, combined with (5.10)–(5.12) and (5.13)–(5.15) we obtain

$$Z_{11}(f_o) = Z_{\text{opt}}^B \tag{5.19}$$

$$Z_{12}(f_o) = Z_{21}(f_o) = -\frac{I_m^F}{I_p^F}\left(Z_{\text{opt}}^B - Z_{\text{opt}}^F\right)$$

$$= -e^{-j\phi}\left(R_{\text{opt}} + j\Delta X_m\right) \tag{5.20}$$

$$Z_{22}(f_o) = Z_{\text{opt}}^P - Z_{12}(f_o)\frac{I_m^F}{I_p^F}$$

$$= Z_{\text{opt}}^P + e^{-j2\phi}\left(R_{\text{opt}} + j\Delta X_m\right) \tag{5.21}$$

where $\Delta X_m = X_B - X_F$. In (5.19)–(5.21), ϕ, which is the phase of the peaking current with reference to the main current, is used as an additional design parameter. It can be shown that only for the case where $\phi = \pm\pi/2$ is perfect linearity achieved over the entire power range; the same result as for a conventional DPA. The impact of ϕ will be further discussed in Section 5.2.3. In addition, using (5.8),

$$V_p^B(f_o) = Z_{21}(f_o)I_m^B(f_o)$$

$$= -e^{-j\phi}\left(R_{\text{opt}} + j\Delta X_m\right)I_m^B, \tag{5.22}$$

will enforce another requirement for proper WeDPA operation based on class B/J theory. The imaginary part in (5.22) must be smaller than, or equal to, the real part (in magnitude) in order to avoid a clipping voltage waveform; hence,

$$|X_B - X_F| \leq R_{\text{opt}}. \tag{5.23}$$

Note that the current analysis assumes the second harmonic terminations are ideally adjusted for perfect class B/J operation and higher order harmonics are short-circuited.

This means that if the fundamental load impedance of the main (or peaking) transistor is equal to $Z(f_o) = R + jX$ at a given power level, then the corresponding second harmonic impedance must be $Z(2f_o) = -j3\pi/8X$ [44]. Thus, one can deduce the requirements of the combiner's Z-parameters at the second harmonic frequency in a similar way, as follows,

$$Z_{11}(2f_o) = -j\frac{3\pi}{8}X_B \tag{5.24}$$

$$Z_{12}(2f_o) = Z_{21}(2f_o) = j\frac{3\pi}{8}e^{-j2\phi}\left|\frac{I_m^F}{I_p^F}(2f_o)\right|\Delta X_m \tag{5.25}$$

$$Z_{22}(2f_o) = -j\frac{3\pi}{8}X_P - Z_{12}(2f_o)e^{-j2\phi}\left|\frac{I_m^F}{I_p^F}(2f_o)\right| \tag{5.26}$$

where $I_m^F/I_p^F(2f_o)$ denotes the ratio of the second harmonic magnitude of the main current to that of the peaking current at full power. This ratio is highly dependent on the exact bias point of the peaking device operating in class C. In practice, however, meeting the second harmonic relations (5.24)–(5.26) is less important than addressing the fundamental impedances described in (5.19)–(5.21).

This novel formulation of the load modulation problem significantly relaxes the design constraints of DPAs and benefits from the advantages of class B/J operation versus class B operation in terms of bandwidth and feasibility. Although the formulation presented here specifically involved class B/J modes, the approach is general and readily applicable to any other continuous mode of operation. In fact, by knowing Z_{opt}^B, Z_{opt}^F, and Z_{opt}^P at the fundamental and harmonic frequencies using waveform engineering theory, one can determine the required Z-matrix of the reduced combining network at those frequencies as follows,

$$Z_{11}(nf_o) = Z_{\text{opt}}^B(nf_o) \tag{5.27}$$

$$Z_{12}(nf_o) = Z_{21}(nf_o)$$
$$= -e^{-jn\phi}\left|\frac{I_m^F}{I_p^F}(nf_o)\right|\left[Z_{\text{opt}}^B(nf_o) - Z_{\text{opt}}^F(nf_o)\right] \tag{5.28}$$

$$Z_{22}(nf_o) = Z_{\text{opt}}^P(nf_o) - Z_{12}(nf_o)e^{-jn\phi}\left|\frac{I_m^F}{I_p^F}(nf_o)\right| \tag{5.29}$$

where n represents the harmonic index. This approach can also be extended to DPAs with a back-off region of larger than 6 dB, simply by modifying the current and impedance profiles. Once the Z-parameters of the reduced combiner are determined, one can reintroduce the third port connected to the load and deduce the Z-parameters of the three-port combining network by applying the unitary conditions, as described in [25].

(a) Voltage of the main device   (b) Voltage of the peaking device
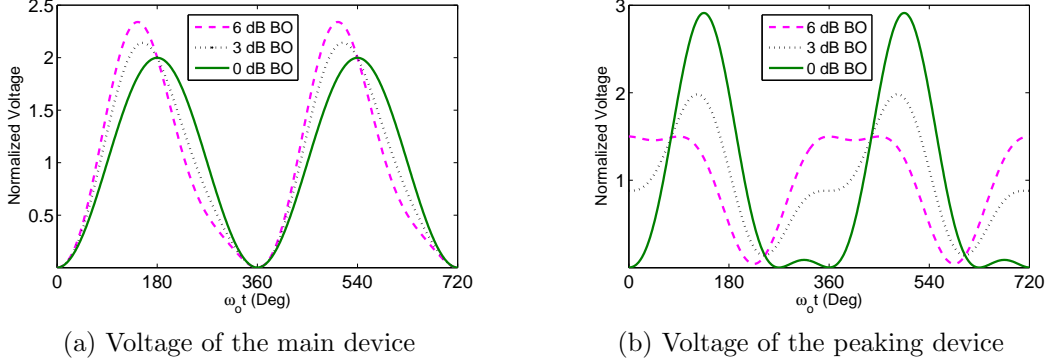
Figure 5.4: Theoretical voltage waveforms of the (a) main , and (b) peaking transistors for {1,0,–1}.

(a) Voltage of the main device   (b) Voltage of the peaking device
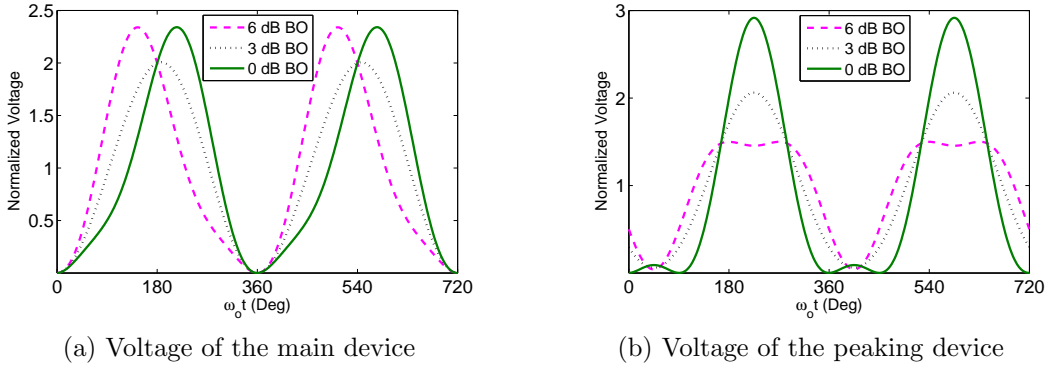
Figure 5.5: Theoretical voltage waveforms of the (a) main , and (b) peaking transistors for {0.5,–0.5,1}.

5.2.2   Analysis of WeDPA

This subsection presents an analysis of the WeDPA concept described in Section 5.2.1. According to (5.19)–(5.21) and (5.24)–(5.26), the reduced combiner's parameters are uniquely determined by setting X_B, X_F, X_P, and ϕ. For simplicity, assume that $\{x_B, x_F, x_P\}$ refers to a given instance in the class B/J continuum (referred to as a "mode" hereafter), where $x_i = X_i/R_{\text{opt}}$ and $i = B, F$, or P. Therefore, $\{0,0,0\}$ refers to the conventional DPA composed of ideal class B/C amplifiers with real fundamental impedances and short-circuited harmonics. Table 5.1 shows different modes of DPA combiner based on the class B/J definition. For the present analysis ϕ is assumed to be $\pi/2$.

Table 5.1: Example of WeDPA modes based on class B/J theory.

	x_B	x_F	x_P
Class B	0	0	0
Class J	2	1	1
Class J*	-2	-1	-1

For a given mode (designated by the triplet $\{x_B, x_F, x_P\}$), and based on the fundamental and second harmonic frequencies from (5.19)–(5.21) and (5.24)–(5.26), the circuit can be simulated using ideal current source models for the devices (see Figure 5.3). Assuming a lossless combining network, the output power delivered by the DPA is calculated as

$$P_{out} = \frac{1}{2}\text{Re}\{V_m(f_o)I_m^*(f_o) + V_p(f_o)I_p^*(f_o)\}. \tag{5.30}$$

Two examples of combiners corresponding to two different modes will be analyzed here: $\{1,0,-1\}$ for Case I and $\{0.5,-0.5,1\}$ for Case II. These modes were selected to illustrate how the voltage waveforms of the devices vary within the class B/J continuum at different power levels. Figures 5.4 and 5.5 illustrate the main and peaking voltage waveforms for the two cases at 6-, 3-, and 0-dB back-off levels. It can be seen that the waveforms of Case I correspond to class B and class J* modes at the peak power for the main and peaking transistors, respectively. In Case II, the operation mode of the main device varies from an instance between class B and J at 6-dB back-off power to its conjugate between class B and J* at full power, so that the waveform looks like a class B shape at the mid-point (i.e., 3-dB back-off level). It can also be observed in Figures 5.4 and 5.5 that the ideal "zero-grazing" conditions are satisfied for all cases; indicating, theoretically, the linear functionality of a WeDPA.

Figure 5.6 shows the fundamental impedance profiles versus normalized drive voltage for the two cases. It is confirmed from Figure 5.6 that the "real" part of the main and peaking impedance profiles follow the exact trend of a conventional DPA [see Figure 5.2(c)]; however, the profiles of the imaginary parts will be different. The efficiency profile associated with the two WeDPA cases (and also other modes represented by $\{x_B, x_F, x_P\}$) is depicted in Figure 5.7. This efficiency profile is identical to that of a conventional class B DPA. In fact, all $\{x_B, x_F, x_P\}$ modes result in an identical performance (in terms of power, efficiency and linearity), and thus can form a three-dimensional (3-D) class B/J design space. It can be inferred from these results that the proposed approach will ensure ideal Doherty behavior for a much larger space than the conventional DPA; a design space which
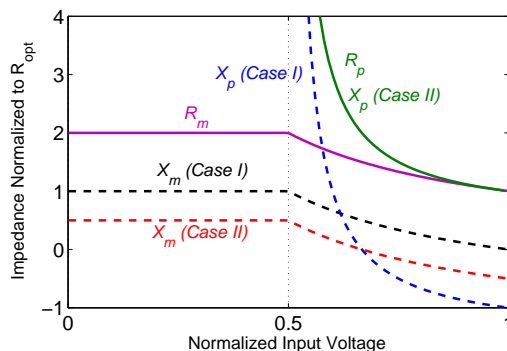
Figure 5.6: Fundamental impedance profiles of the main and peaking devices for Cases I and II.
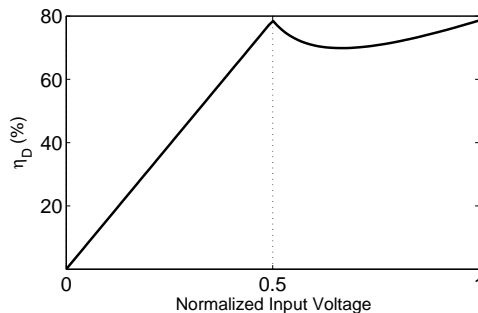
Figure 5.7: Efficiency profile of the WeDPA.

can be harnessed for broadband DPA design. Additionally, unlike previous design methods, no presumptions are made about the combining network topology.

5.2.3 On the Role of Phase Difference

As mentioned in Section 5.2.1, the voltage (or impedance) requirements at the 6-dB back-off and peak powers are satisfied regardless of the ϕ value [see (5.13)–(5.15)]. In other words, for a given ϕ at a specific frequency, there exists a combining network defined by (5.19)–(5.21) and (5.24)–(5.26) that ensures optimum load conditions at the two said power levels. However, simulation results show that for ϕ values other than $\pm 90°$, both efficiency and linearity degrade at mid power levels (between 6-dB back-off and full power). To illustrate this point, Figure 5.8 shows the simulation results using ideal current source models for $\{0,0,0\}$ mode (ideal class B DPA) and different ϕ values ranging from $60°$ to
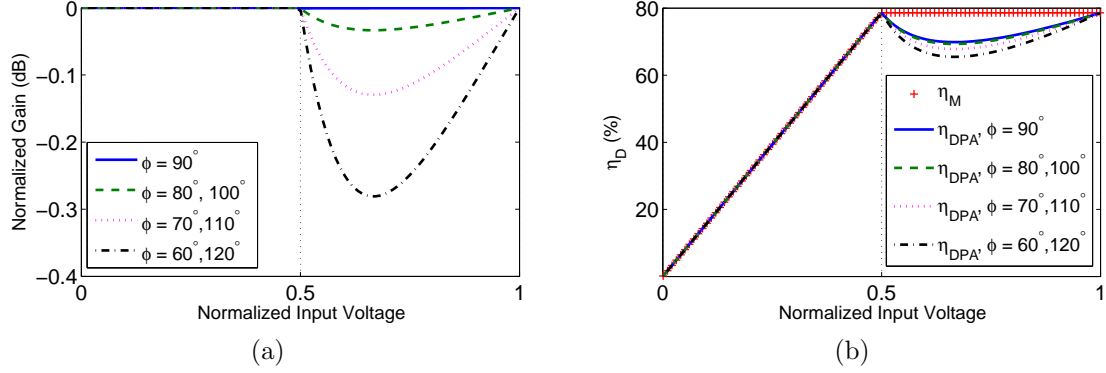
Figure 5.8: (a) Gain and (b) efficiency profiles for {0,0,0} mode and different ϕ values. η_M and η_{DPA} represent the main transistor's and overall efficiency, respectively.
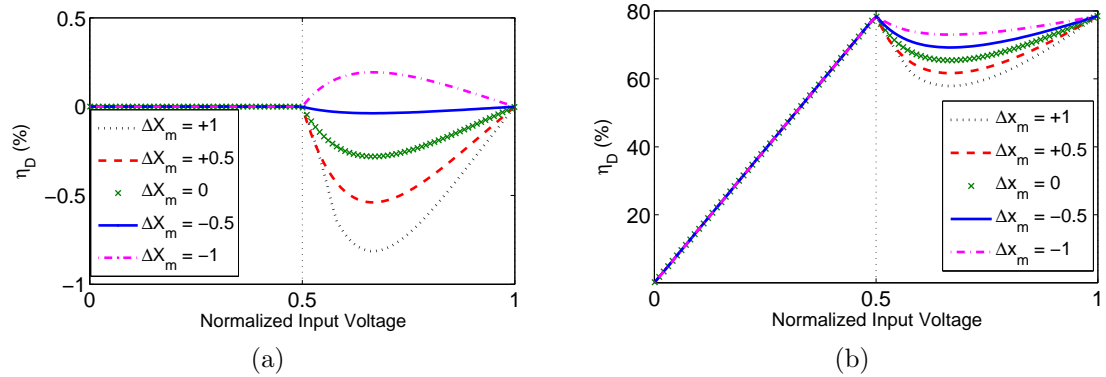
Figure 5.9: (a) Gain and (b) efficiency profiles for different Δx_m values and $\phi = 60°$.

120°, corresponding to an octave bandwidth assuming that assumes ϕ follows a linear relationship with frequency. Note that for each ϕ value a different combining network is assumed which yields optimal performance at 6-dB back-off (and below) and peak power levels. It can be seen that even though the efficiency profile of the main transistor remains unchanged, the overall efficiency and linearity degrade as ϕ deviates from 90°. Investigating further, we observe that only the voltage (or impedance) profile of the peaking transistor deviates from the optimal curve for ϕ values different from 90°.

Additionally, the extent of the performance degradation will depend on the WeDPA mode imposed by $\{x_B, x_F, x_P\}$. Simulation results suggest that the extent of the degradation is a function of the variation of the main transistor's reactance versus power, defined

as $\Delta x_m = (X_B - X_F)/R_{\text{opt}}$, but it is independent of X_P. Figure 5.9 illustrates the efficiency and gain profiles for different Δx_m values in the worst case for an octave-bandwidth DPA, namely $\phi = 60°$. As shown, the DPA combiner associated with $\Delta x_m = -0.5$ is less sensitive to the ϕ variation, both in terms of efficiency and linearity. Repeating this simulation for the upper edge ($\phi = 120°$), shows the opposite behavior, meaning that the output combiner corresponding to $\Delta x_m = +0.5$ is less sensitive to the ϕ variation. This indicates that there exists a "sweet spot" which minimizes the performance degradation of the WeDPA for a given ϕ. Nevertheless, as shown in Figure 5.9, if the ϕ variation is substantial within the frequency band, minimizing the ϕ-dependent performance degradation requires drastic changes in the operation mode of the main transistor. For example, a change from $\Delta x_m = -0.5$ at the lower edge to $\Delta x_m = +0.5$ at the upper edge of the band, limits realizability of the WeDPA and may not be a viable solution. However, according to Figure 5.9, even in the worst case scenario the linearity and efficiency degradations are insignificant and acceptable for most applications.

From a practical point of view, ϕ is dictated by the type of power splitter used at the input of the DPA. The three most common types used are the quadrature hybrid coupler, Lange coupler, and the Wilkinson divider. The first two maintain a relatively constant phase shift of 90° across the bandwidth, whereas the Wilkinson divider (when followed by a $\lambda/4$ line) results in a phase relation proportional to the frequency of operation, or $\phi = \pm 90° \cdot f/f_o$, where f and f_o denote the operating and center frequencies, respectively. Nevertheless, designing a broadband hybrid coupler with sufficient return loss, isolation and symmetry between the two output ports is more difficult than designing a broadband Wilkinson divider. Hence, the Wilkinson divider together with a quarter-wave line is a viable solution up to an octave bandwidth, where ϕ variation is limited to $\pm 30°$ across the band. A Lange coupler, on the other hand, is inherently broadband, but requires additional wire-bonding, complicating the circuit assembly. Nevertheless, it is a potential candidate for DPAs with more than an octave bandwidth [30].

5.3 WeDPA Implementation

Based on the analysis developed in the previous section, a broadband DPA prototype was designed using 25-W and 45-W packaged gallium nitride (GaN) high-electron-mobility transistors (HEMTs) from Cree® as the main and peaking devices, respectively. A larger peaking device was selected to provide the proper current profiles, according to Figure 5.2(a). The design was conducted using the exploded large-signal models for the two transistors. When designing DPAs, conventional or otherwise, it is not always possible
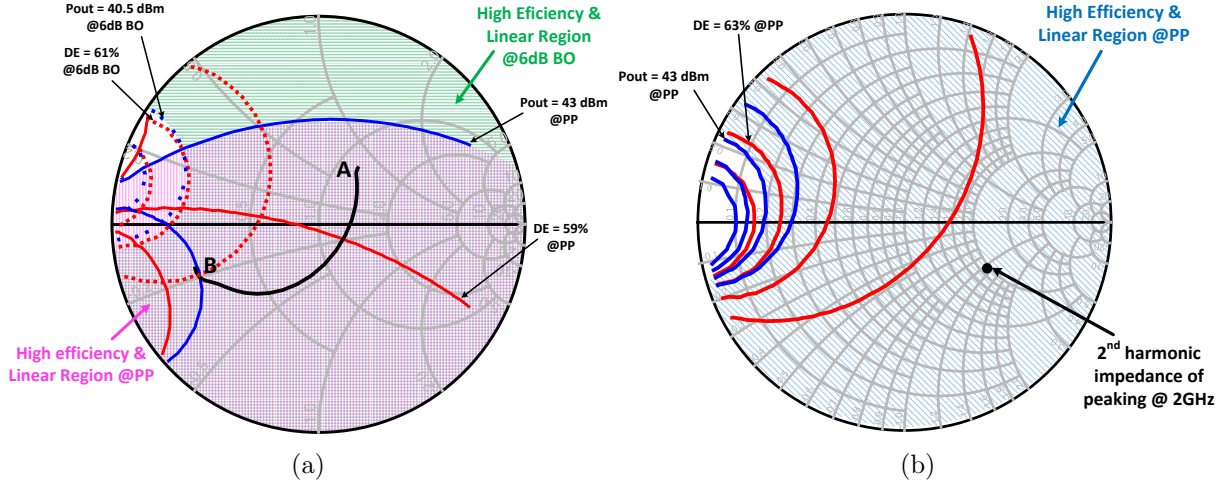
Figure 5.10: Second harmonic load-pull efficiency (red) and power (blue) contours at the package plane for the (a) main device at BO (dotted lines) and PP (solid lines), and (b) peaking device at PP at f_o=2 GHz (50 Ω impedance system). The realized impedance profiles at second harmonic are shown in black. The fficiency and power steps are 5% and 0.5 dB, respectively.

to select devices with the exact sizes required. Other solutions proposed in the literature include using uneven input power division or asymmetrical drain supply voltages [16], however these would compromise the gain and circuit complexity.

To select the WeDPA operation mode within the 3-D class B/J continuum at a given frequency point, two factors were taken into account:

- *Sensitivity to the harmonic mismatch:* According to (5.24)–(5.26), the Z-parameter conditions of the reduced network at the second harmonic frequencies depend on the ratio of the second harmonic current components of the transistors. This ratio is bias-dependent and thus hard to control, particularly in a broadband design scenario. However, in a practical design, depending on the operation mode dictated by the fundamental impedance, the second harmonic termination at the package reference plane can be of secondary importance compared to the fundamental termination (as is also the case for single-device PAs). This means that for some modes, by accepting a small degradation in maximum efficiency (e.g., 10%) and output power (e.g., 0.5– 1.0 dB), only certain regions of the Smith chart would need to be avoided based on harmonic load-pull contours [51, 52]. In other words, the operating mode is not
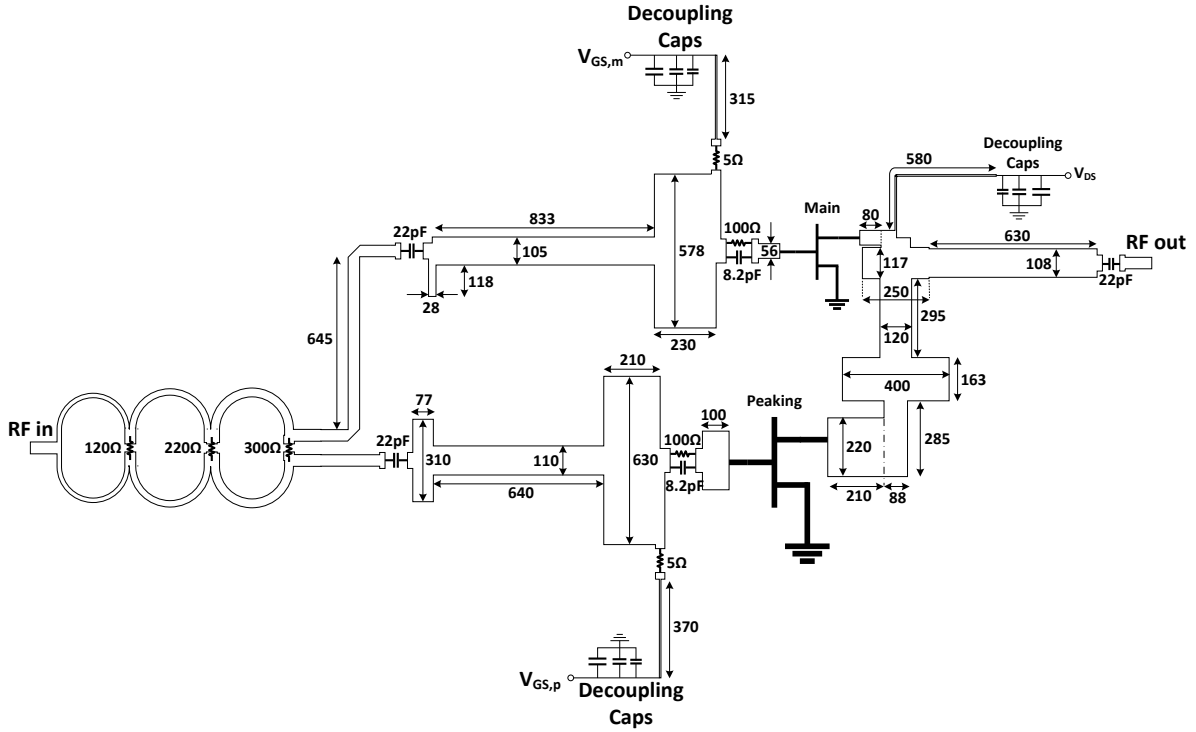
Figure 5.11: Schematic of the designed WeDPA. All dimensions are in mils (0.001 inch). The substrate is RT 6035HTC ($\epsilon_r = 3.6$, $h = 20$ mil) from Rogers.

necessarily chosen to achieve the highest possible efficiency at a given frequency; doing so could result in high sensitivity to the harmonic terminations rendering the design of a wideband combining network an impossible task. Hence, the WeDPA mode (X_B, X_F, and X_P) at a given frequency was selected based on which showed the least efficiency sensitivity to deviation from (5.24)–(5.26) in the second harmonic termination (at the package plane), in order to relax the requirements on the output combining network to enable broadband DPA design.

- ϕ-*dependant degradation:* As discussed in Section 5.2.3, linearity and efficiency deteriorate for a given $\phi \neq 90°$. This becomes particularly important in the case where a Wilkinson power splitter is used at frequencies farther away from the center frequency, toward the edges. However, as discussed, for ϕ variation of less than $\pm 30°$ (equivalent to an octave bandwidth) this deterioration is tolerable. For fractional bandwidths of more than an octave, one can use multi-section quadrature couplers or Lange couplers to split the input signal so as to maintain a relatively constant ϕ

equal to 90° across the band. In this work, a multi-section Wilkinson divider was used, as will be discussed later.

Based on the second harmonic load-pull simulations of different modes within the 3-D class B/J space for the main (at back-off and peak power) and peaking transistors (at peak power), it was found that for both transistors, the sub-space between class B and J* is less sensitive to the second harmonic termination at the package reference plane. It should be noted that the die and package parasitics of the two transistors are incorporated into the output combining network shown in Figure 5.3, using the exploded large-signal transistor models provided by Cree®.

This fact is illustrated in Figure 5.10 which shows the second harmonic load-pull contours (power and efficiency) at the package plane for the main and peaking transistors at 2 GHz. The second harmonic impedance profile of the main transistor is shown as a dotted black line on the left chart, moving from "A" (at back-off power) to "B" (at peak power). The second harmonic of the peaking transistor's impedance at peak power is shown as a black point on the right chart. The main transistor's fundamental impedance at the intrinsic drain moves from $Z_m^B = 27 - 2j$ in back-off to $Z_m^F = 14 - 10j$ at peak power, while the peaking transistor sees a fundamental impedance of $Z_p = 13 - 3j$ at peak power at the intrinsic drain plane.

According to Figure 5.10(a), the green and purple regions represent the safe regions (in terms of high efficiency and linearity) for the main device at back-off and peak power levels, respectively, exactly where points A and B are located. It can be seen that these safe regions include the vast majority of the Smith chart. This means that by making a small sacrifice in efficiency and output power (i.e., linearity), the choice of the second harmonic impedance is significantly relaxed. This allows the designer to focus on synthesizing the fundamental impedances versus frequency, only checking the second harmonic impedances to ensure they are located in the safe regions. According to Figure 5.10(b), the second harmonic impedance of the peaking device at peak power is not sensitive either, as most of the Smith chart is in the safe region (highly efficient and linear). It is worth emphasizing that sensitivity to the second harmonic termination varies for different fundamental impedances corresponding to different modes. In this case, for example, for impedances with positive reactance at the intrinsic drain, the efficiency and power were found to be more sensitive to the second harmonic termination.

Figure 5.11 illustrates the schematic of the implemented WeDPA. The structure of the output combining network was chosen to maintain the impedance conditions imposed by (5.19)–(5.21) over a broad bandwidth. A $\lambda/4$ line ($Z_o = 26\,\Omega$) was used to match the 50 Ω load to a real impedance subsequent to the combining node (similar to a conventional DPA).
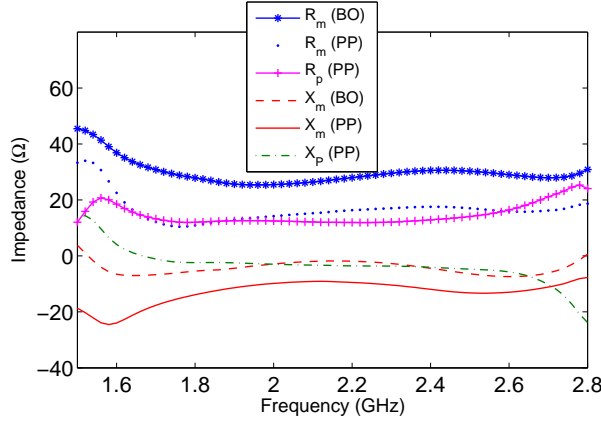
Figure 5.12: Simulated impedances presented to the main and peaking devices at back-off (BO) and peak power (PP) levels.
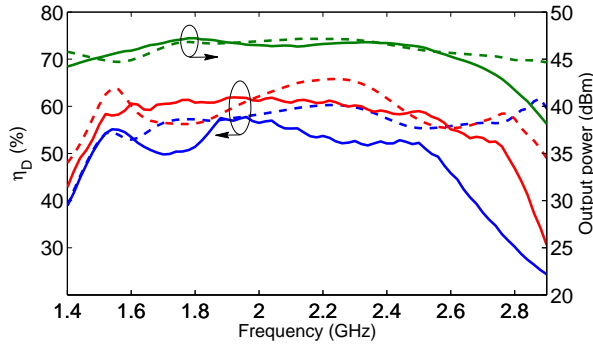
Figure 5.13: Simulated efficiency at 6-dB back-off (blue) and peak power (red) levels and peak output power (green) with (solid line) and without (dashed line) input matching networks.

The shunt inductor, L_1, (implemented as a transmission line) also serves as the biasing feed for both transistors in order to reduce the baseband drain impedance and mitigate memory effects [53, 54]. The main and peaking transistors were biased at $I_{DS} = 200$ mA and $V_{GS} = -4.7$ V, respectively, and $V_{DS} = 30$ V.

The profiles of the impedances at the intrinsic drain of the main and peaking transistors are shown in Figure 5.12 for 1.5–2.8 GHz ($R_{\mathrm{opt}} = 15\Omega$). It can be seen that the real parts of the impedances are maintained within an acceptable range of values from the class B/J continuum, while the imaginary parts are nonzero (see X_m profile at peak power, for instance).

105
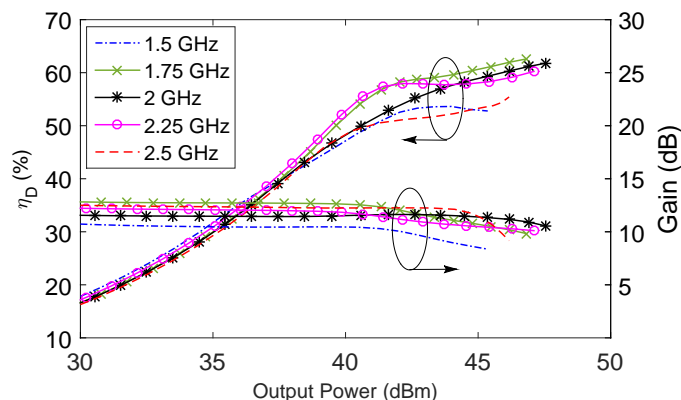
Figure 5.14: Simulated CW drain efficiency and gain versus output power at various frequency points.

Figure 5.13 displays the simulated plots of efficiency and peak output power versus frequency. Note that no input matching network was included at this point and ideal voltage sources have been used to drive the main and peaking amplifiers; thus, the bandwidth is dictated only by the output combiner (dashed lines in Figure 5.13). The WeDPA bandwidth was from 1.5 GHz to 2.9 GHz with 6-dB back-off efficiency of higher than 50%, and peak power degradation of less than 2 dB. This bandwidth (64%) is comparable to those achieved by broadband single-ended class B/J PA prototypes [44, 50]. Nevertheless, as will be shown, the WeDPA's bandwidth was compromised due to the input matching networks.

As seen in Figure 5.11, on the input side a three-section Wilkinson divider with an even power split ratio was designed to cover the frequency range of 1.4–6 GHz. It should be noted that GaN HEMTs exhibit strong nonlinearity at the input due to their inherently nonlinear input capacitance and are sensitive to the second harmonic input termination [55]. Therefore, it is crucial to provide sufficient return loss and isolation between the two output ports of the power divider, at least across the entire fundamental and second harmonic frequency ranges, since it allows for independent design of the main and peaking input matching networks.

The input matching networks were designed so that a relatively flat gain was achieved over the bandwidth. In particular, the main input matching needs to provide a flat gain at low power levels when the peaking device is off (similar to a single-ended PAs). As demonstrated in Chapter 4, in the case of a GaN transistor biased in class C, the nonlinear input capacitance introduces AM/AM distortion [55, 56] which makes it impossible to obtain an identical peaking current versus frequency at all power levels. Nonlinear input

106
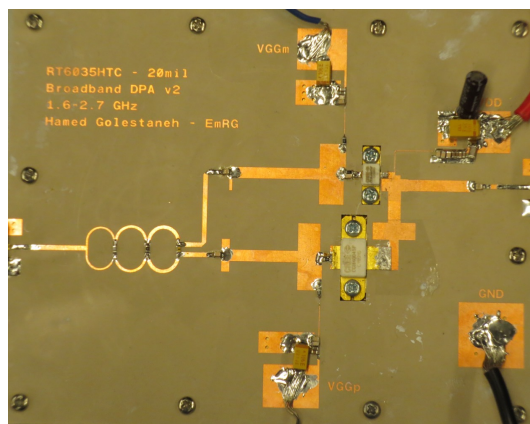
Figure 5.15: Photograph of the fabricated WeDPA.

capacitance of the transistors can also induce significant AM/PM distortion [57], which makes the phase alignment between the two paths possible only at one power level. In other words, at a given operating frequency, ϕ varies versus power, hence, the load modulation will be inadequate and result in linearity and efficiency degradation. These effects can considerably limit the achievable bandwidth of a WeDPA compared to the potential bandwidth of the output combining network. As can be seen in Figure 5.13 (solid lines), the bandwidth of the complete WeDPA including input matching networks was reduced to 1.5–2.6 GHz for the same specifications (more than 50% back-off efficiency and less than 2 dB peak power degradation versus frequency). The problem of input matching design for broadband DPAs using a systematic approach will be addressed in Chapter 6. The CW simulation results versus input power are also shown in Figure 5.14.

5.4   Experimental Results

Figure 5.15 depicts a photograph of the fabricated WeDPA demonstrator. To evaluate the performance of the amplifier, efficiency and output power were measured under CW stimulus. Figure 5.16 illustrates the measured CW plots at 6-dB output back-off and full power. As can be seen, drain efficiency in excess of 43% and 50% was achieved at 6-dB output back-off and peak power, respectively, in the frequency range of 1.5–2.5 GHz. This is equivalent to 50% fractional bandwidth. Also, the peak output power ranged from 45.0–47.2 dBm across the band. Figure 5.17 displays CW efficiency and gain as a function of output power. One can clearly observe the efficiency plateau in the 6-dB back-off power

107
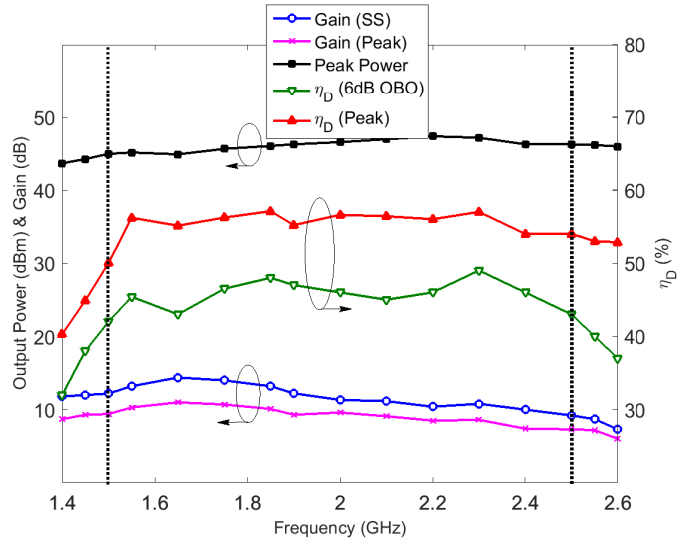
Figure 5.16: Measured CW performance of the WeDPA versus frequency.
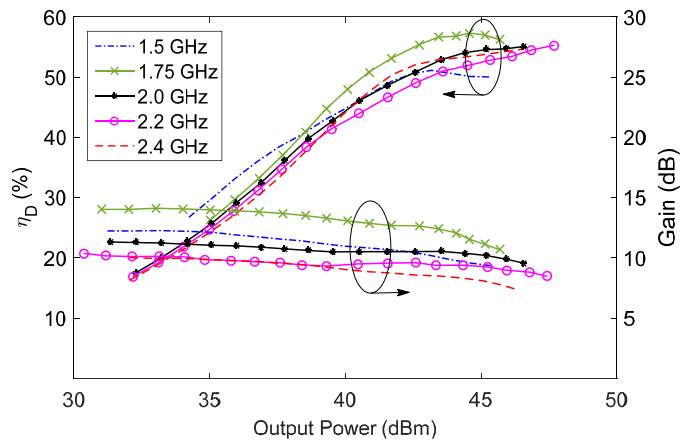
Figure 5.17: Measured CW drain efficiency and gain versus output power at various frequency points.

108
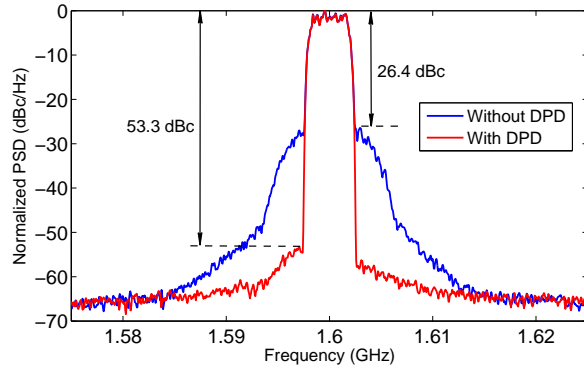
Figure 5.18: Normalized output PSD of the WeDPA with and without memoryless DPD when driven with 5 MHz WCDMA signal at 1.6 GHz.
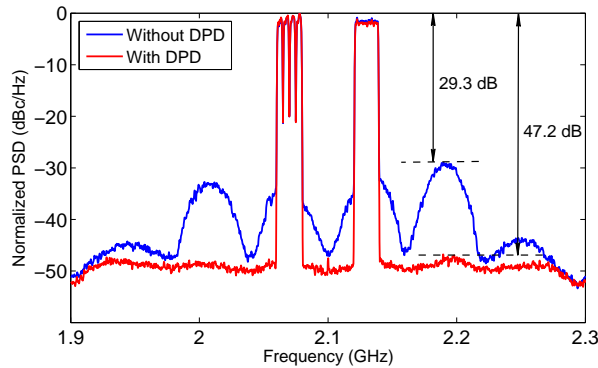
Figure 5.19: Normalized output PSD of the WeDPA with and without DPD when driven with an 80 MHz modulated signal at 2.1 GHz.

range, which verifies proper load modulation over the entire bandwidth.

The linearizabilty of the WeDPA was assessed under single-band, wideband, and multi-band modulated signals. Figure 5.18 displays the output power spectral density (PSD) of the WeDPA under a 5 MHz wideband code division multiple access (WCDMA) signal at 1.6 GHz with 7.1 dB PAPR. An ACLR of 53 dBc was achieved at 37.7 dBm output power using 7th order memoryless digital pre-distortion (DPD) with an average drain efficiency of 46%. The output spectrum of the WeDPA, driven by an 80 MHz signal with 9.4 dB PAPR centered at 2.1 GHz, is depicted in Figure 5.19. After applying the DPD, ACLR improved from 29.3 dBc to 47.2 dBc, while a drain efficiency of 41% was measured at 37.8 dBm output power. Furthermore, analog implementation of the WeDPA allowed

109

(a) Band I – 15-MHz WCDMA (101)

(b) Band II – LTE 15 MHz
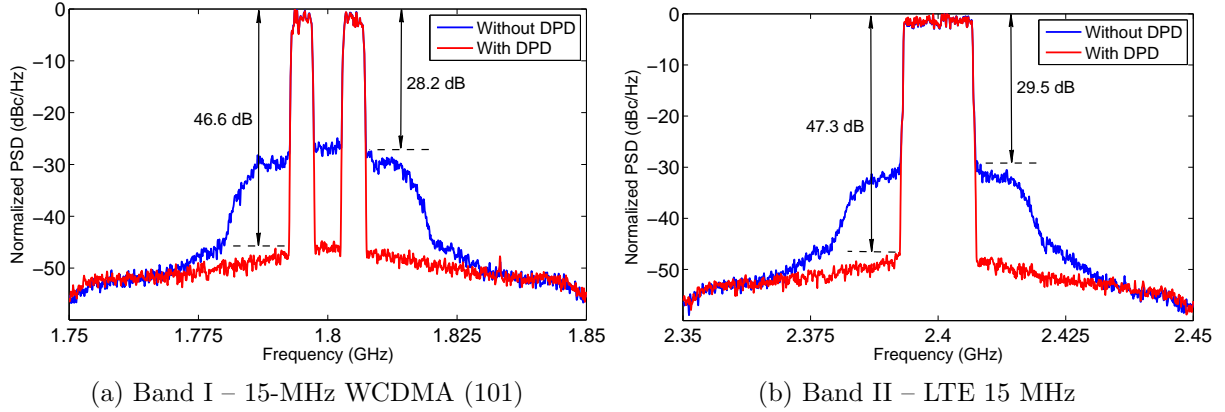
Figure 5.20: Normalized output PSD of the WeDPA with and without DPD when driven with a dual-band modulated signal.

Table 5.2: Benchmarks of Broadband DPAs

	Unit	[16]	[41]	[58][†]	[30][†]	**T.W.**
Freq.	GHz	0.7–1.0	1.6–2.25	1.7–2.4	1.0–2.5	**1.5–2.5**
FBW	%	35	34	36	83	**50**
P_{out}	dBm	49–50.8	51.8–53.1	39–41	40–42	**45–47**
η@6dB OBO	%	52–68	41–61	43–59	35–58	**43–50**
Mod. BW	MHz	20	3.84	20	5	**80**
PAPR	dB	10.5	7.3	9.6	5.4	**9.4**
η_{avg}	%	43	50	47[‡]	35	**41**

[†] GaN bare dies were used.
[‡] ACLR of -40 dBc was reported with DPD.

110

concurrent amplification of multiple modulated signals. To confirm this, the WeDPA was concurrently excited with 15 MHz WCDMA (101) and 15 MHz long-term evolution (LTE) signals at 1.8 GHz and 2.4 GHz, respectively (600 MHz spacing), with a combined PAPR of 9.1 dB. Using a dual-band two-dimensional DPD algorithm [59], the ACLR at the two bands improved to 46.6 dBc and 47.3 dBc (Figure 5.20), and the WeDPA achieved 32% drain efficiency at 37.0 dBm output power. Table 5.2 compares the performance of this WeDPA with state-of-the-art results from the literature. The FBW achieved in this work for a minimum acceptable 6-dB BO efficiency of 40% was the best reported to date among DPAs implemented using packaged transistors. Additionally, linear amplification of wideband and dual-band CA signals were not demonstrated in the references of Table 5.2

5.5   Conclusion

A novel approach to designing broadband DPAs was proposed. The concept of waveform engineering was extended to 2W-DPAs to accommodate complex-valued load modulation. Theoretical and practical aspects of a WeDPA were discussed, and a reformulation of the parameters of the output combining network at the fundamental and harmonic frequencies was provided. To validate the proposed theory, a fully analog RF_{in}/RF_{out} 50-W broadband DPA was developed using GaN packaged devices. The circuit demonstrator maintained more than 43% drain efficiency at 6-dB output power back-off over the frequency range of 1.5–2.5 GHz. The linearizability of the WeDPA was also successfully verified using wideband CA signals of up to 80 MHz, as well as inter-band CA signal stimuli.

Chapter 6

Input Matching Network Design for Broadband DPAs

6.1 Introduction

As discussed in Chapter 5, using waveform engineering techniques one can extend the bandwidth of the output combining network of the DPA to values comparable with single-ended PAs. However, this potentially wideband capability is achieved in reality only if the current profiles assumed in Figure 6.1 are maintained over the bandwidth. Since FETs are VCCSs by nature, proper current profiles need to be generated by providing an appropriate gate-source voltage, i.e., by presenting proper impedances to the gate of the transistors across the bandwidth. This job has to be fulfilled by judicious design of input matching networks (IMNs). Even though this is the case for any PA configuration, in single-ended PAs IMNs are designed primarily for a specific gain. In the case of DPAs, however, synthesizing the current profiles of the main and peaking PAs with minimum AM/AM and AM/PM are the target, and this imposes new challenges on IMN design, specially for peaking device operating in class C. Therefore, even if the output combing network is perfectly designed over a broad frequency range, without well-designed IMNs the Doherty performance will deteriorate, both in terms of efficiency and linearity. Hence, broadband IMN design is just as important as the output combining network, but has never been discussed in the literature to date.

This Chapter expounds input matching design challenges in broadband Doherty power amplifiers. The concept of current contours are then introduced which can be used to
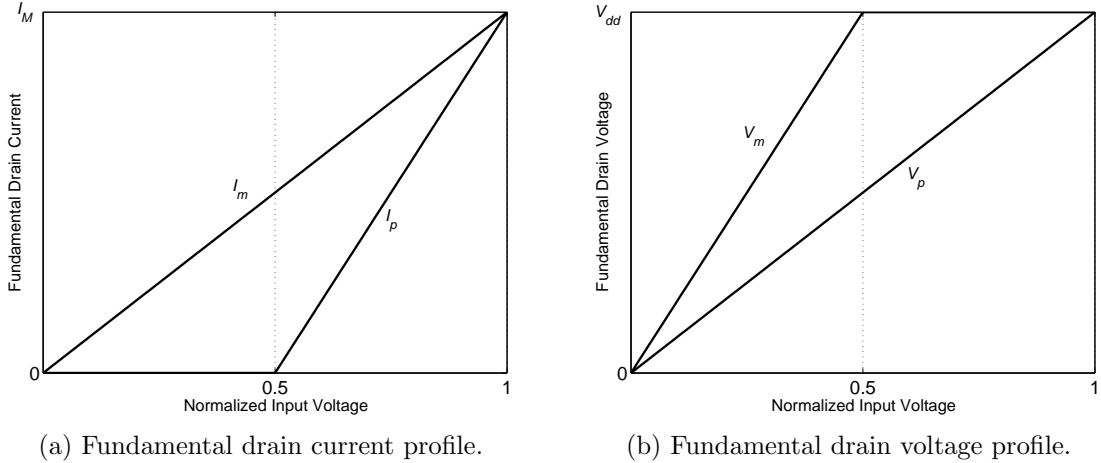
(a) Fundamental drain current profile.

(b) Fundamental drain voltage profile.

Figure 6.1: Fundamental current and voltage profiles in the classical DPA.

tackle the challenges. A systematic approach is then proposed for designing broadband input matching networks for the main and peaking devices in a DPA.

6.2 Broadband Input Matching Design Challenges

Design of IMNs for DPAs is more challenging than a single-ended PA since there are far more requirements for a proper Doherty operation other than just small-signal gain or minimum efficiency degradation versus frequency. First and foremost, the main and peaking PAs must operate harmoniously across the frequency range to maintain the linearity. The main IMN could be designed for a specific gain versus frequency, however, the peaking transistor is required to turn on deliver adequate current at peak power to properly modulate the main transistor's load impedance. Hence, the small-signal gain is meaningless in case of the peaking PA biased in class C.

IMN design is specially more challenging for GaN DPAs due to the strong nonlinearity of the input capacitance. As shown in Chapter 4, this nonlinearity generates AM/PM distortion which is bias-dependent, and hence different for the main and peaking paths. Therefore phase alignment of the two paths is only possible at one power level. This can greatly deteriorate linearity and efficiency as discussed in Chapter 4. Note that this problem is specific to DPAs and does not exist in a single-ended PAs. Furthermore, the input nonlinearity sensitizes the RF performance to the input second harmonic termination.

113

Hence, one should design IMNs not only for proper fundamental impedances, but for second harmonic ones as well. Additionally, the input splitter needs to cover at least fundamental frequencies up to the second harmonic band. This can be achieved by using multi-stage splitters (such as Wilkinson, etc.) to provide enough return loss and isolation between the two paths and minimize the insertion loss (as shown in Chapter 5).

6.3   Concept of Current Contours

This section introduces the concept of current contours which will highly simplify broadband IMN design, and combined with harmonic source-pull contours, form the basis of the proposed design methodology.

Current contours are defined as the locus of points in the Smith chart [$\Gamma_S(f_o)$ plane] that generate the same fundamental drain current at a given frequency and power level (similar to power and PAE contours obtained from source-pull analysis). These contours allow a designer to carefully control the performance of each transistor by mapping the drain current at a specific input power and frequency to the fundamental input impedance, thereby maintaining the current profiles required for Doherty operation. The current contours of class B and C biased transistors are described separately in the following sections.

6.3.1   Class-B Contours

The main PA operating in class-B (or deep AB) is normally biased for best linearity (in terms of AM/AM response). Hence, assuming a linear current profile (as depicted in Figure 6.1, the only important parameter is the slope of the profile or equivalently the current at peak envelope power (PEP). Knowing the peak input power, $I_{m,H}$ represents the minimum acceptable current from the main device at PEP [see Figure 6.2(a)]. An example of constant $I_{m,H}$ contour is illustrated in Figure 6.2(b). Any fundamental impedance selected on the contour will create a current $I_m = I_{m,H}$ at the specified power level. Note that inside (outside) the circle denotes current values higher (lower) than $I_{m,H}$.

6.3.2   Class-C Contours

For the peaking amplifier biased in class C, one possibility is to use an $I_{p,H}$ contour which is similar to the $I_{m,H}$ contour, and represents the minimum acceptable current from the

(a) Definition of $I_{m,H}$.

(b) An example of constant $I_{m,H}$ contour.
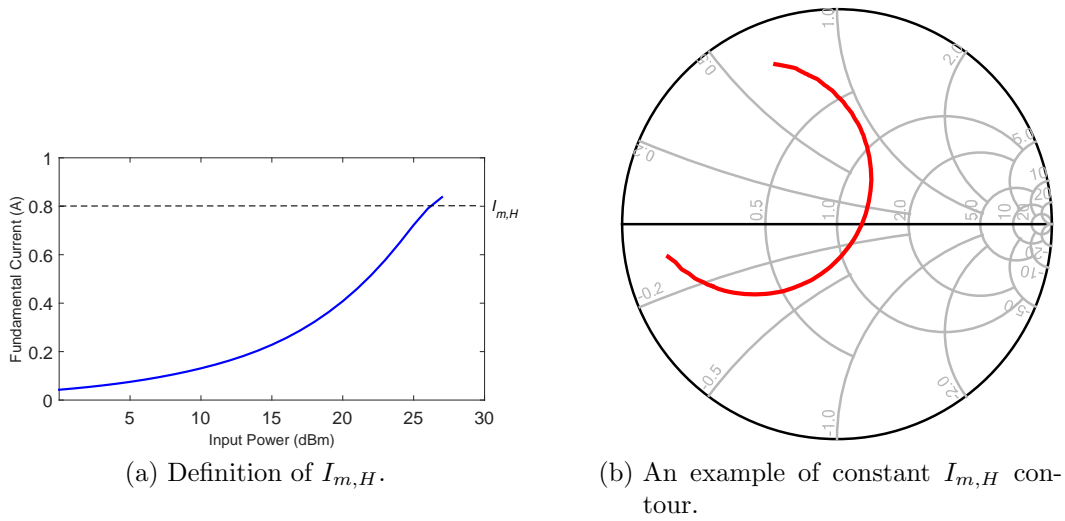
Figure 6.2: Illustration of $I_{m,H}$ contours.
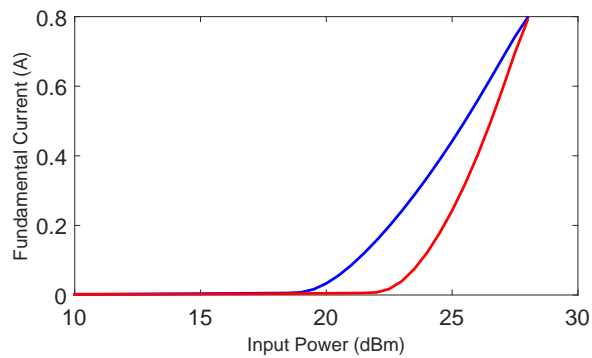
Figure 6.3: An example of input nonlinearity implications in GaN DPAs. Two different fundamental impedances can create similar currents at PEP but have dissimilar turn-on points.

(a) Definition of $I_{p,L}$ and $I_{p,H}$.

(b) An example of constant $I_{p,L}$ and $I_{p,H}$ contour.
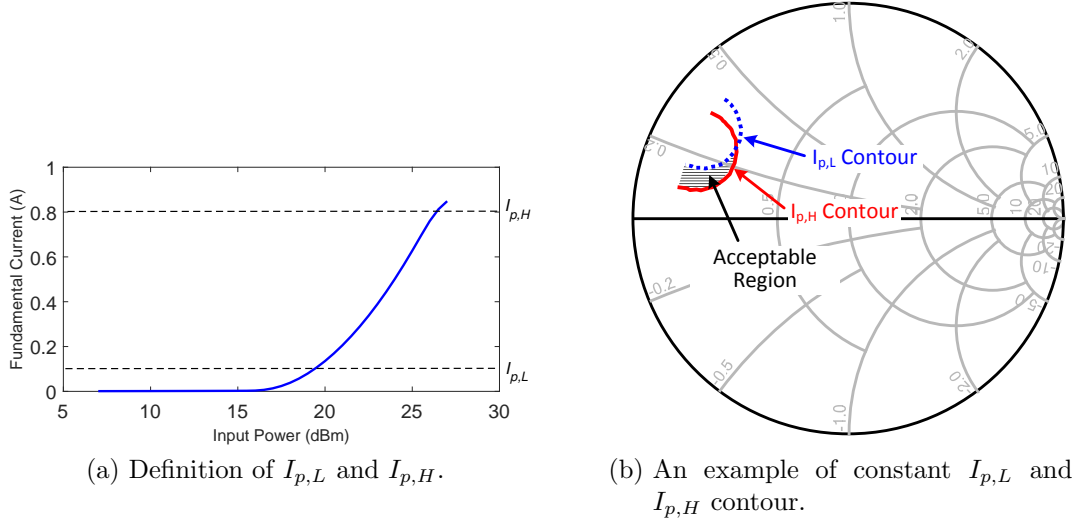
Figure 6.4: Illustration of $I_{p,L}$ and $I_{p,H}$ contours.

peaking device at PEP. However, the peak current is not sufficient to define the peaking current profile, since there is no control of the turn-on point. The current at 6 dB BO, if not carefully monitored, might be too large which will reduce the 6 dB BO efficiency. Furthermore, in the case of GaN PAs, due to the input nonlinearity, two different fundamental impedances with the same peak current might result in different currents at 6 dB BO. An example of such a case is illustrated in Figure 6.3. Note that in this case the harmonic impedances are the same and only fundamental impedances are different, resulting in different turn-on points despite the peak currents being the same.

To control the turn-on point, one can use $I_{p,L}$ contours denoting the maximum acceptable current at 6-dB BO [shown in Figure 6.4(a)]. Usually 10%–20% of the peak current is a suitable range to realize the matching network, specially when a broad bandwidth is targeted. Note that the area enclosed inside the constant $I_{p,H}$ contour ($I_p > I_{p,H}$) at PEP and outside the constant $I_{p,L}$ contour ($I_p < I_{p,L}$) at 6-dB BO is the desired region for the input fundamental impedance, as designated in Figure 6.4.

It is worth mentioning that the overlapping area highly depends on peak input power level (or equivalently the large-signal gain of the peaking PA) as well as the gate biasing point. Hence, maximizing the overlapping area by varying these two parameters can significantly increase the chances of successful IMN design.

116

6.4   Input Matching Design Methodology

Using the concept of current contours as a tool to carefully dictate the currents and the knowledge of the second harmonic termination effects discussed in Chapter 4, a systematic approach is proposed in this section to design broadband IMNs for the main and peaking transistors.

6.4.1   Class-B Matching Design

The steps for designing IMN of the main (class B) PA over a given bandwidth are as follows:

1. Terminate the output with the ideal load ($R_L = R_{opt}$ and short the harmonics) and bias the transistor's gate based on the requirements (usually for the best AM/AM linearity).

2. Use fundamental source-pull simulation at PEP to plot $I_{m,H}$ contours at multiple frequency points within the target bandwidth, assuming ideal harmonic terminations (short circuit) at the intrinsic gate. Note that generally these contours shrink and move counter-clockwise as frequency increases, which indicates that there is a limit on the PA gain beyond which the IMN design is impossible.

3. Determine a realizable fundamental impedance trajectory versus frequency $[Z_S(f_o)]$.

4. Using second harmonic source-pull power and efficiency contours at PEP, complete the impedance trajectory up to the second harmonic $[Z_S(2f_o)]$ by avoiding the sensitive regions on the Smith charts (as illustrated in Section 4.2.2).

5. Realize the impedance profile using synthesis methods such as the simplified real-frequency technique (SRFT) or filter theory.

Note that when the target bandwidth is larger than an octave (e.g. 1.5–3 GHz), the fundamental and second-harmonic frequency ranges will overlap. In this case, the best strategy is to start with the overlapping range (higher edge of the band, i.e., 3 GHz) and choose $Z_S(f_o)$ according to the second-harmonic source-pull contours at $f = f_o/2$ (i.e. 1.5 GHz) together with the $I_{m,H}$ contours.

Figure 6.5(a) shows an example of IMN design for a main (class-B) PA using 15-W GaN HEMT die over the 1–3 GHz frequency range. The impedance trajectory realized
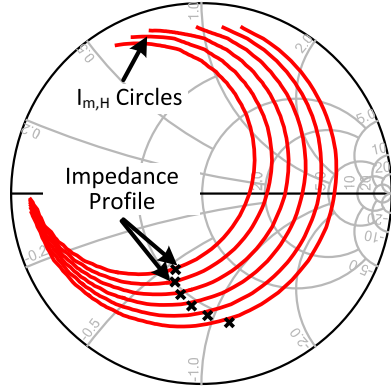
(a) Constant current contours and the input impedance trajectory.
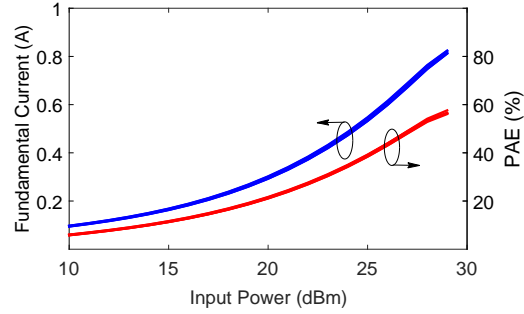
(b) Drain current and efficiency profiles vs. frequency.

Figure 6.5: An example of IMN design for main PA using constant current contours.

using a bandpass topology designed by the SRFT tool in Keysight Advanced Design System (ADS^TM) is shown along with the $I_{m,H} = 0.8\ A$ contours. Figure 6.5(b) depicts the current, gain, and efficiency profiles at different frequencies. As can be seen, the current profile as well as a flat gain are maintained across the bandwidth, and RF performance degradation attributed to the IMN is minimized taking advantage of the developed methodology.

6.4.2 Class-C Matching Design

Similar to the class-B PA, the following steps should be followed for designing a broadband IMN for the peaking (Class C) PA:

1. Starting with an initial gate bias value, $V_{GS,p}$, terminate the output with the ideal load ($R_L = R_{opt}$ and short the harmonics).

2. Use fundamental source-pull simulation to plot $I_{p,L}$ contours at 6-dB BO and $I_{p,H}$ contours at PEP at multiple frequency points within the target bandwidth, assuming ideal harmonic terminations (short circuit) at the intrinsic gate.

3. Designate the overlapping region between $I_{p,L}$ and $I_{p,H}$ contours at each frequency (see Figure 6.4). The design space (overlapping area) can be increased by varying $V_{GS,p}$ and the gain at PEP.

(a) Constant current contours and the input impedance trajectory.

(b) Drain current and efficiency profiles vs. frequency.
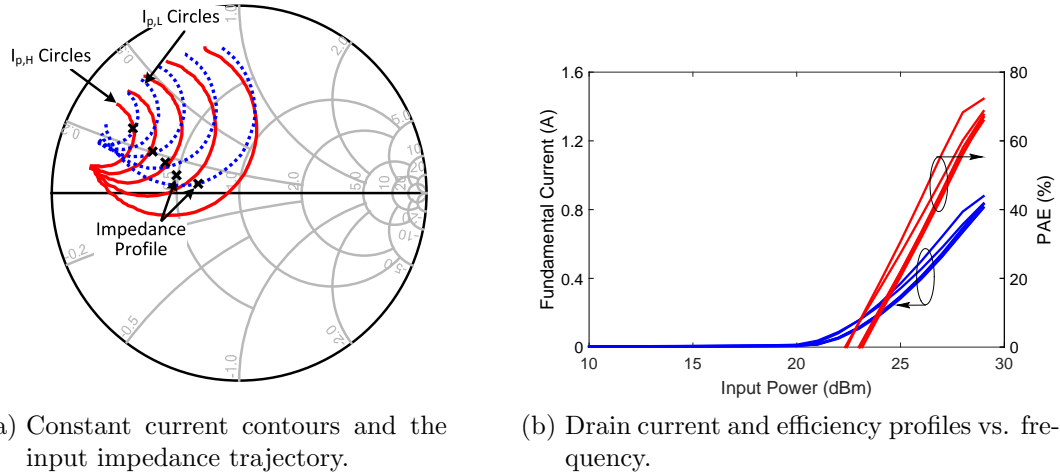
Figure 6.6: An example of IMN design for main PA using constant current contours.

4. Determine a realizable fundamental impedance trajectory versus frequency $[Z_S(f_o)]$.

5. Using second harmonic source-pull power and efficiency contours at PEP, complete the impedance trajectory up to the second harmonic $[Z_S(2f_o)]$ by avoiding the sensitive regions on the Smith charts (as illustrated in Section 4.2.2).

6. Realize the impedance profile using circuit synthesis methods such as simplified real-frequency technique (SRFT) or filter theory.

Figure 6.6(a) shows an example of IMN design for a peaking (class C) PA using 30-W GaN HEMT die over 1–3 GHz frequency range. The impedance trajectory realized using a bandpass topology designed by The SRFT tool in Keysight ADS^{TM}is shown together with $I_{p,L} = 0.15\ A$ and $I_{p,H} = 0.8\ A$ contours. Figure 6.6(b) displays the current and efficiency profiles at various frequencies. Once again, consistent current profile versus frequency with high efficiency at PEP is maintained across the bandwidth, verifying the suitability of the proposed methodology.

It should be noted that due to the intrinsically lower transconductance of the peaking PA compared to the main PA at peak power (even assuming a peaking device twice larger than the main one), it is usually advised to design the peaking IMN first since it is the bottleneck in most cases, in terms of gain flatness and current profile consistency across the target frequency range. The main PA's IMN can be designed afterwards knowing the gain of the peaking PA at peak power.

119
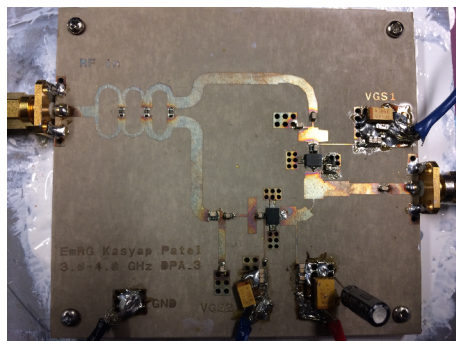
Figure 6.7: Fabricated DPA using the proposed methodology [1].

6.4.3 Bilateral Effect

In the methodology presented in Section 6.3 an ideal load impedance was assumed. In practice, specially when using the WeDPA theory, the output load impedance can be complex and the harmonics are not necessarily shorted. If the transistor is unilateral, changing the output impedance should not affect the source-pull contours (including the current contours). However, due to the feedback capacitance of the device, C_{gd}, RF transistors are bilateral, particularly at frequencies above 2–3 GHz.

To take this effect into account, an iterative approach is needed to co-design the input and output matching networks. A better starting point for such a design procedure could be using the actual load impedance (at fundamental and harmonics) designed over the given frequency range and then following steps 2–5 to complete the first iteration. Then the output combining network can be tweaked for optimal performance and this procedure is repeated. It is worth mentioning that in most cases two iterations is sufficient to obtain an optimal performance over the frequency range.

6.5 Validation

To validate the concept of current contours and the IMN design methodology proposed in this chapter, a broadband DPA prototype was fabricated using two 6-W 0.25 μm packaged GaN HEMTs from Cree® for both the main and peaking transistors [1]. The required input impedances versus frequency were tabulated following the design methodology presented in Section 6.4 and an IMN topology was constructed using the SRFT tool in ADS™. The resultant lumped components were then transformed to a transmission line structure. The

Table 6.1: Comparison of published DPA literature.

Year	Frequency (GHz)	Gain (dB)	P_{out} (dBm)	$\eta_{(6db)}$ min/max (%)
2012 [22]	3–3.6	10	43	38/43
2013 [17]	1.6–2.4	9	42	50/60
2014 [30]	1.05–2.55	8	41	35/58
This Work	**2.7–4.7**	**8**	**41**	**37/47**

output combining network was designed using the broadband transformer-based structure proposed in [29]. A photograph of the fabricated DPA is shown in Figure 6.7. The drain bias for both devices is applied through a low-impedance short-circuit stub included in the output combiner to reduce the drain-induced memory effects, thereby improving linearizability of the DPA as demonstrated in Section 4.3.3. Further details on design and implementation of the DPA can be found in [1].

Figure 6.8 displays the CW measurement results versus frequency. As can be seen, a 6-dB back-off drain efficiency of greater than 37% is achieved over 2.7–4.7 GHz, equivalent to 54% fractional bandwidth. Also, the linearizability of the DPA demonstrator was assessed under single-band and carrier-aggregated modulated signals. Figure 6.9 shows the output spectrum of the DPA when driven with an 80 MHz carrier-aggregated signal with a PAPR of 9.4 dB at 3.3 GHz. After DPD, the ACLR improved from –37 dB to –48 dBc at 30.4 dBm output power. Table 6.1 compares the performance of this DPA to state-of-the-art DPAs. As can be seen, using the proposed current contour methodology similar performance to previous DPAs has been achieved however at a much higher frequency. Using the WeDPA concept proposed in Chapter 5 for broadband output combining network design, the bandwidth can be extended even further.
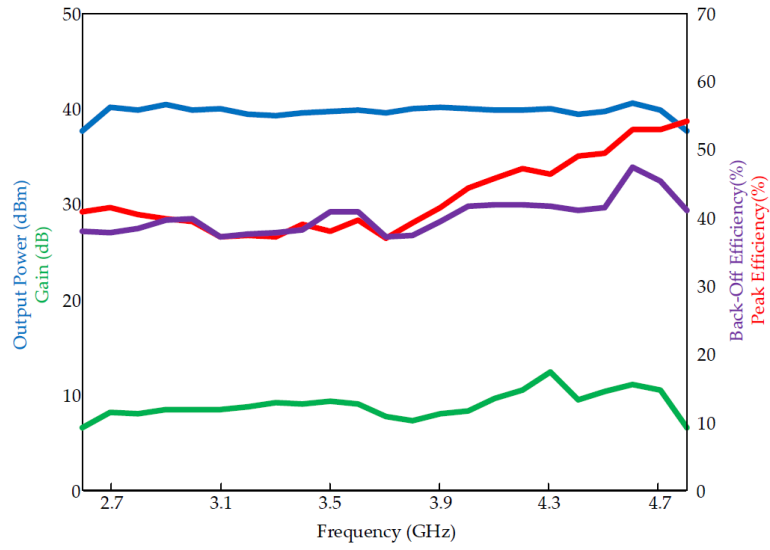
(a)
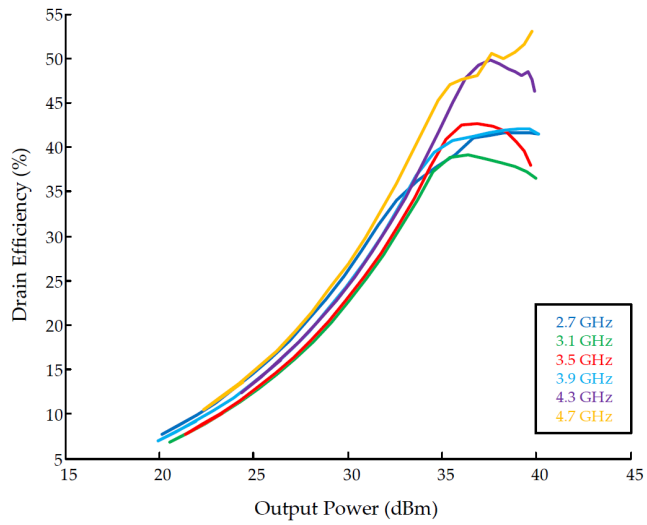
(b)

Figure 6.8: Measured CW performance of the DPA versus (a) frequency, and (b) output power [1].
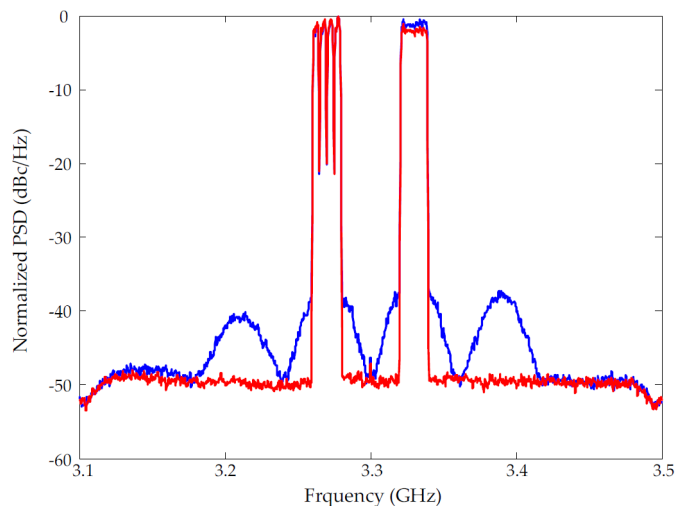
Figure 6.9: Measured output spectrum of the DPA under a 80-MHz modulated signal at
3.3 GHz before (blue) and after (red) DPD [1].

6.6 Conclusion

This chapter proposed a systematic approach to designing IMNs for broadband DPAs.
The output combining network can be designed over a broad frequency range (one octave
or more) using novel techniques proposed in numerous publications, but the challenges of
the IMN design for broadband DPAs remained unaddressed. The requirements of IMNs
for the main PA (biased in class B) and peaking PA (biased in class C) were discussed in
detail. The concept of current contours were introduced as a tool that can be utilized to
systematically design broadband IMNs. Using this tool as well as the linearity analysis
conducted in Chapter 4, a design methodology was proposed for IMNs of the main and
peaking transistors in a 2W-DPA structure. This methodology helps a designer use classical
circuit synthesis techniques in order to achieve consistent current profiles with acceptable
efficiency and linearity for a given bandwidth, instead of relying on blind optimization
of circuit components which does not guarantee the optimum or even acceptable results.
The proposed technique was validated by designing a broadband DPA prototype over 2.7–
4.7 GHz frequency range. The fabricated DPA achieved higher than 37% drain efficiency
and 40 dBm peak output power across the frequency band and was successfully linearized
under wideband modulated signals.

The proposed technique only relies on the required current profiles and hence can be
readily used for IMN design of multi-way DPAs as well.

Chapter 7

Conclusion

Evolution of modern wireless communication networks toward data-centric services has resulted in the emergence of signals with high PAPR, stringent linearity requirements, and large modulation bandwidth. Upcoming wireless standards will utilize carrier aggregated signals to further increase the transmission bandwidth. Hence, RF PAs are required to efficiently and linearly amplify multi-band modulated signals with high PAPR scattered over a broad frequency range. To improve the PA efficiency under high PAPR signals, the DPA technique was selected in this work as a promising solution for multi-standard multi-band base stations. Although DPAs have been primarily used for narrowband scenarios, recent attempts have been made to extend the bandwidth of DPAs to address the challenges of high efficiency multi-band, multi-standard radio systems. Recent attempts made in the literature toward extending the bandwidth of 2W-DPAs achieved $\sim$30% fractional bandwidth over 6-dB back-off power range. This range is not necessarily sufficient for amplifying multi-band modulated signals used in wireless communication systems. 3W-DPAs introduced in the literature targeting higher back-off range (9–12 dB) were implemented only for narrowband scenarios.

The main goal of this thesis was to develop a methodical approach to designing broadband DPAs capable of concurrent amplification of multi-band signals with high PAPR, targeting inter-band and intra-band, non-contiguous scenarios of carrier-aggregated signals. The thesis began with theoretical bandwidth analysis of 2W- and 3W-DPAs using a simple transistor model. This analysis revealed significant bandwidth capability of 3W-DPAs using a novel formulation derived for these DPAs. A proof-of-concept 3W-DPA was designed and fabricated using packaged GaN transistors to improve efficiency over 9-dB OBO power range for 730–980 MHz frequency band (equivalent to 29% fractional

bandwidth). To the author's best knowledge, this was the first broadband 3W-DPA in literature.

Next, bandwidth extension of DPAs to 50% and higher was targeted. As a first step, to ensure linearity and linearizability of broadband 2W-DPAs for multi-band and multi-standard wireless applications, the sources of static and dynamic distortion were investigated both at device and circuit level. This study provided guidelines and methodologies to minimize the sources of distortion from the circuit design stage, rather than designing DPAs merely for maximum efficiency over bandwidth and relying on complex DPD algorithms to correct for nonlinearities, which often fail in concurrent multi-band scenarios.

To further enlarge the bandwidth of 2W-DPA to 50% and beyond, the concept of waveform engineering was proposed in this thesis to reformulate the DPA's output combining network for complex-valued load modulation. This was the main contribution of this work since it significantly broadened the design space and relaxed the Doherty operation requirements, thereby enhancing the achievable bandwidth. A 50-W 2W-DPA prototype was designed using packaged GaN transistors to validate the developed theory. This prototype was successfully linearized under intra- and inter-band carrier-aggregated signals over 1.5–2.5 GHz frequency range. The DPA formulation for complex load modulation was presented for the fist time in this work.

Finally, to address the challenges of broadband IMN design for DPAs, a systematic approach was devised for the first time using the knowledge acquired through studying the sources of distortion in DPAs. Following the proposed IMN design methodology, minimizes the efficiency/linearity degradation across the bandwidth caused by input nonlinearity in GaN HEMTs.

7.1 Summary of Contributions

The key contributions of this thesis can be summarized as follows:

- Developing a novel formulation for 3W-DPAs that maximizes efficiency versus bandwidth over a wide back-off range.

- Detailed analysis of the static and dynamic sources of distortion in 2W-DPAs from device and circuit perspectives and devising circuit techniques to mitigate them, thereby improving linearizability of the 2W-DPAs.

- Developing a waveform engineering approach to designing broadband 2W-DPAs which relaxes the output combiner requirements in terms of fundamental and harmonic matching.

- Developing a systematic IMN approach for broadband DPAs that takes into account the device non-idealities and minimizes efficiency/linearity degradation versus frequency.

7.2   Future Work

The work presented in this thesis can be extended to meet the requirements of next generation (e.g. 5G) base-station PAs. For instance, the suitability of the 2W- and 3W-DPAs proposed in this thesis for higher microwave frequencies (6–30 GHz) as well as millimeter-wave frequencies (30–100 GHz) needs to be investigated. Integration technologies using III-V semiconductors, including GaN and GaAs MMICs can be good candidates for up to 30 GHz frequency range. In addition, advancements in silicon-based technologies, such as bulk CMOS, silicon-on-insulator (SOI), and SiGe heterojunction bipolar transistor (HBT), have enabled high frequency operations up to and above 100 GHz by scaling the gate length (emitter width in HBTs) down to nano-meter dimensions. However, the Q-factor of the passive devices is much lower in Si-based technologies compared to the devices used in this thesis, and even GaN/GaAs MMICs. This combined with low power density and breakdown voltage of CMOS transistors, highly complicates realization of high efficiency and broadband DPAs at mm-wave. Integration of DPA would make it an attractive solution for femto- and pico-cell base stations or even mobile devices. These applications have less stringent linearity requirement compared to macro- and micro-cells (discussed in this thesis). Nevertheless, complicated DPD models (such as the ones used in this thesis) cannot be applied due to the power overhead of the DPD engine. Using a transformer to compensate for the device's output capacitance and reduce the size and complexity of the output combiner is a potential solution for mm-wave DPAs.

Another interesting topic for future would be extending the WeDPA theory presented in Chapter 4 to continuous modes other than class B/J design space (used in this work) to further improve the back-off efficiency, or even extending it to three-way DPAs to enhance efficiency over a wide back-off range and much broader fractional bandwidth than what was reported in Chapter 3.

Lastly, the circuit-level analysis of static distortions as well as memory effects in 2W-DPAs (discussed in Chapter 4) can be extended to 3W-DPAs to ensure their linearizabilty

for multi-standard and multi-band applications, while maintaining high efficiency over a broad FBW and power range.

7.3   List of Publications

Peer-Reviewed Journal Papers

1. **H. Golestaneh**, F. Arfaei Malekzadeh, and S. Boumaiza, "an extended bandwidth three-way Doherty power amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 9, pp. 3318–3328, Sep. 2013. (Based on Chapter 3)

2. S. Boumaiza, **H. Golestaneh**, and M. Naseri, Multispectrum signal transmitters, *IEEE Microwave Magazine* (invited paper), vol. 15, no. 7, pp. S14-S24, Nov. 2014.

3. **H. Golestaneh**, and S. Boumaiza, "A waveform engineering approach to the design of broadband Doherty power amplifiers," Revised version submitted to *IEEE Transactions on Microwave Theory and Techniques*, Oct. 2015. (Based on Chapter 5)

4. K. Patel, **H. Golestaneh**, and S. Boumaiza, "Current contour based IMN design methodology for broadband GaN Doherty power amplifiers," Submitted to *IEEE Transactions on Microwave Theory and Techniques*, Oct. 2015. (Based on Chapter 6)

5. M. Naseri Ali Abadi, **H. Golestaneh**, S. Boumaiza, "Doherty power amplifier with extended bandwidth and improved linearizability under carrier-aggregated signal stimuli," Submitted to *IEEE Microwave and Wireless Components Letters*, Oct. 2015.

Refereed Conference Papers

1. **H. Golestaneh**, F. Arfaei Malekzadeh, and S. Boumaiza, "Three-way Doherty power amplifier for efficient amplification of wideband signals with extended PAPR," in*IEEE Radio and Wireless Symposium*, Newport Beach, CA, Jan. 2014, pp. 136–138. (Based on Chapter 3)

2. S. Boumaiza, **H. Golestaneh**, M. Naseri, A. M. M. Mohamed, and D. Wu, "Broadband Doherty power amplifiers: advances and challenges," in *IEEE WAMICON* (invited paper), Tampa, FL, Jun. 2014.

3. M. Naseri Ali Abadi, **H. Golestaneh**, H. Sarbishaei, and S. Boumaiza,"An extended bandwidth Doherty power amplifier using a novel output combiner," in *IEEE MTT-S International Microwave Symposium Digest*, Tampa, FL, Jun. 2014.

References

[1] K. Patel, "Current contour based design methodology for IMN design of Doherty power amplifiers," Master's thesis, University of Waterloo, 2015. [Online]. Available: https://www.uwspace.uwaterloo.ca/bitstream/handle/10012/9669/Patel_Kasyap.pdf?sequence=1

[2] P. H. Aaen, J. A. Pla, and J. Wood, *Characterization and modeling of LDMOS power FETs for RF power amplifier applications.* Cambridge University Press, 2007.

[3] L. Kahn, "Single-sideband transmission by envelope elimination and restoration," *Proceedings of IRE*, vol. 40, no. 7, pp. 803–806, Jul. 1952.

[4] I. Kim, Y. Y. Woo, J. Kim, J. Moon, J. Kim, and B. Kim, "High-efficiency hybrid EER transmitter using optimized power amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 11, pp. 2582–2593, Nov. 2008.

[5] D. Kim, D. Kang, J. Choi, J. Kim, Y. Cho, and B. Kim, "Optimization for envelope shaped operation of envelope tracking power amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 7, pp. 1787–1795, Jul. 2011.

[6] J. Jeong, D. Kimball, M. Kwak, P. Draxler, C. Hsia, C. Steinbeiser, T. Landon, O. Krutko, L. Larson, and P. Asbeck, "High-efficiency WCDMA envelope tracking base-station amplifier implemented with GaAs HVHBTs," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 10, pp. 2629–2639, Oct. 2009.

[7] J. H. Kim, G. D. Jo, J. H. Oh, Y. H. Kim, K. C. Lee, J. H. Jung, and C. S. Park, "High-efficiency envelope-tracking transmitter with optimized class-F^{-1} amplifier and 2-bit envelope amplifier for 3G LTE base station," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 6, pp. 1610–1621, Jun. 2011.

[8] F. Raab, "High-efficiency linear amplification by dynamic load modulation," in *IEEE MTT-S International Microwave Symposium Digest*, vol. vol.3, Philadelphia, PA, USA, Jun. 2003, pp. 1717–1720.

[9] H. M. Nemati, C. Fager, U. Gustavsson, R. Jos, and H. Zirath, "Design of varactor-based tunable matching networks for dynamic load modulation of high power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 5, pp. 1110–1118, May 2009.

[10] C. M. Andersson, D. Gustafsson, K. Yamanaka, E. Kuwata, H. Otsuka, M. Nakayama, Y. Hirano, I. Angelov, C. Fager, and N. Rorsman, "Theory and design of class-J power amplifiers with dynamic load modulation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 12, pp. 3778–3786, Dec. 2012.

[11] K. Chen and D. Peroulis, "Design of adaptive highly efficient GaN power amplifier for octave-bandwidth application and dynamic load modulation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 6, pp. 1829–39, Jun. 2012.

[12] W. Doherty, "Power amplifier for modulated waves," *Proceedings of IRE*, vol. 24, pp. 1163–1182, Sep. 1936.

[13] S. Cripps, *Advanced Techniques in RF Power Amplifier Design*. Artech House, 2002.

[14] J. Kim, J. Cha, I. Kim, and B. Kim, "Optimum operation of asymmetrical-cells-based linear Doherty power amplifiers-uneven power drive and power matching," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 5, pp. 1802–1809, May 2005.

[15] M. Iwamoto, A. Williams, P.-F. Chen, A. Metzger, L. Larson, and P. Asbeck, "An extended Doherty amplifier with high efficiency over a wide power range," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 12, pp. 2472–2479, Dec. 2001.

[16] D. Wu and S. Boumaiza, "A modified Doherty configuration for broadband amplification using symmetrical devices," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 10, pp. 3201–3213, Oct. 2012.

[17] D. Gustafsson, C. M. Andersson, and C. Fager, "A modified Doherty power amplifier with extended bandwidth and reconfigurable efficiency," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 1, pp. 533–542, Jan. 2013.

[18] J. Qureshi, L. Nan, E. Neo, F. van Rijs, I. Blednov, and L. C. N. de Vreede, "A wide-band 20 W LDMOS Doherty power amplifier," in *IEEE MTT-S International Microwave Symposium Digest*, Anaheim, CA, USA, May 2010, pp. 1504–1507.

[19] K. Bathich, A. Z. Markos, and G. Boeck, "Frequency response analysis and bandwidth extension of the Doherty amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 4 PART 1, pp. 934–944, Apr. 2011.

[20] D. Kang, D. Kim, Y. Cho, B. Park, J. Kim, and B. Kim, "Design of bandwidth-enhanced Doherty power amplifiers for handset applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 12, pp. 3474–3483, Dec. 2011.

[21] G. Sun and R. H. Jansen, "Broadband Doherty power amplifier via real frequency technique," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 1, pp. 99–111, Jan. 2012.

[22] J. M. Rubio, J. Fang, V. Camarchia, R. Quaglia, M. Pirola, and G. Ghione, "3-3.6-GHz wideband GaN Doherty power amplifier exploiting output compensation stages," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 8, pp. 2543–2548, Aug. 2012.

[23] K. Bathich and G. Boeck, "Wideband harmonically-tuned GaN Doherty power amplifier," in *IEEE MTT-S International Microwave Symposium Digest*, Montreal, QC, Canada, Jun. 2012.

[24] Y. Yang, J. Cha, B. Shin, and B. Kim, "A fully matched N-way Doherty amplifier with optimized linearity," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 3, pp. 986–993, Mar. 2003.

[25] W. C. E. Neo, J. 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 Transactions on Microwave Theory and Techniques*, vol. 55, no. 5, pp. 866–878, May 2007.

[26] I. Kim, J. Moon, S. Jee, and B. Kim, "Optimized design of a highly efficient three-stage Doherty pa using gate adaptation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 10, pp. 2562–2574, Oct. 2010.

[27] 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 Transactions on Microwave Theory and Techniques, vol. 53, no. 3, pp. 852–860, Mar. 2005.

[28] J. Qureshi, W. Sneijers, R. Keenan, L. C. N. de Vreede, and F. van Rijs, "A 700-W peak ultra-wideband broadcast Doherty amplifier," in *IEEE MTT-S International Microwave Symposium Digest*, Tampa, FL, USA, Jun. 2014.

[29] D.-T. Wu, J. Annes, M. Bokatius, P. Hart, E. Krvavac, and G. Tucker, "A 350 W, 790 to 960 MHz wideband LDMOS Doherty amplifier using a modified combining scheme," Tampa, FL, USA, Jun. 2014.

[30] R. Giofre, L. Piazzon, P. Colantonio, and F. Giannini, "A closed-form design technique for ultra-wideband Doherty power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 12, pp. 3414–3424, Dec. 2014.

[31] W. Chen, S. Bassam, X. Li, Y. Liu, K. Rawat, M. Helaoui, F. Ghannouchi, and Z. Feng, "Design and linearization of concurrent dual-band Doherty power amplifier with frequency-dependent power ranges," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 10, pp. 2537–2546, Oct. 2011.

[32] P. Saad, P. Colantonio, L. Piazzon, F. Giannini, K. Andersson, and C. Fager, "Design of a concurrent dual-band 1.8-2.4-GHz GaN-HEMT Doherty power amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 6, pp. 1840–1849, Jun. 2012.

[33] A. Grebennikov and J. Wong, "A dual-band parallel Doherty power amplifier for wireless applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 10, pp. 3214–3222, Oct. 2012.

[34] R. Gajadharsing, W. C. E. Neo, M. J. Pelk, L. C. N. de Vreede, and J. Zhao, "3-way Doherty amplifier with minimum output network," U.S. Patent 8 022 760, Jul. 2, 2009.

[35] J. Choi, D. Kang, D. Kim, and B. Kim, "Optimized envelope tracking operation of Doherty power amplifier for high efficiency over an extended dynamic range," *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 6, pp. 1508–1515, Jun. 2009.

[36] J. Moon, J. Son, J. Kim, I. Kim, S. Jee, Y. Y. Woo, and B. Kim, "Doherty amplifier with envelope tracking for high efficiency," in *IEEE MTT-S International Microwave Symposium Digest*, Anaheim, CA, USA, 2010, pp. 1086–1089.

[37] A. Mahmoud Mohamed, S. Boumaiza, and R. Mansour, "Reconfigurable Doherty power amplifier for multifrequency wireless radio systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 4, pp. 1588–98, Apr. 2013.

[38] D. Y.-T. Wu and S. Boumaiza, "A mixed-technology asymmetrically biased extended and reconfigurable Doherty amplifier with improved power utilization factor," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 5, pp. 1946–1956, May 2013.

[39] A. Grebennikov, "A high-efficiency 100-W four-stage Doherty GaN HEMT power amplifier module for WCDMA systems," in *IEEE MTT-S International Microwave Symposium Digest*, Baltimore, MD, USA, Jun. 2011.

[40] M. Akbarpour, M. Helaoui, and F. M. Ghannouchi, "A transformer-less load-modulated (TLLM) architecture for efficient wideband power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 9, pp. 2863–2874, Sep. 2012.

[41] K. Bathich, M. Arnous, and G. Boeck, "Design of 200 W wideband Doherty amplifier with 34% bandwidth," in *Eur. Microw. Conf.*, Nuremberg, Germany, Oct. 2013, pp. 279–282.

[42] H. Golestaneh, F. Arfaei Malekzadeh, and S. Boumaiza, "An extended-bandwidth three-way Doherty power amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 9, pp. 3318–3328, Sep. 2013.

[43] F. Mkadem, S. Boumaiza, J. Staudinger, and J. Wood, "Systematic pruning of Volterra series using Wiener G-functionals for power amplifier and predistorter modeling," in *Proc. IEEE 6th European Microwave Integrated Circuit Conference*, Manchester, UK, Oct. 2011, pp. 482–485.

[44] S. Cripps, P. Tasker, A. Clarke, J. Lees, and J. Benedikt, "On the continuity of high efficiency modes in linear RF power amplifiers," *IEEE Microwave and Wireless Components Letters*, vol. 19, no. 10, pp. 665–667, Oct. 2009.

[45] R. Smith, J. Lees, P. Tasker, J. Benedikt, and S. Cripps, "A design methodology for the realization of multi-decade baluns at microwave frequencies," in *IEEE MTT-S International Microwave Symposium Digest*, Baltimore, MD, USA, Jun. 2011.

[46] V. Carrubba, A. Clarke, M. Akmal, J. Lees, J. Benedikt, P. Tasker, and S. Cripps, "On the extension of the continuous class-F mode power amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 5, pp. 1294–303, May 2011.

[47] H. Sarbishaei, "Concurrent multi-band envelope tracking power amplifiers for emerging wireless communications," Ph.D. dissertation, University of waterloo, 2014. [Online]. Available: https://uwspace.uwaterloo.ca/bitstream/handle/10012/8669/Sarbishaei_Hassan.pdf?sequence=1

[48] S. Cripps, *RF Power Amplifiers for Wireless Communications*. Artech House, 2006.

[49] P. Wright, J. Lees, J. Benedikt, P. Tasker, and S. Cripps, "A methodology for realizing high efficiency class-J in a linear and broadband PA," *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 12, pp. 3196–204, Dec. 2009.

[50] D. Y.-T. Wu, F. Mkadem, and S. Boumaiza, "Design of a broadband and highly efficient 45W GaN power amplifier via simplified real frequency technique," in *IEEE MTT-S International Microwave Symposium Digest*, Anaheim, CA, USA, May 2010, pp. 1090–1093.

[51] X. Fu, D. T. Bespalko, and S. Boumaiza, "Novel dual-band matching network for effective design of concurrent dual-band power amplifiers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 1, pp. 293–301, Jan. 2014.

[52] D. Wu, "A Doherty power amplifier with extended bandwidth and reconfigurable back-off level," Ph.D. dissertation, University of Waterloo, 2013. [Online]. Available: https://uwspace.uwaterloo.ca/bitstream/handle/10012/7385/Wu_Yu-TingDavid.pdf?sequence=1

[53] S. Boumaiza, H. Golestaneh, and M. Naseri Ali Abadi, "Multispectrum signal transmitters," *IEEE Microwave Magazine*, vol. 15, no. 7, pp. S14–S24, Nov. 2014.

[54] M. Naseri Ali Abadi, H. Golestaneh, H. Sarbishaei, and S. Boumaiza, "An extended bandwidth Doherty power amplifier using a novel output combiner," Tampa, FL, USA, Jun. 2014.

[55] H. Sarbishaei, D.-T. Wu, and S. Boumaiza, "Linearity of GaN HEMT RF power amplifiers - a circuit perspective," Montreal, QC, Canada, Jun. 2012.

[56] L. Cotimos Nunes, P. Cabral, and J. Pedro, "AM/AM and AM/PM distortion generation mechanisms in Si LDMOS and GaN HEMT based RF power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 4, pp. 799–809, Apr. 2014.

[57] L. Nunes, P. Cabral, and J. Pedro, "AM/PM distortion in GaN Doherty power amplifiers," Tampa, FL, USA, Jun. 2014.

[58] L. Piazzon, R. Giofre, P. Colantonio, and F. Giannini, "A wideband Doherty architecture with 36% of fractional bandwidth," *IEEE Microwave and Wireless Components Letters*, vol. 23, no. 11, pp. 626–628, Nov. 2013.

[59] S. Bassam, M. Helaoui, and F. Ghannouchi, "2-D digital predistortion (2-D-DPD) architecture for concurrent dual-band transmitters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 10, pp. 2547–2553, Oct. 2011.

APPENDICES

Appendix A

Effect of Nonlinear Input Capacitance on AM/AM and AM/PM

Figure A.1 shows a simplified equivalent circuit model of an RF PA, that can be used to approximate the AM/AM and AM/PM caused by nonlinear input capacitance. First, assume that the effect of feedback capacitance, C_{gd}, is negligible compared to C_{gs}. One can write KCL at ω_o at the input node using (4.5), (4.8), and (4.9) to obtain,

$$
(V_S - \mathbf{V_i}) \cdot Y_S(\omega_o) = \mathbf{I_{gs,f_o}}
$$
$$
= j\omega_o \mathbf{V_i} \left(c_{i1} + \frac{3}{4} c_{i3} V_i^2 \right) \tag{A.1}
$$

where $\mathbf{V_i} = V_i \cdot e^{j\phi_i}$. Solving for $\mathbf{V_i}$) by equating the real and imaginary parts of both sides, one can calculated V_i and ϕ_i as functions of V_S. However, since it requires a tedious mathematical work, in order to show the resulting AM/AM and AM/PM distortions, manipulating (A.1) gives

$$
\mathbf{V_i} = V_S \cdot \frac{1}{1 + j\omega_o Z_S(\omega_o) \left(c_{i1} + \frac{3}{4} c_{i3} V_i^2 \right)}
$$
$$
= V_S \cdot \frac{Y_S(\omega_o)}{Y_S(\omega_o) + j\omega_o \left(c_{i1} + \frac{3}{4} c_{i3} V_i^2 \right)}. \tag{A.2}
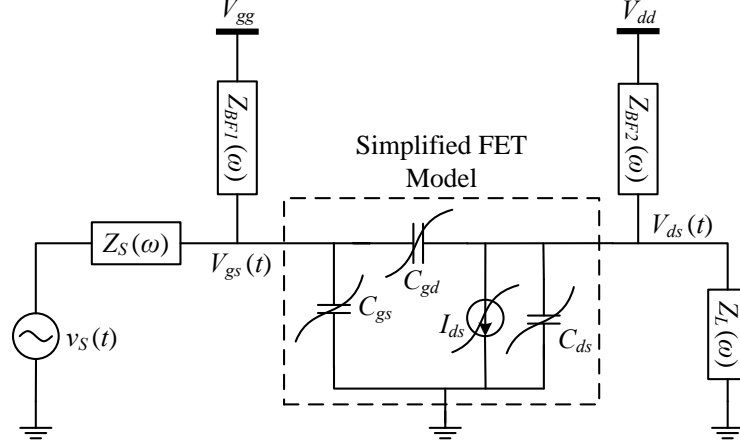$$

Figure A.1: A simplified equivalent circuit of a PA.

Hence, the amplitude and phase distortions are determined as

$$V_i^2 = V_S^2 \cdot \frac{G_S^2(\omega_o) + B_S^2(\omega_o)}{G_S^2(\omega_o) + [B_S(\omega_o) + \omega_o(c_{i1} + 3/4c_{i3}V_i^2)]^2} \qquad \text{(AM/AM)} \qquad \text{(A.3)}$$

$$\phi_i = -\tan^{-1}\left[\frac{B_S(\omega_o) + \omega_o(c_{i1} + 3/4c_{i3}V_i^2)}{G_S(\omega_o)}\right] \qquad \text{(AM/PM)}. \qquad \text{(A.4)}$$

Now, assume that C_{gd} is a non-negligible nonlinear capacitance function of V_{gd}, expressed by (4.13). By revisiting the input KCL at ω_o, we will have

$$(V_S - \mathbf{V_i}) \cdot Y_S(\omega_o) = \mathbf{I_{gs,f_o}} + \mathbf{I_{gd,f_o}} \tag{A.5}$$

where,

$$\begin{aligned}
\mathbf{I_{gd,f_o}} &= j\omega_o C_{gd,f_o}\mathbf{V_f} \\
&= \left(c_{f1} + \frac{3}{4}c_{f3}V_f^2\right)\mathbf{V_f} \\
&\approx \left(c_{f1} + \frac{3}{4}c_{f3}V_o^2\right)(1 + A_V)\mathbf{V_i} \tag{A.6}
\end{aligned}$$

138

where $A_V = -V_o/V_i$ is assumed to be a large constant with no imaginary part. Therefore, the input transfer function can be written as

$$\mathbf{V_i} = V_S \cdot \frac{Y_S(\omega_o)}{Y_S(\omega_o) + j\omega_o C_{in,f_o}} \tag{A.7}$$

where

$$C_{in,f_o} = \left(c_{i1} + \frac{3}{4}c_{i3}V_i^2 \right) + (1 + A_V)\left(c_{f1} + \frac{3}{4}c_{f3}(1 + A_V)^2 V_i^2 \right) \tag{A.8}$$

Thus, the resultant AM/AM and AM/PM are calculated as

$$V_i^2 = V_S^2 \cdot \frac{G_S^2(\omega_o) + B_S^2(\omega_o)}{G_S^2(\omega_o) + [B_S(\omega_o) + \omega_o C_{in,f_o}]^2} \tag{A.9}$$

$$\phi_i = -\tan^{-1}\left[\frac{B_S(\omega_o) + \omega_o C_{in,f_o}}{G_S(\omega_o)} \right]. \tag{A.10}$$

Appendix B

Frequency Dispersion Effect of Impedance Inverter on AM/AM and AM/PM of a Doherty Amplifier

Considering a Doherty configuration at a frequency different from the center frequency, at which the electric length of the impedance inverter is equal to $\theta = \pi/2 + \Delta\theta$ radiant, the voltage phasor across the load is given by

$$
\begin{aligned}
V_L &= V_m \cos\theta - jZ_o I_m \sin\theta \\
&= I_m(Z_m \cos\theta - jZ_o \sin\theta).
\end{aligned}
\tag{B.1}
$$

The impedance seen by the main transistor can be written as

$$
Z_m = Z_o \frac{R_L + jZ_o \tan\theta}{Z_o + jR_L \tan\theta}
\tag{B.2}
$$

where $R_L = R_{opt}/2$ in low power region and $R_L \approx R_{opt}$ at peak power. Therefore, at low powers, it can be approximated as

$$
\begin{aligned}
Z_m^L &= Z_o \frac{0.5 + j\tan(\pi/2 + \Delta\theta)}{1 + j0.5\tan(\pi/2 + \Delta\theta)} \\
&= Z_o \frac{0.5 - j\cot\Delta\theta}{1 - j0.5\cot\Delta\theta} \\
&= Z_o \frac{\tan\Delta\theta - 2j}{2\tan\Delta\theta - j} \\
&= Z_o \left[\frac{2(1 + \tan^2\Delta\theta) - 3j\tan\Delta\theta}{1 + 4\tan^2\Delta\theta} \right].
\end{aligned}
\tag{B.3}
$$

Suppose that $\Delta\theta$ is small such that $\tan^2\Delta\theta \approx (\Delta\theta)^2 \ll 1$, then,

$$
Z_m^L \approx Z_o(2 - j3\Delta\theta).
\tag{B.4}
$$

Using (B.1) and (B.4), the gain change in low power region compared to the ideal case (at center frequency) can be calculated as

$$
\begin{aligned}
|V_L^L| &= Z_o I_m |2\cos\theta - j(3\Delta\theta\cos\theta + \sin\theta)| \\
&= Z_o I_m \sqrt{4\cos^2(\pi/2 + \Delta\theta) + (3\Delta\theta\cos(\pi/2 + \Delta\theta) + \sin(\pi/2 + \Delta\theta))^2} \\
&\approx Z_o I_m \sqrt{4\sin^2\Delta\theta + (\cos\Delta\theta - 3\Delta\theta^2)^2} \\
&\approx Z_o I_m \sqrt{4\Delta\theta^2 + 1 - \Delta\theta^2 - 6\Delta\theta^2} \\
&= Z_o I_m \sqrt{1 - 3\Delta\theta^2}.
\end{aligned}
\tag{B.5}
$$

At 6-dB back-off, $Z_o I_m = V_{dd}/2$, hence,

$$
|V_L^L| \approx \frac{V_{dd}}{2}\sqrt{1 - 3\Delta\theta^2}.
\tag{B.6}
$$

On the other hand, at the peak power, the load impedance of main device is expressed by

$$
Z_m^H \approx Z_o \frac{1 + j\tan(pi/2 + \Delta\theta)}{1 + j\tan(pi/2 + \Delta\theta)} = Z_o.
\tag{B.7}
$$

This means that the impact of frequency dispersion on AM/AM is more significant at low power levels, and the effect is the same whether the frequency is above the $\lambda/4$ line center frequency or below it.

A similar approach can be followed to find the phase distortion. The output phase at low power levels can be expressed as,

$$
\begin{aligned}
\angle V_L^L = \tan^{-1}\left[Z_o \frac{I_m \sin\theta}{V_m \cos\theta} \right] \\
\approx \tan^{-1}\left[\frac{1}{2}\tan(\pi/2 + \Delta\theta) \right] \\
= -\tan^{-1}\left[\frac{1}{2}\cot\Delta\theta \right] \\
= -\cot^{-1}(2\tan\Delta\theta) \\
= -\frac{\pi}{2} + \tan^{-1}(2\tan\Delta\theta) \\
\approx -\frac{\pi}{2} + 2\Delta\theta.
\end{aligned}
\tag{B.8}
$$

Likewise, using , the output phase at peak power is calculated as

$$
\angle V_L^L = \tan^{-1}\left[\frac{Z_o}{Z_m}\tan\theta \right] \approx -\frac{\pi}{2} + \Delta\theta.
\tag{B.9}
$$

The overall AM/PM is thus roughly equal to $\pm\Delta\theta$ at the operating frequency.