Broadband Linearity-Enhanced Doherty Power Amplifier Design Techniques for 5G Sub-6 GHz Applications

by

Emile Traore

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2021

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this 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.

Abstract

The recently deployed fifth generation (5G) cellular networks represent a significant technological advancement over fourth generation (4G) networks. Specifically, new 5G frequency bands were allocated at sub-6 GHz and instantaneous signal bandwidths were increased to satisfy the rapidly growing need for increased data rates. Furthermore, 5G uses more complex modulation schemes to improve spectrum efficiency. Finally, 5G introduced the massive multiple input multiple output (MIMO) concept, where multiple transceivers are used to direct the signal towards specific users, increasing channel capacity.

Conventional power amplifiers (PAs) are not suitable for 5G applications due to the increased signal and system complexity. For example, while the popular Doherty PA (DPA) technique can efficiently amplify signals with complex modulation schemes, conventional DPAs have narrow bandwidths and poor linearity that preclude their use in 5G systems. This has motivated research into techniques to improve DPA bandwidths and linearity for use in 5G networks.

This work focuses on bandwidth and linearity enhancements for sub-6 GHz DPAs realized using discrete components on a printed circuit board. Bandwidth is improved by using broadband architectures for the DPA output combiner network, the absorption of drain parasitics, and a broadband input matching network design. Linearity is enhanced by a proper drain biasing network design and careful selection of transistor source impedances. A 3.3–5.0 GHz DPA using these techniques is fabricated using Cree gallium nitride high electron mobility transistors on a Rogers RO4003C substrate. Measurements indicate 7.8–9.7 dB gain, 38.0–39.2 dBm output power, and drain efficiencies of 48.4–67.1% (saturation) and 38.6–45.1% (6 dB output back-off) over the bandwidth, with good agreement between simulation and measurement results. Under wideband modulated signal excitation, the DPA offers very good linearity, ranging from -51 to -54 dBc at 100 MHz bandwidth and -50 to -52 dBc at 200 MHz bandwidth, with an appropriate digital pre-distortion (DPD) algorithm and number of coefficients. A 2×2 DPA array is evaluated in a fully digital MIMO setup using a 2×2 antenna array with variable spacing to test various antenna crosstalk levels. The DPA array achieves excellent linearity characteristics under 100 MHz bandwidth signals and use of dual-input single-output DPD, with an adjacent channel power ratio below -48 dBc for all antenna coupling levels tested. The DPA remains the ideal choice in 5G MIMO systems when compared to a 2×2 class AB PA array since it can maintain a higher average drain efficiency and similar linearity.

Acknowledgements

I would like to thank my supervisor Professor Slim Boumaiza for his mentorship and support during the past two years. He was instrumental in shaping the direction of my research and led me to gain an entirely new perspective on the design of high-power radio frequency and microwave circuits. He motivated me to surpass myself and accomplish tasks that I never thought I would be capable of and truly made me feel like a valuable member of the EmRG research group.

I would also like to thank all other members of the EmRG research group for technical and moral support during the most trying times of the research process, and for the enlightening and lively discussions we have had during the successful times. In particular, I would like to thank Dr. Jingjing Xia who helped review my designs and provided technical assistance with the equipment, Mohammed Almoneer who assisted me with the modulated signal testbed and digital pre-distortion code, and Hang Yu for being a dependable and cooperative colleague who double-checked all my work inside and outside the lab.

This work would not have been possible without the support from industry and research partners: Keysight for providing the Advanced Design System software and suite of test equipment that were instrumental for accomplishing the work in the EmRG group, Ericsson for contributing financially to the EmRG sub-6 GHz project, and NSERC for funding my research through the CGS M scholarship.

Finally, I would like to extend my thanks to Professors Derek Wright and Peter Levine for reviewing my thesis, and to all the staff at the University of Waterloo who kept the university running and guaranteed students' safety during the COVID-19 pandemic. Without you, this work would not have seen the light of day.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Motivation

Modern wireless networks are faced with a perpetual requirement to increase data rates to fulfill the growing demand for high-quality video streaming (for communication and entertainment purposes) and to accommodate the increasing number of Internet-connected devices, including self-driving vehicles and Internet of Things gadgets. Despite the continuous improvements brought by the fourth generation (4G) Long Term Evolution cellular standard, the growing data requirements of the future would not be satisfied, leading to the advent of fifth generation (5G) cellular networks.

To enable increased data rates, 5G networks employ multiple new strategies. First, higher frequency bands have been allocated for 5G use in the sub-6 GHz space, such as 3.3–3.8 GHz, 3.3–4.2 GHz, and 4.4–5.0 GHz [1]–[4] (new millimeter-wave bands have also been allocated above 24 GHz, which are outside the scope of this work). Figure 1.1 shows the distribution of new 5G frequency bands. Furthermore, wider modulation bandwidths and carrier techniques increase the bandwidth of the signal (up to 100 MHz for the sub-6 GHz bands) [1], [3]–[12]. As data rate is proportional to signal bandwidth, these bands provide higher data rates. Channel capacity for a given bandwidth is further enhanced by more sophisticated modulated schemes (both amplitude and phase modulation), which lead to a higher peak to average power ratio (PAPR) for the signal [1]–[6], [8]–[25]. Finally, 5G introduces the massive multiple input multiple output (MIMO) technique, which is shown in Figure 1.2. This allows the use of beamforming, where the signal can be directed towards a particular user, which improves signal quality and channel capacity. Although MIMO was already present in a few 4G systems, it featured only a few transmitters and receivers, whereas massive MIMO will expand this number up to 64 or 128 [26] or more, greatly improving the effectiveness of this technique.

The more stringent 5G system specifications improvements directly translate into more rigorous requirements for power amplifier (PA) design. Indeed, the PA is arguably the most important component in the transceiver [4], [9], [11], [26]–[27]. As it is the largest and most power-consuming component of the transmitter (between 40–60% of the power in base stations), it is the main determining factor for transmitter size and power efficiency [4], [11], [26]–[27]. Thus, improving PA efficiency leads to reduced energy usage in base stations, reducing costs, cooling needs, and environmental

impact. Furthermore, since the PA is the last active component in the signal chain, its linearity has the greatest effect on the transmitter's overall linearity. This explains the need for improvements in PA design techniques for compliance with 5G standards.

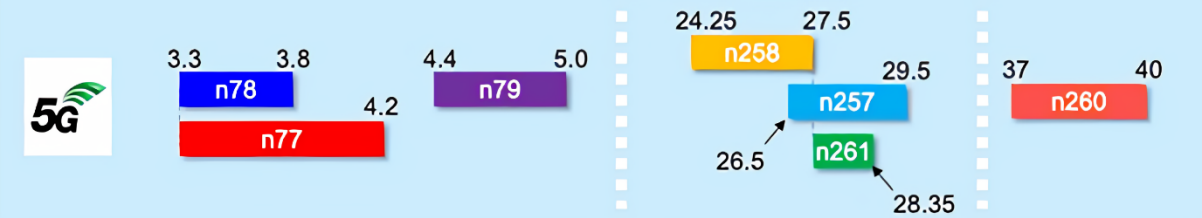


Figure 1.1 Spectrum of 5G frequency bands at sub-6 GHz and millimeter-wave (above 24 GHz) [28]

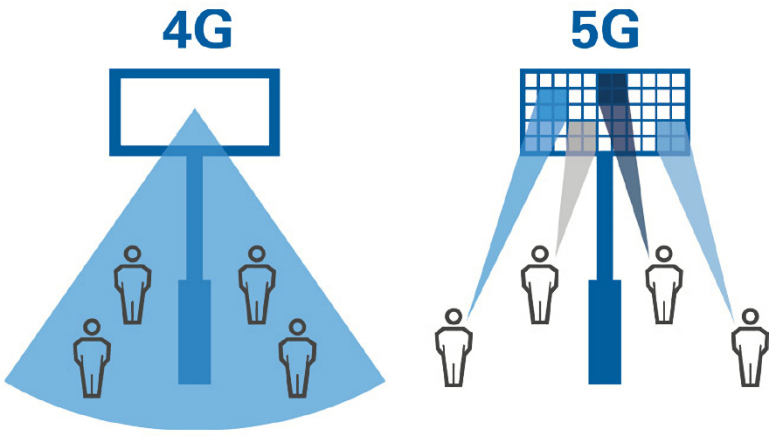


Figure 1.2 Comparison of 4G (no beamforming) and 5G (MIMO beamforming) [29]

1.2 Problem Statement

The advancements in 5G techniques render existing 4G PAs and PA design techniques obsolete. Indeed, 5G places more stringent linearity and efficiency requirements on PAs while requiring the PA to maintain this performance over a wider bandwidth and at higher carrier frequencies. As shown in Figure 1.3 a), the high PAPR of higher-order modulation schemes signifies that the PA must handle signals up to the peak power while meeting the linearity requirements set by the 5G standard [adjacent channel power ratio (ACPR) and normalized mean square error (NMSE)] and remain as efficient as possible at the average power level. However, these specifications are in direct contradiction, as shown in Figure 1.3 b). When a typical PA operates at its peak efficiency, its linearity performance may not meet the defined standard, and when the PA operates at output back-off (OBO) power level and meets the

linearity specifications, its efficiency degrades. Furthermore, in massive MIMO arrays, the PAs will interfere with each other due to crosstalk through the antenna array, introducing a load modulation effect to each of the PA outputs, degrading linearity and efficiency. Digital predistortion (DPD) can alleviate the linearity-efficiency trade-off by enhancing overall system linearity, allowing PAs to operate in a more nonlinear and efficient regime [13], [16].

The quest for an enhanced trade-off between efficiency and linearity has motivated the use of OBO efficiency enhancement techniques for PAs, which increase the power range where a PA operates near its maximum efficiency to handle high PAPR signals with high average efficiency. Of the multiple architectures available, the Doherty PA (DPA) is the most popular due to its inherent simplicity. However, the conventional DPA architecture suffers from significant bandwidth limitations, which are worsened by transistor non-idealities. Furthermore, although the DPA architecture is theoretically perfectly linear, its linearity is significantly compromised by transistor non-idealities in practice. Thus, the improvement of practical DPA bandwidth and linearity is a very active topic for researchers and is the main focus of this work.

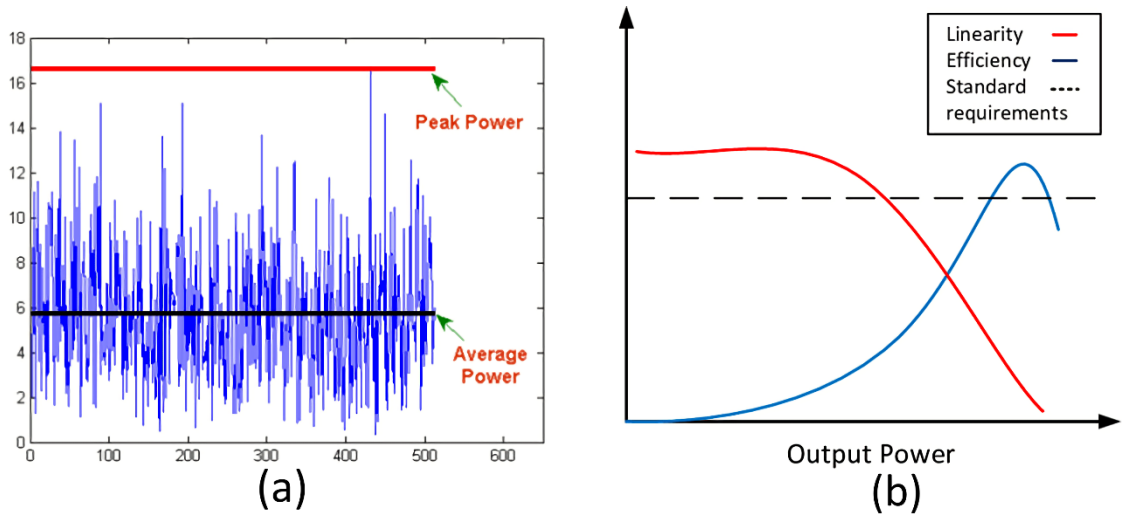


Figure 1.3 a) Peak and average powers of a signal [30], b) PA linearity-efficiency trade-off [31]

1.3 Thesis Organization

This thesis is organized as follows. Chapter 2 begins with a theoretical review of PA-related concepts. It then provides an overview of the different PA classes of operation (basic, high-efficiency, broadband), as well as PA OBO efficiency enhancement techniques including the DPA technique. Finally, the chapter concludes with a literature review of state-of-the-art broadband sub-6 GHz DPAs.

Chapter 3 describes a design methodology for a broadband, linearity-enhanced sub-6 GHz DPA based on discrete components and transmission lines (TLs) on a printed circuit board (PCB). Several bandwidth and linearity enhancement techniques are discussed and applied to a 3.3–5.0 GHz DPA design. Simulation results for the DPA after full electromagnetic (EM) simulation are presented.

Chapter 4 presents measurement results for four different fabricated DPA units and compares them to simulation results. Modulated signal measurements for the individual DPA units and a 2×2 MIMO DPA array are presented, with the devices attached to an antenna array and driven with 5G-candidate orthogonal frequency division multiplexing (OFDM) modulated signals.

Finally, Chapter 5 summarizes the main findings of this work and concludes with a section on potential future work.

Chapter 2
High Power Amplifier Review

2.1 Power Amplifier Basics

2.1.1 Fundamental Power Amplifier Concepts

Radio frequency (RF) PAs amplify signals generated by the RF transmitter to a higher power level, to improve the range of wireless transmission by the antenna. Ideally, the output signal faithfully reproduces the smaller input signal; however, at high power levels, the output signal is affected by PA nonlinear distortion, which is paramount to the study of PAs.

Fundamentally, the PA is an RF amplifier consisting of a transistor with input and output matching networks (IMN and OMN). The IMN and OMN transform the source and load impedances, set to the system characteristic impedance ($Z_0 = 50\ \Omega$), into the impedances seen at the transistor reference planes (Z_S and Z_L) as shown in Figure 2.1. The judicious choice of Z_S and Z_L based on the PA input and output impedances (Z_{in} and Z_{out}) in the nonlinear region is the main aspect of PA design.

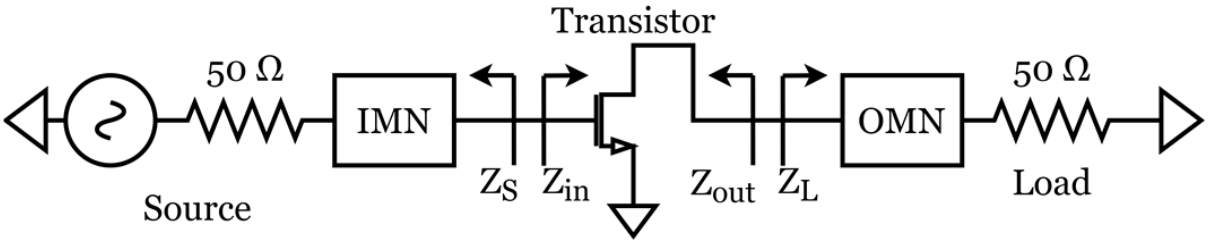


Figure 2.1 Basic schematic of an RF PA

At low input power (P_{in}), the PA output power (P_{out}) is proportional to P_{in}, as the PA operates in linear mode. The small-signal power gain G is thus defined as the ratio of P_{out} to P_{in} for small P_{in}:

$$G = \left.\frac{P_{out}}{P_{in}}\right|_{Pin \approx 0} \tag{2.1}$$

To amplify RF signals, the PA must consume a certain amount of direct current (DC) power, acting as a DC to RF converter. The efficiency with which the PA accomplishes this conversion is known as drain efficiency (η_D) and is given by:

$$\eta_D = \frac{P_{out}}{P_{DC}} \tag{2.2}$$

Drain efficiency is an accurate metric for assessing PA performance when the gain is high. However, high power RF PAs often have low gain and must be driven with a high P_{in}. When P_{in} cannot be neglected, power-added efficiency (PAE) is a more accurate efficiency metric:

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}} = \eta_D \cdot \left(1 - \frac{1}{G}\right) \tag{2.3}$$

Note that as G becomes large, the value of PAE tends towards η_D.

As P_{in} increases, nonlinear distortion starts to occur. The ratio between P_{out} and P_{in} deviates from its small-signal value G. Since the output amplitude is distorted by the input amplitude, this type of distortion is known as amplitude to amplitude (AM-AM). A key PA performance metric is the 1 dB compression point (OP_{1dB}), or the P_{out} level where gain drops by 1 dB:

$$OP_{1dB} = P_{out}|_{P_{out}/P_{in} = G-1} \tag{2.4}$$

The phase of the output signal may also change relative to its small-signal value if the PA exhibits memory effects, a type of distortion known as amplitude to phase (AM-PM). These nonlinearities cause degradation of the signal quality and must be minimized.

Furthermore, as the PA is driven into a nonlinear regime, it emits harmonics, or frequency content at multiples of the signal frequency f_0. The nonlinear PA output signal $y(t)$ can be expressed as a power series of the input signal $x(t)$:

$$y(t) = a_0 + a_1 x(t) + a_2 x^2(t) + a_3 x^3(t) + \cdots \tag{2.5}$$

If $x(t)$ is a sinusoidal signal of the form $A \cdot \cos(2\pi f_0 t) = A \cdot \cos(\omega_0 t)$, $y(t)$ can be expressed as:

$$y(t) = \left(a_0 + \frac{a_2 A^2}{2}\right) + \left(A + \frac{3a_3 A^3}{4}\right)\cos(\omega_0 t) + \frac{a_2 A^2}{2}\cos(2\omega_0 t) + \frac{a_3 A^3}{4}\cos(3\omega_0 t) + \cdots \tag{2.6}$$

The appearance of harmonics at multiples of the center frequency (f_0) is evident, as shown in Figure 2.2 a). However, these terms do not pose a problem because they are far from the desired frequency range and can be filtered out.

However, when the input signal consists of multiple frequencies, equation (2.6) leads to intermodulation distortion (IMD). The output spectrum will then contain terms at frequencies that are the sums and differences of the harmonics as shown in Figure 2.2 b). If the input contains frequencies f_1 and f_2, the second-order IMD terms are $f_1 \pm f_2$ and the third-order IMD terms are $2f_1 \pm f_2$ and $2f_2 \pm f_1$. The $2f_1 - f_2$ and $2f_2 - f_1$ terms are problematic because they fall near f_1, f_2

and cannot easily be filtered out. These terms will cause out-of-band emissions interfering with neighboring communication channels, so it is critical to diminish IMD.

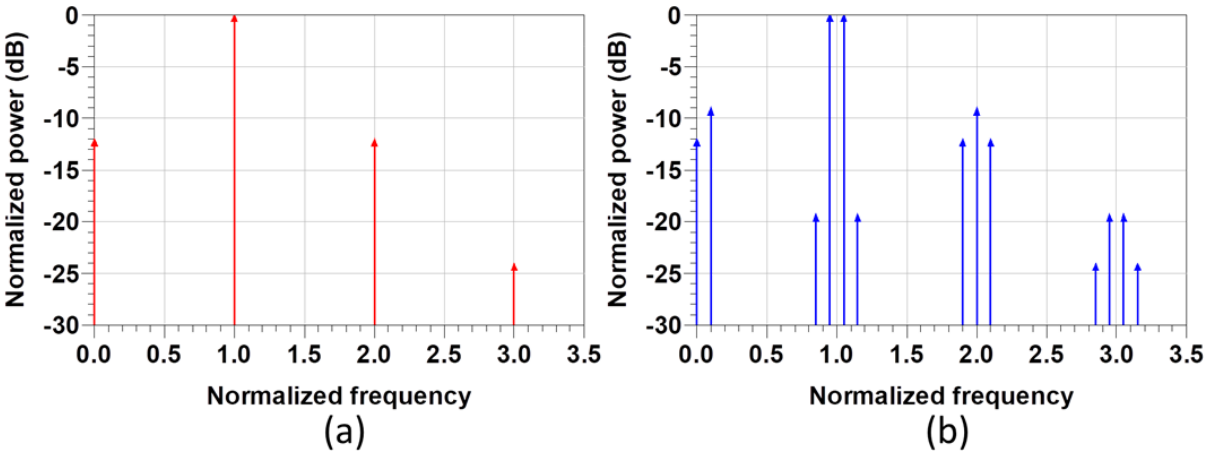


Figure 2.2 Output spectra of a nonlinear PA for: a) Single-tone at f_0, b) Two-tone at $0.95f_0$, $1.05f_0$

2.1.2 Ideal Transistor Model

The majority of RF PAs are based on a transistor operating as a current source. The transistor has three terminals: gate, drain, and source. In the common source configuration, the input is applied at the gate, the output collected at the drain, and the source is grounded. As shown in Figure 2.3, the simplified model of a field-effect transistor (FET) is a voltage-controlled current source. The gate-source voltage V_{gs} is sampled and converted to a drain-source current I_{ds} based on the transconductance (g_m), and I_{ds} is converted to a drain-source voltage V_{ds} based on the load impedance Z_L. Figure 2.4 a) shows the V_{gs} to I_{ds} transfer characteristic, demonstrating the three regions of operation: pinch-off (V_{gs} below pinch-off voltage V_p), linear, and saturation (V_{gs} above $V_{gs,max}$, $I = I_{max}$). In amplifier design, the FET is operated in the linear region to ensure g_m remains constant. Figure 2.4 b) shows the relationship between V_{ds} and I_{ds} at various V_{gs}. V_{ds} must not fall below the knee voltage V_k to ensure linear operation and must not exceed the drain-source breakdown voltage V_{max}. For the subsequent theoretical derivations, the knee voltage will be considered small ($V_k \approx 0$) for simplicity. In real transistors, V_k can be a significant fraction of V_{max}, which varies based on the instantaneous value of V_{gs}, so the analysis is significantly more complex.

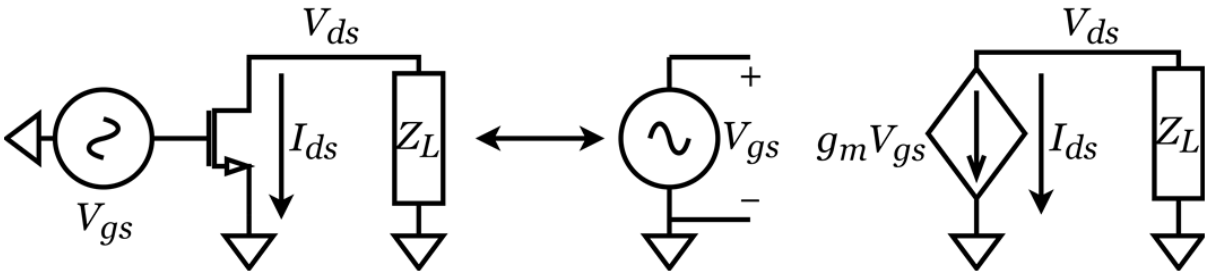


Figure 2.3 Simplified model of a transistor in the common-source configuration

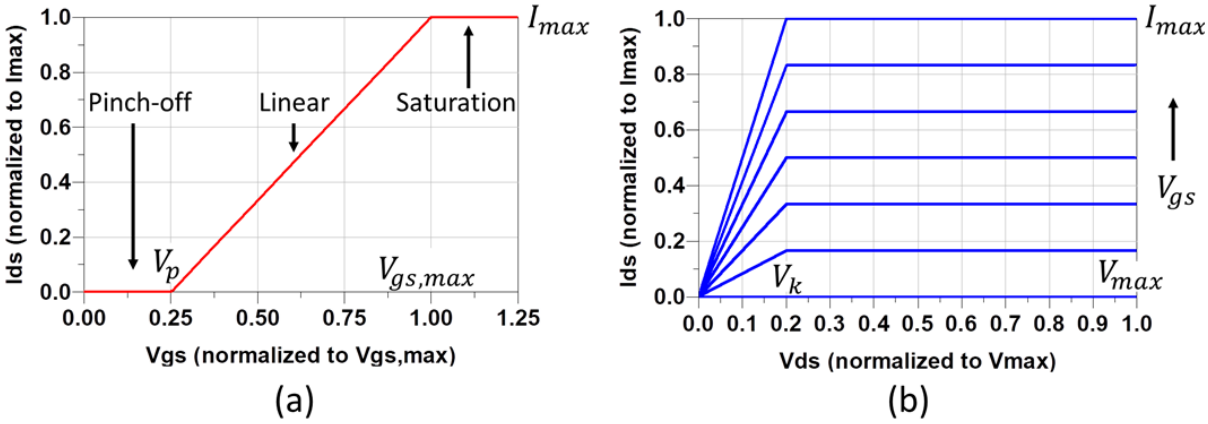


Figure 2.4 Transistor: a) V_{gs} to I_{ds} transfer characteristic, b) V_{ds} vs. I_{ds} at various V_{gs}

2.1.3 Transistor Technologies

According to [32], different RF and microwave applications require different ranges of power and frequency, each satisfied by a different transistor technology. Silicon-based technologies include complementary metal-oxide semiconductor (CMOS), silicon germanium (SiGe), and laterally diffused metal oxide semiconductor (LDMOS). III-V technologies are based on compound semiconductors composed of Group III and Group V elements from the periodic table. The most common ones are gallium arsenide (GaAs) and gallium nitride (GaN).

As shown in Figure 2.5 [32], SiGe and GaAs are mostly used for low-power, high-frequency devices. Although GaAs reaches higher power and frequency than SiGe, silicon-based technologies can integrate the PA with the rest of the transceiver on a single chip [33]–[34]. For high-power applications, the two dominant technologies are LDMOS and GaN, which can sustain much higher operating voltages than other technologies [32], [35]. While LDMOS is limited to lower frequencies, GaN quickly outperforms it as frequencies rise towards the GHz range, due to its significantly higher electron velocities [35]. GaN is also able to sustain higher breakdown voltages and temperatures, enabling

improve power density and smaller devices [35]. Thus, for sub-6 GHz 5G applications, which mainly consist of discrete devices on PCBs, GaN is the optimal technology [32], [35].

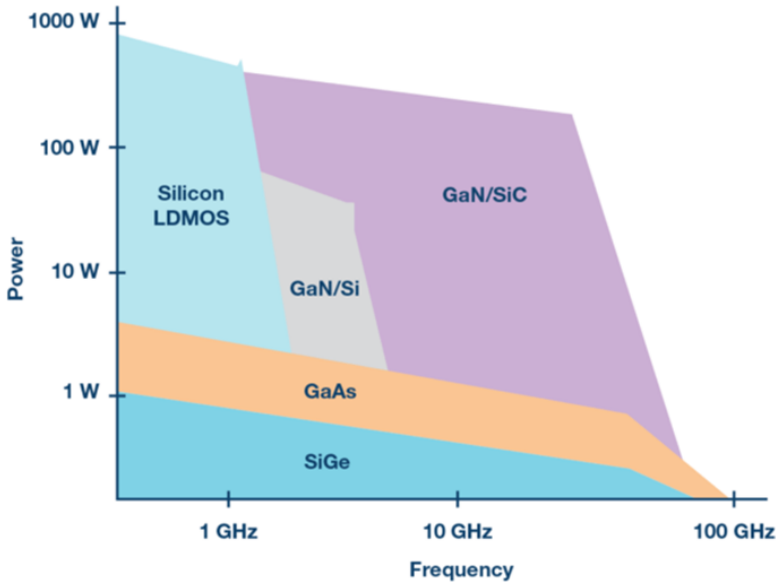


Figure 2.5 Optimum transistor technology vs. frequency and output power range [32]

2.2 Power Amplifier Classes

The different PA classes of operation are distinguished by the biasing point of the transistor and the load network at the transistor drain. The transistor is biased at a quiescent gate voltage $V_{GS,Q}$ and drain voltage V_{DD}, with a corresponding quiescent drain current $I_{DS,Q}$. The various classes offer trade-offs between linearity and efficiency, which are detailed in [36] and will be summarized in this section.

2.2.1 Class A

In class A, $V_{GS,Q}$ is set to the midpoint between V_p and $V_{GS,max}$ ($I_{DS,Q}$ is half of I_{max}), and V_{DD} is set to the midpoint between V_{max} and V_k. Assuming $V_k \ll V_{max}$, the biasing point is given as follows [36]:

$$\left(V_{DD}, I_{DS,Q}\right) = \left(\frac{V_{max}}{2}, \frac{I_{max}}{2}\right) \tag{2.7}$$

This bias point allows the voltage and current to swing from 0 V to V_{max} and 0 A to I_{max}, with amplitudes of $V_{max}/2$ and $I_{max}/2$, respectively. To ensure both V_{ds} and I_{ds} reach their maximum swing simultaneously, the transistor must be terminated with its optimum load resistance (R_{opt}) [36]:

$$R_{opt} = \frac{V_{max}/2}{I_{max}/2} = \frac{V_{max}}{I_{max}} \tag{2.8}$$

In class A, the PA maximum (saturated) output power (P_{sat}), DC power consumption (P_{DC}), and drain efficiency are the following [36]:

$$P_{sat,class\,A} = \frac{1}{2} \cdot \frac{V_{max}}{2} \cdot \frac{I_{max}}{2} = \frac{V_{max} \cdot I_{max}}{8} \tag{2.9}$$

$$P_{DC,class\,A} = V_{DS,Q} \cdot I_{DS,Q} = \frac{V_{max}}{2} \cdot \frac{I_{max}}{2} = \frac{V_{max} \cdot I_{max}}{4} \tag{2.10}$$

$$\eta_{D,class\,A} = \frac{P_{sat}}{P_{DC}} = \left(\frac{V_{max} \cdot I_{max}}{8}\right) / \left(\frac{V_{max} \cdot I_{max}}{4}\right) = 0.5 = 50\% \tag{2.11}$$

Class A PA efficiency is theoretically limited to 50%. Furthermore, since P_{DC} is constant and independent of P_{out}, η_D drops proportionally to P_{out}, which is unsuitable for handling high PAPR signals. However, class A PA linearity is very good because the I_{ds} and V_{ds} are both sinusoidal, and gain is high because the PA is conducting for the entire range of the input signal. Therefore, class A PAs are mostly used for instrumentation applications, and rarely in cellular applications.

2.2.2 Reduced Conduction Angle Classes

The main disadvantage of the class A PA is its low efficiency, especially at OBO. Different classes of operation can be obtained by setting $V_{GS,Q}$ such that $I_{DS,Q}$ is below $I_{max}/2$. For example, class B operates the transistor at $V_{GS,Q} = V_p$ and $I_{DS,Q} = 0$. In class B, the transistor conducts for only half of the input signal waveform and operates in cutoff for the other half. In other words, the conduction angle of the class B PA is $\theta = 180°$, whereas it is $\theta = 360°$ for the class A PA.

Besides classes A and B, there exists a continuum of PA biasing conditions. Class AB PAs are biased such that $0 < I_{DS,Q} < I_{max}/2$ such that $180° < \theta < 360°$. Class C PAs are biased at $V_{GS,Q} < V_p$ such that $\theta < 180°$. In all these cases, the drain is biased at $V_{DD} = V_{max}/2$. Figure 2.6 a) shows the relationship between $V_{GS,Q}$ and $I_{DS,Q}$ for the different classes, and Figure 2.6 b) shows the load lines, which relate the instantaneous voltage and current, $V_{ds}(t)$ and $I_{ds}(t)$, for each class. As the conduction angle decreases, the transistor spends more and more of the cycle at zero current, reducing P_{DC}. The reduced PA classes have different linearity-efficiency trade-offs, which can be directly computed from the conduction angle.

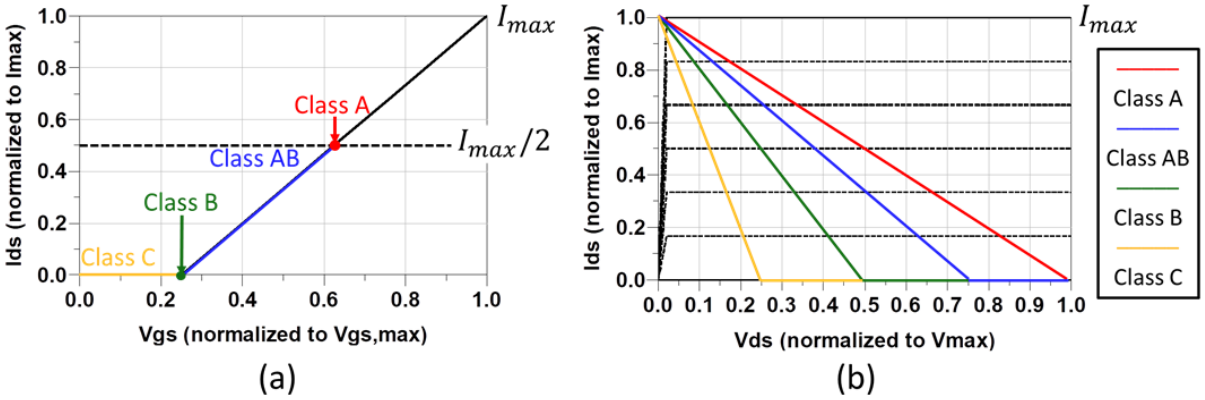


Figure 2.6 PA class A, AB, B, C: a) Biasing point, b) Load line

Figure 2.7 a) shows the time-domain waveforms of $I_{ds}(t)$ for the various classes of operation. The corresponding equation is the following, where ωt represents the phase of the input signal [36]:

$$I_{ds}(t) = \begin{cases} 0, & -\pi < \omega t < -\theta/2 \\ \dfrac{I_{max}}{1-\cos(\theta/2)}\left[\cos(\omega t) - \cos\left(\dfrac{\theta}{2}\right)\right], & -\dfrac{\theta}{2} < \omega t < \dfrac{\theta}{2} \\ 0, & \theta/2 < \omega t < \pi \end{cases} \tag{2.12}$$

By applying Fourier analysis, the drain current can be written as follows [36]:

$$I_{ds}(t) = I_0 + I_1\cos(\omega t) + I_2\cos(2\omega t) + I_3\cos(3\omega t) + \cdots \tag{2.13}$$

$$I_0 = \frac{I_{max}}{2\pi}\cdot\frac{2\cdot\sin(\theta/2) - \theta\cdot\cos(\theta/2)}{1-\cos(\theta/2)} \tag{2.14}$$

$$I_1 = \frac{I_{max}}{2\pi}\cdot\frac{\theta - \sin(\theta)}{1-\cos(\theta/2)} \tag{2.15}$$

$$I_n = \frac{1}{\pi}\int_{-\theta/2}^{\theta/2}\frac{I_{max}}{1-\cos(\theta/2)}[\cos(\omega t) - \cos(\theta/2)]\cdot\cos(n\omega t)\cdot d(\omega t) \tag{2.16}$$

When θ is reduced, harmonics appear in $I_{ds}(t)$. Figure 2.7 b) plots the amplitudes of the first five harmonics against θ, showing that as θ decreases, harmonic distortion increases. To avoid power dissipation at the harmonics, which is undesired because they do not contribute to the useful signal, the load network must feature short-circuits at the harmonic frequencies to force the voltage at the harmonics to zero. Thus, only the DC current I_0 and fundamental current I_1 influence η_D. Interestingly, I_1 is higher for class AB than class A or class B, leading to a higher P_{sat}.

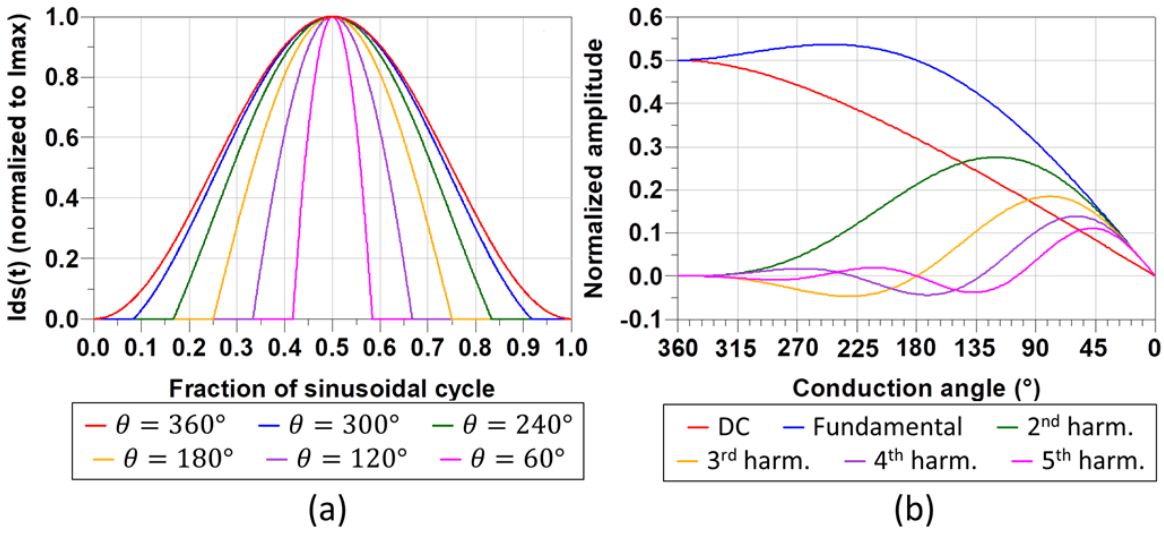


Figure 2.7 a) I_{ds} waveforms vs. conduction angle, b) I_{ds} harmonic amplitudes vs. conduction angle

Based on these parameters, R_{opt}, P_{sat}, P_{DC}, and η_D can be calculated for a general value of θ. R_{opt} is set to simultaneously maximize voltage amplitude ($V_{max}/2$) and current amplitude (I_1). P_{sat} is calculated based on the fundamental voltage and current amplitudes, and P_{DC}, based on the DC voltage and current. This results in the following equations [36]:

$$R_{opt} = \frac{V_{max}/2}{I_1} \tag{2.17}$$

$$P_{sat} = \frac{V_{max}/2 \cdot I_1}{2} = \frac{V_{max} \cdot I_1}{4} \tag{2.18}$$

$$P_{DC} = \frac{V_{max} \cdot I_0}{2} \tag{2.19}$$

$$\eta_D = \frac{P_{sat}}{P_{DC}} = \frac{I_1}{2 \cdot I_0} \tag{2.20}$$

Note that $\theta = 2\pi$ is consistent with the parameters for class A derived previously ($I_0 = I_1 = I_{max}/2$). Furthermore, for the special case of the class B PA ($\theta = \pi$), $I_0 = I_{max}/\pi$ and $I_1 = I_{max}/2$ are obtained. Thus, class B has the same R_{opt} and P_{sat} as class A, but with a significantly lower P_{DC} and higher η_D as shown in equation (2.21). Figure 2.8 a) plots η_D against P_{out} for different biasing points, showing that for smaller values of θ, efficiency is higher at P_{sat} and drops off slower at OBO. In summary, reduced conduction angles classes improve η_D compared to class A [36].

$$\eta_{D,class\,B} = \frac{I_{max}/2}{2 \cdot I_{max}/\pi} = \frac{\pi}{4} \approx 78.5\% \tag{2.21}$$

Figure 2.8 b) plots P_{out} against P_{in} for various biasing points to assess gain and linearity. The gain follows the opposite trend of η_D: it decreases monotonically as θ decreases since the PA conducts for a smaller portion of the input signal. In terms of linearity, for classes A and B, P_{out} is always proportional to P_{in} because θ is constant ($\theta = 360°$ for class A, $\theta = 180°$ for class B). Thus, class A and B are perfectly linear in theory, although class B has a 6 dB lower gain because half of the input signal is not utilized. For classes AB and C, θ changes as the input signal magnitude changes, so they are nonlinear (weakly for class AB, strongly for class C). The characteristics of the different classes of operation of the PA based on conduction angle are summarized in Table 2.1.

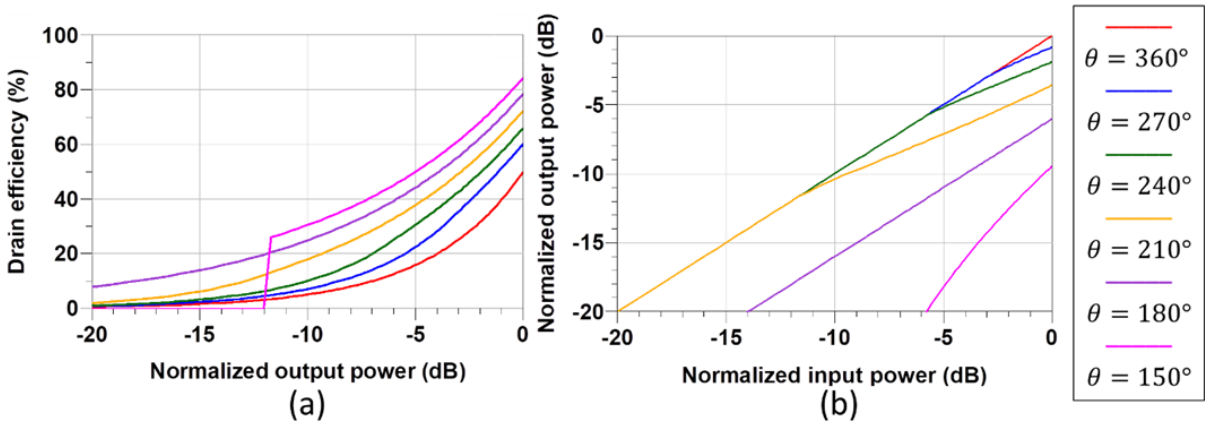


Figure 2.8 a) η_D vs. P_{out} at various conduction angles, b) P_{out} vs. P_{in} at various conduction angles

Table 2.1 Characteristics of conduction angle-based PA classes

Class of operation	Conduction angle	Efficiency		Gain	Linearity (theory)	Output power
A	$\theta = 2\pi$	50%	Poor	Best	Ideal	Good
AB	$\pi < \theta < 2\pi$	50%–78.5%	Fair	Good	Good	Best
B	$\theta = \pi$	78.5%	Good	Fair	Ideal	Good
C	$0 < \theta < \pi$	78.5%–100%	Best	Poor	Poor	Poor

2.2.3 High Efficiency Classes

The reduced conduction angle class PAs presented in the previous section offer trade-offs between linearity and efficiency. However, as η_D tends towards 100% in class C, gain and P_{out} both quickly drop to zero. Therefore, other high efficiency classes were developed, such as classes F and F^{-1} [36].

As described in [27], [36]–[37], class F operation aims to achieve a square wave $V_{ds}(t)$ and a half-sine wave $I_{ds}(t)$, and class F^{-1} is its dual, with a half-sine wave $V_{ds}(t)$ and square wave $I_{ds}(t)$. As shown in Figure 2.9 a) and b), the overlap between the V_{ds} and I_{ds} waveforms of class F and F^{-1} PAs decreases as more and more harmonics are terminated and disappears if all harmonics are terminated.

Class F PAs are biased in class B, to obtain a half-sine I_{ds}, and class F-1 PAs are biased in class A and overdriven to approximate a square wave I_{ds}. V_{ds} waveforms are engineered by setting the harmonic impedances appropriately. In class F, the odd harmonics are open-circuited and even harmonics are short-circuited to obtain a square wave. In class F-1, the even harmonics are open-circuited and odd harmonics are short-circuited to obtain a half-sine wave. The current and voltage waveforms can be written as follows [36]–[37]:

$$I_{ds,class\,F}(t) = I_0 + I_1\cos(\omega t) + I_2\cos(2\omega t) + \cdots = \frac{I_{max}}{\pi} + \frac{I_{max}}{2}\cos(\omega t) + \cdots \tag{2.22}$$

$$V_{ds,class\,F}(t) = V_0 + V_1\cos(\omega t) + V_3\cos(3\omega t) + \cdots = \frac{V_{max}}{2} - \frac{2V_{max}}{\pi}\cos(\omega t) + \cdots \tag{2.23}$$

$$I_{ds,class\,F^{-1}}(t) = I_0 + I_1\cos(\omega t) + I_3\cos(3\omega t) + \cdots = \frac{I_{max}}{2} + \frac{2I_{max}}{\pi}\cos(\omega t) + \cdots \tag{2.24}$$

$$V_{ds,class\,F^{-1}}(t) = V_0 + V_1\cos(\omega t) + V_2\cos(2\omega t) + \cdots = \frac{V_{max}}{\pi} - \frac{V_{max}}{2}\cos(\omega t) + \cdots \tag{2.25}$$

R_{opt}, P_{sat}, P_{DC}, and η_D are the following [36]:

$$R_{opt,class\,F} = \left|\frac{V_1}{I_1}\right| = \frac{2V_{max}/\pi}{I_{max}/2} = \frac{4}{\pi}\cdot\frac{V_{max}}{I_{max}} = \frac{4}{\pi}\cdot R_{opt,class\,B} \tag{2.26}$$

$$R_{opt,class\,F^{-1}} = \left|\frac{V_1}{I_1}\right| = \frac{V_{max}/2}{2I_{max}/\pi} = \frac{\pi}{4}\cdot\frac{V_{max}}{I_{max}} = \frac{\pi}{4}\cdot R_{opt,class\,B} \tag{2.27}$$

$$P_{sat} = \left|\frac{V_1\cdot I_1}{2}\right| = \frac{V_{max}\cdot I_{max}}{2\pi} \tag{2.28}$$

$$P_{DC} = V_0\cdot I_0 = \frac{V_{max}\cdot I_{max}}{2\pi} \tag{2.29}$$

$$\eta_D = \frac{P_{out}}{P_{DC}} = 1 = 100\% \tag{2.30}$$

From equations $(2.28) - (2.30)$, both class F and F^{-1} provide $\eta_D = 100\%$ and increased P_{sat} from the transistor. However, class F is preferred in wireless applications because it provides increased R_{opt}

compared to class B (easier to match to $Z_0 = 50\ \Omega$) and does not require overdriving the transistor like class F^{-1} [37].

In practice, ideal class F and F^{-1} terminations cannot be achieved because of the infinite number of terminated harmonics. Practically, up to three harmonics can be terminated with a single stub matching network and $\lambda/4$ biasing TL [27]. The relationship between the number of terminated harmonics and η_D for class F PAs is given in Table 2.2, showing a 10 percentage point improvement from class B by terminating the third harmonic [36].

Table 2.2 Efficiency of class F PA vs. number of terminated harmonics

		Voltage harmonics			
		1	1, 3	1, 3, 5	1, 3, 5… ∞
Current harmonics	1	0.500	0.563	0.586	0.637
	1, 2	0.667	0.750	0.781	0.849
	1, 2, 4	0.711	0.800	0.833	0.905
	1, 2, 4… ∞	0.785	0.884	0.920	1.000

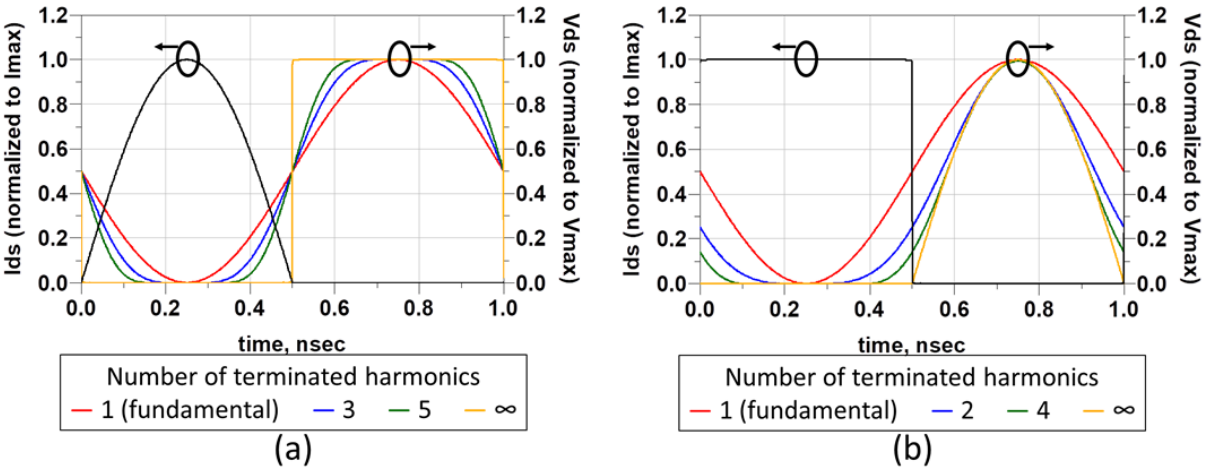


Figure 2.9 $I_{ds}(t)$ and $V_{ds}(t)$ waveforms vs. number of harmonics for: a) Class F, b) Class F^{-1}

2.2.4 Broadband Classes

The reduced conduction angles and high efficiency PAs require open-circuit or short-circuit harmonic terminations, which cannot be maintained over wide bandwidths, so these PAs are narrowband by nature. Indeed, PAs and other RF circuits can be characterized by their fractional bandwidth (FBW), which is the ratio of bandwidth to center frequency, an important figure of merit to determine if a circuit is broadband or narrowband. For example, conventional class F PAs have less than 10% FBW [7]. A

broadband alternative is the class J/J* PA, which maintains the performance of class B without the strict requirement of shorted harmonics [36], [38]–[39], and can achieve above 120% FBW [27].

Class J/J* is based on the class B biasing point, so the drain current waveform equation is obtained from equations (2.13)–(2.16) with the conduction angle set to $\theta = \pi$ [36], [38]–[39]:

$$I_{ds}(t) = I_0 + I_1\cos(\omega t) + I_2\cos(2\omega t) + \cdots = \frac{I_{max}}{\pi} + \frac{I_{max}}{2}\cos(\omega t) + \frac{2}{3\pi}\cos(2\omega t) + \cdots \quad (2.31)$$

However, as shown in [27], [36], [38]–[40], class J/J* allows the fundamental load impedance to deviate from R_{opt} and the second harmonic impedance to take reactive values instead of a strict short-circuit. The other harmonics are still assumed to be short-circuited. The fundamental and second harmonic impedances (Z_1, Z_2) are chosen based on equations (2.32) and (2.33), where $-1 < \alpha < 1$. $\alpha = 0$ corresponds to class B, whereas class J and J* operation refer to the regions where $-1 \leq \alpha < 0$ and $0 < \alpha \leq 1$, respectively.

$$Z_1 = R_{opt} \cdot (1 + j\alpha) \quad (2.32)$$

$$Z_2 = -j\frac{3\pi}{8}\alpha \cdot R_{opt} \quad (2.33)$$

Figure 2.10 a) shows the impedances Z_1 and Z_2 on a Smith chart normalized to R_{opt}. Z_1 has a constant resistive component R_{opt} with a reactive impedance component, and Z_2 has only a reactive component with opposite sign to Z_1. These conditions can be achieved over a broad bandwidth: in fact, the matching network can be designed such that as frequency varies, Z_1 and Z_2 vary along the class J/J* curves. This is known as continuous class B/J/J* design [38]–[39].

The class J/J* load network produces a drain voltage given by equation (2.34) where V_{DD} represents the supply voltage [36], [38]–[39]:

$$V_{ds}(t) = V_{DD}\left[1 - \cos(\omega t) + \alpha \cdot \sin(\omega t) - \frac{\alpha}{2}\sin(2\omega t)\right] \quad (2.34)$$

As shown in [36], [38]–[40], this shape of $V_{ds}(t)$, together with $I_{ds}(t)$ given in equation (2.31), results in P_{out}, linearity, and efficiency identical to a class B PA. Indeed, the first two terms, V_{DD} (DC voltage) and $-V_{DD} \cdot \cos(\omega t)$ (real part of the fundamental voltage), are the same as for class B. Thus, P_{out} and efficiency are identical to class B. The third and fourth terms are purely imaginary and dissipate no power, but shape the drain voltage waveform to avoid intrusion into the knee region to maintain the same linearity as class B. However, class J PAs have limitations. The peak value of $V_{ds}(t)$

is significantly higher than in class B for the same V_{DD}, as shown in Figure 2.10 b), so a continuous class B/J/J* PA design must use transistors with higher V_{max} or decrease V_{DD} which lowers P_{sat}. Furthermore, the maximum theoretical bandwidth is one octave, since the fundamental and second harmonic frequency ranges overlap beyond that point, while they require different termination styles. Also, class J PAs only achieve class B efficiency, not 100% like the class F/F^{-1} PAs.

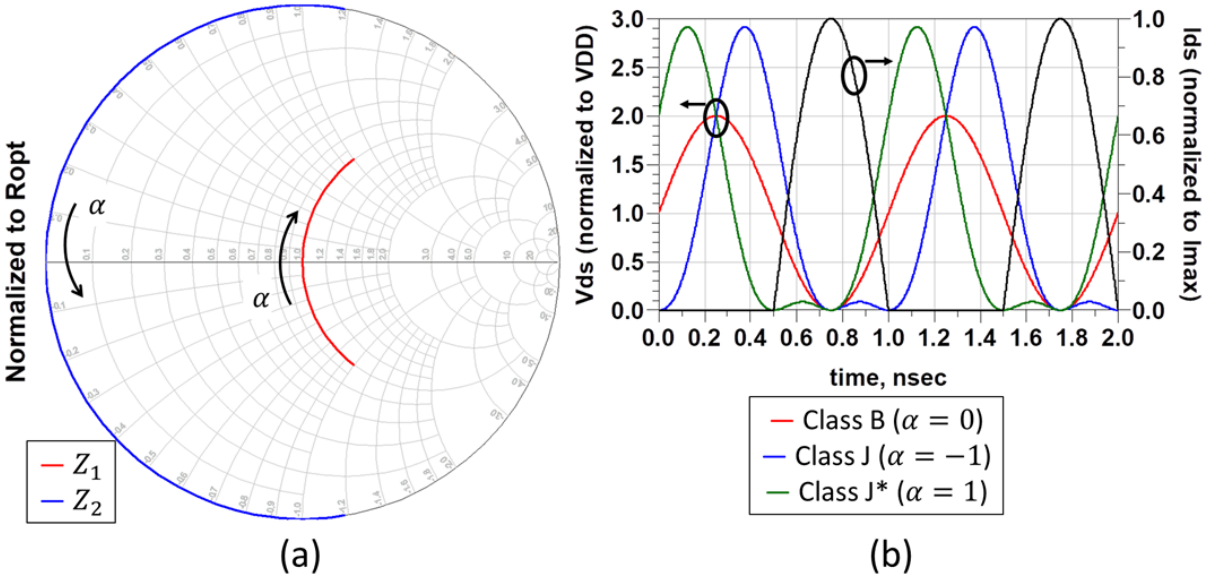


Figure 2.10 Continuous class B/J/J*: a) Z_1 and Z_2 vs. frequency, b) $I_{ds}(t)$ and $V_{ds}(t)$ for various α

Continuous class B/J/J* principle can be applied to class F/F^{-1} PAs, leading to broadband and highly efficient continuous class F/F^{-1} PAs (equivalent efficiency to conventional class F/F^{-1} without strict harmonic open and short circuits). The PA can also be designed to transfer between continuous class F and F^{-1} over its bandwidth. Formulations that allow resistive components for the harmonic terminations [7], [27], [40] also exist, allowing bandwidths exceeding one octave. The detailed theory of continuous class F/F^{-1} PAs is outside of the scope of this work.

2.3 Output Back-off Efficiency Enhancement Techniques

The PA classes presented so far are all targeted maximum η_D at P_{sat}. However, modern communication signals have high PAPR, so the PA spends most of the time at OBO, leading to a rapid drop-off in η_D as P_{out} decreases. This led to the development of techniques that allow the PA to maintain peak η_D at OBO. These techniques are either based on supply modulation, such as envelope elimination and

restoration (EER) or envelope tracking (ET) or based on load modulation, such as the outphasing PA and DPA.

2.3.1 EER and ET Power Amplifiers

EER and ET PAs are described in [36]. Both techniques are based on supply modulation. A diagram of a supply modulated PA is shown in Figure 2.11. The PA V_{DD} is dynamically adjusted based on the envelope of the input signal using a supply modulator (SM). Thus, as signal level decreases from peak power, V_{DD} decreases proportionally, such that the PA always outputs the highest fundamental voltage V_1 allowed by V_{DD} and maintains its peak efficiency across a wide range of power levels.

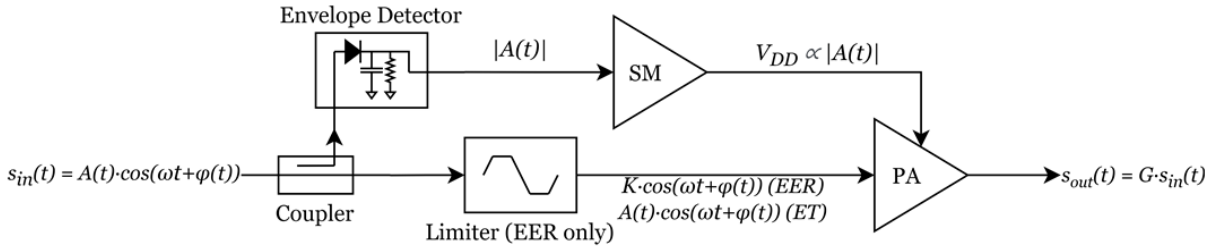


Figure 2.11 Block diagram of a supply-modulated PA (EER/ET)

The main distinction between EER and ET is the input signal to the PA. In EER, the envelope is eliminated through a limiter: the input signal to the PA has constant amplitude and contains only phase information. Since P_{in} is constant, a nonlinear and thus very efficient PA can be used. However, the SM must track the envelope perfectly, which is very difficult to accomplish and is a major limitation of EER. The time alignment of the SM and PA paths must be near-ideal, otherwise the incorrect amplitude will be restored to the PA output signal, causing distortion (third-order IMD approximately -30 dBc for a delay mismatch of 10% of the RF bandwidth) [41]. Thus, EER is limited to signal bandwidths much lower than the hundreds of MHz required by 5G. Furthermore, to maintain the output signal linearity, the supply modulation must be maintained down to 0 V, which requires a very high dynamic range SM.

In contrast, ET does not eliminate the amplitude information from the PA input signal and requires a linear PA, such as class B, which is typically less efficient than nonlinear classes. However, the requirements on the supply tracking are relaxed: the time alignment requirement is not as stringent and the SM can operate at a lower dynamic range, not necessarily down to 0 V. Unfortunately, EER still has practical implementation problems. Accurate delay alignment between SM and PA still plays a

role: if V_{DD} is too low, or too high for the signal, linearity or efficiency will suffer, respectively. Furthermore, a broadband SM is required: the signal envelope is a vector sum of the in-phase and quadrature baseband components, so its bandwidth is theoretically infinite [6], [42]. In practice, the SM bandwidth must be four to eight times the signal bandwidth [6]. Broadband SMs are power-inefficient, so even if the PA is highly efficient, the overall efficiency of the ET system will be degraded due to the low SM efficiency, as shown in equation (2.35) [15]. Thus, ET is also not viable for signal bandwidths in the hundreds of MHz required by 5G.

$$\eta_{total} = \eta_{PA} \cdot \eta_{SM} \tag{2.35}$$

Improvements to the ET PA have been proposed in the literature. In [6], the authors propose to track the envelope squared or to the fourth power, which eliminates the square root function: the required bandwidth of the SM is only two or four times the baseband signal's bandwidth. This improves η_{SM} significantly, at a small linearity penalty. The design in [15] features a soft-switching buck converter that eliminates overlap in the SM drain current and voltage waveforms, improving η_{SM}. However, in general, ET is not a popular OBO efficiency-enhancement technique in state-of-the-art publications in the literature.

2.3.2 Outphasing Power Amplifiers

The outphasing technique is described in [16], [18]. A signal processor decomposes an amplitude- and phase-modulated signal $s_{in}(t)$ into two phase-modulated, constant-amplitude signals $s_1(t)$ and $s_2(t)$ as shown in equations (2.36)–(2.38). $s_{1,2}(t)$ are then amplified using two PAs and combined at the output to form an amplified version of $s_{in}(t)$. Since both PAs always operate at a constant P_{in}, they can be nonlinear and highly efficient topologies. The outphasing output combiner network (OCN) can be either isolated or non-isolated. Figure 2.12 a) shows a block diagram of an outphasing PA.

$$s_{in}(t) = A(t)\cos\big(\omega t + \phi(t)\big) = s_1(t) + s_2(t) \tag{2.36}$$

$$s_{1,2}(t) = \frac{A_{max}}{2} \cdot \cos\left[\omega t + \phi(t) \pm \cos^{-1}\left(\frac{A(t)}{A_{max}}\right)\right] = \frac{A_{max}}{2} \cdot \cos[\omega t + \phi(t) \pm \theta(t)] \tag{2.37}$$

$$\theta(t) = \cos^{-1}\left(\frac{A(t)}{A_{max}}\right) \tag{2.38}$$

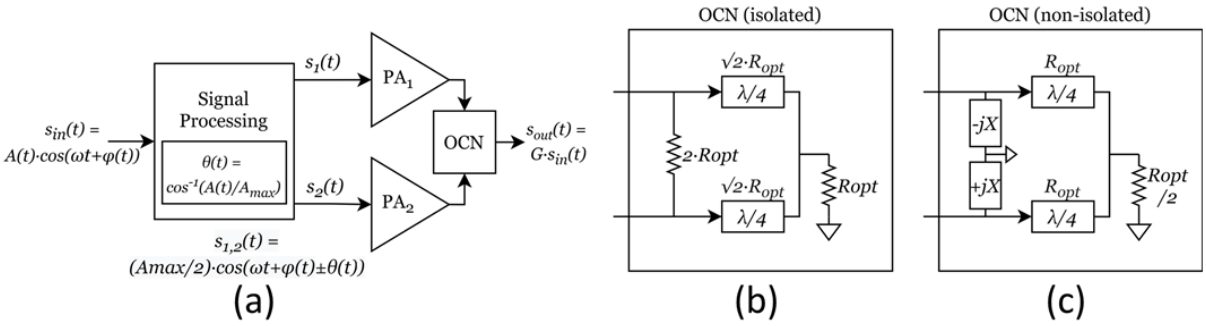


Figure 2.12 Block diagrams of an: a) Outphasing PA, b) Isolated OCN, c) Non-isolated OCN

With an isolated OCN, shown in Figure 2.12 b), each PA sees a load of R_{opt} regardless of the other PA's output signal, so both PAs always operate at peak η_D. However, while the PAs always operate at P_{sat}, P_{out} is proportional to $A^2(t)$. The excess power must be dissipated in the OCN, which is lossy. Thus, no OBO efficiency enhancement results: the only benefit is linearity.

Chireix [43] proposed a non-isolated OCN, shown in Figure 2.12 c). The output impedance of each PA is modulated by the output of the other PA. The impedances seen by each PA are given in equation (2.39) [36], [41], [43]:

$$Z_{1,2} = \frac{R_L}{2}\cdot[1 \mp j\cdot\tan(\theta)] = R_{opt}\cdot[1 \mp j\cdot\tan(\theta(t))] \tag{2.39}$$

As P_{out} decreases, $\theta(t)$ increases and $Z_{1.2}$ both increase, reducing the power output by each PA. Thus, the sum of the PA output powers equals P_{out} and the OCN is lossless. However, the load modulation also produces an imaginary impedance component, which degrades PA efficiency and linearity. Therefore, the Chireix OCN also includes compensating reactances $\pm jX$ to restore the impedances Z_1 and Z_2 to a real value for a specific OBO level. Figure 2.13 a) shows Z_1 and Z_2 for both PAs with various levels of Chireix compensation. The resulting outphasing efficiency profiles, with peaks at the chosen OBO level and at P_{sat}, are shown in Figure 2.13 b).

The Chireix outphasing PA offers good OBO efficiency enhancement when handling high PAPR communication signals [1], [6], [8], [11], [15]–[16], [18], [22], [44]. However, the signal separation is very computationally intensive and requires a separate computer [16] or field-programmable gate array [18]. The baseband processor's power consumption is not accounted for in reported efficiency figures [16], [18], which is not a realistic representation of the outphasing system performance. Furthermore, $\theta(t)$ is a nonlinear function of the baseband signal, so the signal processor's bandwidth requirement is broader than the signal bandwidth, increasing complexity and power consumption [16].

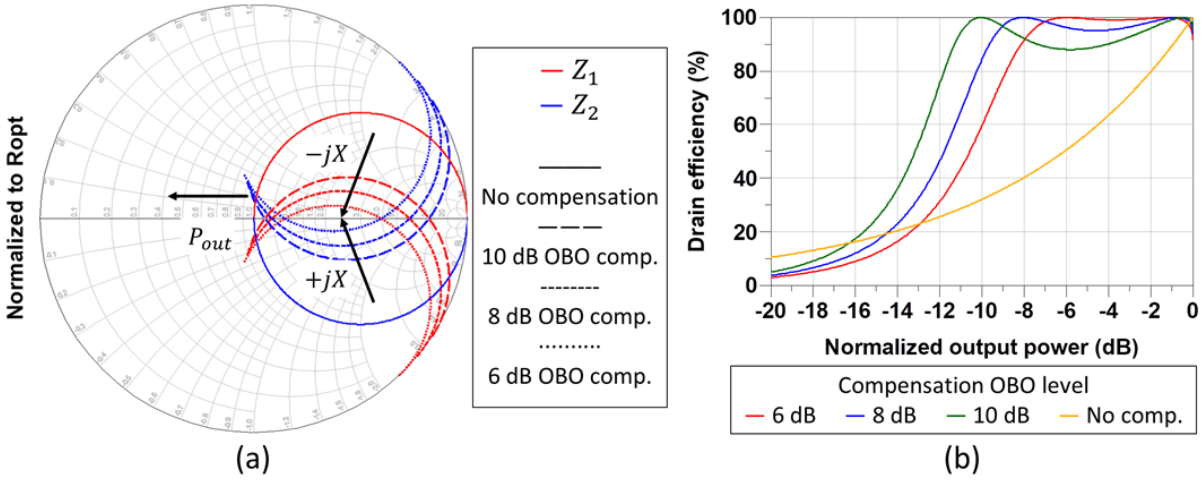


Figure 2.13 Chireix outphasing PA at various OBO levels: a) Load modulation, b) Efficiency profiles

Improvements to the outphasing PA have been proposed in the literature. The outphasing PA in [18] uses broadband continuous class F PAs for improving the overall system bandwidth, and the authors in [16] propose a multi-mode outphasing PA with multiple branches in the Chireix OCN that can be switched on and off, providing OBO efficiency enhancement at various levels. However, the bandwidth limitations and high baseband system complexity per signal chain make the outphasing PA impractical for broadband 5G signals and MIMO applications.

2.3.3 Doherty Power Amplifiers

The DPA is an OBO efficiency enhancement technique based on load modulation occurring entirely in the RF domain without any signal processing. The basic DPA architecture uses two transistors, the main (or carrier), and the auxiliary (or peaking), connected as shown in Figure 2.14 [37].

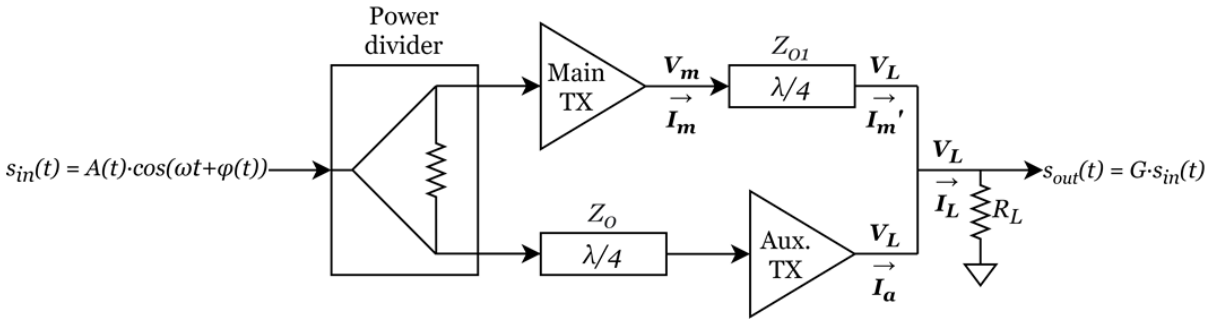


Figure 2.14 Block diagram of a conventional DPA

The basic DPA operation is derived in [37] and presented in this section. The main transistor is designed to reach peak efficiency at a certain OBO level P_{OBO}, while the auxiliary transistor is off. Then, for $P_{OBO} \leq P_{out} \leq P_{sat}$, the auxiliary transistor begins to turn on, modulating the load impedance seen by the main transistor to maintain its peak efficiency until P_{sat}, where both transistors reach peak efficiency. Thus, the overall DPA has efficiency peaks at P_{OBO} and P_{sat}.

Proper DPA operation requires the main and auxiliary transistor currents I_m and I_a to follow the profiles in equations (2.40) and (2.41), respectively. I_m is proportional to the normalized input voltage v_{in} and reaches its peak value $I_{m,sat}$ when $v_{in} = 1$. I_a remains 0 (off) until $v_{in} = 1/\beta$, where β represents the OBO level, then linearly rises towards its peak value $I_{a,sat}$ at $v_{in} = 1$.

$$I_m = I_{m,sat} \cdot v_{in}, \qquad 0 \leq v_{in} < 1 \tag{2.40}$$

$$I_a = \begin{cases} 0, & 0 \leq v_{in} \leq 1/\beta \\ \dfrac{I_{a,sat}}{1 - 1/\beta} \cdot (v_{in} - 1/\beta), & \dfrac{1}{\beta} \leq v_{in} \leq 1 \end{cases} \tag{2.41}$$

To ensure the DPA reaches peak efficiency at P_{OBO} ($v_{in} = 1/\beta$) and P_{sat} ($v_{in} = 1$), the fundamental voltage amplitude of the main transistor (V_m) must be maximized at P_{OBO} and P_{sat}, and that of the auxiliary transistor (V_L), at P_{sat}. Let R_{opt} be the optimum load resistance of the main transistor at P_{sat}. Then, the optimum load resistance of the main transistor at P_{OBO} and aux. transistor at P_{sat} are given by $R_{opt,OBO}$ and $R_{opt,aux}$, respectively.

Since $I_{m,OBO}$ is β times lower than $I_{m,sat}$, the following equation for $R_{opt,OBO}$ is obtained:

$$R_{opt,OBO} = \beta \cdot R_{opt} \tag{2.42}$$

Since $I_{m,sat}$ and $I_{a,sat}$ may not be equal, the values of $R_{opt,aux}$ and R_{opt} are related as follows:

$$R_{opt,aux} = R_{opt} \frac{I_{m,sat}}{I_{a,sat}} \tag{2.43}$$

It can be shown that the circuit in Figure 2.14 can be designed to present the optimum load impedances to the transistors. Consider the ABCD matrix of the $\lambda/4 = 90°$ TL with $Z_0 = Z_{01}$:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} cos\theta & j \cdot Z_{01} \cdot sin\theta \\ j/Z_{01} \cdot sin\theta & cos\theta \end{bmatrix} = \begin{bmatrix} 0 & j \cdot Z_{01} \\ j/Z_{01} & 0 \end{bmatrix} \tag{2.44}$$

Thus, the voltage and current relationships of the DPA can be written as follows, where bolded symbols represent voltage and current phasors:

$$\begin{bmatrix} \boldsymbol{V_m} \\ \boldsymbol{V_L} \end{bmatrix} = \begin{bmatrix} jZ_{01} \cdot \boldsymbol{I'_m} \\ -jZ_{01} \cdot \boldsymbol{I_m} \end{bmatrix} \tag{2.45}$$

The appropriate phases of $\boldsymbol{I'_m}$ and $\boldsymbol{V_L}$ are set using the $\lambda/4$ TLs at the main transistor output and auxiliary transistor input. Thus, the analysis can be performed using only magnitudes. Equation (2.45) becomes the following:

$$\begin{bmatrix} V_m \\ V_L \end{bmatrix} = \begin{bmatrix} Z_{01} \cdot I'_m \\ Z_{01} \cdot I_m \end{bmatrix} \tag{2.46}$$

Applying Kirchhoff's current law and Ohm's law at node V_L results in the following relationship:

$$I'_m = I_L - I_a = \frac{V_L}{R_L} - I_a \tag{2.47}$$

The impedance seen by each transistor is the ratio of its drain voltage to drain current. Using equations (2.46) and (2.47), the impedances Z_m and Z_a can be written as:

$$Z_m = \frac{V_m}{I_m} = \frac{Z_{01}(V_L/R_L - I_a)}{I_m} = \frac{Z_{01}(Z_{01}I_m/R_L - I_a)}{I_m} = \frac{Z_{01}^2}{R_L} - Z_{01}\frac{I_a}{I_m} \tag{2.48}$$

$$Z_a = \frac{V_L}{I_a} = Z_{01}\frac{I_m}{I_a} \tag{2.49}$$

Setting Z_m, Z_a to their optimum values from equations (2.42) and (2.43) results in the following:

$$Z_{m,OBO} = R_{opt,OBO} \rightarrow \frac{Z_{01}^2}{R_L} - Z_{01}\frac{I_{a,OBO}}{I_{m,OBO}} = \frac{Z_{01}^2}{R_L} = \beta R_{opt} \tag{2.50}$$

$$Z_{m,sat} = R_{opt} \rightarrow \frac{Z_{01}^2}{R_L} - Z_{01}\frac{I_{a,sat}}{I_{m,sat}} = R_{opt} \tag{2.51}$$

$$Z_{a,sat} = R_{opt,a} \rightarrow Z_{01}\frac{I_{m,sat}}{I_{a,sat}} = R_{opt}\frac{I_{m,sat}}{I_{a,sat}} \tag{2.52}$$

Solving equations (2.50)–(2.52) results in the following values for Z_{01}, R_L, and I_a/I_m:

$$Z_{01} = R_{opt} \tag{2.53}$$

$$R_L = \frac{R_{opt}}{\beta} \tag{2.54}$$

$$\frac{I_{a,sat}}{I_{m,sat}} = \beta - 1 \tag{2.55}$$

The conventional DPA design uses equal-sized devices with $I_{a,sat} = I_{m,sat}$, with $\beta = 2$. The auxiliary transistor turns on at half the input drive level. Figure 2.15 a) and b) plot the corresponding current and voltage profiles, respectively. V_m rises linearly from $v_{in} = 0$ to $v_{in} = 0.5$, then maintains its peak value, as desired. V_L rises proportionally to v_{in} which shows a useful feature of the DPA: it is theoretically linear, and any nonlinearities will be due to transistor nonidealities.

Figure 2.15 c) plots the load modulation of the main and auxiliary transistors: Z_m varies from $R_{opt,OBO} = 2R_{opt}$ to R_{opt} and Z_a varies from ∞ to $R_{opt,aux} = R_{opt}$. Figure 2.15 d) plots P_{out} for both transistors and the overall DPA against P_{in}, showing that the auxiliary transistor supplies the additional power to guarantee DPA linearity despite the main transistor load modulation. Figure 2.15 e) plots η_D for the main and auxiliary transistor, as well as the overall DPA, against P_{out}. This plot highlights the two efficiency peaks in the DPA, occurring at P_{sat} and P_{OBO}, which is 6 dB below P_{sat}. The efficiency of a class B PA is also plotted for comparison, showing the DPA efficiency enhancement factor of 2 at 6 dB OBO.

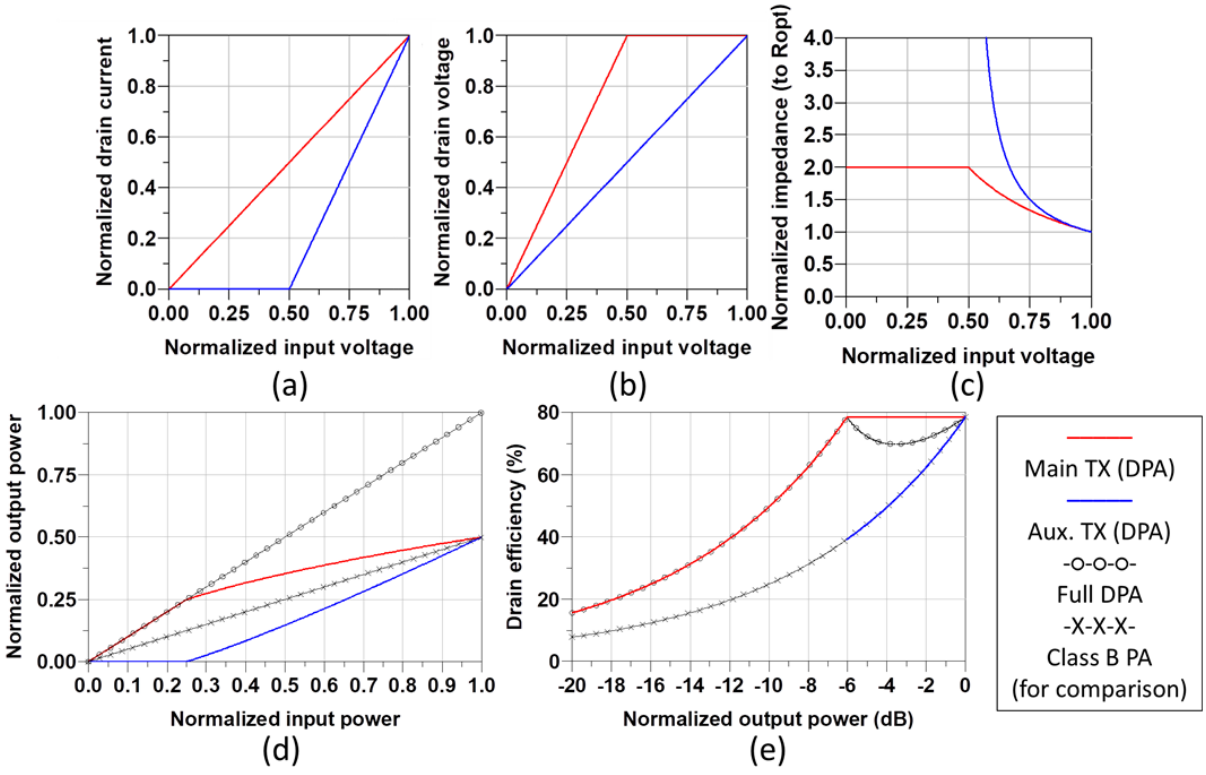


Figure 2.15 Conventional DPA: a) Current profiles, b) Voltage profiles, c) Load modulation, d) Power profiles, e) Efficiency profiles

The DPA architecture is very suitable for amplifying modern high-PAPR communication signals with high power efficiency, especially in base station applications [1]–[6], [8]–[26], [44], [46]. The main advantage of the DPA over other OBO efficiency enhancement techniques is its simplicity: it is a single input, single output architecture requiring no baseband signal processing unlike EER, ET, and outphasing PAs [1], [11]. This is highly beneficial in MIMO systems with multiple low-power PAs: any additional baseband processing would have to be multiplied by the number of PA units, which would take excessive power and space. Therefore, the DPA is the selected OBO efficiency enhancement technique in this work.

However, the conventional DPA formulation presented above suffers from bandwidth limitations: the $\lambda/4$ TL has frequency-dependent characteristics, the output capacitance of the auxiliary transistor cannot be absorbed, and broadband phase alignment is difficult [1], [4], [8], [9], [14], [17], [19]–[21], [23]–[24], [44], [46]. Moreover, the conventional DPA with symmetric transistors is limited to 6 dB OBO efficiency enhancement: for higher PAPR signals, performance is limited [12], [44]. Furthermore, several non-idealities of real transistors lead to high nonlinearity in practical DPAs. The next section will cover practical DPA design challenges.

2.4 Practical DPA Design Challenges

2.4.1 Main and Auxiliary Device Biasing

For synthesizing the main current I_m, which is proportional to v_{in}, a class B bias should be used in theory since it is perfectly linear. However, this assumption is based on an ideal FET transfer characteristic. The transfer characteristic of a practical Cree CGHV1F006 6 W 15 GHz transistor [47] is shown in Figure 2.16 a), exhibiting a soft transition between cutoff, linear, and saturation regions. This leads to a nonlinear g_m as shown in Figure 2.16 b). Due to the soft turn-on, class B is very nonlinear at low power in practice. However, in deep class AB ($V_{GS} \approx -2.7\,V$, near class B), the theoretical nonlinearity of the class AB bias can be used to cancel the practical nonlinearity of the transistor, leading to the flattest g_m curve. This biasing point is selected for the best overall linearity.

The synthesis of I_a is significantly more complex. To obtain the required I_a curve with a turn-on point of $v_{in} = 1/\beta$, the transistor must be biased in class C. However, the peak transconductance of class C is significantly lower than in class B: for example, for $\beta = 2$, I_a will only reach $I_{max}/5$ instead of the required $I_{max}/2$ for a turn-on point of $v_{in} = 1/2$ [37]. Furthermore, the soft turn-on due to the nonlinear g_m is unavoidable in class C.

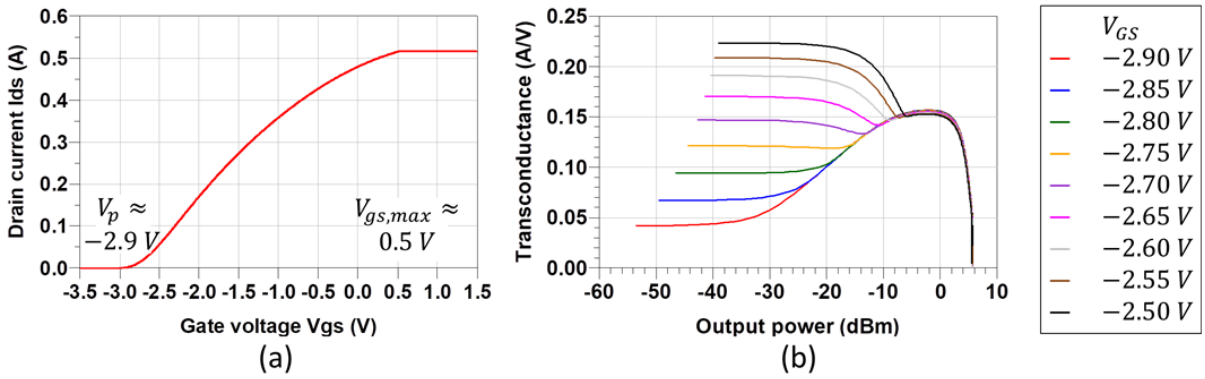


Figure 2.16 CGHV1F006 transistor: a) I_{ds} vs. V_{gs} transfer characteristic, b) g_m vs. P_{out} at various V_{gs}

[37] proposes solutions to the low auxiliary peak current. One option is to use an auxiliary transistor 2.5 times larger than the main so that both transistors reach the same peak current. However, discrete transistors with specific size ratios are usually not available. Furthermore, larger transistors have larger parasitics, which degrade the bandwidth of the circuit. Alternatively, the input power can be split unevenly between the main and auxiliary transistors. By setting the input voltage $V_{in,aux}$ 2.5 times higher than $V_{in,main}$, both devices will reach the same peak current. However, this implies that only 13% of the power reaches the main transistor, reducing gain by 5.6 dB. Also, the uneven power divider is more difficult to realize than symmetric dividers. Another solution is to control the main and auxiliary inputs separately: the mixed-signal DPA replaces the single input RF source and power divider with a dual-channel signal generator to exercise full control over the main and auxiliary input signal magnitude and phase relationships and synthesize the desired current profiles exactly. However, both adaptive biasing and mixed-signal DPA suffer from additional design complexity.

2.4.2 Device Capacitances and Package Parasitics

To represent the high-frequency behavior of the transistor, the simplified transistor model of Figure 2.3 must be augmented by adding parasitics components [37]–[39], as shown in Figure 2.17.

The intrinsic parasitics are part of the transistor die itself, and include gate-source, drain-source, and gate-drain capacitances: C_{gs}, C_{ds}, C_{gd}. These are nonlinear: their value depends on the applied voltages V_{gs} and V_{ds}. Figure 2.18 shows plots C_{gs}, C_{ds}, and C_{gd} vs. V_{gs} for the CGHV1F006 transistor.

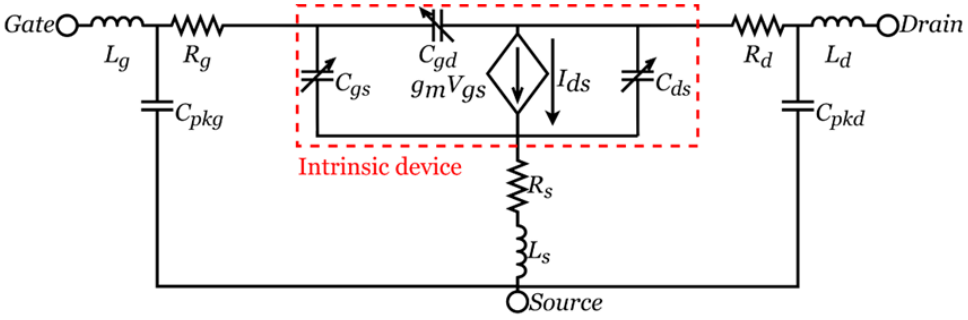


Figure 2.17 RF transistor equivalent circuit model

C_{gs} is highly nonlinear, with an abrupt change between the cutoff and linear regions. This nonlinearity generates harmonics in V_{gs}, so it is crucial to design the gate biasing network (BN) to short-circuit the harmonics. The nonlinear C_{gs} also causes significant AM-AM and especially AM-PM distortion in PAs [37]–[39]. DPAs suffer from additional AM-PM distortion in the high-power region when the auxiliary transistor is turned on [2], [10], [13], [17], [25]. C_{ds} is more linear than C_{gs}. However, its presence is still detrimental for PA design because it shifts the ideally resistive load termination R_{opt} into a reactive region, which is frequency-dependent [37]–[39]. Furthermore, a large value of C_{ds} will decrease the achievable bandwidth of the matching network. It is imperative to explore topologies for the DPA OCN that allow absorption of C_{ds} and other drain parasitics to allow access to the main and auxiliary transistor intrinsic current sources and realize the load modulation over broad bandwidths.

The feedback capacitor C_{gd} is significant despite its small size [37]–[39]. Since it forms a path between the output and input, the transistor is bilateral, which leads to potential instability. Furthermore, C_{gd} creates a dependence of the OMN on the IMN and vice-versa, so the matching network design becomes iterative. C_{gd} also causes an additional capacitance at the transistor input due to the Miller effect, given by equation (2.56), where A_v is the voltage gain of the transistor at the frequency of interest.

$$C_{in} = C_{gs} + C_{gd} \cdot (1 + A_v) \tag{2.56}$$

The extrinsic parasitics are due to the packaging of the device and include bond wire resistance and inductance for all three terminals (L_x and R_x), as well as pad capacitance for the gate and drain (C_{pg} and C_{pd}) [37]–[39]. Given that these are physical components, not model parameters, they are linear components. Thus, they do not cause parasitics, but degrade the bandwidth and cause an additional shift

in the design space of the matching network on the Smith chart. L_s and R_s, also contribute to decreasing the gain due to source degeneration.

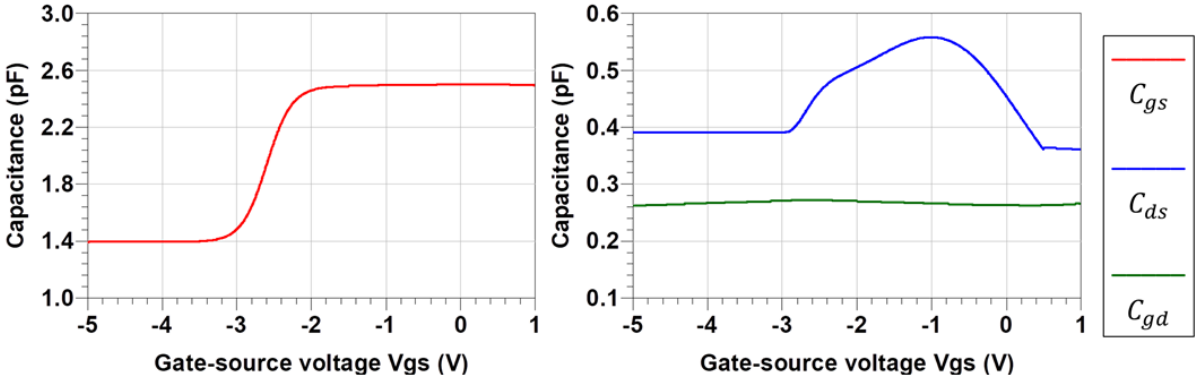


Figure 2.18 CGHV1F006 transistor intrinsic capacitances: a) C_{gs} vs. V_{gs}, b) C_{ds} and C_{gs} vs. V_{gs}

2.4.3 Knee Region Effects

The ideal assumption that $V_k = 0$ implies that the drain voltage can swing from 0 V to V_{max}. Hence, V_{DD} is set to $V_{max}/2$ for maximum voltage swing. However, in a practical PA, V_k represents a significant portion of V_{DD}. For example, for the CGHV1F006 transistor with V_{DD} set to 28 V, $V_k = 6\,V$ or approximately 20% of V_{DD}. To maintain linearity, the voltage must not be allowed to fall below V_k. Thus, the fundamental maximum voltage swing is given by the following [37]–[38]:

$$V_1 = V_{DD} - V_k$$

The value of R'_{opt} and P'_{sat} (optimum load impedance and saturated power when considering the knee region) are thus the following, assuming class B bias [37]–[38]:

$$R'_{opt} = \frac{V_1}{I_{max}/2} = 2\frac{V_{DD} - V_k}{I_{max}} \tag{2.57}$$

$$P'_{sat} = \frac{V_1 \cdot I_{max}/2}{2} = \frac{(V_{DD} - V_k) \cdot I_{max}}{4} = \frac{V_{max} \cdot I_{max}}{8} - \frac{V_k \cdot I_{max}}{4} \tag{2.58}$$

The second term in equation (2.58) shows that the knee region effect reduces P_{sat} (and η_D) significantly compared to the ideal class A or B. It is possible to trade off linearity for efficiency by choosing a load resistance between the original R_{opt} and R'_{opt}, allowing a slight intrusion into the knee region [37]–[38].

2.5 Literature Review of Sub-6 GHz Broadband High-Frequency DPAs

This section highlights the state-of-the-art DPAs in the literature that overcome the bandwidth limitations of the conventional DPA. To restrict the scope, the review focuses on DPAs that feature a single OBO efficiency peak. Multi-way DPAs also exist, with multiple auxiliary transistors that turn on at different OBO levels to provide more than two OBO efficiency peaks [37], [45].

2.5.1 Broadband DPAs

[19] presents two 2.2–3.0 GHz, 40–42 dBm DPAs with 5–6 dB OBO where the OMN of each transistor and OCN are synthesized with the simplified real frequency technique (SRFT). The technique has a limitation: it can only maximize P_{out} and η_D at OBO or at saturation, leading to two distinct designs. Overall, neither of the two designs show proper DPA behavior: the efficiency curves show limited OBO η_D enhancement and performance is variable across the 2.2 to 3.0 GHz band.

[9] extends the work in [19] with a 2.2–3.7 GHz, 43–45 dBm, 6 dB OBO DPA realized using SRFT. The proposed dual-transformation SRFT algorithm can optimize P_{out} and η_D at both OBO and saturation. The resulting DPA achieves good OBO η_D enhancement behavior across the entire bandwidth, with some degradation in the low-frequency bands, which were not fully optimized.

[46] features a 3.0–3.6 GHz, 43–44 dBm broadband DPA with 6 dB OBO. The authors first match the individual transistors to 50 Ω by compensating the transistor parasitics (C_{ds} and L_D). The OCN uses the conventional DPA architecture in Figure 2.14 with an additional $\lambda/4$ TL to match the output from R_L to 50 Ω. However, the DPA behavior falls off quickly across the band: the auxiliary transistor turns on too late at 3.0 GHz and too early at 3.6 GHz, degrading η_D at OBO. Thus, this architecture cannot be considered truly broadband.

[8] features a 1.7–2.6 GHz, 45–46 dBm, 10 dB OBO post-matching DPA. The authors claim that the approach in [46], where the transistors are first matched to 50 Ω, is narrowband by nature. Instead, each transistor's drain parasitics are absorbed into a $\lambda/4$ TL, and an additional $\lambda/4$ TL is added in the auxiliary branch. Then, a high-order, multi-section post-matching network (PMN) matches the output to 50 Ω. This approach guarantees a broadband behavior because the phase behavior of the $\lambda/4$ OMNs is predictable, unlike the OMNs in [46], which require phase offset TLs and only work over narrow bandwidths. Indeed, this work shows very good OBO η_D enhancement and very consistent performance over the entire bandwidth, so it can be considered a broadband DPA.

[21] presents a 1.7–2.8 GHz, 44 dBm, 6 dB OBO DPA with compensating reactance ($\lambda/4$ stub in parallel with R_L). The authors show that this technique broadens the DPA bandwidth at OBO, with a small penalty at P_{sat}. By moving a series compensating reactance in the auxiliary branch, the P_{sat} bandwidth can be restored while keeping the broader OBO bandwidth. This technique can also absorb the drain parasitics of the auxiliary transistor. The resulting DPA shows very consistent OBO η_D enhancement due to the auxiliary transistor parasitic absorption. However, the η_D profile shows incomplete load modulation in the low-frequency band.

[48]–[49] developed a generalized theory for the design of broadband DPAs, with a 2.1–2.6 GHz, 41–45 dBm, 6 dB OBO DPA demonstrator. The OCN consists of a $Z_0 = Z_{01}$, $\lambda/4$ TL in the main path and two $\lambda/4$ TLs in the auxiliary path, one with $Z_0 = Z_{03}$ followed by one with $Z_0 = Z_{02}$. The OCN is terminated with a resistor R_L. This configuration is known as inverted DPA because the main transistor output lags the auxiliary by 90°, unlike the conventional DPA where the auxiliary lags the main. The values of Z_{01}, Z_{02}, Z_{03}, and R_L are based on the parameter α as shown below [48]–[49]:

$$Z_{01} = R_{opt}/\alpha \tag{2.59}$$

$$R_L = \frac{R_{opt}}{2\alpha^2} \tag{2.60}$$

$$\frac{Z_{03}}{Z_{02}} = \alpha \tag{2.61}$$

This theory has two applications for broadband DPAs. First, it can be used to reduce the impedance transformation ratio (ITR) of the main transistor towards unity, which improves bandwidth (the conventional DPA has an ITR of 4). Figure 2.19 shows the relationship between α, ITR, and bandwidth at OBO. Second, the value of α can be used to engineer the value of R_L to 50 Ω based on the transistor's R_{opt}, avoiding the need for a PMN, which is another source of bandwidth limitations. In [48], the authors maximized the bandwidth with $\alpha = 0.5$, but despite the strong theoretical basis, the experimental performance is mediocre (large variability of P_{sat} and η_D across the bandwidth).

[17] pushes the inverted DPA concept further by presenting a 1.5–3.8 GHz, 42–43 dBm, 6 dB OBO DPA, the highest FBW reported in the literature. The transistor parasitics are absorbed into $\lambda/4$ TLs by using low-pass Π networks where one half of the "Π" is represented by the transistors' C_{ds} and L_d, and the other half is transformed into a series TL and shunt stub. An additional $\lambda/4$ TL is added in the auxiliary transistor path. The DPA behavior across the bandwidth is quite inconsistent, with most

frequency bands not exhibiting sufficient OBO η_D enhancement. This demonstrates that proper DPA function is difficult to obtain over a very broad bandwidth.

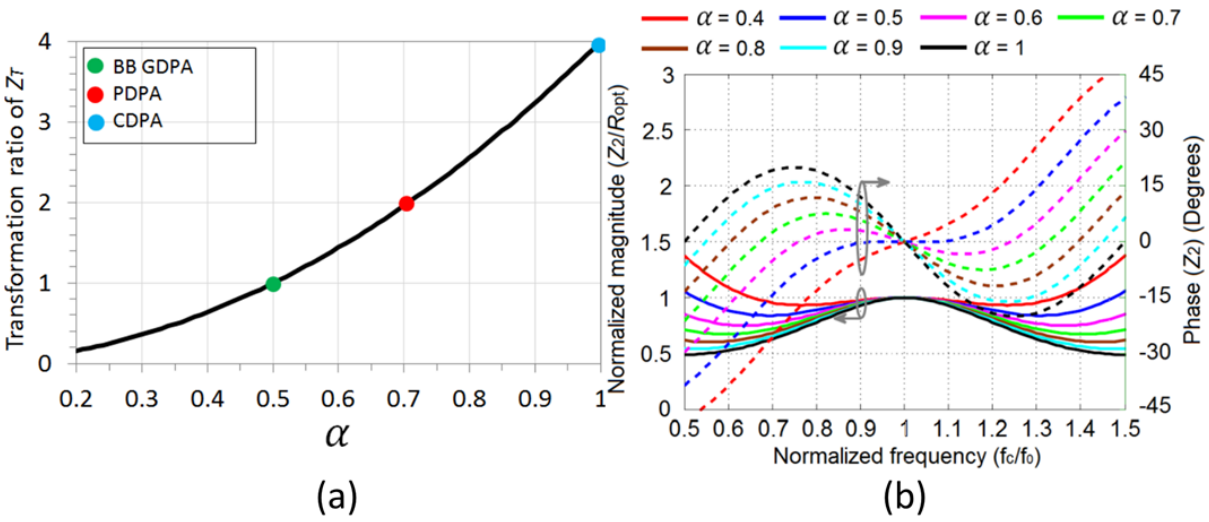


Figure 2.19 Inverted DPA ITR, α, and bandwidth: a) ITR vs. α, b) Bandwidth vs. α [49]

[12] generalizes the inverted DPA concept to accommodate arbitrary transistor sizes and OBO level and demonstrates the theory with a 2.55–3.35 GHz, 44–46 dBm, 8 dB OBO DPA. The characteristic impedances of the $\lambda/4$ TLs in the OCN are given in equations (2.62)–(2.64), where β is the selected OBO level and α is the ratio between auxiliary and main transistor sizes.

$$Z_{01} = \beta \cdot R_{opt} \tag{2.62}$$

$$Z_{02} = \sqrt{\frac{\beta \cdot \alpha}{\beta - 1}} R_{opt} \tag{2.63}$$

$$Z_{03} = \alpha \cdot R_{opt} \tag{2.64}$$

The $\lambda/4$ TLs are realized using OMNs that absorb the transistor drain parasitics. The resulting design achieves a favorable trade-off between bandwidth at OBO and at P_{sat}, with good OBO η_D enhancement over the bandwidth, except in the very low frequency bands. Under wideband modulated signals (100 MHz OFDM, 8.0 dB PAPR), the worst-case ACPR is -26 dBc/-52 dBc before and after DPD, demonstrating the DPA's linearity.

[11] proposes a 2.80–3.55 GHz, 43–45 dBm, 6 dB OBO DPA using a modified inverted DPA topology. The characteristic impedances of the $\lambda/4$ TL-like OMNs are modified to $Z_{01} = Z_{02} = \sqrt{2 \cdot R_{opt}}$, and $Z_{03} = R_{opt}$, with a load resistance $R_L = 2 \cdot R_{opt}$. The phase slopes of the OMNs are

not constant like a $\lambda/4$ TL but are manipulated to achieve a constant value of $Z_{m,OBO} = 2R_{opt}$ over the bandwidth for very consistent OBO performance (but $Z_{m,sat}$ and $Z_{a.sat}$ vary over the bandwidth unlike in the conventional DPA). The resulting DPA shows very consistent performance across the bandwidth at OBO, but performance varies significantly at saturation (up to 2 dB variation in P_{sat}). Under broadband modulated signals (120 MHz OFDM, 7.0 dB PAPR), worst-case ACPR is -30 dBc/ -51 dBc before and after DPD.

[1] features a 3.3–3.8 GHz, 48–49 dBm, 8 dB OBO DPA where the authors apply the continuous class B/J/J* concept to the main transistor to extend the bandwidth. At P_{sat}, the conventional class J/J* fundamental and second harmonic impedances Z_1 and Z_2 given in equations (2.32) and (2.33) are used. The authors also derive the required Z_1 and Z_2 for an arbitrary OBO level β, which are given in equations (2.65)–(2.67) where $-1 < \alpha < 1$. Figure 2.20 plots Z_1 and Z_2 for the class J DPA.

$$Z_{1,OBO} = \frac{2}{\zeta} \cdot R_{opt} \cdot (1 + j\alpha) \tag{2.65}$$

$$Z_{2,OBO} = -j\frac{3\pi}{4 \cdot \zeta}\alpha \cdot R_{opt} \tag{2.66}$$

$$\beta = 6 + |10 \cdot \log_{10}(\zeta)| \tag{2.67}$$

The authors then devise a DPA OCN that decouples fundamental and second harmonic impedance matching, allowing the realization of the desired Z_1 and Z_2. The resulting DPA shows very good OBO η_D enhancement and consistent performance over most of the bandwidth.

[23] improves upon the work in [1] with a 2.7–4.3 GHz, 39 dBm, 6 dB OBO class J DPA. The authors propose an ABCD matrix formulation of the transistors OCNs that can realize the required class J/J* impedances while absorbing the transistor parasitics to reach the drain current source plane, which had not been fully explored in [1]. The resulting design had very consistent OBO η_D enhancement in simulation, but incomplete EM simulation caused the measured performance to degrade. This design's linearity is proven by testing with broadband modulated signals (80 MHz carrier aggregated, 7.0 dB PAPR): the ACPR is -22 dBc and-50 dBc before and after DPD.

[13] proposes a 4.7–5.3 GHz, 39 dBm, 9 dB OBO DPA with reduced AM-PM distortion. First, a conventional DPA based on $\lambda/4$ impedance inverter is designed, and its AM-PM distortion at P_{sat} ($\Delta\theta_C$) is obtained. Then, a new OCN is designed to provide an AM-PM of $-\Delta\theta_c$ to cancel the transistor's AM-PM. The authors provide expressions for design parameters α_c (transistor size ratio),

β_d (OBO level), and $\theta_{M,A}$ (electrical length of TLs in the main and auxiliary transistor branches). Since the transistor sizes are unequal ($\alpha_c \neq 1$), the values of Z_m, Z_a, and R_L are the following [13]:

$$Z_m = R_{opt} \tag{2.68}$$

$$Z_a = \frac{R_{opt}}{\alpha_c} \tag{2.69}$$

$$R_L = \frac{R_{opt}}{1 + \alpha_c} \tag{2.70}$$

An issue with this technique is that α_c and β_d depend on $\Delta\theta_C$, so it is possible that the transistor size ratio required is unavailable, or that the required OBO level is not adequate for the signal's PAPR. Furthermore, θ_M and θ_A take arbitrary values instead $\lambda/4$ and $\lambda/2$, which makes this technique incompatible with the inverted DPA bandwidth extension technique.

The DPA designed in [13] achieved flat AM-PM in the high-power region, as desired, and very consistent performance over the bandwidth. However, the OBO η_D enhancement level was closer to 7 dB due to the use of symmetric devices, which is incompatible with 9 dB OBO. This design has exceptional linearity: under very complex broadband modulated signals (160 MHz intra-band carrier aggregated, 7.4 dB PAPR), worst-case ACPR is as low as -36 dBc/-50 dBc before and after DPD.

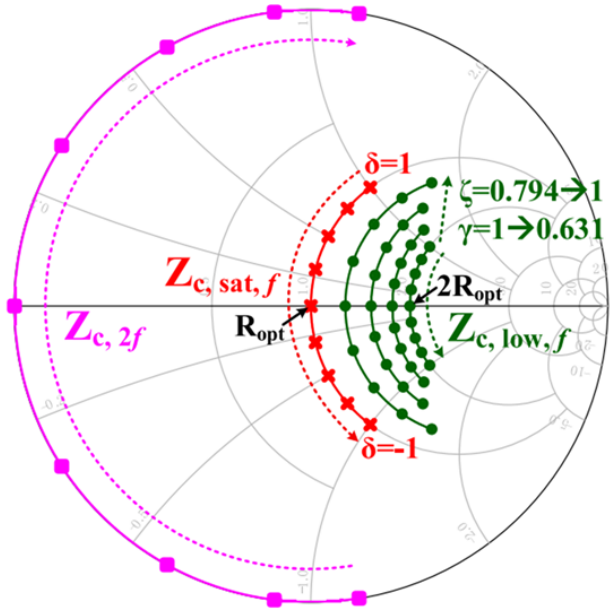


Figure 2.20 Continuous class J/J* DPA main transistor impedances at OBO and saturation [1]

Table 2.3 shows a summary of the broadband DPAs surveyed in this literature review.

Table 2.3 Comparison of state-of-the-art broadband DPAs in the literature

Ref	Bandwidth (GHz)	P_{sat} (dBm)	OBO (dB)	$\eta_{D,sat}$ (%)	$\eta_{D,OBO}$ (%)	Gain (dB)	ACPR (dBc)[a] Bandwidth (MHz) PAPR (dB)			Transistor
[19]	2.2–3.0	39.5–41.8[b]	6	50–67[b]	36–47[b]	7.4–8.8[b]	-23	4.2	10.0	CGH40006
[9]	2.2–3.7	43.0–44.6	6	55–69	45–53	10.2–11.8[c]	-30	20	6.2	CGH40010
[46]	3.0–3.6	43.0–44.0	6	55–66	38–56	11.5–12.8[b]	—			CGH40010
[8]	1.7–2.6	44.6–46.3	10[c]	57–66	41–49[b]	10.2–11.6	-25	20	10.5	CGH40025
[21]	1.7–2.8	44.0–44.5	6[c]	57–71	50–55	12.0–14.5	-29	20	6.5	CGH40010
[48]	2.1–2.6	41.0–45.0	6	60–85	50–63	9.0–10.0	—			CGH40010
[17]	1.5–3.8	42.3–43.4	6	42–63	33–55	10.0–13.8	-42	7.0	9.0	CGH40010
[12]	2.55–3.35	44.3–45.4	8	58–76	48–59	9.2–10.4	-26	100	8.0	CGH40010
[11]	2.8–3.55	43.0–45.0	6	66–78	50–61	8.3–9.1	-30	120	7.0	CGH40010
[1]	3.3–3.75	48.0–48.8	8	58–71	44–55	11.8–13.5	-30	40	8.0	CGHV27030
[23]	2.7–4.3	38.5–39.2	6	48–61	40–43	8.0–10.0[c]	-22	80	7.0	CGHV1F006
[13]	4.7–5.3	39–39.5	9	52–57	30–33	7.7–8.2	-36	160	7.4	CGHV1F006

[a]Worst-case ACPR reported over the bandwidth

[b]Estimated from graph

[c]Determined from the graph of η_D vs. P_{out}

2.5.2 Multi-band DPAs

Multi-band DPAs are an alternative to broadband DPAs in applications that must cover two relatively narrow, but widely spaced bands. Two types of multi-band DPAs are encountered: concurrent multi-band, where the MNs are designed to cover both bands, and reconfigurable multi-band, where the second band is accessible by toggling switches to modify the circuit's behavior. The work in [4] even combines both types of multi-band architectures to achieve a total of six bands, the most in any DPA.

Concurrent multi-band designs such as [2] and [5] realize the transistor OMNs using T or Π topologies, which act as impedance inverters at two distinct, independently selected frequencies. The design in [5] employs short-circuited Π networks instead of the conventional open-circuited ones, while the design in [2] uses multi-stub OMNs. However, both designs are quite narrowband: [2] has only 80–110 MHz per band, and [5] increases this to 250–420 MHz per band. This bandwidth is insufficient when considering the two bands of interest in this thesis, n77 (3.3–4.2 GHz) and n79 (4.4–5.0 GHz). Indeed, multi-band DPAs are not a suitable choice when the two bands are broad and very close together.

Reconfigurable multi-band DPAs either have reconfigurable OMNs ([14]) or PA biasing voltages ([3], [24]). In [14], two auxiliary branches each provide load modulation for 1.8–2.7 GHz and 2.7–3.4 GHz, respectively. The desired branch is selected using switches. This technique is interesting if the two bands are closely spaced like in this thesis, but [14] shows very poor RF performance and inconsistency at the junction between both bands, so it is not a viable approach. In [3] and [24], the reconfigurability is achieved by swapping the roles of the main and auxiliary transistors. In [3], the reconfiguration leads to two additional narrow sidebands near f_0 with poor RF performance compared to the main bands, which is not beneficial in this thesis because 3.3–5.0 GHz can already be covered by a single wide band. In [24], the reconfiguration unlocks a second frequency band centered at $0.5f_0$, which is also not useful in this thesis.

2.5.3 DPA-like Architectures

New state-of-the-art load modulation architectures have also been proposed in the literature. These topologies aim to achieve DPA-like load modulation while avoiding some of the DPA's limitations like the difficulty to extend the OBO level, which requires specific transistor size ratios.

An alternative to the DPA is the load modulated balanced amplifier (LMBA), where the main PA is a balanced amplifier (BA). An auxiliary control amplifier (CA) drives the isolated port of the output hybrid to modulate the impedance seen by the transistors in the BA. The BA in [20] features a single-input, single-output LMBA with a very wide bandwidth (1.8–3.8 GHz), but performance varies significantly over the band and OBO η_D enhancement is weak. The authors in [22] and [44] propose sequential LMBAs with 10 dB OBO. In the sequential LMBA, the CA plays the role of the main amplifier, and the BA is the auxiliary. This technique is naturally suited to realize high OBO level PAs since the BA (two transistors) has a higher current than the CA. Furthermore, OBO η_D enhancement is excellent in both [22] and [44]. However, the main transistor needs to be overdriven for proper OBO η_D enhancement, which shortens its lifespan and reduces the long-term viability of this technique.

Reconfigurable architectures have also been proposed for resiliency against load mismatch and antenna crosstalk, which is particularly important in MIMO applications where the PAs are directly connected to the antenna array. In [10], the authors propose a balanced-to-Doherty (B2D) DPA where the OCN can be switched between a conventional DPA OCN (DPA mode) and a 90° hybrid (balanced mode). DPA mode is the default configuration since it allows OBO η_D enhancement, but balanced mode can restore linearity to acceptable levels at certain load phase angles. In [25], the authors propose

a reconfigurable quasi-balanced (QB) DPA where the OCN can be reconfigured between conventional (series mode) and inverted (parallel mode). Series mode maintains better linearity and efficiency for $R_L < Z_0$ (left side of Smith chart), and parallel mode, for $R_L > Z_0$ (right side of Smith chart). Thus, the proper mode can be selected based on the load phase angle. Despite the obvious benefits for MIMO operation, the designs in [10] and [25] both had limited bandwidth (3.4–3.6 GHz) and are not suitable to cover the 3.3–5.0 GHz range in this thesis.

2.5.4 Literature Review Conclusions

After reviewing broadband, multi-band, and DPA-like architectures, the broadband DPA will be selected for this work. When considering the two closely spaced and relatively wide bands of interest (3.3–4.2 GHz and 4.4–5.0 GHz), it makes the most sense to cover them with a single broadband 3.3–5.0 GHz DPA. The corresponding 41% FBW is easily achievable, even at similar frequency ranges (for example, [23] covered a 2.7–4.3 GHz range, corresponding to a 47% FBW).

Among broadband DPA techniques, a few are unsuitable. Employing optimization-based techniques such as SRFT did not produce the best RF performance in [9], [19] and does not provide a proper insight into the behavior of the circuit, so it is difficult to determine whether DPA-like load modulation is truly occurring across the entire bandwidth. Instead, an approach that is well-grounded in theory is selected, such as the inverted DPA, which provided more tangible bandwidth improvements in [12], [13], [17], [23], [48]. The design will consist of an inverted DPA with low-order impedance inverters and a high-order PMN for the broadest bandwidth [8]. Continuous class B/J/J* termination is not necessary to achieve wide bandwidths as shown in [17], but an effort will be made to achieve purely reactive harmonic terminations to improve efficiency. The AM-PM compensation technique in [13] is difficult to implement over broad bandwidths and combine with other bandwidth-enhancing techniques, so it will likely not be considered.

Chapter 3
3.3–5.0 GHz Linearity-Enhanced DPA Design

This chapter will describe the design of a 3.3–5.0 GHz linearity-enhanced DPA for use in MIMO arrays. It will begin with design specifications for the DPA, then present a design strategy for the DPA. The details of each transistor design (BN, OCN, IMN, power divider) will be described, and the EM-simulated performance of the full DPA will be shown.

3.1 Design Specifications

The design specifications will be based on the conclusions in Chapter 2's literature review, as well as the targeted bandwidth, 3.3–5.0 GHz. For this frequency range, the exotic DPA-like architectures are not suitable because of their high complexity, which makes very high frequency design difficult due to additional parasitics (none of them reach 4.0 GHz). Furthermore, the multi-band DPAs will not be considered because they are best suited for two narrow, widely spaced frequency bands: in this work, the bands of interest are 3.3–4.2 GHz and 4.4–5.0 GHz, with only a 200 MHz gap in the middle, so this situation does not apply. The desired frequency band, 3.3–5.0 GHz, is only 41% FBW, which is easily achievable with conventional broadband architectures. This leaves the broadband DPA. The inverted DPAs employing a PMN architecture offered the best bandwidth characteristics, so this topology will be selected for the OCN. The class J transistor bias might be considered if achieving the desired bandwidth is difficult, and the AM-PM compensation technique might be considered if AM-PM cannot be improved at f_0 using other means.

The frequency range of 3.3–5.0 GHz limits the choice of transistor: the only reported design above 5.0 GHz uses the CGHV1F006 transistor [13]. Since this work must also reach 5.0 GHz, the CGHV1F006 transistor will be selected for its superior high-frequency performance. The transistor choice also influences the OBO level: anything above 6 dB requires a larger auxiliary transistor for optimum DPA performance. Since the Cree 15 GHz family includes only the CGHV1F006 6 W transistor and CGHV1F025 25 W transistor [50] (which is overkill), the design will be based on symmetric CGHV1F006 main and auxiliary transistors and target an OBO level of 6 dB.

The choice of transistor and frequency range influence the selection of the PCB substrate. Since the CGHV1F006 transistor is a high-voltage (40 V), relatively low-power (6 W) device, the optimum load resistance is estimated to be $R_{opt} \approx 133\ \Omega$ according to equation (3.1) (a simplified formula ignoring

knee voltage). Given this high value, a high substrate thickness is required to achieve $Z_0 \approx R_{opt}$ for the TL in the OCN: a 32-mil Rogers RO4003C substrate is selected, with $\epsilon_r = 3.55$ and $\tan\delta = 0.0021$ at 2.5 GHz, which are very good characteristics for a design in this frequency range.

$$R_{opt} = \frac{V_{DD}^2}{2 \cdot P_{out}} \tag{3.1}$$

The targeted performance specifications are based on similar designs in the literature. In particular, [23] is also based on CGHV1F006 transistors for 6 dB OBO and the frequency range (2.7–4.3 GHz, 46%) is very similar to this work, although f_0 is lower. Thus, the targeted specifications will be based on those achieved in [23]: a small-signal gain of 8.0–10.0 dB, P_{sat} of 38.5–39.2 dBm, η_D of 48–61% at P_{sat} and 40–43% at 6 dB OBO. This work will attempt to achieve or exceed these specifications while covering the new 5G sub-6 GHz frequency bands.

3.2 Design Strategy

The first step in the design of any PA is transistor characterization. The CGHV1F006 transistor will be characterized using DC, stability, and load pull simulations, to determine the biasing points and value of R_{opt} for the design. Then, the BNs for the transistor drain and gate will be designed. The rest of the DPA design is performed from right to left, starting from the OCN, then the two transistor IMNs, and finally the input power divider.

Each network is first designed at the schematic level using microstrip TLs and ideal passive components. Then, the passives are be replaced with their equivalent circuit model (supplied by the manufacturer) and the circuit is tuned to recover the desired performance. RF-grade passive components are selected to minimize parasitics. Then, the schematic is converted to a layout and EM-simulated using Keysight ADS (Momentum). When all networks are EM-simulated, the full layout is exported to Ansys HFSS for full 3D EM simulation. Final changes are made to restore the original performance, then the layout is transferred to Altium Designer and fabrication files are exported.

3.3 Transistor Characterization

3.3.1 DC Characterization

The CGHV1F006 transistor class A DC load line is shown in Figure 3.1. This is used to determine R_{opt}, which is the same for class A and B. The pinch-off voltage is $V_p = -3.0\ V$, the knee voltage is

approximately $V_k = 7\ V$ and the corresponding drain current is $I_{max} = 600\ mA$. Based on $V_{DD} = 40\ V$, the maximum voltage swing would be from 7 to 73 V, corresponding to $R_{opt} \approx 110\ \Omega$ as shown in Figure 3.1 a). This results in a 0.3 mm TL width on the selected substrate, which is quite narrow for high current carrying traces at the PA output: realizing the DPA OCN, which requires $Z_0 \approx R_{opt}$ (and possibly higher), would be difficult. Hence, alternatively, the transistor can be biased at $V_{DD} = 28\ V$, resulting in a voltage swing from 7 to 49 V, $R_{opt} \approx 70\ \Omega$ as shown in Figure 3.1 b) and a 1 mm TL width, which relaxes requirements for the OCN realization. The trade-off is a loss of P_{sat} due to the lower voltage swing and is commonly done for this transistor: for example, [13] used $V_{DD} = 30\ V$ and [23], $V_{DD} = 28\ V$.

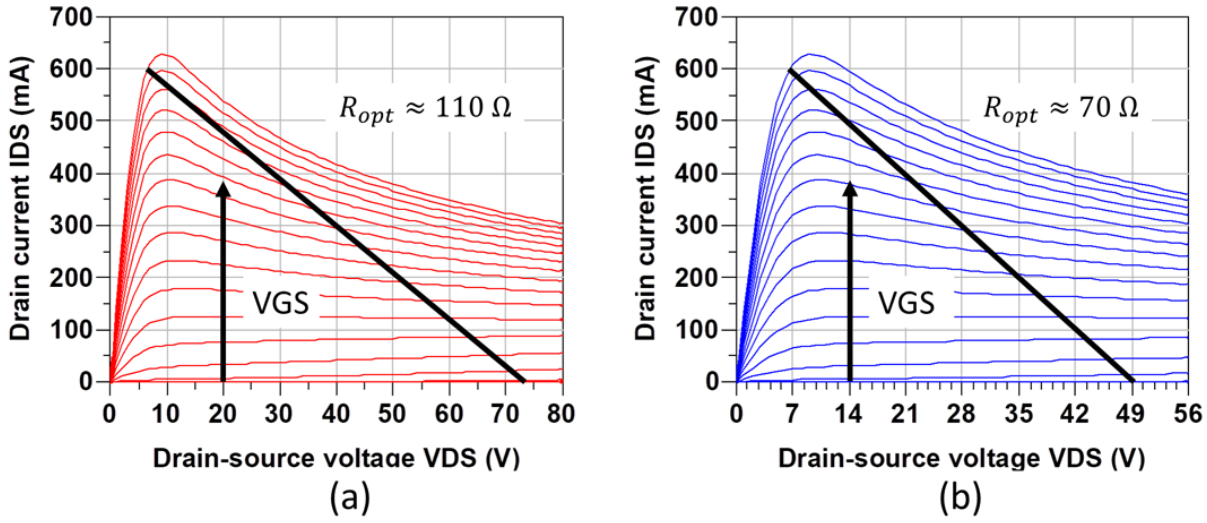


Figure 3.1 CGHV1F006 transistor load line and R_{opt} for: a) $V_{DD} = 40\ V$, b) $V_{DD} = 28\ V$

3.3.2 Stability Characterization

Before performing any large-signal tests, the transistor must be stabilized in small-signal conditions, with the network shown in Figure 3.2. Most importantly, the transistor must be stabilized at low frequencies, which is the function of resistor R_p. Furthermore, in the 3.3–5.0 GHz operating band, the load stability region of the transistor must be well outside the range of impedances that it will experience during operation. However, unconditional in-band stability is not desired because the maximum available gain (MAG) will be too low. In-band stabilization is the function of resistors R_s and capacitor C_s, which shunts R_s to maintain acceptable gain at high frequencies.

For stabilizing the CGHV1F006 transistor, the selected stabilization values are $R_s = 100\ \Omega$, $C_s = 1.4\ pF$, and $R_p = 50\ \Omega$ (note that these are not ideal components but use equivalent circuit models of practical components). As shown in Figure 3.3 a), these component values bring the stability factor at low frequencies (below 1 GHz) well above 1 (unconditional stability), while maintaining sufficient MAG within 3.3–5.0 GHz. For the 3.3–5.0 GHz range, the transistor is only conditionally stable, so source and load stability circles are plotted in Figure 3.3 b) to assess stability. Only loads inside these circles are unstable, and since they are on the left-hand side of the Smith chart, they are not concerning (the impedances seen by the main and auxiliary transistors are within $[R_{opt}, 2 \cdot R_{opt}]$ and $[R_{opt}, \infty]$, respectively. This concludes the stability study.

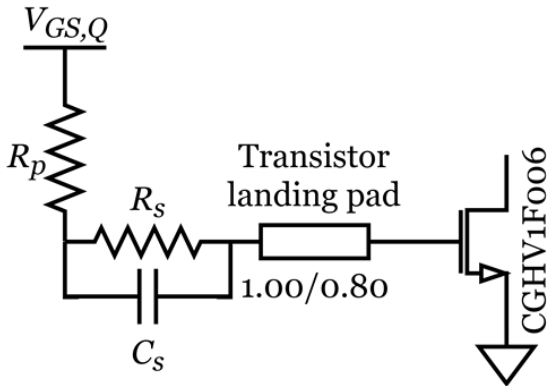


Figure 3.2 CGHV1F006 transistor stabilization circuit

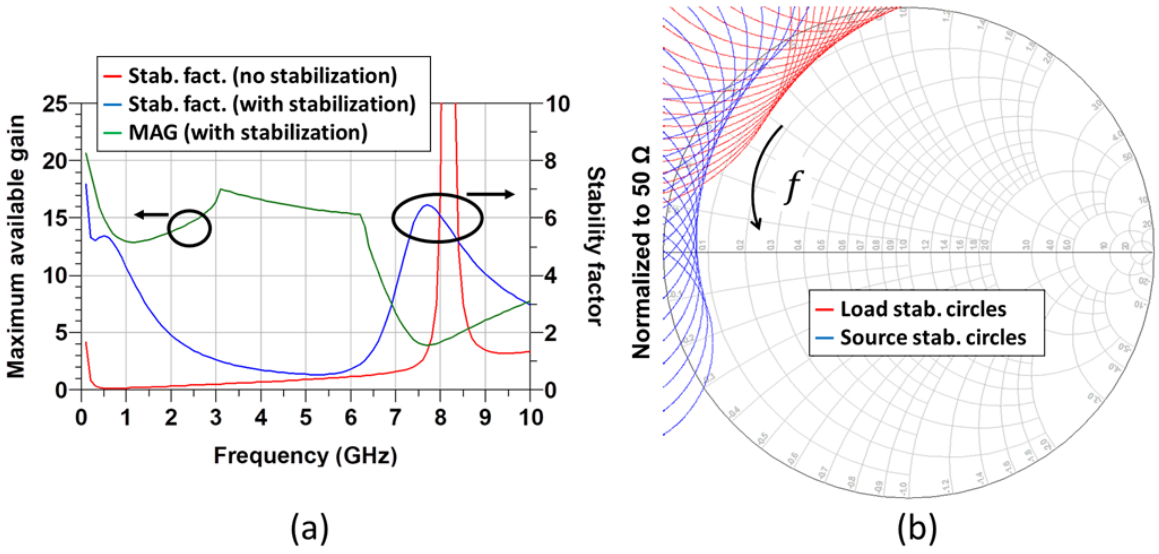


Figure 3.3 CGHV1F006 transistor: a) Stability factor/MAG, b) Load and source stability circles

3.3.3 Load Pull Characterization

After the approximate value of R_{opt} is known from DC analysis, load pull is performed to determine the exact value of R_{opt}, as well as the value of the nonlinear C_{ds}. The drain parasitic network is shown in Figure 3.4, and the values of the linear drain parasitic components (package parasitics) are obtained from ColdFET analysis ($L_{d,1} = 0.64\ nH$, $L_{d,2} = 0.05\ nH$, $C_{pd} = 0.22\ pF$) [23].

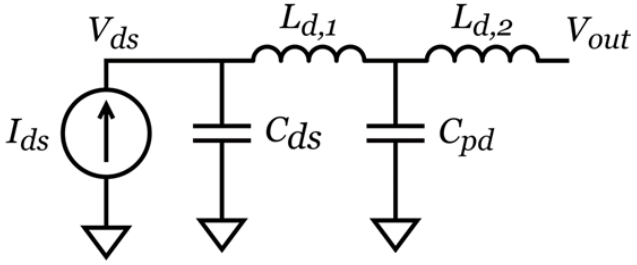


Figure 3.4 CGHV1F006 transistor drain parasitic network

The load pull is performed at the fundamental frequency, with source and load harmonics short-circuited, in accordance with class B bias conditions. The transistor is driven directly with a conjugate-matched source whose value is obtained by source pull. The package parasitics $L_{d,1}$, $L_{d,2}$, and C_{pd} are de-embedded before performing the load pull, so the optimum load impedance Z_{opt} obtained consists of a positive susceptance B_{out} to cancel C_{ds} in parallel with R_{opt}. Thus, R_{opt} and C_{ds} are determined as shown in equations $(3.2) - (3.4)$:

$$Z_{opt} = R + jX \tag{3.2}$$

$$R_{opt} = \frac{1}{G} = \left(\frac{R}{R^2 + X^2}\right)^{-1} \tag{3.3}$$

$$C_{ds} = -B_{out} = \frac{X}{R^2 + X^2} \tag{3.4}$$

Load pull is performed at various frequencies from 3.2 to 5.0 GHz in 0.45 GHz steps and at various power levels, which is important to determine R_{opt} (at saturation) and $R_{opt,OBO}$. The main transistor is biased in class AB, at $V_{GS,Q} = -2.78\ V$. The resulting values of R_{opt} and C_{ds} are shown in Figure 3.5 a) and b), respectively. The obtained plots show good consistency of R_{opt} and C_{ds} over the design bandwidth, especially in the high-power region, which confirms the load pull data. From the plots, the maximum power per transistor was reduced to $P_{sat} = 4\ W$ due to the reduction in V_{DD}, and $R_{opt} \approx 77\ \Omega$. At OBO, where $P_{out} = 2\ W$, $R_{opt,OBO} = 125\ \Omega$, which is not exactly double of R_{opt} due to

nonlinearities such as the knee region effect. Load pull is also performed for the class C-biased auxiliary transistor ($V_{GS,Q} = -6.0\ V$, plots not shown for brevity). This results in $R_{opt,aux} = 88\ \Omega$ since the ratio of currents $I_{a,sat}/I_{m,sat} = 70/88$ (i.e., the auxiliary PA has a lower drain current at $P_{sat} = 4\ W$). Since $I_{a,sat}/I_{m,sat} < 1$, the load modulation ratio $R_{opt,OBO}/R_{opt}$ will be lower than 2. As for C_{ds}, since it is not too nonlinear, a constant value of $C_{ds} = 0.5\ pF$ will be assumed for the design.

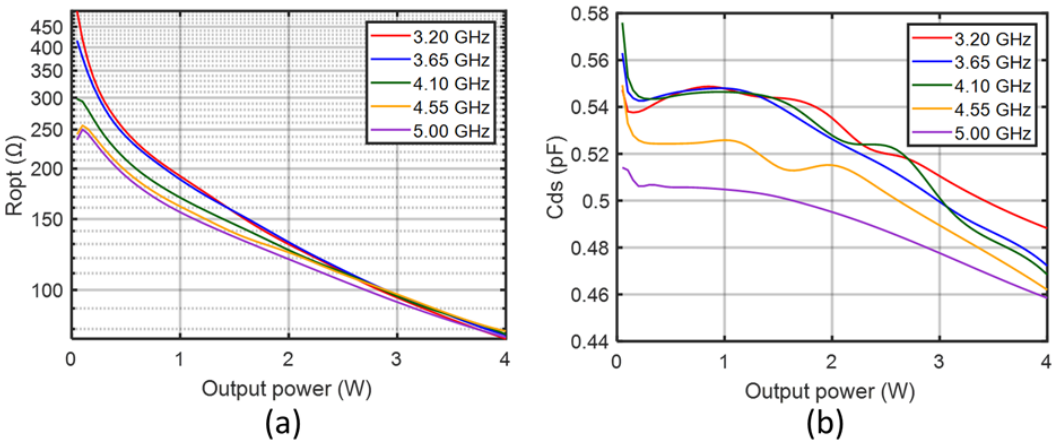


Figure 3.5 CGHV1F006 transistor class AB load pull: a) R_{opt} vs. P_{out}, b) C_{ds} vs. P_{out}

3.4 Biasing Network Design

The design of the BN is critical to the linearity performance of the DPA. Indeed, the BN determines the low-frequency range until which the transistor drain impedance Z_L maintains a low value, which is known as video bandwidth [45]. This is of key importance in broadband modulated signal performance. Indeed, as shown in Figure 2.2 b), when the transistor is subjected to multi-tone input signals f_1 and f_2, a low-frequency IMD voltage component $|f_2 - f_1|$ appears in the $I_{ds}(t)$ waveform. If the drain impedance is high (magnitude comparable to the RF load impedance) in this frequency range, the nonlinearities in $I_{ds}(t)$ will be transferred to $V_{ds}(t)$ as shown in equation (3.5) and drain voltage modulation will occur: the instantaneous drain voltage will vary, which affects the instantaneous performance of the PA and causes significant memory effects [45].

$$V_{ds}(t) = I_{ds}(t) \cdot Z_L \tag{3.5}$$

For modulated signals, the highest $|f_2 - f_1|$ IMD component is equal to the bandwidth of the input signal. Thus, the video bandwidth of the PA should exceed the bandwidth of the input signal. Note that in practical systems, the use of DPD and other signal pre-processing means that the PA input signal

bandwidth is the original input signal multiplied by the oversampling factor, which ranges from 5 to 10. Thus, for 100 MHz modulated signals, the video bandwidth must be 0.5–1 GHz.

According to [45], the conventional BN (high-impedance $\lambda/4$ TL to turn the low-impedance supply node into an open circuit at RF) is unsuitable for broadband modulated signals. This is because the high inductance of the TL resonates with the high capacitance to produce a peak in Z_L at relatively low frequencies. Thus, in this work, the output BNs are connected to the main and auxiliary transistor drains via very short TLs, which have low inductance. However, since there are no $\lambda/4$ TLs to transform the BN impedance into a short-circuit at RF, the OCN will have to be co-designed with the BN. For the transistor input, the IMD effect is much less significant, so the input BNs are connected to the transistor gate and stabilization networks via conventional $\lambda/4$ TLs.

Figure 3.6 a) and b) shows the realized drain and gate BNs, respectively. The drain BN consists of a 0.8 pF capacitor to ensure low impedance at the second harmonic, as well as 10 pF, 470 pF, 4.7 nF, 47 nF, 1.0 µF, and 47 µF (tantalum) capacitors to ensure low baseband impedance, according to [47]. The gate BN does not include the 0.8 pF capacitor since the input harmonic termination was not to be critical in this design. The resulting drain impedances for the main and auxiliary transistors are shown in Figure 3.6 c). A very broad video bandwidth is achieved: at 500 MHz, $Z_{L,main} = 6.5\ \Omega$ and $Z_{L,aux} = 9.2\ \Omega$, and at 1 GHz, $Z_{L,main} = 16.0\ \Omega$ and $Z_{L,aux} = 22.5\ \Omega$ (both well below R_{opt}).

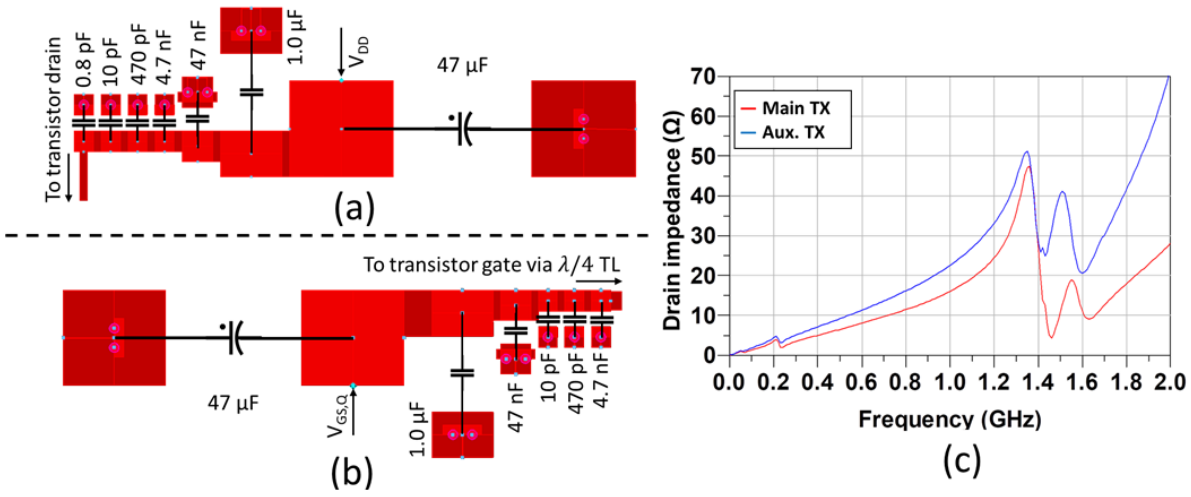


Figure 3.6 CGHV1F006 transistor: a) Drain BN, b) Gate BN, c) Baseband impedance vs. frequency

3.5 Output Combiner Network Design

3.5.1 Output Combiner Network Bandwidth Study

In this section, the theoretical bandwidth of various OCN topologies is compared to determine what FBW can be obtained from the OCN itself, which is the theoretical bandwidth of the DPA. The generalized theory for inverted DPA presented in [48] will be used, both with $\alpha = 1$ (equivalent to a single $\lambda/2$ TL in the auxiliary transistor path), and an optimized value of α for bandwidth. The conventional DPA will also be presented for comparison. Figure 3.7 a), b), and c) shows the OCN topologies: conventional, inverted with $\alpha = 1$, and inverted with optimized α, respectively.

For the conventional DPA, given the values of $R_{opt} = 77\ \Omega$, $R_{opt,aux} = 88\ \Omega$, and $R_{opt,OBO} = 125\ \Omega$, determined from load pull, the DPA equations (2.53)–(2.55) can be used to obtain the values of Z_{01} and R_L. However, with these three values, the equations would have no solution because the system is overdetermined. Thus, $R_{opt} = 70\ \Omega$ is selected instead to provide a solution, a small compromise for the main transistor's P_{out} and η_D at saturation. Solving the equations results in the following values:

$$Z_{01} = Z_{a,sat} \cdot \frac{I_{m,sat}}{I_{a,sat}} = 88\ \Omega \cdot \frac{70}{88} = 70\ \Omega \tag{3.6}$$

$$R_L = \frac{Z_{01}^2}{Z_{m,OBO}} = \frac{(70\ \Omega)^2}{125\ \Omega} = 39\ \Omega \tag{3.7}$$

For the inverted DPA, the theory from [48] is used. For $\alpha = 1$, equation (2.59) gives the value of Z_{01}. To obtain the value of R_L, equation (2.60) is modified given the unequal ratio $I_{a,sat}/I_{msat}$. The values of Z_{02} and Z_{03} are equal according to equation (2.61) and are set to $R_{opt,aux}$, which maximizes OCN bandwidth at saturation. The design parameters are thus the following:

$$Z_{01} = \frac{R_{opt}}{\alpha} = \frac{70\ \Omega}{1} = 70\ \Omega \tag{3.8}$$

$$R_L = \frac{R_{opt}}{\left(1 + I_{a,sat}/I_{m,sat}\right) \cdot \alpha} = \frac{70\ \Omega}{(1 + 70/88) \cdot 1} = 39\ \Omega \tag{3.9}$$

$$Z_{02} = R_{opt,aux} = 88\ \Omega \tag{3.10}$$

$$Z_{03} = \alpha \cdot Z_{02} = 1 \cdot 88\ \Omega = 88\ \Omega \tag{3.11}$$

For the case of $\alpha \neq 1$, the design parameters are selected using equations (3.8)–(3.11), except the value of α is varied. In [48], the authors found that values of $\alpha < 1$ reduce ITR at OBO and consequently improved bandwidth. However, this finding applied to the case where $R_{opt} < Z_0 = 50\ \Omega$. In this work, since $R_{opt} > Z_0$, values of $\alpha > 1$ are found to improve ITR at OBO.

Figure 3.8 shows the magnitudes and phases of impedances $Z_{m,OBO}$, $Z_{m,sat}$, and $Z_{a,sat}$ plotted against frequency (normalized to f_0). Based on the impedances seen by the transistors, the normalized output power can be plotted against normalized frequency to determine which OCN has the best overall bandwidth. Normalized output power for each transistor is calculated using the equation (3.12), where the first case has $|Z| < |Z_{nom}|$ (V_1 not reaching $V_{max}/2$), and the second case, $|Z| > |Z_{nom}|$ (I_1 not reaching $I_{max}/2$). Figure 3.9 a) and b) shows the normalized output power at OBO and P_{sat}, respectively, for the various OCN topologies.

$$p_{out} = \begin{cases} Re[Z]/|Z_{nom}|, & |Z| < |Z_{nom}| \\ Re[Z]/|Z_{nom}| \cdot |Z/Z_{nom}|^2, & |Z| > |Z_{nom}| \end{cases} \tag{3.12}$$

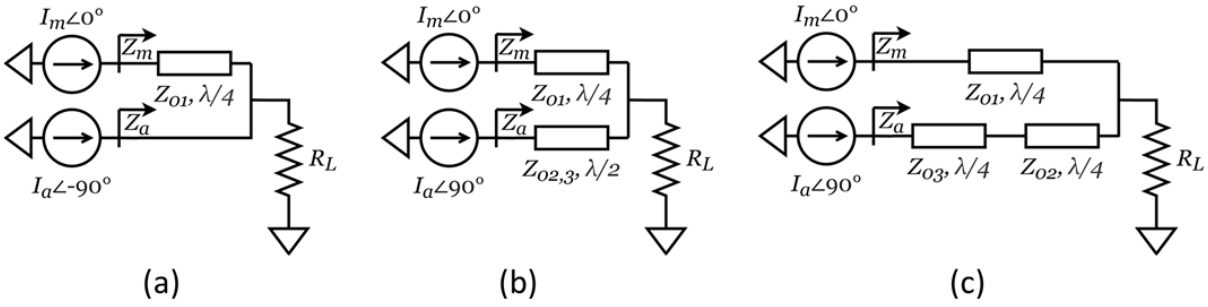


Figure 3.7 DPA OCN topologies: a) Conventional, b) Inverted ($\alpha = 1$), c) Inverted (arbitrary α)

From Figure 3.8 a) and d), it is clear that the behavior of the conventional OCN and inverted OCN with $\alpha = 1$ quickly degrades over the bandwidth. For the conventional OCN, $Z_{m,OBO}$ quickly drops, reducing P_{OBO} due to V_1 not reaching its peak value, and for the inverted OCN ($\alpha = 1$), $Z_{m,OBO}$ quickly rises, reducing P_{OBO} due to I_1 not reaching its peak value. Furthermore, an imaginary component of $Z_{m,OBO}$ also appears as frequency deviates from f_0, which is especially problematic for the inverted OCN because this impedance contributes to early compression but does not lead to usable P_{out}. Thus, in Figure 3.9 a), these topologies have very poor bandwidth when considering normalized P_{OBO}. However, the advantage of these topologies is that OCN bandwidth is unlimited at saturation as seen in Figure 3.8 b), c), e), and f): $Z_{m,sat}$, $Z_{a,sat}$, and normalized P_{sat} remain flat over the bandwidth.

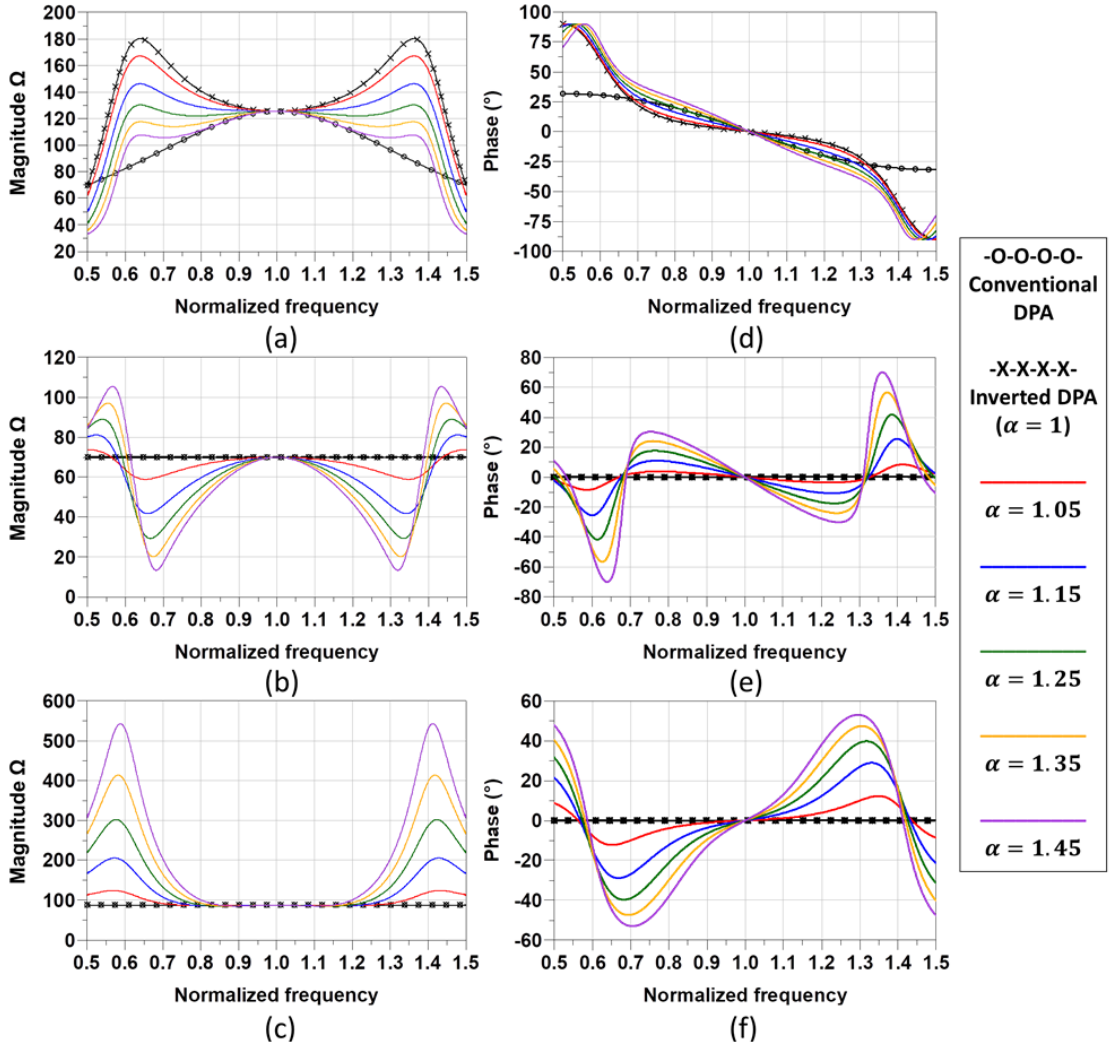


Figure 3.8 DPA OCN realized impedances vs. normalized frequency:
a) $|Z_{m,OBO}|$, b) $|Z_{m,sat}|$, c) $|Z_{a,sat}|$, d) $\angle Z_{m,OBO}$, e) $\angle Z_{m,sat}$, f) $\angle Z_{a,sat}$

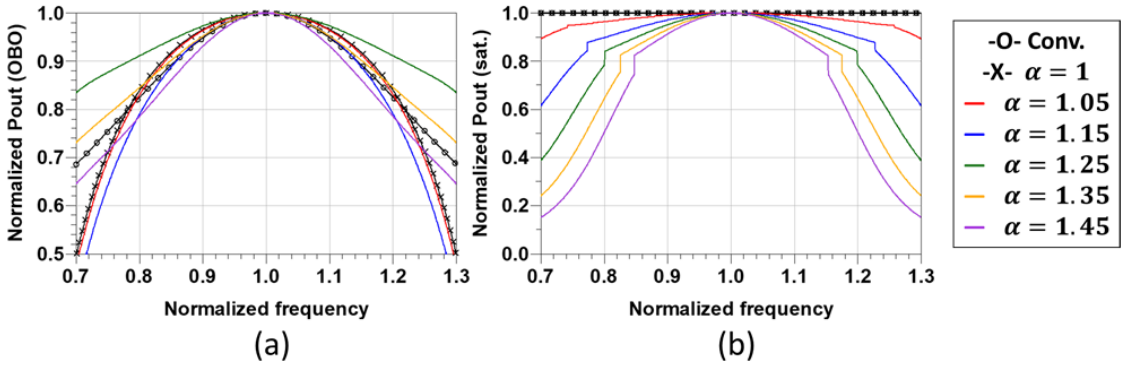


Figure 3.9 DPA OCN normalized P_{out} vs. normalized frequency at: a) OBO, b) P_{sat}

As for the inverted OCN with arbitrary α, the magnitude of $Z_{m,OBO}$ becomes flatter and flatter as α increases. However, the phase slope of $Z_{m,OBO}$ is steeper for increasing values of α (the imaginary component becomes more significant), so actually, there is not a significant improvement in normalized P_{OBO} bandwidth. The only exception is for $\alpha \approx 1.25$, which is a sweet spot for bandwidth defined based on P_{OBO}. However, for values of $\alpha \neq 1$, the normalized P_{sat} is no longer flat and decreases quicker and quicker as α increases. Overall, the best design in terms of trade-off between P_{OBO} vs. P_{sat} bandwidth is the inverted DPA with $\alpha = 1.25$, where P_{OBO} and P_{sat} remain within 91% and 80% of their peak value (respectively) over a 40% FBW.

3.5.2 Output Combiner Network Realization

After determining the theoretical optimum OCN design parameters, the transistor parasitics must be absorbed into the OCN for the actual DPA. For this purpose, ideal current sources with the parasitic network from Figure 3.4 are used to emulate the transistor, which speeds up the optimization significantly without much loss of accuracy. The resulting network topology is shown in Figure 3.10. The main path has an OMN (which absorbs the transistor drain parasitics into a Z_{01}, $\lambda/4$ TL), and the auxiliary path has an OMN (which absorbs the transistor drain parasitics into a Z_{03}, $\lambda/4$ TL) followed by a Z_{02}, $\lambda/4$ TL. These networks are joined together at the R_L node and a PMN matches the $Z_0 = 50\ \Omega$ load resistor to the desired R_L. The initial values are set to $Z_{01} = 56\ \Omega$, $Z_{02} = 88\ \Omega$, and $Z_{03} = 110\ \Omega$ according to equations (3.8)–(3.11) and $\alpha = 1.25$, then optimization is performed.

The resulting design is somewhat different from the original. In particular, the theoretical derivation assumed a frequency-independent PMN, which is not the case in reality. With a frequency-dependent PMN, $\alpha = 0.77$ provides the flattest $Z_{m,OBO}$, $Z_{m,sat}$, and $Z_{a,sat}$. Hence, the PMN is designed to transform $Z_0 = 50\ \Omega$ into $R_L = 66\ \Omega$ according to equation (3.9). The PMN schematic is shown in Figure 3.11 a), and the resulting Z_L is shown in Figure 3.11 b). The OMNs are optimized to compensate for the variation of Z_L over frequency and make $Z_{m,OBO}$, $Z_{m,sat}$, and $Z_{a,sat}$ flat over the bandwidth. Furthermore, a purely reactive second harmonic impedance is desired to avoid harmonic power dissipation and improve η_D (the third harmonic is not controlled due to its low influence on performance [38]). The schematic of the designed OCN is shown in Figure 3.12, including the PMN and the two OMNs. The theoretical and design parameters of the OCN are given in Table 3.1.

The fundamental and second harmonic values of $Z_{m,OBO}$ and $Z_{m,sat}$ are shown in Figure 3.13 a) (normalized to $R_{opt} = 70\ \Omega$), and the values of $Z_{a,sat}$, in Figure 3.13 b) (normalized to $R_{opt,aux} =$

88 Ω). At the fundamental, $Z_{m,sat}$ is very well matched to R_{opt}, with a small reactive component. The value of $Z_{m,OBO}$ is slightly too low ($Re[Z_{m,OBO}] \approx 119\ \Omega$ and $R_{opt,OBO} = 125\ \Omega$), but the resistive component is very consistent over the bandwidth ($Z_{m,OBO}$ nearly lies on a constant resistance circle). However, there is a significant reactive component when the frequency diverges from f_0 (inductive below f_0, capacitive above f_0), which is inevitable with the selected OCN topology. $Z_{a,sat}$ also shows a slightly reduced real value compared to $R_{opt,aux} = 88\ \Omega$, as well as a reactive component.

Table 3.1 Design parameters of the OCN for optimum DPA bandwidth

Parameter	$Z_{m,OBO}$	$Z_{m,sat}$	$Z_{a,sat}$	α	Z_{01}	Z_{02}	Z_{03}	R_L
Theory	125 Ω	70 Ω	88 Ω	1.25	56 Ω	88 Ω	110 Ω	25 Ω
Design				0.77	91 Ω	88 Ω	68 Ω	66 Ω

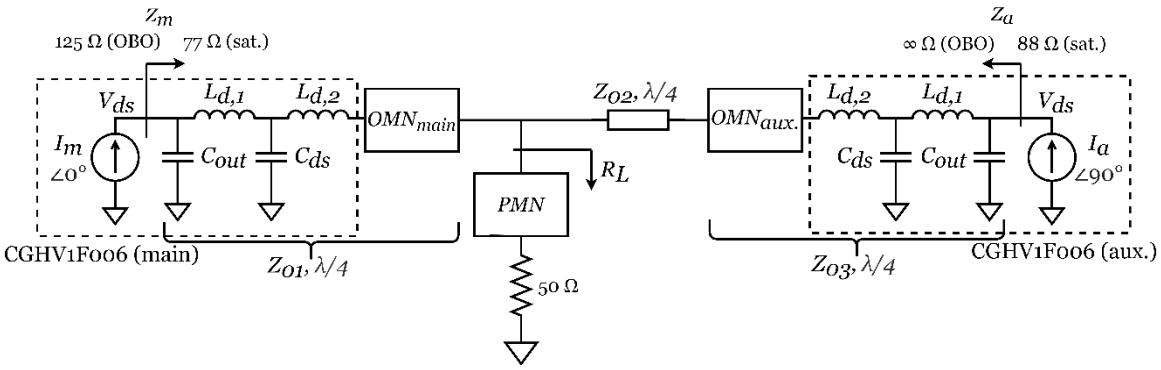


Figure 3.10 DPA OCN topology

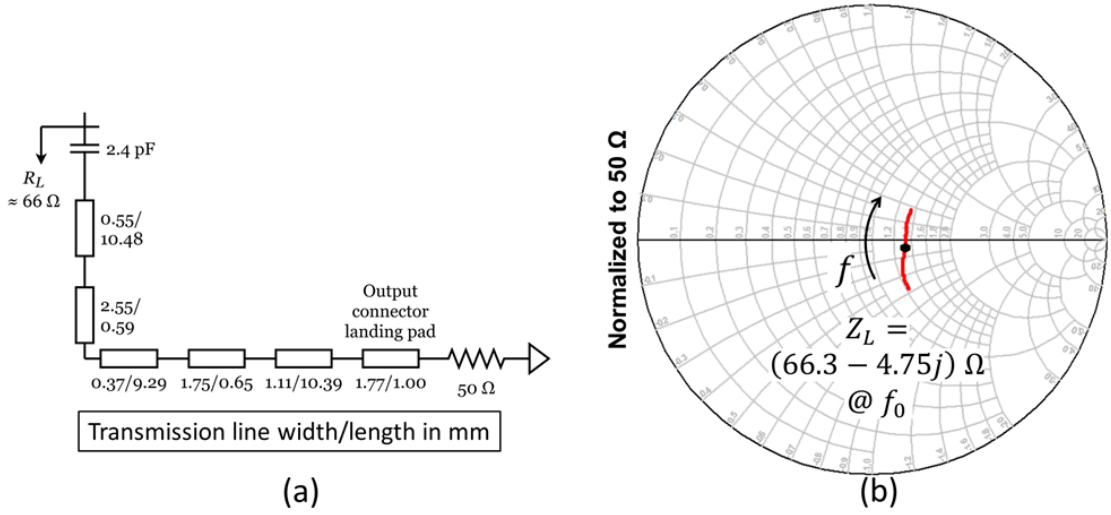


Figure 3.11 a) PMN schematic, b) Realized PMN load impedance Z_L

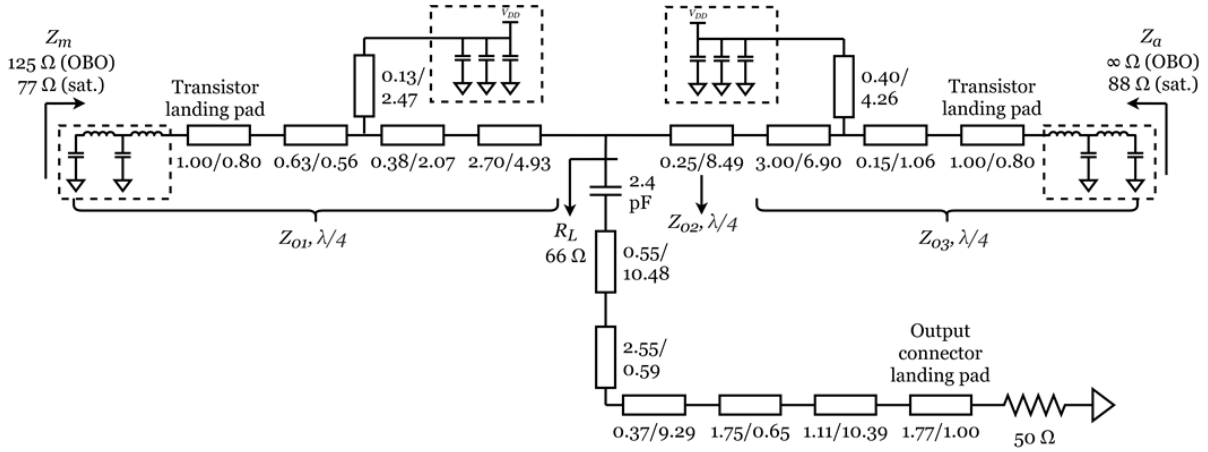


Figure 3.12 OCN schematic

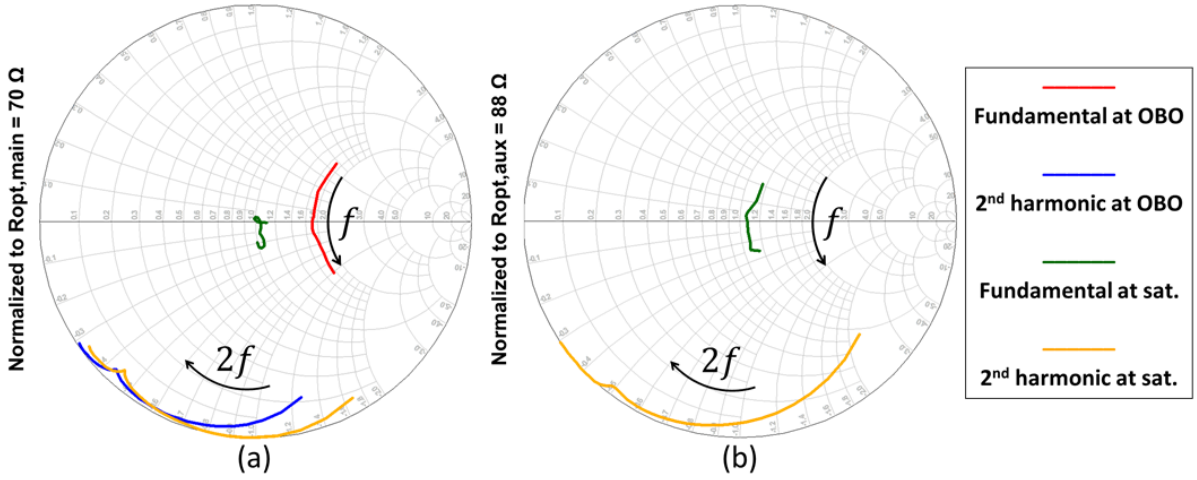


Figure 3.13 Realized PMN transistor impedances at fundamental and second harmonic: a) Z_m, b) Z_a

3.6 Input Section Design

3.6.1 Source Impedance Selection

For single-ended PAs, the main IMN function is to achieve a specific gain, but in DPAs, IMN design is far more critical, since the IMNs are responsible for realizing the current profiles of the main and auxiliary transistor, with the lowest possible AM-AM and AM-PM [45]. Thus, in this work, the source impedances for the main and auxiliary transistor, $Z_{S,m}$ and $Z_{S,a}$ had to be chosen judiciously.

The first design step is to determine the achievable performance of the DPA assuming conjugate matched sources. The schematic, shown in Figure 3.14, is built using two CGHV1F006 transistors

loaded with the OCN designed in the previous section. Since the DPA is symmetric, each transistor is supplied with two equal-power sources with impedances $Z_{S,m} = Z_{in,m}^*$ and $Z_{S,a} = Z_{in,a}^*$. The phase shift between the auxiliary and main sources ($V_{gs,a}$ and $V_{gs,m}$) is assumed to be generated by a $\lambda/4$ TL, which is given in equation (3.13). The main transistor is biased at $V_{GS,Q} = -2.8\ V$ and the auxiliary transistor uses $V_{GS,Q} = -6.0\ V$.

$$\left(\angle V_{gs,a} - \angle V_{gs,m}\right)_{conventional} = 90° \cdot \frac{(f - f_0)}{f_0} + 90° \tag{3.13}$$

Figures 3.15 a), 3.16 a), and 3.17 a) show the resulting AM-AM, AM-PM, and transistor load modulation over 3.2–5.0 GHz, in 450 MHz steps. There are issues, especially at 5.0 GHz. First, the load modulation of the auxiliary transistor brings Z_a outside the Smith chart, which could lead to instability. This is the result of a delay mismatch between the two transistors, since they are biased in different classes (main: class AB, auxiliary: class C). Thus, the phase shift between the auxiliary and main sources is adjusted as shown in equation (3.14), which brings the main and auxiliary transistor output signals in phase. Furthermore, peak AM-PM is problematic, ranging from -27° at 3.2 GHz to 34° at 5.0 GHz. The AM-PM of the OCN, in the DPA load modulation region (P_{out} between 33 and 39 dBm), can be reduced by optimizing the phase of $Z_{s,m}$. Finally, AM-AM is also problematic: gain flatness could be improved (10.6–12.2 dB over the bandwidth) and OP_{1dB} is very poor (as low as 30 dBm at 5.0 GHz). AM-AM is adjusted by optimizing the magnitude of $Z_{s,m}$: if the main transistor is intentionally mismatched, the small-signal gain can be decreased, which can be used to improve gain flatness, as well as OP_{1dB} due to slower gain roll-off. $Z_{s,a}$ remains conjugate matched because the gain of the auxiliary transistor is already low and should not be degraded further.

$$\left(\angle V_{gs,a} - \angle V_{gs,m}\right)_{optimized} = 90° \cdot \frac{(f - f_0)}{f_0} + 110° \tag{3.14}$$

Table 3.2 shows the initial and optimized values of $Z_{s,m}$ and $Z_{s,a}$ for 3.2, 3.65, 4.1, 4.55, and 5.0 GHz. For intermediate frequency points, the values of $Z_{s,m}$ and $Z_{s,a}$ can be interpolated from the values in the table. Figures 3.15 b), 3.16 b), and 3.17 b) show the resulting AM-AM, AM-PM, and transistor load modulation over the 3.2–5.0 GHz range. The load modulation behavior is much better: Z_m and Z_a no longer have any imaginary component at f_0, and Z_a no longer exits the Smith chart like in Figure 3.17 a). The AM-PM is also more manageable, with a range of -25° to 21°. Gain flatness is improved (10.0

to 11.0 dB over the bandwidth, only a 1 dB variation), and OP_{1dB} is improved (above 36 dBm at all frequencies). The trade-off is an acceptable decrease in small-signal gain.

Table 3.2 Initial and optimized values of $Z_{s,m}$ and $Z_{s,a}$ for optimum DPA load modulation

Parameter	$Z_{s,m}$ (Ω)					$Z_{s,a}$(Ω)				
	3.2 GHz	3.65 GHz	4.1 GHz	4.55 GHz	5.0 GHz	3.2 GHz	3.65 GHz	4.1 GHz	4.55 GHz	5.0 GHz
Initial	14.8-j·18.2	10.3-j·10.3	8.0-j·2.3	6.8+j·7.0	8.2+j·19.3	13.2-j·23.6	9.9-j·14.4	7.3-j·5.6	6.5+j·4.2	7.7+j·15.1
Optimized	10.3-j·15.7	10.3-j·3.0	10.9+j·9.1	20.3+j·20.6	15.6+j·38.2					

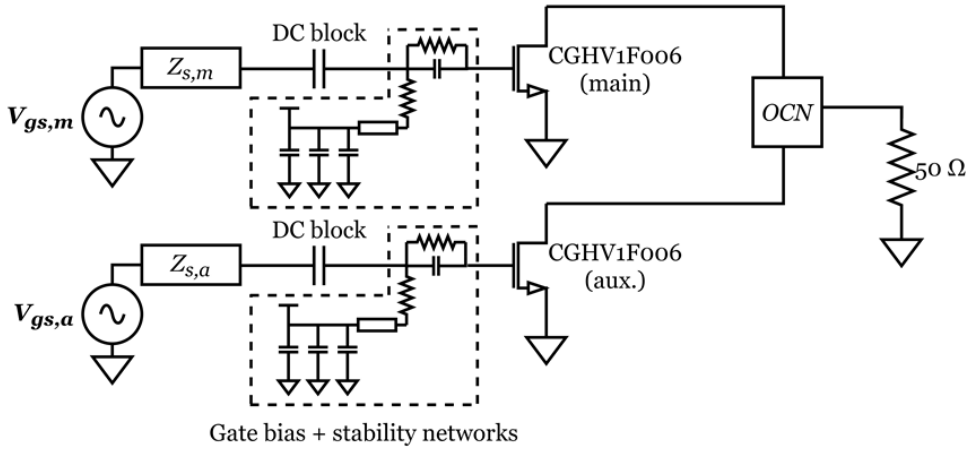


Figure 3.14 Schematic used to determine $Z_{s,m}$ and $Z_{s,a}$

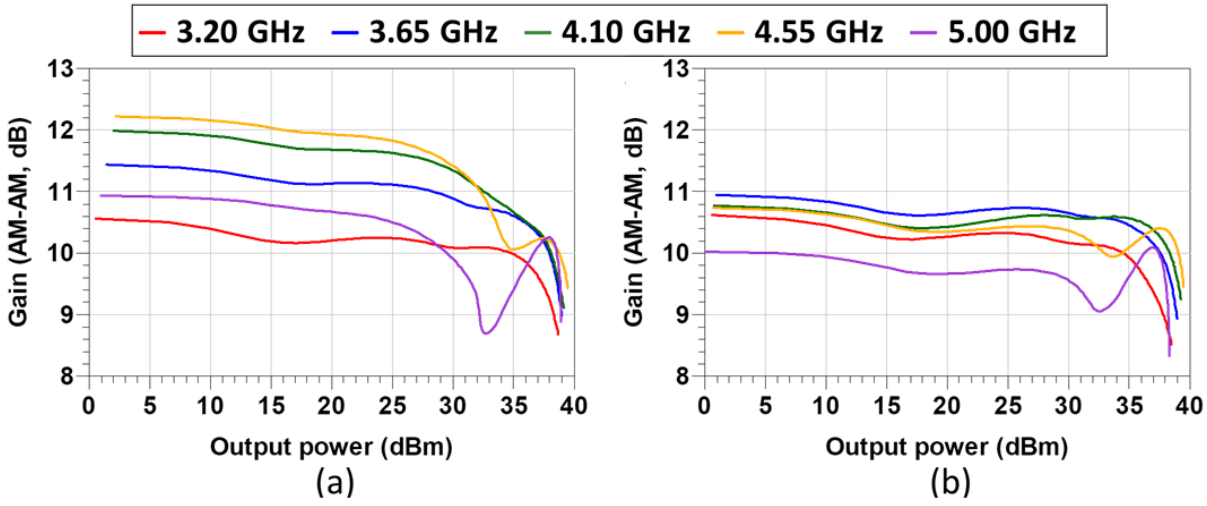


Figure 3.15 DPA AM-AM: a) With initial $Z_{s,m}/Z_{s,a}$, b) With optimized $Z_{s,m}/Z_{s,a}$

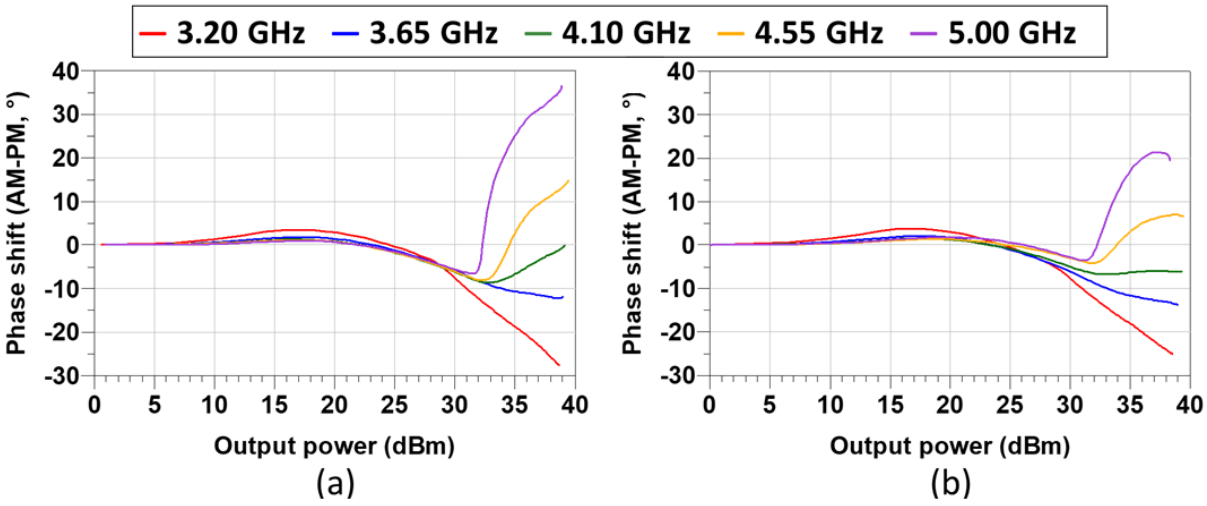


Figure 3.16 DPA AM-PM: a) With initial $Z_{s,m}/Z_{s,a}$, b) With optimized $Z_{s,m}/Z_{s,a}$

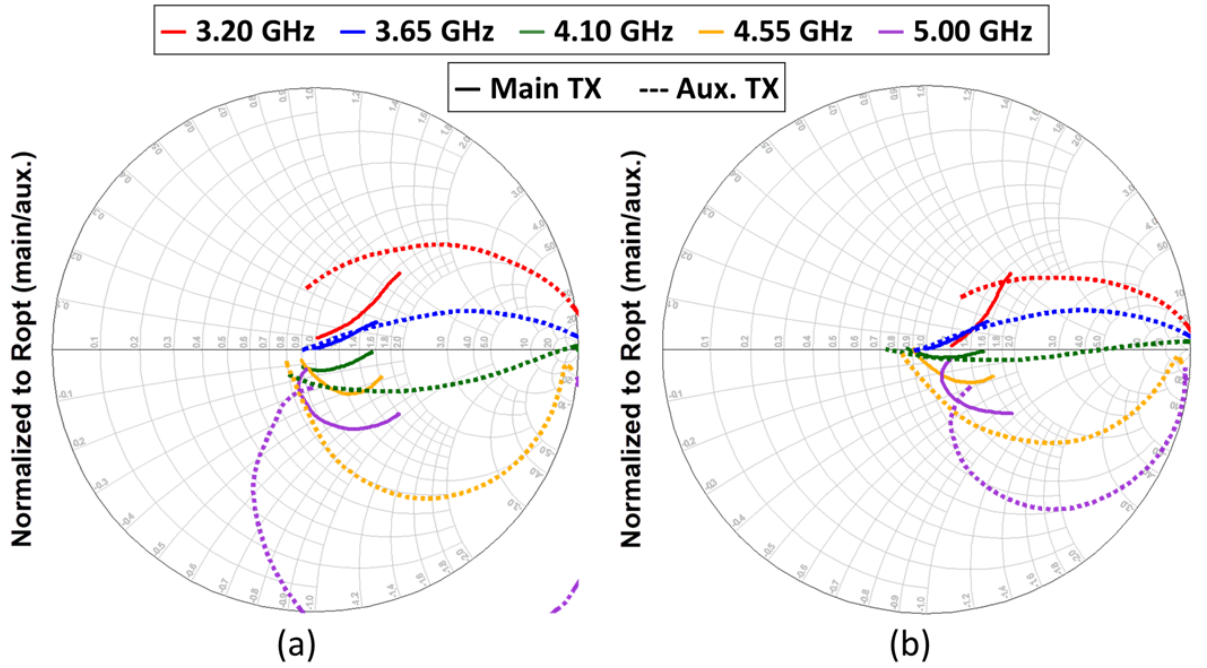


Figure 3.17 DPA load modulation: a) With initial $Z_{s,m}/Z_{s,a}$, b) With optimized $Z_{s,m}/Z_{s,a}$

3.6.2 Input Matching Network Realization

The optimized impedances given in the previous section are very difficult to realize using physical components over the bandwidth because the required $Z_{s,m}$ and $Z_{s,a}$ follow a counterclockwise trajectory on the Smith chart as frequency increases, whereas the impedance of physical networks

always rotates clockwise with frequency [37]. Thus, high-order MNs are introduced to obtain the best matching across the bandwidth. Since there are five frequency points used for the matching (3.2, 3.65, 4.1, 4.55, and 5.0 GHz), a sixth-order filter consisting of five TL sections and a stub is used, to allow matching at all five points with an extra degree of freedom.

Furthermore, the desired phase relationship between $V_{gs,a}$ and $V_{gs,m}$ shown in equation (3.14) must be synthesized by the IMNs. However, the phase slope (90°) is not equal to the phase at f_0 (110°), so it is not achievable with TLs, which have a phase proportional to frequency. Thus, capacitors, which have non-constant τ over frequency, are introduced in the IMN to shape the phase slope. These capacitors also serve as DC blocks.

The component values for the IMNs are obtained using optimization. Figure 3.18 a) and b) shows the schematics of the main and auxiliary transistor IMNs, respectively. The achieved performance of the IMNs is shown in Figure 3.19: a) shows the mismatch between desired and achieved Z_s (return loss), b) shows the corresponding insertion loss, and c) shows the comparison between the desired and achieved phase shifts between the auxiliary and the main transistor paths. Over the bandwidth, the matching achieved corresponds to a 6.4–8.3 dB return loss, which corresponds to a 0.7–1.1 dB insertion loss: the expected small-signal gain is expected to be approximately 0.7–1.1 dB lower than in Figure 3.15 b). The insertion loss ratio is very close to 1 across the bandwidth, showing a good power split between the main and auxiliary transistors. The achieved phase shift is also very close to the desired phase relationship, with a small variation from -3.4° to 3.1° over the entire bandwidth.

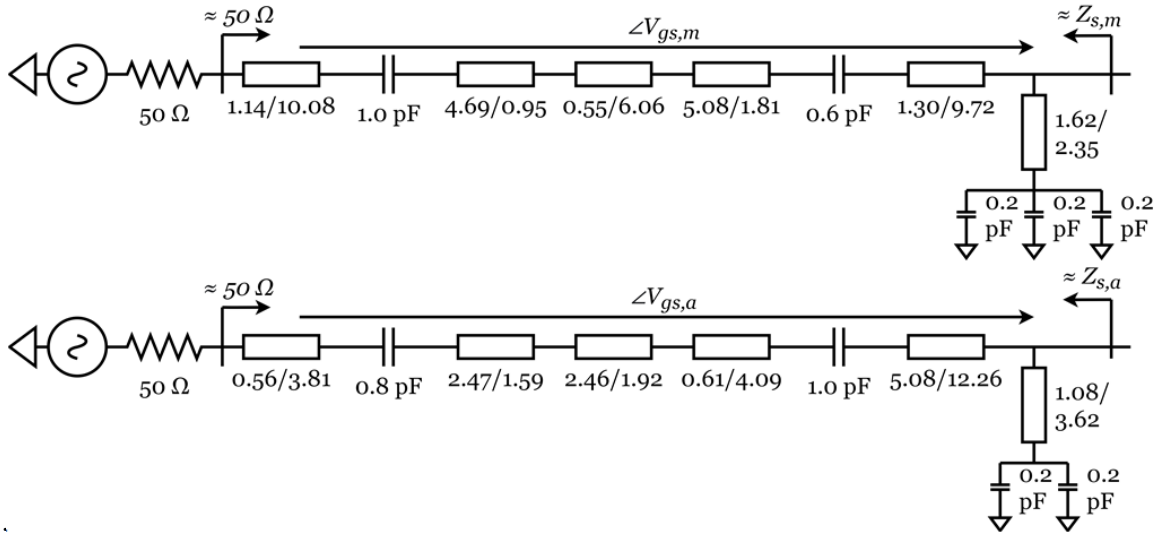


Figure 3.18 IMN schematic for: a) Main transistor, b) Auxiliary transistor

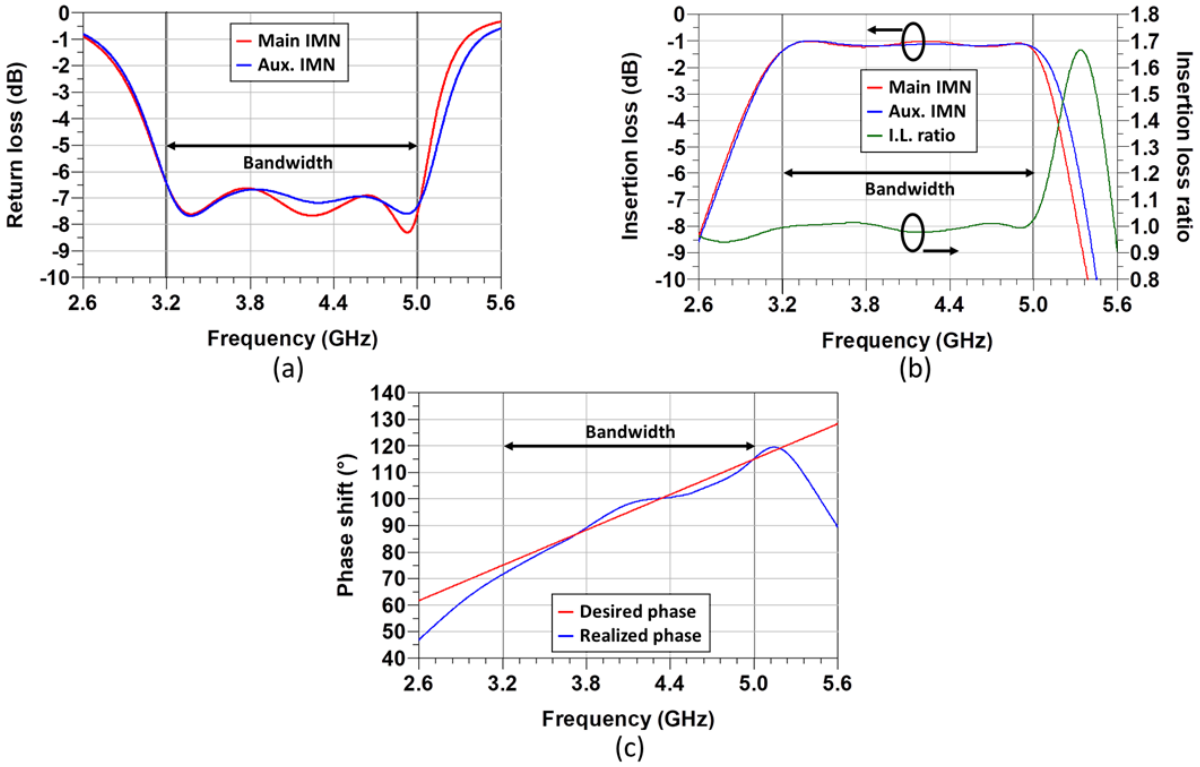


Figure 3.19 IMN performance: a) Return loss, b) Insertion loss, c) Phase relationship

3.6.3 Input Power Divider Design

A Wilkinson power divider providing an in-phase power split is used in this work since the required phase relationship is synthesized entirely by the main and auxiliary IMNs. Since the transistors are equal-sized and IMNs are designed for equal power split, a symmetric Wilkinson can be used, which simplifies the layout. To reduce insertion loss, a single-section Wilkinson is used, as shown in Figure 3.20 a). However, over a 41% FBW (3.3–5.0 GHz), the single-section Wilkinson suffers frequency-dependent behavior that worsens return loss and increases insertion loss at the low and high end of the bands, as shown in Figure 3.20 b). Thus, two additional TL sections are added after the Wilkinson to ease matching, and the rest of the IMN is tuned to restore the bandwidth.

Figure 3.21 shows the schematics of the entire DPA input section. The achieved performance of the input section is shown in Figure 3.22: a) shows the return loss at all ports, b) shows the corresponding insertion loss (including the 3 dB loss due to the equal power split), and c) shows the comparison between the desired and achieved phase shifts between the auxiliary and the main transistor paths. Over the bandwidth, the matching achieved for $Z_{s,m}$ and $Z_{s,a}$ corresponds to a 5.8–8.2 dB return loss, or a

0.7–1.4 dB (excess) insertion loss, slightly worse than with IMNs alone. The power division ratio is also slightly degraded by adding the Wilkinson but remains between 0.92 and 1.17 (fairly close to 1) across the bandwidth. The input return loss is good, 22.7 dB at f_0 and better than 11.9 dB across the bandwidth, so the PA driver will not suffer significant reflection. As for the phase shift, it remains very good, with a -3.6° to 3.9° deviation from the expected over the entire bandwidth. In short, replacing the ideal power split with a Wilkinson had almost no impact on the IMN performance.

The insertion loss ratio is very close to 1 across the bandwidth, showing a good power split between the main and auxiliary transistors. The achieved phase shift is very close to the desired phase relationship, with a small variation from -3.4° to 3.1° over the entire bandwidth. Thus, the load modulation shown in Figure 3.17 b) should be mostly maintained.

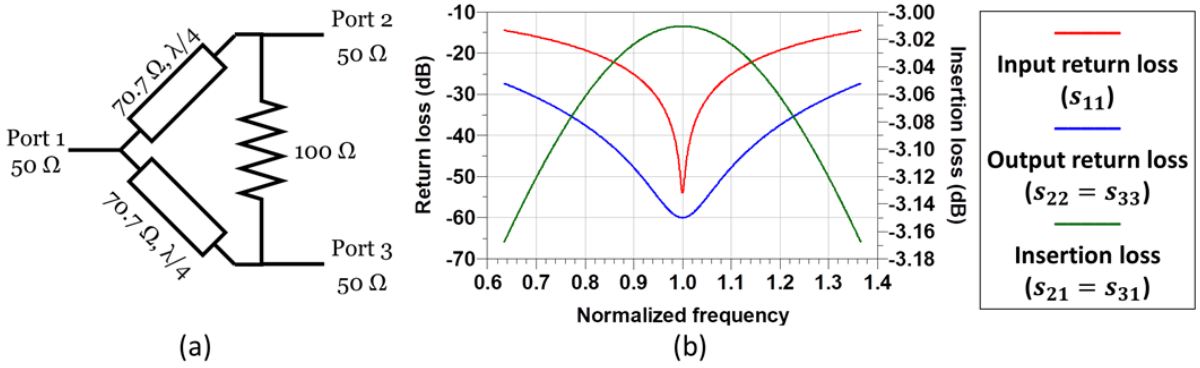


Figure 3.20 Symmetric Wilkinson power divider: a) Schematic, b) S-parameters

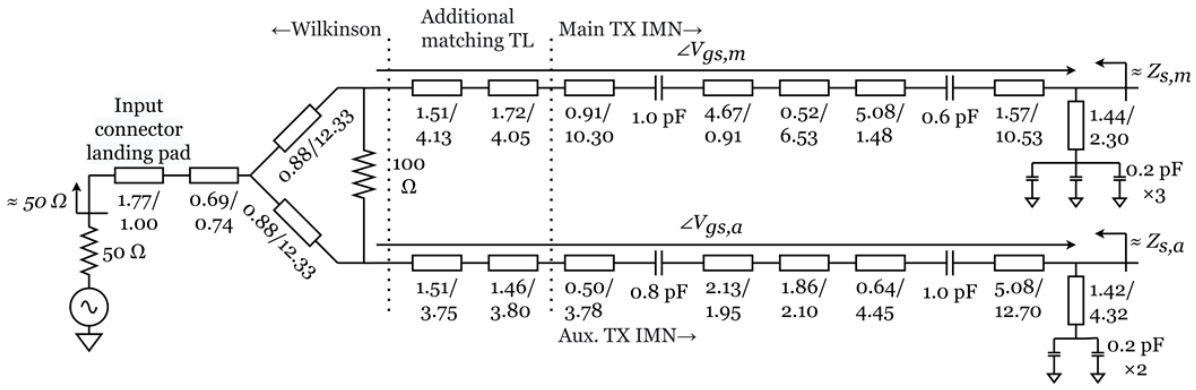


Figure 3.21 Input section schematic

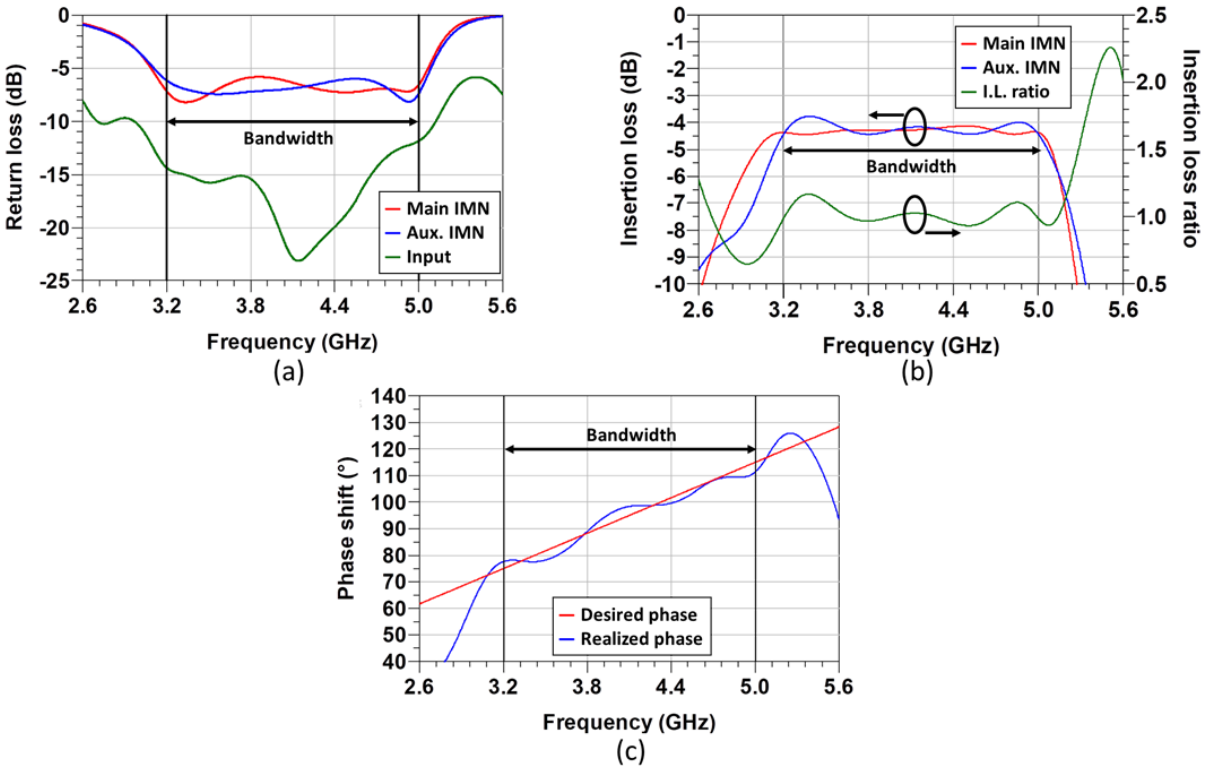


Figure 3.22 Input section performance: a) Return loss, b) Insertion loss, c) Phase relationship

3.7 Full DPA Simulation Results

The bias and stability networks, OCN, IMNs, power divider, and transistors are combined into one layout, which requires the addition of microstrip curved TLs and small adjustment of some TL lengths. The full layout is exported from Keysight ADS to Ansys HFSS for full 3D EM validation, which is more accurate than the 2.5D simulation used in ADS Momentum. Some performance degradation is observed when doing full 3D EM simulation, so the layout and values of passive components are readjusted. The capacitor-loaded stubs in the IMNs are replaced by open-circuited stubs to avoid performance variation due to component tolerances. Figure 3.23 shows the final DPA schematic. After tuning to restore the performance, the layout is exported to Altium Designer for the generation of manufacturing files. Figure 3.24 a) shows the final Altium design as sent for fabrication, and Figure 3.24 b) shows a version with integrated Anaren X4C40K1-20S 20 dB directional coupler to sample the output for DPD. The coupler version was not fabricated at this stage because the coupler S-parameters were not provided over the entire DPA frequency band (including fundamental and harmonics). The EM-simulated performance of the DPA is shown in the following figures, before and after tuning.

Figure 3.25 shows the simulated S-parameters of the DPA from 2.0 to 6.0 GHz: a) shows the small-signal gain (s_{21}) and b) shows the input and output return loss (s_{11} and s_{22}), as well as reverse isolation (s_{12}). The bandwidth (defined by the flat region of s_{21}) shifted down to 3.00–4.50 GHz after full EM simulation but was restored to 3.20–5.20 GHz after circuit tuning (however, the achieved gain was lower overall). The return loss was also improved over the band of interest after tuning, with both s_{11} and s_{22} below -10 dB over the bandwidth (although the sharp resonance in s_{11} expected from Figure 3.22 a) could not be recovered). However, small-signal performance is insufficient to fully characterize high-power PAs, especially not DPAs where the auxiliary transistor is only active at high power.

Figure 3.27 shows the DPA current profiles at the intrinsic drain current source (with the parasitic network shown in Figure 3.4 de-embedded). Plots a) and b) cover 3.3–4.1 and 4.2–5.0 GHz before tuning, and c) and d) cover the same frequency ranges after tuning. The post-tuning version shows a significant improvement in the consistency of the turn-on point of the auxiliary transistor with respect to the normalized input voltage (the range shrinks from $0.25v_{in} - 0.5v_{in}$ to $0.3v_{in} - 0.45v_{in}$ and is a lot more regular over frequency). The auxiliary transistor turn-on point is not exactly $0.5v_{in}$ as per theory, but earlier turn-on is necessary to achieve a sufficiently high $I_{a,sat}$. Overall, after tuning, I_m shows early compression for the lower frequencies (3.3–3.4 GHz), and I_a does not quite reach $I_{a,sat}$ for the lower and higher ends of the frequency band since the auxiliary transistor matching at these frequencies has degraded. Despite these shortcomings, the post-tuning profiles are a lot closer to the ideal case (represented with thick black lines) than pre-tuning, showing a significant improvement.

Figure 3.28 shows the corresponding voltage profiles at the intrinsic drain current source: plots a) and b) cover 3.3–4.1 and 4.2–5.0 GHz before tuning, and c) and d) cover the same frequency ranges after tuning. After tuning, V_m is a lot closer to the desired profile, but 3.3 and 3.4 GHz show V_m higher than expected, indicating intrusion into the knee region and confirming the early compression in these bands. As for V_a, it does not reach the desired peak value in any of the frequency bands, which indicates that $Z_{a,sat}$ achieved was too low compared to $R_{opt,aux}$. Despite these issues, the post-tuning profiles are a lot closer to the ideal case than pre-tuning, showing that performance was improved.

Figure 3.26 shows the load modulation: plots a) and b) represent the main and auxiliary transistor load impedances before tuning, while c) and d) represent the same parameters after tuning. The Smith charts are normalized to R_{opt} of the respective transistor. The full EM simulation shows that the main transistor load modulation was relatively well-maintained at low frequencies (solid lines) but degraded significantly into the capacitive region at high frequencies (dashed lines), which explains the poor

behavior of the voltage and current profiles at these frequencies. Post-tuning, the values of $Z_{m,OBO}$ and $Z_{m,sat}$ lie much closer to the expectation from the initially designed OCN (Figure 3.13 a)). In particular, $Z_{m,OBO}$ lies almost entirely on the constant $1.8 \cdot R_{opt}$ circle, except in the very high bands (above 4.5 GHz). However, at 3.3 and 3.4 GHz, the load modulation is poor ($Z_{m,sat} \gg R_{opt}$), which confirms the early compression (also observed in Figures 3.27 and 3.28). As for the auxiliary transistor, tuning brought the values of $Z_{a,sat}$ much closer to the capacitive region and made the load modulation significantly more consistent for the high frequencies. However, these improvements came at a cost: for the low frequencies, $Z_{a,sat} \ll R_{opt,aux}$, which confirms the observation that V_a did not reach the desired peak value in Figure 3.28. Despite these shortcomings, the post-tuning version shows significantly less variation of $Z_{m,OBO}$, $Z_{m,sat}$, and $Z_{a,sat}$ over frequency compared to pre-tuning.

Figures 3.29, 3.30, 3.31, and 3.32 show AM-AM (gain), AM-PM (phase shift), η_D, and PAE plotted against P_{out}. Plots a) and b) show the low-frequency and high- frequency bands pre-tuning, and c) and d) show the low-frequency and high- frequency bands post-tuning. The pre-tuning design shows a very variable range of P_{sat} (35.5–39.5 dBm), whereas post-tuning made the range much tighter (38.5–39.5 dBm). For AM-AM, the tuning made a huge difference at high frequencies: the pre-tuning case shows a very rapid drop in gain for frequencies above 4.5 GHz (overall 6–11 dB small-signal gain variation over the bandwidth), whereas the post-tuning case shows significantly better gain flatness (9.0–10.5 dB) and saturated gain (above 7 dB for all bands). For AM-PM, the tuning had a very beneficial impact: the range of AM-PM at P_{sat} shrank from [-22°,15°] to [-20°, 8°]. The discontinuity in the AM-PM curve caused by the auxiliary transistor turn-on was minimized, which should aid with DPD linearizability, but could not be fully eliminated over the entire band. Finally, significant improvements are visible in the η_D plots. Before tuning, the turn-on point of the auxiliary transistor (inflection point in the DE curve) is extremely inconsistent, ranging from 28 to 34 dBm. Many low-frequency bands such as 3.3–3.6 GHz showed almost no OBO η_D enhancement over a class B PA. After tuning, the turn-on point is much more consistent at approximately 32 dBm, or 6-7 dB OBO. OBO η_D enhancement is much clearer, with $\eta_{D,OBO} > 40\%$ and $\eta_{D,sat} > 58\%$ for all bands. Low-frequency bands (3.3–3.6 GHz) still show slightly degraded OBO η_D enhancement, likely because the transistor's η_D is poor when $Z_{m,OBO}$ is in the inductive region of the Smith chart. For PAE, the conclusions are largely the same as for η_D, with a very significant improvement post-tune due to improvements in η_D and gain. In addition, note the significant PAE drop-off at P_{sat} compared to OBO ($PAE_{OBO} > 35\%$, $PAE_{sat} > 45\%$), which emphasizes the importance of maintaining high gain in the high-power region.

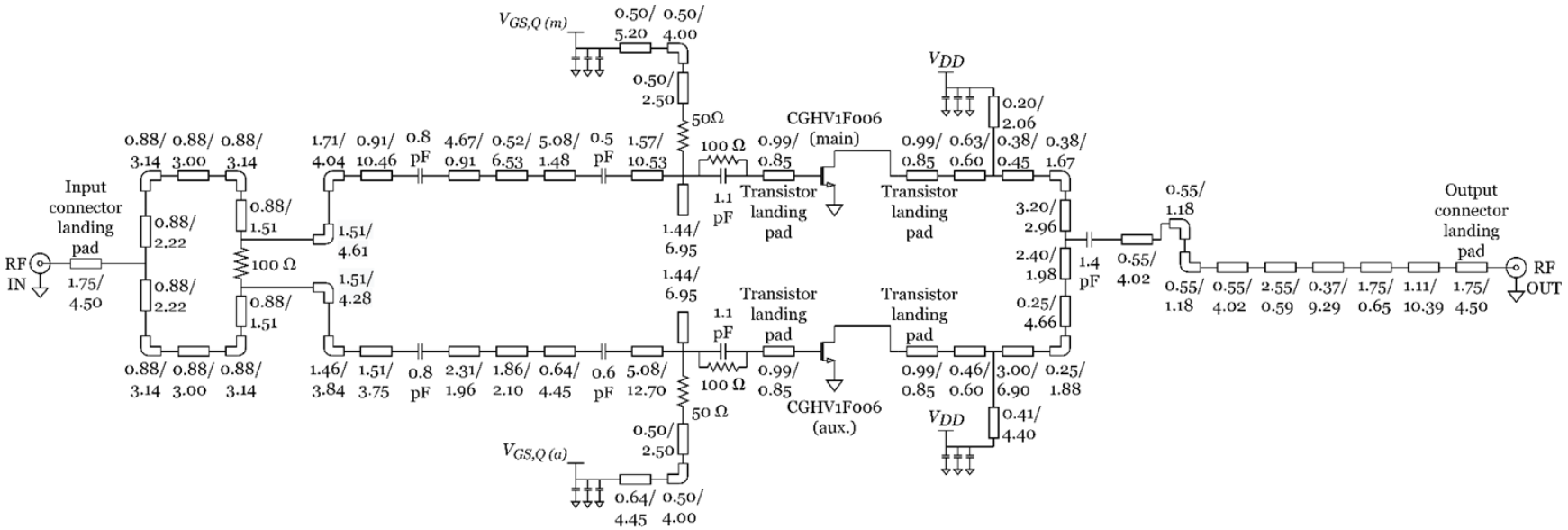
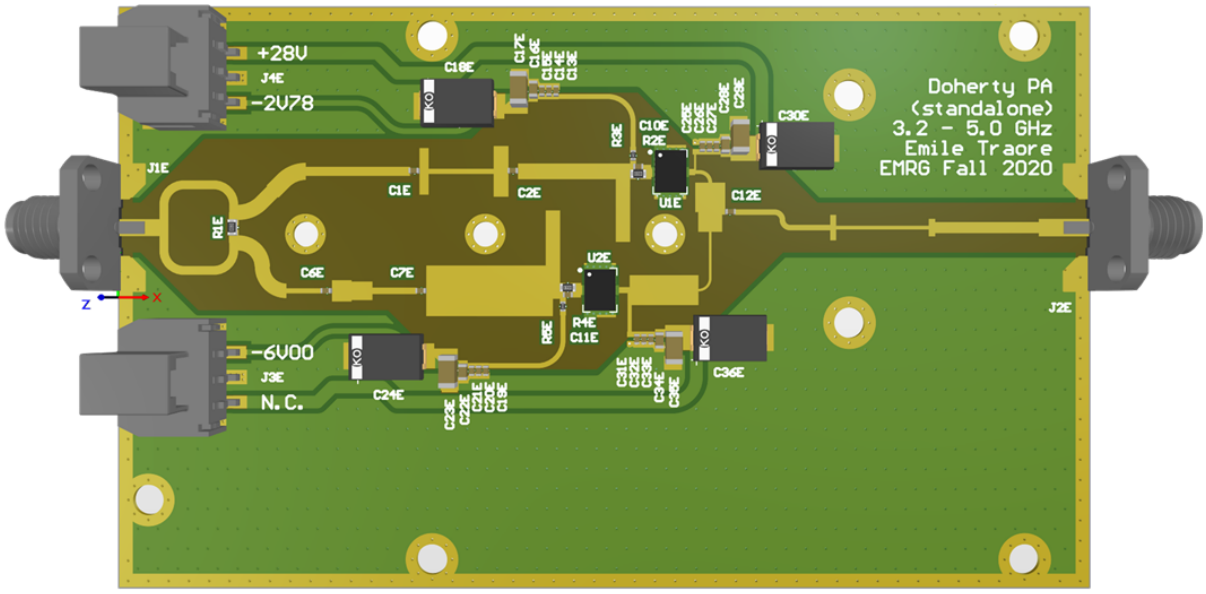
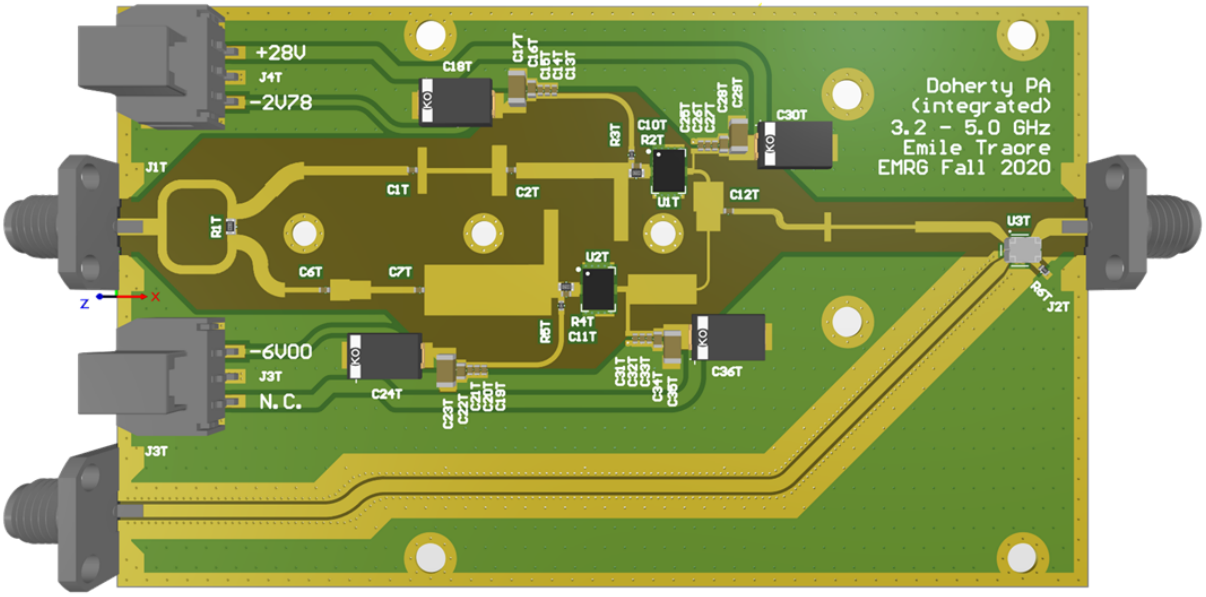


Figure 3.23 Overall DPA schematic



(a)



(b)

Figure 3.24 DPA layout on PCB: a) Without coupler (fabricated), b) With coupler (not fabricated)

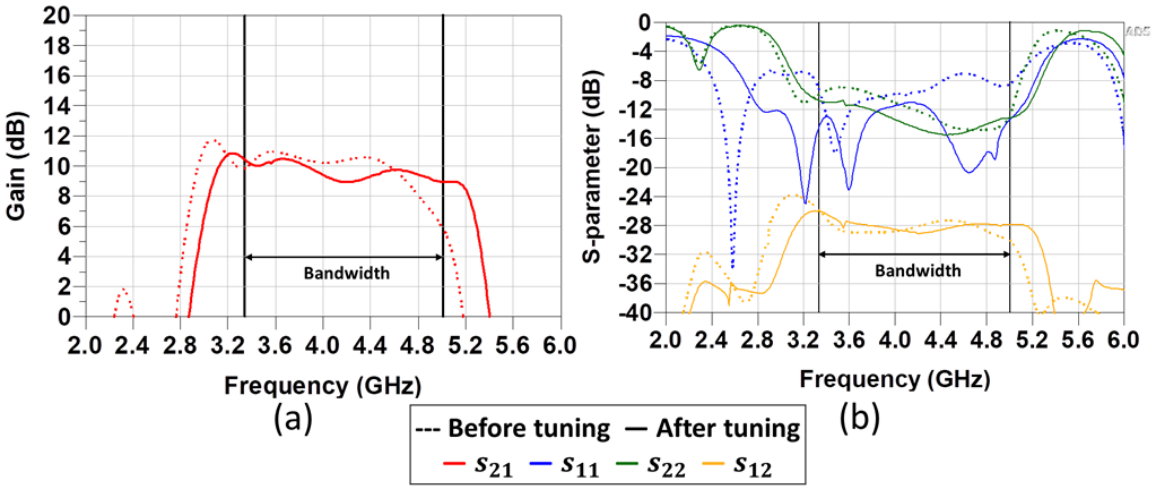


Figure 3.25 Simulated DPA S-parameters: a) s_{21}, b) s_{11}, s_{22}, s_{12}

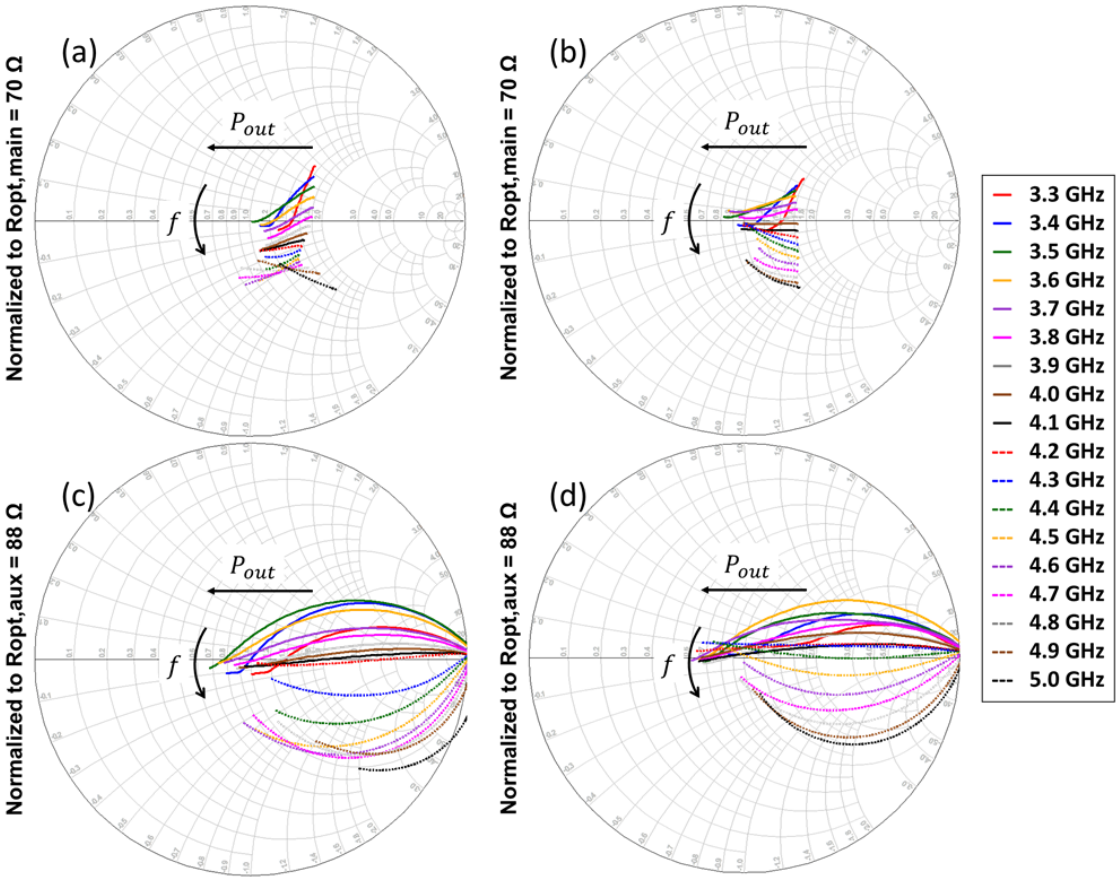


Figure 3.26 Simulated DPA load modulation trajectories vs. frequency and output power:
a) Main transistor pre-tune, b) Main transistor post-tune,
c) Auxiliary transistor pre-tune, d) Auxiliary transistor post-tune

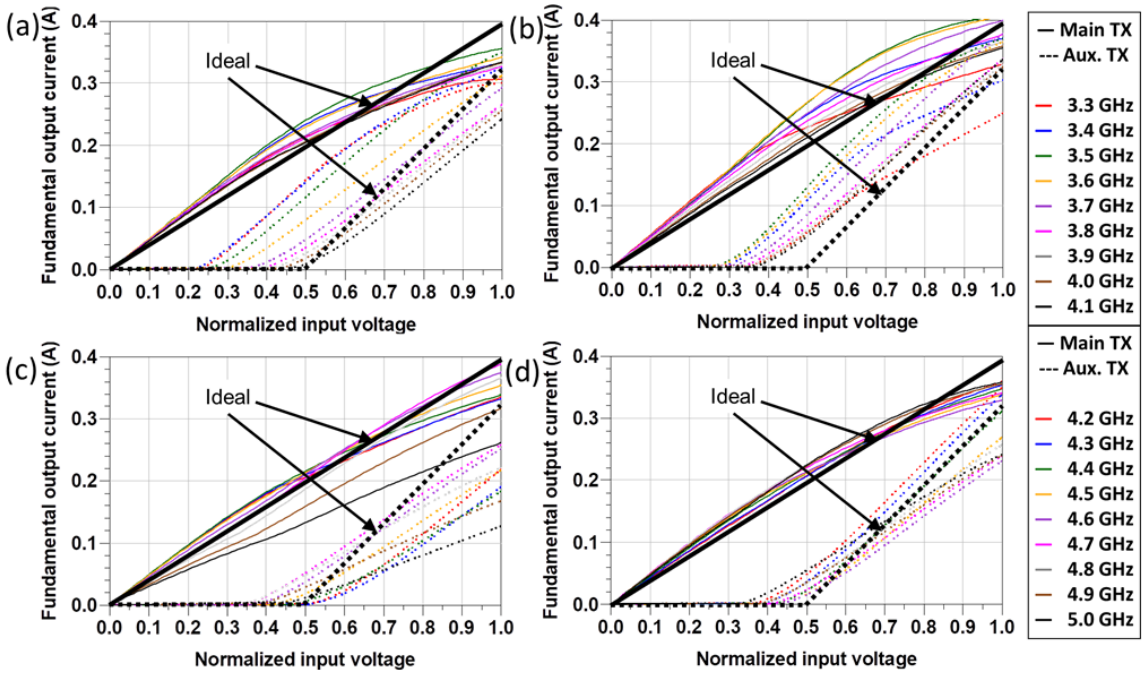


Figure 3.27 Simulated DPA current profiles:

a) Low-f pre-tune, b) Low-f post-tune, c) High-f pre-tune, d) High-f post-tune

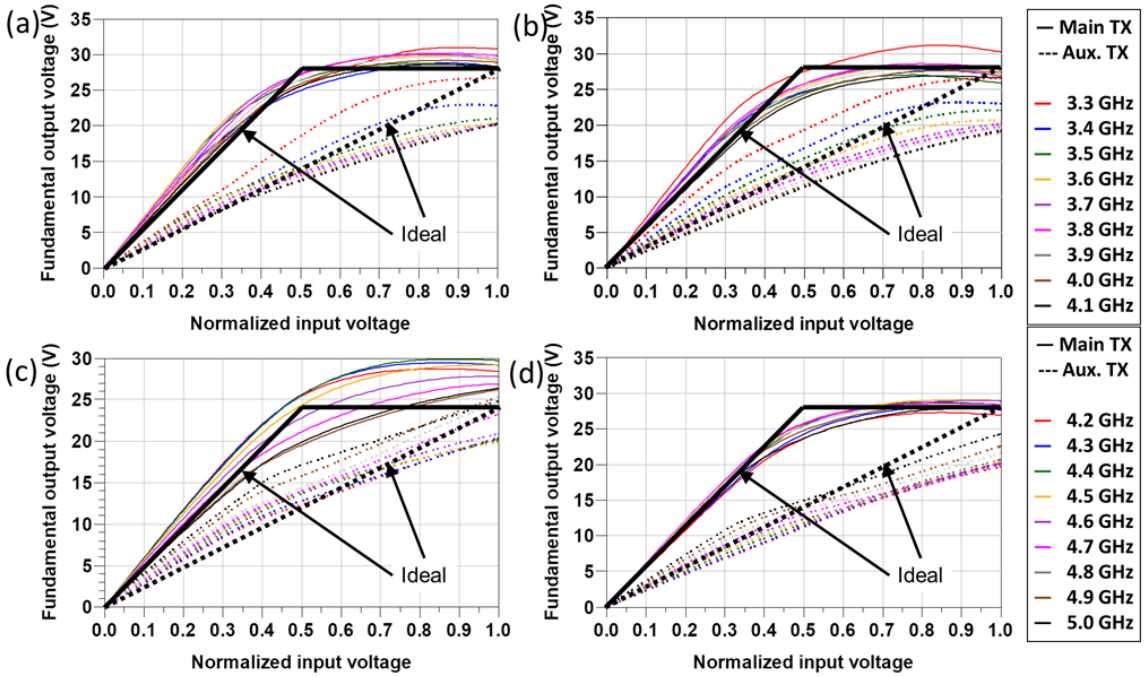


Figure 3.28 Simulated DPA voltage profiles:

a) Low-f pre-tune, b) Low-f post-tune, c) High-f pre-tune, d) High-f post-tune

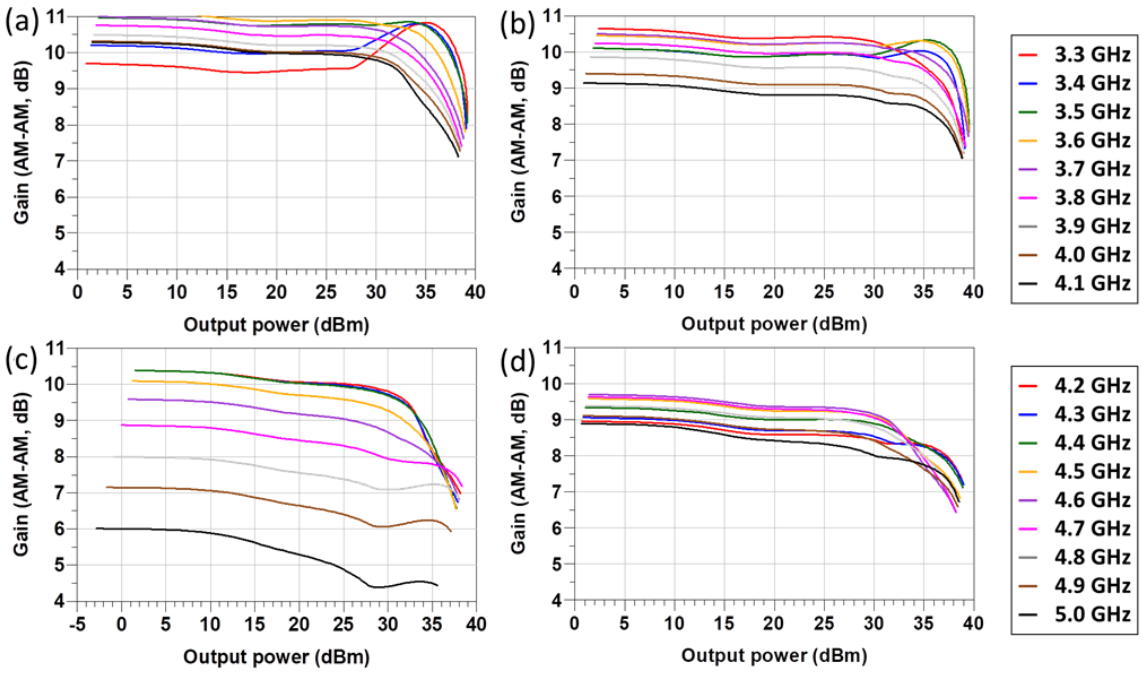


Figure 3.29 Simulated DPA AM-AM:

a) Low-f pre-tune, b) Low-f post-tune, c) High-f pre-tune, d) High-f post-tune

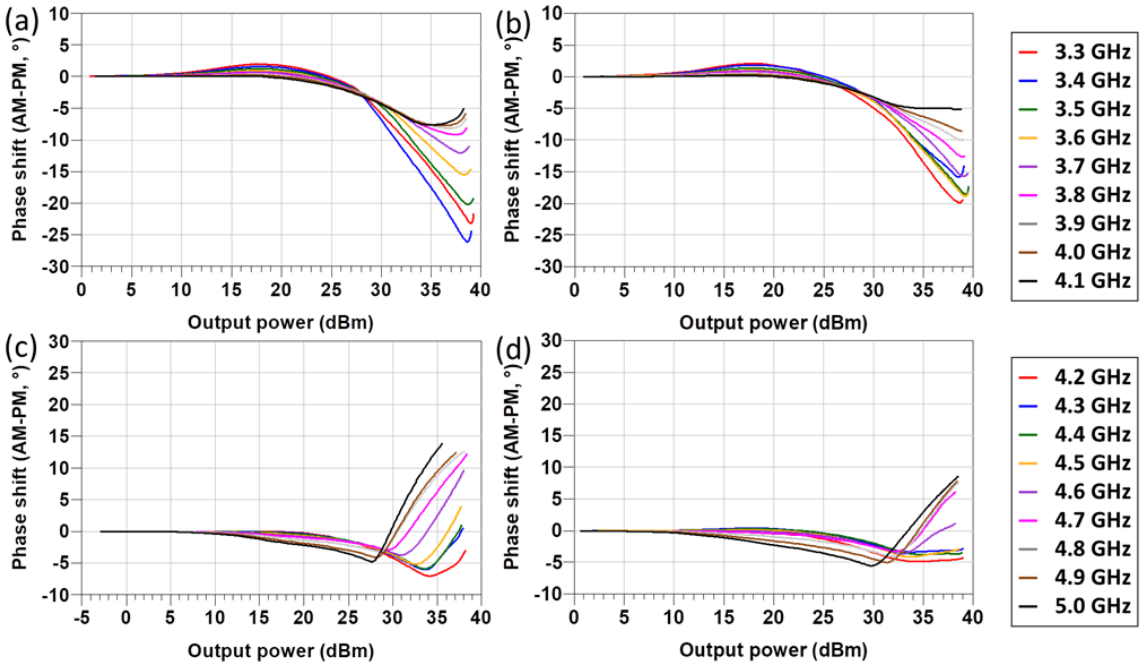


Figure 3.30 Simulated DPA AM-PM:

a) Low-f pre-tune, b) Low-f post-tune, c) High-f pre-tune, d) High-f post-tune

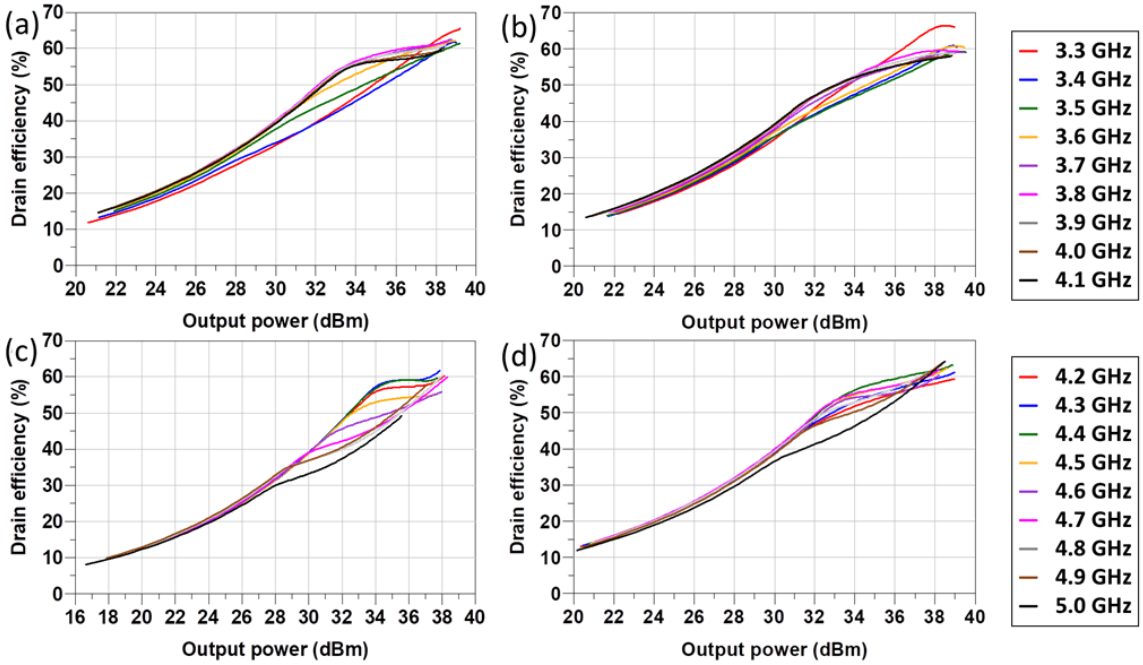


Figure 3.31 Simulated DPA η_D:

a) Low-f pre-tune, b) Low-f post-tune, c) High-f pre-tune, d) High-f post-tune

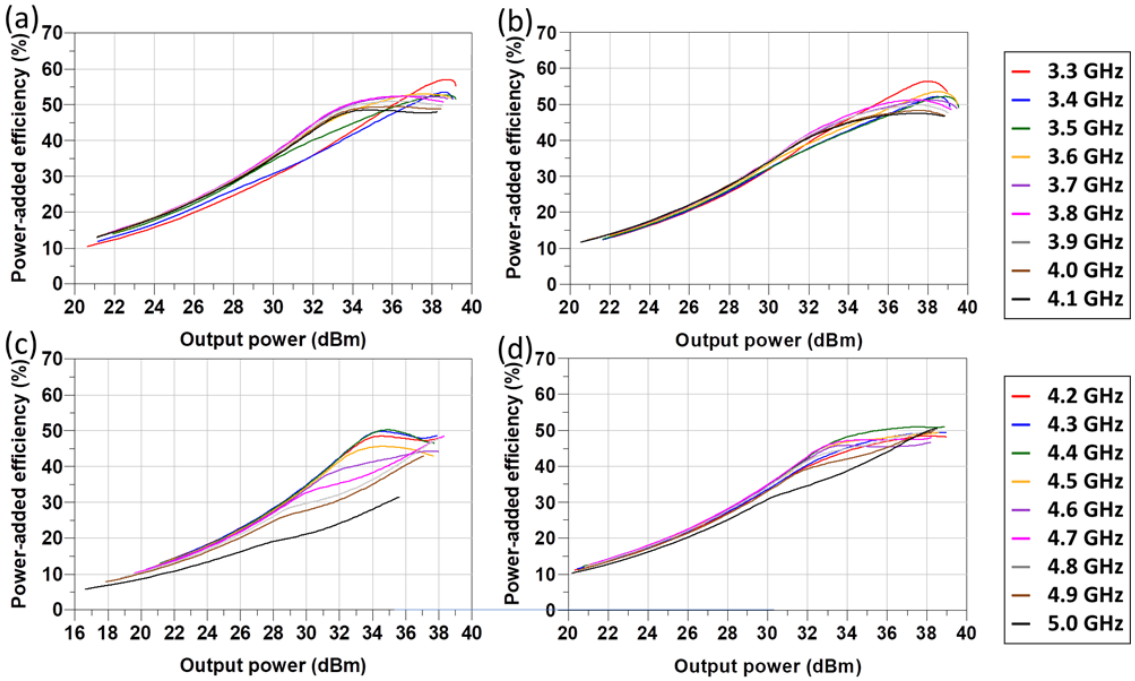


Figure 3.32 Simulated DPA PAE:

a) Low-f pre-tune, b) Low-f post-tune, c) High-f pre-tune, d) High-f post-tune

Chapter 4

3.3–5.0 GHz Linearity-Enhanced DPA Measurements

4.1 Single DPA Measurements

The DPA fabrication files were sent to a third-party PCB manufacturing company for PCB manufacturing and assembly. Figure 4.1 shows the fabricated DPA, of which four copies were built. In this section, each DPA unit is tested individually [S-parameters, continuous-wave (CW), and modulated signals] to compare performance to the simulation, and to check for performance consistency between the four units. In the next section, the four units will be used for modulated signal measurements in a 2×2 MIMO array.

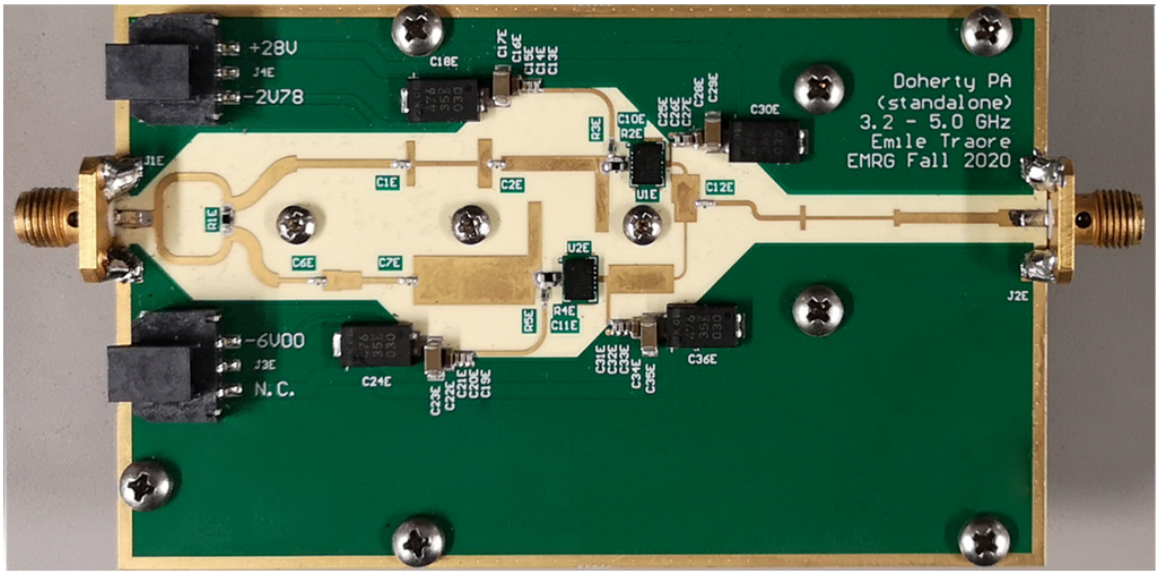


Figure 4.1 Photograph of the fabricated DPA

4.1.1 S-parameter Measurements

The setup shown in Figure 4.2 is used for measuring DPA small-signal performance. The effects of the cables and 30 dB attenuator are de-embedded from the vector network analyzer measurement results using linear calibration to isolate the performance of the DPA only. The four DPAs are biased at $V_{DD} = 28\ V$ and $I_{DS,Q} = 20\ mA$ as in simulation, with $V_{GS,Q}$ for the main transistor ranging from -2.7 V to -2.9 V due to threshold voltage variation between units. The S-parameters (s_{21}, s_{12}, s_{11}, s_{22}) are shown in Figure 4.3 a), b), c), and d), respectively. The S-parameters show good agreement between simulation and measurement, especially for s_{21}, and no frequency shift is observed. s_{11} and s_{22} differ in shape

between simulation and measurement, but the return loss at both ports remains above 10 dB for nearly the entire bandwidth. s_{12} is very flat across the entire bandwidth and increased by 3–4 dB compared to simulation. The consistency between all four units is remarkable, which will be beneficial during MIMO array performance since all units should be nearly identical, which is beneficial for array symmetry.

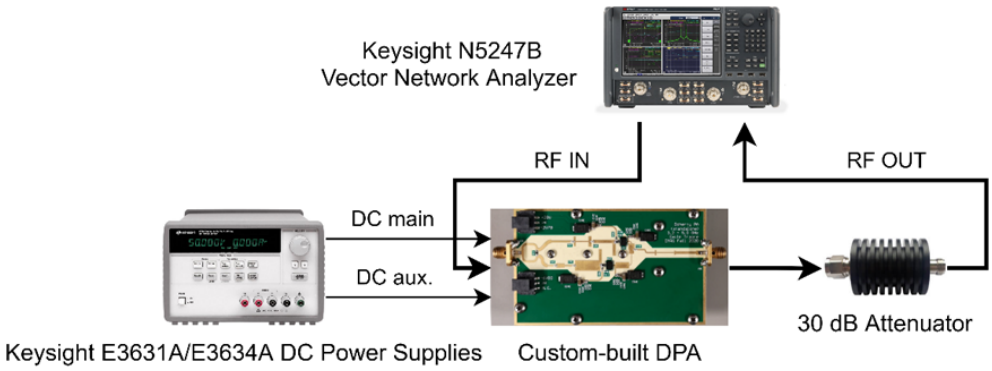


Figure 4.2 Block diagram of the DPA S-parameter measurement setup

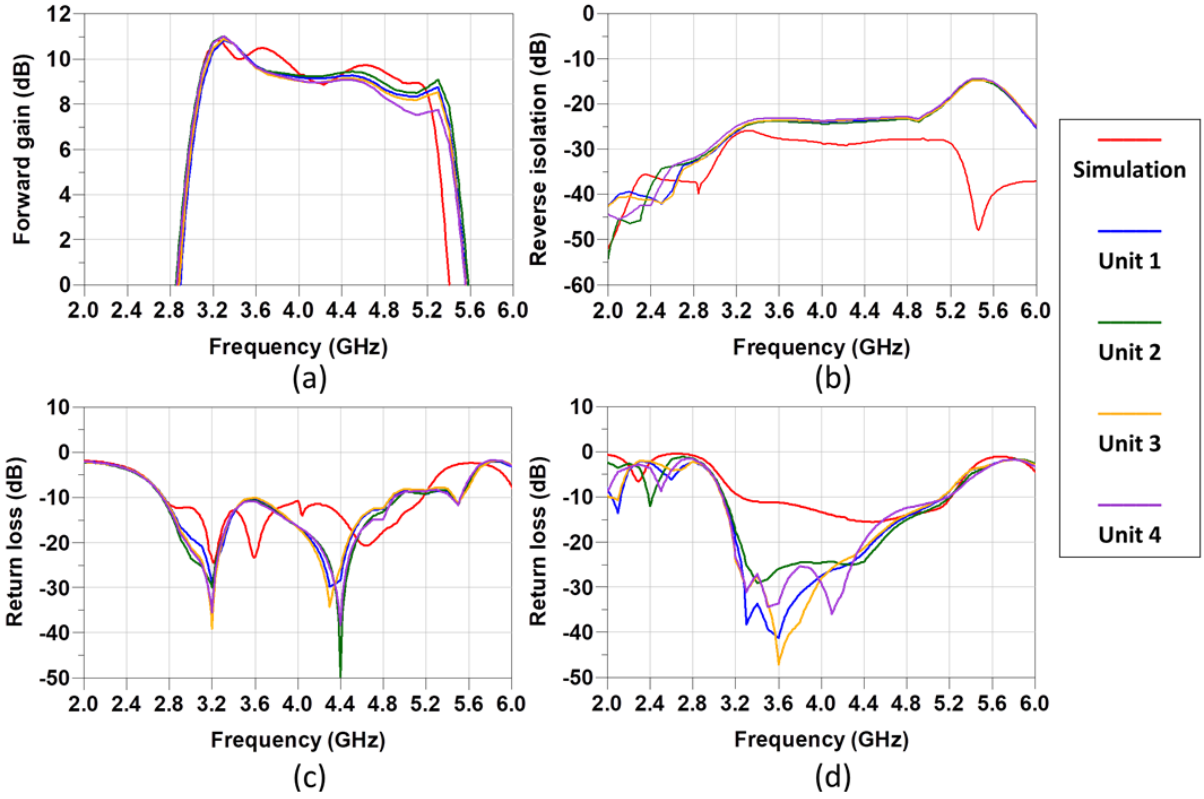


Figure 4.3 Measured DPA S-parameters: a) s_{21}, b) s_{12}, c) s_{11}, d) s_{22}

4.1.2 Continuous-Wave Measurements

The setup shown in Figure 4.4 is used for measuring DPA CW performance. Linear calibration is performed to de-embed the cables, then source and receiver power calibration are used to de-embedded the driver's linear gain and the 30 dB attenuator. The nonlinearities of the ZVE-3W-183+ driver (AM-AM and AM-PM) are de-embedded during data post-processing. The main transistor is biased at $I_{DS,Q} = 20\ mA$. The auxiliary transistor is biased to turn on at 6 dB OBO: units 1, 3, and 4 use $V_{GS,Q} = -6.4\ V$ for 3.3–4.1 GHz and $V_{GS,Q} = -6.0\ V$ for 4.2–5.0 GHz (unit 2 requires a 200 mV higher $V_{GS,Q}$). Figure 4.5 plots AM-AM and AM-PM and Figure 4.6 plots η_D and PAE for one of the DPA units. Figures 4.7 to 4.10 plot small-signal gain, P_{sat}, η_D at P_{sat}, and η_D at 6 dB OBO, respectively, for all four units. The CW performance of all four units compared to the simulation is summarized in Table 4.1. As seen in the aforementioned figures, the CW performance degraded in measurement compared to simulation. Across all units and frequencies, small-signal gain dropped by 0.6–1.2 dB and P_{sat} decreased by 0.1–0.4 dBm°. η_D at P_{sat} reached or exceeded the simulated value in the low frequency bands but fell 11–14 percentage points in the upper bands, and η_D at 6 dB OBO was 4–8 percentage points lower than in simulation. These discrepancies are likely due to inaccurate large-signal transistor performance modeling, especially for the auxiliary transistor, since S-parameter measurements showed excellent agreement between simulation and measurement. Between units, the CW performance of all four units is nearly indistinguishable, which is highly beneficial for array symmetry in MIMO tests. The exception is unit 4, which exhibits early auxiliary transistor turn-on. A slightly lower $V_{GS,Q}$ for the auxiliary transistor would delay the turn-on and improve η_D at OBO, at the expense of large-signal gain and P_{sat}.

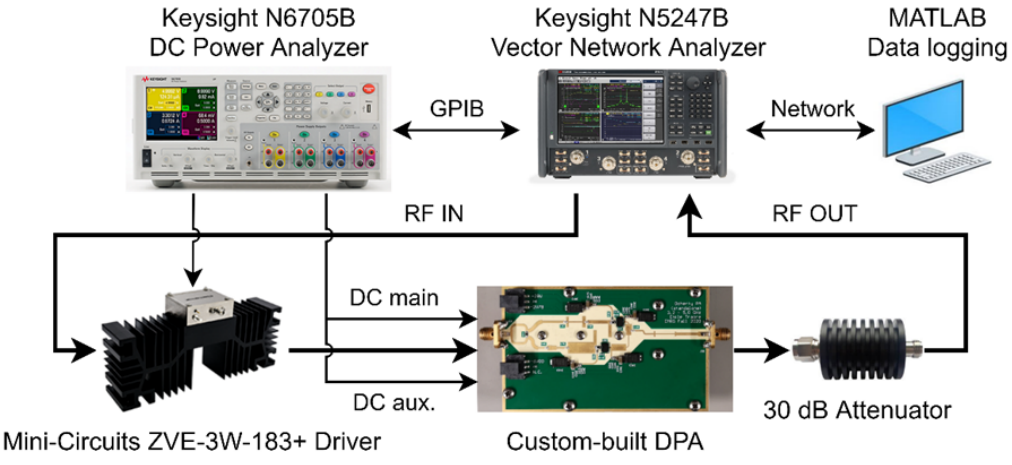


Figure 4.4 Block diagram of the DPA CW measurement setup

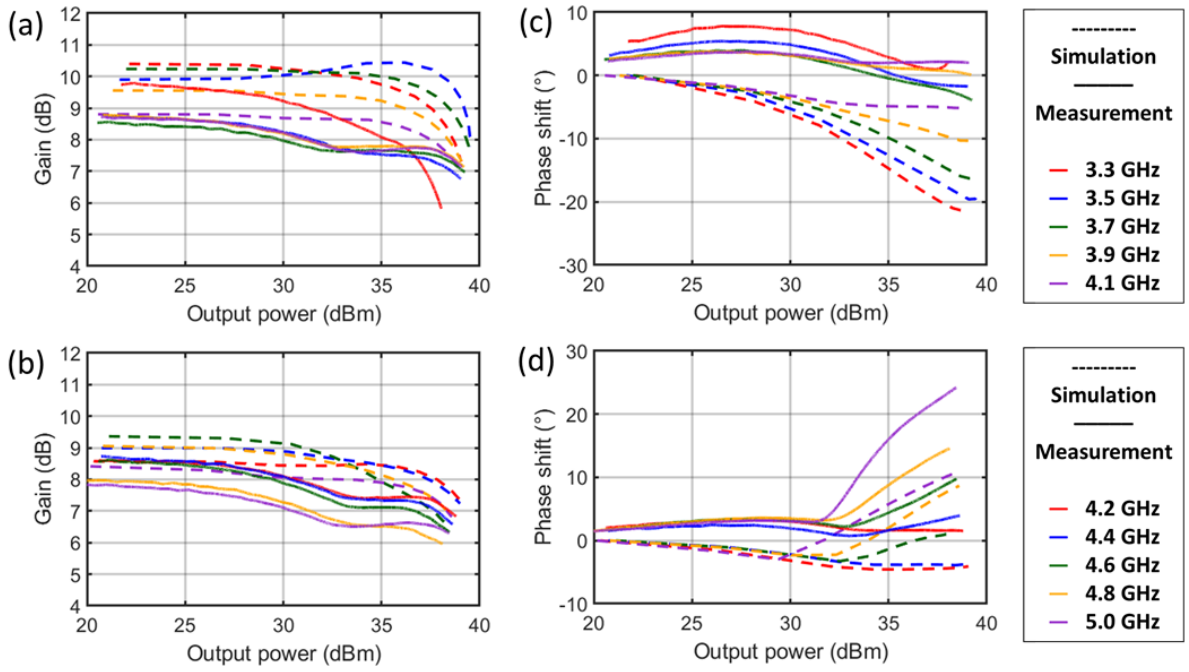


Figure 4.5 Measured DPA linearity performance vs. P_{out}: a) AM-AM over 3.3–4.1 GHz, b) AM-AM over 4.2–5.0 GHz, c) AM-PM over 3.3–4.1 GHz, d) AM-PM over 4.2–5.0 GHz

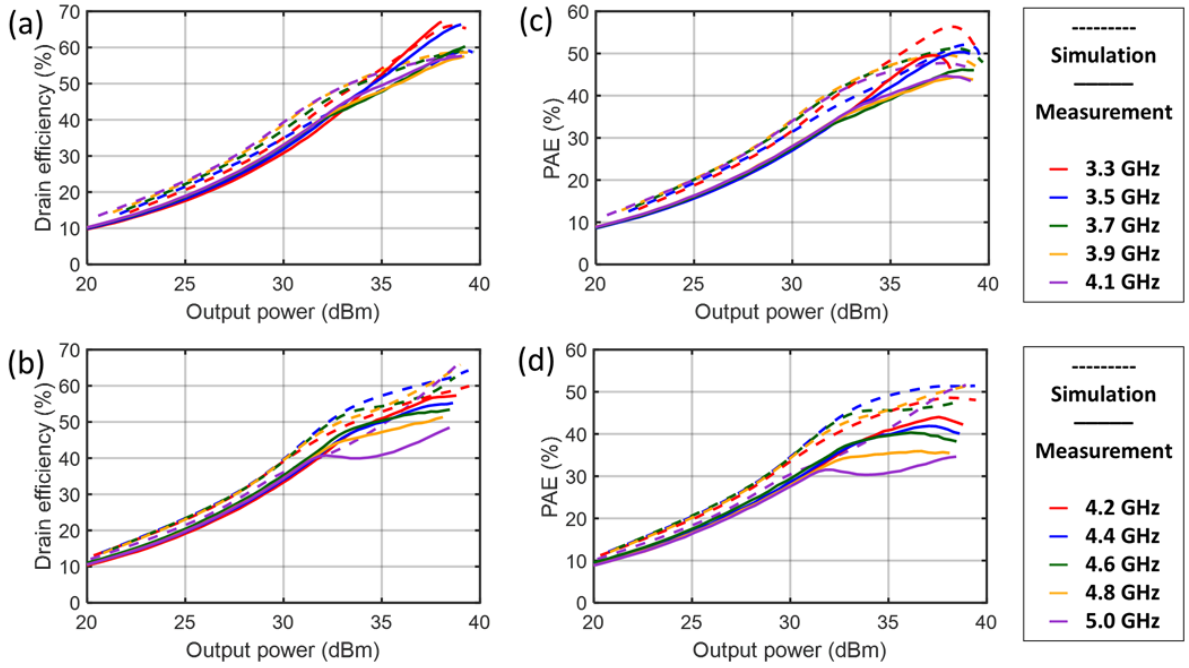


Figure 4.6 Measured DPA efficiency performance vs. P_{out}: a) η_D over 3.3–4.1 GHz, b) η_D over 4.2–5.0 GHz, c) PAE over 3.3–4.1 GHz, d) PAE over 4.2–5.0 GHz

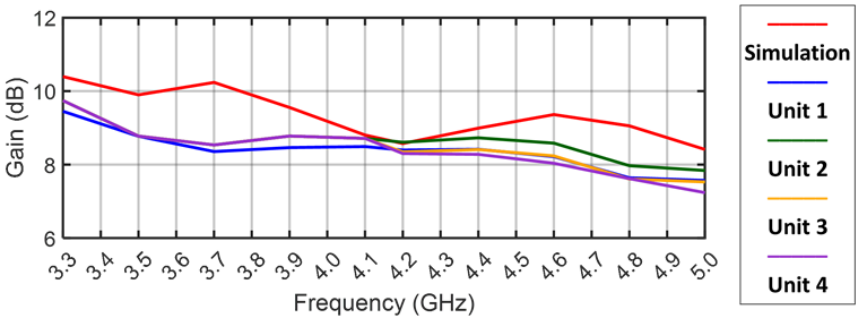


Figure 4.7 Measured DPA small-signal gain vs. frequency

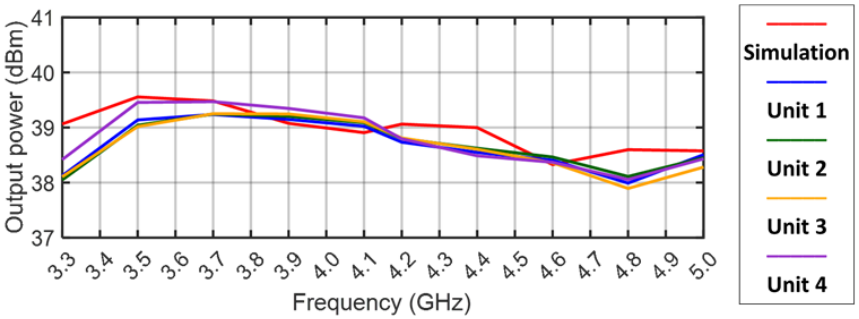


Figure 4.8 Measured DPA P_{sat} vs. frequency

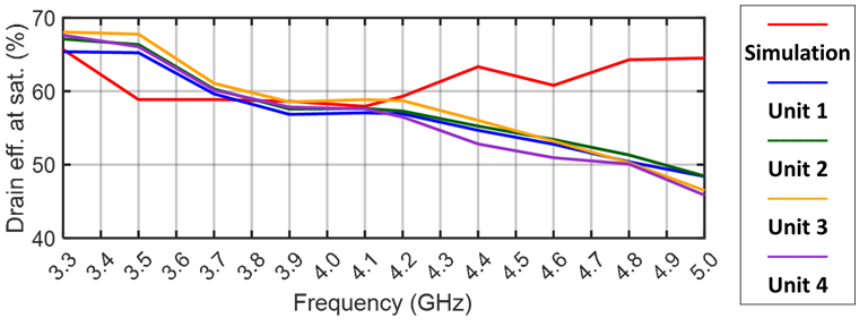


Figure 4.9 Measured DPA η_D at P_{sat} vs. frequency

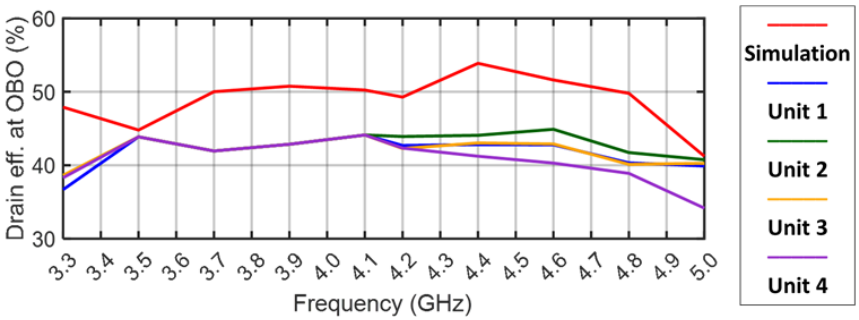


Figure 4.10 Measured DPA η_D at 6 dB OBO vs. frequency

Table 4.1 Comparison of measured DPA CW performance across units and vs. simulation

Unit	Bandwidth (GHz; %)	Gain (dB)	AM-PM (°)	P_{sat} (dBm)	η_D (P_{sat}; 6 dB OBO) (%)		PAE (P_{sat}; 6 dB OBO) (%)	
Sim.	3.3–5.0; 41	8.4–10.4	-20.9–10.9	38.3–39.6	59.3–65.7	42.0–52.3	46.8–53.5	35.4–45.2
1		7.6–9.5	-2.6–25.1	38.0–39.2	48.4–65.3	38.3–44.3	34.4–49.4	30.1–36.3
2		7.8–9.7	-4.0–24.2	38.0–39.2	48.4–67.1	38.6–45.1	34.6–49.9	31.1–37.1
3		7.5–9.7	-2.2–26.1	37.9–39.3	46.5–68.1	38.6–44.5	32.7–50.9	29.6–36.0
4		7.2–9.7	-4.0–25.2	38.1–39.5	45.8–67.6	33.8–43.7	32.7–51.2	25.0–35.9

4.1.3 Modulated Signal Measurements

The setup shown in Figure 4.11 is used to measure DPA modulated signal performance. The DPA is characterized using 100 and 200 MHz bandwidth OFDM signals with 8 dB PAPR, which are similar to realistic 5G signals deployed in the field. All four units are evaluated, with bias voltages set identically to the S-parameter and CW tests. Each DPA unit is paired with a Mini-Circuits ZHL-42+ driver for the entirety of the tests to ensure performance consistency. The RF input is generated from a Keysight M8190A arbitrary waveform generator (AWG) and is attenuated by 3 dB such that the entire dynamic range of the AWG can be utilized, reducing quantization noise.

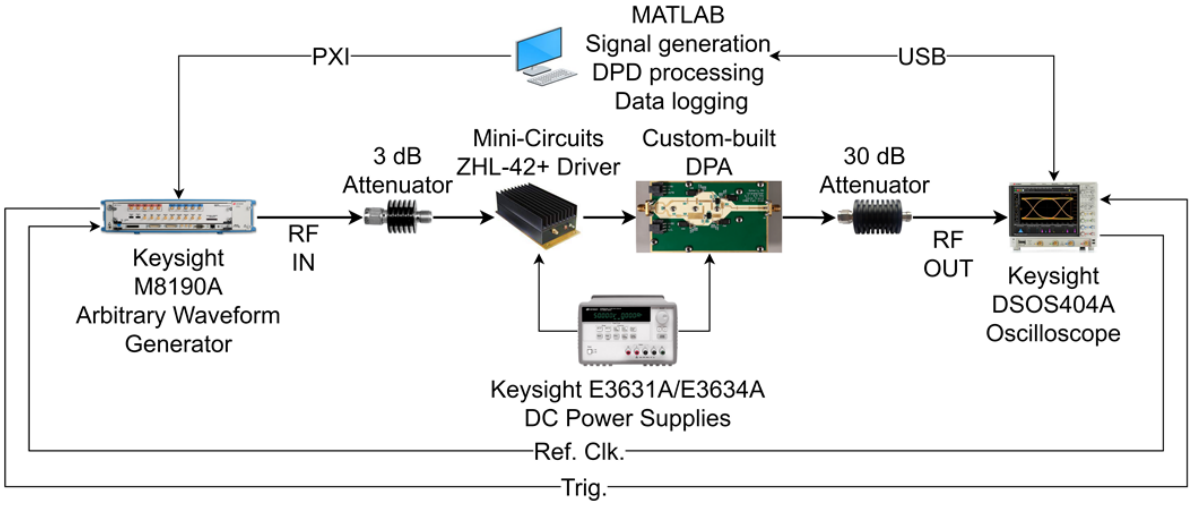


Figure 4.11 Block diagram of the DPA modulated signal measurement setup

First, various DPD algorithms and number of DPD coefficients are evaluated to determine the optimum settings for linearizing the DPA. These tests are conducted with unit 1 and a 100 MHz bandwidth signal centered at 3.5 GHz. The algorithms tested are a memory polynomial-dynamic reduction deviation algorithm (MDDR), as well as coefficient-reduced Volterra (CRV) [51]. The test results are shown in Table 4.2. The DPD algorithm that provides the best trade-off between linearity

and DPD complexity is MDDR with nonlinearity order $N = 7$ or $N = 8$, memory depth $M = 10$, and nonlinear memory depth $NM = 5$ or $NM = 6$, which results in 49 or 68 coefficients, respectively. Overall, the linearity performance of the DPA itself is very good due to the linearity-enhancement techniques employed in its design (AM-PM reduction and drain BN design), which are described in more detail in Chapter 3. When driven near saturation ($P_{out} \approx 31\,dB$, corresponding to $P_{sat} \approx 39\,dBm$ with 8 dB PAPR), the ACPR reaches below -52 dBc with 49 coefficients or -56 dBc with 68 coefficients. This is well below the base station ACPR limit of -45 dBc specified by the 3GPP, which sets the standards for cellular communication [52].

Table 4.2 Comparison of DPA modulated signal performance vs. DPD algorithm

DPD algorithm (N, M, NM, coeff.)	MDDR 7, 10, 5, 49	MDDR 8, 10, 6, 68	CRV 7, 10, 5, 31	CRV 10, 10, 9, 96
P_{out} (dBm) (before/after DPD)	31.0 30.9	31.1 31.0	31.1 31.0	31.1 31.0
NMSE (%) (before/after DPD)	11.0 0.5	10.9 0.4	10.9 0.8	11.0 0.6
ACPR lower (dBc) (before/after DPD)	-33.1 -52.6	-33.5 -56.7	-33.4 -50.7	-33.3 -55.9
ACPR upper (dBc) (before/after DPD)	-31.1 -52.0	-32.2 -56.0	-32.2 -51.0	-32.1 -55.9

Next, a frequency sweep is performed. These tests are conducted with unit 1 and a 100 MHz signal whose center frequency is varied from 3.25 GHz to 3.75 GHz. For consistency, P_{out} is kept consistent at 30 dBm (the highest that can be achieved over the entire band). The DPD algorithm used is MDDR with 49 coefficients. The entire DPA frequency range could not be characterized due to the oscilloscope bandwidth limit (4 GHz), which limits the maximum frequency to 3.75 GHz (due to an oversampling ratio of 5, the modulated signal occupies a bandwidth of ±250 MHz). The test results are shown in Table 4.3. This frequency sweep measurement reveals that the DPA linearity increases as frequency increases towards the DPA f_0 of 4.15 GHz (ACPR is -51 dBc at 3.25 GHz and rises to -57 dBc at 3.75 GHz). This is because as the frequency increases closer and closer to f_0, the DPA linearity improves (AM-PM gets closer and closer to a flat line and P_{sat} increases as shown in Figure 4.5). If modulated signal characterization with a higher frequency range could be realized, the ACPR would likely start decreasing above 4.15 GHz, as linearity becomes worse and worse (AM-PM starts to increase rapidly and P_{sat} starts to decrease).

Table 4.3 Comparison of DPA modulated signal performance vs. signal frequency

Signal frequency (GHz)	3.25	3.35	3.45	3.55	3.65	3.75
Output power (dBm) (before/after DPD)	30.1 29.7	30.1 30.0	30.1 30.1	30.1 30.0	30.1 30.1	30.1 30.1
NMSE (%) (before/after DPD)	11.0 0.7	6.9 0.6	9.9 1.7	7.0 0.4	6.3 0.7	6.3 0.4
ACPR lower (dBc) (before/after DPD)	-31.7 -50.7	-32.9 -53.8	-33.7 -58.4	-34.9 -57.5	-36.5 -57.6	-36.7 -57.4
ACPR upper (dBc) (before/after DPD)	-29.3 -51.3	-30.7 -55.8	-32.1 -57.0	-57.5 -57.7	-57.6 -57.9	-57.4 -57.5

Finally, all four DPA units are characterized with 100 and 200 MHz signals centered at 3.5 GHz to determine the variation in linearity performance between units. The test results are shown in Table 4.4. For the 100 MHz test, MDDR with 49 coefficients is used, resulting in consistent linearity performance between units (ACPR between -51 and -54 dBc across all units when driven near saturation). For the 200 MHz test, MDDR with 68 or 91 coefficients is used, but the higher number of coefficients does not result in an appreciable improvement in ACPR, so the fundamental DPA linearity performance for 200 MHz is established. Average η_D is also recorded for these tests, showing consistent efficiency behavior between units (η_D ranging from 38.0 to 39.5% across DPA units, and not varying significantly between 100 and 200 MHz).

Table 4.4 Comparison of DPA modulated signal performance vs. signal bandwidth for all units

DPA unit	1			2			3			4		
Signal bandwidth (MHz)	100	200	200	100	200	200	100	200	200	100	200	200
Number of coeff.	49	68	91	49	68	91	49	68	91	49	68	91
P_{out} (dBm) (before/after DPD)	30.8 30.7			30.7 30.4			30.7 30.6			31.0 30.9		
NMSE (%) (before/after DPD)	9.4 0.5	18.1 0.8	18.2 0.8	13.4 0.6	17.5 0.8	17.5 0.8	10.5 0.6	17.9 0.9	17.8 0.8	10.0 0.5	17.2 0.7	17.2 0.6
ACPR lower (dBc) (before/after DPD)	-33.0 -52.6	-33.2 -49.5	-33.3 -50.5	-31.8 -54.6	-32.5 -52.4	-32.5 -53.0	-32.5 -51.8	-32.7 -48.8	-32.8 -49.7	-32.6 -53.5	-33.2 -51.9	-33.3 -52.7
ACPR upper (dBc) (before/after DPD)	-31.3 -52.0	-31.3 -50.4	-31.4 -50.0	-30.2 -53.8	-30.0 -50.2	-30.0 -50.8	-30.3 -51.2	-30.5 -49.7	-30.6 -49.5	-30.6 -52.9	-30.9 -51.6	-31.0 -51.6
Average η_D (%)	38.7			38.0			38.3			39.5		

Figures 4.12, 4.13, and 4.14 compare AM-AM, AM-PM, and spectra between 100 and 200 MHz, respectively. The AM-AM and AM-PM curves are much thicker for 200 MHz due to significantly higher memory effects. The spectrum before DPD is much less flat in-band at 200 MHz because the DPA gain is not flat over such a wide bandwidth. After DPD, the 200 MHz spectrum also shows significant residual nonlinearity (sidebands), whereas the spectrum in 100 MHz is much flatter outside of the desired band. Overall, the DPA shows much better behavior under 100 MHz compared to 200.

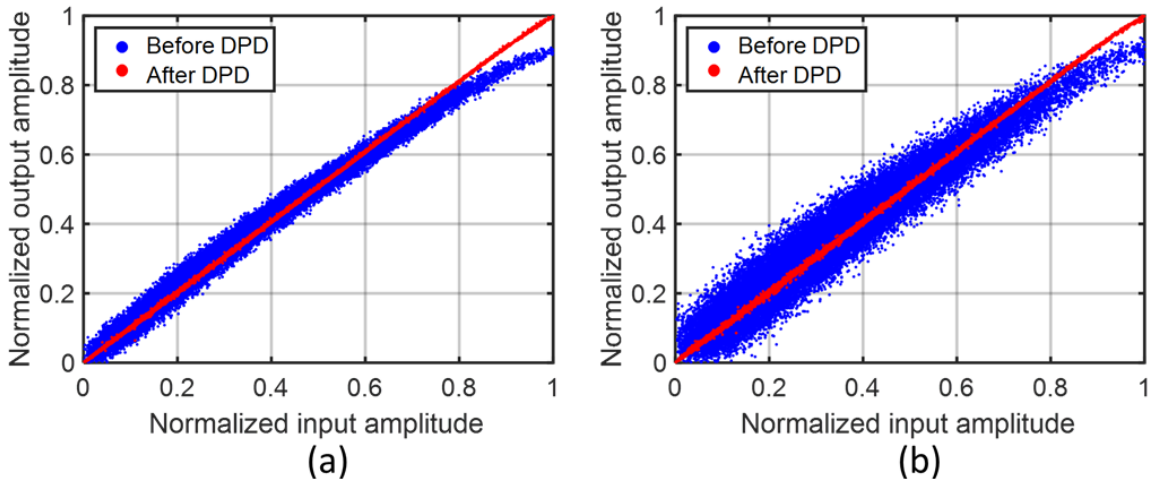


Figure 4.12 AM-AM distortion before and after DPD for: a) 100 MHz, b) 200 MHz

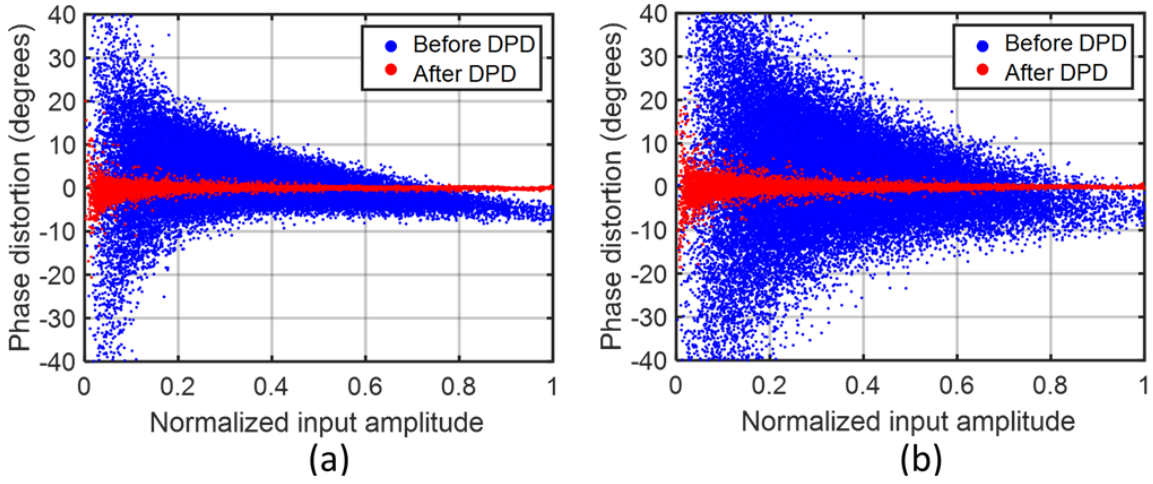


Figure 4.13 AM-PM distortion before and after DPD for: a) 100 MHz, b) 200 MHz

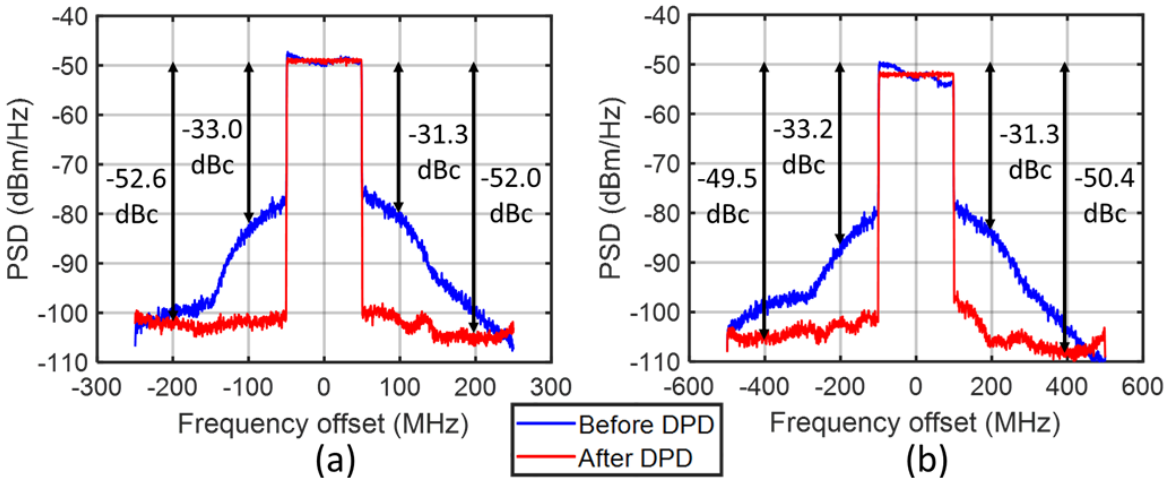


Figure 4.14 Output spectra before and after DPD for: a) 100 MHz, b) 200 MHz

4.2 2×2 DPA Array Measurements

The measurements in this section assess the effects of antenna load mismatch and crosstalk on the linearity and efficiency of the fabricated DPA in a MIMO context, under wideband modulated signals. The DPA is evaluated by connecting it to a custom-designed 2×2 dual-polarized antenna array with variable array spacing for variable crosstalk level between antenna elements. The same study is also performed with a custom-designed class AB PA based on the same transistor to compare the linearity and efficiency of a conventional single-ended PA in a MIMO setup. The study also uses two DPD schemes, conventional single input single output (SISO), as well as dual input single output (DISO) to compare their ability to linearize the PA array.

4.2.1 2×2 Array Test Setup Description

The setup shown in Figure 4.15 is used to measure PA modulated signal performance for both the class AB PA and the DPA presented in this work (DC power supplies are not shown for simplicity). A fully digital beamforming architecture is used, capable of supporting four independent (uncorrelated) OFDM signals (3.4 GHz center frequency, 100 MHz bandwidth, and 8 dB PAPR). In each path, the signal is generated using MATLAB, uploaded to an M8190A AWG, amplified by the Mini-Circuits ZHL-42+ driver, and applied to the PA under test. The PAs are connected to the 2×2 antenna array which acts as the load (since only four PAs are available, only one of the polarizations is used, even though the array could support up to eight PAs with both polarizations). A 20 dB Anaren X4C40K1-20S directional

coupler samples each PA's output signal into a Keysight DSOS404A oscilloscope. The measured signal is then sent back to the PC, which performs offline DPD processing using either SISO or DISO DPD.

Figure 4.16 shows the fabricated dual-polarized antenna array used in this study, which was designed in-house, Figure 4.17 a) shows the S-parameters of a single antenna element, demonstrating good matching (return loss above 10 dB) over a 3.2–3.8 GHz bandwidth. Figure 4.17 b) shows the worst-case coupling between neighboring antenna elements over the three possible antenna spacings in terms of wavelength (0.6 λ, 0.7 λ, and 0.8 λ). A frequency of 3.4 GHz is chosen for the signal, such that antenna coupling does not drop too low to become irrelevant. The resulting coupling values for the three spacings are -21.4, -24.0, and -26.3 dB, respectively.

The SISO DPD scheme is the same conventional DPD used in the previous section single PA measurement. The DPA uses MDDR with 68 coefficients, while the class AB PA uses CRV with 31 coefficients. The DISO DPD scheme was first introduced in [53] to mitigate the combined effects of PA nonlinearity and antenna crosstalk. Figure 4.18 shows the DISO DPD block diagram. The DPD output in each PA path (z_i) is a function of both the PA input signal (x_i) and the output signal of the crosstalk and mismatch (CTMM) block corresponding to this path. The CTMM approximates the crosstalk signal reflected from the antenna array to each PA's output port ($\hat{a}_{2,i}$) as a linear combination of the signals in the other paths. A small number of additional coefficients are required by the DISO DPD algorithm, 12 for the class AB PA (43 total) and 14 for the DPA (84 total).

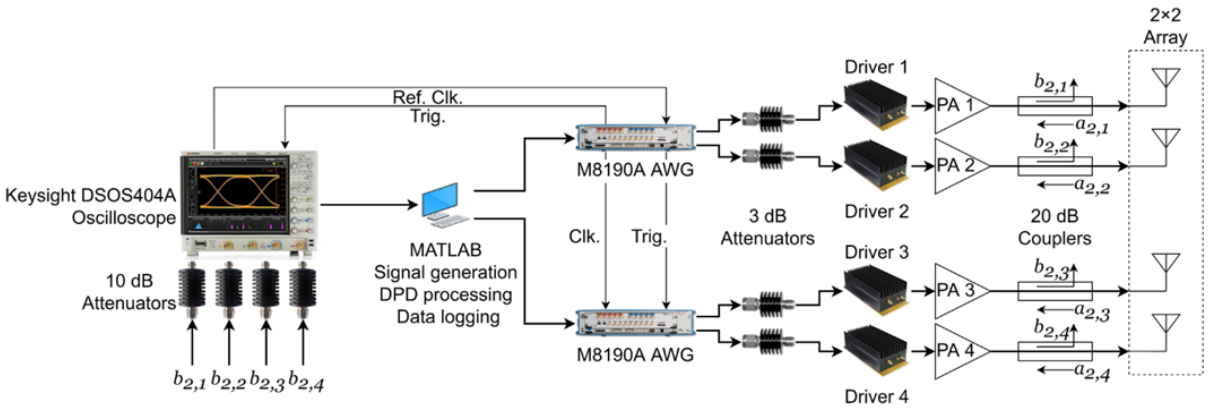


Figure 4.15 Block diagram of the 2×2 PA array modulated signal measurement setup

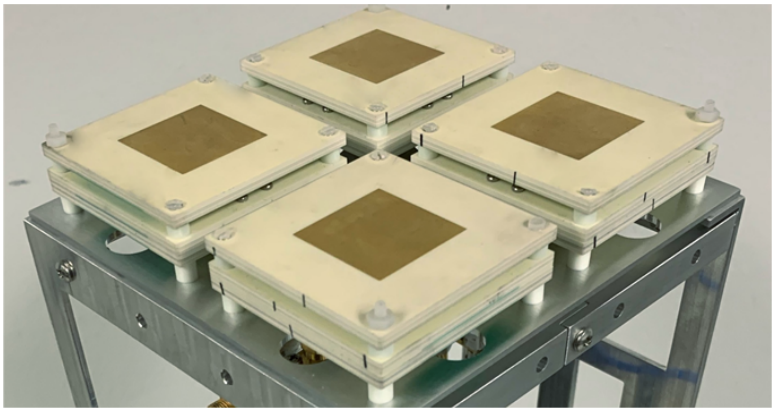


Figure 4.16 Photograph of the fabricated antenna array

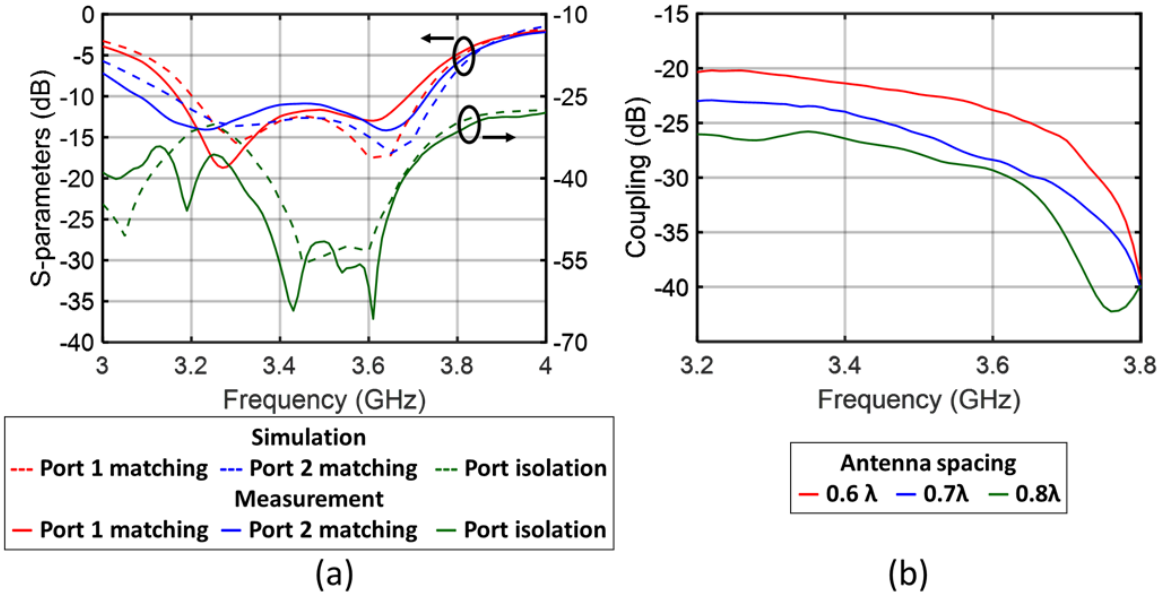


Figure 4.17 Antenna array: a) Single element S-parameters, b) Worst-case coupling between elements

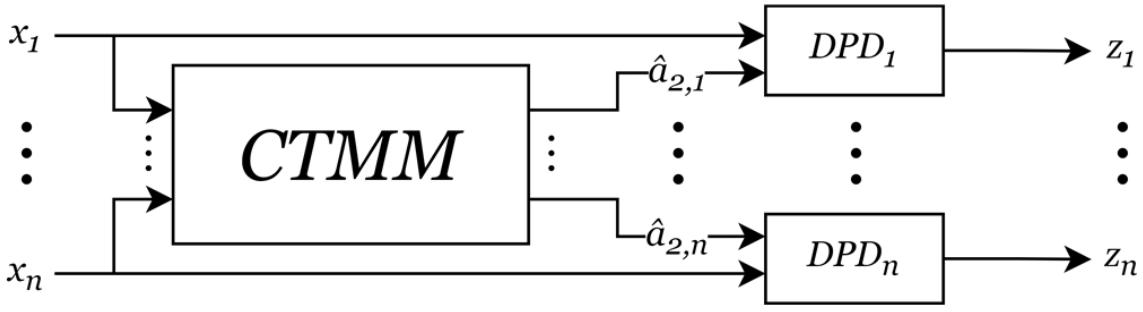


Figure 4.18 Block diagram of DISO DPD including CTMM block

4.2.2 2×2 Array Measurement Results

Table 4.5 shows the measurement results of the 2×2 class AB and DPA arrays. The NMSE and ACPR are calculated for the combined signal, which is the vector sum of all four PA output signals (NMSE is calculated by comparing the combined output signal to the combined input signal, and ACPR, by comparing in-band and out-of-band emissions of the combined signal). This method of measuring ACPR and NMSE yields similar results to an over-the-air (OTA) measurement at boresight (transmitter and receiver antennas facing each other directly). First, only a single chain is turned on, and the PA performance with antenna load is compared to the performance when terminated with 50 Ω (data from the previous section). This test shows performance degradation due to antenna load mismatch, but not crosstalk. Next, all four chains are turned on, and the array performance is compared for 0.6 λ, 0.7 λ, and 0.8 λ spacings, with SISO and DISO DPD. These tests show the combined effects of antenna load mismatch and variable crosstalk levels and the linearity improvement of DISO DPD over SISO DPD.

Table 4.5 Performance summary of the 2×2 class AB PA and DPA arrays under modulated signals

			NMSE of combined signal (%)		ACPR of combined signal (lower/upper) (dBc)		Average η_D(%)	Total P_{out} (dBm)	Number of DPD coeff.
			Before DPD	After DPD	Before DPD	After DPD			
Single AB PA	50 Ω load		8.8	0.4	-33.8 / -31.7	-54.3 / -54.8	29.1	28.8	31
	Antenna		7.9	0.8	-33.1 / -31.5	-49.2 / -50.7	21.1	27.4	
Four chain AB PA	SISO	0.6 λ	9.4	2.7	-34.4 / -32.7	-44.3 / -45.0	21.5	33.4	
		0.7 λ	9.5	2.9	-34.4 / -32.6	-46.5 / -47.0	21.3	33.4	
		0.8 λ	9.4	3.2	-34.4 / -32.7	-47.1 / -47.9	21.0	33.4	
	DISO	0.6 λ	9.2	1.3	-34.4 / -32.7	-48.5 / -49.2	21.3	33.4	43
		0.7 λ	10.2	1.3	-34.7 / -32.9	-49.2 / -49.8	21.2	33.4	
		0.8 λ	9.4	1.3	-34.7 / -32.8	-49.4 / -50.2	21.0	33.4	
Single DPA	50 Ω load		9.9	0.5	-32.8 / -30.6	-54.8 / -54.8	38.3	30.8	68
	Antenna		9.7	0.8	-32.7 / -31.4	-50.0 / -49.9	31.1	28.8	
Four chain DPA	SISO	0.6 λ	9.0	2.2	-31.9 / -32.3	-43.0 / -43.1	31.0	34.6	
		0.7 λ	9.5	2.2	-31.6 / -31.7	-45.8 / -46.1	30.5	34.7	
		0.8 λ	8.9	1.5	-32.7 / -32.2	-48.6 / -48.6	31.4	34.6	
	DISO	0.6 λ	9.1	1.3	-31.8 / -32.2	-48.0 / -49.0	30.5	34.6	84
		0.7 λ	9.3	1.2	-32.2 / -31.6	-49.6 / -50.0	31.2	34.7	
		0.8 λ	9.0	1.4	-32.6 / -32.1	-49.7 / -49.7	30.5	34.5	

The results highlight several interesting trends about PA performance under antenna load mismatch and crosstalk. Load mismatch alone is responsible for significant P_{out} and η_D degradations: 1.2 dB and 8.0 percentage points for the class AB PA, 2.0 dB, and 7.2 percentage points for the DPA. Linearity also degrades, with ACPR increasing from approximately -55 dBc to approximately -50 dBc for both

PAs. These degradations occur because the PAs were designed for 50 Ω terminations. This could be resolved by antenna-PA co-design (optimizing the PA for the load presented by the antenna) if the antenna array parameters are known in advance, or reconfigurable solutions such as the B2D PA [10] and QB DPA [25] if the PA is designed to work with any antenna. The 2×2 array results highlight the effects of crosstalk. P_{out} per PA and η_D vary negligibly regardless of crosstalk level because the average load impedance to the PAs remains constant, despite instantaneous deviations from the mean due to the crosstalk-induced active load modulation. As for linearity, Figure 4.19 shows the output spectra for the class AB PA, and Figure 4.20, for the DPA, at all three array spacings, with both SISO and DISO DPD. With SISO DPD, ACPR degrades as crosstalk increases, ranging from -47 to -44 dBc for the class AB PA and -49 to -43 dBc for the DPA as array spacing varies from 0.8 λ to 0.6 λ. However, DISO DPD can linearize both class AB and DPA arrays to nearly the same level as the single PA with antenna load (ACPR below -48 dBc in all cases), which exceeds the 3GPP ACPR specifications [52].

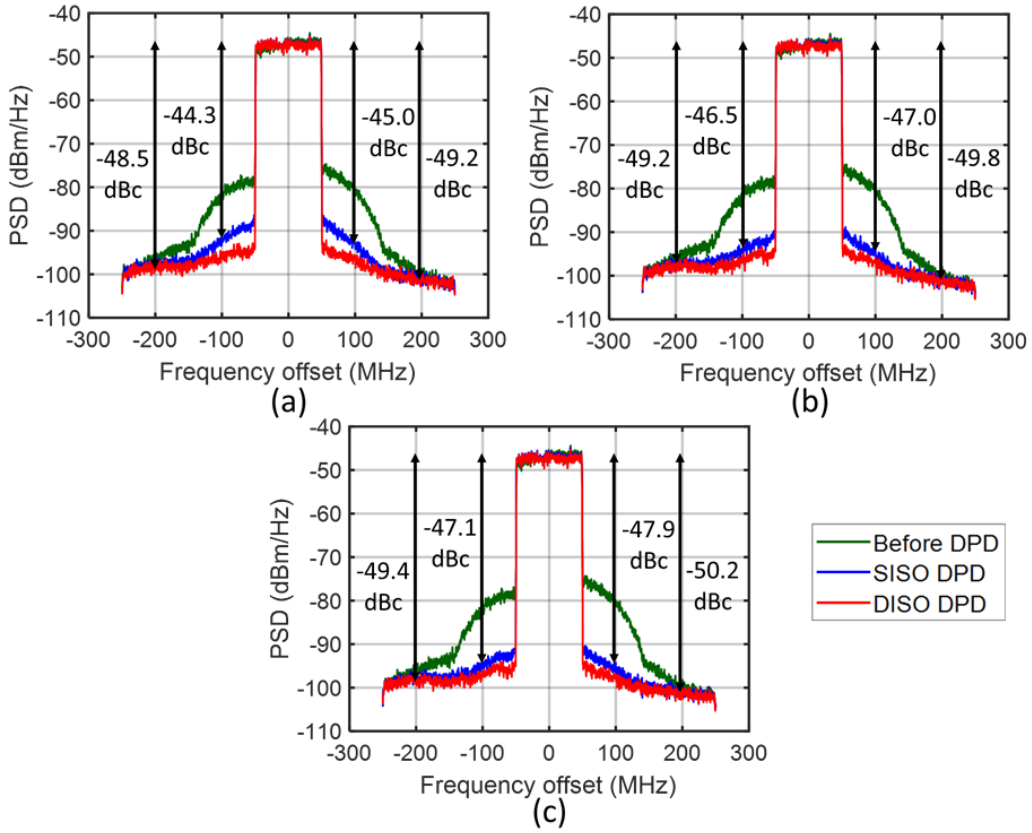


Figure 4.19 Spectra of the class AB PA array for various DPD schemes at: a) 0.6 λ, b) 0.7λ, c) 0.8 λ

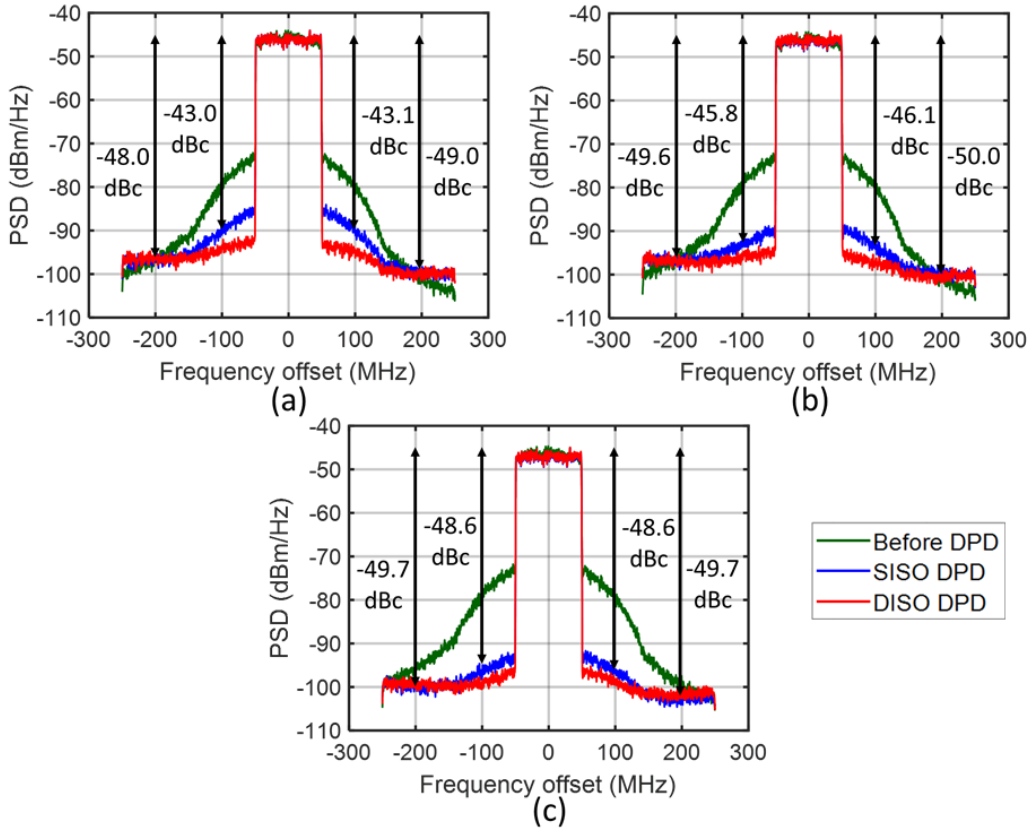


Figure 4.20 Spectra of the DPA array for various DPD schemes at: a) 0.6 λ, b) 0.7λ, c) 0.8 λ

The behavior of individual PAs in the array compared to the overall output signal is shown in the following figures. Figures 4.21, 4.22, and 4.23 compare the AM-AM, AM-PM, and spectra, respectively, for the output of a single DPA in the array and the combined output signal. These results show that the output of each PA in the array is not nearly linearized to the same level as the combined output signal, based on the spread of AM-AM and AM-PM plots, as well as the residual sidebands in the spectrum. Indeed, since the DPD focuses on the combined signal, it only corrects for PA nonlinearities that add up constructively in the output signal, not those that interfere destructively. Thus, the linearization method based on the combined signal is more effective in terms of resource consumption than attempting to linearize each PA individually to the same degree.

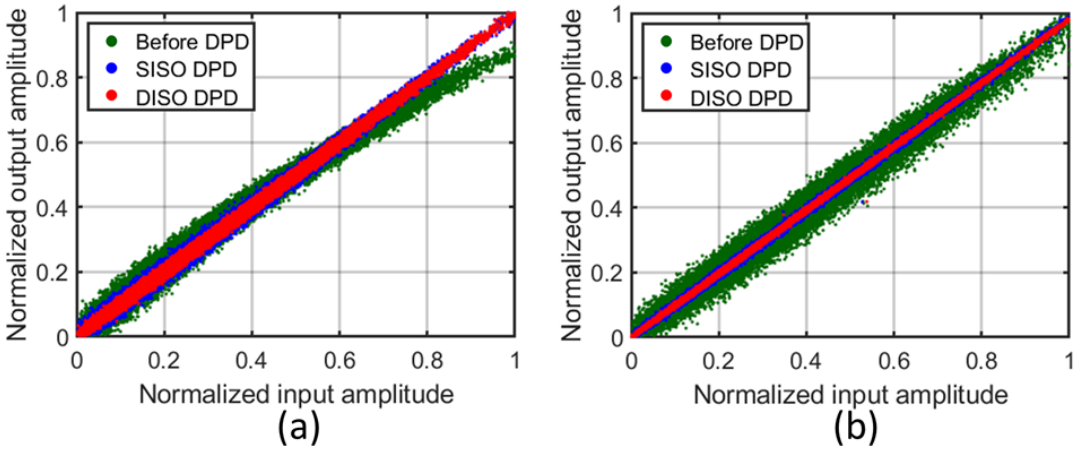


Figure 4.21 2×2 array AM-AM distortion for: a) Single PA output, b) Combined output signal

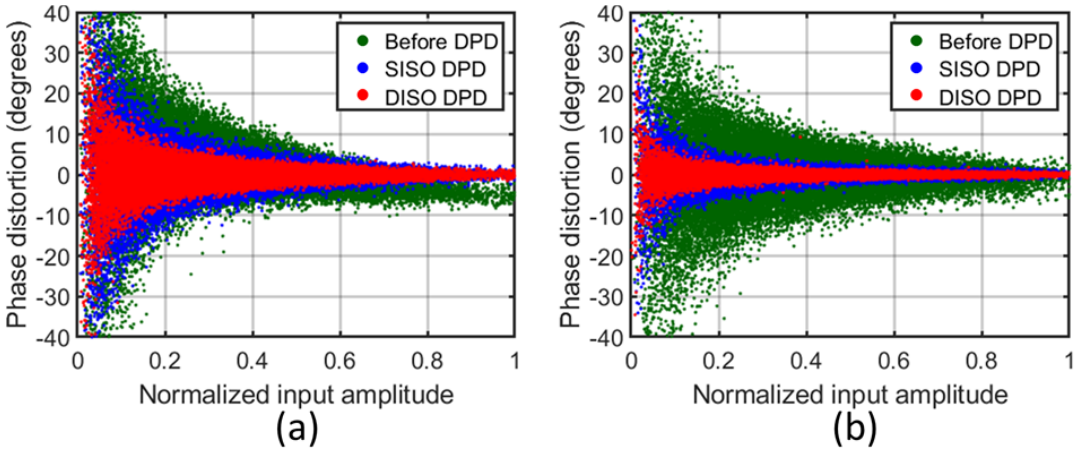


Figure 4.22 2×2 array AM-PM distortion for: a) Single PA output, b) Combined output signal

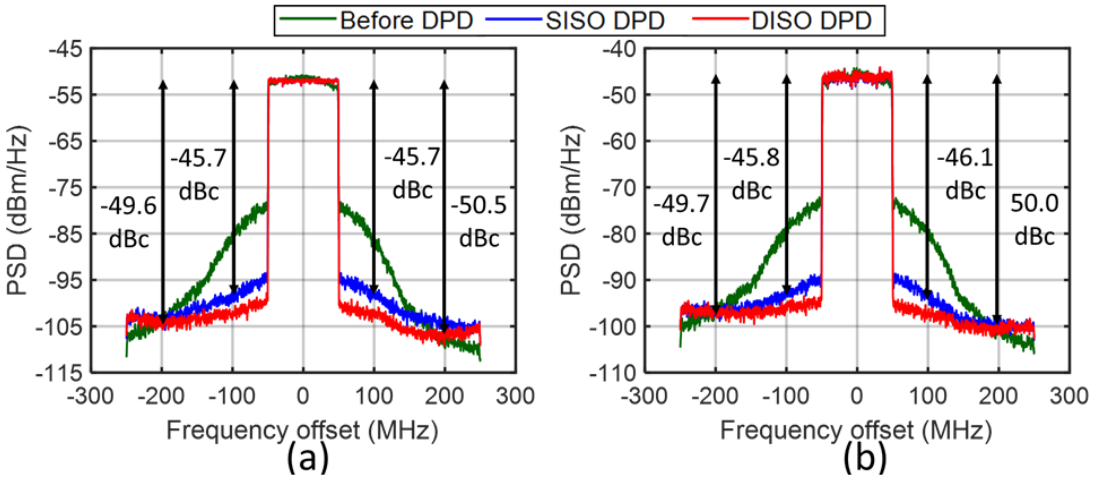


Figure 4.23 2×2 array spectra for: a) Single PA output, b) Combined output signal

Chapter 5
Conclusions and Future Work

5.1 Conclusions

This thesis began with Chapter 1 discussing the implications of 5G cellular communications on PA design. The increased bandwidth requires broadband PA design techniques, the increased PAPR requires OBO efficiency-enhancement techniques, and the emergence of massive MIMO requires improved linearity to comply with linearity standards in the presence of antenna crosstalk. Furthermore, the 5G sub-6 GHz spectrum includes new frequency bands in the 3.3–5.0 GHz range. Thus, existing 4G PAs are not suitable for 5G applications, and new PA design techniques are required.

Next, Chapter 2 discussed various OBO efficiency-enhancement techniques such as the EER and ET PAs, outphasing PA, and DPA. The DPA was shown to be most suitable for MIMO implementation due to its simplicity, as the load modulation occurs in the RF domain, without the need for an external circuit. However, in the subsequent literature review, the surveyed state-of-the-art DPAs did not operate in the 3.3–5.0 GHz range (most were at lower frequencies) and were not evaluated in a MIMO context (as a PA array loaded with an antenna array), so they were not suitable for 5G applications.

Then, Chapter 3 discussed the design procedure for a 3.3–5.0 GHz linearity-enhanced 6 dB OBO DPA designed for 5G applications. The wide FBW of 41% is achieved by using an inverted DPA OCN composed of low-order networks, as well as a high-order PMN. The linearity is enhanced by designing the BNs to provide a low baseband impedance, and by selecting transistor source impedances to reduce AM-PM variation. Simulation results show 8.4–10.4 dB gain, -20.9°–10.9° peak AM-PM, 38.3–39.6 dBm P_{sat}, and drain efficiencies of 59.3–65.7% at P_{sat} and 42.0–52.3% at 6 dB OBO, superior to the results in [23] while also operating at higher frequencies.

Finally, Chapter 4 discussed the measurement results of the four fabricated DPA units. Small-signal measurements demonstrated incredibly consistent DPA performance between units and good agreement between simulation and measurement results (no frequency downshift). CW measurements showed some degradation compared to simulation: 7.8–9.7 dB gain, -4.0°–24.2° peak AM-PM, 38.0–39.2 dBm P_{sat}, and drain efficiencies of 48.4–67.1% at P_{sat} and 38.6–45.1% at 6 dB OBO. However, performance remained competitive with the work in [23] which targeted similar specifications. To the author's best knowledge, this is the only 3.3–5.0 GHz DPA reported in the literature, which demonstrates that all new 5G sub-6 GHz bands can be serviced using a single PA.

Under wideband modulated signal excitation, the DPA offered very good linearity: ACPR below -50 dBc is achievable at 100 or 200 MHz bandwidth, with an appropriate DPD algorithm and number of coefficients. This chapter concluded with a MIMO PA linearity and efficiency study, comparing this work's 2×2 DPA array with a 2×2 class AB PA array under wideband modulated signals. The PA arrays were loaded with a 2×2 antenna array, subjecting the PAs to antenna load mismatch and crosstalk between PAs, which was varied based on antenna spacing. This study revealed that load mismatch is the main factor for performance degradation in MIMO systems, as it worsened average output power and efficiency as well as linearity. Antenna crosstalk mainly affected linearity, as it degraded ACPR under SISO DPD, but DISO DPD was able to successfully linearize the transmitter (ACPR below -48 dBc), even under the highest crosstalk level. Despite the detrimental effects of load mismatch and crosstalk in MIMO systems, the DPA remains the preferred architecture over class AB PAs, since it can achieve a higher drain efficiency and similar linearity. To the author's best knowledge, this is the only work in the literature that compared class AB PA and DPA MIMO performance using 5G-candidate signals and a fully digital beamforming architecture.

5.2 Future Work

This work proved it is possible to enhance the RF performance of DPAs for use in 5G communications, however, several avenues could be pursued in the future to achieve further improvements. Although the fabricated DPA demonstrated a very good agreement with simulation in small-signal measurements (no frequency shift), there were some discrepancies in large-signal measurements, such as AM-AM, AM-PM, and efficiency (which were all worse compared to simulation). Since the discrepancy is in large-signal measurements, this points to an issue with the transistor modeling used in this work. Future refinement of the design methodology could reduce the discrepancy between simulation and measurement results.

In addition, the performance specifications of the DPA could be improved. For example, the drain bias voltage was reduced from 40 V to 28 V to ease matching requirements, but if an intermediate value were selected, the PA output power could be increased. Furthermore, using asymmetric drain bias voltages for the main and auxiliary transistors could enable a higher OBO efficiency enhancement level such as 8 or 9.5 dB, one that would be of more practical relevance than the 6 dB obtained in this work given that 5G-like modulated signals can have a PAPR of 8 dB or higher.

Additional DPA modulated signal measurements could also be performed. Indeed, the entire frequency range could not be characterized in the modulated signal test since the oscilloscope had a limited measurement bandwidth of up to 4 GHz. By using a higher bandwidth oscilloscope or spectrum analyzer to process the received signal, the entire 3.3–5.0 GHz bandwidth could be assessed with modulated signals. Furthermore, only OFDM signals were used, but the DPA should also be tested with carrier aggregated signals, including non-contiguous signals (two frequency bands with a gap in the middle).

Moreover, MIMO array measurements are still in their infancy; a deeper understanding of the array's performance could be gained if additional tests could be performed. For example, the center frequency and bandwidth of the modulated signals could be varied to characterize the system over the entire antenna bandwidth of 3.2–3.8 GHz. Also, OTA measurements using a receiver antenna were not performed, only conductive measurements were taken. OTA measurements could help fully characterize the MIMO transmitter, including its linearity performance (ACPR and NMSE) over different steering angles, not just at boresight. Finally, since the goal is to realize a massive MIMO array, a larger array size (e.g., 4×4 or larger) should be developed and tested. The designed 2×2 transceiver is scalable and could be used in such a configuration.

Bibliography

[1] C. Huang, S. He, and F. You. "Design of Broadband Modified Class-J Doherty Power Amplifier With Specific Second Harmonic Terminations," *IEEE Access*, vol. 6, pp. 2531-2540, 2018.

[2] H. Liu, C. Zhai, and K. M. Cheng. "Novel Dual-Band Equal-Cell Doherty Amplifier Design With Extended Power Back-Off Range," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 3, pp. 1012-1021, Mar. 2020.

[3] M. Li, J. Pang, Y. Li, and A. Zhu, "Ultra-Wideband Dual-Mode Doherty Power Amplifier Using Reciprocal Gate Bias for 5G Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 10, pp. 4246-4259, Oct. 2019.

[4] J. Pang, Z. Dai, Y. Li, M. Li, and A. Zhu, "Multiband Dual-Mode Doherty Power Amplifier Employing Phase Periodic Matching Network and Reciprocal Gate Bias for 5G Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 6, pp. 2382-2397, Jun. 2020.

[5] M. Liu, H. Golestaneh, and S. Boumaiza, "A concurrent 2.15/3.4 GHz dual-band Doherty power amplifier with extended fractional bandwidth," *2016 IEEE MTT-S International Microwave Symposium (IMS)*, San Francisco, CA, USA, May 2016, pp. 1-3.

[6] M. Olavsbråten and D. Gecan, "Bandwidth Reduction for Supply Modulated RF PAs Using Power Envelope Tracking," *IEEE Microwave and Wireless Components Letters*, vol. 27, no. 4, pp. 374-376, Apr. 2017.

[7] H. Huang, B. Zhang, C. Yu, J. Gao, Y. Wu and Y. Liu, "Design of Multioctave Bandwidth Power Amplifier Based on Resistive Second Harmonic Impedance Continuous Class-F," *IEEE Microwave and Wireless Components Letters*, vol. 27, no. 9, pp. 830-832, Sep. 2017.

[8] J. Pang, S. He, C. Huang, Z. Dai, J. Peng, and F. You. "A Post-Matching Doherty Power Amplifier Employing Low-Order Impedance Inverters for Broadband Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 12, pp. 4061-4071, Dec. 2015.

[9] F. Meng, X. Zhu, J. Xia, and C. Yu, "A New Approach to Design a Broadband Doherty Power Amplifier via Dual-Transformation Real Frequency Technique," *IEEE Access*, vol. 6, pp. 48588-48599, 2018.

[10] H. Lyu and K. Chen. "Balanced-to-Doherty Mode-Reconfigurable Power Amplifier With High Efficiency and Linearity Against Load Mismatch," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 5, pp. 1717-1728, May 2020.

[11] M. Li, J. Pang, Y. Li and A. Zhu. "Bandwidth Enhancement of Doherty Power Amplifier Using Modified Load Modulation Network," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 6, pp. 1824-1834, Jun. 2020.

[12] J. Pang, Y. Li, C. Chu, J. Peng, X. Y. Zhou, and A. Zhu. "Extend High Efficiency Range of Doherty Power Amplifier by Modifying Characteristic Impedance of Transmission Lines in Load Modulation Network," *2020 IEEE/MTT-S International Microwave Symposium (IMS)*, Los Angeles, CA, USA, Jun. 2020, pp. 707-710.

[13] X. Fang, A. Chung, and S. Boumaiza, "Linearity-Enhanced Doherty Power Amplifier Using Output Combining Network With Predefined AM–PM Characteristics," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 1, pp. 195-204, Jan. 2019.

[14] P. Jia, F. You, and S. He, "A 1.8–3.4-GHz Bandwidth-Improved Reconfigurable Mode Doherty Power Amplifier Utilizing Switches," *IEEE Microwave and Wireless Components Letters*, vol. 30, no. 1, pp. 102-105, Jan. 2020.

[15] Y. Komatsuzaki, et al., "A High Efficiency 3.6-4.0 GHz Envelope-Tracking Power Amplifier Using GaN Soft-Switching Buck-Converter," *2018 IEEE/MTT-S International Microwave Symposium (IMS)*, Philadelphia, PA, USA, Jun. 2018, pp. 465-468.

[16] A. Banerjee, L. Ding, and R. Hezar, "A High Efficiency Multi-Mode Outphasing RF Power Amplifier With 31.6 dBm Peak Output Power in 45nm CMOS," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 3, pp. 815-828, Mar. 2020.

[17] J. J. Moreno Rubio, V. Camarchia, M. Pirola and R. Quaglia, "Design of an 87% Fractional Bandwidth Doherty Power Amplifier Supported by a Simplified Bandwidth Estimation Method," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 3, pp. 1319-1327, Mar. 2018.

[18] E. Aggrawal and K. Rawat, "Analysis and Design of Chireix Outphasing Switched Mode Power Amplifier," *2018 IEEE/MTT-S Latin America Microwave Conference (LAMC 2018)*, Arequipa, Peru, Dec. 2018, pp. 1-3.

[19] 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.

[20] P. H. Pednekar, E. Berry, and T. W. Barton. "RF-Input Load Modulated Balanced Amplifier With Octave Bandwidth", *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 12, pp. 5181-5191, Dec. 2017.

[21] J. Xia, M. Yang, Y. Guo, and A. Zhu. "A Broadband High-Efficiency Doherty Power Amplifier With Integrated Compensating Reactance," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 7, pp. 2014-2024, Jul. 2016.

[22] J. Pang, et al. "Analysis and Design of Highly Efficient Wideband RF-Input Sequential Load Modulated Balanced Power Amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 5, pp. 1741-1753, May 2020.

[23] Y. Li, et al. "Two-Port Network Theory-Based Design Method for Broadband Class J Doherty Amplifiers," *IEEE Access*, vol. 7, pp. 51028-51038, 2019.

[24] Y. Komatsuzaki, R. Ma, S. Sakata, K. Nakatani and S. Shinjo. "A Dual-Mode Bias Circuit Enabled GaN Doherty Amplifier Operating in 0.85-2.05GHz and 2.4-4.2GHz," *2020 IEEE/MTT-S International Microwave Symposium (IMS)*, Los Angeles, CA, USA, 2020, pp. 277-280.

[25] H. Lyu, Y. Cao, and K. Chen. "Linearity-Enhanced Quasi-Balanced Doherty Power Amplifier With Mismatch Resilience Through Series/Parallel Reconfiguration for Massive MIMO," *IEEE Transactions of Microwave Theory and Techniques*, vol. 69, no. 4, pp. 2319-2335, Apr. 2021.

[26] G. Lv, W. Chen, X. Liu, F. M. Ghannouchi and Z. Feng, "A Fully Integrated C-Band GaN MMIC Doherty Power Amplifier With High Efficiency and Compact Size for 5G Application," *IEEE Access*, vol. 7, pp. 71665-71674, 2019.

[27] S. Y. Zheng, Z. W. Liu, X. Y. Zhang, X. Y. Zhou, and W. S. Chan, "Design of Ultrawideband High Efficiency Extended Continuous Class-F Power Amplifier," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 6, pp. 4661-4669, Jun. 2018.

[28] L. Basenese, *Resonant (RESN): 5 Reasons to Let It Ride on 5G*, SmarterAnalyst,1 Mar. 2019. Accessed on 15 Jun. 2021. [Online]. Available: https://www.smarteranalyst.com/bloggers-corner/resonant-resn-5-reasons- to-let-it-ride-on-5g/

[29] C. Schroeder, *Exploring the Future of 5G*, Microwave Journal, 19 Nov. 2018. Accessed on 15 Jun. 2021. [Online]. Available: https://www.microwavejournal.com/blogs/23-5g-and-beyond-national-instruments/post/31371-exploring-the-future-of-5g

[30] Techplayon, *What is PAPR (Peak to average power ratio), Why it matters to Power Amplifier*, Techplayon, 8 Jun. 2017. Accessed on 15 Jun. 2021. [Online]. Available: https://www.techplayon.com/papr-peak-average-power-ratio-matters-power-amplifier/

[31] L. L. Tafuri, "Linearity and Efficiency Enhancement Techniques for Mobile Communication Power Amplifiers," PhD Thesis, Aalborg University, Aalborg, Denmark, 2014. Accessed on:

15 Jun. 2021. [Online]. Available: https://www.researchgate.net/publication/340686583_Linearity_and_Efficiency_Enhancement_Techniques_for_Mobile_Communication_Power_Amplifiers

[32] K. Fogarty, *GaN Versus Silicon for 5G*, Semiconductor Engineering, 15 Aug. 2019. Accessed on 13 May 2021. [Online]. Available: https://semiengineering.com/gan-versus-silicon-for-5g/

[33] R. Szweda, T. O'Connell, and M. Racanelli, "Experts see SiGe holding its own," *III-Vs Review*, vol. 17, no. 6, pp. 32-35, 2004.

[34] J. DeLisle, *Semiconductor Technology Offers Exciting Possibilities*, Microwaves & RF, 11 Feb. 2014. Accessed on 13 May 2021. [Online]. Available: https://www.mwrf.com/technologies/semiconductors/article/21845511/semiconductor-technology-offers-exciting-possibilities

[35] NXP, "NXP Goes with GaN," *Compound Semiconductor*, vol. 17, no. 6, pp. 23-24, Aug. 2011.

[36] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*, 2nd ed. Artech House Microwave Library, 2006.

[37] D. Y. Wu, "A Doherty Power Amplifier with Extended Bandwidth and Reconfigurable Back-off Level," PhD thesis, University of Waterloo, Waterloo, Canada, 2013. Accessed on: May 20, 2021. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/7385

[38] M. Liu, "Concurrent Dual-band Doherty Power Amplifiers for Carrier Aggregation," MASc thesis, University of Waterloo, Waterloo, Canada, 2017. Accessed on: May 21, 2021. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/12036

[39] Y. Li, "Two Port Network Theory Based Design Method of Broadband Doherty Power Amplifier," MASc thesis, University of Waterloo, Waterloo, Canada, 2018. Accessed on: May 22, 2021. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/13892

[40] Q. Tang, Y. Li, and W. Li, "Over Second Octave Power Amplifier Design Based on Resistive

Resistive Series of Continuous Class-F/F−1 Modes," *IEEE Microwave and Wireless Components Letters*, vol. 27, no. 5, pp. 494-496, May 2017.

[41] F. H. Raab, "Intermodulation distortion in Kahn-technique transmitters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, no. 12, pp. 2273-2278, Dec. 1996.

[42] F. Wang, A. H. Yang, D. F. Kimball, L. E. Larson, and P. M. Asbeck, "Design of wide-bandwidth envelope-tracking power amplifiers for OFDM applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 4, pp. 1244-1255, Apr. 2005.

[43] H. Chireix, "High Power Outphasing Modulation," *Proceedings of the Institute of Radio Engineers*, vol. 23, no. 11, pp. 1370-1392, Nov. 1935.

[44] Y. Cao and K. Chen, "Pseudo-Doherty Load-Modulated Balanced Amplifier With Wide Bandwidth and Extended Power Back-Off Range," *IEEE Transactions on Microwave Theory And Techniques*, vol. 68, no. 7, pp. 3172-3183, Jul. 2020.

[45] H. Golestaneh, "Broadband Doherty Power Amplifiers with Enhanced Linearity for Emerging Radio Transmitters," PhD thesis, University of Waterloo, Waterloo, Canada, 2016. Accessed on: Jun. 19, 2021. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/10200\

[46] 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.

[47] Wolfspeed, a Cree company, "CGHV1F006S 6 W, DC - 15 GHz, 40 V, GaN HEMT," CGHV1F006S datasheet, 2013 [Revised Dec. 2020].

[48] A. Barakat, M. Thian, V. Fusco, S. Bulja and L. Guan, "Toward a More Generalized Doherty Power Amplifier Design for Broadband Operation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 3, pp. 846-859, Mar. 2017.

[49] A. Bakarat, M. Thian, and V. Fusco. “Towards generalized Doherty power amplifier design for wideband multimode operation,” *2015 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*, Porto de Galinhas, Brazil, 2015, pp. 1-5.

[50] Wolfspeed, a Cree company, “CGHV1F025S 25 W, DC - 15 GHz, 40 V, GaN HEMT,” CGHV1F025S datasheet, 2014 [Revised Jul. 2020].

[51] F. Mkadem, M. C. Fares, S. Boumaiza, and J. Wood, “Complexity-reduced Volterra series model for power amplifier digital predistortion,” *Analog Integrated Circuits and Signal Processing*, vol. 79, no. 2, pp. 331-343, Feb. 2014.

[52] *5G NR Base Station (BS) radio transmission and reception*, 3GPP TS 38.104 version 16.4.0 Release 16, Jul. 2020.

[53] K. Hausmair, P. N. Landin, U. Gustavsson, C. Fager, and T. Eriksson, “Digital Predistortion for Multi-Antenna Transmitters Affected by Antenna Crosstalk,” IEEE Transactions on Microwave Theory and Techniques, vol. 66, no. 3, pp. 1524–1535, Mar. 2018.