A Novel Power-Scalable Wideband Power Amplifier Linearization Technique

by

Yushi Hu

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

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Thomas Johnson Associate Professor, School of Engineering, University of British Columbia
Supervisor:	Slim Boumaiza Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
Internal Members:	John Long Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
	David Nairn Associate Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
Internal-External Member:	Patricia Nieva Professor, Dept. of Mechanical and Mechatronics Engineering, University of Waterloo

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

Global mobile traffic is expected to continue to increase at an astonishing rate in the future, due to the ever-increasing number of mobile phone subscribers and the adoption of smart devices which generate significantly more mobile traffic. To satisfy this growth in demand, it is envisioned that future 5th Generation (5G) mobile networks will utilize lower powered small-cell base stations and base stations with large antenna arrays to greatly improve network coverage and capacity. A power amplifier (PA) is a critical component in a base station's transmitter, required to boost the signal power such that it is high enough for transmission to the intended receiver. The design of the PA for 5G base stations, however, presents new challenges to designers.

When driven with modern wideband communication signals, the PA must be both efficient and linear in order to minimize power consumption, improve reliability, maintain transmission accuracy, and avoid interference with neighbouring signals. In conventional high-powered macrocell base station designs, the aforementioned requirements are usually satisfied using a two-step procedure. First, the PA is designed using a Doherty power amplifier (DPA) topology, which has high efficiency, but poor linearity. Then, digital predistortion (DPD) linearization techniques are applied to ensure that the DPA attains the required linearity performance. However, for the lower-powered PAs needed in small cells and large antenna arrays, the relatively high power overhead of DPD techniques, which does not scale down as the power range of the PA decreases, make them unattractive PA linearization solutions.

In response, a new PA linearization technique is proposed and developed in this thesis. It is based on the design and addition of a linearization amplifier (LA), an approach which can help the PA attain the required linearity even when it is driven with modern communication signals with very wide bandwidths. Of particular note, the LA's power consumption is relatively low, it scales with the PA's power range, and it does not increase with signal bandwidth. These qualities make it highly suitable for use with PAs in future 5G small-cell base stations and base stations with large antenna arrays.

To validate the proposed technique's effectiveness, a prototype circuit was designed, fabricated and applied to a high peak efficiency 6 W class AB PA with a centre frequency of 850 MHz. When stimulated by a wideband 40 MHz signal, the PA's adjacent channel leakage ratio (ACLR) was improved by up to 13 dB after the addition of the LA. This enabled the PA to achieve an ACLR of about -45 dBc without the use of any other linearization techniques. Significant ACLR improvements were also observed for signals with even wider bandwidths of up to 160 MHz. Moreover, it was shown that the LA could

be used in conjunction with a simple predistorter to further improve the efficiency and linearity of the class AB PA.

Next, the LA is augmented with a conventional DPA design to form a new linear DPA topology that was able to achieve a better linearity-efficiency trade-off compared to the linearized class AB PA. To accomplish this, a study of the interactions between the LA and DPA circuitries was conducted and a design strategy was developed to determine the circuit parameters that maximized ACLR improvement while minimizing power consumption. For validation purposes, this strategy was applied to design a proof-of-concept prototype with a centre frequency of 800 MHz and a peak envelope power of 12 W. With the addition of the LA, a more than 11 dB improvement of the ACLR was obtained at the prototype's output when it was driven with signals with up to 40 MHz of modulation bandwidth: an ACLR of about -45 dBc or better was achieved over wide average power range. As expected, the efficiency of the linear DPA topology remained significantly higher than the linearized class AB PA for all signals tested.

Another challenge faced in particular by PAs in a large antenna array, is that it will experience dynamic load impedance variations due to antenna coupling. This unwanted variation in the load impedance can cause instability and significant distortions at the output of the PA that is difficult to remedy using conventional techniques. To address these issues, it is shown in the last part of this thesis that the LA can be used to mitigate this problem by minimizing the amount of load impedance variation seen by the PA due to antenna coupling, such that it remains closer to its optimal value, and by maintaining excellent linearization across a wide range of load impedance values.

Acknowledgements

I would like to say how grateful I am to Prof. Slim Boumaiza for giving me the opportunity to conduct research under his supervision during my time as a PhD student. His support, insights and guidance have helped me develop as a researcher during this exciting period of time. Aside from my supervisor, I would also like to thank my committee members, Prof. Johnson, Prof. Long, Prof. Nairn, and Prof. Nieva for reviewing my thesis and providing valuable feedback.

I am very grateful for the generous financial support I received through the Ontario Graduate Scholarship program and the University of Waterloo, and I would like to acknowledge the support of Wolfspeed Inc. and Keysight Inc. for donating the large-signal transistor models and Advanced Design System software, respectively.

Thanks should also go to Yujuan Zhang for providing me with the model of the antenna array used in Section 5.4, and I would like to thank Sarah Wallace, Dr. Hamed Golestaneh, and Dr. Xiao-Hu Fang for proofreading some of the sections of this thesis.

I have been blessed to have wonderful friends during my time as a PhD student, and they have made my life more rich and enjoyable in many ways. For example, their constant support, encouragements, and prayers have given me strength to persevere through the numerous failures I encountered; their willingness to discuss anything with me has helped me grow and laugh; and their constantly asking me when I will be finished with my PhD has helped me become more focused and organized. I look forward to many more years of friendship with all of them.

I was fortunate to be born to great parents who have made countless sacrifices for me. I can never thank them enough. My wife, Dr. Sharon Liang, has been patient with me even when I was stressed and became difficult to deal with. She has helped me become a better person, and I am forever grateful that I am able to share my life with her. I would also like to thank my son, Gabriel Hu, who is the cutest baby in the world. He has been a great source of joy and motivation to me. Last, but certainly not least, I would like to thank God for giving me peace, hope, and purpose throughout this time. None of this would be possible without Him.

Table of Contents

List of Tables **xi**

List of Figures **xii**

Acronyms **xviii**

1 Introduction **1**

1.1 Motivation 1

1.2 Thesis Objective 4

1.3 Thesis Outline 5

2 PA Efficiency Enhancement and Linearization Techniques **6**

2.1 PA Efficiency 6

2.1.1 Ideal Field Effect Transistor 7

2.1.2 Efficiency of Conventional PA Topologies 7

2.2 Back-off Efficiency Enhancement Techniques 10

2.2.1 Envelope Tracking 10

2.2.2 Outphasing 11

2.2.3 Doherty Amplifier 13

2.3 PA Distortions 15

2.3.1 Effects of PA Distortions on Communication Signals 16

2.3.2 Sources of Distortion 18
2.3.3 Additional Sources of Distortions in Back-Off Efficiency Enhancement Topologies 22
2.4 PA Linearization Techniques 25
2.4.1 Feedback 26
2.4.2 Feedforward 26
2.4.3 Predistortion 29
2.4.4 Derivative Superposition 31
2.4.5 Circuit Level Linearization Techniques 32
2.5 Summary 34

3 Power-Scalable Wideband PA Linearization Technique 36
3.1 Proposed PA Linearization Technique 36
3.1.1 Linearity Improvement Mechanisms for Intermodulation Distortions 38
3.1.2 Linearity Improvement for In-band Distortions 41
3.2 Linearization Amplifier Design 42
3.2.1 LA Realization Using Negative Feedback 42
3.2.2 Power Consumption of LA and Overall Efficiency 44
3.3 Main PA and LA Co-design 45
3.3.1 PA and LA design 46
3.3.2 Linearization Analysis Using a Two-tone Input 46
3.3.3 Linearization Under Modulated Signals 49
3.4 Measurement Results 49
3.5 Joint Linearization with RF Predistortion 53
3.6 Conclusion 57

4 Doherty Distortion Reduction Using a Linearization Amplifier 58

4.1 Proposed Topology . . . 58

4.1.1 Linearization Improvement by the Impedance Inverter . . . 59

4.1.2 Correction Current Reduction . . . 62

4.1.3 Comparison to Linear N-Way DPAs . . . 63

4.1.4 Effects of LA on Carrier and Peaking Transistors . . . 64

4.2 DPA and LA design and Simulation . . . 68

4.2.1 Linearization with CW Stimulus . . . 70

4.2.2 Linearization with Two-Tone Stimulus . . . 73

4.2.3 Linearization Under Modulated Signals . . . 74

4.3 Measurements . . . 75

4.4 Conclusion . . . 80

5 Desensitizing the PA to the Effects of Antenna Coupling 82

5.1 Antenna Coupling in Massive MIMO and Its Effects on the PA's Load Impedance . . . 82

5.2 Analysis of Load Impedance Variation with LA . . . 84

5.3 Linearization of Massive MIMO PAs in the Presence of Antenna Coupling Using the LA . . . 88

5.4 Simulated Performance with a Four-Element Array Antenna . . . 90

5.4.1 Performance with Various Input Signal Phases . . . 91

5.4.2 Performance with Various Input Signal Magnitudes . . . 93

5.5 Conclusion . . . 95

6 Conclusion 96

6.1 Summary of Contributions . . . 97

6.2 Future Work . . . 98

6.3 List of Relevant Publications . . . 99

References **100**

APPENDICES **108**

A IMD3 Improvement Breakdown **109**

List of Tables

1.1 Small cells and macrocell base station range and power levels [11] 3

2.1 Linear DPA performance 35

3.1 Measured PA performance with and without LA at 850 MHz 52

4.1 Maximum Linearity Improvement with the LA 78

4.2 DPA Performance When -45 dBc ACLR was Achieved 79

4.3 Comparison with Other Linear DPAs 80

4.4 Performance When -45 dBc ACLR is achieved 80

A.1 Calculated and simulated IMD3 reduction with tones at 840 and 860 MHz and an output power of 30 dBm. 111

List of Figures

1.1 Illustration of (a) small cells in HetNet and (b) large antenna arrays used in a transmitter in a massive MIMO system. 2

2.1 Illustration of PA efficiency. 6

2.2 Ideal FET model characteristic. 7

2.3 A (a) simplified class A PA schematic and its (b) gate and drain waveforms with a sinusoidal input. V_k is assumed to be zero to further simplify the analysis. 8

2.4 Class A DE with the PDF of a 20 MHz WCDMA signal with 7.6 dB PAPR. 9

2.5 A (a) simplified class B PA schematic and its (b) gate and drain waveforms with a sinusoidal input. V_k is assumed to be zero to further simplify the analysis. 9

2.6 Simplified ET system block diagram. 11

2.7 Simplified outphasing system. 12

2.8 Simplified DPA topology and its ideal performances. 14

2.9 Illustration of a) 16-QAM constellation, b) AM-AM and AM-PM of a linear (dashed line) and nonlinear (solid line) PA, and c) error vector. 16

2.10 Expected output spectrum distortions with a two-tone input. 17

2.11 Illustration of (a) ACLR and (b) dynamic distortions. 18

2.12 Simplified PA schematic for quasistatic distortion analysis. 18

2.13 Frequency domain equivalent circuit of the PA. 19

2.14 Simulated (a) I_{ds} versus V_{gs} of a 6 W GaN transistor and its (b) large-signal short-circuit transconductance. 20

2.15 Simulated (a) I_{ds} versus V_{ds} of a $6\,W$ GaN transistor and (b) an illustration of distortion due to the knee region. 21

2.16 Illustration of the (a) generation and (b) effects of unwanted low frequency drain modulation. 22

2.17 Variation of the (a) phase and (b) magnitude of DPA gain due to frequency dispersion of the quarter-wave transformer. 24

2.18 Drain voltage of the carrier transistor for various (a) phase and (b) gain offsets between the peaking and carrier transistor current. ($\Delta\phi$ is how much $\angle I_p$ deviates from its ideal value.) 25

2.19 Illustration of negative feedback. 25

2.20 Feedforward amplifier configuration. 27

2.21 Predistortion system configuration. 28

2.22 Simplified schematic of a cubic predistorter. 29

2.23 DPD based transmitter architecture. 30

2.24 Simplified schematic of a DS amplifier. 32

2.25 The a) circuit diagram and b) AM-PM of the DPA proposed in [66]. Reprinted from [66]. 33

2.26 Graphical representation of the method proposed in [68]. Reproduced from [68]. 34

3.1 Simplified topology of the LA and PA. 37

3.2 (a) The PA and its Norton equivalent with the (b) Norton equivalent of the PA and LA. 38

3.3 Expected output frequency spectrum with: (a) an output power of $P_o\,$dBm with PA only, (b) output power of $(P_o - G)\,$dBm with PA only, and (c) output power of $P_o\,$dBm with PA and LA. 39

3.4 Desired values for $\alpha(2f_1 - f_2)$ and $\beta^{-1}(2f_1 - f_2)$ are highlighted in (a) and (b), respectively. 41

3.5 LA realized through negative feedback and its Norton equivalent 43

3.6 LA schematic with negative feedback 43

3.7 Complete schematic of the PA and LA 47

3.8 Simulated small-signal transconductance of the LA 47
3.9 Simulated small-signal output impedances of the (a) LA and (b) PA. . . . 48
3.10 Simulated a) $\beta^{-1}(f)$ profile and b) IMD3 of the PA with a two-tone input signal with tones 1 and 2 at 840 and 860 MHz, respectively 48
3.11 Simulated and normalized (a) magnitude and (b) phase of the gain of the PA with a two-tone input signal with tones 1 and 2 at 840 MHz and 860 MHz, respectively 49
3.12 Simulated PA output spectrum with a 40 MHz bandwidth modulated signal 50
3.13 Photograph of a) fabricated PA with LA and b) measurement setup 50
3.14 Measured PA output frequency spectrum with various modulated signals. . 51
3.15 Topology of joint linearization 53
3.16 RF predistorter schematic and calculated performance 54
3.17 Diagram of measurement setup 54
3.18 Measured performance of the different linearization configurations with a 15 MHz WCDMA signal 55
3.19 Measured performance of the different linearization configurations with a 40 MHz WCDMA signal 56

4.1 Block diagram of the linear DPA topology. 59
4.2 Equivalent circuit schematic of the DPA (a) without and (b) with the LA. 59
4.3 Calculated a) $|\kappa(f_{ob})|$ for various $Z_{DPAt}(f_{ob})$ values with $Z_{LA}(f_{ob}) = R_L/3$, $I_{LA}(f_{ob}) = 0.01\ A$, $R_L = 20\,\Omega$, and $V_L(f_{ob}) = 1V = I_{DPAt}(f_{ob})[Z_{DPAt}(f_{ob})||R_L]$ and b) LA correction current for various $Z_{DPAt}(f_i)$ values with $V_L(f_i) - V_L^{'}(f_i) = 1\ V$ and $R_L = 20\ \Omega$. 61
4.4 $\kappa(f_{ob})$ versus (a) $I_{LA}(f_{ob})/I_{DPAt}(f_{ob})$ and (b) phase of $I_{LA}(f_{ob}) - I_{DPAt}(f_{ob})$ (with their magnitudes being equal). 63
4.5 Equivalent circuit schematic of the DPA (a) without and (b) with the LA. 65
4.6 Calculated $|V_C^{'}(f_i)|/V_{MAX}$ with various magnitudes of $I_{LA}(f_i)$. 65
4.7 Calculated (a) $|V_C^{'}(f_i)|/V_{MAX}$ and (b) $\Re[V_C^{'}(f_i)]/V_{MAX}$ with $Z_{LA}(f_i) = R_L/3$, phase of $I_C(f_i)$ set to 0°, and various phases of $I_{LA}(f_i)$. 66

4.8 Calculated (a) $|V_C^{'}(f_i)|/V_{MAX}$ and (b) $\Re[V_C^{'}(f_i)]/V_{MAX}$ with $|Z_{LA}(f_i)| = |R_L/3|$, phase of $Z_{LA}(f_i)$ set to 30°, phase of $I_C(f_i)$ set to 0°, and various phases of $I_{LA}(f_i)$. 66

4.9 Simplified schematic of the proposed linear DPA topology. 67

4.10 Simulated $\kappa_{denom}(f)^2$ versus frequency. 68

4.11 Schematic of the fabricated linear DPA. 68

4.12 Simulated phase of V_L(800 MHz) and $V_L^{'}$(800 MHz) for different t_d values. . 70

4.13 Simulated (a) PAE, of the linear DPA, with the LA's power consumption factored in, for various t_d values, and the (b) phase of LA's open circuit voltage and V_L(800 MHz) for different t_d values. 71

4.14 Simulated (a) real part of $V_C^{'}$(800 MHz) and (b) magnitude of the correction current from the LA for different t_d values. 71

4.15 Simulated (a) $P_{out}/P_{avs\text{-}DPA}$, where $P_{avs\text{-}DPA}$ is the power available at the input of the DPA, the (b) DPA's PAE, with the LA's power consumption factored in, for various G_o settings, and (c) the magnitude of $V_C^{'}$(800 MHz) for various G_o settings. 72

4.16 Simulated (a) DE of the DPA (without the LA consumption), DE_{DPA}, versus the output power of the DPA only (without the power from the LA), $P_{out-DPA}$ and (b) magnitude of the correction current from the LA for different G_o values. 72

4.17 Simulated IMD3 with (solid line) and without (dashed line) the LA with input tones at 795 and 805 MHz 73

4.18 Simulated (a) $P_{out}/P_{avs\text{-}DPA}$ and (b) phase deviation of the output voltage for the case with and without the LA and t_d and G_o of 2.56 ns and 0.9, respectively. 74

4.19 Simulated DPA (a) output spectrum and (b) output ACLR versus frequency when the DPA is driven by a 15 MHz signal with 8.3 dB PAPR. 75

4.20 The (a) picture of fabricated fully analog linear DPA and (b) measured DPA output spectrums when stimulated by a 15 MHz modulation bandwidth signal with 8.3 dB PAPR and 32 dBm output power. 75

4.21 Measured and co-simulated ACLR when the DPA is stimulated by a 15 MHz modulation bandwidth signal with 8.3 dB PAPR for the case with (solid line) and without (dashed line) the LA. 76

4.22 Measured DPA output spectrums when stimulated by a (a) 20 MHz, (b) 30 MHz and (c) 40 MHz WCDMA signal at the average output power level were -45 dBc of ACLR is achieved. 76

4.23 Measured (a) ACLR, (b) EVM, and c) PAE of the DPA with (solid line) and without (dashed line) the LA versus average output power. Note that the PAE figures with LA include the power consumption of the LA. 77

4.24 Measured (a) gain and (b) phase distortion when stimulated by a 15 MHz WCDMA signal for the case with (blue) and without (magenta) the LA. . 79

5.1 Illustration of a) digital beamforming reproduced from [8], where different colors represent different signals, and b) an antenna array with multiple PAs reproduced from [73]. 83

5.2 Equivalent circuit of the PA and LA used for load impedance variation analysis. 84

5.3 Calculated Z_L' (solid line) and Z_L (dashed line) for various $VSWR_L$ values. The impedances are plotted on a Smith chart that is normalized to Z_{Lo}. . . 86

5.4 The (a) topology of the DPA and LA and the (b) simulated carrier transistor impedances for various $VSWR_{load}$ values for the case with (solid line) and without (dashed line) the LA. The impedances are plotted on a Smith chart that is normalized to 50 Ω. 86

5.5 Simulated impedance seen by the (a) carrier and (b) peaking transistor for the case with (solid line) and without (dashed line) the LA and $|\Gamma_{load}| = 0.2$. 87

5.6 Simulated impedance seen by the carrier and peaking transistor for the case with (solid line) and without (dashed line) the LA and $|\Gamma_{load}| = 0.5$. 87

5.7 Calculated $|\kappa_{denom}(800MHz)|^2$ for various $VSWR_L$ values. 88

5.8 Simulated AM-AM, AM-PM, and IMD3s across Z_{load} for the case with (solid line) and without (dashed line) the LA. 89

5.9 Simulated AM-AM, AM-PM, and IMD3s across Z_{load} for the case with (solid line) and without (dashed line) the LA. 90

5.10 Illustration of the simulation setup used to observe the performance of an 2×2 array antenna at a far-field point. 91

5.11 Simulated AM-AM, AM-PM, and IMD3s at the far-field port for the case with (solid line) and without (dashed line) the LA for various beam angles. 92

5.12 Impedance seen by the carrier transistor in the array antenna PAs for the case with and without the LA as θ is swept from zero to 30° in steps of 10°. 93

5.13 Impedance seen by the carrier transistor in the PA1 for the case with (circle) and without (triangle) the LA for various $P_{out1-bo}$ and $\angle V_{in1}$ values. 93

5.14 Simulated AM-AM, AM-PM, and IMD3s at the output of PA1 for the case with (solid line) and without (dashed line) the LA for various V_{in1} magnitudes. $\angle V_{in1}$ was set to 180° for this test. 94

A.1 The a) simulated IMD3L and IMD3U with tones 1 and 2 at 840 MHz and 860 MHz, respectively, for the case with (solide line) and without (dashed line) the LA and filter, and an b) illustration of the setup used for the simulation. 110

A.2 Simulation setup used to find the PA and LA's Norton currents at the IMD3 frequencies 110

A.3 Simulated IMD3L and IMD3U with tones 1 and 2 at 840 MHz and 860 MHz, respectively, for the case with (solid line) and without (dashed line) the LA. Also plotted is the calculated IMD3L and IMD3U (short dashes). 111

Acronyms

5G 5th Generation.

ACLR adjacent channel leakage ratio.

ADS Advanced Design System.

AM-AM amplitude-to-amplitude.

AM-PM amplitude-to-phase.

AWG arbitrary waveform generator.

CMOS complimentary metal-oxide-semiconductor.

CW continuous wave.

DC direct current.

DE drain efficiency.

DPA Doherty power amplifier.

DPD digital predistortion.

DS derivative superposition.

EA envelope amplifier.

EM electromagnetic.

ET envelope tracking.

EVM error vector magnitude.

FET field-effect transistor.

FFW Feedforward.

GaN gallium nitride.

IC integrated circuit.

IMD intermodulation distortions.

IMD2 second-order intermodulation distortion.

IMD3 third-order intermodulation distortion.

IMD3L lower third-order intermodulation distortion.

IMD3U upper third-order intermodulation distortion.

IMD5 fifth-order intermodulation distortion.

LA linearization amplifier.

LTE Long-Term Evolution.

MIMO multiple-input, multiple-output.

MMIC monolithic microwave integrated circuit.

mmWave millimeter wave.

PA power amplifier.

PAE power added efficiency.

PAPR peak-to-average power ratios.

PDF probability distribution function.

QAM quadrature amplitude modulation.

RF radio frequency.

RFPD RF predistortion.

TOR transmitter observation receiver.

VSWR voltage standing wave ratio.

WCDMA Wideband Code Division Multiple Access.

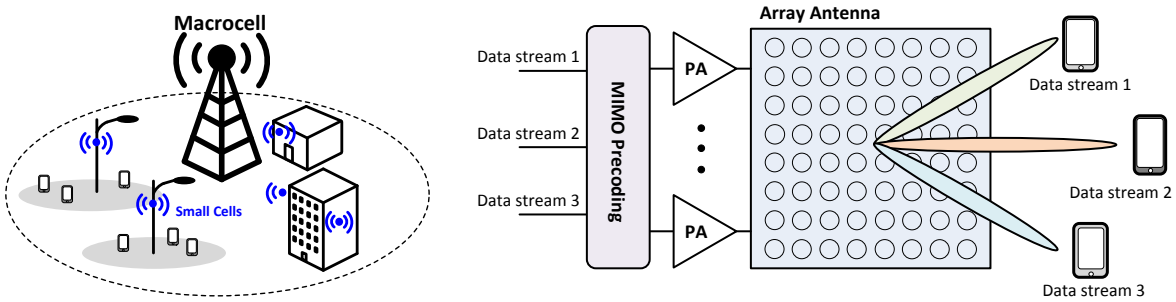
Chapter 1

Introduction

1.1 Motivation

Currently, there are more than 5 billion unique mobile phone subscribers in the world. This means that more than two-thirds of the world's population now has a mobile subscription and can access a mobile network. It is easy to see why mobile networks have become such an integral part of our lives. For example, they have allowed us to stay connected with each other and access the internet across a very wide geographic area. In the future, it is expected that the number of mobile subscribers will continue to increase, reaching 5.9 billion by the year 2025 [1]. At the same time, more and more mobile users are adopting smart devices like smartphones, which generate significantly more mobile traffic compared to basic-featured cell phones because they allow the users to enjoy a variety of bandwidth-hungry mobile applications such as video streaming. For instance, an average smartphone in 2018 uses about 5.6 GB of data per month, and it is estimated that this number will increase to 21 GB per month by the end of 2024 due to increased video viewing time, more embedded video content, and higher video resolutions with more demanding formats (e.g., augmented/virtual reality) [2]. Both the rising number of smartphone users and the increasing data usage by these smart devices have driven global mobile traffic to grow at an astonishing rate. According to a report by Ericsson, global mobile data traffic in Q3 2018 grew close to 79 percent year-on-year, and is expected to be five times higher, reaching 136 exabytes per month by the end of 2024 [2].

In response, mobile networks have had to evolve rapidly to accommodate this explosive growth in data traffic. At present, mobile networks are moving towards a heterogeneous solution, termed HetNet where small-cell base stations (small cells) are deployed alongside



(a) (b)

Figure 1.1: Illustration of (a) small cells in HetNet and (b) large antenna arrays used in a transmitter in a massive MIMO system.

traditional macrocell base stations to improve network coverage and capacity [3]. Compared to macrocell base stations, small cells are smaller, shorter ranged, lower powered, less expensive, and can be more easily installed. This allows them to be sited in densely populated urban areas to improve network capacity (more cells covering fewer users), and in areas not covered by the macrocell base stations (such as certain rural or indoor areas) to improve network coverage. Another advantage of using small cells is that the average distance between the user and the base station is now shorter. This helps to improve both the data rate and energy efficiency [4].

The addition of small cells alone, however, is not enough to satisfy the orders of magnitude increase in mobile data traffic predicted for the future. Hence, it has been envisioned that 5th Generation (5G) systems will make use of massive MIMO technology at sub-6-GHz frequencies to enable each base station to serve multiple, spatially separated users using the same time and frequency resource. By doing this, the capacity of the network will be increased without requiring additional base stations or spectrum [5]. Massive MIMO technology relies on the use of a large number of antennas in an antenna array to accomplish spatial multiplexing where each spatial channel carries independent information as illustrated in Fig. 1.1(b). It is also anticipated that 5G systems will expand into the underutilized millimeter wave (mmWave) frequencies where a large amount of spectrum is still available [6]. This will allow service providers to significantly increase their channel bandwidth, translating to higher data transfer rates. A key enabler for mmWave transmission is the use of massive MIMO systems with large antenna arrays to focus the transmitted signal into a narrow beam aimed at the user. This could effectively increase the gain of the transmitting and receiving antennas and partially overcome the higher path-loss at

Table 1.1: Small cells and macrocell base station range and power levels [11]

Cell Type	Cell Radius (km)	Output Power (W)
Pico Cell	0.1 to 0.2	0.25 to 1
Micro Cell	0.2 to 2	1 to 10
Macro Cell	8 to 30	10 to >50

mmWave frequencies [7]. Note that, compared to sub-6-GHz massive MIMO systems, it is anticipated that mmWave massive MIMO systems will use analog/hybrid beamforming instead of full digital beamforming. This means mmWave massive MIMO will be less flexible and will have lower user multiplexing capability [8].

Notice from Fig. 1.1(b) that in a large antenna array, each antenna needs to be driven by a power amplifier (PA). The PA is a critical component in a base station's transmitter because it is used to boost the power of the communication signal such that it is high enough for transmission to the receiver. The PA must operate efficiently, as they are the largest energy consumer in the transmitter chain. Furthermore, they should operate linearly in order to minimize signal distortion and interference with signals in neighbouring frequency bands. To achieve both these goals, a two-step procedure is usually taken. Firstly, the PA is designed using topologies or classes of operation which optimize peak power and back-off power efficiency at the cost of linearity. It is important for the PA to have high efficiency at power back-off because modern communication signals have high peak-to-average power ratios (PAPR) [9]. As a result, the PA is expected to operate in back-off power the majority of the time. Then, linearization techniques are introduced to ensure that the PA attains the required linearity performance. Predistortion techniques, in particular digital predistortion (DPD) variations [10], have become the preferred linearization approach for high powered macrocell base station PAs. These approaches have benefited from advancements in digital signal processors and allow for excellent linearization.

The recent trend towards small cells and large antenna arrays, however, means that the PAs will now be lower powered than those in conventional macrocell base stations. This could render existing DPD techniques unattractive due to the relatively high power overheads that do not scale down with power range. To illustrate this point, the typical cell radius and PA power of several different categories of small-cell and macrocell base stations are given in Table 1.1. The base station's output power can vary from tens and hundreds of watts for a macrocell down to a few watts for a pico or micro cell. Now, a typical DPD system can require a few watts of power consumption [12]. So, it would not significantly decrease the efficiency of a high powered macrocell PA because its power consumption is

small compared to the macrocell's output power. However, for low powered PAs the power consumption of DPD becomes relatively large and would significantly decrease the PA's overall efficiency, giving a poor trade-off between linearity and efficiency. This problem is exacerbated when we consider using DPD to linearize large antenna array systems. Firstly, the power range of the PAs in a large antenna array is further reduced. The greater the number of antennas, the lower the power required from each PA. Secondly, the cost of DPD would be multiplied by the number of PAs in the system if each PA is linearized using DPD. Lastly, DPD techniques for large antenna array systems must be even more complex because, in addition to the nonlinearity of each PA, the coupling between the antennas add another source of distortion in the antenna array system that must be mitigated by the DPD in order to attain the required linearity.

1.2 Thesis Objective

In light of the fact that conventional linearization techniques used for macrocell PAs are no longer attractive for PAs in small cells and large antenna arrays, the objective of this thesis is to develop a PA linearization technique which is more suitable for these emerging applications. The following steps will be taken towards the achievement of this goal:

1. Develop a novel linearization technique which possesses the following attributes:
    - Its power consumption should be relatively low and *scalable* compared to the PA's output power. In other words, the power consumption of the linearization circuitry should decrease as the power range of the PA decreases. This ensures that the overall efficiency of the PA is not degraded when the power range of the PA is scaled down.
    - It should enable the PA to meet the linearity requirements of modern communication standards when driven with wideband signals.
2. Verify the new linearization technique by developing a design strategy and applying it to a conventional class AB PA as a proof of concept.
3. Develop a design strategy to apply the proposed linearization technique to a high back-off efficiency Doherty power amplifier (DPA) topology to further improve the PA's linearity-efficiency trade-off.
4. Investigate the proposed technique's capacity to counter the unwanted effects of coupling between the antennas in a large antenna array system.

1.3 Thesis Outline

Chapter 2 shows that conventional PA designs will have low efficiency when driven with modern communication signals, highlighting the need for PA topologies with high efficiency at power back-off. An overview of some popular back-off efficiency enhancement topologies are given here. The effects and sources of PA distortions are also briefly discussed, to emphasize the importance of PA linearization techniques, and an overview of existing PA linearization techniques is provided.

Chapter 3 introduces a novel PA linearization technique suitable for low powered PAs that is based on the design and addition of a linearization amplifier (LA). The technique is analyzed in detail and used to linearize a class AB PA. The design of a proof of concept prototype is also presented, and measurement results of the prototype are shown to validate the technique's effectiveness. In addition, a joint-linearization setup is created by using the LA in conjunction with a simple predistorter, and measurement results demonstrate that the joint-linearization configuration is able to further improve linearity and reduce power consumption.

In Chapter 4, the LA is augmented with a high back-off efficiency DPA topology to form a new linear DPA topology. A study of the interaction between the LA and the DPA is detailed here, as well as the design method developed from the study. The design of a proof-of-concept prototype is also presented, and measurement results show that the new linear DPA topology is capable of achieving high efficiency and linearity when stimulated by wideband modulated signals.

Next, in Chapter 5, the LA is studied for its capacity to remedy the unwanted dynamic load impedance variations caused by antenna coupling in large array antennas. Initial simulation results are provided to validate the analysis.

Finally, Chapter 6 concludes the thesis by summarizing the contributions of this work and discussing possible directions for future research.

Chapter 2

PA Efficiency Enhancement and Linearization Techniques

2.1 PA Efficiency

The power efficiency of the PA characterizes how well it converts the power supplied to it into useful output signal power. This is a critical design parameter because the PA is one of the largest power consumers in the transmitter, and a low efficiency PA can consume an excessive amount of power. Furthermore, the wasted energy is converted into heat which can reduce the PA's reliability and require additional cooling. One common measure of power efficiency is drain efficiency (DE), defined as the ratio between the desired output signal power P_{out} to the input DC power P_{DC}. Another measure of power efficiency, power added efficiency (PAE), also takes into consideration the input signal power P_{in}. The two measures of efficiency are given in Fig. 2.1. For the analysis in this section it will be assumed that the PA's gain is high such that $P_{out} \approx P_{out} - P_{in}$. Hence, the two efficiency measures converge, and DE alone can be used to characterize efficiency.

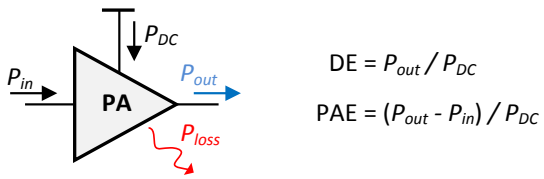
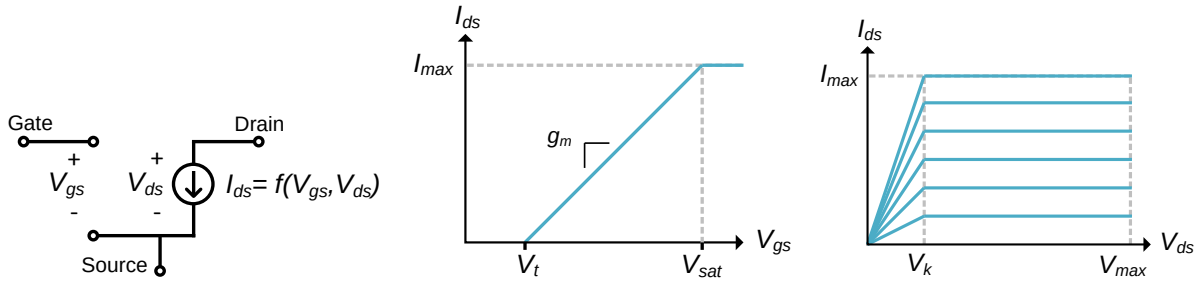


Figure 2.1: Illustration of PA efficiency.



(a) Ideal FET model (b) I_{ds} versus V_{gs} ($V_{ds} > V_k$) (c) I_{ds} versus V_{ds}

Figure 2.2: Ideal FET model characteristic.

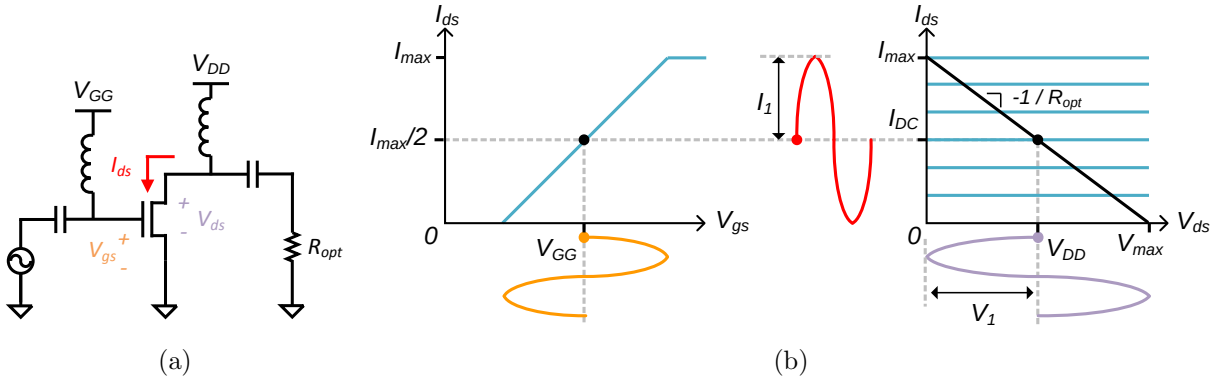
2.1.1 Ideal Field Effect Transistor

The field-effect transistor (FET) is the most common transistor used in base station PA designs, and the efficiency analysis in this section will use, for simplicity and clarity, the ideal FET model shown in Fig. 2.2(a) [13]. The relationship between the gate source voltage (V_{gs}) and drain source current (I_{ds}) is given in Fig. 2.2(b) for a drain source voltage (V_{ds}) greater than the knee voltage V_k. As can be seen, when $V_t < V_{gs} < V_{sat}$, where V_t is the threshold voltage, the ideal transistor behaves as a voltage-controlled current source with a transconductance of g_m. Unlike small-signal amplifiers, the PA is designed to utilize the maximum current and voltage swing of a transistor. Hence, the saturation current (I_{max}), breakdown voltage (V_{max}), as well as V_k are specified. For efficiency analysis, V_k will be assumed to be zero to further simplify the analysis.

In reality, the transistor will include parasitic elements, g_m is not constant, V_k is not zero, and the input and output impedances are not infinite. These nonidealities will be become more critical when we look at the sources of PA distortions later on in Section 2.3.2.

2.1.2 Efficiency of Conventional PA Topologies

Class A One of the simplest PA topologies is the class A PA. The schematic of the class A PA and its waveforms with a sinusoidal input are shown in Fig. 2.3. Notice that the gate (V_{GG}) and drain (V_{DD}) bias are set such that the direct current (DC) current and voltage at the drain are I_{max} and $V_{max}/2$, respectively, and the load resistance R_{opt} is set to V_{max}/I_{max} to ensure maximum V_{ds} swing and, therefore, maximum possible output power from the transistor. The DC power consumption (P_{DC-A}) and output power (P_{out-A}) of



(a) (b)

Figure 2.3: A (a) simplified class A PA schematic and its (b) gate and drain waveforms with a sinusoidal input. V_k is assumed to be zero to further simplify the analysis.

the class A PA can be found to be $V_{max}I_{max}/4$ and $V_1I_1/2$, respectively, where V_1 and I_1 are the amplitudes of V_{ds} and I_{ds}, respectively, at the sine wave frequency. Then, the DE of the class A PA (DE_A) can be derived to be

$$DE_A = \frac{P_{out-A}}{P_{DC-A}} = \frac{2V_1I_1}{V_{max}I_{max}}. \tag{2.1}$$

The efficiency of the class A PA is highest at maximum output power ($V_1 = V_{max}/2$, $I_1 = I_{max}/2$, and $DE_A = 50\,\%$). At power back-off, the DE of the class A PA will decrease linearly with output power. The reason why DE drops so quickly is because P_{DC-A} stays constant regardless of its output power. In fact, P_{DC-A} stays constant at $V_{max}I_{max}/4$ even when there is no output power. The DE of a class A PA is next plotted versus power back-off in Fig. 2.4 as well as the probability distribution function (PDF) of a 20 MHz Wideband Code Division Multiple Access (WCDMA) signal with 7.6 dB PAPR. Notice how the majority of the signal is at large power back-off, where the efficiency of the PA is low, and the expected DE can be found to be 9.6 % when this WCDMA signal is transmitted through a class A PA. This means that for a mere 1 W of output power more than 10 W of DC power is required by the PA, and more than 9 W of power is wasted.

Class B The gate bias of the PA can be reduced to improve the PA's peak efficiency performance, leading to more efficient PA operation classes such as the Class B PA. A simplified schematic of a class B PA and its waveforms with a sinusoidal input are shown in Fig. 2.5. V_{GG} is now set such that the DC value of I_{ds} is zero when there is no signal input, and as a result, the transistor will only conduct for half of the signal period. This is

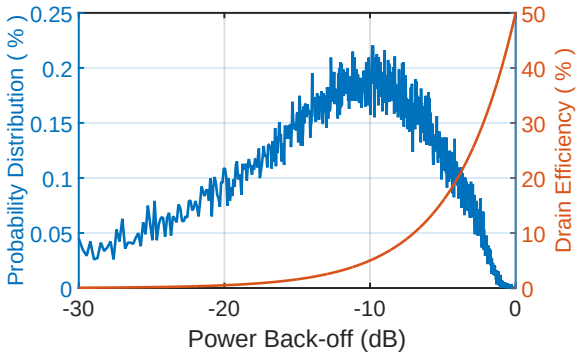
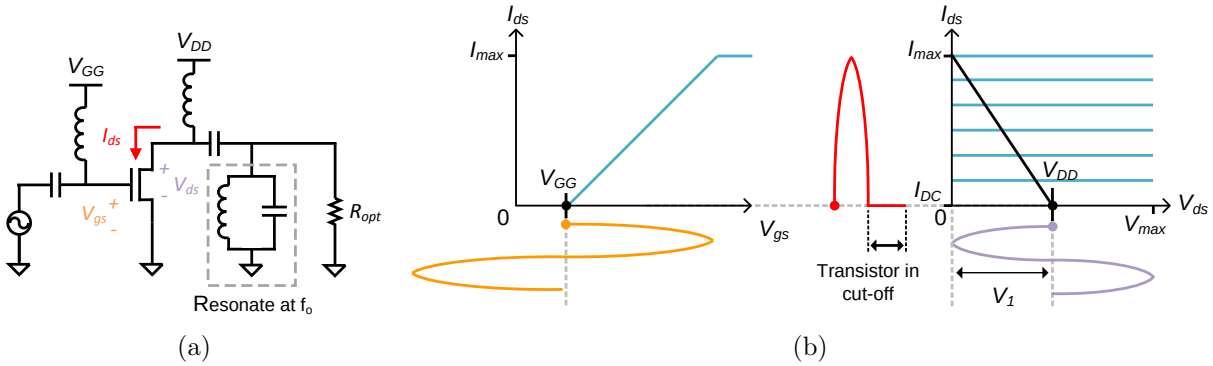


Figure 2.4: Class A DE with the PDF of a 20 MHz WCDMA signal with 7.6 dB PAPR.



(a) (b)

Figure 2.5: A (a) simplified class B PA schematic and its (b) gate and drain waveforms with a sinusoidal input. V_k is assumed to be zero to further simplify the analysis.

what enables the class B PA to attain a higher efficiency, but I_{ds} will now have significant components at harmonic frequencies which should be shorted using, for example, a tank circuit with a resonant frequency at the fundamental frequency f_o.

The DE of the class B PA can be found through Fourier analysis of the I_{ds} waveform. The relationship between the amplitude of the fundamental frequency and DC component of I_{ds} (I_{DC}) can be found to be $I_1/I_{DC} = \pi/2$, and the DE of the class B PA can be expressed as

$$DE_B = \frac{(V_1 I_1)/2}{V_{DC} I_{DC}} = \frac{\pi}{4}\frac{V_1}{V_{DC}} = \frac{\pi}{4}\frac{V_1}{V_{max}/2}. \tag{2.2}$$

At peak output power, $V_1 = V_{max}/2$, and the DE of the class B PA is 78.5 %. This is a significant improvement from the class A PA (50 %). Moreover, notice that the class B

PA's DE does not decrease linearly with power as does the class A PA. Rather, it decreases linearly with V_1, which is proportional to the square root of the output power. If we use a class B PA to amplify the WCDMA signal (whose PDF was shown in Fig. 2.4), we can expect a DE of about 38.4 %. This means that the class B PA would consume four times less power than the class A PA.

Class AB and class C PAs also have a lower gate bias, and as a result, better efficiency than the class A PA. The class C PA is even more efficient than the class B PA, but it has the poorest linearity, gain, and maximum output power. The class AB PA, on the other hand, has a higher gain than the class B PA, but it has a lower efficiency. See [14] for a more detailed discussion and comparison between these classes of operation.

2.2 Back-off Efficiency Enhancement Techniques

We saw in the previous subsection that the DE of the PA at power back-off can be improved by going to different classes of operation (for example, from class A to class B). The back-off efficiency of these conventional PAs can be further improved using the back-off efficiency enhancement techniques we will discuss next.

2.2.1 Envelope Tracking

Notice from Eq. (2.2) that if V_{DC} can be made to decrease with V_1, such that they are always equal, then DE of a class B PA can be maintained at 78.5 % at all output powers. This is the basic idea behind drain modulation techniques where the drain bias of the PA's transistor is modulated with the envelope of the signal to improve power back-off efficiency. The envelope tracking (ET) technique is one of the more popular drain modulation techniques, and a block diagram of the ET system is given in Fig. 2.6. Drain modulation in the ET system is achieved through an envelope amplifier (EA) which modulates the drain supply voltage of the PA according to the signal's magnitude. The EA topology is usually based on a linear and efficient hybrid topology which comprises of a parallel linear voltage source and a highly efficient switched current source [15], and extensive research has been done to further increase its efficiency. For example, by using dual-switch supply, up to 80 % efficiency has been achieved by the EA when transmitting WCDMA signals [16]. By increasing the EA efficiency, the efficiency of the overall ET system is increased as well since the efficiency of the ET system is given by $\eta_{ET} = \eta_{EA}\eta_{PA}$, where η_{EA} and η_{PA} is the efficiency of the EA and PA, respectively, in the ET system.

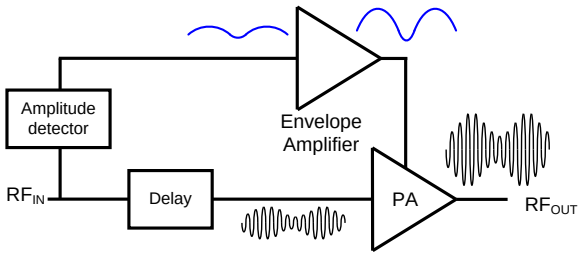


Figure 2.6: Simplified ET system block diagram.

Excellent efficiency performance has been achieved by ET systems. For instance, the system in [17] attained about 60 % efficiency with a WCDMA signal.

An advantage of ET is that it is agnostic to the carrier frequency of the PA, and the radio frequency (RF) bandwidth of the ET system is only limited by the RF bandwidth of the PA employed. Hence, if broadband PA topologies such as class J [18] or push pull [19] were used to design the PA then the ET system can achieve high-efficiency performance across an extensive frequency range. The instantaneous modulation bandwidth of a base station ET system, however, is limited by the slew rate or bandwidth of the EA. The envelope of communication signal using IQ modulation is given by $\sqrt{I^2 + Q^2}$, where I and Q are the in-phase and quadrature components, respectively. The bandwidth of the envelope signal will be 3-4 times larger than the signal due to the nonlinear transformation, and this high bandwidth requirement can degrade the efficiency and accuracy of the EA [20]. Fortunately, the EA output does not have to faithfully follow the actual envelope in an ET system. Hence, envelope shaping functions have been proposed and used to reduce the bandwidth of the envelope signal [20–22]. This can simplify the EA design and improve its efficiency, but nevertheless, with the modulation bandwidth of communication signals set to dramatically increase in the future the difficulty of designing an efficient EA with very high bandwidth can severely limit the application of ET systems.

2.2.2 Outphasing

The principle of outphasing modulation, which was proposed in the 1930's [23], can be illustrated using the simplified outphasing system shown in Fig. 2.7. If an amplitude and phase modulated signal is given to the input, such that $S_{in}(t) = A(t)cos(\omega t + \phi(t))$, then

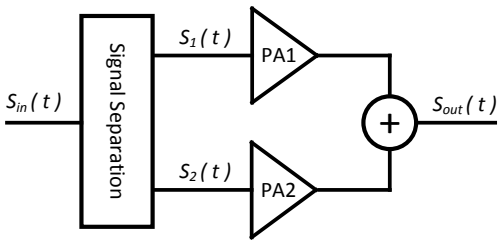


Figure 2.7: Simplified outphasing system.

the PA inputs are fixed amplitude signals which can be written as follows:

$$\begin{aligned} S_1(t) &= cos(\omega t + \phi(t) + cos^{-1}(A(t))) \\ S_2(t) &= cos(\omega t + \phi(t) - cos^{-1}(A(t))) \end{aligned} \tag{2.3}$$

If the PAs have a gain of G_{pa}, then the output voltage is equal to

$$S_{out}(t) = G_{pa}(S_1(t) + S_2(t)) = 2G_{pa}A(t)cos(\omega t + \phi(t)). \tag{2.4}$$

Notice how $A(t)$ is effectively encoded into input signals as a differential phase shift, and then recovered at the output using the combiner. For example, the maximum output amplitude is achieved when $S_1(t)$ and $S_2(t)$ are in phase, and the output amplitude is zero when they are antiphase. As a result, the input and output of each PA only needs to be phase modulated, and highly efficient nonlinear PAs, like a saturated class B PA, can now be used, ideally, without compromising the output's linearity.

Either an isolated or non-isolated power combiner can be utilized to combine the power of the two PAs in the outphasing system. With a non-isolated combiner, however, the load impedance seen by the PAs will become more reactive as the two outputs become more out of phase. Although no real power is dissipated by the reactive load, the PA that is used to supply reactive power to the reactive load will consume a certain amount of real power. Thus, the outphasing PA's efficiency will be lower at low power levels. This problem can be somewhat remedied by adding a shunt reactance to tune out the reactance seen by the PA for a given amplitude, but this will only maximize the efficiency around that particular amplitude. Nevertheless, through careful selection of the shunt reactance the average efficiency can be improved by as much as a factor of two compared to a class B amplifier [24]. This approach does, however, compromise the linearity of the outphasing system due to the impedance mismatch between the two PAs [25].

Alternatively, isolated combiners, like the Wilkinson power divider or a hybrid combiner, can be used instead which ensures that the both PAs see a constant resistive load

at all signal levels. While this improves the linearity of the outphasing system, both PAs will now deliver full power at all output power levels. As a result, efficiency will rapidly decrease at power back-off [26]. Work has been done to try to remedy this problem by replacing the load resistor at the difference port of a 180° hybrid combiner with a converter that can reuse some of the wasted power at power back-off by returning it to the power supply of the PAs [27]. While this technique can improve the PA's efficiency, one of the undesirable consequences is that the impedance of the converter is a function of input power, and can only be matched at one power level. This, in turn, can affect the isolation of the combiner and generate unwanted distortions.

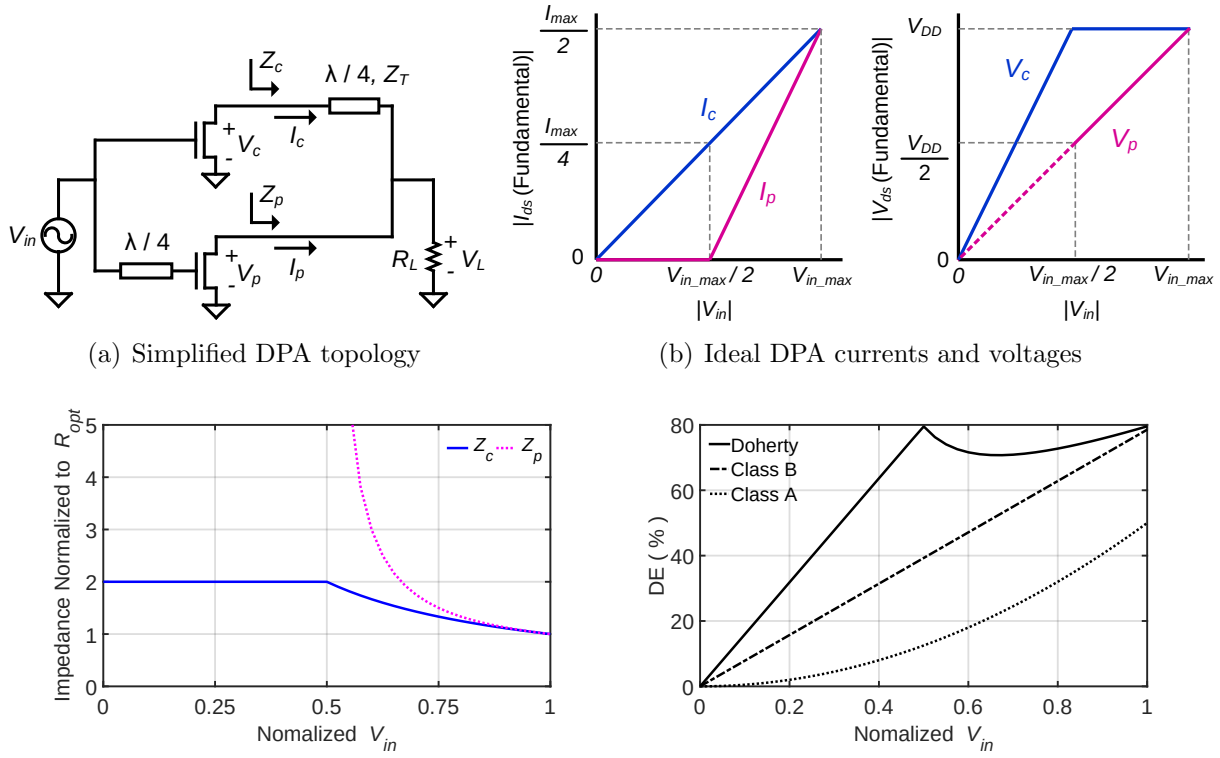
Lastly, a couple of other caveats of the outphasing system should be mentioned here. First the amplitude of PA's input signals are constant versus output power in a conventional outphasing system. Hence, the gain of the outphasing amplifier is reduced at power back-off, and this can significantly decrease the overall PAE of the outphasing system if the gain of the PAs are not high enough [28]. Secondly, the bandwidths of the input signals are expanded due to the inverse cosine operation, and will stretch into the adjacent channels. Any mismatches in the phase and gain of the two PAs means that these out-of-band components will not be perfectly cancelled, and will thus introduce new distortions at the output. In most practical situations, additional calibration and a feedback loop are required to compensate for these mismatches [29].

2.2.3 Doherty Amplifier

While drain modulation improves back-off efficiency by decreasing the drain bias, load modulation techniques, on the other hand, improves back-off efficiency through increasing the load resistance, and thus V_1, at lower output powers. This, as can be seen from Eq. (2.2), can also increase the class B PA's efficiency. The DPA topology is an example of a PA topology which utilizes load modulation to enhance its power back-off efficiency [30], and a simplified DPA topology is shown in Fig. 2.8(a). The impedance seen by the carrier transistor, Z_c, can be derived to be

$$Z_c = \frac{Z_T^2}{R_L} - Z_T \frac{I_p}{I_c}, \tag{2.5}$$

where Z_T is the characteristic impedance of the quarter-wave transformer, and I_p and I_c are the magnitude of the fundamental currents from the peaking and carrier transistors, respectively. Notice that we can control Z_c by controlling the value I_p. The higher I_p is, the lower the Z_c is.



(a) Simplified DPA topology

(b) Ideal DPA currents and voltages

(c) Ideal DPA impedance

(d) Ideal DPA efficiency

Figure 2.8: Simplified DPA topology and its ideal performances.

Next, the ideal DPA voltages and currents are plotted in Fig. 2.8(b), and the ideal impedance seen by the carrier and peaking transistor, Z_p, are plotted in Fig. 2.8(c). At low output power levels, when $|V_{in}|$ is below $V_{in\_max}/2$, the peaking transistor is off, and Z_c stays constant at a value of $2R_{opt}$. At higher output powers the peaking transistor turns on, and I_p will cause Z_c to decrease according to Eq. (2.5). The purpose of decreasing Z_c is to keep the fundamental voltage at the drain of the carrier transistor V_c at a constant value of V_{DD}. If V_c increases beyond V_{DD} the transistor will become strongly saturated[1], and the output will be heavily distorted as a result.

The DE of the DPA is calculated and plotted in Fig. 2.8(d) for the case when both transistors are biased for class B operation. When V_{in} is less than $V_{in\_max}/2$, the peaking amplifier is off and the DE of the DPA increases linearly from zero to 78.5 %: Z_c is double that of a conventional class B PA in this operation region ($2R_{opt}$ instead of R_{opt}). Hence,

[1]In practice, V_c should be less than V_{DD} to avoid strong saturation due to the non-zero value of V_k.

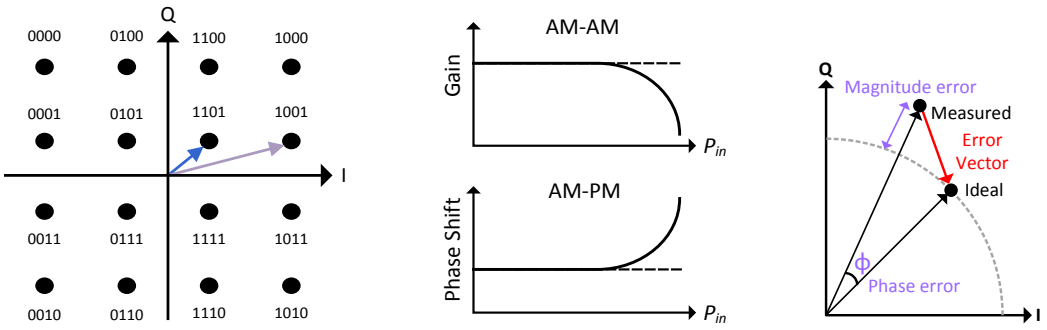
the amplitude of the fundamental voltage at the drain of the carrier transistor (V_1) is also doubled. As a result, the carrier transistor's DE is two times higher than the maximum DE of a class B amplifier when $V_{in} = V_{in\_max}/2$ according to Eq. (2.2). When V_{in} is higher than $V_{in\_max}/2$, the DE of the DPA becomes a function of both the carrier and peaking amplifier's power consumption. While the DE of the carrier transistor stays constant at 78.5 %, the efficiency of the peaking transistor increases from 39.25 to 78.5 % as V_{in} increases from $V_{in\_max}/2$ to V_{in_max}. The DE of a class B and class A PA is also plotted on Fig. 2.8(d) for comparison. Notice that at 6 dB power back-off ($V_{in} = V_{in\_max}/2$), the DE of the DPA is 2 and 6.28 times higher than the DE of the class B and class A PA, respectively.

The RF bandwidth of conventional DPAs is limited by its use of a narrowband quarter-wave line. However, recent works on three-way DPAs [31], digital DPAs [32,33], and DPAs with modified design methods have significantly increased its bandwidth capability [34–37]. For example, three-way DPAs can achieve over 55 % efficiency at an output back-off of up to 9 dB over a wide frequency range from 0.73-0.98 GHz [31]. In another example, the authors in [34] utilized a wideband impedance inverter to push the DPA's fractional bandwidth to be greater than 50 % (470-803 MHz).

2.3 PA Distortions

To improve spectral efficiency, modern communication signals use modulation schemes like quadrature amplitude modulation (QAM) which also carries information in the signal amplitude. For example, the constellation diagram of a 16-QAM modulation scheme is given in Fig. 2.9(a). Notice how the symbols representing the bit stream 1101 and 1001 have different magnitudes. As a result, the modulated signal will have a varying amplitude, and this has led to the requirement for PAs which are highly linear versus input signal amplitude such that it maintains the accuracy of the transmitted information and does not generate interference for neighbouring signals.

Unfortunately, PAs have sources of distortions which prevent them from achieving the required linearity. Hence, linearization techniques are necessary in achieving the required PA linearity. In the following subsections, the effects and sources of PA distortion as well as existing linearization techniques used to remedy PA distortions are discussed in brief.



(a) (b) (c)

Figure 2.9: Illustration of a) 16-QAM constellation, b) AM-AM and AM-PM of a linear (dashed line) and nonlinear (solid line) PA, and c) error vector.

2.3.1 Effects of PA Distortions on Communication Signals

A PA can have quasi-static as well as dynamic distortions (commonly referred to as memory effects). The Quasi-static distortions of a PA are captured by its steady-state performance, and can be measured using a continuous wave (CW) signal. A perfectly linear PA should have an output voltage amplitude which is a scalar multiple of the input signal. If the input signal is given by $A\cos(\omega t)$. Then the linear output can be written as $KA\cos(\omega t + \phi)$, where K and ϕ are the constant gain and phase shift of the linear amplifier, respectively. In a nonlinear amplifier, the output is now given as $K(A)A\cos(\omega t + \phi(A))$, where the gain $K(A)$ and phase shift $\phi(A)$ changes versus input signal amplitude, leading to AM-AM and AM-PM distortions, respectively, as shown in Fig. 2.9(b).

Both AM-AM and AM-PM distortions will cause constellation points to deviate from their ideal values, thus reducing signal transmission quality. The amount of deviation caused by the PA distortions can be characterized by the error vector magnitude (EVM), which is defined in percentage as

$$EVM(\%) = 100\%\sqrt{P_{error}/P_{ref}}, \tag{2.6}$$

where P_{error} and P_{ref} are the root-mean-square power of the error vector and ideal signal vector, respectively. The error vector is illustrated in Fig. 2.9(c) for one symbol. The higher EVM is, the further away the measured constellation point is from its ideal location.

Quasi-static distortions will also cause the frequency spectrum of the signal to expand and generate new frequency components. This can be seen through a two-tone signal test

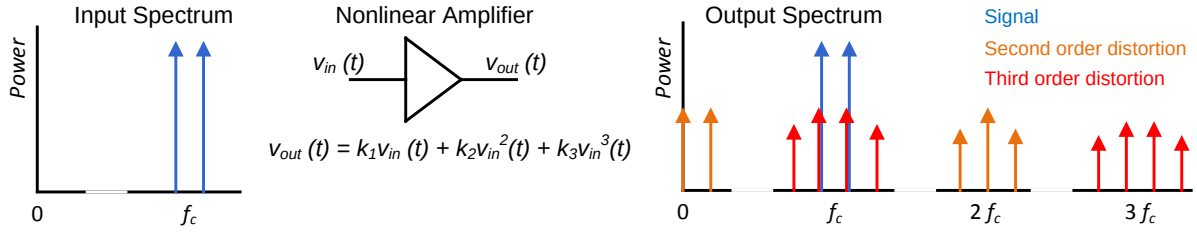


Figure 2.10: Expected output spectrum distortions with a two-tone input.

with the following input: $v_{in}(t) = A\,cos(\omega_1 t) + A\,cos(\omega_2 t)$, where A is the magnitude of both tones, and ω_1 and ω_2 are the frequency of tone 1 and 2, respectively. For a weakly nonlinear PA with AM-AM distortion, its output can be approximately modeled as

$$v_{out}(t) = k_1 v_{in}(t) + k_2 v_{in}(t)^2 + k_3 v_{in}(t)^3 \ldots, \tag{2.7}$$

where $v_{out}(t)$ and $v_{in}(t)$ are the output and input voltages of the PA, respectively, at time t, k_1 is the small signal gain, and k_2 and k_3 are the coefficient for the second and third-order nonlinear terms, respectively. Substituting the two-tone input into Eq. (2.7), it can be shown that new frequency components are generated at a frequency of $m\omega_1 + n\omega_2$, where $m + n$ is the order of the distortion [38] (e.g., the third-order term will generate distortions at ω_1, ω_2, $2\omega_2 \pm \omega_1$, $2\omega_1 \pm \omega_1$, $3\omega_1$, and $3\omega_2$). Fig. 2.10 shows the expected output spectrum with up to third-order distortions. In addition to harmonic distortions ($2\omega_1$, $2\omega_2$, $3\omega_1$, and $3\omega_2$), intermodulation distortions (IMD) are generated which are a combination of two or more frequencies. The most troublesome IMDs are the third-order intermodulation distortion (IMD3) at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ because they are usually the largest IMDs and are difficult to filter out because they are close in frequency to the output signal. With a modulated signal, the spreading of the signal into adjacent frequency bands can be characterized by the adjacent channel leakage ratio (ACLR), which gives the ratio between the leakage power in the adjacent channel to the power in the main signal bandwidth. ACLR in dB is given by

$$ACLR = P(B_1) - P(B_2), \tag{2.8}$$

where $P(B_1)$ and $P(B_2)$ is the total power in dBm for the main signal bandwidth B_1 and adjacent channel B_2, respectively, as illustrated in Fig. 2.11(a). The smaller the value of ACLR, the smaller the amount of power leakage, and the less the transmission will interfere with a neighbouring signal.

In addition to quasi-static distortions, the PA also has dynamic distortions which are not only a function of the current input, but also of previous inputs. In the frequency

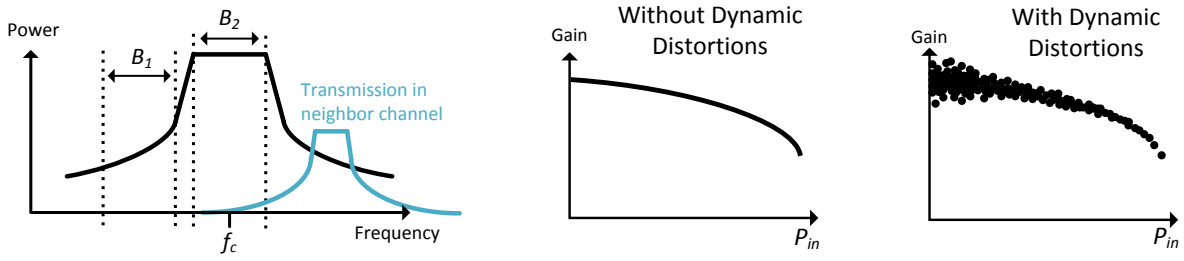


Figure 2.11: Illustration of (a) ACLR and (b) dynamic distortions.

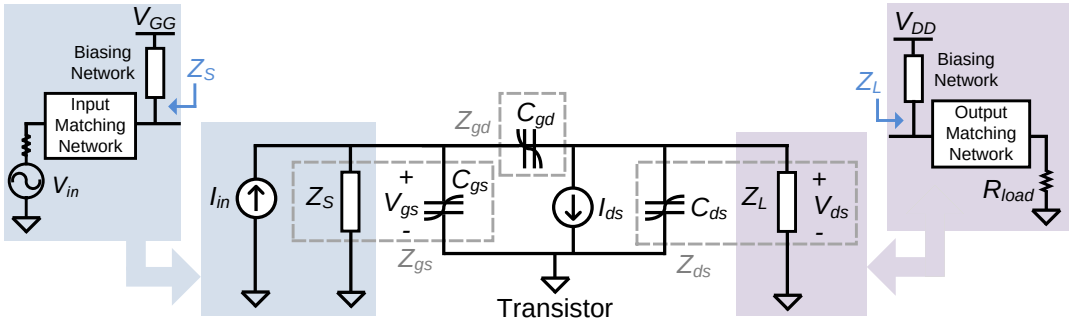


Figure 2.12: Simplified PA schematic for quasistatic distortion analysis.

domain, dynamic distortions can modify the magnitudes and phases of the intermodulation distortions. IMD imbalance for example is one of the consequences of dynamic distortions [39]. In the time domain, dynamic distortions will cause the AM-AM and AM-PM of the PA to demonstrate hysteresis effects where, for a given input, there can be many possible outputs depending on what the previous inputs are as illustrated in Fig. 2.11(b).

2.3.2 Sources of Distortion

In this section, a brief overview of some of the significant sources of PA distortion is given to highlight the need for PA linearization technique. First, to study sources of quasi-static distortions, a simplified PA with a large-signal transistor model is drawn in Fig. 2.12. This circuit will be analyzed in the frequency domain, and an equivalent frequency domain circuit is given in Fig. 2.13 for a specific operation point where $G_m(f) = \frac{I_{ds}(f)}{V_{gs}(f)}$, and $I_{ds}(f)$ and $V_{gs}(f)$ is the drain source current and gate source voltage of the transistor, respectively, at frequency f. Since $I_{in}(f)$ is a linear function of the input voltage, and the transistor's

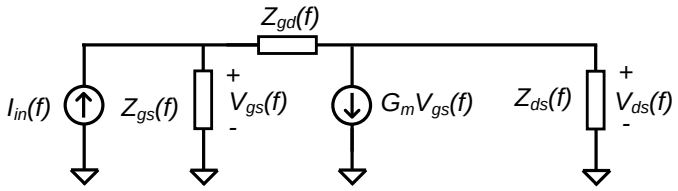


Figure 2.13: Frequency domain equivalent circuit of the PA.

drain source voltage $V_{ds}(f)$ is a linear function of the output voltage across the load, the linearity of the PA can thus be analyzed through the relationship $V_{ds}(f)/I_{in}(f)$. It can be found that the magnitude and phase of $V_{ds}(f)/I_{in}(f)$ can be expressed as

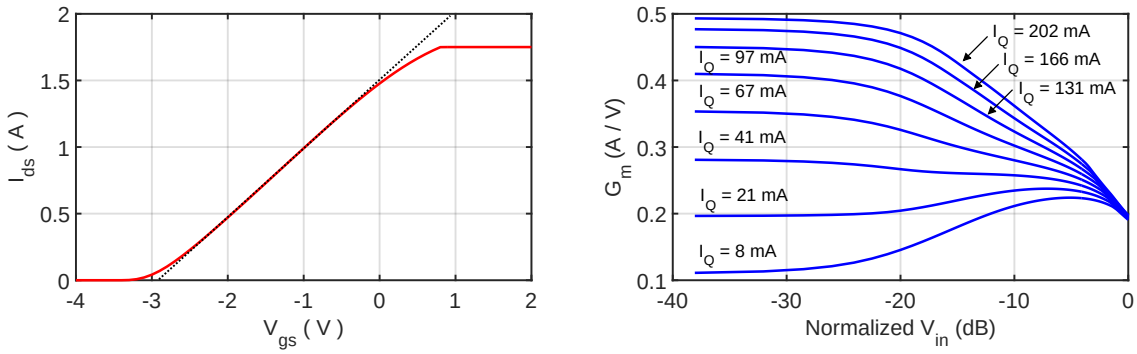
$$\left|\frac{V_{ds}(f)}{I_{in}(f)}\right| \approx \frac{|Z_{gs}(f)||Z_{ds}(f)||Z_{gd}(f)||G_m|}{|Z_{ds}(f)+Z_{gd}(f)+Z_{gs}(f)+Z_{gs}(f)Z_{ds}(f)G_m|}, \tag{2.9}$$

and

$$\angle\left(\frac{V_{ds}(f)}{I_{in}(f)}\right) \approx 180^o + \angle Z_{gs}(f) + \angle Z_{ds}(f) + \angle Z_{gd}(f) + \angle G_m - \angle(Z_{ds}(f)+Z_{gd}(f)+Z_{gs}(f)+Z_{gs}(f)Z_{ds}(f)G_m), \tag{2.10}$$

respectively, where $Z_{gs}(f)$, $Z_{gd}(f)$, and $Z_{ds}(f)$ are the impedance from the gate to source, gate to drain, and drain to source, respectively, at frequency f. The PA is linear if the magnitude and phase of $V_{ds}(f)/I_{in}(f)$ remains constant versus input power. However, PAs will have sources of distortion which changes the magnitude and phase of $V_{ds}(f)/I_{in}(f)$ versus input power. As few of these sources are briefly discussed next and a 6 W gallium nitride (GaN) transistor (CGH40006) will be used as an example.

Nonlinear Transconductance The simulated drain source current I_{ds} of the 6 W transistor is plotted in Fig. 2.14(a) versus V_{gs}. Note from this graph that, unlike the ideal FET behavior shown in Fig. 2.2(b), I_{ds} increases gradually at first when V_{gs} is just above V_t and it gradually saturates near its maximum value. These nonidealities causes the large-signal short-circuit transconductance, G_m, of a real transistor to be more nonlinear versus input power. To demonstrate this, the G_m of the 6 W transistor is simulated and plotted in Fig. 2.14(b) for different quiescent currents (I_Q). Notice how G_m varies with input power, and the amount of variation is highly dependent on I_Q. Recall from equations (2.9) and (2.10) that a variation of G_m can lead to both AM-AM and AM-PM nonlinearity.

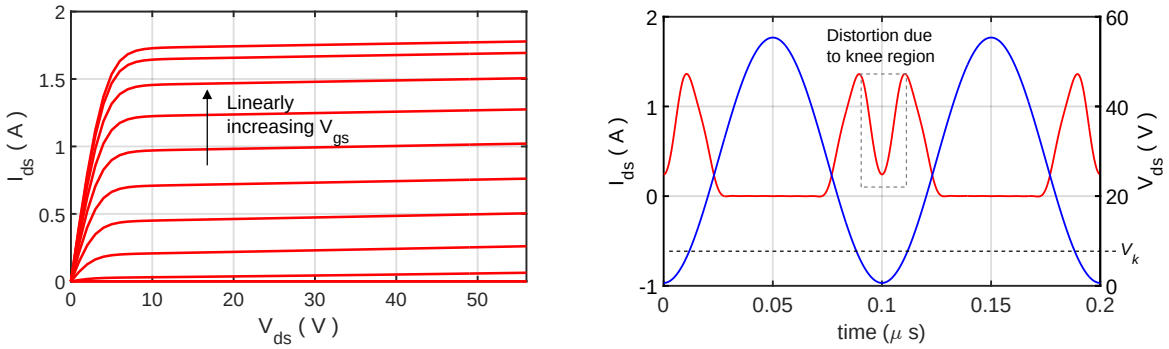


(a) (b)

Figure 2.14: Simulated (a) I_{ds} versus V_{gs} of a 6 W GaN transistor and its (b) large-signal short-circuit transconductance.

Knee Voltage Realistic transistors will have a non-zero knee voltage V_k, and in the I_{ds} versus V_{ds} plot shown in Fig. 2.15(a) we can see that I_{ds} drops rapidly when V_{ds} goes below V_k. To demonstrate how this can cause of distortions, the 6 W GaN transistor was loaded with a 50 Ω output impedance at the fundamental frequency while the harmonic frequencies are short-circuited. Next, the transistor was driven by a sinusoidal input, and $i_{ds}(t)$ and $v_{ds}(t)$ are plotted in Fig. 2.15. Notice how $i_{ds}(t)$ dips when $v_{ds}(t)$ is pushed below V_k. This dip in the $i_{ds}(t)$ waveform will cause a steep decrease in the magnitude of the gain, and thus AM-AM distortions. Going below V_k can also cause AM-PM distortions depending on the harmonic terminations [40].

Non-linear Capacitance The parasitic gate source and drain source capacitance of GaN transistors are strongly nonlinear versus the voltage across them. As a result, they can cause both AM-AM and AM-PM distortions in the PA. For example, the gate source capacitance C_{gs} affects both the magnitude and phase of $Z_{gs}(f)$, and a nonlinear C_{gs} can lead to a change in both the magnitude and phase of $Z_{gs}(f)$ versus input voltage. This, in turn, can cause both AM-AM and AM-PM distortions as can be seen from equations (2.9) and (2.10). Note that it is possible to minimize the effects that C_{gs} nonlinearity has on the AM-AM and AM-PM by carefully biasing the transistor and shorting the second harmonic impedance at the gate. This is, however, difficult to achieve in practice over a wide bandwidth [41].

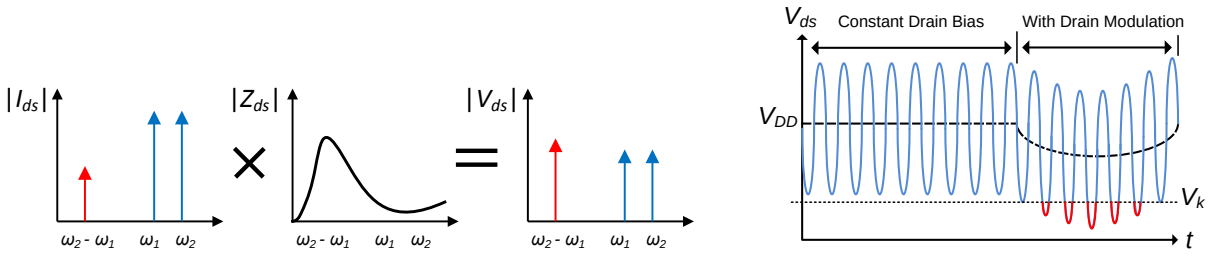


(a) (b)

Figure 2.15: Simulated (a) I_{ds} versus V_{ds} of a $6\,W$ GaN transistor and (b) an illustration of distortion due to the knee region.

Baseband Drain Impedance This source of distortion can be illustrated using a two-tone test. When the PA's driven by two tones at ω_1 and ω_2, the PA's nonlinear transistor's intrinsic current source generates a low frequency second-order intermodulation distortion (IMD2) current at $\omega_2 - \omega_1$. This current then creates an unwanted IMD2 drain voltage when it is driven through the impedance seen by the intrinsic current source $Z_{ds}(\omega_2 - \omega_1)$ as illustrated in Fig. 2.16(a). This unwanted voltage will add to the PA's distortions in a few different ways. Firstly, if the output capacitance is nonlinear, then this IMD2 voltage will modulate the value of the output capacitance, leading to dynamic AM-AM and AM-PM distortions. Secondly, this IMD2 voltage can also be fed back to the input where it can modulate the nonlinear input capacitor as well. Finally, and perhapes most detrimentally, this IMD2 voltage is superimposed onto the RF signal, and as a result the transistor's drain voltage can be sporadically pushed below the knee voltage as illustrated in Fig. 2.16(b). These sporatic knee region intrusions can cause strong dynamic distortions that greatly reduce the linearity and linearizabilty of the PA [42].

Thermal and Trapping Effects Thermal and trapping effects are device level effects that can cause dynamic distortions in the PA. Thermal effects occur when the temperature of the transistor is not constant over time. Since the gain of the PA is dependent on the temperature, this would result in the PA having time dependent behaviours. Moreover, the transistor device can have trapping effect, where the capture and release of carriers can cause unwanted current collapse and drain-current transients [43]. Both thermal and trapping effects become negligible, however, when the thermal and charge trapping time



(a) (b)

Figure 2.16: Illustration of the (a) generation and (b) effects of unwanted low frequency drain modulation.

constants is much larger than the timescale of the signal envelope [44].

Frequency Dispersive Gain and Group Delay Variation of the PA's gain and group delay can also cause dynamic distortions. This means that when the PA is driven by a wideband signal, the gain and group delay experienced by each frequency component will be different, and the output signal will be distorted as a result. This source of dynamic distortion can be minimized by carefully designing the PA's matching network to desensitize the PA's AM-PM to frequency and power, and to maintain constant group delay versus power and frequency within the modulated signal bandwidth [41].

2.3.3 Additional Sources of Distortions in Back-Off Efficiency Enhancement Topologies

Aside from the aforementioned sources of distortion, back-off efficiency enhancement topologies can introduce additional sources of distortion which can further degrade output linearity. In particular, we will focus on the distortions in the DPA because it seems to be the most promising topology for future base stations.

Load Modulation As was shown in Fig. 2.8(c), the impedance seen by the drain of the carrier transistor in the DPA changes versus input power after the peaking amplifier is turned on. This load modulation ensures that the carrier transistor is not pushed further into saturation at high power, and is therefore critical to the DPA's linear operation. Recently, however, authors in [45] have shown that this change of load impedance is actually

a source of AM-PM distortion in DPAs. To analyze this, they used a simplified model to show that the phase of the output voltage depends on the impedance seen by the drain of the main transistor. This can also be seen using Eq. (2.10). The phase of the PA's output depends on the magnitude and phase of Z_{ds}, which is varied versus input signal in the DPA. Since the variation of this impedance versus input voltage is inherent in the DPA operation, some AM-PM distortion is DPA is consequently unavoidable. It should also be noted here that the DPA's AM-PM distortions can be reduced by purposely mismatching the source impedance [45]. This, however, comes at the cost of lower PA gain.

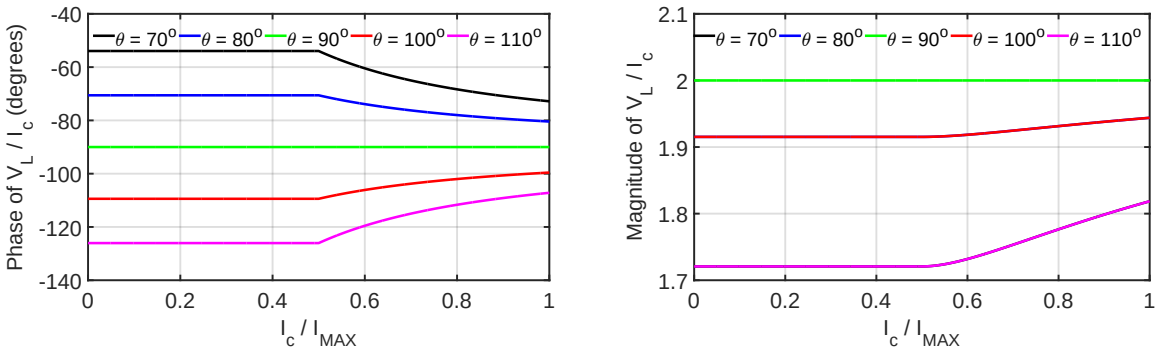
Quarter-Wave Transformer The DPA relies on a quarter-wave transformer to transform the carrier transistor into a voltage source at the combining node. As a result, the DPA is ideally linear because the output voltage remains linearly proportional to the current of the carrier transistor I_c, which is ideally linearly proportional to the input voltage. This quarter-wave transformer, however, only acts like an ideal impedance inverter at its design frequency, and we will next look at what happens when the DPA linearity when the frequency of operation deviates from the quarter-wave transformer's design frequency by studying the ratio between the output voltage V_L to the carrier transistor current I_c as shown in Fig. 2.8(a). Note that I_c is ideally proportional to the input voltage. Hence, if V_L/I_c is linear, then the DPA is linear. At low power, the peaking amplifier is off, and V_L/I_c can be expressed as

$$\frac{V_L}{I_c} = \frac{1}{j(1/Z_T)sin(\theta) + cos(\theta)}, \tag{2.11}$$

where θ is the electrical length of the quarter-wave transformer. Note that, no matter what value θ is, V_L/I_c is constant versus I_c. Hence, the DPA is ideally linear at low power. When the peaking transistor turns on, the current of the peaking transistor I_p can be expressed as a function of I_c as $I_p = -2j(I_c - 0.5I_{MAX})$ where I_{MAX} is the maximum value of I_c. Then, the ratio between V_L and I_c when the peaking transistor is on can be expressed as

$$\frac{V_L}{I_c} = R_L\left[\frac{1 - sin(\theta)(R_L/Z_T)(2 - I_{MAX}/I_c)}{j(R_L/Z_T)sin(\theta) + cos(\theta)} - j(2 - I_{MAX}/I_c)\right]. \tag{2.12}$$

At the design frequency of the quarter-wave transformer, $\theta = 90°$, and equations (2.11) and (2.12) both simplifies to $V_L/I_c = -jZ_T$. Hence, the DPA stays linear for its entire operation region. This is not the case, however, for other frequencies when θ is no longer $90°$. V_L/I_c is plotted for different θ value in Fig. 2.18. Notice how the magnitude and phase of V_L/I_c only remains constant at the quarter-wave transformer's design frequency, where $\theta = 90°$. Otherwise, the DPA will have both AM-AM and AM-PM distortions due to the change in electrical length of the quarter-wave line versus frequency.



(a) (b)

Figure 2.17: Variation of the (a) phase and (b) magnitude of DPA gain due to frequency dispersion of the quarter-wave transformer.

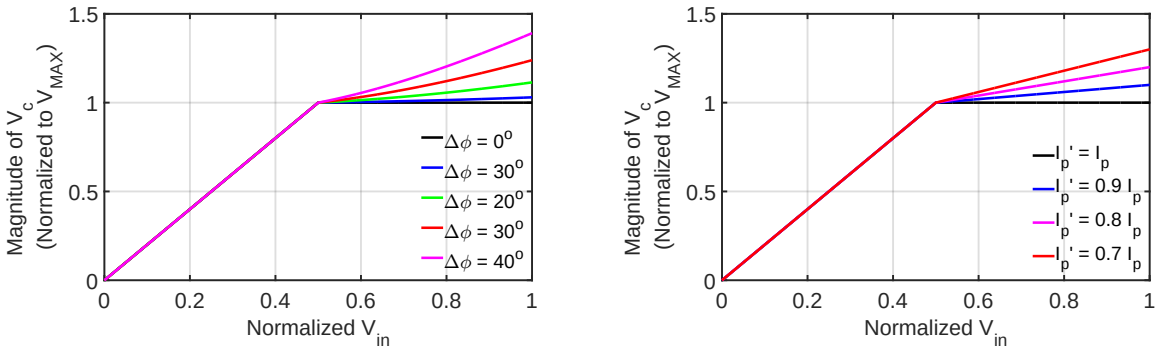
Phase and Magnitude of the Peaking Transistor Current As was shown in section 2.2.3, the ideal DPA operation depends on the correct relationship between I_p and I_c. Otherwise, Z_c can deviate from its ideal values at high power levels and cause AM-AM and AM-PM distortions. Recall that Z_c can be written as

$$Z_c = \frac{Z_T^2}{R_L} - \frac{jZ_T I_p}{I_c}. \qquad \text{(Eq. 2.5 revisited)}$$

If the relationship between I_p and I_c is correct, then Z_c will follow the ideal DPA impedance profile shown in Fig. 2.8(c). Recall that the peaking transistor turns on at higher power levels to reduce Z_c such that the voltage at the drain of the carrier transistor V_c remains constant at a maximum value of V_{MAX} as shown in Fig. 2.8(b). Otherwise, the transistor will be driven into its knee region if V_c is pushed to go beyond V_{MAX}. This will cause I_c (and hence V_L, which is linearly proportional to I_c) to be strongly nonlinear as a result.

Note that the second term in Eq. (2.5), $jZ_T I_p/I_c$, is real when the phase difference between I_p and I_c is 90°. In practice, there could be phase misalignment between the carrier and peaking transistor due to, for example, the AM-PM distortions in each transistor and the variation of AM-PM versus frequency for a broadband PA. When the phase difference deviates from 90° then $jZ_T I_p/I_c$ will have an imaginary part. As a result, the magnitude of Z_c will be higher at high power levels and the drain voltage of the main transistor will be pushed beyond V_{MAX} as shown in Fig. 2.18(a).

The magnitude of the I_p can also deviate from its ideal value due to practical reasons. For example, the peaking amplifier is usually biased in class C, which will have a slow



(a) (b)

Figure 2.18: Drain voltage of the carrier transistor for various (a) phase and (b) gain offsets between the peaking and carrier transistor current. ($\Delta\phi$ is how much $\angle I_p$ deviates from its ideal value.)

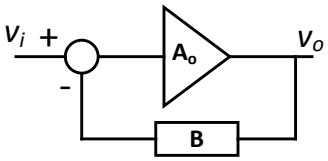


Figure 2.19: Illustration of negative feedback.

turn-on at low power levels. In addition, the point at which the peaking amplifier turns on can vary versus frequency and the magnitude of I_p can be reduced if the peaking amplifier is turned on at a higher input voltage. If the magnitude of I_p is decreased to be a value of I_p', then the magnitude of the second term in Eq. (2.5) will be reduced as a result. This will, in turn, increase the value of Z_c and the magnitude of V_c will be pushed beyond V_{MAX} as shown in Fig. 2.18(b).

2.4 PA Linearization Techniques

The distortions discussed in the previous section must be corrected by linearization techniques in order for the PA to achieve sufficient linearity. A few of the popular existing linearization techniques are discussed next.

2.4.1 Feedback

In feedback linearization techniques, the output is constantly compared with the input and corrected. As a result, it can remedy both quasi-static and dynamic distortions. In RF feedback linearization, the RF signal is directly fed back to the input. In this technique, the RF gain is traded for better linearity. This can be illustrated through the block diagram in Fig. 2.4. The overall gain of the amplifier with feedback is then given by

$$\frac{v_o}{v_i} = A = \frac{A_o}{1 + BA_o}, \tag{2.13}$$

where A is the gain of the amplifier with negative feedback (closed loop gain), A_o is the gain of the amplifier used in the feedback loop (open loop gain), and B is the feedback transfer function. The sensitivity of the closed loop gain to the open loop gain is

$$\frac{d(A)A_o}{d(A_o)A} = \frac{1}{1 + BA_o}. \tag{2.14}$$

Note that, the higher the factor $1 + BA_o$ is, the less sensitive the closed loop gain is to the open loop gain. If $1 + BA_o$ is large, then even if the open loop gain changes, the closed loop gain with negative feedback can remain almost constant, thereby making the amplifier more linear. However, a high $1 + BA_o$ value decreases the overall gain of the amplifier as seen from Eq. 2.13. Hence, there is a trade-off between linearity and gain.

At low frequencies it is easy to design amplifiers with excessively high gain. Hence it is practical to use negative feedback to improve amplifier linearity at the cost of lowering gain. However, at RF high gain is more difficult to achieve. One can use multistage PAs to increas gain, but multistage PAs will have large group delays. This means that the required phase of the feedback signal (180°) can only be maintained in a narrow frequency band. Outside of this narrow bandwidth, the phase of the feedback signal quickly changes, and the benefit of negative feedback diminishes [13]. Furthermore, the amplifier can become unstable as the change in feedback signal phase can lead to positive feedback.

The authors in [46] carefully designed the phase characteristic of the feedback path to achieve performance improvements at back-off between 3.4 and 4.2 GHz. For example, more than 7 dB improvement in IMD3 was achieved at 3 dB input back-off at 3.7 GHz. This, however, came at the cost of 7 dB gain reduction.

2.4.2 Feedforward

Feedforward (FFW) correction is a very effective distortion reduction technique which is unconditionally stable and able to compensate for both quasi-static and dynamic distor-

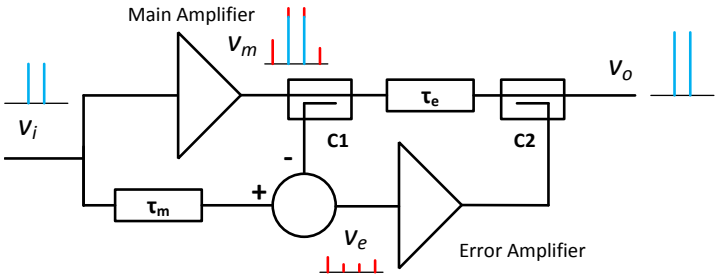


Figure 2.20: Feedforward amplifier configuration.

tions. It has been shown that, by using the FFW linearizer, a DPA's output ACLR can be improved by more than 20 dB [47]. The configuration of a FFW amplifier is given in Fig. 2.20. The input signal is fed to the main amplifier and an error cancellation path. The signal traveling through the main amplifier is amplified to be V_m, and due to the nonlinearity of the main amplifier, V_m will have significant distortion components. This distorted signal is then sampled by the coupler C1 and subtracted from the input signal to produce an error signal V_e, which is an attenuated version of the error generated by the main amplifier. V_e is then amplified by an error amplifier to its original magnitude, and subtracted from V_m to produce a clean output signal. Delay elements with τ_m and τ_e delays are use to compensate for the group delay of the main and error amplifier, respectively. They must be carefully tuned for optimal error cancellation.

One of the disadvantages of FFW amplifiers is that its overall power efficiency is usually low due to the high power consumption of the error amplifier and the insertion loss of the delay line and coupler in the output path. The error amplifier can have a significant amount of power consumption due to its low efficiency and relatively high power output. To start with, the reasons why the error amplifier's efficiency η_{error} is low is discussed (e.g., η_{error} was assumed to be only 5 % by the authors in [47]). Firstly, the error amplifier should be linear so that it can accurately amplify V_e. Any distortions produced by the error amplifier are directly injected to the output of the FFW amplifier. To achieve a good linearity, the error amplifier is usually biased in class A, which is the most linear but least efficient class of operation. Secondly, the PAPR of V_e is often higher than the signal itself [26]. This forces the error amplifier to operate even deeper into its low efficiency power back-off region. Aside from its efficiency, the power consumption of the error amplifier also depends on the amount of output power required for distortion cancellation. The higher the power of the main amplifier distortions and coupling ratio of C2, the higher the power required from the error amplifier. For example, the authors in [48] used a coupling ratio of 10 dB, and the average power of the distortions generated by the main amplifier was about 11 dBm.

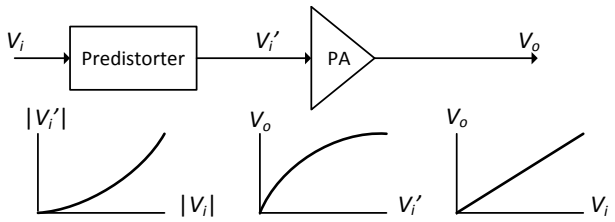


Figure 2.21: Predistortion system configuration.

Hence, the error amplifier should be linear with an average output power of 21 dBm. They ended up using an error amplifier with a peak power of 41 dBm, which is only 8 dB lower than the main amplifier's peak power. Another major reason why the efficiency of FFW amplifiers are generally low is due to the insertion loss of the delay line and coupler in the output path of the main amplifier. For the same FFW amplifier output power, the main amplifier will have to provide more power if this loss is increased (For example, with an insertion loss of 1.5 dB, the main amplifier will have to generate about 40 % more power compared to the case with no insertion loss).

The efficiency of the FFW amplifier has been improved through various ways. By using a DPA as the main amplifier, the authors in [47] have increased PAE by 2.2 % compared to the case when the main amplifier is a class AB PA. However, PAE was still relatively low at 10.9 % when the FFW amplifier was driven with a WCDMA signal. To further increase the efficiency of the FFW amplifier, authors in [49] have reduced the power consumption of the error amplifier by replacing it with a DPA topology as well. Although the error amplifier is now more nonlinear, they were still able to get the required linearity through crest factor reduction and offsetting the alignment (gain and phase) of the input signal to both the main amplifier and error amplifier such that their distortions cancel each other. By using a nonlinear error amplifier, the efficiency of the FFW amplifier was improved to be 23.4 % – 25.4 % with a 8.3 dB PAPR WCDMA signal. However, it remains to be seen if this technique will work with wider bandwidth signals and for different power levels. The efficiency of the FFW amplifier has also been improved by reducing the loss of the delay element after the main amplifier. The authors in [48] used a negative group delay circuit to eliminate the delay (and hence loss) required at the output of the main amplifier. As a result, this FFW amplifier was able to achieve an efficiency of 19.5 % with a WCDMA signal. The circuit used to implement the negative group delay, however, requires about 1 W of power consumption. This can significantly reduce overall efficiency if used with lower powered PAs. Moreover, this technique is not suitable for wideband signals because the bandwidth of the negative group delay circuit is limited to about 30 MHz.

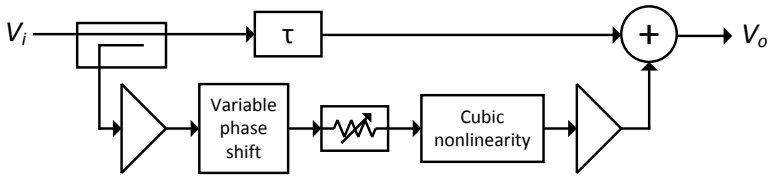


Figure 2.22: Simplified schematic of a cubic predistorter.

2.4.3 Predistortion

In predistortion linearization, a predistorter is used to generate a distorted input signal with distortion characteristics which are complementary to the PA's. This ensures that the PA with the predistorter is linear overall as illustrated in Fig. 2.21. The predistorter can be implemented using either analog/RF or digital circuits.

RF Predistortion An advantage of RF predistortion (RFPD) is that the predistorter is usually simple and low cost. One of the simplest RF predistorter can be built using a series diode, which was applied to a PA in [50] to improve the ACLR of a modulated signal by up to 5 dB. A more detailed analysis of this predistorter is given in Chapter 3. However, to the best of the author's knowledge, this series diode predistorter has not been used for the linearization of DPAs. This is likely due to the fact that the DPA's nonlinearity is usually stronger than a conventional PA. Hence, it can be difficult for the series diode based predistorter to generate the necessary predistorted signal.

To obtain better linearization, a cubic predistorter can be used which gives the designer more control over the predistorted signal. An example of a cubic predistorter is given in Fig. 2.22. The input signal is split into a main and secondary path: The main path has a time delay element which ensures that when the two signals are combined at the end they are properly time aligned while the secondary path is responsible for generating the predistortion component to be combined with the input signal. The cubic nonlinearity can be generated using anti-parallel diodes with a circulator [51] or 90° hybrid [52], and the cubic nonlinearity can then be shaped by phase and amplitude control elements. Using a cubic predistorter, the authors in [51] achieved 10 dB reducing in distortion noise for a 256-QAM signal. Predistorters similar to the cubic predistorter have been used in the past for the linearization of DPA amplifiers. For example, the authors in [53] used such a predistorter to significantly improve the linearity of the DPA by up to 14.6 dB with a four carrier WCDMA signal. This technique, however, requires the careful tuning of both the predistorter and the gate bias of the peaking transistor in the DPA for each power level.

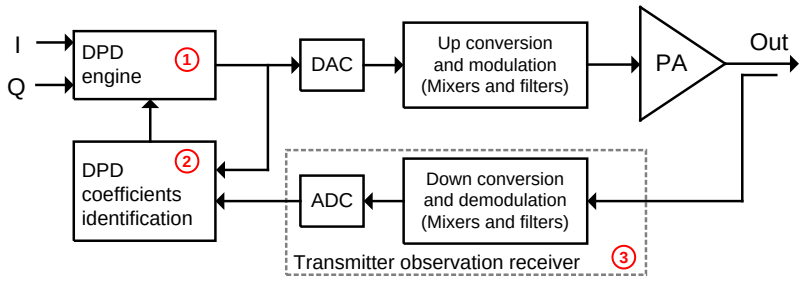


Figure 2.23: DPD based transmitter architecture.

Moreover, it does not correct for the dynamic distortions in the DPA. As a result, the effectiveness of the predistorter can degrade when wider bandwidth signals are used. In light of this, the authors in [54] have developed and used a multi-path predistorter where additional non-linear paths are added with various delay differences to compensate for the memory effects in the PA. This improves the linearity with wideband signals at the cost of increased complexity and power consumption.

Lastly, note that RF predistorters can be easily used in conjunction with other DPA linearization technique to further improve linearity. For example, the authors [55] used a cubic predistorter to improve the linearity of a 3-way DPA by another 1-2 dB.

Digital Predistortion Predistortion can also be performed in a digital manner. The DPD linearization technique utilizes digital circuitry to model the PA behaviour and construct the required predistortion signal. DPD has become the preferred linearization technique for macrocell base station PAs because it has benefited from the advancements in digital signal processors technologies and allow for excellent linearization with both quasi-static and dynamic distortions [10]. For example, the DPA in [56] was driven by a 100 MHz Long-Term Evolution (LTE) signal, and ACLR was improved by almost 20 dB after the application DPD.

The DPD based transmitter architecture is shown Fig. 2.23. The output signal is down converted to baseband and fed back through a transmitter observation receiver (TOR) which must have a bandwidth that is five times higher than the input signal bandwidth in order to capture the distortions generated by the nonlinear PA [57]. The feedback signal is then used by the DPD coefficient identification block and the DPD engine to model the PA's nonlinearity and generated an predistortion signal which accurately counteracts the PA's nonlinearity. For high-powered macrocell base station PAs, the power consumption and cost of the DPD blocks are negligible. However, the cost and power consumption

of the DPD architecture will significantly increase as the input signal bandwidth widens. For example, to linearize an LTE signal, whose modulation bandwidth can reach up to 100 MHz [58], an expensive 1 Gsps analog-to-digital converter with high power consumption is needed by the TOR to achieve a bandwidth of 500 MHz. In addition, the DPD's engine and coefficient identification blocks will require high-speed digital circuits with the same bandwidth as the TOR. They will further increase the power and cost overhead of DPD. Lastly, it should be noted that the DPD architecture's power overhead does not scale down with the power range of the PA. This means that the smaller the power of the PA, the more impractical it is to use DPD.

2.4.4 Derivative Superposition

A simplified schematic of a derivative superposition (DS) amplifier is given in Fig. 2.24. The DS method works by adding additional transistors in parallel with the main transistor and by adjusting their gate biases and widths such that their IMD currents cancel the main transistor's when summed together [59]. In [60], a similar technique was used, and the drain currents of the transistors were expressed as

$$i_{DS} = I_{dc} + g_m v_{gs} + \frac{g_m^{'}}{2!} v_{gs}^2 + \frac{g_m^{''}}{3!} v_{gs}^3 + \cdots, \tag{2.15}$$

where I_{dc} is the DC drain source current, g_m is the transconductance of the transistor, and $'$ and $''$ are the first and second derivatives with respect to the gate to source voltage v_{gs}, respectively. $g_m^{''}$ affects the value of the amplifier's IMD3 and should be minimized. In [60], it was shown that the main transistor M1 can have a nonzero and negative $g_m^{''}$ value around its bias point. In light of this, another auxiliar transistor (M2) was added and its bias and size was chosen such that its $g_m^{''}$ is positive and cancels the negative $g_m^{''}$ of the main transistor. Finally, a second auxiliary transistor (M3) was added to achieve cancellation across a wider input range. Looking at the results of the linearization versus input power in [59], however, it can be seen that IMD3 cancellation is maximized only around one power level in the DS amplifier, and degrades quickly towards peak power level. This is likely due to the fact that the distortions from the transistors changes versus input power such that they no longer achieve optimal cancellation at higher power levels. A new form of the DS technique was later developed to linearize high powered signals near the 1 dB compression point [61]. Once again, however, the cancellation is highly dependent on the power level. The linearity of the DS amplifier was better than a class A amplifier at high power, but worse at lower power levels.

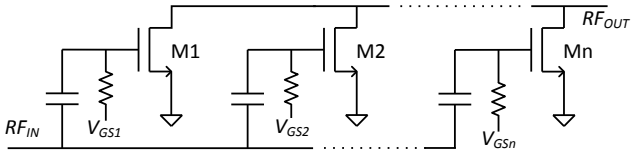


Figure 2.24: Simplified schematic of a DS amplifier.

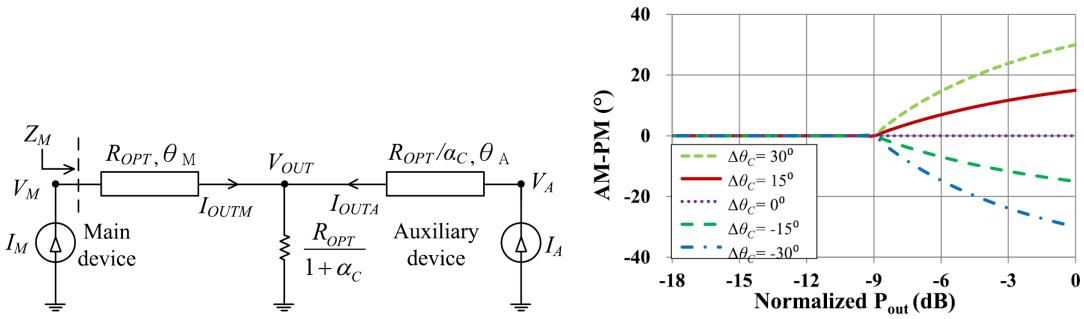
2.4.5 Circuit Level Linearization Techniques

Recent works have led to the development of different design methods that enables the PA to achieve a linear output without the assistance of other linearization blocks. In the following subsections, we will focus on methods which exploit the DPA topology to simultaneously achieve high linearity and efficiency.

Optimizing the Carrier and Peaking Transistor Bias A common DPA linearization technique is to adjust the gate bias of the carrier and peaking amplifier such that the distortions of the peaking amplifier cancels the distortions from the carrier amplifier. For example, this technique is utilized in [62] to improve the ACLR by about 12 dB. This method of distortion cancellation has also been utilized by N-way [55,63] and asymmetrical [64] DPAs to improve linearity performance.

There are, however, certain costs and limitations with this popular DPA linearizaiton technique. Firstly, the optimal gate bias required for distortion cancellation might not coincide with the bias for best efficiency. As a result, there is a trade-off between linearity and efficiency. Secondly, the effectiveness of this linearization technique is only maximized around a certain power level. Theoretically, there should be no distortion cancellation at low power levels when the peaking transistor is off. Furthermore, the distortions from the carrier and peaking transistor changes versus power. Hence, the level of cancellations will degrade versus power. The authors in [65] adaptively changes the value of the peaking amplifier's gate bias versus power to achieve better linearity over a wider power range. This approach, however, comes at the expense of additional power tracking bias supply circuits. Lastly, it should be noted that the distortions generated by the carrier and peaking transistor in the DPA is bandwidth dependent as a result of memory effects [64]. Hence, the effectiveness of this type of DPA linearization will likely degrade when the input is a modulated signal with a wide modulation bandwidth.

Current Scaled Doherty Amplifier Recently, the authors in [67] analyzed the output combining network of the DPA as a three port network and its parameters (as well as the



(a) (b)

Figure 2.25: The a) circuit diagram and b) AM-PM of the DPA proposed in [66]. Reprinted from [66].

input phase delay) are solved to simultaneously achieve high efficiency and linearity. For example, they found that it was possible to counteract the phase distortion of the DPA at high power by controlling the current ratio, load terminations, and the phase difference between the carrier and peaking transistor's current. Using their proposed technique, they were able to achieve 40 % PAE and -41 dBc ACLR with a 8.6 dB PAPR 10 MHz LTE signal. Similarly, the authors in [66] also played with the ratio of between the peaking and carrier transistor, and they developed a method to determine the combiner network parameters to synthesize a predetermined AM-PM characteristic that is the inverse of the carrier transistor's, thereby reducing the DPA's overall phase distortion. The combiner design used in [66] is shown in Fig. 2.25(a). Through the design of θ_M, θ_A, α_C, and β_d (where $\beta_d R_{OPT}$ is the impedance of Z_M at power back-off), the phase of the transmission coefficient (V_{OUT}/I_M) of the combining network, θ_C, can be set to generate different AM-PM profiles as can be seen in Fig. 2.25(b).

Both these techniques, however, mainly improve the linearity of the DPA at high power levels when the peaking transistor is on because they rely on the peaking transistor's current to correct for phase distortions at the output. This is evident when looking at the plot of ACLR versus signal power in [67]. The ACLR actually degrades as the average output power is decreased. Furthermore, the correction of AM-AM is achieved only through the design of the input matching network, which could lead to a compromise in gain.

Nonlinear Drivers As discussed in Section 2.3.3, the inherent load modulation that happens in the DPA can lead to AM-PM distortions. The authors in [68] have compensated for this source of distortion through the design of the driver stage. The idea is illustrated

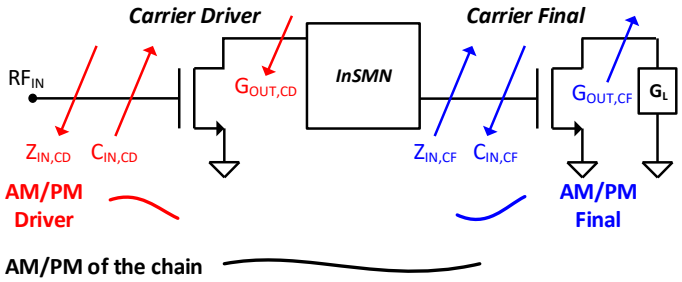


Figure 2.26: Graphical representation of the method proposed in [68]. Reproduced from [68].

in Fig. 2.26. At higher powers, output conductance seen by the carrier transistor, $G_{out,CF}$, increases. This causes the input impedance of the carrier transistor, $Z_{in,CF}$, to increase, which will in turn cause the output conductance ($G_{out,CD}$) and input impedance ($Z_{in,CF}$) of the driver stage to both decrease. Note that this impedance variation is opposite that of the carrier transistor. As a result, the driver will have an AM-PM that is also the opposite of the carrier transistor and, when cascaded together, the AM-PM distortion of the entire chain can be minimized. Using this technique, the authors were able to preserve the gain and efficiency of the DPA while effectively mitigating its phase distortion. However, this approach does not reduce the DPA's gain distortion, and as a result the ACLR of their final design was not better than the required -45 dBc.

2.5 Summary

At the beginning of this chapter, it was shown that conventional PAs will have low efficiency when used with modern communication signals with high PAPR. Hence, it is necessary to use back-off efficiency enhancement techniques, and a few common back-off efficiency enhancement techniques were surveyed. Out of all the techniques examined, the DPA amplifier was most promising because, unlike ET, it can support very wideband signals, and its combiner does not introduce as many complexities as the outphasing technique. Although the DPA's RF bandwidth was inherently limited by its quarter-wave transmission line, recent modifications have enabled it to achieve more than an octave of bandwidth [34].

It was also shown in this chapter that the DPA topology has multiple sources of distortion which must be remedied using linearization techniques. DPD has become the preferred linearization technique for macrocell base station DPAs because it offers excellent linearization and can support wideband modulated signals. However, DPD is no longer

Table 2.1: Linear DPA performance

	Topology	Centre freq. (GHz)	Mod. BW. (MHz)	Average P_{out} (dBm)	PAE (%)	ACLR (dBc)
[47]	2-way w. FFW	2.14	5	44.8	10.9	-60
[53]	2-way w. RFPD[1]	2.14	20	40.8	34.8[4]	-45
[54]	2-way w. RFPD[2]	2.14	5	34	18.9	-53.8
[55]	3-way w. RFPD	2.14	5/15	40/40	10.4/10.3	-51/-49.1
[65]	2-way w. DS[3]	2.14	5	39.4	31.8	-45
[63]	2/3/4-way	2.14	5 and 10	32/31/34	27.9/14.8/19.3	-30.1/-43.2/-38.4
[62]	2-way	2.14	20	44.8	26.9	-43.2
[67]	2-way	2.14	5/10/20	35.4/35.4/35.4	40/40/40	-41.5/-40.8/-40.5
[68]	2-way	7	56	32	40	-36
[66]	2-way	5	40/80/160	32	33.8[4]	-43.8/-44.1/-43.1

attractive for use with PAs in small cells and large antenna arrays due to its relatively high power overhead which does not scale down with the PA's power range. Some alternative methods of linearizing the DPA amplifier were also discussed in this chapter, and the performance of a few linearized DPA designs are summarized in Table 2.1. In [47], it can be seen that the FFW linearization technique is highly effective at canceling distortions, but its error amplifier power consumption and insertion losses limit the overall achievable power efficiency. The predistorters used in RFPD techniques, on the other hand, have lower power overheads compared to the FFW system. The RF predistorter, however, must either be carefully tuned for each power level or used in conjunction with another linearization technique for sufficient linearization [53] [55]. Moreover, the RF predistorter usually doesn't remedy the effects of dynamic distortions, and RF predistorters which try to reduce dynamic distortions are highly complex and more power hungry [54]. Various DPA design methods have also been proposed which optimizes its linearity [62, 63, 66–68]. These methods, however, are usually only very effective for a limited power or frequency range.

[1]The predistorter and the gate bias of the peaking transistor were carefully tuned for each power level.
[2]The predistorter uses multiple nonlinear paths.
[3]The drain and gate bias are controlled versus average power.
[4]Estimated from DE and gain plot.

Chapter 3

Power-Scalable Wideband PA Linearization Technique

In light of the limitations of conventional linearization techniques, a novel linearization technique is proposed for the lower-powered PAs in small cells and large antenna arrays. Aside from achieving the required amount of distortion reduction, the technique's power overhead should be low and scalable with the power range of the PA, and it should be able to support modern wideband communication signals.

The linearization technique proposed in this section achieves the aforementioned requirements by using a parallel addition of an LA at the PA's output. Contrary to the DS technique, which relies on adjusting the transconductance (biasing and sizing) of the parallel transistors, the proposed approach uses a LA where in addition to its transconductance, its output impedance is also judiciously chosen to be conducive to distortion reduction over a wide frequency range. Moreover, the LA is designed to contribute to the in-band output power, to further improve the linearity. Detailed theoretical analysis of the proposed technique is provided, as well as a practical way to realize LA. In addition, the design of a proof-of-concept prototype is detailed, and measurement results are presented to validate the technique's effectiveness.

3.1 Proposed PA Linearization Technique

The topology of the proposed linearization technique is shown in Fig. 3.1. The input signal $v_{in}(t)$ is split in two, where the first part is used to feed the PA, while the second is first

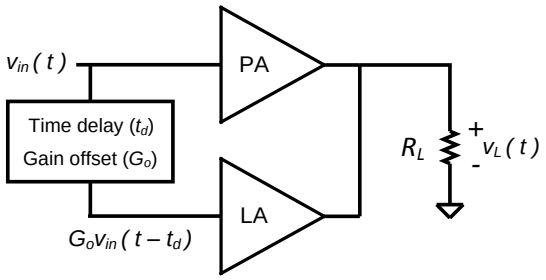


Figure 3.1: Simplified topology of the LA and PA.

time delayed by t_d and weighted by a scaling factor G_o before feeding the LA. A linear output voltage across the output load $v_L(t)$ can be attained through proper design of the LA and adjustment of t_d and G_o.

The following analysis of the proposed technique has been carried out in the frequency domain using the Norton equivalent circuits of both amplifiers for a given bias, input, and frequency[1]. The Norton circuit of the PA and the output load is illustrated in Fig. 3.2(a) where $I_{PA}(f)$ and $Z_{PA}(f)$ denote the Norton current and output impedance of the PA (including the load resistance R_L), respectively, at frequency f. In Fig. 3.2(b) the linearization amplifier's equivalent Norton circuit is added in parallel at the PA's output, where $I_{LA}(f)$ and $Z_{LA}(f)$ are LA's Norton current and output impedance, respectively, at frequency f. Based on figure 3.2(a) and 3.2(b), the PA's output voltage before and after addition of the LA, $V_L(f)$ and $V_L^{'}(f)$, respectively, can be expressed as

$$V_L(f) = I_{PA}(f)Z_{PA}(f), \tag{3.1}$$

and

$$V_L^{'}(f) = (I_{PA}(f) + I_{LA}(f))\frac{Z_{PA}(f)Z_{LA}(f)}{Z_{PA}(f) + Z_{LA}(f)}. \tag{3.2}$$

The ratio between the Norton currents and impedances of the LA and PA for the frequency f can be defined by $\alpha(f)$ and $\beta(f)$, respectively, where

$$\alpha(f) = \frac{I_{LA}(f)}{I_{PA}(f)} \quad \text{and} \quad \beta(f) = \frac{Z_{LA}(f)}{Z_{PA}(f)}. \tag{3.3}$$

[1]Note that this analysis makes the assumption that the PA and LA's output impedance and Norton current does not changes versus the load impedance. While this is not exactly true in practice, this approximation is reasonable for distortion reduction analysis as long as the distortion components (the IMD3 distortions for example) are relatively small in magnitude.

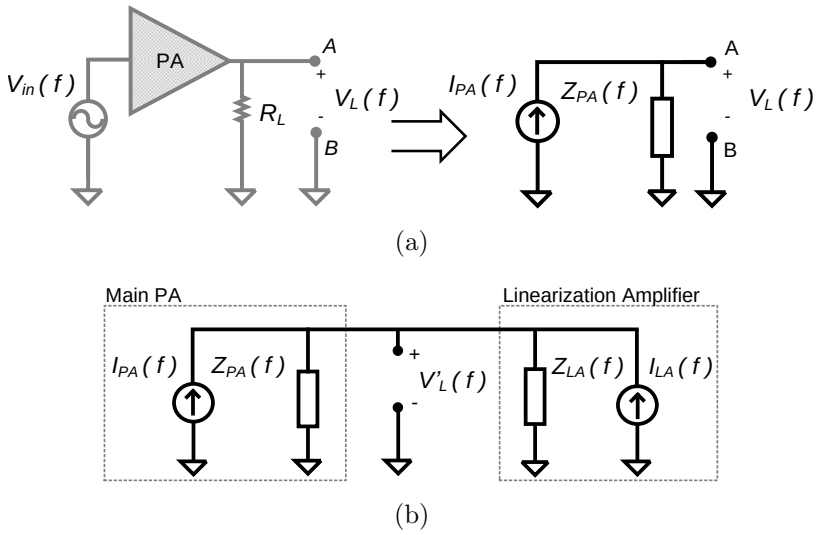


Figure 3.2: (a) The PA and its Norton equivalent with the (b) Norton equivalent of the PA and LA.

Then, by substituting equations (3.1) and (3.3) into Eq. (3.2), the output voltage after the addition of the LA can be expressed as

$$V_L^{'}(f) = V_L(f)\frac{1+\alpha(f)}{\beta^{-1}(f)+1}. \tag{3.4}$$

3.1.1 Linearity Improvement Mechanisms for Intermodulation Distortions

In this subsection, the LA's effects on the PA's lower third-order intermodulation distortion (IMD3L) will be discussed. The PA's IMD3L can be expressed as

$$IMD3L = 20\log\left|\frac{V_L(2f_1-f_2)}{V_L(f_1)}\right|. \tag{3.5}$$

Signal Power Enhancement The distortion voltage $V_L(2f_1-f_2)$ depends on the main PA's fundamental output power at f_1 and f_2. The lower the main PA's fundamental output power, the less distortion it generates, and the smaller $V_L(2f_1-f_2)$ is. Therefore, LA can yield IMD3L improvement if it can lower the portion of the fundamental output

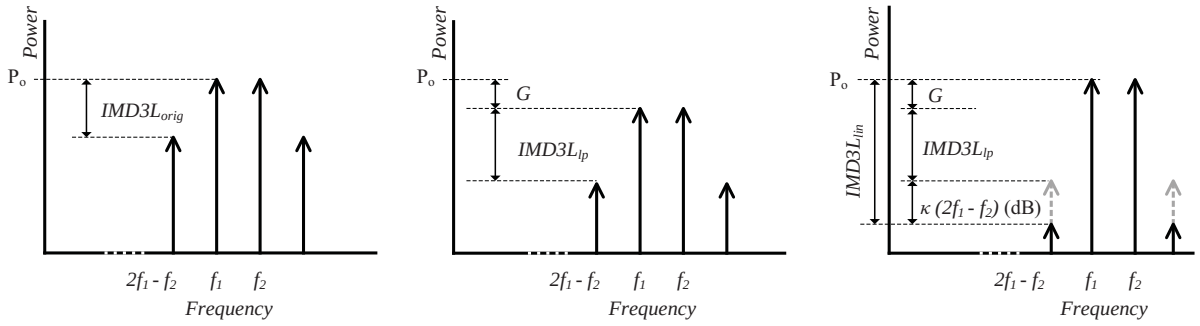


Figure 3.3: Expected output frequency spectrum with: (a) an output power of P_o dBm with PA only, (b) output power of $(P_o - G)$ dBm with PA only, and (c) output power of P_o dBm with PA and LA.

power required from the main PA. This can be accomplished through the setting of the LA's Norton current at f_1 and f_2 such that it contributes to the output power at those frequencies. For example, the LA can add to the output power at f_1 and increase it by G dB if the following equation is satisfied

$$\alpha(f_1) = 10^{\frac{G}{20}} - 1 + \beta(f_1)^{-1} 10^{\frac{G}{20}}. \tag{3.6}$$

In this case, the main PA would be supplying S_{LA} dB less power compared to the case without the LA. For the case where the impedances and currents are real, S_{LA} can be found to be

$$S_{LA} = -10 \log \left[1 + \frac{R_L(1 - \alpha(f_1)\beta(f_1))}{(1 + \alpha(f_1)) Z_{LA}(f_1)} \right]. \tag{3.7}$$

Note that S_{LA} is not equal to G. This is because adding the LA will increase the effective load impedance seen by the PA at the fundamental frequency f_1, and the amount of power supplied by the PA will be changed as a result.

As illustrated in Fig. 3.3(a), if the main PA is the only contributor to attain an output power of P_o dBm at f_1 and f_2, it will yield an IMD3L represented by $IMD3L_{orig}$. If an LA is to be added to the main PA, while maintaining the same output signal power P_o dBm, the main PA will only be driven to supply $(P_o - G)$ dBm of output signal power as shown in Fig. 3.3(b). Consequently the IMD3L of the main PA is now $IMD3L_{lp}$, which is improved over $IMD3L_{orig}$ because $V_L(2f_1 - f_2)$ has been decreased. If $V_L(2f_1 - f_2)$ for the case of the PA supplying P_o dBm and $(P_o - G)$ dBm of output power is given by $V_{L-orig}(2f_1 - f_2)$

and $V_{L-lp}(2f_1 - f_2)$, respectively, then

$$V_{L-lp}(2f_1 - f_2) = V_{L-orig}(2f_1 - f_2)10^{(IMD3L_{lp}-G-IMD3L_{orig})/20}. \tag{3.8}$$

Distortion Reduction According to Eq. (3.4), $V_{L-lp}(2f_1 - f_2)$ can be further reduced through the proper design of $I_{LA}(2f_1 - f_2)$ and $Z_{LA}(2f_1 - f_2)$ according to the following expression

$$\left|V'_{L-lp}(2f_1 - f_2)\right| = \kappa(2f_1 - f_2)\left|V_{L-lp}(2f_1 - f_2)\right| \tag{3.9}$$

where

$$\kappa(2f_1 - f_2) = \frac{|1 + \alpha(2f_1 - f_2)|}{|\beta^{-1}(2f_1 - f_2) + 1|}. \tag{3.10}$$

Equations (3.9) and (3.10) show that to minimize $|V'_{L-lp}(2f_1 - f_2)|$ the LA should be configured so that $\kappa(2f_1 - f_2) < 1$. This can be achieved by minimizing the numerator and maximizing the denominator of $\kappa(2f_1 - f_2)$ such that

$$|1 + \alpha(2f_1 - f_2)| \ll 1 \tag{3.11}$$

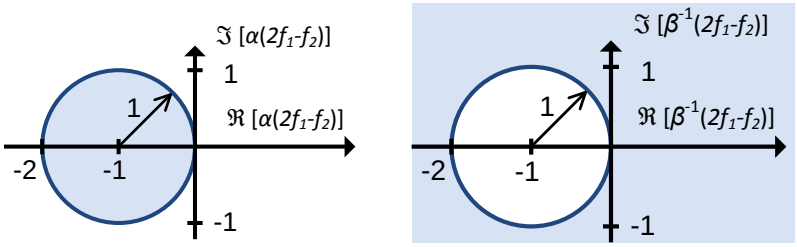
and

$$|\beta^{-1}(2f_1 - f_2) + 1| \gg 1. \tag{3.12}$$

Equation (3.11) can be satisfied when $\alpha(2f_1 - f_2)$ is designed to be well within the unity circle centred around -1. This region is designated in Fig. 3.4(a). The closer $\alpha(2f_1 - f_2)$ is to -1, the lower $\kappa(2f_1 - f_2)$ is. Optimally, if $\alpha(2f_1 - f_2) = -1$, (this is when the Norton current of the LA is equal in magnitude, but opposite in phase to the Norton current of the main PA) then $\kappa(2f_1 - f_2)$ goes to zero. However, this is not always possible, so it is also important to maximize the denominator of $\kappa(2f_1 - f_2)$ as given in Eq. (3.12). This can be achieved when $\beta(2f_1 - f_2)$ is designed to be well outside the unity circle centred around -1. This region is designated in Fig. 3.4(b). The further away $\beta(2f_1 - f_2)$ is from -1, the lower $\kappa(2f_1 - f_2)$ is. For example, in the case where Z_{PA} and Z_{LA} are both real, then the smaller Z_{LA} is, the further away $\beta(2f_1 - f_2)$ is from -1, and the smaller $V_{L-lp}(2f_1 - f_2)$ is.

The frequency spectrum of the PA output after the addition of the LA is illustrated in Fig. 3.3(c), and the IMD3L with LA, in dB, is denoted $IMD3L_{lin}$. Compared to Fig. 3.3(b), the output powers at f_1 have been restored to P_o dBm, while the distortion powers at $2f_1 - f_2$ has been further reduced by $\kappa(2f_1 - f_2)$. The total reduction in IMD3L due to addition of the LA can be expressed as

$$\Delta IMD3L_{lin} = IMD3L_{orig} - IMD3L_{lin} = \\ G - 20\log(\kappa(2f_1 - f_2)) + IMD3L_{orig} - IMD3L_{lp}. \tag{3.13}$$



(a) (b)

Figure 3.4: Desired values for $\alpha(2f_1 - f_2)$ and $\beta^{-1}(2f_1 - f_2)$ are highlighted in (a) and (b), respectively.

Although the preceding analysis was conducted for IMD3L, the same results can be obtained for the upper third-order intermodulation distortion (IMD3U). Also, Eq. (3.13) has been validated through simulation, and the results are presented in Appendix A.

In summary, we showed through analysis in this subsection that the LA can improve the linearity of the PA by enhancing its output signal power and by reducing its distortions through the design of the LA's output impedance and Norton currents.

3.1.2 Linearity Improvement for In-band Distortions

The Norton current of the PA and LA at an in-band frequency f_1 can be expressed as follows:

$$\begin{aligned} I_{PA}(f_1) &= g_{m-pa}V_{in}(f_1) + I_{PA-d}(f_1) \\ I_{LA}(f_1) &= g_{m-la}V_{in}(f_1) + I_{LA-d}(f_1) \end{aligned} \tag{3.14}$$

where g_{m-pa} and g_{m-la} are the small signal transconductances of the PA and LA's Norton current, respectively, and $I_{PA-d}(f_1)$ and $I_{LA-d}(f_1)$ are the distortion caused by the main PA and the LA's non-linearity, respectively. The output voltage at f_1 before the addition of the LA is given by

$$V_L(f_1) = g_{m-pa}V_{in}(f_1)Z_{PA}(f_1) + I_{PA-d}(f_1)Z_{PA}(f_1). \tag{3.15}$$

If g_{m-la} is set to be

$$g_{m-la} = g_{m-pa}(10^{\frac{G}{20}} - 1 + \beta(f_1)^{-1}10^{\frac{G}{20}}), \tag{3.16}$$

then the output voltage with the LA, denoted by $V_L^{'}(f_1)$, is given by

$$V_L^{'}(f_1) = 10^{\frac{G}{20}}(g_{m-pa}V_{in}(f_1)Z_{PA}(f_1)) + \kappa_d(f_1)(I_{PA-d}(f_1)Z_{PA}(f_1)). \tag{3.17}$$

Note that the linear part of the PA's load voltage ($g_{m-pa}V_{in}(f_1)Z_{PA}(f_1)$) is now G dB higher with the addition of the LA while the distorted part of the PA's load voltage ($I_{PA-d}(f_1)Z_{PA}(f_1)$), on the other hand, is now multiplied by a factor $\kappa_d(f_1)$ with the addition of the LA, which can be written as

$$\kappa_d(f_1) = \frac{1 + \frac{I_{LA-d}(f_1)}{I_{PA-d}(f_1)}}{1 + \beta(f_1)^{-1}}. \tag{3.18}$$

The lower the value of $|\kappa_d(f_1)|$, the lower the in-band distortion is with the addition of the LA. Note that $|\kappa_d(f_1)|$ is similar to $\kappa(f_1)$ in Eq. (3.10) except in the place of $\alpha(f_1)$, the numerator is now a function of the ratio $I_{LA-d}(f_1)/I_{PA-d}(f_1)$. Therefore, the same distortion cancellation analysis can be applied here, and the in-band distortion can be reduced by the addition of LA.

3.2 Linearization Amplifier Design

3.2.1 LA Realization Using Negative Feedback

The LA designer should have some control over the LA's Norton current and output impedance such that they can be designed to satisfy constraints given in equations (3.11) and (3.12). One possible way to realize the LA is through the use of negative feedback on a single common source transistor. The ideal topology of this LA design and its Norton equivalent is shown in Fig. 3.5. If $R_s g_m \gg 1$, then the output resistance of the LA is given by

$$Z_{LA} = \frac{R_s + R_f}{1 + R_s g_m} \approx \frac{1 + \frac{R_f}{R_s}}{g_m}. \tag{3.19}$$

If it is assumed that $R_f g_m \gg 1$, then the Norton current source of the LA is given by

$$I_{LA} = \frac{V_{in}(1 - R_f g_m)}{R_f + R_s} \approx -\frac{V_{in} g_m}{1 + \frac{R_s}{R_f}}. \tag{3.20}$$

The larger R_s/R_f is, the smaller the output resistance and the Norton current of the LA will be.

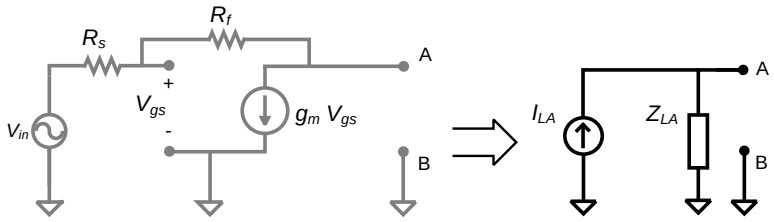


Figure 3.5: LA realized through negative feedback and its Norton equivalent

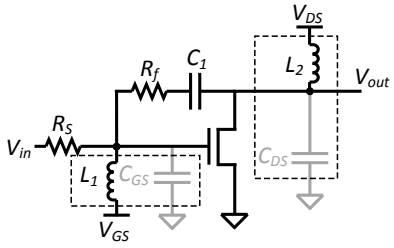


Figure 3.6: LA schematic with negative feedback

The partial derivative of the output impedance and Norton current with respect to the transconductance of the transistor is

$$\frac{\partial(Z_{LA})}{\partial(g_m)} \approx \frac{-1-\frac{R_f}{R_s}}{g_m^2} \quad \text{and} \quad \frac{\partial(I_{LA})}{\partial(g_m)} = \frac{-V_{in}}{1+\frac{R_s}{R_f}}, \tag{3.21}$$

respectively. A high R_s/R_f value can reduce the sensitivity of the LA's output impedance and Norton current to variations in the transistor's nonlinear transconductance. Hence, using negative feedback also improves the linearity of LA.

Lastly, the parasitics of the transistor should be considered as they will affect the effectiveness of the feedback at high frequencies. For example, a large capacitance between the gate and source (C_{GS}) can cause I_{LA} to approach zero and Z_{LA} to approach R_f at high frequencies. In light of this, the inductors L_1 and L_2 are added to resonate out the gate- and drain-source capacitances, respectively, around the centre frequency of the PA as illustrated in the simplified schematic of the final LA topology shown in Fig. 3.6. The capacitance between the gate and the drain of the transistor (C_{GD}) can also affect the feedback and reduce I_{LA} at high frequencies. Hence, it is desirable to select a transistor which has a very small gate-drain capacitance. Finally, a DC blocking capacitor, C_1, is added in the feedback branch to isolate the gate and drain bias of the LA. Its value should be chosen such that it presents a low impedance at the PA's centre frequency.

3.2.2 Power Consumption of LA and Overall Efficiency

The LA introduces additional DC power consumption P_{LA-DC} to the overall system given as

$$P_{LA-DC} = V_{LA-DS} I_{LA-DS}, \tag{3.22}$$

where V_{LA-DS} and I_{LA-DS} are the DC components of the LA transistor's drain to source voltage and current, respectively. V_{LA-DS} is equal to the DC drain supply voltage of the LA transistor, and to ensure that the LA transistor's drain voltage is not pushed below its knee voltage V_k, V_{LA-DS} should be chosen such that $V_{LA-DS} \geq V_k + V_{peak}$, where V_{peak} is the PA's output voltage's peak amplitude. I_{LA-DS} is mainly determined by the amount of current supplied by the LA for fundamental power contribution. For example, in the case of a CW input signal, if ideal class B behaviour is assumed for the LA transistor, then I_{LA-DS} is given by

$$I_{LA-DS} = \frac{2I_{out}(1 - 10^{-\frac{S_{LA}}{10}})}{\pi}, \tag{3.23}$$

where I_{out} is the fundamental current through the output load. If V_k is relatively small compared to V_{peak}, then the DC power consumption of LA can be written as

$$P_{LA-DC} \approx \frac{2(V_{peak} I_{out})(1 - 10^{-\frac{S_{LA}}{10}})}{\pi}. \tag{3.24}$$

According to Eq. (3.24), P_{LA-DC} is directly proportional to $V_{peak} I_{out}$. Since $V_{peak} \propto \sqrt{P_{max}}$, where P_{max} is the maximum capable output power of the PA in watts, and $I_{out} \propto \sqrt{P_{out}}$, where P_{out} is the output power of the PA in watts, it can be stated that the DC power consumption of the LA will scale with the factor $\sqrt{P_{max} P_{out}}$. If the ratio between P_{out} and P_{max} remains the same, then the DC power consumption of the LA scales linearly with the PA's power range.

With LA contributing to the fundamental output power, the total DE is now given by

$$\eta_{total} = \frac{P_{out}}{P_{LA-DC} + P_{PA-DC}} = \frac{P_{out}}{\frac{P_{LA-out}}{\eta_{LA}} + \frac{P_{PA-out}}{\eta_{PA}}}, \tag{3.25}$$

where P_{LA-out} and P_{PA-out} are the output power supplied by the PA and LA, respectively, and η_{PA} and η_{LA} are the DE of the PA and LA, respectively. As the output power supplied by main amplifier is S_{LA} dB less than P_{out} when the LA is used, the output power from the PA and LA can be written instead as

$$P_{PA-out} = P_{out} 10^{\frac{-S_{LA}}{10}} \quad \text{and} \quad P_{LA-out} = P_{out}(1 - 10^{\frac{-S_{LA}}{10}}), \tag{3.26}$$

respectively. Then, by substituting Eq. (3.26) into Eq. (3.25), the efficiency can be expressed alternatively as

$$\eta_{total} = \frac{\eta_{LA}\eta_{PA}}{10^{\frac{-S_{LA}}{10}}\eta_{LA} + (1 - 10^{\frac{-S_{LA}}{10}})\eta_{PA}}. \tag{3.27}$$

If the power contribution of LA is very low, then S_{LA} goes to zero, and $\eta_{total} \approx \eta_{PA}$. In this condition, the main PA would be more saturated, and thus most efficient. However, if the LA is used to contribute to the output power, such that $S_{LA} > 0$, then the efficiency of the overall amplifier is both a function of η_{PA} and η_{LA}. Because the PA is now less saturated, η_{PA} will be reduced as a result, and η_{total} will be decreased.

The PAE of the overall system is also affected by the addition of the LA. This is because the total input power is now comprised of both the input power to the PA and LA. Thus it is important that the LA has a sufficiently high transconductance ($I_{LA}(f)/V_{in}(f)$), else it can significantly decrease the overall gain, and hence the PAE. It should also be noted that the input power to the PA is reduced after the addition of the LA because the PA now only has to supply a portion of the output power.

3.3 Main PA and LA Co-design

In light of the previous analysis, the following method was developed for the design of a prototype system:

1. Design a high efficiency main PA with the necessary RF bandwidth.
2. Design an LA with the following goals:
   - An output impedance that places $\beta^{-1}(f)$ as far away from -1 as possible.
   - A sufficiently high transconductance ($I_{LA}(f)/V_{in}(f)$).
3. Connect the LA output to the load of the main amplifier as shown in Fig. 4.1 and implement a time delay and scaling factor for the LA input, given by t_d and G_o, respectively.
4. With a two-tone signal as input, tune the LA's Norton current and output impedance in simulation by adjusting t_d, G_o, and the LA's bias to improve IMD3 while avoiding excessive power consumption.
5. Verify the system's performance with modulated signals by performing co-simulation.

3.3.1 PA and LA design

Both the PA and the LA were realized on a 1.9 mm thick ROGERS RT/duroid 6006 substrate. The PA was designed as a class AB PA using a 6 W Wolfspeed GaN transistor (CGH40006), and the simplified real frequency technique [69] was used to achieve a frequency of operation between 700 and 1000 MHz. The LA was designed using the same 6 W transistor. A smaller sized transistor could be used for the design of the LA since it will have a lower output power than the PA, but 6 W was the smallest commercial GaN transistor available from Wolfspeed.

The ratio between the LA's resistors, R_f/R_s, was used to control the output impedance $Z_{LA}(f)$ and transconductance of the LA to satisfy condition 2 outlined at the beginning of Section 3.3. The best setting for R_f/R_s is not obvious: Maximizing R_f/R_s can minimize $Z_{LA}(f)$ and help to push $\beta^{-1}(f)$ far away from -1. However, it will decrease the transconductance, which can severely degrade the PA's overall gain. Moreover, the value of R_f should not be too small or there will be a significant amount of power dissipation through it. At the same time, R_f should be smaller than the impedance magnitude of the parasitic gate drain capacitance it is in parallel with.

The gate bias of the LA's transistor should also be carefully chosen as it affects affects the LA's output impedance, transconductance, and efficiency. The LA transistor's transconductance, g_m, is reduced by having a low gate bias. This, in turn, will increase $Z_{LA}(f)$ and decrease the LA's transconductance, given previously in Eq. (3.19) and Eq. (3.20), respectively. On the other hand, a high gate bias will lower the efficiency of the LA.

A schematic of the final design is given in Fig. 3.7. The layout was simulated using Keysight's Momentum electromagnetic (EM) simulator and Keysight's Advanced Design System (ADS) was used for EM-Circuit co-simulation of the entire circuit. The simulated small-signal transconductance of the LA is plotted in Fig. 3.8 and the small-signal output impedances of the LA and PA are plotted in Fig. 3.9 for the entire bandwidth of the PA. The value of $\beta^{-1}(f)$ is plotted in Fig. 3.10(a). Note that $\beta^{-1}(f)$ remained outside of the unity circle centred at -1 for the entire frequency band of interest, thus satisfying the constraint specified by eq. (3.12).

3.3.2 Linearization Analysis Using a Two-tone Input

A two-tone signal was sent through the PA and LA and the output was obtained using Harmonic Balance simulation. The frequencies of the two tones, f_1 and f_2 were set at

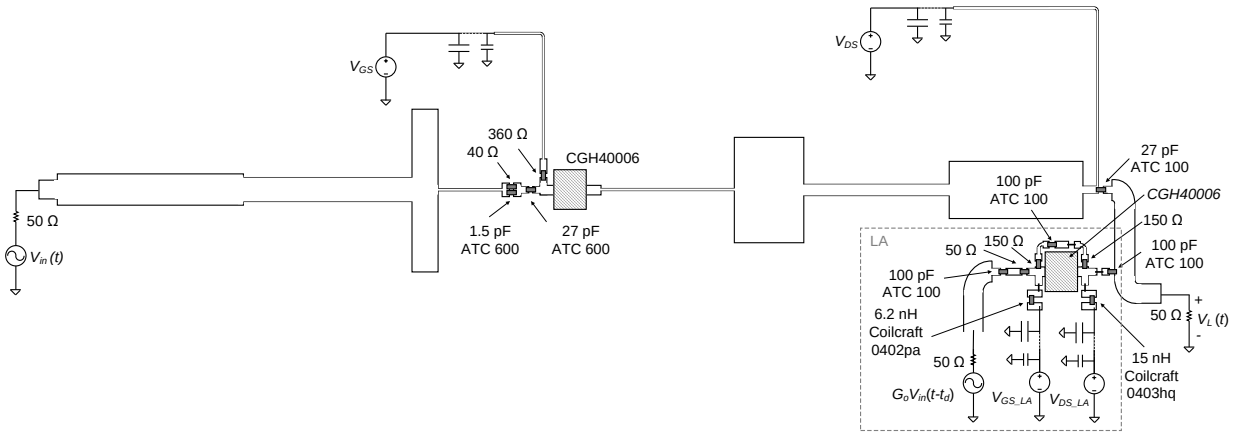


Figure 3.7: Complete schematic of the PA and LA

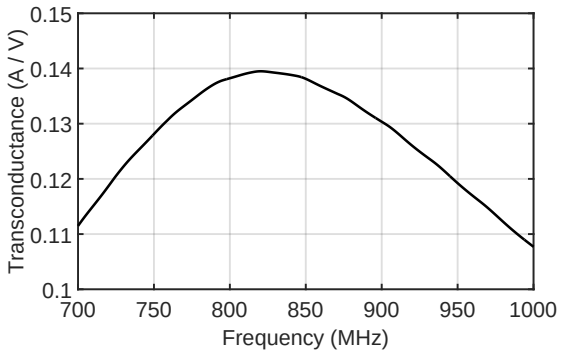
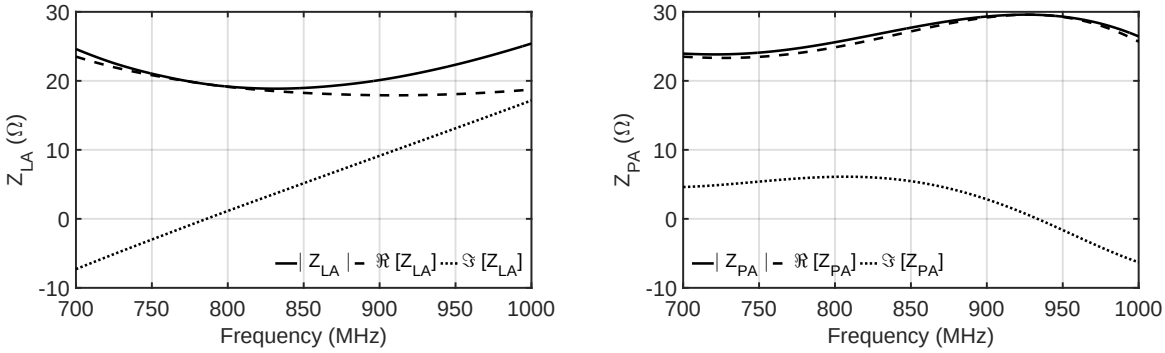


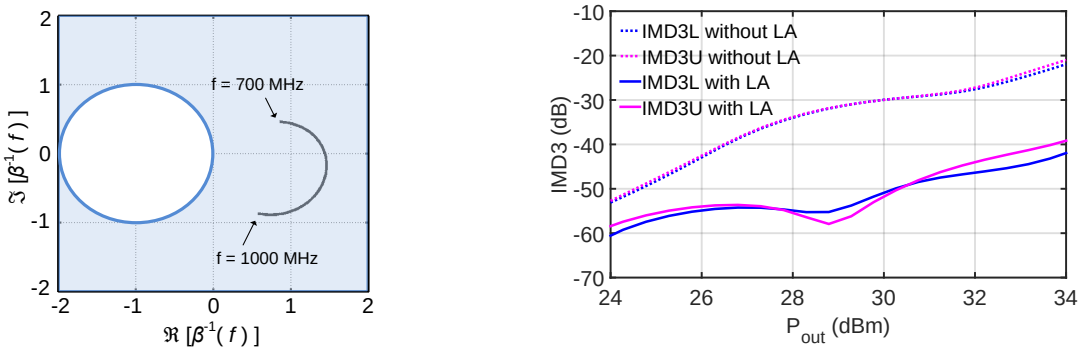
Figure 3.8: Simulated small-signal transconductance of the LA

840 and 860 MHz, respectively. Then, IMD3L and IMD3U at the output was minimized by tuning the following values. First, while t_d is initially adjusted to compensate for the difference in group delay between the LA and PA, it can also be slightly offset to tune the phase of LA's IMD currents, and therefore the phase of $\alpha(f)$ at an IMD frequency. However, this should be done sparingly as the phase of the output signal at f_1 and f_2 will be offset as well: This, in turn, will degrade how effective the fundamental powers are combined. In addition to t_d, the LA transistor's gate bias was tuned to set both the phases and magnitudes of the LA's IMD frequency Norton currents such that the criterion specified by Eq. (3.11) can be achieved. Lastly, G_o can also be adjusted to control the magnitude of $I_{LA}(f)$ at f_1 and f_2 for the purpose of adjusting the LA's fundamental power contributions as shown in Eq. (3.6). Note that the setting of G_o affects the phase and



(a) (b)

Figure 3.9: Simulated small-signal output impedances of the (a) LA and (b) PA.

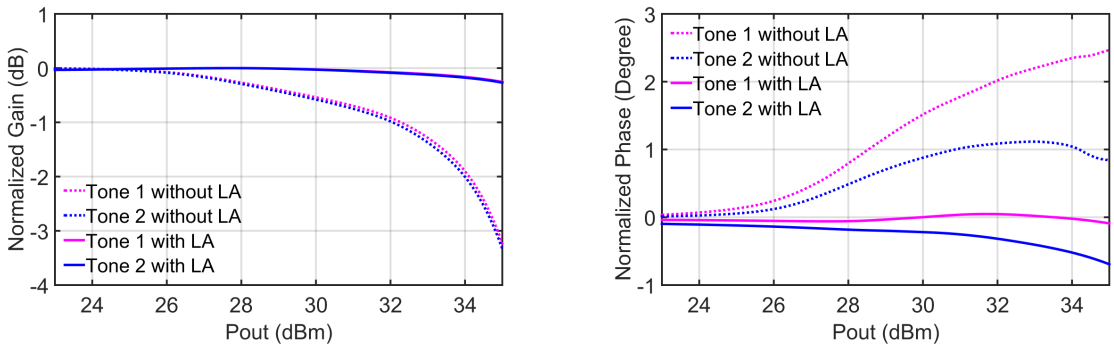


(a) (b)

Figure 3.10: Simulated a) $\beta^{-1}(f)$ profile and b) IMD3 of the PA with a two-tone input signal with tones 1 and 2 at 840 and 860 MHz, respectively

magnitude of $I_{LA}(f)$ at IMD frequencies as well.

The IMD3 before and after the addition of the LA is plotted in Fig. 3.10(b) where the LA was biased in deep class AB, with a bias current of 15 mA, while t_d and G_o were set to 1.32 ns and 2, respectively. The IMD3 of the PA improved by up to 20 dB after the addition of the LA. Also plotted is the normalized magnitude and phase of gain of the two tones in Fig. 3.11. The magnitude and phase variation of the gain was significantly reduced, which demonstrates the technique's ability to mitigate both AM-AM and AM-PM nonlinearities. However, the absolute value of the gain was reduced by the additional



(a) (b)

Figure 3.11: Simulated and normalized (a) magnitude and (b) phase of the gain of the PA with a two-tone input signal with tones 1 and 2 at 840 MHz and 860 MHz, respectively

input power required by the LA. For example, the gain was decreased from about 17.5 to 12.5 dB at 30 dBm output power.

3.3.3 Linearization Under Modulated Signals

Using the design values found through the two-tone signal simulations, the PA and LA were then driven by wideband modulated signals. The design was simulated using Keysight's ADS RF and Ptolemy co-simulations. A modulated signal with a bandwidth of 40 MHz was used to drive the PA. The signal consisted of a 15 MHz bandwidth LTE signal in band 1, and a 15 MHz bandwidth WCDMA signal in band 2. The carrier frequencies of the two bands were 25 MHz apart. The PAPR of the signal was 8.4 dB, and the output power was kept at 29.3 dBm. The frequency spectrum at the output of the PA is plotted in Fig. 3.12 for the cases with and without the LA. After the addition of LA, the ACLR was improved by 15.2 dB from -31.9 dBc to -47.1 dBc and the EVM was improved from 5.1 % to 2.5 %.

3.4 Measurement Results

A picture of the fabricated PA and LA is shown in Fig. 3.13(a). The PA and the LA both used a 28 V drain supply, and the biasing currents for the PA and the LA were 100 mA and 10 mA, respectively. Keysight's M8190A arbitrary waveform generator (AWG) was used to generate the RF input signal to the PA and LA through its channel 1 and 2, respectively.

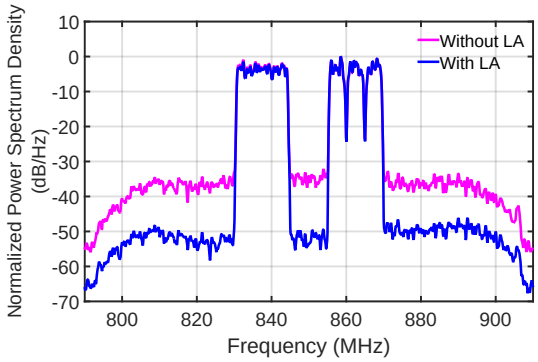
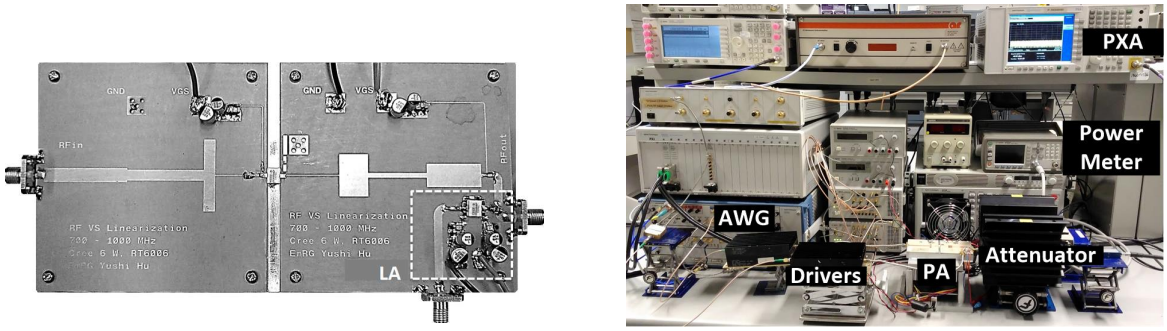


Figure 3.12: Simulated PA output spectrum with a 40 MHz bandwidth modulated signal

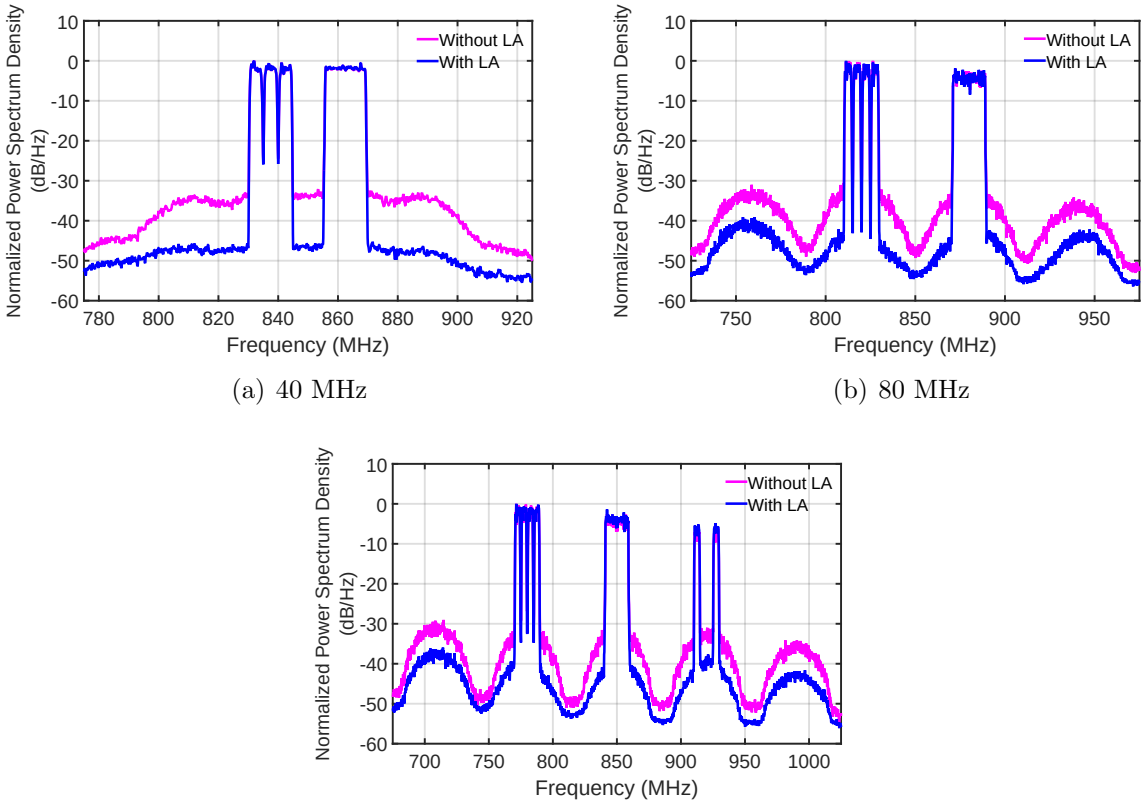


(a) (b)

Figure 3.13: Photograph of a) fabricated PA with LA and b) measurement setup

The time delay t_d was realized and tuned by delaying the input signal uploaded to the AWG's channel 2, while the gain G_o was realized and adjusted by controlling the output amplitude of channel 2 relative to channel 1. The output from channel 1 and channel 2 were then amplified through two RF drivers (ZHL-42). A picture of the complete measurement setup is shown in Fig. 3.13(b).

The same 40 MHz bandwidth signal used for co-simulation was used to drive the PA, and the performance with and without the LA was measured. The output frequency spectrum is plotted in Fig. 3.14(a) for an average output power of 29.1 dBm. The ACLR was improved by up to 13.2 dB, while the EVM was improved from 6.5 % to 3.7 %. Next, two modulated signals with very wide bandwidths were used to drive the PA and the performance was measured with and without the LA. The first modulated signal had a bandwidth of 80 MHz and a PAPR of 9.6 dB. The signal consisted of a 20 MHz bandwidth



(a) 40 MHz

(b) 80 MHz

(c) 160 MHz

Figure 3.14: Measured PA output frequency spectrum with various modulated signals.

WCDMA signal in band 1 and a 20 MHz bandwidth LTE signal in band 2. The carrier frequencies of the two bands were 60 MHz apart. The output frequency spectrum is plotted in Fig. 3.14(b) for an average output power of 28.1 dBm. The second modulated signal had a bandwidth of 160 MHz and a PAPR of 8.9 dB was used to drive the PA. The signal consisted of a 20 MHz bandwidth WCDMA signal in band 1, a 20 MHz bandwidth LTE signal in band 2, and another 20 MHz bandwidth WCDMA signal in band 3. The output frequency spectrum is plotted in Fig. 3.14(c) for an average output power of 28.8 dBm. The linearity improvements achieved for the three modulated signals are summarized in Table 3.1.

Although significant linearity improvements were achieved for every signal, it should be noted that the ACLR improvement does degrade with increasing bandwidth. One of the

Table 3.1: Measured PA performance with and without LA at 850 MHz

Signal Bandwidth (MHz)	PAPR (dB)	Average P_{out} (dBm)	t_d (ns)	G_o	DE (PAE) without / with LA (%)	Band	ACLR without/with LA (dBc)	ACLR improvement (dB)
40	8.4	29.4	1.42	2.1	25.9 (25.4) / 17.4 (16.4)	1	-33.1 / -45.7	12.6
						2	-31.6 / -44.8	13.2
80	9.6	28.1	1.43	2.1	21.9 (21.5) / 15 (14.2)	1	-31 / -42.2	11.2
						2	-31.3 / -40.4	9.1
160	8.9	28.8	1.38	2.2	24.7 (24.2) / 17.3 (16.3)	1	-31.9 / -39.5	7.6
						2	-29 / -39	10
						3	-24.9 / -32.9	8

reasons is the LA's bandwidth was limited by the large delay through its negative feedback path. In the prototype design a discrete transistor was used, and the feedback path needed to be routed around the transistor. This issue can be mitigated by realizing the topology in an integrated circuit (IC) technology. Moreover, the bandwidth of this linearization technique is limited by the mismatch in group delay between the main PA and the LA across frequency.

According to Fig. 3.14, IMD3 distortions generated by the mixing of the two modulated signals were also improved. For the modulated signal with 80 MHz bandwidth, IMD3 distortions around 760 and 940 MHz were reduced by 8.1 and 8 dB, respectively. For the modulated signal with 160 MHz bandwidth, its IMD3 distortions around 710 MHz and 990 MHz were reduced by 7 and 7.5 dB, respectively. The reduction of these IMD3 distortions is critical for the operation of the entire transceiver system. When sufficiently large, IMD3 distortions can leak into the receiver path and cause receiver desensitization [70]. Furthermore, fifth-order intermodulation distortion (IMD5) is also reduced after the addition of the LA. In Fig. 3.14(a), it can be seen that the IMD5 distortions around 787.5 and 912.5 MHz were reduce by about 6 and 7.6 dB with the LA.

The DE of the combined amplifier decreased by about 7-8 percentage points after the addition of the LA. Note that the power consumption overhead of the LA in this prototype is still significantly less than a conventional DPD solution. For example, in the case where the PA is driven by a modulated signal with 40 MHz bandwidth, the use of LA added about $1.6\,W$ of DC power consumption, while a conventional DPD system would require several watts of power (estimated to be 5.4 W in [12] for a 30 MHz bandwidth system). Moreover, this DC power consumption should scale down with the PA's maximum power as previously discussed. Therefore, the lower the PA's power rating, the more beneficial the present solution is in terms of power consumption when compared with DPD. Finally,

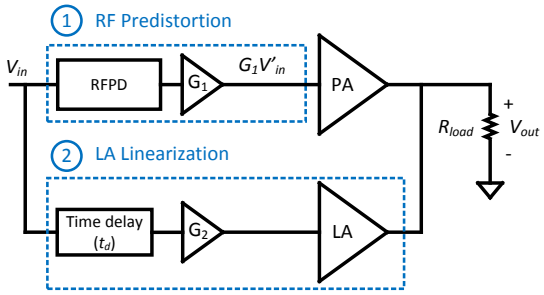


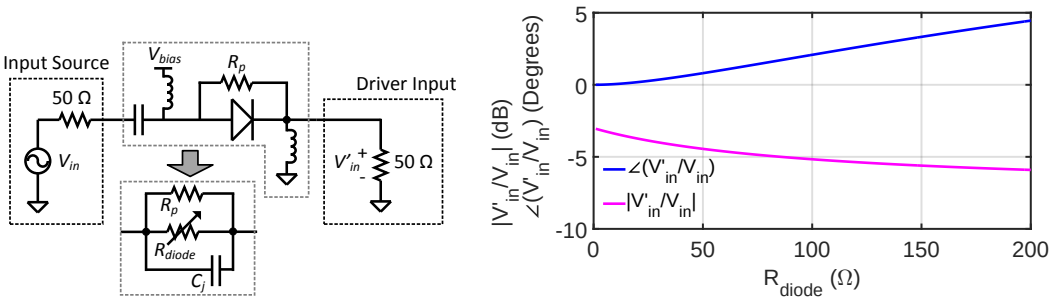
Figure 3.15: Topology of joint linearization

the power consumption of the LA does not increase with higher bandwidth output signals. The DE stayed constant when the input signal bandwidth increased from 40 MHz to 160 MHz. A DPD solution on the other hand will require a higher digital clock rate for wider bandwidth output signals, and the overall system's complexity as well as power and hardware costs will increase significantly. While DPD is capable of achieving better linearization, the proposed technique provide a lower power and lower cost alternative that is practically viable for future applications involving small cells and antenna array systems.

3.5 Joint Linearization with RF Predistortion

As discussed in section 2.4.3, the amount of linearization provided by a simple RF predistorter is not sufficient by itself when there is strong nonlinearity. However, it is a low cost technique which, when used with the proposed linearization technique, can further improve PA linearity and reduce the power consumption of the LA. The topology of the resulting joint RFPD and LA linearization is shown in Fig. 3.15. Block one is used to implement coarse static RFPD linearization at the PA input. Then, block two realizes the LA based linearization. For the RF predistorter, a simple configuration with a series diode is used based on a previous work [50]. The schematic and equivalent circuit of the predistorter is given in Fig. 3.16(a), where C_j is the junction capacitance of the diode, and R_{diode} is the diode's variable resistance. When the diode is biased with a very low forward current, an increase in the input power will decreases R_{diode}. The change in R_{diode} will then vary the ratio between the RF predistorter's input and output voltages, given by $V_{in}(f)$ and $V'_{in}(f)$, respectively, which can be expressed as

$$\frac{V'_{in}(f)}{V_{in}(f)} = \frac{50}{100 + \frac{R_d}{1+(2\pi f C R_d)^2} - \frac{j2\pi f C R_d^2}{1+(2\pi f C R_d)^2}}, \tag{3.28}$$



(a) Series diode linearizer

(b) Calculated magnitude and phase of V'_{in}/V_{in} at 850 MHz for the case when C_j =0.2 pF and $R_p = 300\Omega$

Figure 3.16: RF predistorter schematic and calculated performance

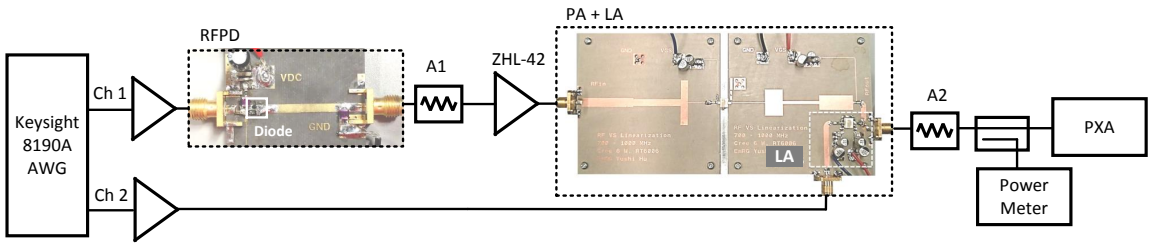
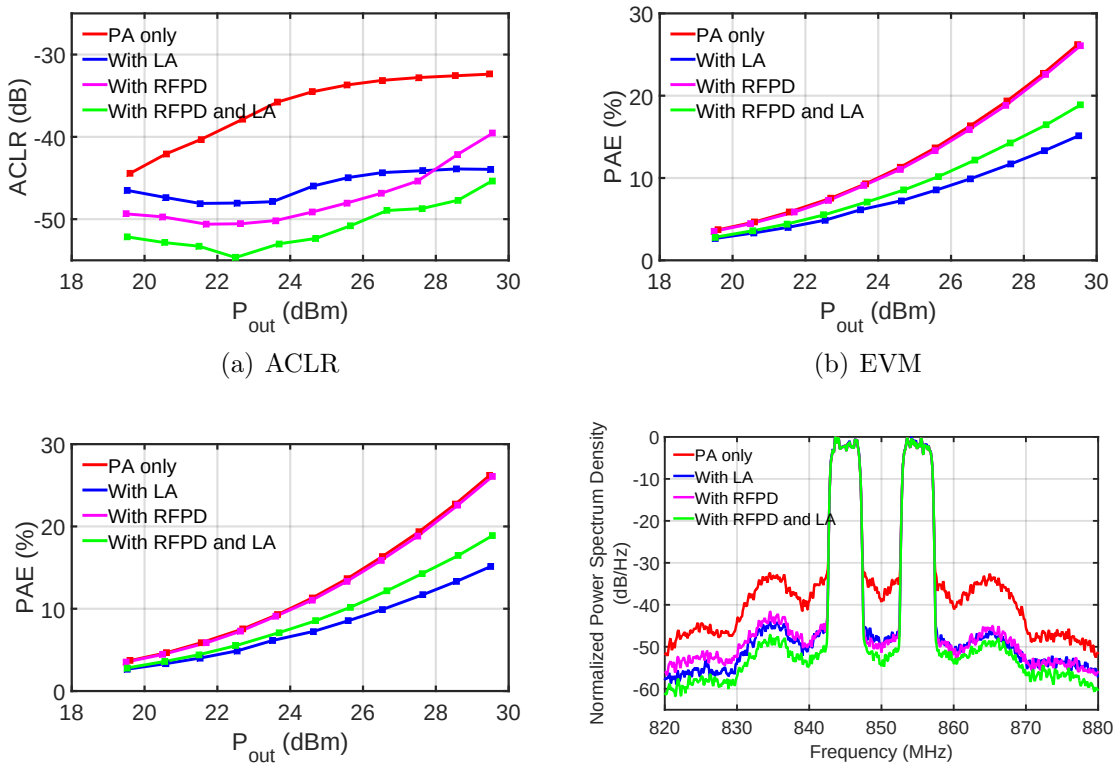


Figure 3.17: Diagram of measurement setup

where $R_d = R_{diode} \parallel R_p$. The phase and magnitude of $V'_{in}(f)/V_{in}(f)$ is plotted versus R_{diode} in Fig. 3.16(b) at 850 MHz for the case when $C_j = 0.2$ pF (junction capacitance of Skyworks' SMS7621 Schottky diode) and R_p is 300 Ω. Note that the phase decreases while the magnitude increases as R_{diode} decreases. Since R_{diode} will decrease with increasing input power, $V'_{in}(f)/V_{in}(f)$'s magnitude and phase will experience expansion and compression, respectfully, versus input power. The bias voltage V_{diode}, R_p, and G_1 are all then used to adjust the phase of magnitude of $V'_{in}(f)/V_{in}(f)$ such that it approaches the inverse of the PA's AM-AM and AM-PM, respectively.

A prototype of the proposed system is implemented and measured as a proof of concept. The same PA and LA is used as shown in Fig 3.7, while the RF predistorter was built using Skyworks' SMS7621 schottky diode with a V_{bias} of 0.1 V. The measurement setup is shown in Fig. 3.17. The input signals to the RF predistorter and LA were generate through channel 1 and 2 of Keysight's 8190A AWG. t_d and G_2 are tuned through delaying and amplifying the magnitude of channel 2's output, respectively. t_d was fixed at 1.42 ns, while G_2 was equal to 1.3 and 2.5 for the configurations with and without RFPD, respectively.



(a) ACLR

(b) EVM

(c) PAE

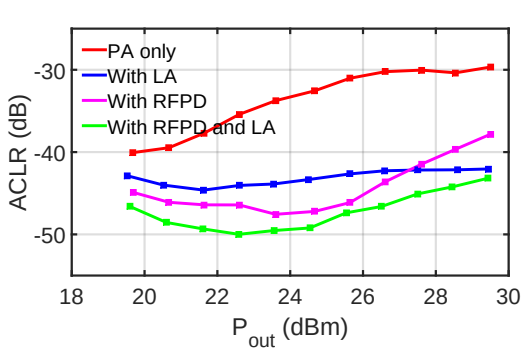
(d) Output spectrum with 28.5 dBm output power

Figure 3.18: Measured performance of the different linearization configurations with a 15 MHz WCDMA signal

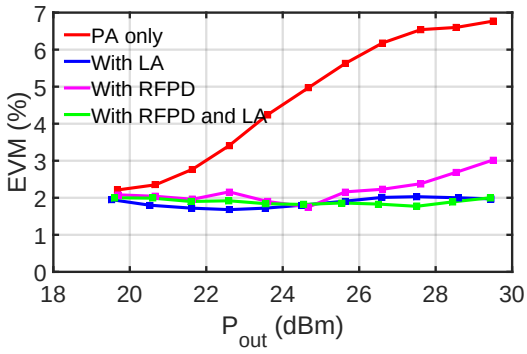
Lastly, G_1 is set to be about 18 dB through the value of the attenuator A1.

First, a 15 MHz bandwidth WCDMA modulated signal with a PAPR of 8.3 dB was used to drive the PA to 29.9 dBm of output power. The ACLR, EVM, and PAE of four linearization configurations are measured versus average output power and plotted in Fig. 3.18. The following are the four configurations measured for comparison: PA only, PA with LA, PA with the RFPD, and PA with RFPD and LA. The measured output spectrum with 28.5 dBm average output power is also shown in Fig. 3.18(d) for the four configurations.

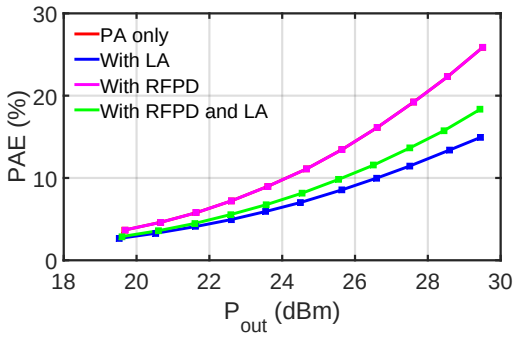
The benefit of using RFPD and LA together becomes apparent as this configuration gave the best ACLR and EVM performance at every power level. It managed up to 5.8 dB and 6.6 dB more ACLR improvement when compared to the configuration with only RFPD or LA, respectively. It is observed that while the configuration with only the



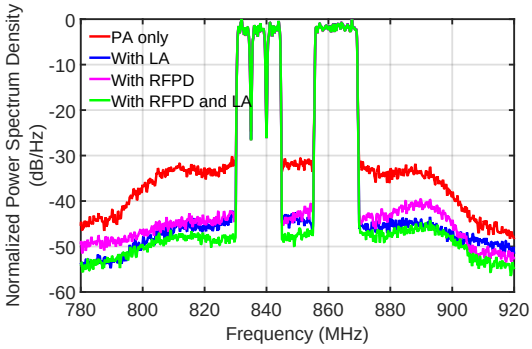
(a) ACLR



(b) EVM



(c) PAE



(d) Output spectrum with 28.5 dBm output power

Figure 3.19: Measured performance of the different linearization configurations with a 40 MHz WCDMA signal

RFPD attained good linearization at low power levels, at peak power, when there is strong nonlinearity, it was limited to only about 7 dB of ACLR improvement. On the other hand, the configuration with only the LA achieved good linearization at peak power, but the level of ACLR improvement does not increase at power back-off. Compared to the configuration with only the PA, the RFPD and LA combination attained over 13 dB and 5 percentage point of ACLR and EVM improvement, respectively, at peak power, and up to 18 dB and 5.2 percent point of ACLR and EVM improvement, respectively, at power back-off.

In addition to the improvements in distortion reduction, the configuration with both the RFPD and LA yielded better efficiency than the configuration with only LA. The PAE of the RFPD and LA combination is 3.8 percentage points higher at peak power than the configuration where only the LA was used. The reason for this is because, when the RFPD

is cascaded with the PA, the LA no longer needed to be driven to contribute significant power to the output for the best linearity results. Consequently, both the LA's input power and power consumption were decreased.

Lastly, the linearization configurations were also tested with a wideband modulated signal with 40 MHz bandwidth and the ACLR, EVM, and PAE performance of the different configurations are measured versus average output power and plotted in Fig. 3.19. The measured output spectrum for the four cases with 28.5 dBm average output power is also shown in Fig. 3.19(d). Similar to the results with the 15 MHz signal, the linearity performance was best when the RFPD was used with the LA. The configuration with both RFPD and LA managed up to 5.4 dB and 6 dB more in ACLR improvement when compared to the case with only RFPD or LA, respectively. Compared to the configuration with only the PA, the RFPD and LA combination attained 13.5 dB and 4.8 percentage point of ACLR and EVM improvement, respectively, at peak power, and up to 16.7 dB of ACLR improvement at power back-off.

3.6 Conclusion

In this chapter, a novel linearization technique based on the addition of an LA was proposed and analyzed. As a proof of concept, an LA was designed for a broadband class AB PA with a centre frequency of 850 MHz. Measurement results showed that the LA can significantly improve the PA's linearity even when it is amplifying signals with a modulation bandwidth as wide as 160 MHz. In the case where the PA was driven by a modulated signal with 40 MHz bandwidth, for example, the PA's ACLR was improved by up to 13 dB with the addition of the LA, allowing it to achieve an ACLR of -45 dBc without any additional linearization techniques. Furthermore, the power overhead of the proposed linearization technique does not increase with the input signal bandwidth and are scalable such that it remains low compared to the PA's power range. This makes the proposed technique very suitable for use in small cells and large antenna arrays. Moreover, it was shown that, for the case of a class AB PA, an RF predistorter can be used in conjunction with the LA to reduce power consumption and further improve the linearity of the PA.

Chapter 4

Doherty Distortion Reduction Using a Linearization Amplifier

In the previous chapter a new linearization technique which is suitable for PAs in small cell base stations and large antenna arrays was proposed based on the design and addition of an LA. However, the overall efficiency was relatively low because the technique was applied to a class AB PA, which has low back-off efficiency, and because of the power consumption of the LA. In this chapter, the LA will be augmented with a DPA design to form a new linear DPA topology which can achieve both high efficiency and linearity when stimulated by wideband modulated signals. A suitable design strategy is first developed which accounts for the interaction between the LA and the DPA, and, in comparison to the method used in the previous section, an important difference is that the proposed topology now allows for the design of the output impedance of the DPA, which is utilized as a new design parameter to improve linearity and minimize the LA's power consumption.

4.1 Proposed Topology

The proposed linear DPA topology, composed of the DPA and LA, is shown in Fig. 4.1. The input signal $V_{in}(t)$ is split in two. One part is used as input to the DPA, while the other is time delayed by t_d and weighted by a scaling factor G_o before feeding the LA. A linear output voltage across the load, R_L, can be attained through proper design of the LA and adjustment of t_d and G_o. The DPA and LA are connected using an impedance inverter with a characteristic impedance of about R_L. The impedance inverter is used to

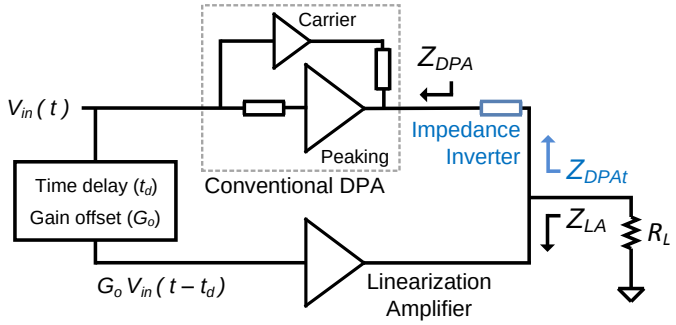


Figure 4.1: Block diagram of the linear DPA topology.

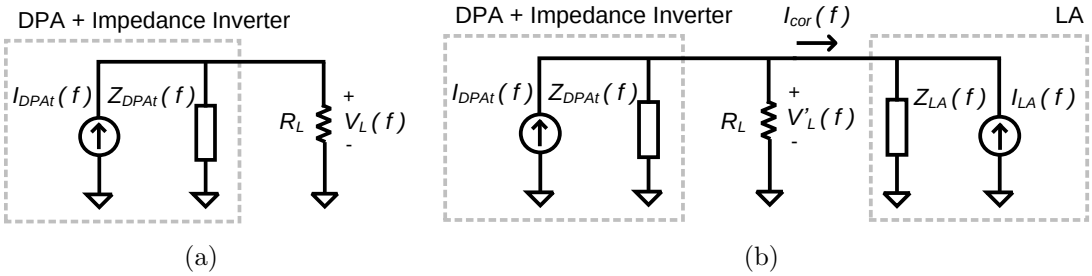


Figure 4.2: Equivalent circuit schematic of the DPA (a) without and (b) with the LA.

transform the output impedance of the DPA from the conventional output impedance of $Z_{DPA}(f)$ to a higher impedance of $Z_{DPAt}(f)$. In the following subsections, the motivation for this design of the DPA's output impedance is detailed.

4.1.1 Linearization Improvement by the Impedance Inverter

In this subsection, the proposed linear DPA topology will be analyzed and the pertinence of the additional impedance inverter will be explained. For this analysis, the schematics given in Fig. 4.2 will be used, which are simplified versions of Fig. 4.1 for the case with and without the LA. $I_{DPAt}(f)$ and $Z_{DPAt}(f)$ are the Norton current and output impedance, respectively, of the DPA and impedance inverter series combination at frequency f for a given bias and input. Notice that, compared to the equivalent circuit schematic in Fig. 3.2, the load impedance R_L is now separated from the $Z_{DPAt}(f)$ in the analysis. This is done so that the effects $Z_{DPAt}(f)$ has on the overall performance can be better studied.

At an out-of-band frequency[1], f_{ob}, the factor $\kappa(f_{ob})$, which relates the load voltage with and without the LA, denoted as $V_L^{'}(f_{ob})$ and $V_L(f_{ob})$, respectively, can be written as

$$\kappa(f_{ob}) = \frac{V_L^{'}(f_{ob})}{V_L(f_{ob})} = \frac{1 + \frac{I_{LA}(f_{ob})}{I_{DPAt}(f_{ob})}}{1 + \frac{R_L \| Z_{DPAt}(f_{ob})}{Z_{LA}(f_{ob})}}. \tag{4.1}$$

The lower the value of $|\kappa(f_{ob})|$, the lower the out-of-band distortion component is with the addition of the LA.

At an in-band frequency f_i the Norton currents of the DPA and LA can be written as follows:

$$\begin{aligned} I_{DPAt}(f_i) &= g_{m-dpa} V_{in}(f_i) + I_{DPAt-d}(f_i) \\ I_{LA}(f_i) &= g_{m-la} V_{in}(f_i) + I_{LA-d}(f_i) \end{aligned} \tag{4.2}$$

where g_{m-dpa} and g_{m-la} are the small signal transconductances of the DPA and LA's Norton current, respectively, and $I_{DPAt-d}(f_i)$ and $I_{LA-d}(f_i)$ are the distortion parts of the DPA and LA's Norton current, respectively. Then, the load voltage at f_i, without the LA, can be written as

$$V_L(f_i) = g_{m-dpa} V_{in}(f_i)(R_L \parallel Z_{DPAt}(f_i)) + I_{DPAt-d}(f_i)(R_L \parallel Z_{DPAt}(f_i)). \tag{4.3}$$

If g_{ml} is set based on the factor G such that

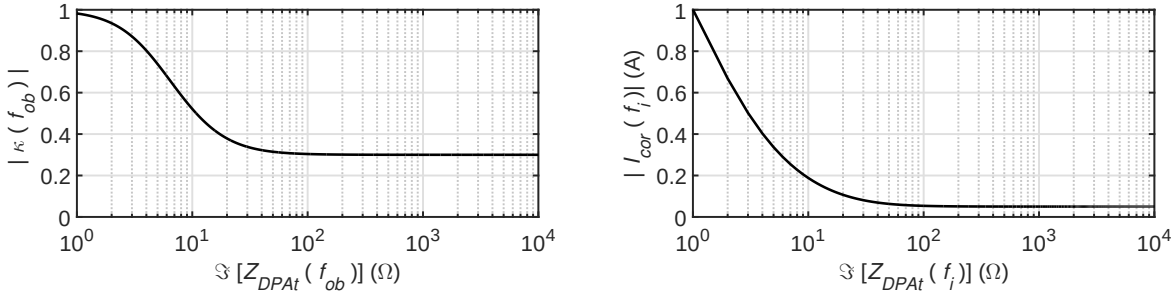
$$g_{m-la} = g_{m-dpa}\left(10^{\frac{G}{20}} - 1 + 10^{\frac{G}{20}} \frac{R_L \parallel Z_{DPAt}(f_i)}{Z_{LA}(f_i)}\right), \tag{4.4}$$

then the output voltage with the LA, denoted by $V_L^{'}(f_i)$, is given by

$$V_L^{'}(f_i) = 10^{\frac{G}{20}}[g_{m-dpa} V_{in}(f_i)(R_L \parallel Z_{DPAt}(f_i))] + \kappa_d(f_i) I_{DPAt-d}(f_i)(R_L \parallel Z_{DPAt}(f_i)). \tag{4.5}$$

Note that the linear part of the DPA's load voltage ($g_{m-dpa} V_{in}(f_i)(R_L \parallel Z_{DPAt}(f_i))$) is now G dB higher with the addition of the LA. In the last chapter, G was set to be greater than zero such that the LA contributes to the in-band power to further reduce distortion. However, it will be shown later that this approach is not suitable with the DPA topology due to the undesirable load modulation this would cause. Instead, G will be set to be about zero in this topology, through the adjustment of G_o and t_d, such that the LA does not significantly change the linear output power of the DPA.

[1]For example, if we used the two-tone analysis done in Chapter 3, then the IMD3 frequencies $2f_1 - f_2$ and $2f_1 - f_2$ are both out-of-band frequencies.



(a) (b)

Figure 4.3: Calculated a) $|\kappa(f_{ob})|$ for various $Z_{DPAt}(f_{ob})$ values with $Z_{LA}(f_{ob}) = R_L/3$, $I_{LA}(f_{ob}) = 0.01\ A$, $R_L = 20\ \Omega$, and $V_L(f_{ob}) = 1V = I_{DPAt}(f_{ob})[Z_{DPAt}(f_{ob})||R_L]$ and b) LA correction current for various $Z_{DPAt}(f_i)$ values with $V_L(f_i) - V_L^{'}(f_i) = 1\ V$ and $R_L = 20\ \Omega$.

The distorted part of the DPA's load voltage ($I_{DPAt-d}(f_i)(R_L \parallel Z_{DPAt}(f_i))$), on the other hand, is now multiplied by a factor $\kappa_d(f_i)$ with the addition of the LA, which can be written as

$$\kappa_d(f_i) = \frac{1 + \frac{I_{LA-d}(f_i)}{I_{DPAt-d}(f_i)}}{1 + \frac{R_L \| Z_{DPAt}(f_i)}{Z_{LA}(f_i)}}. \tag{4.6}$$

The lower the value of $|\kappa_d(f_i)|$, the lower the in-band distortion is with the LA.

The analysis thus far has shown that in-band and out-of-band distortions can be reduced through minimizing the value of $|\kappa(f_{ob})|$ and $|\kappa_d(f_i)|$, respectively, while the linear part is preserved if G is set to 0 dB. In other words, the LA will behave like an attenuator which only attenuates the distortions. $|\kappa(f_{ob})|$ and $|\kappa_d(f_i)|$ can be minimized through maximizing the denominator in equations (4.1) and (4.6). In the previous chapter, this was achieved solely through the design of $Z_{LA}(f)$. However, according to equations (4.1) and (4.6), their denominators could be further increased through the design of the DPA's output impedance, $Z_{DPAt}(f)$. In this topology, the value of $Z_{DPAt}(f)$ will be utilized, in addition to $Z_{LA}(f)$, to maximize the denominator of equations (4.1) and (4.6), thereby further reducing output distortion.

The value of $Z_{DPAt}(f)$ is controlled in this topology using the impedance inverter between the DPA and LA. First, it should be noted that without this impedance inverter $Z_{DPAt}(f)$ is simply the output impedance of the conventional DPA. In the conventional DPA design, a quarter-wave transformer is in series with the infinite output impedance of the carrier transistor, transforming it into a voltage source (zero output impedance).

Hence $Z_{DPAt}(f) = 0$ and $I_{DPAt}(f) = \infty$. As a result, $\kappa(f) = \kappa_d(f) = 1$, and no distortion reduction is possible. With the impedance inverter, $Z_{DPAt}(f)$ becomes the impedance looking into the impedance inverter and the conventional DPA, and the impedance inverter transforms $Z_{DPAt}(f)$ from the zero output impedance of the conventional DPA to ∞. Assuming that $Z_{LA}(f)$ is real, then the denominators in equations (4.1) and (4.6) are maximized when $Z_{DPAt}(f) = \infty$, and can help reduce both $\kappa(f)$ and $\kappa_d(f)$. To demonstrate this, $|\kappa(f_{ob})|$ is calculated and plotted in Fig. 4.3(a) while $Z_{DPAt}(f_{ob})$ is swept from $j1$ to $j10000\ \Omega$. During the sweep, $I_{DPAt}(f_{ob})$ is also adjusted such that the output distortion voltage without the LA, $V_L(f_{ob})$, is kept at $1\ V$. Notice that the value of $|\kappa(f_{ob})|$ is decreased as $Z_{DPAt}(f_{ob})$ is increased. For example, $|\kappa(f_{ob})|$ is more than halved as $Z_{DPAt}(f_{ob})$ is increased from $j1$ to $j100\ \Omega$. However, beyond $j100\ \Omega$, the change in $|\kappa(f_{ob})|$ due to a change in $Z_{DPAt}(f_{ob})$ is small as $|\kappa(f_{ob})|$ approaches a value of $|1+I_{LA}(f_{ob})R_L/V_L(f_{ob})|/|1+R_L/Z_{LA}(f_{ob})|$, which is equal to 0.3 in this particular case.

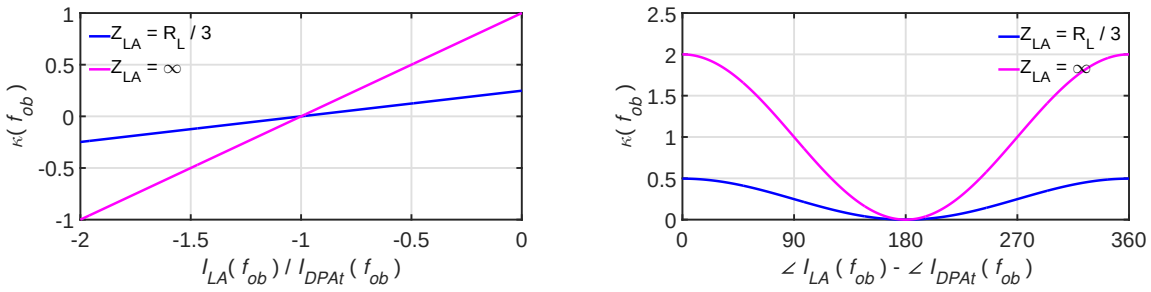
In summary, the analysis in this section showed that using an impedance inverter at the output of the DPA to achieve a high DPA output impedance can enhance both in-band and out-of-band distortion reduction.

4.1.2 Correction Current Reduction

In this section, the impact of $Z_{DPAt}(f)$ on the amount of current the LA has to source or sink for distortion correction is examined at an in-band frequency, f_i. The simplified schematics in Fig. 4.2 will be used for this analysis. When the LA corrects the load voltage from $V_L(f_i)$ to a more ideal value of $V_L^{'}(f_i)$, the magnitude of the current the LA will need to source or sink for this correction, $|I_{cor}(f_i)|$, can be expressed as

$$|I_{cor}(f_i)| = \frac{|V_L(f_i) - V_L^{'}(f_i)|}{|Z_{DPAt}(f_i) \parallel R_L|}. \tag{4.7}$$

From Eq. (4.7), it can be seen that $I_{cor}(f_i)$ will be a nonzero value when $V_L(f_i)$ does not equal $V_L^{'}(f_i)$. In other words, $I_{cor}(f_i)$ will be nonzero if the LA corrects either or both the magnitude and phase of the load voltage. Since most of $I_{cor}(f_i)$ will flow through the LA's transistor, the higher the magnitude of $I_{cor}(f_i)$, the higher the power consumption of the LA's transistor. Therefore, $|I_{cor}(f_i)|$ should be minimized, and this can be achieved, according to Eq. (4.7), by maximizing the value of $|Z_{DPAt}(f_i) \parallel R_L|$. To demonstrate this, $|I_{cor}(f_i)|$ has been plotted in Fig. 4.3(b) as $Z_{DPAt}(f_i)$ is swept is swept from $j1$ to $j10000\ \Omega$. As expected, $|I_{cor}(f_i)|$ is decreased as $Z_{DPAt}(f)$ is increased. For example, $|I_{cor}(f_i)|$ is decreased by a factor of about 16 as $Z_{DPAt}(f)$ is increased from $j1$ to $j100\ \Omega$.



(a) (b)

Figure 4.4: $\kappa(f_{ob})$ versus (a) $I_{LA}(f_{ob})/I_{DPAt}(f_{ob})$ and (b) phase of $I_{LA}(f_{ob}) - I_{DPAt}(f_{ob})$ (with their magnitudes being equal).

With the impedance inverter, $Z_{DPAt}(f_i)$ equals ∞, and $|Z_{DPAt}(f_i) \parallel R_L|$ equals its highest possible value of R_L. Thus, the impedance inverter can minimize the power consumption of the LA, and as a result improve the overall efficiency of the linear DPA.

4.1.3 Comparison to Linear N-Way DPAs

Before proceeding, it is worth comparing the proposed topology to a linear N-way DPA discussed in Section 2.4.5. In a N-way DPA, additional peaking amplifiers are added in parallel to the first peaking amplifier which turns on at high output powers to improve the DPA's back-off efficiency. To enhance the linearity of an N-way DPA, the gate biases of the peaking amplifiers are adjusted such that the peaking transistors' distortion currents cancel the carrier transistor's distortion current [63]. This linearization method is also similar to the DS linearization technique discussed in Section 2.4.4.

The function of the LA in this topology, on the other hand, is solely to linearize the DPA, and it is always on such that it can perform linearization across the entire signal power range. Importantly, note that a key difference between the LA and the extra peaking amplifier in the linear N-way DPA is that the output impedance of the LA is much smaller than the peaking amplifier's, which behaves like a current source with a high output impedance. The LA's lower output impedance will enable it to be more effective at linearizing the DPA across a wide operation range, and this is demonstrated in Fig. 4.4, where $\kappa(f_{ob})$ is plotted for various $I_{LA}(f_{ob})$ magnitudes and phases for the case when $Z_{LA}(f_{ob})$ is equal to $R_L/3$ and ∞. The case where $Z_{LA}(f_{ob}) = \infty$ demonstrates what happens with the N-way DPA's second peaking amplifier. Note that when the distortion

current from the DPA is equal to the negative value of the distortion current from the LA ($I_{LA}(f_{ob}) = -I_{DPAt}(f_{ob})$) then $\kappa(f_{ob}) = 0$, and there is perfect distortion cancellation regardless what the $Z_{LA}(f_{ob})$ is. However, this is difficult to achieve and maintain for a wide operation range. Hence, it is important to study what happens when $I_{LA}(f_{ob})$ deviates from $-I_{DPAt}(f_{ob})$. When $I_{LA}(f_{ob})$ deviates from $-I_{DPAt}(f_{ob})$ the magnitude of the numerator of Eq. (4.1), and hence the magnitude of $\kappa(f_{ob})$, is increased as a result. The rate at which $|\kappa(f_{ob})|$ increases with respect to the magnitude of the numerator of Eq. (4.1) is determined by the magnitude of the denominator of Eq. (4.1), and can be simply expressed as

$$\frac{d|\kappa(f_{ob})|}{d\left|1 + \frac{I_{LA}(f_{ob})}{I_{DPAt}(f_{ob})}\right|} = \frac{1}{\left|1 + \frac{R_L \| Z_{DPAt}(f_{ob})}{Z_{LA}(f_{ob})}\right|}. \tag{4.8}$$

The rate in Eq. (4.8) is equal to 1 and 0.25 when $Z_{LA}(f_{ob})$ is ∞ and $R_L/3$, respectively. Hence, $|\kappa(f_{ob})|$ will increase slower when $Z_{LA}(f_{ob}) = R_L/3$. This shows that the design of $Z_{LA}(f_{ob})$ can help keep $\kappa(f_{ob})$ low for a larger range of $I_{LA}(f_{ob})$ values. For example, observe in Fig. 4.4 that $\kappa(f_{ob}) = 1$ when the phase difference between $I_{LA}(f_{ob})$ and $I_{DPAt}(f_{ob})$ is 90° for the case when $Z_{LA}(f_{ob}) = \infty$. Thus, there is no more distortion reduction. On the other hand, if $Z_{LA}(f_{ob}) = R_L/3$, then $\kappa(f_{ob}) = 0.25$, and there is still a significant amount of distortion reduction.

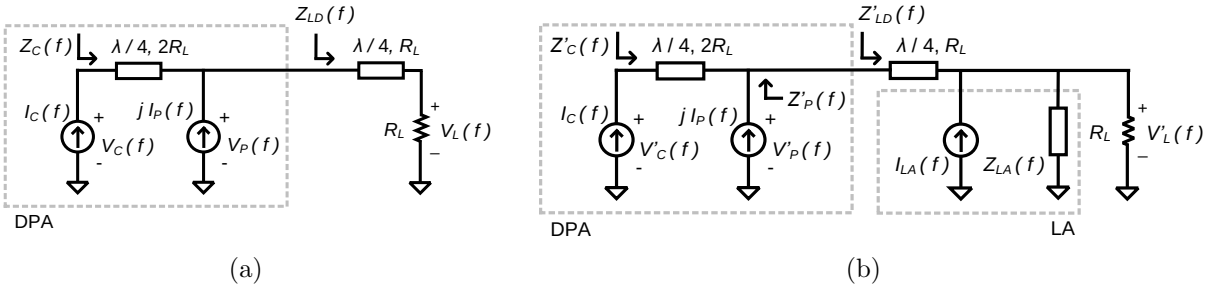
4.1.4 Effects of LA on Carrier and Peaking Transistors

When LA corrects the DPA's output it will change the load impedance seen by the DPA, $Z'_{LD}(f)$, from its original value of R_L. The effects this will have on the DPA operation will be analyzed at an in-band frequency, f_i, where the change is most critical. This analysis will be conducted using the equivalent schematics of the proposed linear DPA topology given in Fig. 4.5, where $I_P(f_i)$ and $I_C(f_i)$ are the current magnitudes of the DPA's peaking and carrier transistor, respectively, and a quarter-wave transmission line is used for the impedance inverter. The impedance seen by the DPA's carrier transistor for the case with and without the LA will be denoted as $Z'_C(f_i)$ and $Z_C(f_i)$, respectively, and they can be expressed as

$$Z_C(f_i) = 4R_L - 2R_L \frac{I_P(f_i)}{I_C(f_i)}, \tag{4.9}$$

and

$$Z'_C(f_i) = \frac{2R_L Z_{LA}(f_i)}{R_L + Z_{LA}(f_i)} \left(2 - \frac{I_{LA}(f_i)}{I_C(f_i)}\right) - 2R_L \frac{I_P(f_i)}{I_C(f_i)}. \tag{4.10}$$



(a) (b)

Figure 4.5: Equivalent circuit schematic of the DPA (a) without and (b) with the LA.

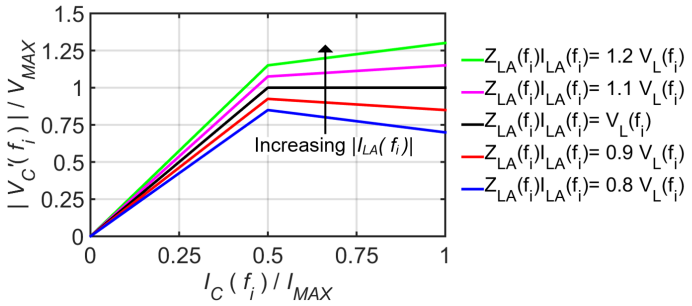


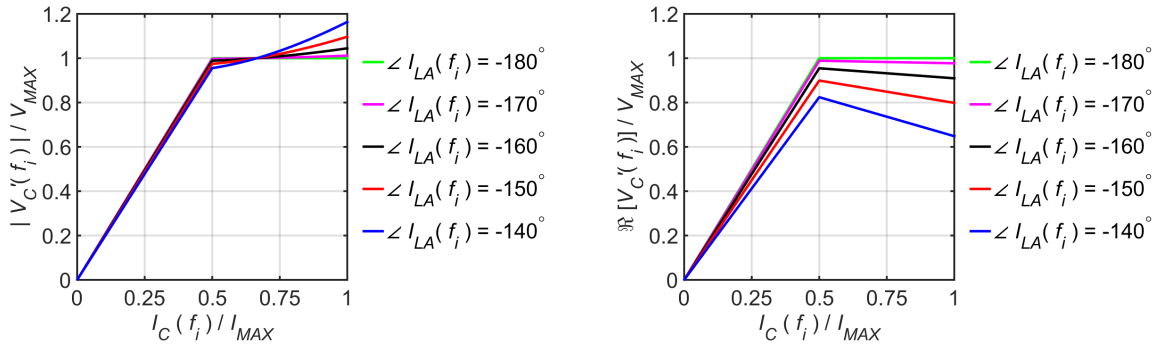
Figure 4.6: Calculated $|V_C^{'}(f_i)|/V_{MAX}$ with various magnitudes of $I_{LA}(f_i)$.

Note that $Z_C^{'}(f_i) = Z_C(f_i)$ if the open circuit voltage of the LA, given by $Z_{LA}(f_i)I_{LA}(f_i)$, is equal to the value of $V_L(f_i)$, given by $-2R_LI_C(f_i)$. In other words, the LA will not change the load seen by the carrier transistor if

$$I_{LA}(f_i) = \frac{-2R_LI_C(f_i)}{Z_{LA}(f_i)}. \tag{4.11}$$

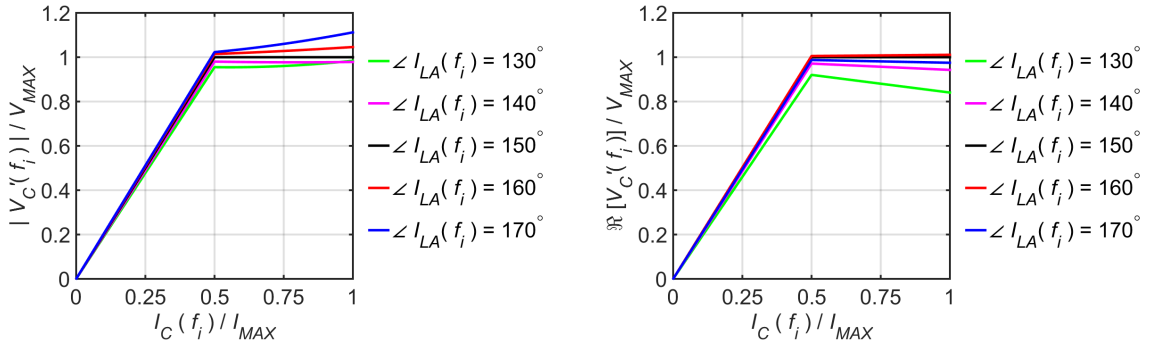
Otherwise, if $I_{LA}(f_i)$ does not satisfy the relationship given in Eq. (4.11), then the load impedance seen by the carrier transistor can be altered from its conventional value.

If only the magnitude of $Z_{LA}(f_i)I_{LA}(f_i)$ and $V_L(f_i)$ differs such that $|Z_{LA}(f_i)I_{LA}(f_i)| > |V_L(f_i)|$, then $|Z_C^{'}(f_i)| > |Z_C(f_i)|$. As a result, when the LA is added, the voltage swing at the drain of the carrier transistor will be increased. This will drive the transistor into its strongly non-linear knee region, leading to less effective distortion correction. Conversely, if $|Z_{LA}(f_i)I_{LA}(f_i)| < |V_L(f_i)|$, then $|Z_C^{'}(f_i)| < |Z_C(f_i)|$, and the voltage swing at the drain of the carrier transistor will be decreased instead. This will cause a decrease in the DPA's power efficiency. To illustrate the aforementioned effects, the voltage swing at the drain of



(a) (b)

Figure 4.7: Calculated (a) $|V_C^{'}(f_i)|/V_{MAX}$ and (b) $\Re[V_C^{'}(f_i)]/V_{MAX}$ with $Z_{LA}(f_i) = R_L/3$, phase of $I_C(f_i)$ set to 0°, and various phases of $I_{LA}(f_i)$.



(a) (b)

Figure 4.8: Calculated (a) $|V_C^{'}(f_i)|/V_{MAX}$ and (b) $\Re[V_C^{'}(f_i)]/V_{MAX}$ with $|Z_{LA}(f_i)| = |R_L/3|$, phase of $Z_{LA}(f_i)$ set to 30°, phase of $I_C(f_i)$ set to 0°, and various phases of $I_{LA}(f_i)$.

the carrier transistor after the addition of the LA, given by $V_C^{'}(f_i)$, is calculated and plotted for different magnitudes of $I_{LA}(f_i)$ in Fig. 4.6. $Z_{LA}(f_i)$ is assumed to be $R_L/3$ for this plot. As expected, $V_C(f_i)$ differs from the ideal value as the magnitude of $Z_{LA}(f_i)I_{LA}(f_i)$ deviates from $V_L(f_i)$. Note that $V_C^{'}(f_i)$ here is allowed to go beyond V_{MAX} in the calculated results, where V_{MAX} is the maximum voltage swing the transistor drain can support before going into its ohmic region. In practice, however, $V_C^{'}(f_i)$ will be more restricted because beyond V_{MAX} the transistor would be strongly saturated. When there is a phase difference between $Z_{LA}(f_i)I_{LA}(f_i)$ and $V_L(f_i)$, one of the consequence is that the phase of the first

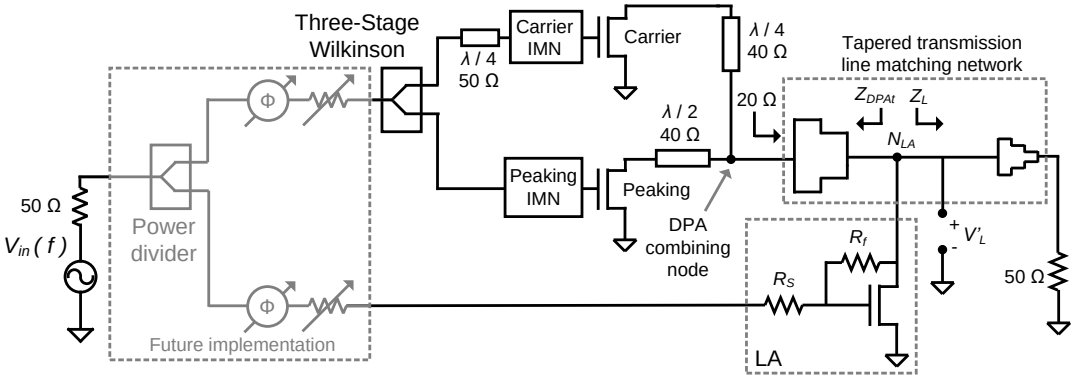


Figure 4.9: Simplified schematic of the proposed linear DPA topology.

term in Eq. (4.10) will changed such that it is out of phase with the second term, whose magnitude becomes non-zero when the peaking transistor turns on, and the two terms will no longer add in phase. This means $|V_C^{'}(f_i)|$ will no longer be constant at high powers when the peaking transistor is turned on, as is the case with conventional DPAs. This is illustrated in Fig. 4.7(a), where $|V_C^{'}(f_i)|$ is calculated and plotted for different $\angle I_{LA}(f_i)$ values for the case when $Z_{LA}(f_i) = R_L/3$. When $I_C(f_i)$ is higher than $I_{MAX}/2$, a phase difference more than 10° can significantly increase the value of $|V_C(f_i)|$, thus pushing the carrier transistor into the knee region. Furthermore, Fig. 4.7(b) shows that the real part of $V_C^{'}(f_i)$ drops significantly when there is more than 10° of phase shift. This is another consequence of varying the phase of the first term in Eq. (4.10). Varying this term's phase will change the phase of $Z_C^{'}(f_i)$ (hence the phase of $V_C^{'}(f_i)$) such that the real part of $V_C^{'}(f_i)$ will be smaller than its magnitude, and the output power of the carrier transistor can be decreased as a result. This decrease in the carrier transistor's output power will cause a drop in the efficiency of the DPA itself, and, moreover, the missing power must now be supplemented by the less efficient LA.

It should be mentioned that the change in $V_C^{'}(f_i)$ in response to a change in $\angle I_{LA}(f_i)$ is dependent on the value of $Z_{LA}(f_i)$. To demonstrate this, the plots in Fig. 4.7 are replotted in Fig. 4.8 with $\angle Z_{LA}(f_i)$ changed to 30°. Then, according to Eq. (4.11), $\angle I_{LA}(f_i)$ should be 150° such that LA does not change the load impedance of carrier transistor. Notice that the behavior of $V_C^{'}(f_i)$ in Fig. 4.8 differs from the behavior shown in Fig. 4.7. When $\angle I_{LA}(f_i)$ is increased in Fig. 4.8(b), the real part of $V_C^{'}(f_i)$ does not change by a significant amount.

The load impedance seen at the output of the peaking transistor, $Z_P(f_i)$, can be expressed as $Z_P(f_i) = 2R_L I_C(f_i)/I_P(f_i)$. Since $Z_P(f_i)$ is solely determined by the currents

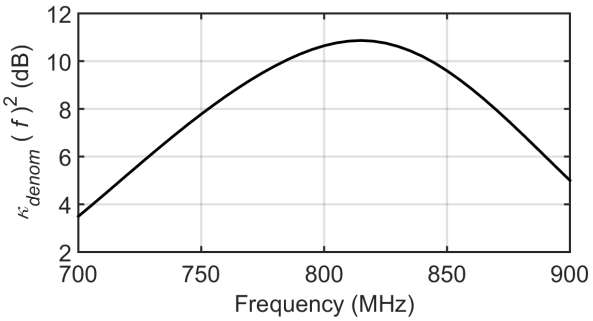


Figure 4.10: Simulated $\kappa_{denom}(f)^2$ versus frequency.

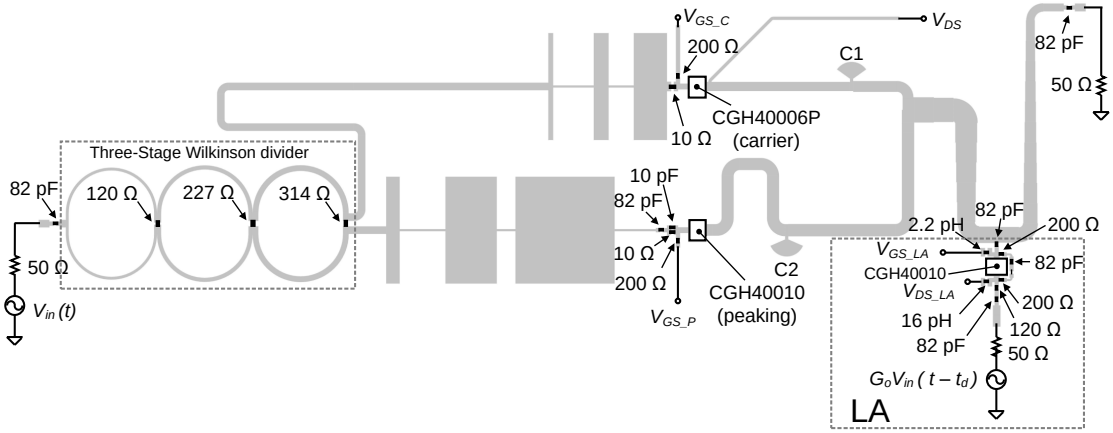


Figure 4.11: Schematic of the fabricated linear DPA.

of the carrier and peaking transistors, it should therefore not be affected by the addition of the LA.

4.2 DPA and LA design and Simulation

A prototype of the proposed linear DPA topology is designed as a proof of concept. The layout of the pototype was simulated using Keysight's Momentum EM simulator and Keysight's ADS was used for EM-Circuit co-simulation of the entire circuit. The simplified schematic of the prototype is shown in Fig. 4.9. The DPA was designed for a peak envelope power of 12 W and a centre frequency of 800 MHz. The prototype uses two Wolfspeed GaN transistors: A 6 W model (CGH40006P) and a 10 W model (CGH40010)

was utilized for the carrier and peaking transistor in the DPA, respectively. The half-wave line after the peaking amplifier was utilized to achieve wideband impedance inversion [34]. The impedance seen by the DPA at the combining node was chosen to be 20 Ω for optimal power and efficiency performance. This impedance was synthesized using the 50 Ω load impedance and a tapered transmission line, which achieves real-to-real impedance matching.

The output impedance seen looking into the DPA at the combining node is $5 + j7$ Ω at 800 MHz. Note that this is different from the ideal zero output impedance. Hence, an impedance inverter with 20 Ω characteristic impedance should not be used in practice as it would result in a Z_{DPAt}(800 MHz) value of only $27 - j38$ Ω. Instead, the LA is inserted at the node N_{LA}, where the tapered transmission line between N_{LA} and the combining node transforms the output impedance of the DPA to a high impedance value of Z_{DPAt}(800 MHz) $= 123 + j4$ Ω. Recall from Section 4.1.1 and 4.1.2 that $Z_{DPAt}(f)$ should be maximized for better distortion reduction and lower LA power consumption.

The LA design uses feedback to control its output impedance and improve its linearity as discussed in the previous chapter. The size of the LA's transistor was chosen with the following considerations. First, the LA transistor's maximum current should be high enough so that it can supply the current needed to correct the DPA's distortions. Particularly at peak power levels, when the most amount of correction current is required. It is also important to note that the larger the transistor size, the higher its transconductance. Recall that a high transconductance can lower both the LA's output impedance and output distortion. However, the transistor's parasitic capacitance will be larger as well with a larger device, and it will be harder to limit their impacts on the linearization performance. In light of this trade-off, a 10 W GaN transistor (CGH40010) was chosen, and a 2.2 and 16 pH inductor is added at the gate and drain of the transistor, respectively, to resonate out its input and output capacitance. R_f, R_s, and the bias of the LA transistor are used to control $Z_{LA}(f)$ and maximize $\kappa_{denom}(f)$, where $\kappa_{denom}(f)$ is given by

$$\kappa_{denom} = 1 + \frac{Z_L(f) \parallel Z_{DPAt}(f)}{Z_{LA}(f)}, \tag{4.12}$$

and is the denominator of $\kappa(f)$ and $\kappa_d(f)$ given in equation (4.1) and (4.6), respectively. A high $\kappa_{denom}(f)$ value means that the LA's output impedance is conducive to distortion reduction, and can reduce both $\kappa(f)$ and $\kappa_d(f)$. However, there is a trade-off between how high κ_{denom} is and the power consumption of the LA. For example $Z_{LA}(f)$, and hence $\kappa_{denom}(f)$, can be reduced by increasing the gate bias of the LA's transistor, which increases its transconductance. This, however, would increase the power consumption of the LA as well. $Z_{LA}(f)$ can also be reduced through decreasing the ratio between R_f and R_s, but

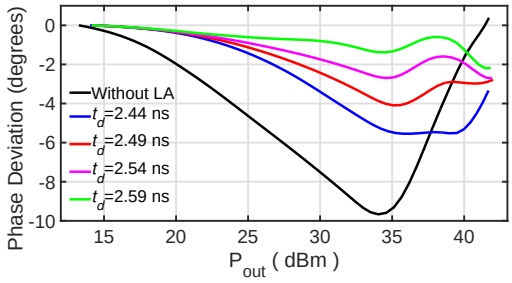


Figure 4.12: Simulated phase of $V_L(800\text{ MHz})$ and $V_L^{'}(800\text{ MHz})$ for different t_d values.

this would lower the gain of the LA, and increase the power required to drive the LA. In the final design $Z_{LA}(800\text{ MHz}) = 8.1 + j4.6\ \Omega$, which results in a $\alpha(800\text{ MHz})$ magnitude of 3.33. Note here that $\kappa(f)$ and $\kappa_d(f)$ affects the distortion voltage, and should be squared to see its effect on the distortion power. Hence κ^2_{denom} is plotted in Fig. 4.10 for frequencies between 700 and 900 MHz. κ^2_{denom} remains high for frequencies around 800 MHz.

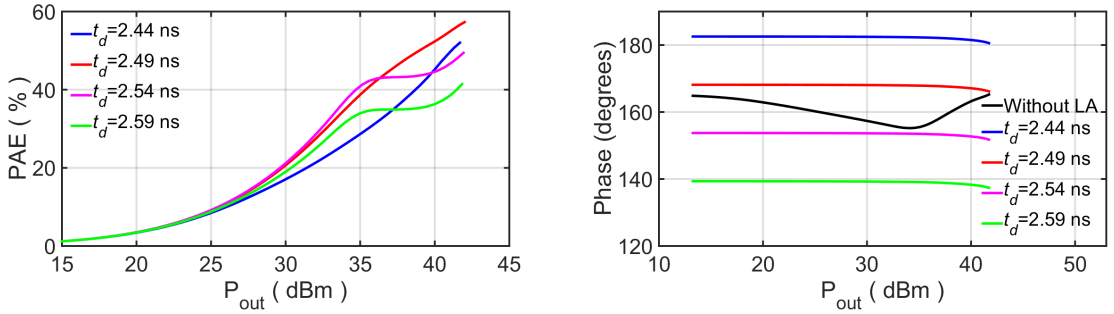
In the final design $Z_L(800\text{ MHz})$ equals $26.7 - j6\ \Omega$. This impedance is higher than the load impedance seen by the DPA at the combining node. Thus the voltage swing at N_{LA} is higher than the voltage swing at the combining node, and the LA transistor's drain must be biased higher than the peaking transistor to avoid being pushed into the knee region.

The schematic of the final linear DPA prototype design is given in Fig. 4.11. C1 and C2 are used to account for the parasitic output capacitance of the carrier and peaking transistor. C1, the carrier transistor's output capacitance, and the transmission line between the two are used to implement a quasi-lumped quarter-wave transmission line [71]. Similarly, C2 is used to implement a quasi-lumped quarter-wave transmission line at the output of the peaking transistor, which is then connected to a quarter-wave transmission line to form a half-wave transmission line.

4.2.1 Linearization with CW Stimulus

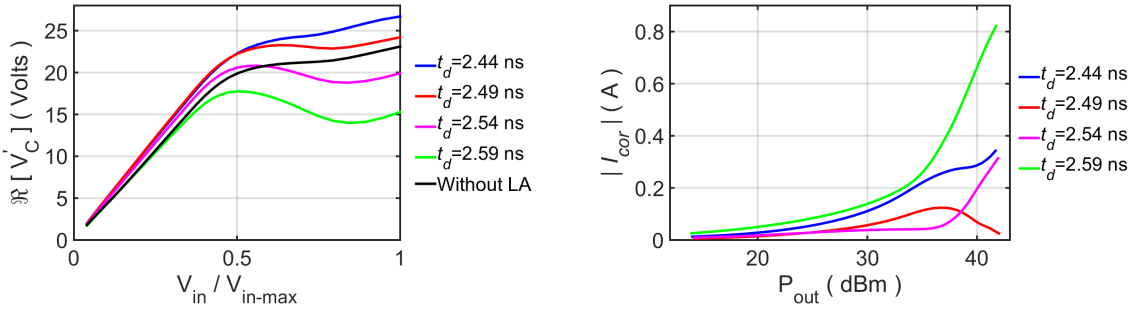
First the DPA and LA combination was given a CW input at the centre frequency of 800 MHz, and the value of $I_{LA}(800\text{ MHz})$ was adjusted through the setting of t_d and G_o. The following guidelines were followed in light of the LA's effects on the impedance seen by the carrier transistor.

The phase of $I_{LA}(800\text{ MHz})$ was adjusted by tuning the value of t_d to achieve good phase linearization without significantly decreasing the DPA's PAE. According to Fig. 4.12, the



(a) (b)

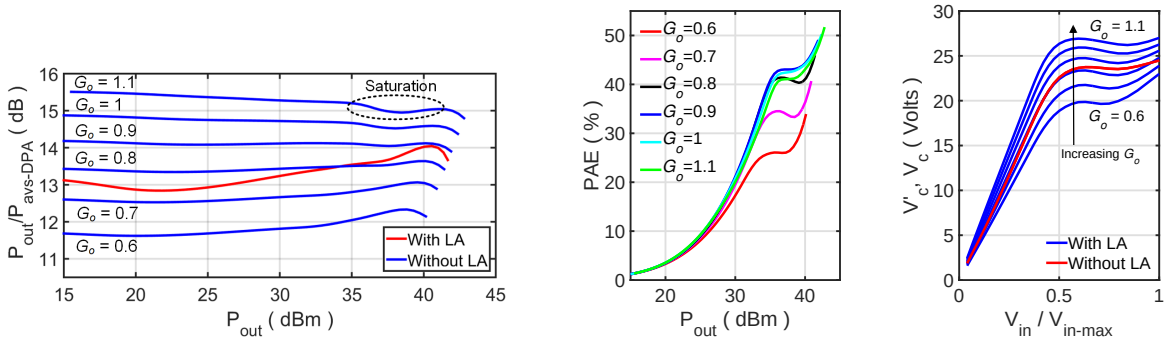
Figure 4.13: Simulated (a) PAE, of the linear DPA, with the LA's power consumption factored in, for various t_d values, and the (b) phase of LA's open circuit voltage and V_L(800 MHz) for different t_d values.



(a) (b)

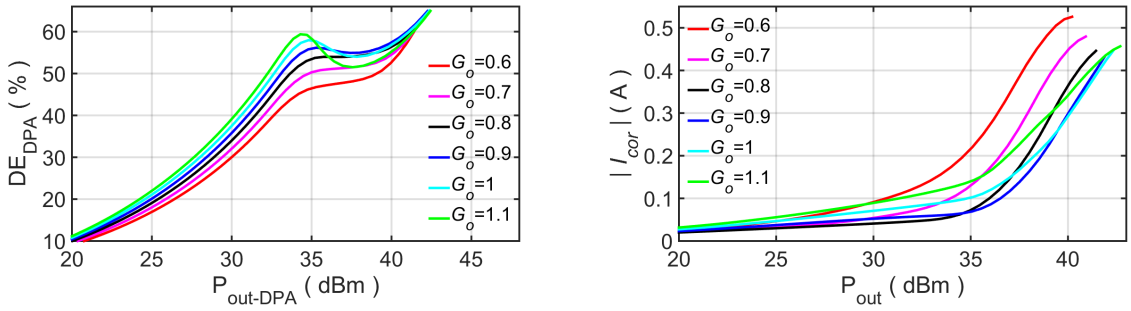
Figure 4.14: Simulated (a) real part of $V_C^{'}$(800 MHz) and (b) magnitude of the correction current from the LA for different t_d values.

DPA output will have the least amount of phase deviation when t_d is 2.59 ns. However, the PAE of the DPA is low at this t_d setting, and the reason for this can be seen in Fig. 4.13. The phase of the LA's open circuit voltage is about 20° lower than V_L(800 MHz) when t_d is 2.59 ns, and this can cause a change in the carrier transistor's load impedance once the LA is added. This is confirmed in Fig. 4.14, where the real part of $V_C^{'}$(800 MHz) is plotted for various t_d settings. Note that Fig. 4.14(a) shows similar behavior as the plot in Fig. 4.8(b) because the phase and magnitude of Z_{LA}(800 MHz) is about the same as the phase and magnitude of $Z_{LA}(f_i)$ assumed in Fig. 4.8. According to Fig. 4.14, when t_d is 2.59 ns



(a) (b) (c)

Figure 4.15: Simulated (a) $P_{out}/P_{avs\text{-}DPA}$, where $P_{avs\text{-}DPA}$ is the power available at the input of the DPA, the (b) DPA's PAE, with the LA's power consumption factored in, for various G_o settings, and (c) the magnitude of $V_C^{'}$(800 MHz) for various G_o settings.



(a) (b)

Figure 4.16: Simulated (a) DE of the DPA (without the LA consumption), DE_{DPA}, versus the output power of the DPA only (without the power from the LA), $P_{out-DPA}$ and (b) magnitude of the correction current from the LA for different G_o values.

the real part of $V_C^{'}$(800 MHz) is significantly decreased. Hence, as previously discussed, the overall efficiency will be decreased as well. In Fig. 4.13(a), it is also shown that the PAE is low when t_d is at its lowest sweep value of 2.44 ns. This is because although the LA will not have to supply as much real power, it will need to supply a large amount of reactive current when the phase difference between the LA's open circuit voltage and V_L is large. This trend can be seen in the plot of the correction current coming from the LA, I_{cor}(800 MHz), in Fig. 4.14(b).

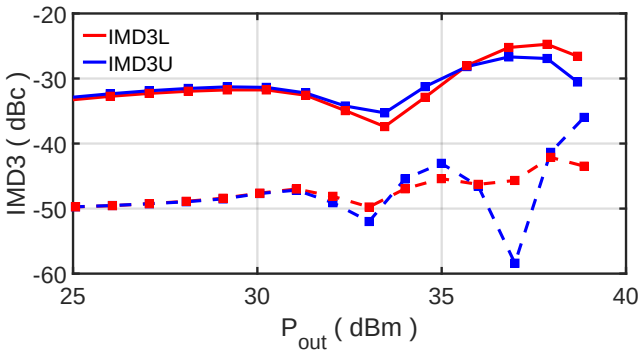


Figure 4.17: Simulated IMD3 with (solid line) and without (dashed line) the LA with input tones at 795 and 805 MHz

The value of G_o was set such that $|I_{LA}(800\text{ MHz})|$ does not saturate the carrier transistor or significantly decrease the DPA's efficiency. In Fig. 4.15(a), the G_o is swept from 0.6 to 1.1. It can be observed that the DPA's gain, with the LA, can experience compression at high power when G_o is close to 1.1. On the other hand, when G_o is below 0.7 the PAE of the DPA will drop rapidly as shown in Fig. 4.15(b). This can be explained using Fig. 4.15(c), which plots the voltage swing at the carrier transistor's drain. Fig. 4.15(c) shows a behavior which is similar to the one predicted in Fig. 4.6, where the voltage swing $V_C'(800\text{ MHz})$ increases with G_o. When G_o is too high, the carrier transistor is pushed into saturation, hence the linearization is not as effective. On the other hand, if G_o is too low, the voltage swing at the carrier transistor will be decreased, and the DPA will no longer achieve its maximum efficiency. This can be seen in Fig. 4.16(a), where it can be observed that the DE of the DPA (without the LA's power consumption) decreases with the value of G_o. Moreover, Fig. 4.16(b) shows that the further away G_o deviates from 0.8 and 0.9, the higher the amount of correction current from the LA. This is because when G_o is very high or low the LA will increase or decrease the overall gain by sourcing or sinking a large amount of current, respectively.

4.2.2 Linearization with Two-Tone Stimulus

The simulations with CW signal gave a range for what t_d and G_o should be within. The DPA and LA combination was then stimulated by a two-tone input with tone frequencies at 795 and 805 MHz, and the IMD3 is minimized through fine tuning of $I_{LA}(f)$ and $Z_{LA}(f)$ through the adjustment of t_d, G_o, and the LA's transistor's bias. The simulated IMD3 with

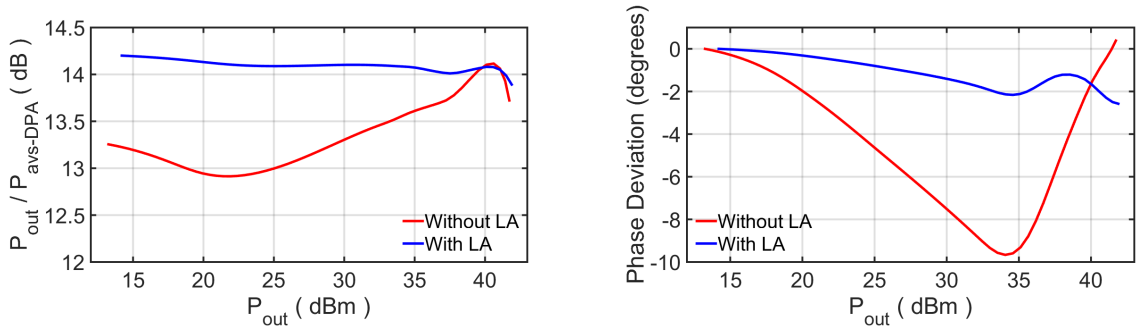


Figure 4.18: Simulated (a) $P_{out}/P_{avs\text{-}DPA}$ and (b) phase deviation of the output voltage for the case with and without the LA and t_d and G_o of 2.56 ns and 0.9, respectively.

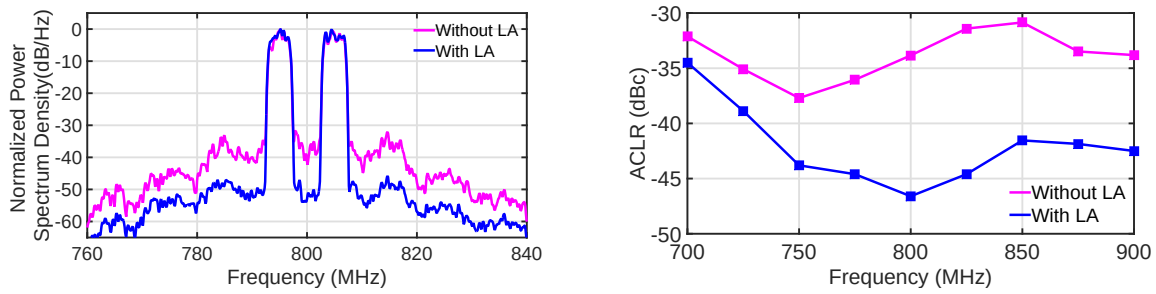
and without LA is plotted in Fig. 4.17 for a t_d and G_o value of 2.56 ns and 0.9, respectively. It can be observed that significant improvement was achieved at all power levels.

The DPA's gain and output phase with an 800 MHz CW input signal was then simulated again with the refined t_d and G_o values of 2.56 ns and 0.9, respectively, and the results are plotted in Fig. 4.18. With the LA, the variation in gain has been reduced from 1.2 to 0.3 dB, and the phase variation is decreased from 10.1° to 2.5°.

4.2.3 Linearization Under Modulated Signals

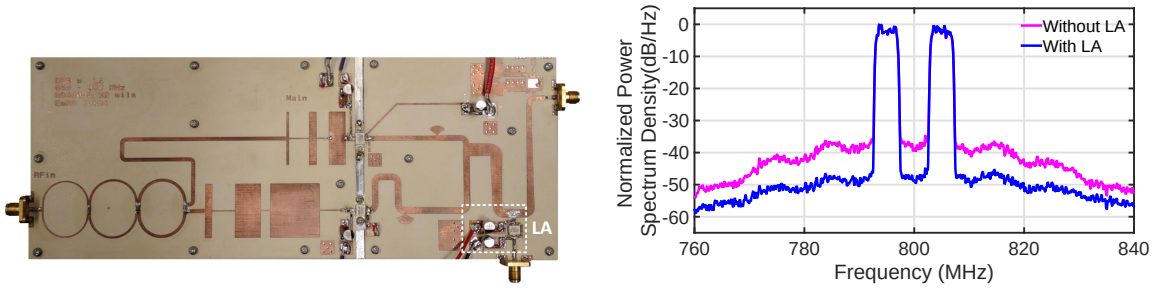
Next, the DPA was driven by a modulated signal in simulation using the design values found in the two-tone signal linearization subsection. The design was simulated using Keysight's ADS RF and Ptolemy co-simulations. A WCDMA modulated signal with a bandwidth of 15 MHz and PAPR of 8.3 dB was used to drive the PA, and the average output power was kept at 33 dBm. The frequency spectrum at the output of the DPA is plotted in Fig. 4.19 for the cases with and without the LA. With the LA, the ACLR improved by 13.7 dB from -32.1 to -45.8 dBc, and the EVM was improved from 4.9 % to 1.5 %. The PAE, however, was decreased from about 44 % to 30 % when the LA was added. This efficiency drop is mainly due to the extra power required by the LA.

The co-simulation was then repeated at different centre frequencies from 700 to 900 MHz. The value of t_d and G_o were adjusted to account for the change in the DPA and LA's gain and group delay versus frequency, and the ACLR performance at each frequency is plotted in Fig. 4.19(b). Note from Fig. 4.19(b) that the most amount of linearization is achieved



(a) (b)

Figure 4.19: Simulated DPA (a) output spectrum and (b) output ACLR versus frequency when the DPA is driven by a 15 MHz signal with 8.3 dB PAPR.



(a) (b)

Figure 4.20: The (a) picture of fabricated fully analog linear DPA and (b) measured DPA output spectrums when stimulated by a 15 MHz modulation bandwidth signal with 8.3 dB PAPR and 32 dBm output power.

around 800 MHz as predicted by the plot of $\alpha(f)^2$ in Fig. 4.10.

4.3 Measurements

A photograph of the fabricated linear DPA prototype is shown in Fig. 4.20(a). The bias current of the carrier transistor is 33 mA and the peaking transistor's gate is biased at -5.8 V. The drain bias of the carrier and peaking transistor is 28 V and the LA transistor's drain bias voltage and current is 34 V and 60 mA, respectively. Keysight's M8190A AWG was used to generate the input signals to the DPA and LA, and t_d and G_o were kept at

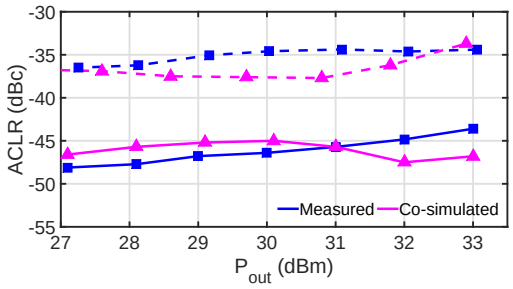


Figure 4.21: Measured and co-simulated ACLR when the DPA is stimulated by a 15 MHz modulation bandwidth signal with 8.3 dB PAPR for the case with (solid line) and without (dashed line) the LA.

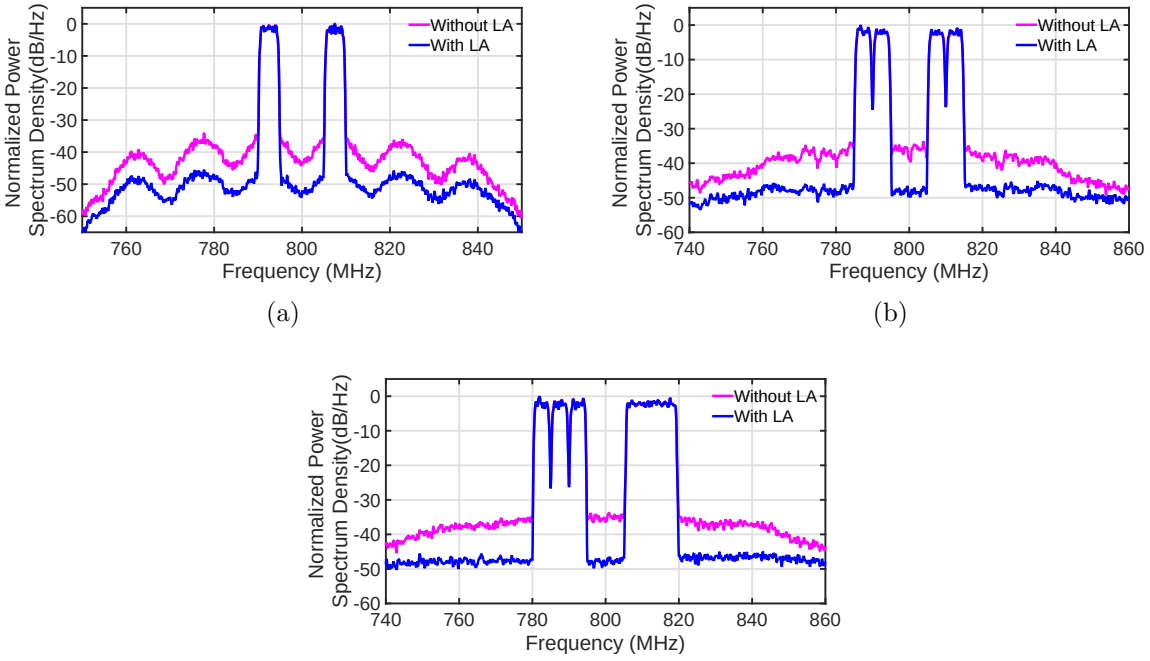


Figure 4.22: Measured DPA output spectrums when stimulated by a (a) 20 MHz, (b) 30 MHz and (c) 40 MHz WCDMA signal at the average output power level were -45 dBc of ACLR is achieved.

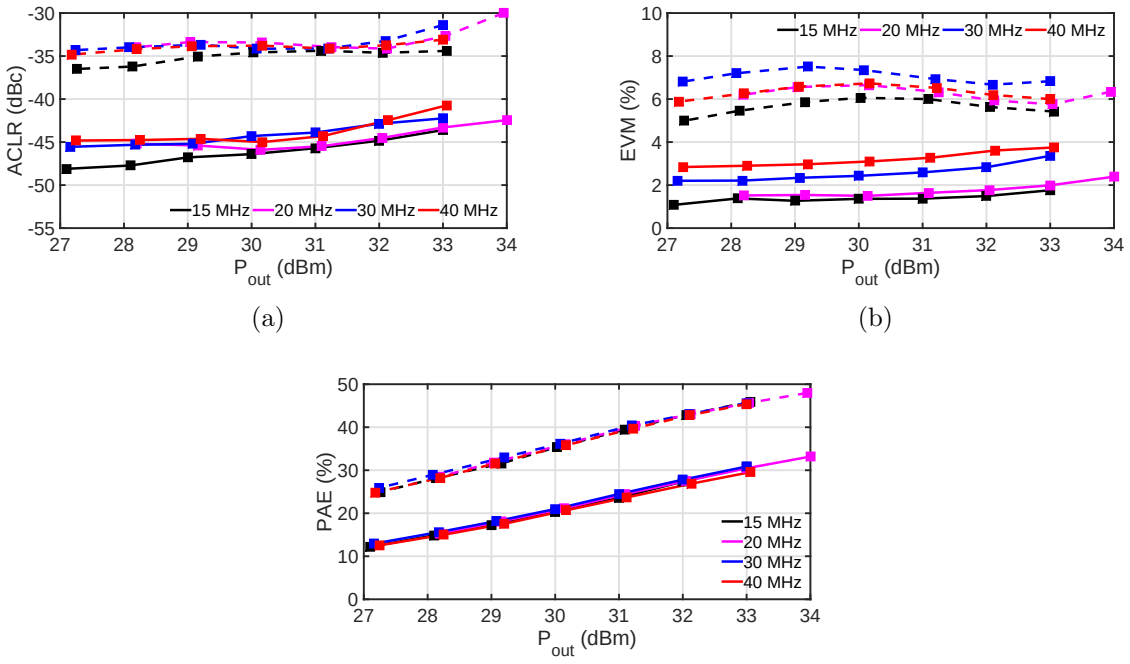


Figure 4.23: Measured (a) ACLR, (b) EVM, and c) PAE of the DPA with (solid line) and without (dashed line) the LA versus average output power. Note that the PAE figures with LA include the power consumption of the LA.

2.52 ns and about 0.9, respectively, for all the measurements. In this proof of concept test, the inputs to the DPA and LA are generated by separate AWGs. It is possible though to generate the inputs to both the LA and DPA from a single input signal source by using a power divider, variable phase shifters and variable attenuators as shown in Fig. 4.9. Moreover, the same phase shifters and attenuators can also be utilized in the future for beamforming.

The performance of the prototype was first measured using the same 15 MHz input signal used in the co-simulation. The output spectrum for the case with and without the LA is plotted in Fig. 4.20(b), where the average output power was 32 dBm. Note that ACLR was improved by 10.3 dB from -34.7 dBc to about -45 dBc, and the EVM was improved by 4.1 percentage points from 5.6 % to 1.5 %. This improvement is very comparable to the co-simulation results. In fact, even at other power levels, the co-simulated linearization results are close to the measurement results as can be seen in Fig. 4.21. Significant linearization

Table 4.1: Maximum Linearity Improvement with the LA

Signal bandwidth (MHz)	15	20	30	40
ACLR improvement (dB)	11.8	12.5	11.5	11.2
EVM improvement (percentage points)	4.7	5.2	5.2	3.6

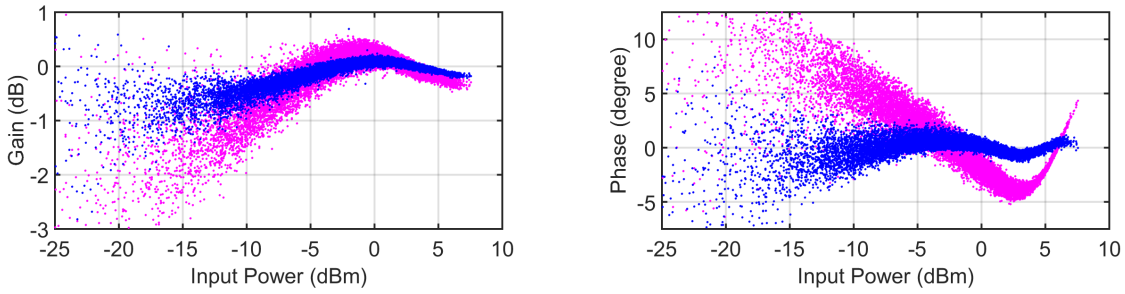
was achieved at all power levels. This, as mentioned in Section 4.1.3, is due to the low $Z_{LA}(f)$ value achieved in the design which can help enable distortion reduction across a wide range of operation conditions.

The DPA output was also measured at different average output power levels with up to 40 MHz bandwidth modulated signals. The PAPR of the 20, 30, and 40 MHz signal is 7.1, 8.6, and 8.4 dB, respectively. The output spectrums of the DPA with the different signals tested are plotted in Fig. 4.22 for the case with and without the LA. The ACLR, EVM, and PAE versus different average power levels are summarized in Fig. 4.23. After the addition of the LA, significant linearity improvement was achieved at all power levels. The maximum ACLR and EVM improvements are summarize in Table 4.4 for the different signals. More than 11 dB of maximum ACLR improvement is achieved for all the signals. The PAE, however, is reduced due to the LA's power consumption. Note though that the LA's power consumption is still much less than the power consumption of a conventional DPD system. For example, the LA consumes about 2 W of power at 32 dBm average output power. This is much less than the power consumption of a DPD system, which is estimated to be 5.4 W in [12] for a 30 MHz bandwidth system. Moreover, the LA's power will scale down with the power rating of the PA as discussed in the previous chapter. Hence, the lower the DPA's power rating, the more beneficial the present topology is in terms of power consumption. It is also worth mentioning here that the PAE of the topology did not increase with an increase in signal bandwidth. Lastly, it is important to point out that the efficiency of the linear DPA is significantly higher than the efficiency of the linearized class AB PA in the last chapter. For example, with the same 40 MHz signal the PAE of the linear DPA is more than 13 percentage points higher than the linearized class AB PA at the maximum output power level.

As a result of the LA's linearization, the ACLR now passes the -45 dBc requirement for all the modulated signals without a large back-off in average output power. On the other hand, when the LA was not used, the ACLR stays higher than -36.5 dBc for all the average output powers tested. Table 4.2 summarizes the performance of the linear DPA at the average output power levels where -45 dBc ACLR is achieved, and the corresponding output spectrums are plotted in Fig. 4.22. For example, when driven by the 15 MHz

Table 4.2: DPA Performance When -45 dBc ACLR was Achieved

Signal bandwidth (MHz)	15	20	30	40
PAPR (dB)	8.3	7.1	8.6	8.4
Average output power (dBm)	32	31.5	29	30
ACLR improvement with the LA (dB)	10.2	11.5	11.5	11.2
EVM (%)	1.5	1.7	2.3	3.1
PAE (%)	27.4	25.9	19	20.7



(a) (b)

Figure 4.24: Measured (a) gain and (b) phase distortion when stimulated by a 15 MHz WCDMA signal for the case with (blue) and without (magenta) the LA.

signal, the ACLR is now better than -45 dBc when the average output power is 32 dBm, and the PAE at that power level is 27.4 %. The gain and phase distortion of the DPA are also measured and plotted in Fig. 4.24. Both gain and phase distortions are significantly decreased with the LA, and the EVM was improved from 5.6 % to 1.5 %.

The performance of the prototype is compared with other linear DPA designs in Table 4.3. All results are achieved without the use of DPD, and the designs that reported better than -45 dBc ACLR are highlighted in Table 4.4. Out of these designs, this work demonstrated the highest modulation bandwidth, especially relative to the centre frequency. Although the designs in [65] and [53] achieved better PAE, they required either drain and gate tracking versus average power, or the predistorter and gate bias needed to be carefully tuned for each power level.

Table 4.3: Comparison with Other Linear DPAs

	Topology	Centre freq. (GHz)	Mod. BW. (MHz)	Average P_{out} (dBm)	PAE (%)	ACLR (dBc)
[47]	2-way w. FFW	2.14	5	44.8	10.9	-60
[53]	2-way w. RFPD[1]	2.14	20	40.8	34.8[4]	-45
[54]	2-way w. RFPD[2]	2.14	5	34	18.9	-53.8
[55]	3-way w. RFPD	2.14	5/15	40/40	10.4/10.3	-51/-49.1
[65]	2-way w. DS[3]	2.14	5	39.4	31.8	-45
[63]	2/3/4-way	2.14	5 and 10	32/31/34	27.9/14.8/19.3	-30.1/-43.2/-38.4
[62]	2-way	2.14	20	44.8	26.9	-43.2
[67]	2-way	2.14	5/10/20	35.4/35.4/35.4	40/40/40	-41.5/-40.8/-40.5
[68]	2-way	7	56	32	40	-36
[66]	2-way	5	40/80/160	32	33.8[4]	-43.8/-44.1/-43.1
This work	2-way with LA	0.8	15/20/ 30/40	33/34/33/33	30.5/33.2/ 30.9/29.6	-44.6/-42.5/ -42.3/-40.7

Table 4.4: Performance When -45 dBc ACLR is achieved

	Centre freq. (GHz)	Mod. BW. (MHz)	Average P_{out} (dBm)	PAE
[47]	2.14	5	47	14
[65]	2.14	5	39.4	31.8
[53]	2.14	20	40.8	34.8[4]
[55]	2.14	5/15	42.6/42.2	15.3/14.2
[54]	2.14	5	36	27
This work	0.8	15/20/30/40	32/31.5/29/30	27.4/25.9/19/20.7

4.4 Conclusion

In this section, a linear DPA was proposed by augmenting conventional Doherty techniques with an LA amplifier. A prototype was fabricated and measurements confirmed the topology's capability for linear and efficient amplification with wideband modulated signals with up to 40 MHz bandwidth. For example, when stimulated by a 15 MHz WCDMA signal, the linear DPA achieved -45 dBc of ACLR at 32 dBm output power with 27.4 % PAE. In fact, the proposed linear DPA was able to achieve about -45 dBc of ACLR or better

[1]The predistorter and the gate bias of the peaking transistor was carefully tuned for each power level
[2]The predistorter uses multiple nonlinear paths
[3]The drain and gate bias are controlled versus average power
[4]Estimated from DE and gain plot

over a wide average power range for all the modulated signals. Compared to the linearized class AB PA in the last chapter, the linear DPA is able to achieve a significantly higher efficiency for all signals tested. And, because it still uses the LA for linearization, the topology's power overhead will scale down with the PA's output power range and does not increase with signal bandwidth. All the aforementioned reasons make this linear DPA very promising for future 5G small cells and large antenna arrays.

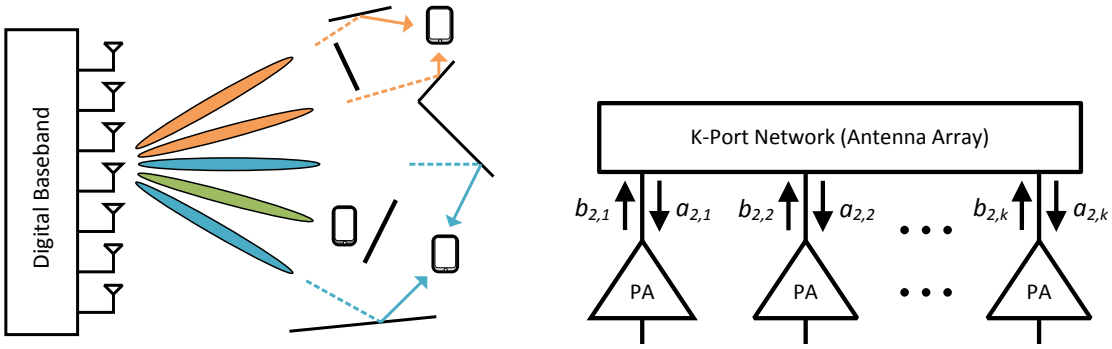
Chapter 5

Desensitizing the PA to the Effects of Antenna Coupling

As discussed in Chapter 1, it is envisioned that future 5G systems will make use of massive MIMO technology at sub-6-GHz frequencies to increase the capacity of the network without requiring additional base stations or spectrum. This is achieved by implementing digital beamforming using a large antenna array where each antenna has its own transceiver and data converter. This way, any signal can be sent from any antenna, and this gives digital beamsteering a lot of flexibility. As a result, a number of beams can now be superimposed together and adapted to multi-path and frequency-selective fading such that many users can share the same time-frequency resource as illustrated in Fig. 5.1(a). Notice how the propagation paths can be line-of-sight or non-line-of-sight. For the case of line-of-sight propagation, a narrow beam is formed to concentrate power in the direction of the user. For non-light-of-sight propagation, the transmitted waveforms are chosen such that they superimpose constructively at the user's location [72].

5.1 Antenna Coupling in Massive MIMO and Its Effects on the PA's Load Impedance

In practice, there will not be perfect isolation between the antenna elements in a large and dense antenna array. As a result, unwanted coupling between the antenna elements can cause significant distortions at the output of the PAs and at the receiver in the far-field of the array antenna that cannot be compensated for using conventional linearization



(a) (b)

Figure 5.1: Illustration of a) digital beamforming reproduced from [8], where different colors represent different signals, and b) an antenna array with multiple PAs reproduced from [73].

techniques [74, 75]. Consider the large antenna array system illustrated in Fig. 5.1(b). The wave coupled to each PA is given by

$$a_{2,i} = \sum_{k=1}^{K} S_{i,k} b_{2,k}, \tag{5.1}$$

where $a_{2,i}$ is the reflected wave at the ith device, $b_{2,k}$ is the incident wave at the output of the kth device, and $S_{i,k}$ is the scattering parameters of the antenna array. The reflection coefficient seen by each PA can then be found to be

$$\Gamma_{L,i} = \sum_{k=1}^{K} S_{i,k} \frac{b_{2,k}}{b_{2,i}}. \tag{5.2}$$

As can be seen from Eq. (5.2), the load seen by each PA is a function of the outputs of the other PAs in the array and the mutual coupling of the antennas [73]. For example, we can expect the load impedance seen by each PA to vary dynamically in a massive MIMO system because the output from each PA changes depending on user locations and the environment. This unwanted load impedance variation can affect the linearity performance of the PAs and can be particularly troublesome for DPAs. As discussed in the previous chapter, the dynamic load modulation in the DPA (and hence its linearity), is highly sensitive to its load impedance. For example, a change in the load impedance can increase the impedance seen by the DPA's carrier transistor. This could, in turn, push it into its strongly nonlinear knee region over a large range of power levels.

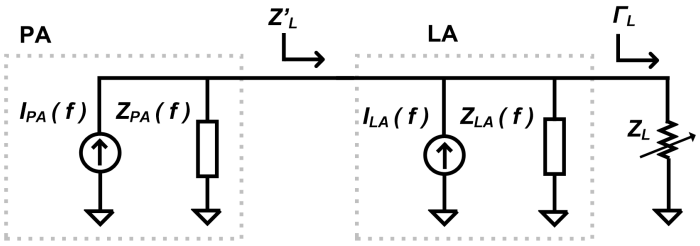


Figure 5.2: Equivalent circuit of the PA and LA used for load impedance variation analysis.

While it is possible to mitigate the effects of antenna coupling by adding circulators between the PAs and the antennas to better isolate the two, this approach is not ideal because circulators are bulky and expensive. Alternatively, one can try to linearize each PA individually using DPD techniques which accounts for antenna coupling. For example, a dual-input DPD model was developed in [76] to compensate for PA nonlinearity, antenna crosstalk, and mismatch. This method, however, can be inefficient for large antenna array systems because the cost of DPD would be multiplied by the large number of PAs in the system. In light of this, there has been recent work to linearize the entire array using only a single polynomial-based DPD [77] or to model and linearize the array as a single-input-single-output system [78]. These techniques, however, either don't consider the crosstalk or mismatch [77], or they rely on the use of a reference antenna in the antenna's far-field which is not always practical. Moreover, these linearization techniques were only designed for array antenna systems with analog beamforming, where only the phases of the input signals are varied. In digital beamforming in massive MIMO systems, the amplitude of each input signal is also dynamically changed to produce the beams required. This can further complicate linearization.

In response to the problem of antenna coupling in large antenna arrays, the following subsections will analyze the LA's capacity to counteract its undesirable effects.

5.2 Analysis of Load Impedance Variation with LA

In this subsection, it will be demonstrated that the LA can be used to significantly reduce the amount of load impedance variation caused by antenna coupling in a massive MIMO system. This will be shown using the the simplified schematic in Fig. 5.2. Note that Z'_L is the impedance seen looking into the LA and the load impedance Z_L, and it is the impedance seen by the PA when the LA is added. Z'_L can be found to be

$$Z'_L = \frac{(\frac{I_{PA}}{I_{LA}} + 1)Z_{PA}Z_{LA}Z_L}{\frac{I_{PA}}{I_{LA}}Z_{PA}(Z_L + Z_{LA}) - Z_L Z_{LA}}. \tag{5.3}$$

Assuming that G in eq. (3.6) is set to 0 dB[1], such that $Z'_L = Z_L$ when Z_L is equal to the optimal load impedance of the PA Z_{Lo}, then Z'_L can be found to be

$$Z'_L = \frac{Z_L(Z_{LA}Z_{PA} + Z_{LA}Z_{Lo} + Z_{PA}Z_{Lo})}{(Z_{LA}Z_{PA} + Z_{LA}Z_{Lo} + Z_{PA}Z_L)}. \tag{5.4}$$

Next, to see how Z'_L changes when the load impedance deviates from its original value due to antenna coupling, the derivative of Z'_L with respect to Z_L is found at the point when $Z_L = Z_{Lo}$ to be

$$\frac{d(Z'_L)}{d(Z_L)}(Z_{Lo}) = \frac{1}{1 + \frac{Z_{Lo}Z_{PA}}{Z_{LA}(Z_{Lo}+Z_{PA})}}. \tag{5.5}$$

If the magnitude of the numerator in Eq. (5.5) is greater than one, then the magnitude of the impedance variation seen by the PA is effectively reduced by the addition of the LA. It can also be seen that the amount of desensitization depends on the value of Z_{LA}. For example, if $Z_{LA} = 0$, then the derivative in Eq. (5.5) equals zero, and Z'_L will not vary at all in response to a change in Z_L. This is expected because LA would be an ideal voltage source in that case, and the voltage at the PA's output is constant no matter what Z_L is. As a result, the impedance seen by the PA remains constant because the current at the output of the PA also stays the same.

For the case of the DPA in chapter 4, the magnitude of the derivative in Eq. (5.5) can be solved to be about 0.31. Hence, we should expect Z'_L to vary much less than Z_L. To illustrate this, Z'_L is calculated using Eq. (5.4) for different Z_L values and plotted in Fig. 5.3. For the variation Z_L, $|\Gamma_L|^2$ was set such that the voltage standing wave ratio (VSWR) value, $VSWR_L$[3], is equal to values from 1 to 4, and the phase of Γ_L was varied over 360°. Notice how Z'_L stays much closer to the optimal impedance Z_{Lo} (the centre of the Smith chart) compared to Z_L. Hence, it is expected that the LA can be very effective in reducing the amount of load impedance variation seen by the PA in the presence of antenna coupling.

To verify the analysis above, the DPA and LA topology detailed in chapter 4 is simulated at a low power level (when the peaking transistor is off), and its load impedance Z_{load} is

[1]This is achieved by setting $\frac{I_{PA}}{I_{LA}} = \frac{Z_{LA}(Z_{PA}+Z_{Lo})}{Z_{PA}Z_{Lo}}$

[2]$\Gamma_L = (Z_L - Z_{Lo})/(Z_L + Z_{Lo})$

[3]$VSWR_L = (1 + |\Gamma_L|)/(1 - |\Gamma_L|)$

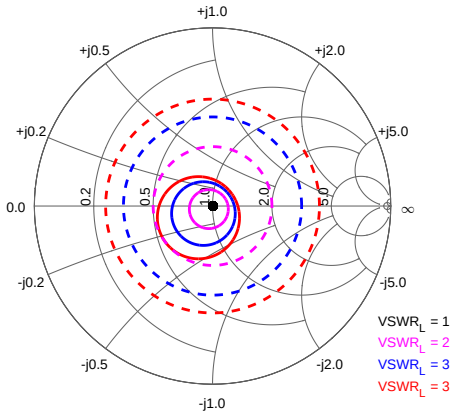
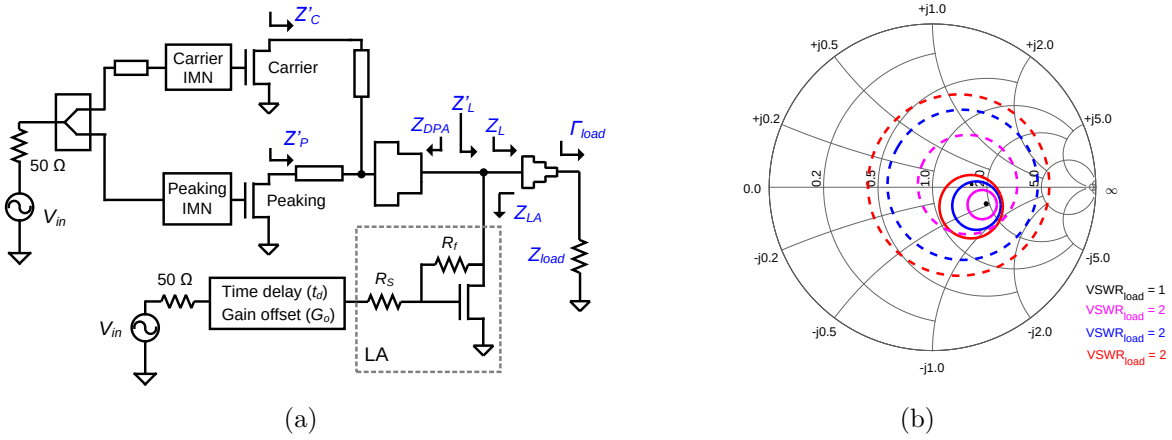


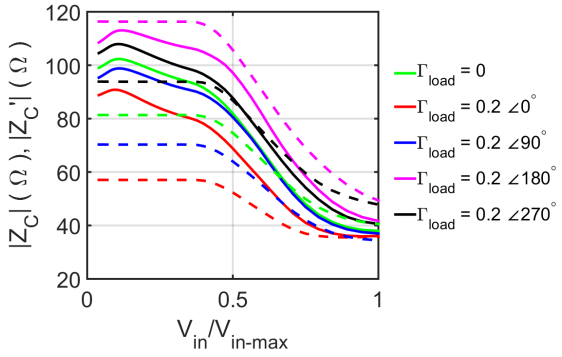
Figure 5.3: Calculated Z'_L (solid line) and Z_L (dashed line) for various $VSWR_L$ values. The impedances are plotted on a Smith chart that is normalized to Z_{Lo}.



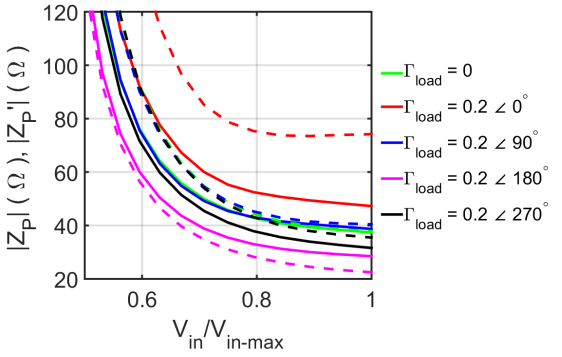
(a) (b)

Figure 5.4: The (a) topology of the DPA and LA and the (b) simulated carrier transistor impedances for various $VSWR_{load}$ values for the case with (solid line) and without (dashed line) the LA. The impedances are plotted on a Smith chart that is normalized to 50 Ω.

varied across a wide range of values. The topology of the DPA and LA is redrawn in Fig. 5.4(a) with all the relevant impedance marked, and the impedance seen by the carrier transistor is plotted in Fig. 5.4(b) for various Z_{load} values for the case with (Z_c) and without (Z'_c) the LA. For the variation of Z_{load}, the magnitude of the reflection coefficient,

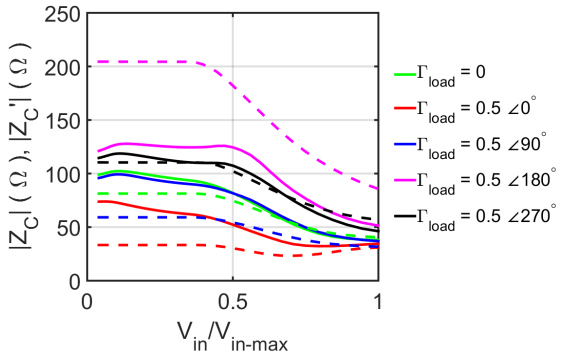


(a) $|Z_c|$ and $|Z_c'|$

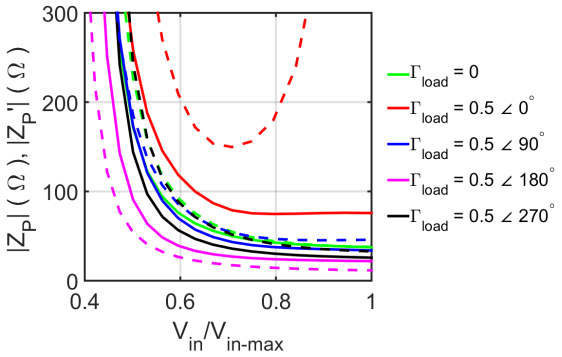


(b) $|Z_p|$ and $|Z_p'|$

Figure 5.5: Simulated impedance seen by the (a) carrier and (b) peaking transistor for the case with (solid line) and without (dashed line) the LA and $|\Gamma_{load}| = 0.2$.



(a) $|Z_c|$ and $|Z_c'|$



(b) $|Z_p|$ and $|Z_p'|$

Figure 5.6: Simulated impedance seen by the carrier and peaking transistor for the case with (solid line) and without (dashed line) the LA and $|\Gamma_{load}| = 0.5$.

Γ_{load}[4], is set such that VSWR at the load, $VSWR_{load}$[5], is equal to values from 1 to 4, and the phase of Γ_{load} was varied over 360°. Notice how the impedance seen by the carrier transistor remains much closer to the optimal value ($VSWR_{load} = 1$) with the LA added.

Next, the magnitude of impedance seen by the carrier and peaking transistor is plotted versus different input voltages for a range of different Z_{load} values in Fig. 5.5. As expected,

[4] $\Gamma_{load} = (Z_{load} - 50)/(Z_{load} + 50)$
[5] $VSWR_{load} = (1 + |\Gamma_{load}|)/(1 - |\Gamma_{load}|)$

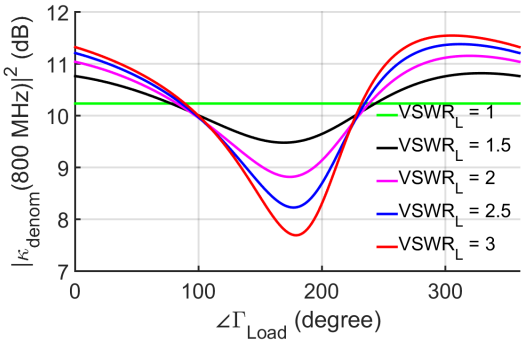
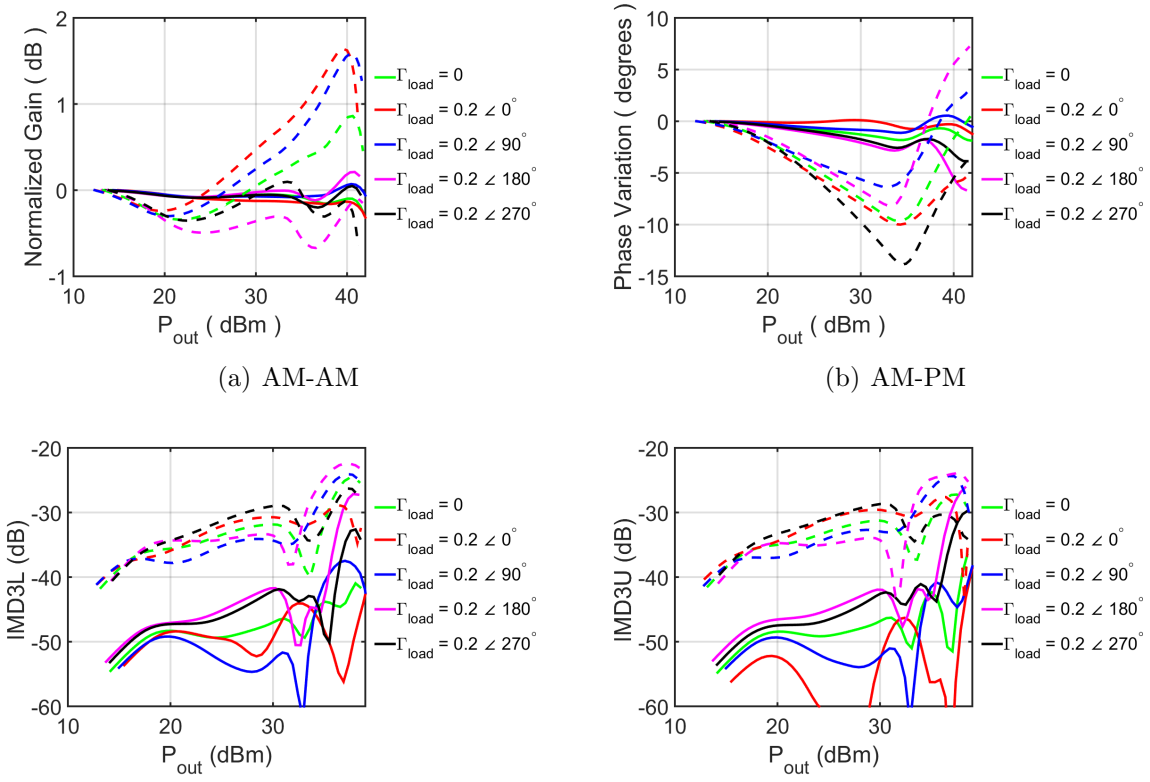


Figure 5.7: Calculated $|\kappa_{denom}(800MHz)|^2$ for various $VSWR_L$ values.

the LA was able to significantly reduce the variation seen by both impedances such that they stayed much closer to the original optimal value. For example, at low power levels the impedance seen by the carrier transistor varied by about 60 Ω without the LA, but only by about 20 Ω with the LA added. At peak power, the impedance seen by the peaking transistor varied by about 52 Ω without the LA, but it only varied by about 19 Ω with the LA added. Finally, the impedance variation was worsened by increasing $|\Gamma_{load}|$ to be 0.5, and the impedances are plotted again versus input voltage in Fig. 5.6. Once again, the LA was able to significantly reduce the amount of impedance variation seen by both the carrier and peaking transistor. For example, the impedance seen by the carrier transistor varied by about 170 Ω without the LA, but only by about 50 Ω with the LA added.

5.3 Linearization of Massive MIMO PAs in the Presence of Antenna Coupling Using the LA

Aside from reducing the load impedance caused by antenna coupling, the LA also improves the linearity performance of PAs in massive MIMO systems by maintaining a high level of linearization under a wide range of load impedance values. Recall the term $\kappa_{denom}(f)$ given in Eq. (4.12). The larger this term is, the more the LA is able to reduce the distortions in the DPA. Although $\kappa_{denom}(f)$ is a function of Z_L, it is shown in Fig. 5.7 that the magnitude of $\kappa_{denom}(f)$ can still remain high versus the Z_L deviations tested. Hence, we can expect that the LA is still able to achieve significant linearization for these various load impedance values. To verify this, the DPA and LA configuration detailed in chapter 4 was simulated again with various Z_{load} values, and the AM-AM and AM-PM distortions



(a) AM-AM

(b) AM-PM

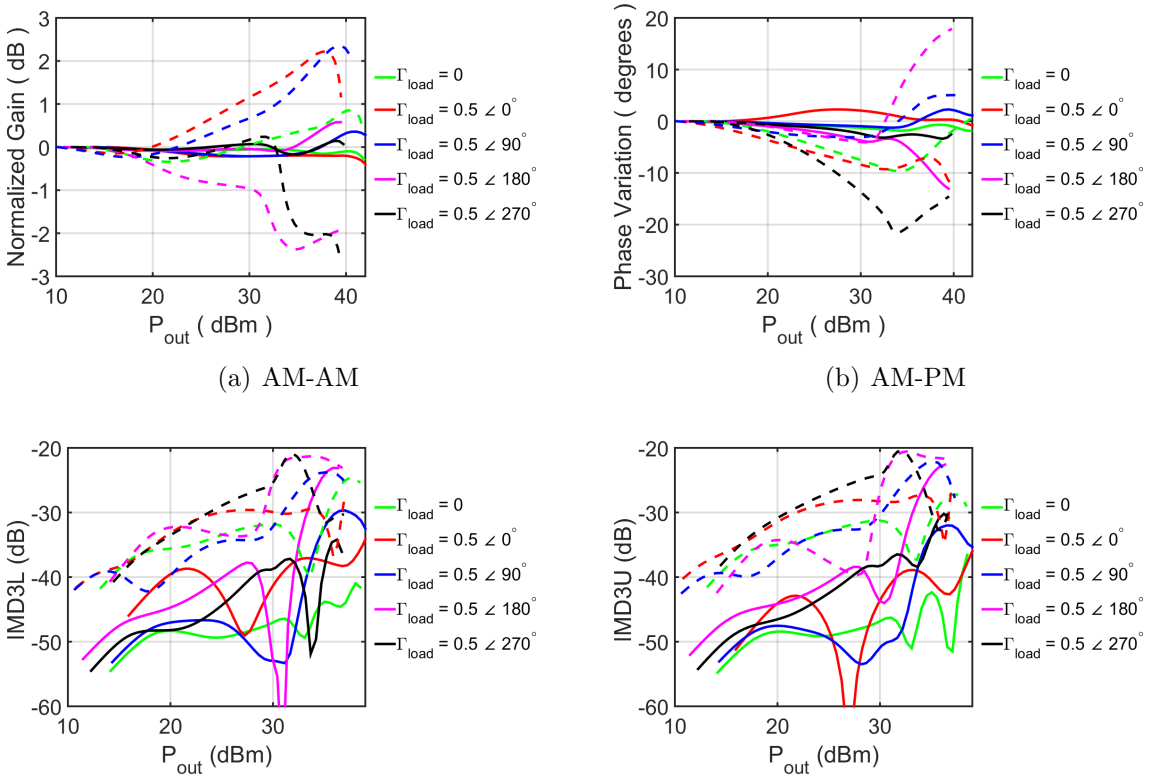
(c) IMD3L

(d) IMD3U

Figure 5.8: Simulated AM-AM, AM-PM, and IMD3s across Z_{load} for the case with (solid line) and without (dashed line) the LA.

are plotted in Fig. 5.8(a) and Fig. 5.8(b), respectively. Without the LA, the AM-AM and AM-PM profiles changes significantly versus load impedance. For example, in the case of AM-AM, the amount of gain variation can go from 0.7 to 1.9 dB for the case when Γ_{load} is equal to $0.2\angle 180°$ and $0.2\angle 90°$, respectively. On the other hand, the AM-AM and AM-PM profile for the case with the LA is much smaller and more consistent. For instance, AM-AM variation for all the load impedance tested stayed within 0.3 dB for the case when the LA was used.

Next, a two-tone input with tone frequencies at 795 and 805 MHz was used, and the IMD3 performances with various Z_{load} values for the case with and without the LA are plotted in Fig. 5.8(c) and 5.8(d). Significant IMD3 improvement was obtained with all the load impedance values tested. Finally, the amount of impedance variation was then



(a) AM-AM

(b) AM-PM

(c) IMD3L

(d) IMD3U

Figure 5.9: Simulated AM-AM, AM-PM, and IMD3s across Z_{load} for the case with (solid line) and without (dashed line) the LA.

increased by raising $|\Gamma_{load}|$ to be 0.5, and the AM-AM, AM-PM, and IMD3s are plotted in Fig. 5.9 for the case with and without the LA. Similarly, significantly improvement in linearity is achieved for all load impedances when the LA was added.

5.4 Simulated Performance with a Four-Element Array Antenna

As an initial test of LA's effectiveness in a large antenna array system, four identical copies of the DPA developed in Chapter 4 (PA1, PA2, PA3, and PA4) were used in simulation to drive a 2×2 array antenna as illustrated in Fig. 5.10. The input signals going into the

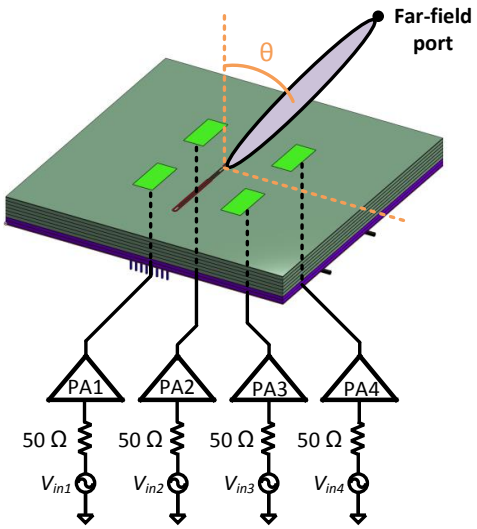
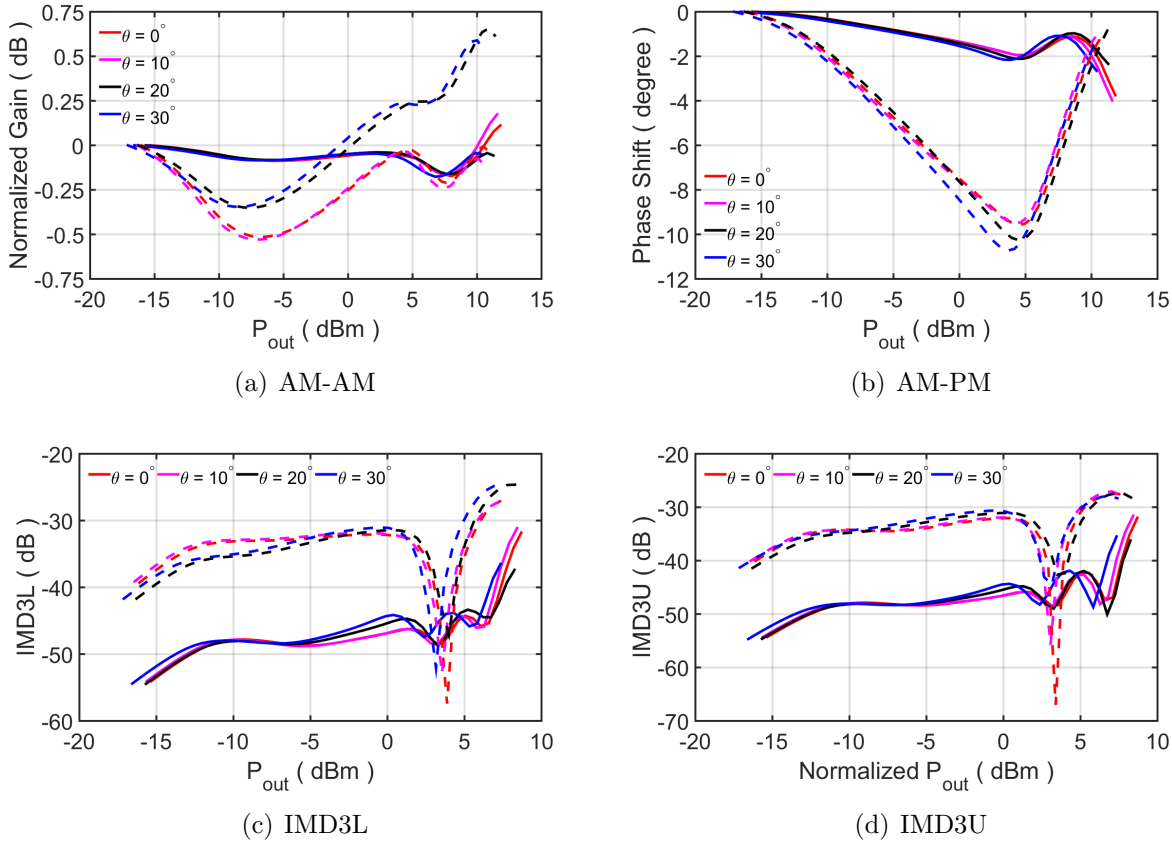


Figure 5.10: Illustration of the simulation setup used to observe the performance of an 2×2 array antenna at a far-field point.

DPAs are termed V_{inx} where x is the PA number. The 2×2 array antenna was modeled by a 5-port scattering parameter using the ANSYS HFSS 3D electromagnetic simulation software where the fifth port is placed at a far-field location in the direction of the beam formed by the antenna.

5.4.1 Performance with Various Input Signal Phases

First, the beam angle θ was swept from zero to 30° by keeping the phase of V_{in3} and V_{in4} fixed at zero and varying the phase of V_{in1} and V_{in2} from zero to 90°. For each beam angle, the normalized AM-AM and AM-PM at the far-field port is plotted in Fig. 5.11 for the case with and without the LA. Without the LA, the behaviour of both AM-AM and AM-PM varies for different beam angles because the impedance seen by each PA changes versus the beam angle due to antenna coupling as shown in Fig. 5.12. A change in load impedance will, in turn, affect the behaviour of the PA's nonlinearity. The fact that the nonlinearity changes versus θ also means that, if a type of predistortion linearization was used, then the predistorter will need to be dynamically adapted for the best output linearity. On the other hand, the AM-AM and AM-PM at the far-field port stays very consistent versus beam angle when the LA was used in the PAs because of two reasons: First, the load impedance seen by the DPA stays much more consistent as can be seen from Fig. 5.12.



(a) AM-AM

(b) AM-PM

(c) IMD3L

(d) IMD3U

Figure 5.11: Simulated AM-AM, AM-PM, and IMD3s at the far-field port for the case with (solid line) and without (dashed line) the LA for various beam angles.

Secondly, the LA is able to achieve a significant amount of linearization across a wide load impedance range as discussed in the previous subsection.

Finally, a two-tone input voltage was applied to each PA with tones at 795 and 805 MHz, and the IMD3L and IMD3U at the far-field port is plotted in Fig. 5.11(c) and Fig. 5.11(d), respectively. With the LA, significant IMD3L and IMD3U improvements were achieved for all beam angles tested.

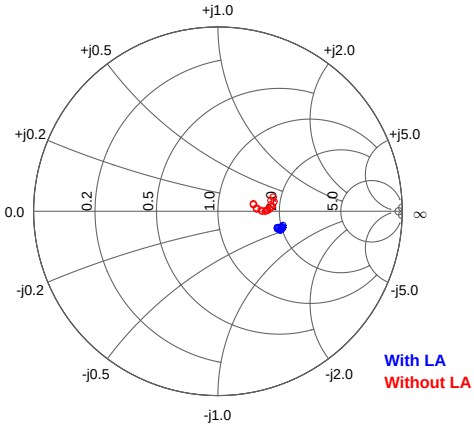


Figure 5.12: Impedance seen by the carrier transistor in the array antenna PAs for the case with and without the LA as θ is swept from zero to 30° in steps of 10°.

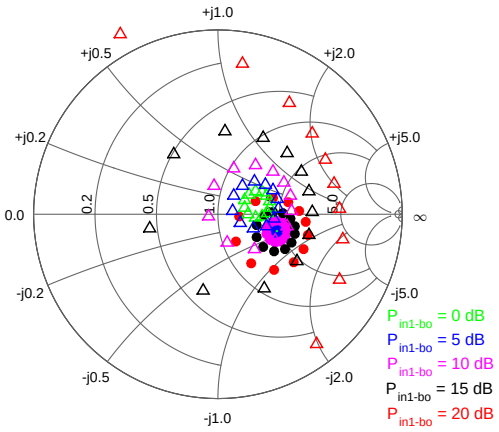
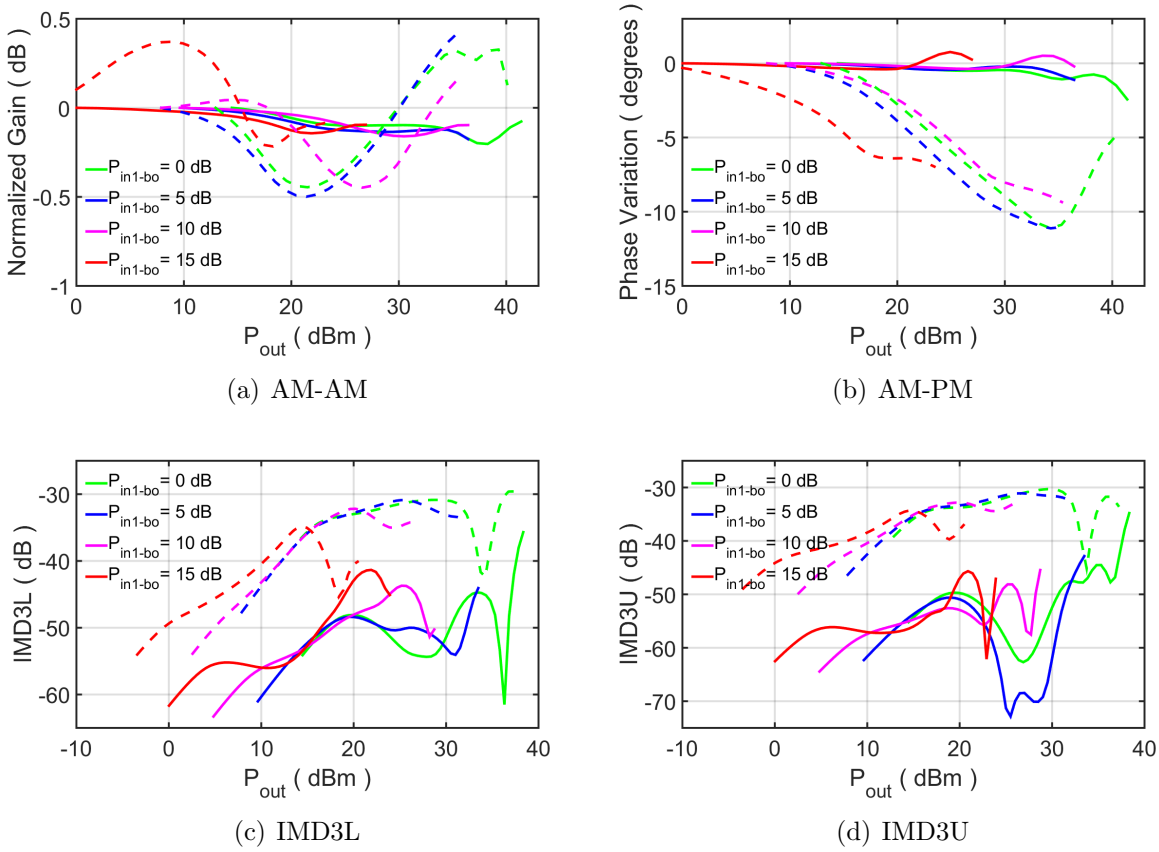


Figure 5.13: Impedance seen by the carrier transistor in the PA1 for the case with (circle) and without (triangle) the LA for various $P_{out1-bo}$ and $\angle V_{in1}$ values.

5.4.2 Performance with Various Input Signal Magnitudes

In the previous test, only the phase of the input signals is varied to achieve different beam angles. However, the magnitude of the input signals can also fluctuate in a massive MIMO system, and this can further complicate the load variations caused by antenna coupling.



(a) AM-AM (b) AM-PM

(c) IMD3L (d) IMD3U

Figure 5.14: Simulated AM-AM, AM-PM, and IMD3s at the output of PA1 for the case with (solid line) and without (dashed line) the LA for various V_{in1} magnitudes. $\angle V_{in1}$ was set to 180° for this test.

To illustrate this, the magnitude of V_{in1} was reduced such that it is $P_{out1-bo}$ dB lower than the magnitude of the input signal going into the other PAs (the input signal going into the other PAs are set to be the same value and kept constant in this test). Moreover, the phase of V_{in1} was also varied from 0 to 330°, and the impedance seen by the carrier transistor in PA1 at a low power level is plotted in Fig. 5.13 for the case with (Z'_{c1}) and without (Z_{c1}) the LA. Notice how, without the LA, there is significant variation in the value of Z_{c1}. In fact, the real part of Z_{c1} even becomes negative at some points[6]. This can be highly problematic because it can cause PA1 to become unstable. With the LA, on the other

[6]Note that there are two more Z_{c1} points that have a negative real part. They are not shown in the plot because they are far from the passive part of the Smith chart.

hand, Z'_{c1} stays much closer to its optimal value, and the real part of Z'_{c1} stays positive for all the cases tested.

It is also expected that the LA will improve the linearity of PA1 when the magnitude of its input signal is varied across a wide range. To show this, the AM-AM and AM-PM at the output of PA1 are plotted for the case when $\angle V_{in1} = 180°$ in Fig. 5.14. The IMD3s at the output of PA1 are also plotted with input tones at 795 and 805 MHz. Without the LA, the linearity of PA1 varies significantly versus $P_{out1-bo}$ due to the large variation of load impedance cause by antenna coupling. With the LA, however, significant linearity improvements were achieved for all the input power levels tested. For example, the AM-AM and AM-PM distortions stayed within 0.2 dB and 2.5 °, respectively, for all $P_{out1-bo}$ values tested when the LA was used.

5.5 Conclusion

In this section, it was shown that unwanted antenna coupling in large antenna arrays can cause the load impedance of the PAs to vary dynamically from its optimal value. As a result, the PA can become unstable and significant distortions can be introduced at the output of the PA which cannot be easily remedied using conventional linearization techniques. In response, analysis and simulations were performed to demonstrate that the LA can be used to counteract this problem by significantly reducing the amount of load variation seen by the PA, such that it remains closer to its optimal value, and by maintaining a high amount of linearization over a wide range of load impedance values. For example, with the LA added, the variation in the impedance seen by the carrier transistor decreased from 60 to 20 Ω and AM-AM distortions decreased from up to 1.9 dB to less than 0.3 dB for a given change in the load impedance.

Lastly, as an initial test of its effectiveness, an LA was applied to each of the four DPAs that were used to drive a 2×2 array antenna. It was shown in simulation that, without the LA, the DPAs will experience a large amount of load impedance variation when either the magnitude or phase of the input voltages is changed for beamforming purposes. This, in turn, will affect the PA's linearity performance and stability. With the LA, however, each DPA now sees a much lower amount of load variation, and the linearity and stability of the DPAs were significantly improved for all conditions tested. For instance, with the LA added, the AM-AM distortions decreased from up to 0.9 dB to 0.2 dB while the AM-PM distortions decreased from up to 11 ° to 2.5 ° for a given amount of change in the magnitude of one of the input voltages.

Chapter 6

Conclusion

In this thesis, a novel wideband PA linearization technique based on the design and addition of an LA was proposed and developed. It was shown that the technique's power overhead is relatively low, scales with the power range of the PA, and does not increase with signal bandwidth, making this approach very attractive for use with future 5G small-cell base stations and large antenna arrays.

The proposed linearization technique was first applied to a high peak efficiency 6 W class AB PA with a centre frequency of 850 MHz, and the technique's effectiveness was validated by measurement results which showed that the LA can significantly improve the PA's linearity even when it is amplifying signals with a modulation bandwidth as wide as 160 MHz. For example, in the case where the PA was driven by a modulated signal with 40 MHz bandwidth, the PA's ACLR was improved by up to 13 dB with the addition of the LA, allowing the PA to achieve an ACLR of -45 dBc without any additional linearization techniques. Moreover, it was shown that the LA can also be used in conjunction with a simple RF predistorter to further improve the efficiency and linearity of the class AB PA.

Next, the LA was augmented with a conventional DPA design to form a new linear DPA topology that achieves a better linearity-efficiency trade-off compared to the linearized Class AB PA. A study of the interactions between the LA and DPA circuitries was conducted, and a design strategy was developed, to determine the circuit parameters that maximize the ACLR improvement while minimizing the LA's power consumption. A prototype was designed and fabricated with a peak envelope power of 12 W and centre frequency of 800 MHz. Measurements of the prototype confirmed the topology's capability for linear and efficient amplification with wideband modulated signals with up to 40 MHz bandwidth. For example, when stimulated by a 15 MHz WCDMA signal, the linear DPA

achieved -45 dBc of ACLR at 32 dBm output power with a PAE of 27.4 %. In fact, the proposed linear DPA was able to achieve about -45 dBc of ACLR or better over a wide average power range for all modulated signals tested, and the efficiency of the linear DPA remaining significantly higher than the linearized class AB PA, as expected.

Lastly, it was shown that another important benefit of the LA is that it can mitigate the distortions and instability caused by unwanted antenna coupling present in large antenna arrays. It does so by reducing the amount of load impedance variation seen by the PA, such that the load impedance remains closer to its optimal value, and by maintaining a significant amount of linearization under a wide range of load impedance conditions.

6.1 Summary of Contributions

The goal of this thesis was to develop a new linearization method suitable for PAs used in future small cells and large antenna arrays. This goal was achieved through the following contributions:

- Proposed and developed a new wideband PA linearization technique based on the design and addition of an LA. The LA is very suitable for linearizing small cell and large array antenna PAs because its power consumption is relatively low and scales with the PA's power range.
- Verified the effectiveness of the proposed technique by developing a design strategy to linearize a class AB PA when it is driven with wideband modulated signals.
- Proposed and developed a linear DPA topology by augmenting a DPA design with an LA. The interaction between the LA and the DPA was studied, and based on the findings, a design strategy was devised to maximizes distortion reduction while minimizing the LA's power consumption, This further improved the PA's linearity-efficiency trade-off.
- Showed through analysis and simulation that the LA can be used to mitigate the unwanted load impedance variation and distortions caused by coupling between the antennas in a large antenna array system.

6.2 Future Work

Implementation at Higher Frequencies In the near future, operators are expected to deploy 5G networks in the 3.3 - 4.2 GHz and 4.4 - 5 GHz frequency ranges [5]. Hence, it is pertinent to increase the frequency of the LA design such that it can be used to linearize PAs in those frequency ranges.

In the LA prototypes designed thus far a centre frequency of 800 MHz was used for practical implementation with discrete components. At higher frequencies, however, the parasitics associated with discrete transistors and interconnects (in particular, the long interconnects in the LA's feedback path that needed to be routed around the discrete transistor) will be more pronounced and can make the LA less effective and more susceptible to instability. To remedy these drawbacks, the LA can be implemented in an IC technology which will allow for much shorter interconnects and smaller parasitics. The choice of which technology will depend on the power level targeted. For higher powered small-cell PAs, the GaN monolithic microwave integrated circuit (MMIC) process has excellent power capabilities, and can be utilized to provide a few watts of power [79]. For lower-powered small cells and large antenna arrays on the other hand, implementation in a complimentary metal-oxide-semiconductor (CMOS) process is an attractive option as it is low cost, can be produced in high volume, and allows for excellent integration with other transmitter components [80].

Massive MIMO System Prototype In Section 5.4, it was shown through simulation that the LA can reduce the load impedance variations and distortions caused by antenna coupling in large antenna arrays used in a massive MIMO system. The next step is to validate the analysis and simulations through measurements. For this, a massive MIMO system with a large antenna array should be built with an LA and DPA connected to the input of each antenna. Then, the signal at the intended user locations should be measured, and the linearity of the received signal will be compared between the case with and without the LA enabled. Finally, the user locations and environment should be varied to confirm that significant linearity improvement can be achieved across a wide range of conditions when the LA is added.

6.3 List of Relevant Publications

Journal Papers

- **Y. Hu** and S. Boumaiza, "Doherty Power Amplifier Distortion Correction Using an RF Linearization Amplifier," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 5, pp. 2246-2257, May 2018. (Based on Chapter 4.)
- **Y. Hu** and S. Boumaiza, "Power-Scalable Wideband Linearization of Power Amplifiers," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 5, pp. 1456-1464, May 2016. (Based on Chapter 3.)

Conference Papers

- **Y. Hu** and S. Boumaiza, "Joint RF pre-distortion and post-distortion linearization of small cell power amplifiers," *2017 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, Phoenix, AZ, 2017, pp. 66-69. (Based on Chapter 3.)

References

[1] GSM Association, "The mobile economy 2018," 2018. Available at https://www.gsma.com/mobileeconomy/wp-content/uploads/2018/02/The-Mobile-Economy-Global-2018.pdf.

[2] Ericsson, "Ericsson mobility report," November 2018. Available at https://www.ericsson.com/assets/local/mobility-report/documents/2018/ericsson-mobility-report-november-2018.pdf.

[3] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, pp. 22–30, June 2011.

[4] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.

[5] D. Schnaufer and B. Peterson, "Realizing 5G sub-6-GHz massive MIMO using GaN." Available at https://www.mwrf.com/semiconductors/realizing-5g-sub-6-ghz-massive-mimo-using-gan.

[6] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, pp. 101–107, June 2011.

[7] W. Roh, J. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, pp. 106–113, February 2014.

[8] E. Bjornson, L. Van der Perre, S. Buzzi, and E. G. Larsson, "Massive MIMO in sub-6 GHz and mmwave: Physical, practical, and use-case differences," *IEEE Wireless Commun.*, pp. 2–10, 2019.

[9] M. Steer, "Beyond 3G," *IEEE Microw. Mag.*, vol. 8, pp. 76–82, Feb 2007.

[10] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electron. Lett.*, vol. 37, pp. 1417–1418, Nov 2001.

[11] T. Nguyen, "Small cell networks and the evolution of 5G (part 1)." Available at https://www.qorvo.com/design-hub/blog/small-cell-networks-and-the-evolution-of-5g.

[12] "Rf predistortion vs. digital predistortion." Available at https://www.maximintegrated.com/en/products/comms/wireless-rf/rf-predistorters/predistortion-linearization/rfpal-technology/rfpd-vs-dpd.html.

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

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

[15] F. Wang, D. Kimball, J. Popp, A. Yang, D. Lie, P. Asbeck, and L. Larson, "Wideband envelope elimination and restoration power amplifier with high efficiency wideband envelope amplifier for WLAN 802.11g applications," in *2005 IEEE MTT-S International Microwave Symposium Digest*, pp. 4 pp.–, June 2005.

[16] C. Hsia, A. Zhu, J. Yan, P. Draxler, D. Kimball, S. Lanfranco, and P. Asbeck, "Digitally assisted dual-switch high-efficiency envelope amplifier for envelope-tracking base-station power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 59, pp. 2943–2952, Nov 2011.

[17] C. Hsia, D. Kimball, S. Lanfranco, and P. Asbeck, "Wideband high efficiency digitally-assisted envelope amplifier with dual switching stages for radio base-station envelope tracking power amplifiers," in *2010 IEEE MTT-S International Microwave Symposium Digest (MTT)*, pp. 672–675, May 2010.

[18] P. Wright, J. Lees, J. Benedikt, P. Tasker, and S. Cripps, "A methodology for realizing high efficiency class-J in a linear and broadband PA," *IEEE Trans. Microw. Theory Techn.*, vol. 57, pp. 3196–3204, Dec 2009.

[19] J. Yan, Y.-P. Hong, S. Shinjo, K. Mukai, and P. Asbeck, "Broadband high PAE GaN push-pull power amplifier for 500MHz to 2.5 GHz operation," in *2013 IEEE MTT-S International Microwave Symposium Digest (IMS)*, pp. 1–3, June 2013.

[20] J. Jeong, D. F. Kimball, M. Kwak, C. Hsia, P. Draxler, and P. M. Asbeck, "Wideband envelope tracking power amplifiers with reduced bandwidth power supply waveforms and adaptive digital predistortion techniques," *IEEE Trans. Microw. Theory Techn.*, vol. 57, pp. 3307–3314, Dec 2009.

[21] J. F. Miranda, M. Olavsbråten, K. M. Gjertsen, and W. Caharija, "First and second order polynomial functions for bandwidth reduction of dynamic biasing signals," in *2010 IEEE International Conference on Wireless Information Technology and Systems*, pp. 1–4, Aug 2010.

[22] H. Sarbishaei, , B. Fehri, and S. Boumaiza, "Concurrent dual-band envelope tracking power amplifier for carrier aggregated systems," in *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, pp. 1–4, June 2014.

[23] H. Chireix, "High power outphasing modulation," *Proceedings of the Institute of Radio Engineers*, vol. 23, pp. 1370–1392, Nov 1935.

[24] F. Raab, "Efficiency of outphasing RF power-amplifier systems," *IEEE Trans. Commun.*, vol. 33, pp. 1094–1099, Oct 1985.

[25] A. Birafane and A. B. Kouki, "On the linearity and efficiency of outphasing microwave amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 52, pp. 1702–1708, July 2004.

[26] F. Raab, P. Asbeck, S. Cripps, P. Kenington, Z. Popovic, N. Pothecary, J. Sevic, and N. Sokal, "Power amplifiers and transmitters for RF and microwave," *IEEE Trans. Microw. Theory Techn.*, vol. 50, pp. 814–826, Mar 2002.

[27] R. Langridge, T. Thornton, P. M. Asbeck, and L. E. Larson, "A power re-use technique for improved efficiency of outphasing microwave power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 47, pp. 1467–1470, Aug 1999.

[28] I. Hakala, D. K. Choi, L. Gharavi, N. Kajakine, J. Koskela, and R. Kaunisto, "A 2.14-GHz Chireix outphasing transmitter," *IEEE Trans. Microw. Theory Techn.*, vol. 53, pp. 2129–2138, June 2005.

[29] X. Zhang, L. E. Larson, P. M. Asbeck, and P. Nanawa, "Gain/phase imbalance-minimization techniques for LINC transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 49, pp. 2507–2516, Dec 2001.

[30] W. H. Doherty, "A new high efficiency power amplifier for modulated waves," *Proceedings of the Institute of Radio Engineers*, vol. 24, pp. 1163–1182, Sept 1936.

[31] H. Golestaneh, F. Malekzadeh, and S. Boumaiza, "An extended-bandwidth three-way Doherty power amplifier," *IEEE Trans. Microw. Theory Techn.*, vol. 61, pp. 3318–3328, Sept 2013.

[32] C. Andersson, D. Gustafsson, J. Chani Cahuana, R. Hellberg, and C. Fager, "A 1-3 GHz digitally controlled dual-RF input power-amplifier design based on a Doherty-outphasing continuum analysis," *IEEE Trans. Microw. Theory Techn.*, vol. 61, pp. 3743–3752, Oct 2013.

[33] R. Darraji, F. Ghannouchi, and M. Helaoui, "Mitigation of bandwidth limitation in wireless Doherty amplifiers with substantial bandwidth enhancement using digital techniques," *IEEE Trans. Microw. Theory Techn.*, vol. 60, pp. 2875–2885, Sept 2012.

[34] J. H. Qureshi, W. Sneijers, R. Keenan, L. C. N. deVreede, and F. van Rijs, "A 700-W peak ultra-wideband broadcast Doherty amplifier," in *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, pp. 1–4, June 2014.

[35] D.-T. Wu and S. Boumaiza, "A modified Doherty configuration for broadband amplification using symmetrical devices," *IEEE Trans. Microw. Theory Techn.*, vol. 60, pp. 3201–3213, Oct 2012.

[36] D.-T. Wu, J. Annes, M. Bokatius, P. Hart, E. Krvavac, and G. Tucker, "A 350 W, 790 to 960 MHz wideband LDMOS Doherty amplifier using a modified combining scheme," in *2014 IEEE MTT-S International Microwave Symposium (IMS)*, pp. 1–4, June 2014.

[37] R. Giofre, L. Piazzon, P. Colantonio, and F. Giannini, "An ultra-broadband GaN doherty amplifier with 83% of fractional bandwidth," *IEEE Microw. Wireless Compon. Lett.*, vol. 24, pp. 775–777, Nov 2014.

[38] D. M. Pozar, *Microwave engineering; 3rd ed.* Hoboken, NJ: Wiley, 2005.

[39] J. Sevic, K. Burger, and M. Steer, "A novel envelope-termination load-pull method for ACPR optimization of RF/microwave power amplifiers," in *1998 IEEE MTT-S International Microwave Symposium Digest*, vol. 2, pp. 723–726 vol.2, June 1998.

[40] H. Sarbishaei, D. Y.-T. Wu, and S. Boumaiza, "Linearity of GaN HEMT RF power amplifiers - a circuit perspective," in *2012 IEEE MTT-S International Microwave Symposium Digest (MTT)*, pp. 1–3, June 2012.

[41] H. Sarbishaei, *Concurrent Multi-Band Envelope Tracking Power Amplifiers for Emerging Wireless Communications*. PhD thesis, University of Waterloo, 2014.

[42] Y. Hu, H. Sarbishaei, and S. Boumaiza, "Broadband power amplifier linearity enhancement under carrier aggregated stimulus using an auxiliary impedance branch," in *2015 IEEE International Wireless Symposium (IWS)*, pp. 1–4, March 2015.

[43] S. C. Binari, P. B. Klein, and T. E. Kazior, "Trapping effects in gan and sic microwave fets," *Proceedings of the IEEE*, vol. 90, pp. 1048–1058, June 2002.

[44] H. Golestaneh, *Broadband Doherty Power Amplifiers with Enhanced Linearity for Emerging Radio Transmitters*. PhD thesis, University of Waterloo, 2016.

[45] L. Piazzon, R. Giofrè, P. Colantonio, and F. Giannini, "Investigation of the AM/PM distortion in Doherty power amplifiers," in *2014 IEEE Topical Conference on Power Amplifiers for Wireless and Radio Applications (PAWR)*, pp. 7–9, Jan 2014.

[46] A. Ezzeddine, H.-L. Hung, and H. Huang, "An MMAC C-band FET feedback power amplifier," *IEEE Trans. Microw. Theory Techn.*, vol. 38, pp. 350–357, Apr 1990.

[47] K.-J. Cho, J.-H. Kim, and S. Stapleton, "A highly efficient Doherty feedforward linear power amplifier for W-CDMA base-station applications," *IEEE Trans. Microw. Theory Techn.*, vol. 53, pp. 292–300, Jan 2005.

[48] H. Choi, Y. Jeong, C. D. Kim, and J. S. Kenney, "Efficiency enhancement of feedforward amplifiers by employing a negative group-delay circuit," *IEEE Trans. Microw. Theory Techn.*, vol. 58, pp. 1116–1125, May 2010.

[49] R. N. Braithwaite and A. Khanifar, "High efficiency feedforward power amplifier using a nonlinear error amplifier and offset alignment control," in *2013 IEEE MTT-S International Microwave Symposium Digest (MTT)*, pp. 1–4, June 2013.

[50] K. Yamauchi, K. Mori, M. Nakayama, Y. Itoh, Y. Mitsui, and O. Ishida, "A novel series diode linearizer for mobile radio power amplifiers," in *IEEE MTT-S International Microwave Symposium Digest, 1996.*, vol. 2, pp. 831–834 vol.2, June 1996.

[51] N. Imai, T. Nojima, and T. Murase, "Novel linearizer using balanced circulators and its application to multilevel digital radio systems," *IEEE Trans. Microw. Theory Techn.*, vol. 37, pp. 1237–1243, Aug 1989.

[52] T. Nojima and T. Konno, "Cuber predistortion linearizer for relay equipment in 800 MHz band land mobile telephone system," *IEEE Trans. Veh. Technol.*, vol. 34, pp. 169–177, Nov 1985.

[53] Y. Lee, S. Lee, and Y. Jeong, "Linearity improvement of Doherty amplifier with analog predistorter for WCDMA applications," in *2006 European Microwave Conference*, pp. 1205–1208, Sept 2006.

[54] Y. Lee, M. Lee, S. Kam, and Y. Jeong, "Advanced design of high-linearity analog predistortion doherty amplifiers using spectrum analysis for WCDMA applications," in *2010 IEEE Radio and Wireless Symposium (RWS)*, pp. 140–143, Jan 2010.

[55] B. Shin, J. Cha, J. Kim, Y. Y. Woo, J. Yi, and B. Kim, "Linear power amplifier based on 3-way Doherty amplifier with predistorter," in *2004 IEEE MTT-S International Microwave Symposium Digest (IEEE Cat. No.04CH37535)*, vol. 3, pp. 2027–2030 Vol.3, June 2004.

[56] C. Ma, W. Pan, S. Shao, C. Qing, and Y. Tang, "A wideband Doherty power amplifier with 100 MHz instantaneous bandwidth for LTE-advanced applications," *IEEE Microw. Wireless Compon. Lett.*, vol. 23, pp. 614–616, Nov 2013.

[57] F. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microw. Mag.*, vol. 10, pp. 52–64, Dec 2009.

[58] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, "LTE-advanced: next-generation wireless broadband technology [invited paper]," *IEEE Wireless Commun.*, vol. 17, pp. 10–22, June 2010.

[59] D. Webster, J. Scott, and D. Haigh, "Control of circuit distortion by the derivative superposition method [MMIC amplifier]," *IEEE Microw. Guided Wave Lett.*, vol. 6, pp. 123–125, March 1996.

[60] B. Kim, J.-S. Ko, and K. Lee, "A new linearization technique for MOSFET RF amplifier using multiple gated transistors," *IEEE Microw. Guided Wave Lett.*, vol. 10, pp. 371–373, Sept 2000.

[61] D. Webster and D. Haigh, "Low-distortion MMIC power amplifier using a new form of derivative superposition," *IEEE Trans. Microw. Theory Techn.*, vol. 49, pp. 328–332, Feb 2001.

[62] K.-J. Cho, W.-J. Kim, J.-H. Kim, and S. P. Stapleton, "Linearity optimization of a high power Doherty amplifier based on post-distortion compensation," *IEEE Microw. Wireless Compon. Lett.*, vol. 15, pp. 748–750, Nov 2005.

[63] Y. Yang, J. Cha, B. Shin, and B. Kim, "A fully matched N-way Doherty amplifier with optimized linearity," *IEEE Trans. Microw. Theory Techn.*, vol. 51, pp. 986–993, March 2003.

[64] J. Kim, B. Fehri, S. Boumaiza, and J. Wood, "Power efficiency and linearity enhancement using optimized asymmetrical Doherty power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 59, pp. 425–434, Feb 2011.

[65] Y.-S. Lee, M.-W. Lee, and Y.-H. Jeong, "Linearity-optimized power tracking GaN HEMT Doherty amplifier using derivative superposition technique for repeater systems," in *2008 IEEE MTT-S International Microwave Symposium Digest*, pp. 427–430, June 2008.

[66] X. Fang, A. Chung, and S. Boumaiza, "Linearity-enhanced Doherty power amplifier using output combining network with predefined AM-PM characteristics," *IEEE Trans. Microw. Theory Techn.*, pp. 1–10, 2018.

[67] W. Hallberg, M. Özen, D. Gustafsson, K. Buisman, and C. Fager, "A Doherty power amplifier design method for improved efficiency and linearity," *IEEE Trans. Microw. Theory Techn.*, vol. 64, pp. 4491–4504, Dec 2016.

[68] R. Giofré and P. Colantonio, "A high efficiency and low distortion 6 W GaN MMIC Doherty amplifier for 7 GHz radio links," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, pp. 70–72, Jan 2017.

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

[70] H.-T. Dabag, H. Gheidi, S. Farsi, P. Gudem, and P. Asbeck, "All-digital cancellation technique to mitigate receiver desensitization in uplink carrier aggregation in cellular handsets," *IEEE Trans. Microw. Theory Techn.*, vol. 61, pp. 4754–4765, Dec 2013.

[71] J. H. Qureshi, N. Li, W. C. E. Neo, F. van Rijs, I. Blednov, and L. C. N. de Vreede, "A wide-band 20W LMOS Doherty power amplifier," in *IEEE MTT-S Int. Microw. Symp. Dig.*, pp. 1504–1507, May 2010.

[72] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[73] F. M. Barradas, T. R. Cunha, and J. C. Pedro, "Digital predistortion of RF PAs for MIMO transmitters based on the equivalent load," in *2017 Integrated Nonlinear Microwave and Millimetre-wave Circuits Workshop (INMMiC)*, pp. 1–4, April 2017.

[74] C. Fager, X. Bland, K. Hausmair, J. Chani Cahuana, and T. Eriksson, "Prediction of smart antenna transmitter characteristics using a new behavioral modeling approach," in *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, pp. 1–4, June 2014.

[75] C. Fager, K. Hausmair, K. Buisman, K. Andersson, E. Sienkiewicz, and D. Gustafsson, "Analysis of nonlinear distortion in phased array transmitters," in *2017 Integrated Nonlinear Microwave and Millimetre-wave Circuits Workshop (INMMiC)*, pp. 1–4, April 2017.

[76] K. Hausmair, P. N. Landin, U. Gustavsson, C. Fager, and T. Eriksson, "Digital predistortion for multi-antenna transmitters affected by antenna crosstalk," *IEEE Trans. Microw. Theory Techn.*, vol. 66, pp. 1524–1535, March 2018.

[77] N. Tervo, J. Aikio, T. Tuovinen, T. Rahkonen, and A. Parssinen, "Digital predistortion of amplitude varying phased array utilising over-the-air combining," in *2017 IEEE MTT-S International Microwave Symposium (IMS)*, pp. 1165–1168, June 2017.

[78] E. Ng, Y. Beltagy, P. Mitran, and S. Boumaiza, "Single-input single-output digital predistortion of power amplifier arrays in millimeter wave RF beamforming transmitters," in *2018 IEEE/MTT-S International Microwave Symposium - IMS*, pp. 481–484, June 2018.

[79] C. H. Kim, S. Jee, G. Jo, K. Lee, and B. Kim, "A 2.14-GHz GaN MMIC Doherty power amplifier for small-cell base stations," *IEEE Microw. Wireless Compon. Lett.*, vol. 24, pp. 263–265, April 2014.

[80] H. Wang, S. Kousai, K. Onizuka, and S. Hu, "The wireless workhorse: Mixed-signal power amplifiers leverage digital and analog techniques to enhance large-signal RF operations," *IEEE Microw. Mag.*, vol. 16, pp. 36–63, Oct 2015.

APPENDICES

Appendix A

IMD3 Improvement Breakdown

The total reduction in IMD3L due to the addition of the LA, $\Delta IMD3L_{lin}$, was expressed as

$$\Delta IMD3L_{lin} = (G + IMD3L_{orig} - IMD3L_{lp}) - 20\log(\kappa(2f_1 - f_2)). \tag{A.1}$$

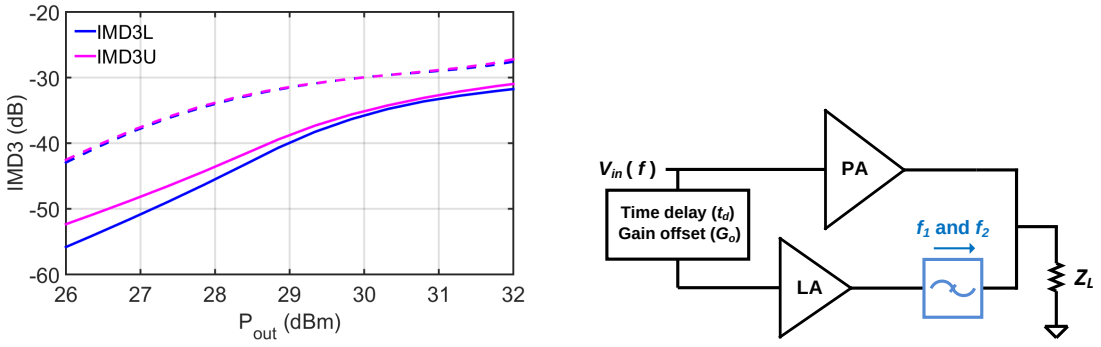
in Eq. (3.13), where $\kappa(2f_1 - f_2)$ is given as

$$\kappa(2f_1 - f_2) = \frac{|1 + \alpha(2f_1 - f_2)|}{|\beta^{-1}(2f_1 - f_2) + 1|} \tag{A.2}$$

in Eq. (3.10), where $\alpha(f)$ and $\beta(f)$ are the ratios between the LA and PA's Norton current and output impedance, respectively, at frequency f. In this section, simulations using Keysight's ADS harmonic balance simulator are used to validate this expression.

Recall that the LA was able to improve IMD3 through two ways. One of the ways was through signal power enhancement. To demonstrate this, a two-tone input signal with tones at 840 MHz (f_1) and 860 MHz (f_2) was applied to the LA and PA and an ideal filter which only passes currents at f_1 and f_2 was inserted between the LA and PA output as shown in Fig. A.1(b). As a result, the LA will only affect the fundamental tones. IMD3 with and without the LA is then plotted in Fig. A.1(a). After the addition of the LA and the filter, the power of the output signal was increased by around 2 dB (G=2 dB). As a result, the PA's IMD3 curves shifted 2 dB to the right and 2 dB down, and the IMD3 at 30 dBm output power improved by 5.8 dB and 5.2 dB for IMD3L and IMD3U, respectively. (This is the value of $G + IMD3L_{orig} - IMD3L_{lp}$ and $G + IMD3U_{orig} - IMD3U_{lp}$, respectively.)

The LA also improves the IMD3 through distortion reduction, and the amount of distortion reduction was found using the relationship for κ given in Eq. (3.10). For this,



(a) (b)

Figure A.1: The a) simulated IMD3L and IMD3U with tones 1 and 2 at 840 MHz and 860 MHz, respectively, for the case with (solide line) and without (dashed line) the LA and filter, and an b) illustration of the setup used for the simulation.

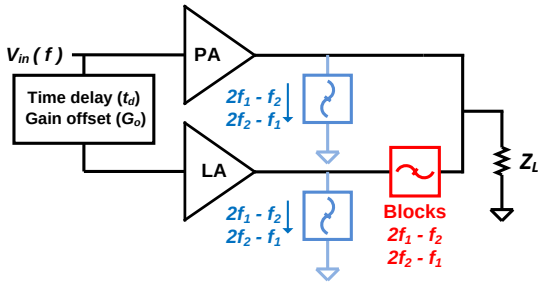


Figure A.2: Simulation setup used to find the PA and LA's Norton currents at the IMD3 frequencies

the small-signal output impedance of both the PA and LA were found through simulation, while the Norton currents at the IMD3 frequencies were found using the simulation setup shown in Fig. A.2. An ideal filter was added between the PA and LA to block the currents at the IMD3 frequencies, and an ideal filter was added at the PA and LA output to short the IMD3 currents to ground. Then, the IMD3 currents at the output of the PA and LA were measured to estimate their Norton currents at the IMD3 frequencies.

The simulated Norton currents and output impedance of the PA and LA were obtained and recorded in Table A for 30 dBm output power. These values are then used to calculate κ using Eq. (3.10), which is then used to calculate $\Delta IMD3L_{lin}$ using Eq. (3.13). (The equation was also adapted to calculate the expected improvement in IMD3U.) From this

Table A.1: Calculated and simulated IMD3 reduction with tones at 840 and 860 MHz and an output power of 30 dBm.

	IMD3L	IMD3U
I_{LA}	$0.00332\angle 52.9A$	$0.00343\angle -1A$
I_{PA}	$0.00475\angle -148.2A$	$0.00454\angle -162.9A$
$\alpha = I_{LA}/I_{PA}$	$0.699\angle -158.9$	$0.755\angle 161.9$
Z_{LA}	$18.9\angle 8.2\Omega$	$20\angle 22.1\Omega$
Z_{PA}	$26.4\angle 13.4\Omega$	$28.8\angle 8.2\Omega$
$\beta = Z_{LA}/Z_{PA}$	$0.72\angle -5.2$	$0.7\angle 13.9$
κ	-14.9 dB	-16.4 dB
$G + IMD3_{orig} - IMD3_{lp}$	5.8 dB	5.2 dB
Calculated $\Delta IMD3_{lin}$	20.7 dB	21.6 dB
Simulated $\Delta IMD3_{lin}$	20.6 dB	21.4 dB

calculation, improvements of 20.7 dB and 21.6 dB were expected for IMD3L and IMD3U, respectively. This estimation was very close to the level of improvement actually achieved by the addition of the LA in simulation, which was 20.6 dB for IMD3L and 21.4 dB for IMD3U.

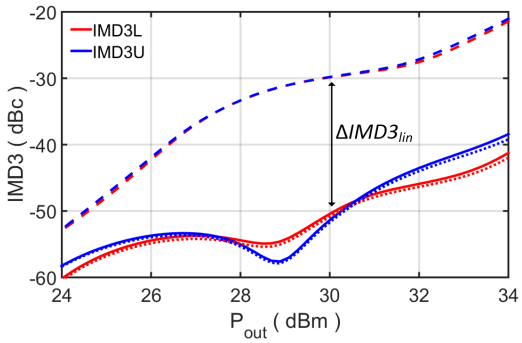


Figure A.3: Simulated IMD3L and IMD3U with tones 1 and 2 at 840 MHz and 860 MHz, respectively, for the case with (solid line) and without (dashed line) the LA. Also plotted is the calculated IMD3L and IMD3U (short dashes).

This comparison was also done over a wide range of other power levels, and the calculated and simulated IMD3 values are plotted in Fig. A.3. Notice how excellent agreement is obtained between calculation and simulation, thus validating the expression given in Eq. (3.13).